The World Journal of Biological Psychiatry, 2015; Early Online: 1–11

ORIGINAL INVESTIGATION

Common and rare variants of microRNA genes in autism spectrum disorders

CLAUDIO TOMA[1,2], BÀRBARA TORRICO[1,2], AMAIA HERVÁS[3,4], MARTA SALGADO[3,4], ISABEL RUEDA[4], RAFAEL VALDÉS-MAS[5], JAN K. BUITELAAR[6,7], NANDA ROMMELSE[7,8], BARBARA FRANKE[8,9], CHRISTINE FREITAG[10], ANDREAS REIF[11], LUIS ALBERTO PÉREZ-JURADO[2,12,13], AGATINO BATTAGLIA[14], LUIGI MAZZONE[15], ELENA BACCHELLI[16], XOSE S. PUENTE[5] & BRU CORMAND[1,2,17]

[1]*Departament de Genètica, Universitat de Barcelona, Spain,* [2]*Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Spain,* [3]*Child and Adolescent Mental Health Unit, Hospital Universitari Mútua de Terrassa, Spain,* [4]*Developmental Disorders Unit (UETD), Hospital Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain,* [5]*Department of Biochemistry and Molecular Biology, University of Oviedo-IUOPA, Spain,* [6]*Radboud University Medical Center, Donders Institute for Brain, Cognition and Behaviour, Department of Cognitive Neuroscience, Nijmegen, The Netherlands,* [7]*Karakter Child and Adolescent Psychiatry University Centre, Nijmegen, The Netherlands,* [8]*Radboud University Medical Center, Donders Institute for Brain, Cognition and Behaviour, Department of Psychiatry, Nijmegen, The Netherlands,* [9]*Radboud University Medical Center, Donders Institute for Brain, Cognition and Behaviour, Department of Human Genetics, Nijmegen, The Netherlands,* [10]*Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital, Frankfurt, Germany,* [11]*Department of Psychiatry, Psychosomatics, and Psychotherapy, University of Wuerzburg, Germany,* [12]*Unitat de Genètica, Universitat Pompeu Fabra, Barcelona, Spain,* [13]*Hospital del Mar Research Institute, IMIM, Barcelona, Spain,* [14]*Stella Maris Clinical Research Institute for Child and Adolescent Neuropsychiatry, Calambrone, Pisa, Italy,* [15]*Child Neuropsychiatry Unit, Department of Neuroscience, I.R.C.C.S. Children's Hospital Bambino Gesù, Rome, Italy,* [16]*Department of Pharmacy and Biotechnology, University of Bologna, Italy and* [17]*Institut de Biomedicina de la Universitat de Barcelona (IBUB), Spain*

Abstract

Objectives. MicroRNAs (miRNAs) are post-transcriptional regulators that have been shown to be involved in disease susceptibility. Here we explore the possible contribution of common and rare variants in miRNA genes in autism spectrum disorders (ASD). *Methods.* A total of 350 tag SNPs from 163 miRNA genes were genotyped in 636 ASD cases and 673 controls. A replication study was performed in a sample of 449 ASD cases and 415 controls. Additionally, rare variants in 701 miRNA genes of 41 ASD patients were examined using whole-exome sequencing. *Results.* The most significant association in the discovery sample was obtained for the miR-133b/miR-206 cluster (rs16882131, $P = 0.00037$). The replication study did not reach significance. However, the pooled analysis (1,085 cases and 1,088 controls) showed association with two miRNA clusters: miR-133b/miR-206 (rs16882131, $permP = 0.037$) and miR-17/miR-18a/miR-19a/miR-20a/miR-19b-1/miR92a-1 (rs6492538, $permP = 0.019$). Both miR-133b and miR-206 regulate the *MET* gene, previously associated with ASD. Rare variant analysis identified mutations in several miRNA genes, among them miR-541, a brain-specific miRNA that regulates *SYN1*, found mutated in ASD. *Conclusions.* Although our results do not establish a clear role for miRNAs in ASD, we pinpointed a few candidate genes. Further exome and GWAS studies are warranted to get more insight into their potential contribution to the disorder.

Key words: *autistic disorder; genetic association studies; biological psychiatry; MicroRNAs; exome*

Correspondence: Claudio Toma, Neuroscience Research Australia, Margarete Ainsworth Building, Barker Street, Randwick NSW 2031, Sydney, Australia. Tel: + 61-29399-1890. E-mail: c.toma@neura.edu.au; Bru Cormand, Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 643, Edifici Prevosti, planta 3, 08028 Barcelona, Catalonia, Spain. Tel: + 34-93-4021013. E-mail: bcormand@ub.edu

(Received 5 December 2014; accepted 4 March 2015)

ISSN 1562-2975 print/ISSN 1814-1412 online © 2015 Informa Healthcare
DOI: 10.3109/15622975.2015.1029518

Introduction

Recently, large-scale technologies for genome analysis have allowed the identification of a large number of non-coding RNAs (ncRNA) that represent a new challenge for our understanding of gene regulation. ncRNAs are involved in many biological processes and can be grouped in different classes: small interfering RNAs (siRNAs), piwi-interacting RNAs (piRNAs), long ncRNAs and microRNAs (miRNAs). miRNAs are short RNA molecules 19–25 nucleotides in length, that work as post-transcriptional regulators of gene expression by mRNA degradation or translational repression. Approximately two-thirds of human genes are estimated to be regulated by miRNAs (Friedman et al. 2009). Deep sequencing data have allowed the identification of an increasing number of miRNA genes in the human genome, which total number has been estimated as 1,881 in a recent annotation, 296 considered as highly confident (http://www.mirbase.org).

The biogenesis of miRNAs starts with a precursor transcript that consists of a multiple hairpin structure (pri-miRNA) that is processed to become a second precursor of ~100 bp in length (pre-miRNA) and exported to the cytoplasm. The pre-miRNA is further processed into a double-stranded RNA of ~22 bp. Finally, only one strand acts as the functional active form (mature miRNA), and binds to the specific target site on the mRNA thereby exerting the post-transcriptional repression. The binding of the mature miRNA to the 3′-UTR of the target mRNA is achieved by a specific conserved region of ~7 bp (seed region) (Farh et al. 2005). MiRNA genes can be located between genes (intergenic) or within a host gene (intragenic), usually in an intron, and they can be either isolated or clustered.

