Schizophrenia

ARTICLE OPEN

Gene co-expression architecture in peripheral blood in a cohort of remitted first-episode schizophrenia patients

Natalia Rodríguez1,2,48, Patricia Gassó2,3,48, Albert Martínez-Pinteno1,2, Álex-González Segura1,2, Gisela Mezquida1,2,3,4, Illumína Corripio1,2,3,4,5,7,8, Àlex-González Segura1,2, Gisela Mezquida1,2,3,4, Iluminada Corripio1,2,3,4,5,7,8, Iñaki Zorrilla1,2,3,4,5,7,8, Marta Martín1,2,3,4,5,7,8, Roberto Rodríguez-Jiménez1,2,3,4,5,7,8, Fernando Contreras23,24, Miquel Bioque2,3,4,25, Anna Butjosa2,3,4,25, Anna González-Pinto5,9,10,11, Esther Berrocoso26,27, Marta Martin8,12,13, Roberto Rodriguez-Jimenez8,14,15, Iluminada Corripio8,16,17, Salvador Sarró18,19, Anna Butjosa20,21,22, Fernando Contreras23,24, Miquel Bioque2,3,4,25, Manuel-Jesús Cuesta26,27, Mara Parellada28, Ana González-Pinto29,9,10,11, Esther Berrocoso26,27, Miquel Bernardo2,3,4,25, Sergi Mas2,3,4,25, and 2EPS group*

A better understanding of schizophrenia subtypes is necessary to stratify the patients according to clinical attributes. To explore the genomic architecture of schizophrenia symptomatology, we analyzed blood co-expression modules and their association with clinical data from patients in remission after a first episode of schizophrenia. In total, 91 participants of the 2EPS project were included. Gene expression was assessed using the Clarion S Human Array. Weighted-gene co-expression network analysis (WGCNA) was applied to identify modules of co-expressed genes and to test its correlation with global functioning, clinical symptomatology, and premorbid adjustment. Among the 25 modules identified, six modules were significantly correlated with clinical data. These modules could be clustered in two groups according to their correlation with clinical data. Hub genes in each group showing overlap with risk genes for schizophrenia were enriched in biological processes related to metabolic processes, regulation of gene expression, cellular localization and protein transport, immune processes, and neurotrophin pathways. Our results indicate that modules with significant associations with clinical data showed overlap with gene sets previously identified in differential gene-expression analysis in brain, indicating that peripheral tissues could reveal pathogenic mechanisms. Hub genes involved in these modules revealed multiple signaling pathways previously related to schizophrenia, which may represent the complex interplay in the pathobiological mechanisms behind the disease. These genes could represent potential targets for the development of peripheral biomarkers underlying illness traits in clinical remission stages after a first episode of schizophrenia.

Schizophrenia (2022) 8:45; https://doi.org/10.1038/s41537-022-00215-1

INTRODUCTION

Schizophrenia is a severe, lifelong mental disorder that can lead to significant functional impairment. It is a multifactorial disorder, where both genetic and environmental factors are important for its etiology. The largest genetic studies of schizophrenia have highlighted its polygenic basis, with many genetic loci, mostly with small effect size. Clinical symptomatology and response to antipsychotics, the cornerstone medication treatment for schizophrenia, varies largely between patients. Developing a better understanding of schizophrenia subtypes could contribute to the development of personalized medicine. For this, it is necessary to identify biomarkers to stratify the patients according to clinical attributes (i.e., the symptoms) that allow such medical care to be directed more effectively.

Within this setting, there is a need to understand how genomic variation could explain the existing heterogeneity in symptoms, treatment response, course, and outcome in schizophrenia, and to determine if genomic data could be a useful tool to define patients or groups of patients for clinical purposes. To understand how genomic variation contributes to schizophrenia heterogeneity and the biological insights behind clinical stratification, we have to move from genetic variants to genes, and from genes to networks. Genes act in coordination, affecting the function of other genes in order to influence a particular phenotype through cellular pathways that are intertwined in complex networks. Network analysis could define disorder-associated molecular pathology and has allowed the identification of molecular pathways involved in schizophrenia risk. Transcriptomic

