Accepted Manuscript

Time dependent expression of the blood biomarkers *eif2d* and *tox* in patients with schizophrenia


PII: S0889-1591(19)30342-3
DOI: https://doi.org/10.1016/j.bbi.2019.05.015
Reference: YBRBI 3705

To appear in: *Brain, Behavior, and Immunity*

Received Date: 31 March 2019
Revised Date: 5 May 2019
Accepted Date: 8 May 2019


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
TIME DEPENDENT EXPRESSION OF THE BLOOD BIOMARKERS EIF2D AND TOX IN PATIENTS WITH SCHIZOPHRENIA

Javier Gilabert-Juan\textsuperscript{a,b,c,d,*}, Guillermo López-Campos\textsuperscript{a}, Noelia Sebastiá-Ortega\textsuperscript{a,b,c,d}, Sonia Guara-Ciurana\textsuperscript{a}, Fulgencio Ruso-Julve\textsuperscript{c,f}, Carlos Prieto\textsuperscript{g}, Benedicto Crespo-Facorro\textsuperscript{c,f}, Julio Sanjuán\textsuperscript{c,d,h}, María Dolores Moltó\textsuperscript{a,b,c,d}

\textsuperscript{a} Department of Genetics. Universitat de València. Valencia, Spain
\textsuperscript{b} Neurobiology Unit, Cell Biology Department, Interdisciplinary Research Structure for Biotechnology and Biomedicine (BIOTECMED). Universitat de València. Valencia, Spain
\textsuperscript{c} Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Spain
\textsuperscript{d} INCLIVA, Valencia, Spain
\textsuperscript{e} Centre for Experimental Medicine, Queen’s University Belfast, Belfast, UK
\textsuperscript{f} University Hospital Marqués de Valdecilla. IDIVAL. Department of Psychiatry. School of Medicine, University of Cantabria, Santander, Spain.
\textsuperscript{g} Servicio de Bioinformática, Nucleus, Universidad de Salamanca. Salamanca, Spain
\textsuperscript{h} Unit of Psychiatry. Universitat de València, Valencia, Spain

* Corresponding author at Department of Genetics. University of Valencia, Valencia, Spain. Tel. +34963543400. Fax. +34963543029. Email address: javier.gilabert@uv.es

# Present address: Centre for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS/UMR 7241 - INSERM U1050, Paris, France
Abstract

Background: During last years, there has been an intensive search for blood biomarkers in schizophrenia to assist in diagnosis, prognosis and clinical management of the disease. Methods: In this study, we first conducted a weighted gene coexpression network analysis to address differentially expressed genes in peripheral blood from patients with chronic schizophrenia (n = 30) and healthy controls (n = 15). The discriminating performance of the candidate genes was further tested in an independent cohort of patients with first-episode schizophrenia (n = 124) and healthy controls (n = 54), and in postmortem brain samples (cingulate and prefrontal cortices) from patients with schizophrenia (n = 34) and healthy controls (n = 35). Results: The expression of the *Eukaryotic Translation Initiation Factor 2D (EIF2D)* gene, which is involved in protein synthesis regulation, was increased in the chronic patients of schizophrenia. On the contrary, the expression of the *Thymocyte Selection-Associated High Mobility Group Box (TOX)* gene, involved in immune function, was reduced. EIF2D expression was also altered in first-episode schizophrenia patients, but showing reduced levels. Any of the postmortem brain areas studied did not show differences of expression of both genes. Conclusions: *EIF2D and TOX* are putative blood markers of chronic patients of schizophrenia, which expression change from the onset to the chronic disease, unraveling new biological pathways that can be used for the development of new intervention strategies in the diagnosis and prognosis of schizophrenia disease.
1. Introduction

During recent decades, great efforts have been made to understand the molecular basis of schizophrenia (SZ). As a result, a growing list of genetic polymorphisms, rare genetic variants, de novo mutations, epigenetic changes and other genomic alterations showing association with SZ has been provided(1-3). However, these findings have not been translated into suitable biomarkers for clinical use because most markers account for low disease risk and those based on gene expression in postmortem brain tissues cannot be analyzed since brain biopsies are not acceptable for clinical evaluation.