MiRNAs are crucial in regulating many biological processes, and some of them have been implicated in synaptic plasticity, brain development and memory functions (Krichevsky et al. 2003; Sempere et al. 2004; Li and Jin 2010; Griggs et al. 2013). Understanding the role of miRNAs in the mammalian brain is crucial for us to gain insight into their potential involvement in neurological and psychiatric diseases. An individual miRNA may fine-tune tens of genes or entire biological pathways, and for this reason they may have important implications for complex disorders. It is probable that common and rare variants of miRNA genes play a significant role in phenotypic variability, including psychiatric disorder susceptibility. The most convincing evidence for the contribution of miRNAs to neuropsychatric diseases are the recent reports of involvement of the common variant rs1625579 of miRNA-137 in schizophrenia (Schizophrenia Psychiatric GWAS Consortium 2011; Lett et al. 2013). This miRNA is involved in neuronal maturation and migration. Although several studies have investigated the role of miRNA genes in schizophrenia, little has been done to explore their potential role in autism spectrum disorders (ASD), except for a few studies that have analysed miRNA expression. Three reports evaluated the expression of miRNAs in lymphoblastoid cell lines of autistic probands (Talebizadeh et al. 2008; Sarachana et al. 2010; Ghahramani et al. 2011), while another study identified dysregulated miRNAs in the cerebellum of ASD patients (Abu-Elneel et al. 2008). Although those studies represent the first efforts to explore miRNA genes in autism, their results are not conclusive, probably due to the limited sample size, the lack of replication and the tissues studied, which may not systematically mimic neural conditions.

Here we present, to the best of our knowledge, the first study of the potential contribution of common and rare variants in miRNA genes to ASD susceptibility. A case–control association study was performed by genotyping 350 SNPs that tag common variability of 163 miRNA genes, selected for having validated target genes, with a function in the brain or because they have been associated with psychiatric disease. The sample used for this study consisted of 1,085 ASD cases and 1,088 controls from four European populations. Rare variant analysis was performed using whole-exome sequencing data from 101 individuals from 30 ASD families.

Materials and methods

ASD patients and clinical assessment

The sample used in the case–control association study consisted of 1,085 ASD cases and 1,088 control individuals from four European populations: Spain, The Netherlands, Germany and Italy. They were organized into a discovery sample and a replication sample (Table I). The sample used for exome sequencing included 101 individuals comprising 41 ASD patients and their parents. Twenty-one probands belonged to 10 multiplex families, whereas 20 probands were from singleton families. The affected individuals included in this study met DSM-IV TR criteria for autistic disorder, Asperger's disorder or pervasive developmental disorder not otherwise specified (PDD-NOS) and were assessed using the ADI-R (Autism-Diagnostic Interview-Revised) and where possible also the ADOS (Autism Diagnostic Observation Schedule) diagnostic instruments (Lord et al. 1994, 2000). All individuals were European and Caucasian. Controls were anonymous healthy individuals that matched with patients for gender

Table I. Summary of the European ASD and control samples used in the case-control association study.

	Discovery (350 SNPs)		Replication (28 SNPs)		Pooled sample (28 SNPs)	
	Cases (M)	Controls (M)	Cases (M)	Controls (M)	Cases (M)	Controls (M)
Spanish	311 (87)	322 (87)	93 (87)	109 (87)	404 (87)	431 (87)
Dutch	247 (78)	269 (78)	–	–	247 (78)	269 (78)
German	78 (89)	82 (89)	124 (90)	131 (90)	202 (90)	213 (89)
Italian	–	–	232 (82)	175 (81)	232 (82)	175 (81)
Total	636 (84)	673 (84)	449 (85)	415 (85)	1,085 (84)	1,088 (84)

M, Percentage of male individuals.

ratio and population of origin (Table I). Additional clinical information about cases and controls for each country can be found in previous publications (Van Steijn et al. 2012; Toma et al. 2013; Bacchelli et al. 2015). The study was approved by the relevant ethical committees and written informed consent was obtained from all parents according to the Helsinki declaration.

DNA extraction

Genomic DNA was isolated from blood lymphocytes using standard salting-out methods, or alternatively from saliva using the Oragene DNA Extraction kit (DNA Genotek, Ontario, Canada), following the manufacturer's protocol.

Selection of miRNA genes and SNPs

The miRNA genes included in this study were selected on the basis of their validated target genes in Tarbase v5 (http://diana.cslab.ece.ntua.gr/tarbase) and miRecords (http://mirecords.biolead.org). Individual miRNAs were selected if one of their target genes was associated with psychiatric disease or had a specific function in the brain. The selection also included miRNA genes with a known neuronal function, involved in synaptic plasticity or implicated in ASD or schizophrenia. The final selection included 163 miRNA genes, 37 of which are isolated and 126 organized in clusters. The complete list of miRNA genes and details of the selection are shown in Supplementary Table 1 (available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518). The selection of tag SNPs was performed by considering the patterns of linkage disequilibrium (LD) in the CEU panel using HapMap project data (http://www.hapmap.org; phases 1, 2, 3; release 28). The elements regulating miRNA expression are complex and may diverge from the canonical regions observed in protein-coding genes. Approximately 35% of intragenic miRNAs have associated promoters that could drive transcription independently of their host gene promoter, suggesting that miRNA expression may be regulated by multiple promoters (Monteys et al. 2010). Also, the distance between the identified transcription start site (TSS) and its corresponding miRNA may range from 100 bp to 10 kb (Wang et al. 2010). In order to ensure that all the miRNA regulatory elements were included, the region covered by the tag SNP selection included 10 kb upstream and 5 kb downstream from each miRNA gene. Additionally, we considered 5 kb of the host gene promoter region for intragenic miRNAs. A total of 329 tag SNPs were selected using the Tagger implementation in HaploView 4.2 (Barrett et al. 2005). The parameters used for Tagger followed these criteria: $r^2 \geq 0.8$ and minor allele frequency (MAF)≥ 0.05. In addition, 21 SNPs were included in the study because they were located within the pre-miRNA gene, with no restrictions on MAF or validation status (Supplementary Table 2 available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518). Genomic annotations were obtained from the GRCh37/hg19 assembly of the UCSC Human Genome Browser.

SNP genotyping in the discovery sample

A total of 350 SNPs were genotyped at the Barcelona node of the National Genotyping Center (CeGen, www.cegen.org) using VeraCode technology (Illumina, San Diego, CA, USA) (Lin et al. 2009). Fifteen HapMap individuals (5 trios) were genotyped and used as controls and for genotyping clustering. We obtained 100% genotyping concordance with the HapMap individuals and 10 duplicates used in the assay. After exclusion of SNPs and samples that failed, the average call rate was 99.5%. From the 350 SNPs initially genotyped in our study, 302 were finally assessed for association, whereas 48 (14%) were excluded for one of the following reasons: monomorphic SNPs ($n = 11$), deviation from Hardy–Weinberg equilibrium (HWE; threshold set at $P < 0.01$ in our control population) ($n = 5$), MAF < 0.05 ($n = 5$) and SNPs failing in the assay ($n = 27$) (Supplementary Table 3 available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518). Genetic stratification was

excluded in the discovery sample using the STRUCTURE v2.3 software on 44 unlinked SNPs (Pritchard et al. 2000). SNPs were selected from those genotyped in our study that were in HWE, had a MAF > 0.05 and showed no association with the disease phenotype in any of the tested population groups. The analysis was performed under the admixture model, with a length of the burning period and a number of MCMC repeats of 10,000 and performing three independent runs at each K value (from 1 to 4), with K referring to the number of groups to be inferred (Supplementary Table 4 available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518).

