Genetic Analysis of the ZNF804A Gene in Mexican Patients With Schizophrenia, Schizoaffective Disorder and Bipolar Disorder

LUCIA MUNCH ANGUIANO  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz  
https://orcid.org/0000-0003-4888-4645

Beatriz Camarena (bcamarenam@gmail.com)  
Departamento de Farmacogenética, Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Mexico City, Mexico  
https://orcid.org/0000-0001-7737-501X

Jesica Nieto-Quinto  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz

Patricia de la Torre  
UNAM Instituto de Investigaciones Biomedicas: Universidad Nacional Autonoma de Mexico Instituto de Investigaciones Biomedicas

Juan Pedro Laclette  
National Autonomous University of Mexico Institute of Biomedical Research: Universidad Nacional Autonoma de Mexico Instituto de Investigaciones Biomedicas

Harumi Hirata-Hernandez  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz

Sandra Hemández-Muñoz  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz

Alejandro Aguilar-Garcia  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz

Claudia Becerra-Palars  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz

Doris Gutierrez-Mora  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz

Hiram Ortega-Ortiz  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz

Raul Escamilla-Orozco  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz

Ricardo Saracco-Alvarez  
Instituto Nacional de Psiquiatria Ramon de la Fuente Muniz

Ismael Bustos-Jaimes  
Universidad Nacional Autonoma de Mexico Facultad de Medicina
Abstract

**Background:** Evidence suggests that Schizophrenia (SCZ), Schizoaffective Disorder (SAD) and Bipolar Disorder (BPD) share genetic risk variants. *ZNF804A* gene has been associated with these disorders in different populations. The aim was to analyze the rs1344706 SNP and identify common and rare variants in a targeted region of the *ZNF804A* gene in SCZ, BPD and SAD Mexican patients compared with a control group.

**Methods:** We genotyped the rs1344706 in 228 Mexican patients diagnosed with SCZ, SAD and BPD, and 295 controls. An additional sample of 167 patients and 170 controls was analyzed to identify rare and common variants using the Sanger-sequence analysis of a targeted region of *ZNF804A* gene.

**Results:** We replicated the association between the rs1344706 and SCZ, SAD and BPD. In the sequence analysis, we did not identify rare variants; however, we identified three common variants: rs3046266, rs1366842 and rs12477430. A comparison of the three identified variants between SCZ, SAD and BPD did not show statistical differences.

**Conclusions:** Our findings support the evidence suggesting that *ZNF804A* gene is involved in the etiology of SZC, SAD and BPD. Future studies are needed to identify the pleiotropic effect of this gene in psychiatric disorders.

Background

Schizophrenia (SCZ), Schizoaffective Disorder (SAD) and Bipolar Disorder (BPD) are chronic and serious mental disorders that have been recognized as an important public health matter. These disorders are important worldwide contributors to years lived with disability, largely due to the young age at onset, high disability burden, and the often-chronic course of illness. Patients with SCZ, SAD and BPD demand a disproportionate amount of health resources and have premature mortality because of comorbid cardiovascular diseases or suicide (Correll et al. 2017).

SCZ and BPD are complex disorders and different genetic studies have found that genes represent around the 80% of the variance of the population at risk. However, the genetic risk variants (either common or rare) of these disorders are still under scrutiny. Previous studies have suggested that there is an overlap between SCZ, SAD and BPD genetic vulnerability, suggesting that these disorders shared genetic risk variants. *ZNF804A* gene has been importantly associated with SCZ, SAD and BPD in different populations, specifically the common variants of the gene (Sun et al. 2015). Nevertheless, rare variants of *ZNF804A* may also play an important role in the risk of developing these disorders. In 2010, a study in UK population did not find a significant association between rare variants of the *ZNF804A* gene and SCZ (Dwyer et al. 2010).

*ZNF804A* gene is known to be involved in different neuronal processes that are distorted in patients with SZC and BPD, such as embryonic neurodevelopment, neuroprotection, neurogenesis, neuronal migration,
synaptic formation and function (Chang et al. 2017; Deans et al. 2016; Hill et al. 2012; Hill & Bray 2012). Studies have found evidence that sustain that ZNF804A protein might be implicated in pre-messenger RNA processing, RNA translational control and other genes expression regulation, including some genes involved in SZC, synaptic formation, neurodevelopment and inflammation (Chapman et al. 2018; Chen et al. 2015; Girgenti et al. 2012; Zhou et al. 2017).