1Department of Clinical Foundations, Pharmacology Unit, University of Barcelona, Barcelona, Spain. 2Institut d’Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain. 3Barcelona Clínic Schizophrenia Unit (BCSU), Neuroscience Institute, Hospital Clinic de Barcelona, Barcelona, Spain. 4Department of Psychiatry, Complejo Hospitalario de Navarra, Pamplona, Spain. 5IdiSNA, Navarra Institute for Health Research, Pamplona, Spain. 6Department of Psychology, University of Cádiz, Cádiz, Spain. 7Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, IISGM, School of Medicine, Universidad Complutense, Madrid, Spain. 8Centro de Investigación Biomédica en Red en salud Mental (CIBERSAM), Madrid, Spain. 9Hospital Universitario de Alava, Vitoria, Spain. 10BIOARABA Health Research Institute, Vitoria, Spain. 11University of the Basque Country, Vitoria, Spain. 12Hospital del Mar Medicar Research Institute (IMIM), Barcelona, Spain. 13Autonomous University of Barcelona, Barcelona, Spain. 14Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas 12), Madrid, Spain. 15CogPsy Group, Universidad Complutense de Madrid (UCM), Madrid, Spain. 16Psychiatry Department, Institut d’Investigació Biomèdica-Sant Pau (IIB-SANT PAU), Hospital de la Santa Creu i Sant Pau, Barcelona, Spain. 17Universitat Autònoma de Barcelona (UAB), Barcelona, Spain. 18FIDMAG Germans Hospitalaries Research Foundation, Barcelona, Spain. 19School of Medicine, Universitat Internacional de Catalunya, Barcelona, Spain. 20Department of Psychiatry, Hospital Universitario Ramón y Cajal, IRYCIS, Universidad de Alcalá, Madrid, Spain. 21Parc Sanitari Sant Joan de Déu, Institut de Recerca Sant Joan de Déu, Barcelona, Spain. 22Hospital Sant Joan de Déu, Barcelona, Spain. 23Bellvitge Biomedical Research Institute IDIBELL, Department of Psychiatry, Bellvitge University Hospital, Hospital de Llobregat, Barcelona, Spain. 24Biomedical Research Network of Mental Health Network (CIBERSAM), Barcelona, Spain. 25Department of Medicine, University of Barcelona, Barcelona, Spain. 26Neuropsychopharmacology and Psychobiology Research Group, Department of Psychology, University of Cádiz, Cádiz, Spain. 27Instituto de Investigación e Innovación Biomédica de Cádiz, INIBICA, Hospital Universitario Puerta del Mar, Cádiz, Spain. 48These authors contributed equally: Natalia Rodríguez, Patricia Gassó. *A list of authors and their affiliations appears at the end of the paper. **email: bernardo@clinic.cat; sergimash@ub.edu

Published in partnership with the Schizophrenia International Research Society.
network analysis based on gene co-expression identified transcriptional networks dysregulated in schizophrenia and its cellular architecture. Network analysis can interrogate multiple levels of molecular organization and provide a biological interpretation to clinical phenotypes.

Therefore, to explore the genomic architecture of schizophrenia symptomatology, we analyzed blood co-expression modules, i.e., clusters of genes with highly correlated expression, in a cohort of remitted first-episode schizophrenia patients with <5 years of evolution, and their association with clinical data, including global functioning, clinical symptomatology, and premorbid adjustment.

RESULTS

WGCNA results and validation of the gene co-expression network

Table 1 shows the demographic, clinical, and pharmacological characteristics of the 91 participants in the present study.

| Table 1. Demographic, clinical and pharmacological data of the 91 participants in the present study. |
| --- |
| Baseline |  |
| N | 91 |
| Age, mean ± SD | 25.3 ± 5.85 |
| Age at first diagnosis, mean ± SD | 24.1 ± 5.7 |
| Gender, male, N (%) | 62 (67.4) |
| Ethnicity, Caucasian, N (%) | 81 (88.0) |
| Functioning |  |
| CGI (mean ± SD) | 3.2 ± 1.2 |
| GAF (mean ± SD) | 69 ± 14.5 |
| FAST (mean ± SD) | 20.2 ± 15.7 |
| Psychiatric symptoms |  |
| PANSS positive (mean ± SD) | 9.6 ± 3.2 |
| PANSS negative (mean ± SD) | 13.9 ± 5.4 |
| PANSS general (mean ± SD) | 25.6 ± 7.9 |
| PANSS total (mean ± SD) | 49.1 ± 14.8 |
| Affective symptoms |  |
| YMRS (mean ± SD) | 0.8 ± 1.5 |
| MADRS (mean ± SD) | 6.9 ± 6.3 |
| PAS (mean ± SD) | 45.7 ± 21.4 |
| Antipsychotic |  |
| Aripiprazol, N (%) | 34 (40.5) |
| Paliperidone, N (%) | 24 (28.6) |
| Risperidone, N (%) | 13 (15.4) |
| Olanzapine, N (%) | 12 (14.3) |
| Clozapine, N (%) | 6 (7.1) |
| Quetiapine, N (%) | 3 (3.6) |
| Amsulpride, N (%) | 3 (3.6) |
| Clotiapine, N (%) | 2 (2.4) |
| Asenapine, N (%) | 1 (1.2) |
| Haloperidol, N (%) | 1 (1.2) |
| Ziprasidone, N (%) | 1 (1.2) |
| Co-medication |  |
| Antidepressant, Yes, N (%) | 24 (28.6) |
| Anxiolytic, Yes, N (%) | 13 (15.5) |
| Lithium, Yes, N (%) | 2 (2.4) |
| Antiepileptic, Yes, N (%) | 5 (5.9) |
| Antiparkinsonian, Yes, N (%) | 11 (13.1) |