SNP genotyping in the replication sample

The six miRNA genes or clusters selected for the replication sample were those that showed association with ASD in at least one of the subsamples in the discovery phase (SNP with $P < 0.01$). A total of 28 SNPs was selected for genotyping at the Santiago de Compostela node of the National Genotyping Center using iPlex-Sequenom technology (Sequenom, San Diego, CA, USA). The SNP genotyping success rate was 100% and the average call rate was 99.7%.

Statistical analysis

Single-marker analysis. HWE and MAF were analysed with the SNPassoc R package (González et al. 2007). Markers that were successfully genotyped and passed quality control filters were assessed in the case–control association study under the additive model using the PLINK package (Purcell et al. 2007). A quantile–quantile plot (Q–Q plot) was generated with the ggplot2 R library (Wickham 2009). Correction for multiple testing was performed in the joint analysis using 10,000 permutations (PLINK package).

Haplotype analysis. To minimize multiple testing and type I (α) errors, we restricted the haplotype-based association study to only those LD blocks that included SNPs surviving multiple testing corrections. The haplotype-based analysis was performed using the HaploView v4.2 software. The same software was used to perform the permutation tests (Barret et al. 2005).

Meta-analysis. The study was performed using the rmeta R package (www.cran.r-project.org/web/packages/rmeta/index.html) and PLINK (Purcell et al. 2007). The Q-statistic was used to assess heterogeneity among the studies. The pooled OR was estimated using a fixed-effects model.

Gene ontology and pathway analysis

Genes regulated by the miRNAs that showed association in the joint analysis (discovery + replication) were analysed via gene ontology (GO) enrichment analysis using the FatiGO software from the Babelomics suite (Medina et al. 2010). We also analysed the pool of miRNA genes carrying rare variants by analysing the "top bio functions" categories of the ingenuity pathway analysis (IPA) using the core analysis method, where only experimentally observed data were used. The comparison between transmitted and untransmitted rare variants in miRNA genes was performed using the comparative analysis method.

Screening for rare variants of miRNA genes showing association

The two clusters showing association in the joint sample (discovery + replication) were also investigated for rare variants. Screening for rare variants was performed in the Spanish sample (241 ASD patients), which was the only population group showing association of both clusters in the joint analysis. Primers and PCR conditions are shown in Supplementary Table 5 (available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518).

Exome sequencing and analysis

Exome enrichment was performed on 3 μg of blood genomic DNA. The TruSeq DNA Sample Preparation kit (Illumina) and the NimbleGen SeqCap EZ Human Exome Library v1.0 kit (Roche), which targets, among others, 701 miRNA genes, were used according to the manufacturer's instructions. The resulting exome libraries were applied to an Illumina flowcell for cluster generation. Paired-end sequencing was performed on a HiSeq2000 instrument (Illumina) using 76-base reads. Reads were aligned to the reference genome (GRCh37) using Burrows–Wheeler analysis (BWA) with the sampe option (Li and Durbin 2009), and a BAM file was generated using SAMtools (Li et al. 2009). SNP calling was performed using a combination of SAMtools and Sidrón as previously described (Puente et al. 2011). Statistics for the "callability" (percentage of the exome covered by at least 10 reads with mapping quality ≥ 30 and base quality ≥ 20) of each sample are shown in Supplementary Table 6 (available online

at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518). Common variants, defined as those present in dbSNP135 with a minor allele frequency ≥ 1%, were filtered out. All rare variants were validated by performing standard PCR reactions and Sanger sequencing. The potential pathogenic role of rare variants affecting the secondary structure of the mature miRNA was evaluated by calculating changes in free energies with UNAFold (http://mfold.rna.albany.edu).

Results

Of the 350 SNPs from 163 miRNA genes and clusters (Supplementary Table 3 available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518), 302 SNPs were assessed after quality control in a total of 636 cases and 673 gender-matched controls from Spanish, Dutch and German populations (Table I). Suggestive associations ($P < 0.01$) were observed for five SNPs in the following miRNA genes: miRNA-219-1 (rs107822), cluster miR-133b/miR-206 (rs16882131 and rs17578851), cluster miR-106b/miR-93/miR-25 (rs4729575) and cluster miR-17/miR-18a/miR-19a/miR-20a/miR-19b-1/miR92a-1 (rs6492538). The most significant result was achieved for SNP rs16882131 at cluster miR-133b/miR-206 ($P = 0.00038$). Figure 1 shows a quantile–quantile plot of all the *P* values obtained under the additive model. The *P-values* for all SNPs in individual populations and in the whole discovery sample are detailed in Supplementary Table 3 (available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518).

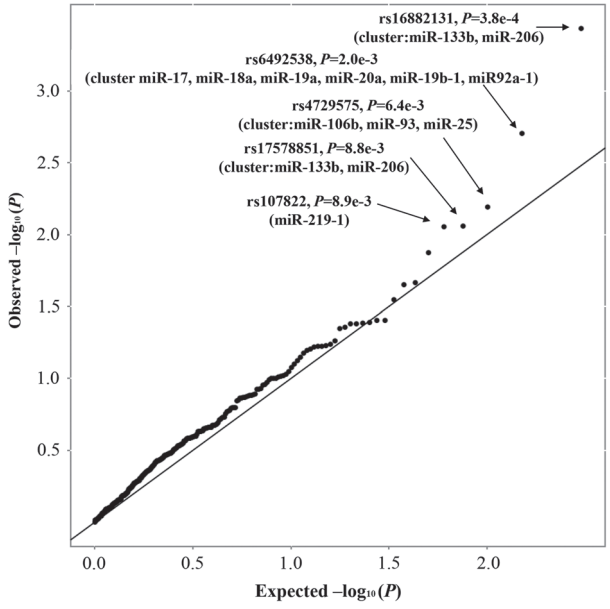

Figure 1. Quantile–quantile plot of 302 *P*-values obtained in the case–control association study (discovery study) of 636 ASD patients and 673 controls. Only markers with *P*-values below 0.01 are indicated.