ZNF804A gene is located on chromosome 2q32.1 and is organized in 4 exons that encode for a protein with a zinc finger domain in the N-terminus. This domain allows the protein to bind to DNA, RNA or other proteins so it can regulate gene expression, RNA processing and protein translation. The coding sequence for ZNF804A gene consists in 542 nt in exon 1, 114 nt in exon 2, 131 nt in exon 3 and 3710 nt in exon 4 (Ensembl Genome Browser 2021).

Several common variants have been identified in the ZNF804A gene. The rs1344706 SNP (Single Nucleotide Polymorphism) is the most importantly common variant that has been associated with SZC, SAD and BPD. This SNP is located in the second intron of the gene and consists in a substitution of adenine for cytosine. O'Donovan et al. (2008) found an association between SCZ and BPD and rs1344706 in a GWAS, proposing the A allele as a risk factor (O'Donovan et al. 2008). The association of rs1344706 with SZC, SAD and BDP has been replicated in different studies and populations (Sun et al. 2015). It has been suggested that rs1344706 SNP may alter the gene expression of the ZNF804A modifying transcription activity or splicing factors binding to mRNA, or by acting as a cis variant that interacts with other polymorphisms that are in linkage disequilibrium (Hess & Glatt, 2014). In addition, rs1344706 has been associated with alterations in white or gray matter, brain connectivity and cognitive functions in patients with SCZ or BPD and in controls (Chang et al. 2017; Wang et al. 2019). Also, this variant has been associated with symptoms severity and response to treatment in both disorders (Wickramasinghe et al. 2015). Furthermore, rs1344705 has also been associated with psychosis-proneness by indirectly impacting on a tendency to misconstrue external or internal cues, which is hypothesized as central feature for presenting psychotic experiences (Stefanis et al. 2013).

Another ZNF804A SNP that has been associated with SCZ and BPD is the rs1366842, located in exon four and consists in a change of cytosine to adenine, resulting in a threonine to lysine change in the amino acid sequence (Baek et al. 2017). Therefore, the aim of the present study was to analyze the rs1344706 of the ZNF804A gene and identify common or rare variants by sequence analysis of a 352 bp targeted region around the rs1366842 SNP in Mexican patients with SCZ, BPD and SAD compared with a control group.

**Methods**

**Sample**

We included a total of 228 Mexican psychiatric outpatients who met the DSM-IV-TR and DSM5 criteria for SCZ, SAD and BPD type I, using the Mini-International Neuropsychiatric Interview (MINI) (Bobes, 1998)
and corroborated by at least two certified psychiatrists of the Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz. Inclusion criteria were: 1) Diagnostic Stability of two years; 2) Age ≥ 18 years; 3) Age of onset of the disorder < 50 years; and 4) Mexican patients, with parents and grandparents born in Mexico.

Also, we included 295 controls with the following inclusion criteria: Mexican subjects with parents and grandparents born in Mexico; individual without personal or family history of mental disorders evaluated with the MINI scale.

For the sequence analysis, we include an additional sample of 167 patients composed of 73 (44%) patients with SCZ, 23 (14%) with SAD and 71 (42%) with BPD, with the following characteristics: BPD patients with psychotic symptoms at least once in their life, family history of SCZ, SAD or BPD type I (at least one relative with a maximum of third degree), and patients carriers of A allele of the rs1344706 SNP. Also, we included 170 controls. All participants in the additional samples met the previous inclusion criteria.

**Genetic analysis**

**Genotyping**

Genomic DNA was extracted from blood lymphocytes using the Genomic DNA Purification Kit FlexiGene® (QIAGEN). Genotyping of rs1344706 was performed using a TaqMan SNP Genotyping Assay by-Design (C_2834835_10). Genotypes were analyzed by the TaqMan allele specific assay method using the ABI Prism® 7500 Sequence Detection System according to the manufacturer’s protocols (Applied Biosystems Inc., Foster City, CA) with a reaction mixture consisting of 20 ng of genomic DNA, 2.5 µl of TaqMan Universal Master Mix, and 0.125 µl of TaqMan SNP Genotyping Assay (Applied Biosystems Inc.). After denaturing at 95 °C for 10 min, 50 cycles of PCR were performed under the following conditions: denaturing at 95°C for 15 s, annealing and extension at 60°C for 1min.

