

Association study of schizophrenia with variants in miR-137 binding sites

^{1,2}David Curtis, ^{1,3,4}Warren Emmett

1. UCL Genetics Institute, University College London
2. Centre for Psychiatry, Barts and the London School of Medicine and Dentistry
3. The Francis Crick Institute
4. Department of Molecular Neuroscience, UCL Institute of Neurology

Correspondence: David Curtis, Darwin Building, Gower Street, London WC1E 6BT, UK. Phone 00 44 20 7679 2212. Fax 00 44 20 7791 5201. d.curtis@ucl.ac.uk

Conflict of interest

The authors declare they have no conflict of interest.

Abstract

There is strong cumulative evidence for the involvement of miR-137 and its targets in the aetiology of schizophrenia. Here we test whether variants, especially rare variants, in miR-137 binding sites are associated with schizophrenia in an exome-sequenced sample of 4225 cases and 5834 controls. Only a small proportion of binding sites were covered by the capture system which had been used. A weighted burden test using the 372 detected variants demonstrated an excess among cases significant at $p=0.024$. The sample size is too small to implicate individual variants or genes but overall this finding does provide some further support for the hypothesis that disruption of miR-137 binding sites can increase the risk of schizophrenia, perhaps by leading to over-expression of the target gene. We recommend that future exome sequencing studies should cover the untranscribed regions of genes, which contain the microRNA binding sites, in order that this potentially important pathogenic mechanism can be adequately investigated.

Keywords

Schizophrenia; miR-137; microRNA.

Introduction

MicroRNAs are small RNA molecules which can bind to specific sites in the 3' untranslated region (UTR) of transcripts of certain other genes, and these genes are described as the targets of the microRNA. As discussed recently (Olde Loohuis et al., 2017) markers for both the gene for the microRNA miR-137 and for the genes it targets demonstrate association with schizophrenia (Kwon et al., 2013; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).

Experimentally, reducing or increasing miR-137 expression in rat hippocampal neurons identifies sets of regulated genes involved in neurodevelopmental processes and neuronal maturation, supporting the proposition that it acts as a critical gene network hub contributing to the pathophysiology of schizophrenia (Olde Loohuis et al., 2017). When microRNA binds to the 3' UTR of a gene it suppresses transcription, so a variant in the binding site could interfere with this process and lead to over-expression. Here, we investigate whether variants in the binding sites of miR-137 are more commonly found in exome-sequenced schizophrenia cases than controls.

Material and methods

The dataset analysed was downloaded from the dbGaP data repository and consisted of whole exome sequence variants from a Swedish schizophrenia association study containing 4968 cases and 6245 controls (Genovese et al., 2016). As described elsewhere (Curtis, 2017), the dataset was subjected to quality control (QC) procedures including the removal of subjects who appeared to have a substantial Finnish component to their ancestry to leave a sample of 4225 cases and 5834 controls which appeared to be ethnically well-matched. When carrying out exome-sequencing, only the coding parts of the genome are sequenced using a capture procedure which is intended to cover the exons of all genes and which may partially cover the neighbouring UTRs. As reported in the original paper, the subjects were sequenced in 12 waves. For all but the first wave the Agilent SureSelect Human All Exon v.2 Kit was used for the hybrid-capture procedure whereas for the first wave an earlier version was used. Predicted binding sites for miR-137 were obtained from microRNA.org (Betel et al., 2010). Excluding the Y chromosome, there were 8139 predicted binding sites, of which 1139 were covered by SureSelect. Variants in these regions were extracted and analysed by SCOREASSOC, which performs a weighted

burden test such that very rare variants are weighted 10 times higher than a common variant with observed MAF=0.5 (Curtis, 2012). Each subject receives a score consisting of the weighted sum of the variant alleles possessed by that subject at all sites and a t test is used to compare the scores between cases and controls.

Results

In the binding sites covered, 372 variants were found which passed QC procedures. On average, the weighted burden test scores were indeed higher for cases than controls ($t=2.3$, 10057 df, $p=0.024$), indicating that subjects with schizophrenia on average are more likely to have variants, especially rare variants, in miR-137 binding sites than controls. However the effect is at best modest and most variants occurred in only one or two subjects, while for others there were only small differences in frequencies between cases and controls. Thus, if the effect is real it is not possible to implicate specific genes or variants. For the purposes of illustration, it may be worth noting the results for two genes for which the difference was most marked, as shown in Table 1. A variant at 10:106027165, rs7589 in the 3' UTR of GSTO1, was heterozygous in 5 out of 4224 cases and none out of 5833 controls while the adjacent variant at 10:106027166, rs375727967, was heterozygous in a single additional case. A variant at 19:58773876, in a miR137 binding site in the 3'UTR of ZNF544, was heterozygous in 17 out of 4205 cases and 13 out of 5814 controls, yielding an odds ratio of 1.8. Another variant in the same binding site at 19:58773869, rs199600767, was heterozygous in an additional 2 cases and 1 control. The results for all variants are presented in Supplementary Table 1.