Blood is a good source of samples for disease assessment in living patients because this tissue is easily accessible and is in contact with tissues and cells throughout the body. Many studies have showed that alteration of metabolism and cellular functions in the central nervous system (CNS), as well as disturbances in neurotransmitter and hormonal systems are related with altered function and metabolism of blood cells (4, 5). This connection results from a complex network involving the nervous, endocrine and immune systems and has been investigated in several neuropsychiatric disorders. Different studies have compared gene expression in peripheral blood (PB) and brain to evaluate the usefulness of PB as a source of biomarkers that mirror the state of disease in the brain (6-8). Results from transcriptomic, proteomic and epigenomic analysis support the view that some genes, which expression is dysregulated in blood, might reflect specific changes that characterizes the abnormal biological processes in CNS (9-12). Therefore, in recent years, there has been an intensive search for this type of biomarker as a useful and cheaper alternative for SZ (13).
Here, we sought to identify gene expression markers in PB that may be useful as biomarkers of development and disease evolution in SZ. We first performed an expression microarray analysis in blood samples from male SZ patients with several years of illness duration for comparison to controls. To prioritize biomarkers, a systems biology approach was used based on weighted gene coexpression network analysis (WGCNA), and the discriminating performance of candidate genes was further tested in an independent cohort of patients with first-episode SZ.

2. Methods

2.1 Peripheral sample cohorts

Thirty unrelated male patients with SZ were recruited at the Psychiatric Out-patient Unit of the Clinic Hospital of Valencia University, Spain (cohort 1). The control group (CNT) consisted of 15 age-matched healthy males without a history of drug abuse or familial background of mental disorders. A second sample of 124 patients with first-episode SZ (62 males) and 54 healthy controls (30 males) were recruited from the University Hospital Marqués de Valdecilla in Santander, Spain (cohort 2). Detailed description of these cohorts is shown in the supplementary material and supplementary Table 1. This study was performed according to the guidelines of each institution involved and was approved by the Local Ethics Committee of each participating center.

2.2 Postmortem sample
Frozen RNA from the postmortem frontal cortex and cingulate cortex was obtained from the Stanley Medical Research Institute Array Collection (http://www.stanleyresearch.org/brain-research/array-collection). Postmortem brain samples were from donor subjects with SZ (n = 35) and controls with no history of psychosis (n = 34). The distributions of age and gender in the two groups remained similar. The tissue characteristics of this sample are detailed in Supplementary Table 2.

2.3 Expression microarray analysis

PB of subjects from cohort 1 was collected in RNA-stabilizing PAXgene tubes (Qiagen, Izasa, Spain). Total RNA was extracted from the obtained cells using the PAXgene blood RNA Kit (Qiagen, Izasa, Spain) according to the manufacturer's protocol (supplementary Table 1). Gene expression analysis was performed using the GeneChip Human Gene 2.0 ST Array (Affymetrix, Santa Clara, CA, USA) covering more than 750,000 unique 25-mer oligonucleotide features constituting over 40,000 RefSeq transcripts. Hybridization experiments and microarray data generation were conducted in the Central Unit for Research in Medicine Faculty-INCLIVA, University of Valencia (Spain). A gene expression profile was obtained for each subject in cohort 1.

2.4 Gene coexpression network analysis

WGCNA package in R(14) was used to calculate the coexpression network from the reduced gene set (24618 probes). We initially identified the value that transformed the data into an approximate scale-free topology. Next, we used Pearson’s correlation coefficient, a minimum module size of 30 genes and a minimum height value of 0.25 to merge modules (groups of highly correlated genes). For the functional annotation of the network modules associated with the different traits, we used the Database for
Annotation, Visualization and Integrated Discovery (DAVID) (http://david.abcc.ncifcrf.gov/home.jsp). We considered only those results with a false discovery rate of 0.05 to be significantly enriched. DAVID analysis results were visualized using the EnrichmentMap 2.0.1(15) plugin for Cytoscape 3.3.0 (http://cytoscape.org). For detailed description see supplementary material.

