Schizophrenia Research xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Rare variant analysis in multiply affected families, association studies and functional analysis suggest a role for the *ITGB4* gene in schizophrenia and bipolar disorder

N.L. O'Brien [a,1], A. Fiorentino [a,1,2], D. Curtis [b,c], C. Rayner [b], C. Petrosellini [a], M. Al Eissa [a], N.J. Bass [a], A. McQuillin [a,*], S.I. Sharp [a]

[a] UCL Molecular Psychiatry Laboratory, Division of Psychiatry, University College London, London, UK
[b] UCL Genetics Institute, University College London, London, UK
[c] Centre for Psychiatry, Barts and the London School of Medicine and Dentistry, London, UK

ARTICLE INFO

Article history:
Received 11 December 2017
Received in revised form 22 February 2018
Accepted 1 March 2018
Available online xxxx

Keywords:
Sequencing
Genotyping
Psychosis
Genetic risk

ABSTRACT

Recent results imply that rare variants contribute to the risk of schizophrenia. Exome sequence data from the UK10K project was used to identify three rare, amino acid changing variants in the *ITGB4* gene which segregated with schizophrenia in two families: rs750367954, rs147480547 and rs145976111. Association analysis was carried out in the exome-sequenced Swedish schizophrenia study and in UCL schizophrenia and bipolar cases and controls genotyped for these variants. A gene-wise weighted burden test was performed on a trio sample of schizophrenia cases and their parents. rs750367954 was seen in two Swedish cases and in no controls. The other two variants were commoner in cases than controls in both Swedish and UCL cohort samples and an overall burden test was significant at $p = 0.0000031$. The variants were not observed in the trio sample but *ITGB4* was most highly ranked out of 14,960 autosomal genes in a gene-wise weighted burden test. The effect of rs147480547 and rs145976111 was studied in human neuroblastoma SH-SY5Y cells. Cells transfected with both variants had increased proliferation at both 24 and 48 h ($p = 0.013$ and $p = 0.05$ respectively) compared to those with wild-type *ITGB4*. Taken together, these results suggest that rare variants in *ITGB4* which affect function may contribute to the aetiology of schizophrenia and bipolar disorder.

1. Introduction

The results of exome sequence studies imply a role for rare variants in the aetiology of schizophrenia (SCZ). Most studies to date have implicated pathways and sets of genes rather than individual genes (Curtis and UK10K Consortium, 2016; Genovese et al., 2016; Purcell et al., 2014; Singh et al., 2017), however, a combined analysis of case control and trio sequence data has implicated rare loss of function variants in the *SETD1A* gene ($p = 3.3 \times 10^{-9}$) in the aetiology of SCZ (Singh et al., 2016).

Analysis of sequencing data from multiply affected extended pedigrees can be valuable for identifying extremely rare variants that are transmitted to affected individuals (Curtis, 2011) and this approach has been successful in identifying variants in the *PLD3* (phospholipase D3) gene in late-onset Alzheimer's disease (Cruchaga et al., 2014). Recently a whole genome-sequencing study identified two different rare, protein-truncating variants in the *RBM12* gene cosegregating with psychosis in two different pedigrees (Steinberg et al., 2017). Here we report the discovery and follow-up of rare, protein-changing variants which were observed to cosegegrate with SCZ in subjects from two multiply-affected families which were sequenced as part of the UK10K project (Muddyman et al., 2013; UK10K Consortium et al., 2015).

2. Material and methods

2.1. Subjects

Four independent datasets were used. These comprised:

2.1.1. UK10K cohort

We used British subjects from the UK10K dataset comprising 1392 with SCZ and 982 with severe childhood obesity (SCOOP) not known to have mental illness which have been described elsewhere (Curtis and UK10K Consortium, 2016). UK10K subjects were exome sequenced with coverage of 72× as described in full elsewhere (UK10K Consortium

* Corresponding author at: University College London, Molecular Psychiatry Laboratory, Division of Psychiatry, Gower Street, London WC1E 6BT, UK.
E-mail address: a.mcquillin@ucl.ac.uk (A. McQuillin).
[1] Drs O'Brien and Fiorentino contributed equally to the work presented.
[2] Current address: UCL Institute of Ophthalmology, University College London, London, UK.

https://doi.org/10.1016/j.schres.2018.03.001

et al., 2015). Among the UK10K SCZ subjects were 48 subjects from 16 published and unpublished families multiply affected with SCZ that had been collected at UCL (Kalsi et al., 1994; Sherrington et al., 1988). Exome data was available for between 2 and 5 affected members per pedigree.

2.1.2. Swedish SCZ study

The Swedish SCZ study consisted of the 2545 controls and 2545 cases with SCZ from Sweden for whom whole exome sequence data is available *via* dbGaP (Purcell et al., 2014). Detailed sample descriptions have been provided previously (Purcell et al., 2014).

2.1.3. Bulgarian trio sample

The Bulgarian trio sample consisted of probands with SCZ and their parents (Fromer et al., 2014; Rees et al., 2015). The sample comprised whole exome sequence data from 591 trios, consisting of probands with SCZ and their parents, five of whom were also affected (Fromer et al., 2014; Rees et al., 2015). The short read files were downloaded from dbGaP along with family structure and phenotype information.