We sought to replicate the positive associations found in the discovery sample in a European replication sample from Italy, Spain and Germany. We investigated six miRNA genes or clusters, including four miRNA regions showing association in the whole discovery sample and two showing association in only one population: the miR-129-2 gene (rs11606426, $P = 0.0047$, Spanish sample) and miR-15a/miR-16-1 cluster (rs2177313, $P = 0.0002$, German sample). Thus, a total of 28 SNPs were genotyped and assessed in 449 cases and 415 gender-matched controls (Table I). The most significant result in the replication study, considering the whole sample, was achieved for the miR-106b/miR-93/miR-25 cluster (rs13242458, $P = 0.011$). However, none of the 28 SNPs under study survived Bonferroni correction. The *P-values* obtained in the replication study (whole sample and individual populations) are summarized in Supplementary Table 7 (available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518).

In order to increase the statistical power to detect genetic association of modest effects, we performed a joint analysis by pooling the discovery and replication samples, resulting in an association study of 28 SNPs in 1,085 ASD cases and 1,088 controls. This study identified two SNPs that remained significant after applying the Bonferroni correction for multiple testing and also permutations: rs16882131 from the miR-133b/miR-206 cluster ($P = 0.00158$, $permP = 0.037$) and rs6492538 from the miR-17/miR-18a/miR-19a/miR-20a/miR-19b-1/miR92a-1 cluster ($P = 0.00075$, $permP = 0.019$) (Figure 2 and Supplementary Table 8 (available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518)). We also performed meta-analysis considering the four European populations for the two SNPs that showed association in the joint analysis. No heterogeneity between the studies was detected (Q-statistic rs16882131: $P = 0.21$; Q-statistic rs6492538: $P = 0.58$). The meta-analysis under the fixed effects model (Mantel–Haenszel) showed an OR of 0.8 for rs16882131 ($P = 0.0017$, 95% CI = 0.71–0.92), whereas for rs6492538 we obtained an OR of 0.79 ($P = 0.0016$, 95% CI = 0.68–0.91). The meta-analysis Forest plot for both SNPs is shown in Figure 3. Haplotype analysis was only performed for the miR-133b/miR-206 cluster, where the associated marker rs16882131 lies in an LD block of 9 kb that includes both miR-206 and miR-133b (Table II and Figure 4). The rs6492538 marker

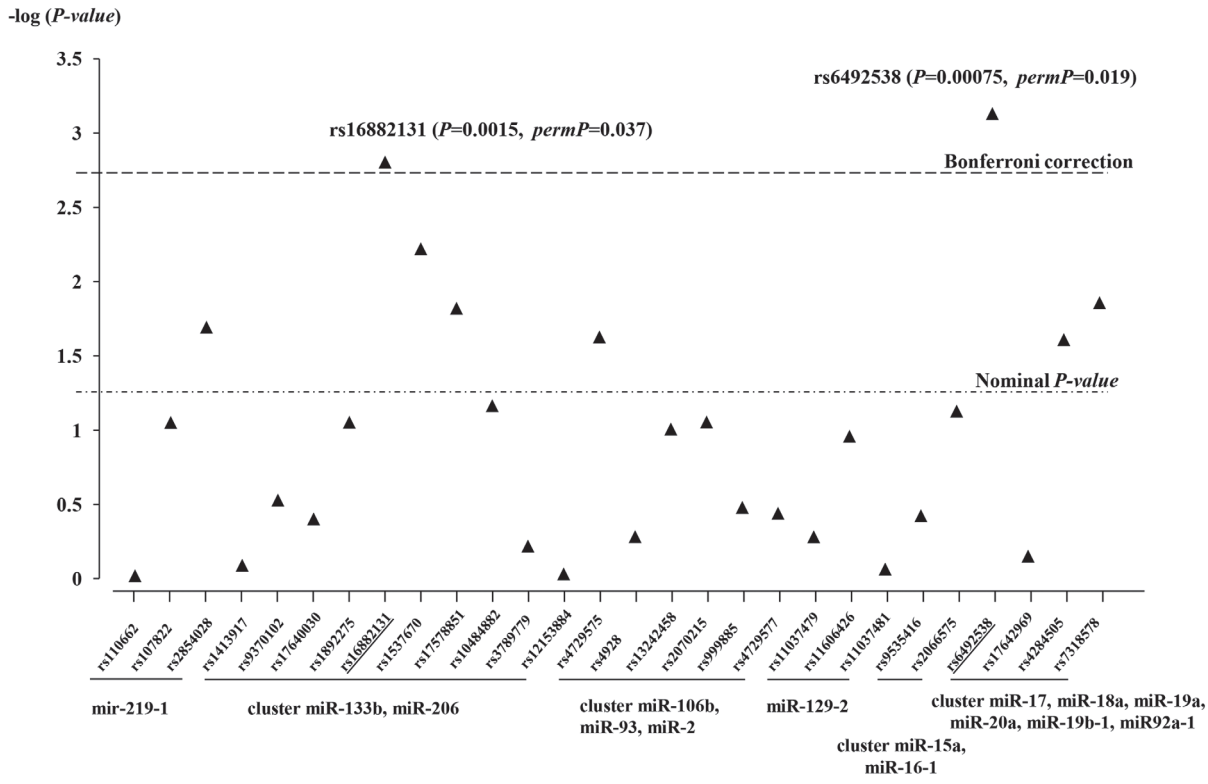

Figure 2. Results from the case-control association study of 28 tag SNPs in the pooled sample (discovery sample and replication sample) of 1,085 ASD cases and 1,088 controls. The *x*-axis shows the studied miRNA tagSNPs, whereas the *y*-axis indicates the significance of the association as –log (*P*-value). The dashed lines represent the thresholds for statistical significance: nominal *P*-value ($P = 0.05$) and Bonferroni correction ($P = 0.0017$). SNPs that survived multiple testing corrections in the pooled sample are underlined.

in the other cluster is a singleton SNP, located 9 kb upstream from the cluster of pre-miRNAs, and no LD blocks were estimated in this region. Multimarker analysis using the whole sample of 1,085 ASD cases and 1,088 controls identified a risk haplotype associated with autism that included six markers (rs1892275,rs16882131,rs1537670,rs17578851, rs10484882, rs3789779). The allelic combination T-C-T-C-C-C was overrepresented in the autistic sample ($P = 0.00017$) and still significant after applying permutations ($permP = 0.0017$). The Spanish and German samples showed several SNPs associ-