2.5 Reverse transcription quantitative PCR (RT-qPCR)

Total RNA of subjects from cohort 1 was converted into cDNA with Expand Reverse Transcriptase and hexanucleotide mix (Roche Applied Science, Indianapolis, USA), following the experimental conditions as previously described (16).

Amplification was conducted with an ABI PRISM 7700 Sequence Detector (Applied Biosystems) as previously described (16), using 5X HOT FIREPol EvaGreen qPCR Mix Plus (Solis BioDyne, Estonia) and gene-specific primers (Supplementary Table 3).

2.6 RNA sequencing (RNA-seq)

Total RNA was extracted from the PB of individuals from cohort 2. mRNA was isolated from the total RNA and was fragmented once transformed into cDNA. Fragments of 300 bp on average were selected to construct the libraries for sequencing. Pair-end sequences of 70 nucleotides for each end were produced. Then it was sequenced using Illumina HiSeq instruments (San Diego, CA, USA).

RT-qPCR and RNA-seq analyses are further described in the supplementary material.

2.7 Statistical analyses

Demographic, clinical and tissue features of cohorts were compared using ANOVA and \(t\)-test. Statistical significance was indicated at \(p < 0.05\). To identify potential blood
biomarkers for SZ contributing to the differentiation between control individuals and patients, a binary logistic regression was carried out with the gene expression data. A significance test of the Wald statistic was conducted to investigate genes whose expression is significantly associated in the final equation (p < 0.05). All statistical analyses were performed using SPSS v14.0 software (SPSS, Chicago, IL).

3. Results

3.1 Gene expression array analysis

Whole transcriptome analysis in PB samples of cohort 1 was performed to identify potential biomarkers in SZ. We focused on male subjects to minimize sex differences in gene expression. We identified 10 modules of correlated genes (genes with similar expression attending to the different groups studied and variables of the sample), each one of them was coded with an arbitrary color, and an eleventh gray module containing the genes that were not part of any particular module. The functional annotation of each module is shown in Supplementary Table 4.

The main purpose of the analysis was to identify genes whose expression was associated with the clinical condition (CNT versus SZ). We specified a significance threshold of p-value ≤ 0.05 for the correlation between connected-gene modules and traits. According to this threshold, four modules (turquoise, purple, yellow and blue) significantly correlated with the clinical condition (Figure 1), which means that there are differences in gene expression levels among CNT versus SZ. In addition, two modules (magenta and blue) were correlated with the patient condition, which accounts for differences in gene expression levels between SZ patients with and
without auditory hallucinations. Three modules (turquoise, purple and yellow) were associated with medication (supplementary Table 5); two modules (black and pink) correlated with RNA purity and one module (black) with the subject’s age. Importantly, we did not identify any highly connected gene module specifically related to chlorpromazine dosage under the thresholds applied (Figure 1).

---

**Figure 1**: Relationship between the different traits analyzed (RNA purity measured as A260/A280 and A260/230; Clinical condition; Patient condition; Age; Medication; Chlorpromazine equivalent) and gene modules using the WGCNA approach. Gene modules are designated by different colors on the left, color scheme on the right (blue-red) denotes the correlation (-1 to 1) between module and trait, and the $p$-value of the correlation in each interaction is indicated in brackets.
3.2 Selection of candidate genes as potential blood biomarkers in schizophrenia

We selected modules that were significantly correlated with the clinical or the patient conditions (with a cutoff value of \( p \leq 0.02 \) at least in one of these conditions) but that showed no correlation with the other traits analyzed. Gene expression in the blue module (Figure 1) is associated with the clinical condition \((p = 0.02)\) and patient condition \((p = 0.03)\). Purple and turquoise modules showed correlations with the clinical condition \((p = 0.02 \text{ and } 0.01, \text{ respectively})\). In these two modules, medication might have an effect on the expression of the included genes \((p = 0.03 \text{ and } p = 0.01, \text{ respectively})\). No significant effect was observed for the remaining confounding factors. Finally, the magenta module is associated with the patient condition \((p = 0.02)\).