One sample with connectivity less than –5 was removed (Supplementary Fig. 1A). We identified 24 modules of co-expressed genes (Supplementary Fig. 1C). The inferred modules showed different sizes ranging from 41 (Dark gray module) to 5627 genes (Turquoise module). A further 1901 genes were assigned into the Gray module, which represents the genes that were not co-expressed based on gene dissimilarity. Importantly, the organization of our co-expression modules was robustly defined in our cohort (Supplementary Fig. 2) and significantly overlapped with the modules originally reported in peripheral blood described by Gudmundsdottir et al. We identified 24 modules of the 24 identified in our analysis (excluding the Gray module) showed significant overlap with 41 of the 55 modules identified by Gudmundsdottir et al.

Identification of co-expression networks related to clinical data

We found that MEs of six modules were significantly correlated with clinical data after correcting for multiple testing (p < 0.007) (Table 2 and Supplementary Table 2). These modules could be grouped according to their correlations with global functioning, clinical symptomatology, and premorbid adjustment in two clusters (Fig. 1). Cluster 1 included the Blue, Green, and Cyan modules, showing positive correlations with global functioning measured with the CGI, GAF, and FAST scores, negative correlations with psychotic symptoms measured with the PANSS scale and its subscales (positive, negative, and general), and positive correlations with premorbid adjustment measured with the PAS scale (the Blue module was not significant for the latter). Cluster 2 was formed by the Red, Turquoise, and Magenta modules, which showed an opposite pattern to Cluster 1 modules, with negative correlations with global functioning, positive correlations with psychotic symptoms, and negative correlations with premorbid adjustment.

Five additional modules (Yellow, Black, Grey60, Cyan, and Darkgreen) showed significant correlations with some clinical variables. However, these modules did not show a significant correlation between GS and MM. The lack of a GS–MM correlation could indicate that only a submodule relates to the trait or suggests that the association should be considered as more tentative, needing further validation or evidence. For this reason, these modules were not selected for further analysis.

Identification of hub genes and external validation of relevant modules

For each module, "hub genes" (more centralized genes in the network) were defined according to their module membership (MM > 0.8) and their gene significance (GS > 0.3) (Supplementary Table 3). Hub genes from each module were combined with the hub genes of the other modules included in the same defined cluster and used to create gene sets. These gene sets were tested for overlap with gene sets previously associated with schizophrenia in gene expression studies in brain samples. The 1331 hub genes from Cluster 1 showed significant overlap with genes involved in several modules identified in DLPFC of schizophrenia patients in the studies of Fromer et al. (five modules) and Gandal et al. (one module) (Table 3). Among these modules, M11 and M13 modules from Fromer et al. were significantly associated with schizophrenia in the original study. The 403 hub genes from Cluster 2 showed significant overlap with four modules reported by Fromer et al. and four models reported by Gandal et al. Among these modules, M2 from Fromer et al. was significantly associated with schizophrenia in the original study but also significantly enriched with genes associated with schizophrenia in genetic studies that include genome-wide association studies but also copy number variants and rare variants studies. The Turquoise module from Gandal et al. was significantly associated.
with schizophrenia risk. Moreover, the M2 and the Turquoise modules were enriched with neuronal cell markers.

Finally, the gene set enrichment analysis of the hub gene sets defined for each cluster is shown in Fig. 2.