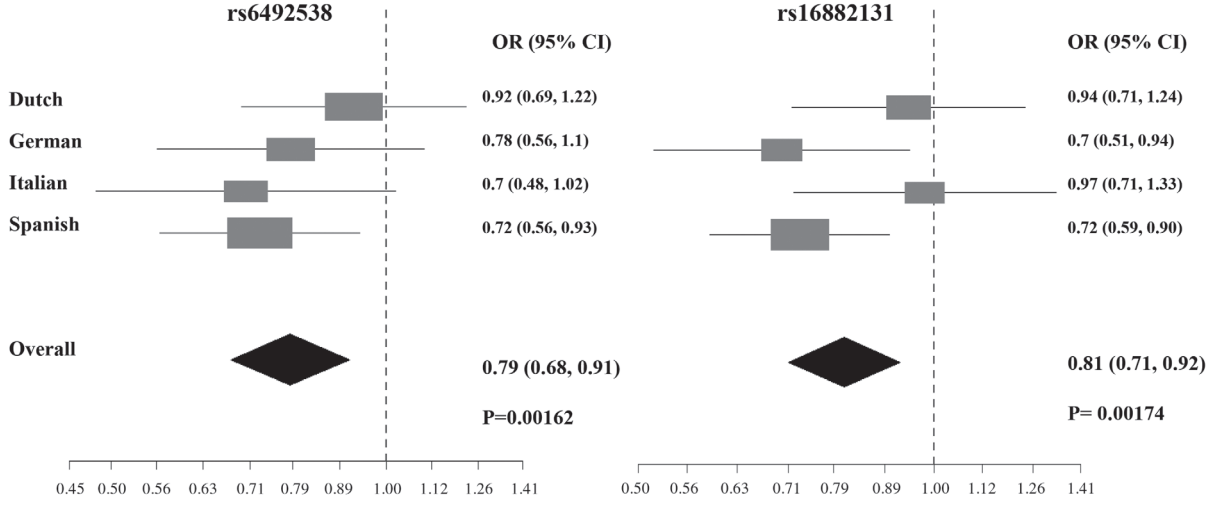

Figure 3. Forest plot for the odds ratio (95% CI) of markers rs6492538 (allele C) and rs16882131 (allele C) found to be associated in the pooled analysis. Odds ratios are shown for each European population under study and in the combined sample (Mantel–Haenszel meta-analysis). *P*-values from the meta-analysis are indicated at the bottom for each SNP.

Table II. Haplotype analysis of block 2 (miR-206/miR133b cluster) in 1,085 ASD patients and 1,088 controls.

Allele combinations[a]	Cases (%)	Controls (%)	*P*-value (permP)	Odds ratio (95% CI)
TCTCCC	835 (43.6)	725 (37.7)	0.0001768 (0.0017)	1.28 (1.25; 1.45)
TCTCCT	482 (25.2)	482 (25.0)	0.9254 (NS)	1 (0.87; 1.16)
TTTTCC	340 (17.7)	399 (20.7)	0.01945 (NS)	0.82 (0.7; 0.97)
CCTCCC	149 (7.7)	175 (9.1)	0.1441 (NS)	0.84 (0.67; 1.06)
TTCCTC	106 (5.5)	141 (7.3)	0.02382 (NS)	0.74 (0.57; 0.96)

A total of 10,000 permutations were performed to correct for multiple testing (permP).
NS, not significant; CI, confidence interval; OR, odds ratio.
[a]Marker haplotypes: rs1892275, rs16882131, rs1537670, rs17578851, rs10484882 and rs3789779.

ated in this cluster ($P < 0.01$) (Figure 4 and Supplementary Table 8 (available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518)). The Tarbase v5 and MiRecords websites were used to select all validated genes regulated by the eight miRNAs belonging to the two associated clusters. In the resulting list of 44 genes several GO categories were overrepresented and some of them may be relevant for ASD, such as "axon" (*ad-P* = 1.8E-06), "neuron projection" (*ad-P* = 1.6E-05) and "cell migration" (*ad-P* = 8.7E-12), among others (Supplementary Table 9 available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518).

We also performed mutational screening of the miRNA genes from the associated clusters to investigate the putative presence of pathogenic rare variants. A Spanish sample of 241 ASD patients was screened for these pre-miRNA genes: miR-133b, miR-206, miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1 and miR92a-1. The only rare variant found in this sample was identified in a proband and consisted of a T > C transition in miR-133b at position chr6:52013809, predicted to be irrelevant in the miRNA secondary structure ($\Delta\Delta G = 0$, Table III). We also investigated the presence of rare variants in 701 pre-miRNA genes sequenced by WES in a sample of 101 individuals from 30 ASD families. The filtered and Sanger-validated variants are listed in Table III. All the rare variants identified were heterozygous. The number of rare variants transmitted from parents to affected probands in miRNA genes was 18, whereas the number of untransmitted variants was 12 (Supplementary Table 10 available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518). However, the difference between the number of transmitted and untransmitted miRNA variants considering the total number of rare variants in both groups was not significant ($P = 0.22$). We explored the potential involvement in autism of the pool of transmitted rare variants of miRNA genes considering the category "Diseases and Disorders" from the IPA "top biological functions". The most significantly overrepresented categories in our list of miRNA genes were "hereditary disorder" and "developmental disorder", among others (Table IV). The same two categories were found among the top hits when we compared, with IPA, the lists of miRNA genes including variants that were transmitted and those that were not transmitted, selecting the comparisons that showed the largest divergence: "hereditary disorder" (transmitted: 5.63E-07–1.49E-02; untransmitted: 9.58E-05–2.15E-02), "developmental disorder" (transmitted: 6.01E-07–1.49E-02; untransmitted: 9.58E-05–2.15E-02) (Supplementary Table 11 available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518).

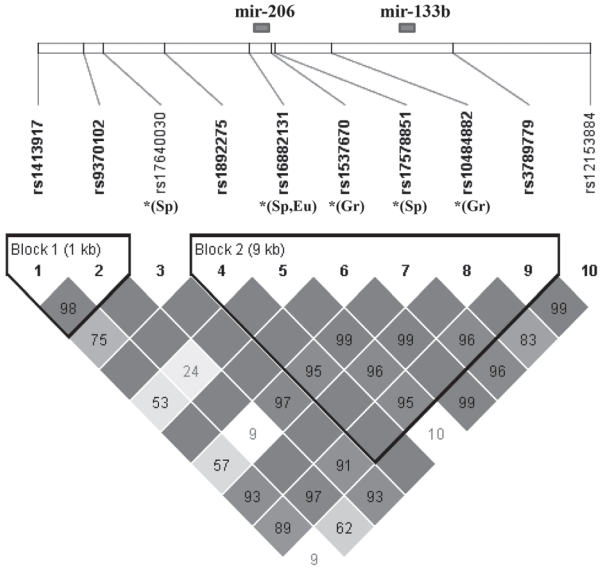

Figure 4. Linkage disequilibrium plot for the miR-206/miR-133b cluster calculated using the whole sample under study. D′ values are indicated (tones from white to dark grey indicate D′ values from 0 to 1). The genomic locations of the miRNAs is indicated as grey blocks at the top. The SNPs with *P*-value < 0.01 are indicated with an asterisk (*). In parenthesis, the population where association with ASD was found (Sp, Spanish; Gr, German; Eu, total European sample).