Next, we chose genes from the selected modules (Supplementary Figures 1-4), whose expression differed in the clinical or patient condition traits with high statistical significance \((p < 0.001 \text{ in the blue and turquoise modules because they included a high number of genes; } p < 0.01 \text{ for purple and magenta modules, with fewer genes})\). A total of 14 genes were selected; \(CPQ, MSL1, NRD1, SRPK1, CMTM5, APOBEC3G, NKG7, SLAMF6, TOX, C12orf73, EIF2D, RNPS1, KPL29P2\) and \(SIGIRR\) (supplementary Table 6). These genes are involved in different biochemical pathways, including proteolysis, chromatin organization and RNA splicing (turquoise module); chemotaxis (magenta module); cytidine deamination, inflammatory process, protein binding, regulation of immune response and DNA regulation (purple module); and gene expression, translation initiation, signal transduction and unknown functions (blue module). Expression of these genes was measured by RT-qPCR and compared between the
control group and the SZ patients in cohort 1 (Supplementary Figure 5). We found differences in two of them. The expression of the *Eukaryotic Translation Initiation Factor 2D (EIF2D)* gene was increased in SZ patients versus controls (fold change = 2.05, \( p = 0.008 \), Figure 2A). On the contrary, the expression of the *Thymocyte Selection-Associated High Mobility Group Box (TOX)* gene was reduced in the SZ group (fold change = 0.6, \( p = 0.017 \), Figure 2A).

![Figure 2. A) Boxplots representing the quartiles of the relative gene expression values for the EIF2D and TOX in cohort 1; controls (CNT) are compared with schizophrenia patients (SZ). * \( p < 0.05 \). B) Boxplots in cohort 2, CNT are compared with the SZ patients at the onset of the disease (SZ 0), after 3 months (SZ 3 M) and 1 year of disease onset (SZ 1Y), *** \( p < 0.001 \).](image-url)
Next, we studied the expression of both genes in prefrontal and cingulated post-mortem brain regions, which have been largely associated with SZ symptoms, and no significant changes were observed (Supplementary Figure 6).

### 3.3 Predictive Model Equation

Using binary logistic regression, we obtained a predictive equation in which \(EIF2D\) and \(TOX\) gene expression could differentiate the SZ patients versus controls in cohort 1. The final equation model is 
\[-0.106 + [2.957 \times EIF2D] + [-4.198 \times TOX],\]
where \(EIF2D\) and \(TOX\) respectively indicate the expression levels of these genes. Applying this equation to subjects in cohort 1, SZ patients were distinguished from healthy controls with 86.5% accuracy, with 0.5 as a cutoff value (Supplementary Table 7).

### 3.4 Gene expression biomarkers in a new cohort

We assessed the predictive value of the 14 selected genes by measuring their expression in a new cohort (cohort 2). Gene expression was analyzed with RNA-seq data at disease onset (0), after 3 months (3M) and after 1 year (1Y) of disease onset. In this case, we also found significant differences in \(EIF2D\) expression between controls and patients, but \(EIF2D\) mRNA levels were lower in patients than controls, contrary to what we observed in cohort 1. This reduction was found at the 3 time points tested (fold change at disease onset = 0.88; after 3 months = 0.87; after 1 year = 0.87, \(p < 0.0001\) in all cases (Figure 2B). \(TOX\) expression was reduced in the SZ patients as in
cohort 1, although this change showed a trend at time point 0 and was not significant at 3M or 1Y (Figure 2B).

Next, we investigated a new predictive model equation using the expression data of the 14 selected genes in the cohort 2. We again found that EIF2D is included in a predictive model at each of 3 time points, but not TOX (Supplementary Table 8). Nevertheless, the predictive values of EIF2D expression in cohort 2 were found to be in opposite directions relative to those in cohort 1.