**Sequence analysis**

Sequencing of the targeted region of exon four of the ZNF804A was performed in a sequence of 352 bp region spanning from 184937315 bp to 184937667 bp. The fragment was amplified using the oligonucleotide forward: 5'-AATGCTGGAATATCTATTGGGAAC-3' and reverse: 5'-GTGAATGTTGTCTACGTTTTCGAT-3', designed with the Primer-BLAST software(Ye et al., 2012). A 2% agarose gel electrophoresis was performed using 1X TAE as a running buffer and SYBR® Safe DNA Gel Stain as a dye to visualize the PCR product under UV light. The amplified product was purified using the QIAquick® PCR Purification Kit (QIAGEN) and its concentration and purity were evaluated by spectrophotometric analysis (NanoDrop ™ 2000). The BigDye ™ Terminator v3.1 Cycle Sequencing Kit reagent was used to perform the Sanger Sequencing, and the Centri-Sep ™ Spin Columns were used to remove the reaction residues. The sequences were obtained in the 3500XL Genetic Analyzer sequencer and, to identify the genetic variants, the sequencing results were compared with the reference sequence using the BioEdit Sequence Alignment Editor v7.1.11 software, the BLAST alignment tool, and the website
of the National Center for Biotechnology Information-NCBI. All genotyping and sequencing were performed blind to sample outcome.

**Statistical Analysis**

The categorical data were analyzed using Chi-square or Fisher’s exact test. Differences in dimensional variables were evaluated by Student’s t-test. Data are presented as means ± SD, frequencies and percentages. Analysis was performed using IBM SPSS Statistics 20 program. Analysis of linkage disequilibrium (LD) and haplotype analysis was performed with the SHEsis software (Shi & He, 2005).

**Results**

**Clinical and demographic characteristics**

We included a total sample of 228 patients, composed of 101 (44.3%) patients with SZC, 40 (17.5%) patients with SAD and 87 (38.2%) with BPD. Of these, 126 (55.2%) were females and 102 (44.8%) were males. Also, we included 295 controls, 143 (48.5%) were females and 152 (51.5%) males. Regarding sex there was no difference between cases and controls ($X^2 = 3.27, p = 0.124$).

Most of the patients were single (70.2%) and 28.1% were unemployed. Average age and years of education were $41.1 \pm 11.8$ years and $12.4 \pm 3.9$ years, respectively. 86.8% of the patients had family history of psychiatric disorders, the most frequent was mayor depressive disorder (37.7%), followed by alcohol use disorders (31.1%). 64.9% of the patients had comorbidity with other psychiatric disorder. In the case of SZC, mayor depressive disorder was the most frequent comorbidity (55.4%). When considering all the patients, the most frequent comorbid conditions were anxiety disorders (28.9%) and substance use disorders (26.3%). 61.4% and 24.1% of the patients reported previous suicidal ideation and attempt, respectively. The age of onset of the disorder, the years of course and the number of hospitalizations were $22.9 \pm 8.6$, $18.2 \pm 10.5$, and $2 \pm 2.2$, respectively.

**Genetic Analysis**

The genotype and allele frequencies of the rs1344706 in the patients with SCZ, SAD and BPD showed significant differences compared with the control subjects (Table 1). In particular, A allele was associated with an increased risk for presenting SCZ, SAD and BPD ($OR = 1.35$, CI 95% = 1.054–1.739, $p = 0.0208$). The SNP was in Hardy–Weinberg equilibrium ($p \geq 0.05$).
Table 1
Genotype and allele frequencies of rs1344706 in Mexican patients and controls

<table>
<thead>
<tr>
<th>Groups</th>
<th>Genotype frequencies, n (%)</th>
<th>Allele frequencies, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
<td>AC</td>
</tr>
<tr>
<td>Patients (n = 228)</td>
<td>91 (0.4)</td>
<td>107 (0.47)</td>
</tr>
<tr>
<td>Controls (n = 295)</td>
<td>99 (0.34)</td>
<td>133 (0.45)</td>
</tr>
</tbody>
</table>

Genotypes: $\chi^2 = 6.4$, $p = 0.041$; Alleles: $\chi^2 = 5.3$, $p = 0.0208$

In the sequence analysis, we identified a total of 3 genetic variants that had been previously reported in other populations. Of these, 2 were SNPs and 1 was a duplication. We did not identify rare variants in the sample analyzed in the present study. In Table 2, we showed the common variants identified in the sequence analysis among the groups (Table 2).