Table III. Rare variants identified in miRNA genes by whole-exome sequencing (WES) in 41 ASD probands from multiplex and singleton families.

Chr	Position	Expected	Found	Pre-miRNA gene	Family type	ΔΔG (WT-Mut)	Conservation	Variant present in the mature miRNA
1	231155581	G	A	miR-1182	S	−0.9	No	No
4	38869742	A	G	miR-574[&]	M-2	0.4	No	No
4	113569539	CACTTACTT	CACTT	**miR-302C**	M-1	2.9	Yes	Yes
5	95414907	C	T	**miR-583**	S	2.4	Yes	No
5	100152227	A	G	miR-548P	S	0	No	No
6	52013809	T	C	miR-133b°	S	0	Yes	No
7	5535469	G	A	**miR-589**	M-1	−4.8	Yes	*
7	129414541	T	G	miR-96	S	1	Yes	*
9	126164854	A	G	miR-601	S	−2.1	No	Yes
13	50623110	T	C	miR-16-1	M-1	0.6	No	No
14	101500166	C	T	miR-495	S	0	No	No
14	101340858	CTT	CTTT	miR-337	M-2	−1	Yes	Yes (miR-337-3p)
14	101530906	T	G	**miR-541**	S	4.1	Yes	Yes
15	63116185	A	G	miR-190°	S	0	Yes	Yes
20	34041869	A	C	miR-1289-1	S	3.8	No	No
20	62572830	C	T	**miR-1914**	S	4.6	Yes	*
22	23165356	C	T	**miR-650[&]**	S	2	Yes	No
X	151128107	A	T	**miR-452**	M-1	4	Yes	No
X	146342088	CCAATC	CCAATCTCA	**miR-509-1**	M-1	3.6	Yes	Yes (miR-509-3p)

The only variant identified in the screening of mutations in the two miRNA clusters found associated with ASD is also included (miR-133b).
M, multiplex family, -1, only in one affected sib, -2, in both affected sibs. S, singleton family.
ΔΔG (WT-Mut), Difference in free energy change between wild-type and mutant.
*, Complementary region of the mature miRNA.
°, variant identified in the targeted mutation screening of the associated clusters;
&, the same variant does not segregate with ASD in other families (Supplementary Table 9).
MiRNAs in bold are those with variants showing strong conservation and a significant change in the prediction of the secondary structure (ΔΔG > 2).

The potential pathogenic role of the miRNA rare variants transmitted to the affected probands was evaluated by considering: (1) the change in free energy for the secondary structure between the wild-type and the mutated pre-miRNA gene; (2) the nucleotide conservation status; and (3) the location within the mature miRNA or seed region. A schematic representation of the pre-miRNA secondary structures (loop, mature miRNA, seed region and MIR*) is shown in Supplementary Figure 1 (available online at http://informahealthcare.com/doi/abs/10.3109/15622975.2015.1029518). However, no variants were found in the miRNA seed region in our study. The most interesting changes with a potential pathogenic effect were found in the following miRNA genes: miR-302C, miR-583, miR-589, miR-541, miR-1914, miR-650, miR-452 and miR-509-1 (Table III).

Discussion

Common and rare variants of miRNAs may have an important impact on the fine-tuning of nervous system pathways and contribute to genetic burden in psychiatric disorders. Although several studies in ASD identified miRNAs differentially expressed in blood and brain (Abu-Elneel et al. 2008; Talebizadeh et al. 2008; Sarachana et al. 2010; Ghahramani et al.

Table IV. Disease and disorders found overrepresented in the "top biological function" of ingenuity pathway analysis (IPA).

Diseases and disorders	Number of miRNA genes included	*P*-value range
Hereditary disorder	6	5.63E-07 – 1.49E-02
Skeletal and muscular disorders	4	5.63E-07 – 4.03E-02
Developmental disorder	5	6.01E-07 – 1.49E-02
Connective tissue disorders	3	1.12E-05 – 2.77E-02
Inflammatory disease	5	1.12E-05 – 2.77E-02

2011), no study to date has explored the genetic variability of this class of genes. Here, we present the first comprehensive study of common and rare variants of miRNA genes potentially involved in ASD.

Common miRNA variants were explored through a two-stage case–control association study. The discovery study suggested association of miRNA genes or clusters with ASD (Figure 1). However, none of them could be replicated in the follow-up sample. In order to increase the statistical power of our study we performed a joint analysis of 28 SNPs genotyped in both stages, consisting of 1,085 cases and 1,088 controls (Nsengimana and Bishop 2012). We identified two SNPs in two different clusters showing a significant association with ASD that survived correction for multiple testing: rs16882131 (miR-133b/miR-206 cluster) and rs6492538 (miR-17/miR-18a/miR-19a/miR-20a/miR-19b-1/miR92a-1 cluster). MiR-133b and miR-206, located in the same cluster, are strong candidate genes for ASD on the basis of their functional role. Both miRNAs regulate the proto-oncogene *MET* (Taulli et al. 2009; Yan et al. 2009; Hu et al. 2010), which has consistently been associated with autism (Campbell et al. 2006; Sousa et al. 2009; Thanseem et al. 2010). Although the miR-133b/miR-206 cluster was initially considered to have muscle-specific expression, recent studies indicate that miR-133b regulates differentiation and maturation of dopaminergic neurons (Hebert and de Strooper 2009), whereas miR-206 regulates *BDNF*, a key regulator of synaptic plasticity (Lee et al. 2012), and is highly expressed in rat cerebellum (Olsen et al. 2009). Moreover, genetic studies have reported an association of rs17578796, located in the vicinity of miR-206, with schizophrenia (Hansen et al. 2007). This variant was not included in our study, but is located in the miR-133b/miR-206 LD block 2, which showed a suggestive associated haplotype (Table II and Figure 4). It is noteworthy that the association in this cluster was not indexed by the same variant in the Spanish and German populations, but driven by flip-flopping associated markers. However, rs16882131 was the only marker in this cluster that survived multiple testing. This variant, 214 bp upstream from the pre-miRNA-206 gene, is located in a potential regulatory region and may be implicated in the complex transcriptional elements that regulate the uncoupled expression of mir-206 and miR-133b (Cesana et al. 2011). We also found a significant association for rs6492538 in cluster miR-17/miR-18a/miR-19a/miR-20a/miR-19b-1/miR92a-1. This variant is located approximately 9 kb upstream from the miRNA genes.