3.5 EIF2D changes with disease duration

We found that EIF2D expression increased in SZ patients in cohort 1 but decreased in SZ patients in cohort 2 regarding to the corresponding controls. Cohort 1 includes men with chronic SZ, while cohort 2 is composed of first-episode patients of both sexes evaluated over 1 year from the disease onset (Supplementary Table 1). To examine the potential effect of sex on EIF2D expression, we compared EIF2D mRNA levels only in men from cohort 2. We found a result similar to that described before for all samples, including men and women. Next, we compared EIF2D expression with respect to disease duration in SZ patients in cohort 1 and observed a positive correlation trend between EIF2D mRNA levels and years of disease ($p = 0.0794; r = 0.35$) (Figure 3A). Then, we grouped the SZ patients into those with a disease duration of less (n = 17) and more (n = 9) than 10 years. When comparing the EIF2D expression between both subgroups of SZ patients, we found a trend of an increase in the expression levels of EIF2D ($p = 0.0634$) in patients suffering from the disease more than 10 years compared with the SZ patients from the other subgroup (Figure 3B).
Figure 3: EIF2D expression with disease duration in SZ patients in cohort 1. A) Correlation between EIF2D mRNA level and the years of the disease in the SZ patients in cohort 1. B) Differences in EIF2D expression in SZ patients with less than 10 years of the disease compared with SZ patients with more than 10 years of the disease.

4. Discussion

The identification of PB biomarkers for SZ risk assessment has become a new and promising area of translational investigation in psychiatry. Our results revealed that the expression levels of EIF2D and TOX could discriminate between SZ patients and controls with 86.5% accuracy in the cohort 1.

EIF2D encodes a factor involved in a noncanonical translation initiation mechanism of protein synthesis that is believed to operate mainly in cells under stress (17). The
expression of the yeast orthologue gene rises under some stresses, such as anoxia, oxidative stress and accumulation of unfolded proteins in the endoplasmic reticulum. In these situations, the activity of eIF2, which participates in the canonical eukaryotic translation initiation pathway, is suppressed (17). In mammalian cells, eIF2 function is also blocked by different stress conditions (18). Our results of a two-fold increase in EIF2D mRNA levels in SZ patients relative to the levels in controls suggest the presence of cellular stressors in SZ patients. A growing body of data supports the implication of oxidative and endoplasmic reticulum stress in SZ (reviewed in(19)). Most studies analyzing markers of oxidative status, such as glutathione, in blood and in neural tissues point out oxidative imbalance in this disease (20, 21).

On the other hand, EIF2D has been involved in translation control of cellular mRNAs having short regulatory upstream open reading frames in the 5'UTR (22, 23). A large number of mammalian proteins are regulated by this mechanism, including many proteins needed for cell growth and proliferation (22). Other noncanonical initiation factors, DENR (density regulated protein) and MCT-1 (product of malignant T cell-amplified sequence 1 oncogene), also promote the translation reinitiation process through the DENR-MCTS1 complex (22, 23). Interestingly, missense mutations in DENR have been reported in unrelated patients with autism spectrum disorder (24). These results highlight the importance of proper mRNA translation reinitiation processes for brain development. We can speculate that higher EIF2D doses above its physiological level could also interfere with the translation reinitiation mechanism. In this line, increasing the levels of eIF4E in mice resulted in exaggerated cap-dependent translation and autistic-like behaviors (25). Finally, EIF2D is located on 1q32.1, a
chromosomal locus in which several studies have reported genetic linkage with SZ(26, 27).

TOX is an evolutionarily conserved member of the HMG-box family of transcription factors that has been studied mainly in the immune system and hematopoiesis (28, 29). TOX is necessary for the development of some cells of the adaptive and innate immune systems including CD4\(^+\) T cells, natural killer cells and lymphoid tissue inducer cells (30, 31). Interestingly, a high number of studies related dysregulation of immune and inflammatory function with SZ (13, 32). Remarkably, inflammation is closely interconnected with oxidative stress, being both a cause and a consequence, and vice versa (33). In addition, Tox expression in the mouse thymus is calcineurin dependent (34), and the calcineurin pathway has been proposed to be implicated in the pathogenesis of SZ (35). In neural tissues, Tox functions as an important regulator of neural stem cell proliferation and dendritogenesis, and its expression in the brain is also controlled by calcineurin (36). However, the changes in expression of EIF2D and TOX that we found in PB were not observed in the prefrontal and cingulated regions analyzed when comparing SZ patients with controls (Supplementary Figure 6), but these genes might be dysregulated in specific cell types and/or other different brain areas. In line, regionally specific changes in gene expression in SZ have been found in postmortem brain tissue samples (37).