Table 2
Genetic variants identified by sequence analysis in a candidate region within exon 4 of ZNF804A gene in Mexican patients and controls

<table>
<thead>
<tr>
<th>Reference dbSNP</th>
<th>Position</th>
<th>Nucleotide Change</th>
<th>Aminoacid Change</th>
<th>Prediction</th>
<th>MAF patients</th>
<th>MAF controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3046266</td>
<td>185802211</td>
<td>c. 2090_2092dup</td>
<td>p.Thr697dup</td>
<td>Not reported</td>
<td>0.29</td>
<td>0.29</td>
</tr>
<tr>
<td>rs1366842</td>
<td>185802243</td>
<td>c. 2120C &gt; A/T</td>
<td>p.Thr707Lys</td>
<td>Benign</td>
<td>0.29</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>p.Thr707Ile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12477430</td>
<td>185802363</td>
<td>c. 2240G &gt; A</td>
<td>p.His747Arg</td>
<td>Benign</td>
<td>0.18</td>
<td>0.21</td>
</tr>
</tbody>
</table>

MAF = Minor Allele Frequency

The SNP based-analysis of the common variants identified in the sequence analysis did not show differences between cases and controls (Table 3). The three variants were in Hardy–Weinberg equilibrium ($p \geq 0.05$). Also, the analysis of the common variants by sex did not show differences between cases and controls (data not shown).
Table 3
Genotype and allele frequencies of the common variants identified in the sequence analysis

<table>
<thead>
<tr>
<th>Group</th>
<th>Genotype frequencies</th>
<th>Allele frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3046266¹</td>
<td>ACAACAA/ACAACAA ACAACAA/ACA AAACA</td>
<td>ACAACAA ACAACAA</td>
</tr>
<tr>
<td>Patients</td>
<td>80 (0.48) 78 (0.47) 9 (0.05)</td>
<td>238 (0.71) 96 (0.29)</td>
</tr>
<tr>
<td>Controls</td>
<td>83 (0.49) 74 (0.43) 13 (0.07)</td>
<td>240 (0.71) 100 (0.29)</td>
</tr>
<tr>
<td>rs1366842²</td>
<td>AA AC CC</td>
<td>A C</td>
</tr>
<tr>
<td>Patients</td>
<td>80 (0.48) 78 (0.47) 9 (0.05)</td>
<td>238 (0.71) 96 (0.29)</td>
</tr>
<tr>
<td>Controls</td>
<td>83 (0.49) 74 (0.43) 13 (0.07)</td>
<td>240 (0.71) 100 (0.29)</td>
</tr>
<tr>
<td>rs12477430³</td>
<td>GG GA AA</td>
<td>G A</td>
</tr>
<tr>
<td>Patients</td>
<td>109 (0.65) 55 (0.33) 3 (0.02)</td>
<td>273 (0.82) 61 (0.18)</td>
</tr>
<tr>
<td>Controls</td>
<td>109 (0.64) 52 (0.31) 9 (0.05)</td>
<td>270 (0.79) 70 (0.21)</td>
</tr>
</tbody>
</table>

Genotype and allele frequencies:

¹χ² = 0.86, p = 0.65; χ² = 0.04, p = 0.84
²χ² = 0.86, p = 0.65; χ² = 0.04, p = 0.84
³χ² = 3.06, p = 0.22; χ² = 0.58, p = 0.45

We observed a haplotype block composed of rs3046266, rs1366842 and rs12477430 (D′=1, r² = 1). The haplotype analysis did not show statistical differences between cases and controls (data not shown).

Discussion

Previous studies reported association between the ZNF804A gene and SZC, SAD and BPD, particularly with the common variant rs1344706. Our finding supports that in Mexican population the rs1344706 is also associated with SCZ, SAD and BPD. It has been suggested that this SNP can alter the expression of the ZNF804A gene at different stages of development, affecting neuronal growth, dendritic spine maintenance, and synaptic function; processes that have been related to the pathophysiology of these disorders (Deans et al. 2016; Hill et al. 2012; Hill & Bray, 2012; Tao et al. 2014). Also, different studies have described that the ZNF804A gene can modify the expression of other genes involved in neurogenesis, synaptogenesis, neurotransmission and inflammation (Chapman et al. 2018; Chen et al. 2015; Deans et al. 2016; Zhou et al. 2017). Therefore, it is possible that rs1344706 SNP, by modifying ZNF804A gene expression may affect the expression of other genes, modifying neuronal processes that would lead to dysfunctional neural circuits and structural alterations reported in patients with SCZ, SAD and BPD. These findings are supported by studies that found an association between rs1344706 and
dysfunction in brain connectivity or alterations in gray and white matter (Chang et al. 2017; Wang et al. 2019). However, more studies are still needed to understand how ZNF804A gene is related to the pathophysiology of SZC, SAD and BPD. For example, little is known about the function of the different mRNA isoforms or proteic isoforms encoded by this gene, as well as the change in their expression because of the rs1344706 variants. An allelic specific binding of different transcription and splicing factors has been related to this SNP (Hess et al. 2015; Hess & Glatt, 2014). Nevertheless, several results come from studies in silico, so more experimental studies are required.