The associations found in these clusters may highlight regulatory allelic variants altering expression, transcript tertiary structure, processing and transport of the miRNA (Wu et al. 2008). Alterations in miRNA dosages of these two clusters may perturb neuronal homeostasis as shown by the enriched GO categories of the validated target genes of these miRNAs. Whereas SNPs flanking miRNA genes may have a regulatory impact on their expression, variants in the pre-miRNA may have strong consequence for hairpin stability, or abolish interactions with target gene sites when they occur in the functional seed regions.

Genes where common variation has been identified as a risk factor for ASD could also be involved in the disorder by means of rare variation. Thus, we sequenced the eight pre-miRNA genes from the two associated clusters in the Spanish sample, the only cohort showing association signals in both clusters. We identified only one variant in the 133b mature miRNA gene, predicted to be neutral.

The first Mendelian disease caused by a mutation in the seed region of a miRNA, responsible for nonsyndromic progressive hearing loss, was reported a few years ago (Mencia et al. 2009). Point mutations in the miRNA seed regions may have dramatic consequences by abolishing the interaction between the miRNA and tens or hundreds of target genes, misregulating specific physiological pathways. In this study we used exome sequencing data from 101 individuals, including 41 autistic probands and their parents, to explore the link of ASD with rare aetiological variants in 701 pre-miRNA genes, especially in their seed regions. The analysis of miRNA-associated categories showed that the pool of miRNAs bearing rare variants that are transmitted to ASD siblings may have interesting implications for autism. However, we failed to identify variants in the seed regions, although several changes were found in mature miRNAs or were predicted to alter hairpin stability. Recently, rare and common variants in pre-miRNA genes, but located outside the seed regions, were reported to affect the correct maturation and were associated with Mendelian or complex diseases (Li et al. 2010; Solda et al. 2012). In our study we found several rare variants in pre-miRNA genes that may have a functional impact on miRNA processing (Table III). For most of these miRNAs no data were available concerning their target genes or their function. The most interesting result was a mutation in a conserved position of miR-541, predicted to decrease miRNA stability by modifying the secondary structure due to the enlargement of a bulge on the hairpin (Supplementary Figure 2 available online at http://informahealthcare.com/doi/abs/10.

3109/15622975.2015.1029518). This variant also flanks a Drosha cleavage site that may affect the actual mature miRNA cleavage (Li et al. 2010). MiR-541 belongs to the brain-specific cluster, miR-379-410 (Fiore et al. 2009), and has been reported to be involved in neurite outgrowth by regulating the X-linked gene *SYN1* (Zhang et al. 2011). Recently, loss-of-function mutations in *SYN1* were reported in individuals with autism and epilepsy (Fassio et al. 2011) Thus, variants affecting correct miR-541 processing may also be involved in autism susceptibility. Notwithstanding, functional studies are needed to establish whether this variant has a specific functional role.

In summary, here we present the first genetic study of ASD that focuses on miRNA gene variability. The results show that common allelic variants of two clusters (miR-206/miR-133b and miR-17/miR-18a/miR-19a/miR-20a/miR-19b-1/miR92a-1) and rare variants of mature miRNA genes may confer susceptibility to ASD. Further replication studies and inspection of miRNA genes in larger exome or genome datasets are warranted to elucidate the role of this crucial class of regulatory genes in this disorder.

Acknowledgments

We are grateful to all the patients and their families for their kind participation. We thank all the clinical collaborators that contributed to the diagnosis of the probands. We would like to thank M. Monfort and M. Torres for technical support at the Spanish National Genotyping Center (CeGen) in Barcelona and in Santiago de Compostela, respectively and M. Bayés from the National Center for Genomic Analysis (CNAG), Barcelona. Financial support was received from "Fundació La Marató de TV3" (092010), "Fundación Alicia Koplowitz", AGAUR (2014SGR932, 2014SGR1468), the Spanish "Ministerio de Economía y Competitividad" (SAF2012-33484, SAF2010-21165), the Spanish "Fondo de Investigación Sanitaria" (FIS PI-1002512 and PI-1302841), the University of Bologna (RFO), Saarland University (grant #T2042103-01 to C. Freitag), the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement numbers 278948 (TACTICS) and 602805 (AGGRESSOTYPE), the European Community's H2020 Research and Innovation Program under the Marie Sklodowska-Curie grant agreement number 643051 (MiND), and the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115300. Collecting the Dutch samples was supported by grants from Karakter Child and Adolescent Psychiatry (to J. Buitelaar) and a grant by the Netherlands Organisation for Scientific Research (NWO grant #91610024 to N. Rommelse). CT was supported by the European Union (Marie Curie, PIEF-GA-2009-254930) and BT by AGAUR (Generalitat de Catalunya).

Statement of Interest

None to declare.

References

Abu-Elneel K, Liu T, Gazzaniga FS, Nishimura Y, Wall DP, Geschwind DH, et al. 2008. Heterogeneous dysregulation of microRNAs across the autism spectrum. Neurogenetics 9(3):153–161.

Bacchelli E, Battaglia A, Cameli C, Lomartire S, Tancredi R, Thomson S, et al. 2015. Analysis of *CHRNA7* rare variants in autism spectrum disorder susceptibility. Am J Med Genet A 167A:715–723. doi:10.1002/ajmg.a.36847

Barrett JC, Fry B, Maller J, Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21(2):263–265.

Campbell DB, Sutcliffe JS, Ebert PJ, Militerni R, Bravaccio C, Trillo S, et al. 2006. A genetic variant that disrupts MET transcription is associated with autism. Proc Natl Acad Sci USA 103(45):16834–16839.

Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. 2011. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell 147(2):358–369.

Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, et al. 2005. The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. Science 310(5755): 1817–1821.

Fassio A, Patry L, Congia S, Onofri F, Piton A, Gauthier J, et al. 2011. SYN1 loss-of-function mutations in autism and partial epilepsy cause impaired synaptic function. Hum Mol Genet 20(12):2297–2307.

Fiore R, Khudayberdiev S, Christensen M, Siegel G, Flavell SW, Kim TK, et al. 2009. Mef2-mediated transcription of the miR379-410 cluster regulates activity-dependent dendritogenesis by fine-tuning Pumilio2 protein levels. EMBO J 28(6): 697–710.

Friedman RC, Farh KK, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19(1):92–105.

Ghahramani Seno MM, Hu P, Gwadry FG, Pinto D, Marshall CR, Casallo G, et al. 2011. Gene and miRNA expression profiles in autism spectrum disorders. Brain Res 1380:85–97.

Gonzalez JR, Armengol L, Sole X, Guino E, Mercader JM, Estivill X, et al. 2007. SNPassoc: an R package to perform whole genome association studies. Bioinformatics 23(5): 644–645.