Discussion

Given that the weighted burden test is statistically significant, our results do provide some further support for the hypothesis that abnormalities in miR-137 functionality could be involved in the aetiology of schizophrenia. Disruption of a microRNA binding site could lead to increased gene expression, providing a mechanism for a dominant, gain of function effect. However, we would emphasise that the results are by no means conclusive and we did not observe a large excess of variants among cases. With the sample sizes used it is not possible to assign risk to individual variants and most, including

those in *GSTO1* and *ZNF544*, are too rare to have been imputed in large GWAS samples (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). All six subjects with a variant in the *GSTO1* binding site have schizophrenia. This gene codes for an omega class glutathione S-transferase and a previous study showed that patients with schizophrenia have reduced glutathione levels in cerebrospinal fluid (Do et al., 2000), although a recent association study of schizophrenia with markers for *GSTO1* and other glutathione related genes was negative (Matsuzawa et al., 2009). These negative results would be expected if only rare variants at critical locations, including perhaps microRNA binding sites, exerted a substantial effect. However, as Table 1 shows, the variant which occurs in five cases, rs7589, is very common in African subjects in ExAC, making it implausible that it confers a major risk for schizophrenia. *ZNF544* belongs to the C2H2-type zinc-finger family and is involved in gene transcription. In a genome-wide association study of ADHD traits an intronic SNP, rs260461, was significant at $p=10^{-5}$ (Lasky-Su et al., 2008) and in a methylome-wide study a CpG island near *ZNF544* was found to be hyper-methylated at birth in subjects with a high trajectory for ADHD symptoms (Walton et al., 2017) but there does not seem to be other external evidence to implicate this gene. We certainly do not make any strong claim that our study provides much support for the theory that disturbances of either glutathione metabolism or *ZNF544* function are involved in schizophrenia aetiology but rather we highlight these results to illustrate the nature of the findings produced by studies of very rare variants detected by sequencing.

We believe that the hypothesis that disruption of miR-137 binding sites could increase the risk of schizophrenia is attractive. However, with the most informative data currently available we can obtain at best modest support for this. Although larger sample sizes might provide more conclusive results, a major limitation of the current study is that only a minority of binding sites were covered by the capture procedure. We strongly recommend that future exome sequencing studies should routinely obtain full coverage for UTRs as well as exons so that the role of variants potentially affecting microRNA binding can be properly evaluated.

Acknowledgments

The datasets used for the analysis described in this manuscript were obtained from dbGaP at <http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession number *phs000473.v2.p2*. Samples used for data analysis were provided by the Swedish Cohort Collection supported by the NIMH grant R01MH077139, the Sylvan C. Herman Foundation, the Stanley Medical Research Institute and The Swedish Research Council (grants 2009-4959 and 2011-4659). Support for the exome sequencing was provided by the NIMH Grand Opportunity grant RCMH089905, the Sylvan C. Herman Foundation, a grant from the Stanley Medical Research Institute and multiple gifts to the Stanley Center for Psychiatric Research at the Broad Institute of MIT and Harvard. The work of W.E. is supported by the Wellcome Trust (103760/Z/14/Z) and the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001002), the UK Medical Research Council (FC001002), and the Wellcome Trust (FC001002). The projects which produced the datasets studied were subject to ethical approval and no separate ethical approval was required for the analyses presented here.

Author contributions

DC and WE jointly formulated the study design and DC carried out the analyses.

Conflict of interest

The authors declare they have no conflict of interest.