It is important to analyze if rs1344706 in LD with other SNPs localized in the same gene and/or the interaction between ZNF804A and other genes may be related with the etiology of these disorders. A bioinformatic study found an interaction between rs1344706 and rs11685645 of Myt1L gene, and between rs1344706 and rs1836796 of NCAM1 gene, concluding that these SNPs may be implicated in the early development of central nervous system (Hess et al. 2015; Riley et al. 2010). Also, interaction study between rs1344706 and a SNP of CACNA1C reported association with increase activation in brain regions of SCZ and BPD patients (Tecelão et al. 2019). In addition, it was reported an interaction effect of rs1344706 and COMT variant on brain activity in healthy individuals (Cui et al. 2019). However, more studies that analyze the interaction between rs1344706 and other SNPs are required. On the other hand, we do not know how these gene variants interact with the environment, in order to understand the biological relationship between SZC, SAD and BPD and the ZNF804A gene.

Genetic studies carried out in Caucasian and Asian populations reported association with the rs1388842; however, we did not find an association in Mexican patients with SCZ, BPD and SAD. Also, we did not find an association between rs3046266 and rs12477430 with these disorders. The present study replicates the previous findings of negative association between rs12477430 and SCZ in Irish population (Riley et al. 2010).

Furthermore, within a targeted 352 bp region located on exon four of ZNF804A gene we did not identify single rare variants in SCZ, BPD and SAD Mexican patients. Similarly, Dwyer S. et al. (2010) in a sequencing study of exons 2, 3 and 4 of ZNF804A gene did not identified rare variants in SZC patients of UK population.

Our results replicate that ZNF804A gene is involved in the genetic risk of SCZ, SAD and BPD. However, much research is still required, as obviously the mechanisms involved in the development of these highly heterogeneous disorders are more complex. The differences between these disorders and even between individuals with the same disorder are probably explained by the numerous combinations of the different genetic and environmental risk factors that interact with each other; coupled with the sum of different additional modifying factors, either genetic and/or environmental, that do not influence the risk of developing a disorder, but influence its clinical presentation; explaining the difficulty to understand the biological role of the ZNF804A gene in this disorders.

In addition, we consider that in the future it would also be important to study the role of ZNF804A gene in other psychiatric disorders, principally neurodevelopmental disorders, or those with psychotic symptoms,
such as autism, primary delusional disorder, major depressive disorder with psychotic symptoms and substance-induced psychotic disorders.

The present study has several limitations. First, the small sample size in our analysis resulted in limited power to identify rare variants, principally ultra-rare variants. Second, we only sequenced a targeted region of exon four of the gene, consequently the existence of rare variants in other regions of the ZNF804A gene in SZC, SAD and BPD is possible; therefore, our study did not exclude the involvement of rare variants in the risk of presenting these disorders in Mexican population. It will be important sequencing the complete gene and replicate the association between rs1344706 and SZC, SAD and BPD in a large sample size.

Conclusions

In conclusion, our findings support the evidence suggesting that ZNF804A gene is involved in the etiology of SZC, SAD and BPD. Future studies are needed to identify the pleiotropic effect of this gene in psychiatric disorders.

Abbreviations

SCZ
Schizophrenia; SAD: Schizoaffective Disorder; BPD: Bipolar Disorder; SNP: Single Nucleotide Polymorphism; MINI: Mini-International Neuropsychiatric Interview.

Declarations

Acknowledgments

We thank all patients and controls that participated in this study.

Authors’ contributions

LMA and BC design the study. LMA, HHH, CBP, DGM, HOO, REO and RSA contribute to the clinical evaluation of the patients. LMA, JNQ, PT, JPL, SHM and IBJ carried out the genetic analysis. LMA, BC, JNQ and SHM performed the statistical analysis. LMA, BC and JNQ wrote and edit the manuscript. All authors read and approved the final manuscript.

Funding

CONACYT-FOSISS Grant 261459 funded this study. LMA was supported by a postgraduate grant from CONACYT (Number: 424842) at Master’s and Doctorate in Medical Sciences Program, UNAM.

Availability of data and materials
The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

**Ethics approval and consent to participate**

The Institutional Review Board of the Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz approved this study. All subjects that participated in this study previously signed and informed consent.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

1 Dirección de Servicios Clínicos, Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Ciudad de México, México.
2 Departamento de Farmacogenética Psiquiátrica, Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Ciudad de México, México.

3 Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Ciudad de México, México.
4 Subdirección de Investigaciones Clínicas, Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Ciudad de México, México.
5 Laboratorio de Fisicoquímica e Ingeniería de Proteínas, Departamento de Bioquímica, Facultad de Medicina, Universidad Nacional Autónoma de México, Ciudad Universitaria, Ciudad de México, México.

**References**