Griggs EM, Young EJ, Rumbaugh G, Miller CA. 2013. MicroRNA-182 regulates amygdala-dependent memory formation. J Neurosci 33(4):1734–1740.

Hansen T, Olsen L, Lindow M, Jakobsen KD, Ullum H, Jonsson E, et al. 2007. Brain expressed microRNAs implicated in schizophrenia etiology. PLoS One 2(9):e873.

Hebert SS, De Strooper B. 2009. Alterations of the microRNA network cause neurodegenerative disease. Trends Neurosci 32(4):199–206.

Hu G, Chen D, Li X, Yang K, Wang H, Wu W. 2010. miR-133b regulates the MET proto-oncogene and inhibits the growth of colorectal cancer cells in vitro and in vivo. Cancer Biol Ther 10(2):190–197.

Krichevsky AM, King KS, Donahue CP, Khrapko K, Kosik KS. 2003. A microRNA array reveals extensive regulation of microRNAs during brain development. RNA 9(10):1274–1281.

Lee ST, Chu K, Jung KH, Kim JH, Huh JY, Yoon H, et al. 2012. miR-206 regulates brain-derived neurotrophic factor in Alzheimer disease model. Ann Neurol 72(2):269–277.

Lett TA, Chakravarty MM, Felsky D, Brandl EJ, Tiwari AK, Goncalves VF, et al. 2013. The genome-wide supported microRNA-137 variant predicts phenotypic heterogeneity within schizophrenia. Mol Psychiatry 18(4):443–450.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14): 1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078–2079.

Li X, Jin P. 2010. Roles of small regulatory RNAs in determining neuronal identity. Nat Rev Neurosci 11(5):329–338.

Li L, Meng T, Jia Z, Zhu G, Shi B. 2010. Single nucleotide polymorphism associated with nonsyndromic cleft palate influences the processing of miR-140. Am J Med Genet A 152(4):856–862.

Lin CH, Yeakley JM, McDaniel TK, Shen R. 2009. Medium- to high-throughput SNP genotyping using VeraCode microbeads. Methods Mol Biol 496:129–142.

Lord C, Risi S, Lambrecht L, Cook EH, Jr, Leventhal BL, DiLavore PC, et al. 2000. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 30(3):205–223.

Lord C, Rutter M, Le Couteur A. 1994. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 24(5):659–685.

Medina I, Carbonell J, Pulido L, Madeira SC, Goetz S, Conesa A, et al. 2010. Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. Nucleic Acids Res 38(Web Server issue):W210–213.

Mencia A, Modamio-Hoybjor S, Redshaw N, Morin M, Mayo-Merino F, Olavarrieta L, et al. 2009. Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. Nat Genet 41(5):609–613.

Monteys AM, Spengler RM, Wan J, Tecedor L, Lennox KA, Xing Y, et al. 2010. Structure and activity of putative intronic miRNA promoters. RNA 16(3):495–505.

Nsengimana J, Bishop DT. 2012. Design considerations for genetic linkage and association studies. Methods Mol Biol 850:237–262.

Olsen L, Klausen M, Helboe L, Nielsen FC, Werge T. 2009. MicroRNAs show mutually exclusive expression patterns in the brain of adult male rats. PLoS One 4(10):e7225.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Puente XS, Pinyol M, Quesada V, Conde L, Ordonez GR, Villamor N, et al. 2011. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature 475(7354):101–105.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81(3):559–575.

Sarachana T, Zhou R, Chen G, Manji HK, Hu VW. 2010. Investigation of post-transcriptional gene regulatory networks associated with autism spectrum disorders by microRNA expression profiling of lymphoblastoid cell lines. Genome Med 2(4):23.

Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. 2011. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 43(10): 969–76.

Sempere LF, Freemantle S, Pitha-Rowe I, Moss E, Dmitrovsky E, Ambros V. 2004. Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. Genome Biol 5(3):R13.

Solda G, Robusto M, Primignani P, Castorina P, Benzoni E, Cesarani A, et al. 2012. A novel mutation within the MIR96 gene causes non-syndromic inherited hearing loss in an Italian family by altering pre-miRNA processing. Hum Mol Genet 21(3):577–585.

Sousa I, Clark TG, Toma C, Kobayashi K, Choma M, Holt R, et al. 2009. MET and autism susceptibility: family and case-control studies. Eur J Hum Genet 17(6):749–758.

Talebizadeh Z, Butler MG, Theodoro MF. 2008. Feasibility and relevance of examining lymphoblastoid cell lines to study role of microRNAs in autism. Autism Res 1(4):240–250.

Taulli R, Bersani F, Foglizzo V, Linari A, Vigna E, Ladanyi M, et al. 2009. The muscle-specific microRNA miR-206 blocks human rhabdomyosarcoma growth in xenotransplanted mice by promoting myogenic differentiation. J Clin Invest 119(8): 2366–2378.

Thanseem I, Nakamura K, Miyachi T, Toyota T, Yamada S, Tsujii M, et al. 2010. Further evidence for the role of MET in autism susceptibility. Neurosci Res 68(2):137–141.

Toma C, Hervás A, Balmaña N, Salgado M, Maristany M, Vilella E, et al. 2013. Neurotransmitter systems and neurotrophic factors in autism: association study of 37 genes suggests involvement of DDC. World J Biol Psychiatry 14(7): 516–27.

Van Steijn DJ, Richards JS, Oerlemans AM, de Ruiter SW, van Aken MA, Franke B, et al. 2012. The co-occurrence of autism spectrum disorder and attention-deficit/hyperactivity disorder symptoms in parents of children with ASD or ASD with ADHD. J Child Psychol Psychiatry 53(9):954–963.

Wang G, Wang Y, Shen C, Huang YW, Huang K, Huang TH, et al. 2010. RNA polymerase II binding patterns reveal genomic regions involved in microRNA gene regulation. PLoS One 5(11):e13798.

Wickham. 2009. ggplot2: elegant graphics for data analysis. New York: Springer.

Wu M, Jolicoeur N, Li Z, Zhang L, Fortin Y, L'Abbe D, et al. 2008. Genetic variations of microRNAs in human cancer and their effects on the expression of miRNAs. Carcinogenesis 29(9):1710–1716.

Yan D, Dong Xda E, Chen X, Wang L, Lu C, Wang J, et al. 2009. MicroRNA-1/206 targets c-Met and inhibits rhabdomyosarcoma development. J Biol Chem 284(43):29596–29604.

Zhang J, Liu LH, Zhou Y, Li YP, Shao ZH, Wu YJ, et al. 2011. Effects of miR-541 on neurite outgrowth during neuronal differentiation. Cell Biochem Funct 29(4):279–286.

Supplementary material available online

Supplementary Tables 1 to 11 and Figures 1, 2.