To further assess the predictive value of the 14 selected genes for SZ risks, we analyzed their expression in a second independent cohort of patients and controls. This sample comprised first-episode SZ patients and then, its comparison with the chronic SZ subjects from cohort 1 could also reveal potential PB biomarkers dependent on the
disease evolution. In cohort 2, \textit{TOX} expression was also reduced with respect to controls, but it was not significant. In this cohort, we also found significantly altered expression of \textit{EIF2D} between the first-episode SZ patients and control subjects; however, \textit{EIF2D} mRNA levels were reduced in this case. This finding contrasts with the increased \textit{EIF2D} mRNA levels in the chronic patients of cohort 1. The reduction of \textit{EIF2D} expression in patients of cohort 2 was similar in the three time points tested during the first year of disease duration, suggesting that medication probably did not affect \textit{EIF2D} expression. The different results obtained for this gene in the SZ patients of the cohorts 1 and 2 might be explained by the disease state. In line with these findings, we observed a trend toward increased \textit{EIF2D} expression in patients suffering from the disease for more than 10 years compared with patients with less than 10 years of illness. Therefore, the expression level of this gene may be indicative of redox status in PB revealing higher redox dysregulation after several years of disease progression. This result agrees with lower antioxidant response found in SZ patients with longer duration of illness (38). So, \textit{EIF2D} expression may show a stage of disease progression. Future longitudinal studies in the first-episode SZ patients are needed to address this question together with molecular approaches to identify the underlying mechanism.

One of the limitations of our work is the sample size, but it is comparable to that used in other studies looking for biomarkers in human tissues (39, 40). Then further analysis should be followed up in larger cohorts at different time points and in different postmortem brain regions to compare the results. In addition, we cannot exclude the possibility that medication may have an effect on gene expression levels. It could be, in part, the cause of the differences found between the two cohorts, since the cohort 1 is
composed by long term medicated chronic patients and the cohort 2 is composed by naïve patients. Nevertheless, the significant reduction of $EIF2D$ expression in cohort 2 was maintained during the first year of the disease, when the patients are already medicated. Moreover, we tried to reduce the effect of this variable in the first discrimination point, excluding the nodes of genes associated with medication. Finally, different approaches have been used to study gene expression (RNA microarray, RT-qPCR and RNA sequencing), which could result in some variability in the final results. Nevertheless, all the techniques are evaluating the differences in gene expression. Also, appropriate controls were chosen to discard any alteration of the techniques and all methods have been normalized to minimize such technical deviations. In conclusion, we can underline the potential implication of protein synthesis regulation, oxidative stress and immune processes in the development and prognosis of SZ, marked by the expression of $EIF2D$ and $TOX$ genes in our samples. Our results support the use of antioxidants in SZ as an adjunctive treatment to standard antipsychotic medication.

Acknowledgments

This work was supported by Fondo de Investigación Sanitaria, Ministerio de Economía y Competitividad, Spain (PI10/01399, PI13/00447; PI17/00402, co-financed by FEDER) to J. Sanjuan and M.D. Moltó; Generalitat Valenciana PROMETEO Excellence Program, Spain (PROMETEO2016/082) to J Sanjuán. J Gilabert-Juan and N. Sebastiá-Ortega were recipients of research contracts from CIBERSAM, Spain. The RNA samples donated by
the Stanley Medical Research Institute Brain Collection were courtesy of Drs. Michael B. Knable, E. Fuller Torrey, Maree J. Webster, and Robert H. Yolken. The authors also thank the collaboration of the staff members of the hospitals.

The Authors have declared that there are no conflicts of interest in relation to the subject of this study.
5. References


37. Scarr E, Udawela M, Dean B. Changed frontal pole gene expression suggest altered interplay between neurotransmitter, developmental, and inflammatory pathways in schizophrenia. NPJ Schizophr. 2018;4(1):4.
Highlights

- The $EIF2D$ and $TOX$ genes, involved in protein synthesis and immune function are blood markers of schizophrenia.
- $EIF2D$ gene expression in blood changes with disease duration.
- $EIF2D$ as a potential marker of schizophrenia disease status.