2.1.4. UCL case-control samples

The UCL case-control samples sample and the recruitment methods have been previously described (Al Eissa et al., 2017; Fiorentino et al., 2014). Briefly, the sample comprised 1917 bipolar disorder (BP) participants, 1304 SCZ participants and 1348 control participants recruited from the UK.

2.2. Variant selection and bioinformatic prediction of variant impact

Custom-written software was used to annotate and identify rare, possibly functional variants shared between affected pedigrees members in the UCL samples included in UK10K. The impact of these variants was predicted using the PolyPhen-2 (Adzhubei et al., 2013) and Sort Intolerant from Tolerant (SIFT)(Kumar et al., 2009) bioinformatics tools.

2.3. Sanger sequencing in families F047 and F158

Exome sequence data was available for three members of family F158 (subjects 3, 5 and 7) and for two members of family F047 (subjects 3 and 5; Fig. 1) from the UK10K sample. Transmission of rs750367954 (allele T) in family F158 and of rs147480547 (allele A) and rs145976111 (allele T) in family F047 was sought in all additional family members with available DNA by PCR and Sanger sequencing (Fig. 1). There was a limited amount of DNA available for individual 4 from family F047 and it was not possible to verify the genotype status of rs147480547 in this individual (Fig. 1).

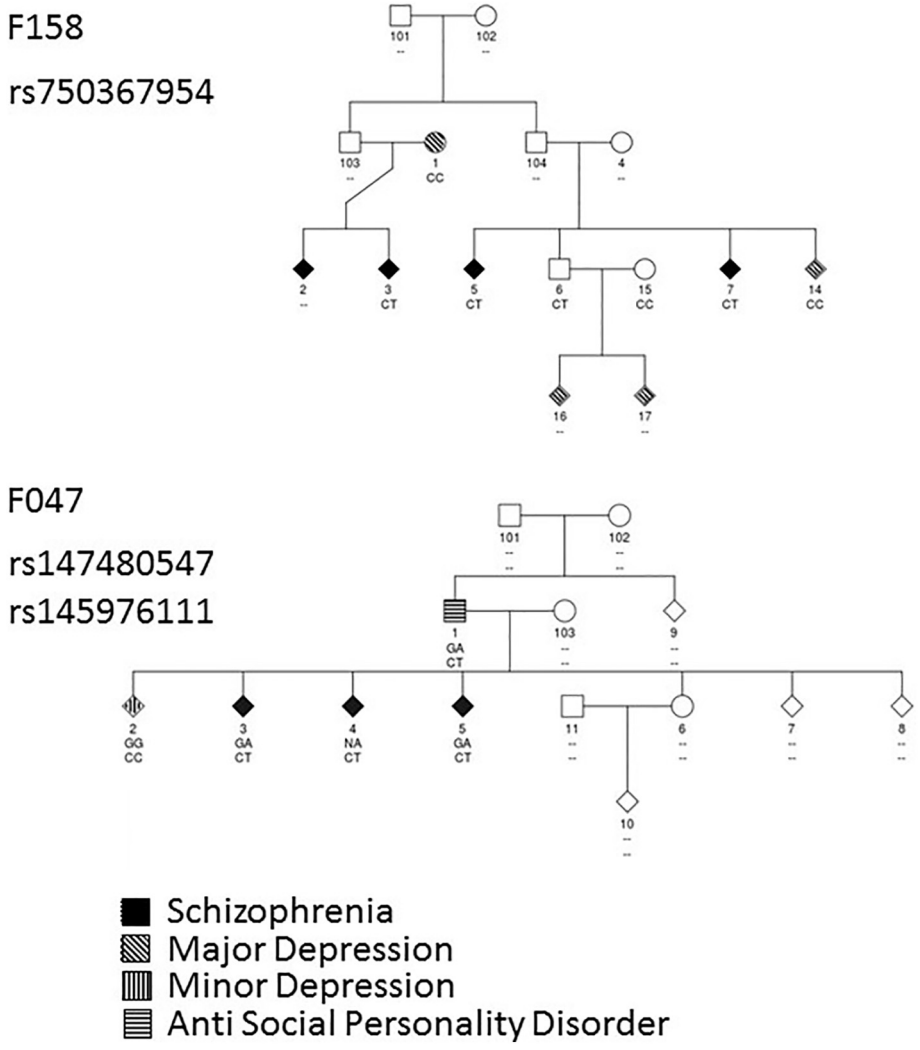

Fig. 1. Non-synonymous *ITGB4* variants in two families multiply affected by schizophrenia and other psychopathology. unk Genotype not determined; filled shapes represent individuals suffering from schizophrenia; shaded shapes represent other psychiatric diagnoses as indicated.

Please cite this article as: O'Brien, N.L., et al., Rare variant analysis in multiply affected families, association studies and functional analysis suggest a role for the *ITGB4* gene in schizophrenia and bipolar disorder..., Schizophr. Res. (2018), https://doi.org/10.1016/j.schres.2018.03.001