References

- Betel, D., Koppal, A., Agius, P., Sander, C., Leslie, C., 2010. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.* 11, R90. <https://doi.org/10.1186/gb-2010-11-8-r90>
- Curtis, D., 2017. Construction of an exome-wide risk score for schizophrenia based on a weighted burden test. *Ann. Hum. Genet.*
- Curtis, D., 2012. A rapid method for combined analysis of common and rare variants at the level of a region, gene, or pathway. *Adv Appl Bioinform Chem* 5, 1–9.
- Deeks, J., Higgins, J.P.T., 2007. *Statistical Algorithms in Review Manager 5*. [WWW Document]. URL https://www.researchgate.net/publication/252093205_Statistical_Algorithms_in_Review_Manager_5
- Do, K.Q., Trabesinger, A.H., Kirsten-Krüger, M., Lauer, C.J., Dydak, U., Hell, D., Holsboer, F., Boesiger, P., Cuénod, M., 2000. Schizophrenia: glutathione deficit in cerebrospinal fluid and prefrontal cortex in vivo. *Eur. J. Neurosci.* 12, 3721–8.
- Genovese, G., Fromer, M., Stahl, E.A., Ruderfer, D.M., Chambert, K., Landén, M., Moran, J.L., Purcell, S.M., Sklar, P., Sullivan, P.F., Hultman, C.M., McCarroll, S.A., 2016. Increased burden of ultra-rare

protein-altering variants among 4,877 individuals with schizophrenia. *Nat. Neurosci.* 19, 1433–1441. <https://doi.org/10.1038/nn.4402>

Kwon, E., Wang, W., Tsai, L.-H., 2013. Validation of schizophrenia-associated genes CSMD1, C10orf26, CACNA1C and TCF4 as miR-137 targets. *Mol. Psychiatry* 18, 11–2. <https://doi.org/10.1038/mp.2011.170>

Lasky-Su, J., Neale, B.M., Franke, B., Anney, R.J.L., Zhou, K., Maller, J.B., Vasquez, A.A., Chen, W., Asherson, P., Buitelaar, J., Banaschewski, T., Ebstein, R., Gill, M., Miranda, A., Mulas, F., Oades, R.D., Roeyers, H., Rothenberger, A., Sergeant, J., Sonuga-Barke, E., Steinhausen, H.C., Taylor, E., Daly, M., Laird, N., Lange, C., Faraone, S. V., 2008. Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* 147B, 1345–1354. <https://doi.org/10.1002/ajmg.b.30867>

Lek, M., Karczewski, K.J., Minikel, E. V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D.N., Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B., Won, H.-H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G., Exome Aggregation Consortium, O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D.N., Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B., Won, H.-H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G., Exome Aggregation Consortium, 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291. <https://doi.org/10.1038/nature19057>

Matsuzawa, D., Hashimoto, K., Hashimoto, T., Shimizu, E., Watanabe, H., Fujita, Y., Iyo, M., 2009. Association study between the genetic polymorphisms of glutathione-related enzymes and schizophrenia in a Japanese population. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* 150B, 86–94. <https://doi.org/10.1002/ajmg.b.30776>

Olde Loohuis, N.F.M., Nadif Kasri, N., Glennon, J.C., van Bokhoven, H., Hébert, S.S., Kaplan, B.B., Martens, G.J.M., Aschrafi, A., 2017. The schizophrenia risk gene MIR137 acts as a hippocampal gene network node orchestrating the expression of genes relevant to nervous system development and function. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* 73, 109–118. <https://doi.org/10.1016/j.pnpbp.2016.02.009>

Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511, 421–427. <https://doi.org/10.1038/nature13595>

Walton, E., Pingault, J.-B., Cecil, C.A.M., Gaunt, T.R., Relton, C.L., Mill, J., Barker, E.D., 2017.

Epigenetic profiling of ADHD symptoms trajectories: a prospective, methylome-wide study. *Mol. Psychiatry* 22, 250–256. <https://doi.org/10.1038/mp.2016.85>

Table 1. Genotype counts and allele frequencies for variants in miR137 binding sites in the 3' UTRs of *GTSO1* and *ZNF544*. No subject was heterozygous for more than one variant. WT = homozygous wild type, het = heterozygous. (Totals are less than the overall sample sizes because some individual genotypes failed QC.) Odds ratios (OR) with standard errors were calculated according to the Statistical Algorithms in Review Manager 5 (Deeks and Higgins, 2007). Also shown are variant allele frequencies in non-Finnish European and African populations, as provided by ExAC (Lek et al., 2016).

	Position	Controls WT	Controls het	Cases WT	Cases het	OR (95% CI)	MAF, European (Non- Finnish)	MAF, African
<i>GTSO1</i>								
rs7589	10:106027165	5833	0	4219	5	15.2 (0.8-275.1)	0.00082	0.26
rs375727967	10:106027166	5833	0	4223	1	4.1 (0.2-101.7)	0.000030	0
<i>ZNF544</i>								
rs199600767	19: 58773869	5812	1	4200	2	2.8 (0.3-30.5)	0.00013	0
rs200237099	19: 58773876	5801	13	4188	17	1.8 (0.9-3.7)	0.0012	0