**DISCUSSION**

We performed an analysis of gene co-expression architecture in peripheral blood in a cohort of remitted first-episode schizophrenia patients with <5 years of evolution. We identified six modules of co-expressed genes with significant associations with clinical data that showed significant overlap with genes previously associated with schizophrenia in brain samples. These modules were clustered into two groups according to their association and the direction of these associations, and the hub genes of each cluster were enriched with functional ontologies.

The purpose of the study was to identify and prioritize co-expressed gene sets based on their association with clinical data.

| Module        | Functionality | Symptomatology | PAS |
|---------------|---------------|-----------------|-----|
|               | CGI           | GAF             | FAST| PANSS positive | PANSS negative | PANSS general | PANSS total | YMRS | MADRS | PAS   |
| Meturquoise   | 0.190         | −0.255*         | 0.162 | 0.283*         | 0.229*         | 0.313**       | 0.311**     | −0.011 | 0.088 | 0.081 |
| Memagenta     | 0.244**       | −0.211          | 0.315** | 0.017         | 0.226*         | 0.209         | 0.197       | −0.141 | 0.138 | 0.243* |
| Mecyan        | −0.291*       | 0.308**         | −0.374** | −0.164       | −0.273*        | −0.366**      | −0.329**    | 0.214   | −0.192 | −0.243* |
| Mebue         | −0.283*       | 0.320**         | −0.207 | −0.287*       | −0.294**       | −0.332**      | −0.346**    | −0.011  | −0.061 | −0.179 |
| Megreen       | −0.320**      | 0.315**         | −0.411** | −0.389**     | −0.306**       | −0.482**      | −0.451**    | 0.198   | −0.200 | −0.320** |
| Mered         | 0.399**       | −0.442**        | 0.370** | 0.232**       | 0.380**        | 0.387**       | 0.394**     | −0.186  | 0.313** | 0.469** |

*p < 0.05; **p < 0.007.

CGI Clinical Global Impression Scale, GAF Global Assessment of Functioning Scale, FAST Functional Assessment Staging Test, PANSS Positive and Negative Syndrome Scale, YMRS Young Mania Rating Scale, MADRS Montgomery–Asberg Depression Rating Scale, PAS Premorbid Adjustment Scale.

**Fig. 1** Heatmap of the Pearson correlation coefficient between module eigengenes (MEs) and clinical information. The color of the cell reflects the size of the correlation coefficient, as shown in the legend on the right.
and their overlap with genes previously associated with schizophrenia in large datasets. The purpose of the study was to identify and prioritize co-expressed gene sets based on their association with clinical data and their overlap with genes previously associated with schizophrenia in large datasets, in order to identify peripheral biomarkers that could ultimately translate into the clinical practice with the stratification of patients according to the underlying illness traits in clinical remission stages after a first-episode schizophrenia.

We clustered modules with significant correlations with clinical variables to provide a more clinical interpretation of our results. Some modules showed significant associations with the same clinical variables and directions, although we could not detect a strong correlation among them. These modules could include genes involved in similar biological processes that could explain its effect on patients’ functionality, symptomatology, and premorbid adjustment, although these genes were clustered in different co-expression modules. The functional analysis and interpretation were focused on the hub genes of these clusters. There is an ongoing debate in the literature regarding the importance of hub genes, central nodes in network architecture, and its importance in functional analysis. On this regard, several authors reported that intramodular hub genes from coexpression networks instead of whole network hubs are more significantly related to a trait and more often of clinical importance. Moreover, selecting intramodular hubs in relevant modules often leads to gene lists with cleaner biological annotations after functional enrichment analysis evaluation. This is relevant for studying candidate biological processes associated with the trait of interest.

Lower expression of genes in Cluster 1 correlated with poor global functioning, the worst severity of psychotic symptoms, and low premorbid adjustment. The hub genes of this group participated in the chemical reactions and pathways involving compounds such as catecholamine (norepinephrine and dopamine) and serotonin, the neurotransmitters that have been identified in schizophrenia. The hub genes of this group are involved in the regulation of gene expression, protein localization, and signal transduction pathways. The enrichment analysis of the hub genes from Module Cluster 1 showed significant overlap with those reported by Fromer et al. and Gandal et al.

| Study | Module | Overlapping genes | Corrected P-value |
|-------|--------|-------------------|------------------|
| Cluster 1 | Fromer M10s | 42 | 0.049 |
| | Fromer M11s | 94 | 1.30 × 10^{-7} |
| | Fromer M13s | 81 | 3.49 × 10^{-7} |
| | Fromer M6s | 58 | 0.001 |
| | Fromer M8s | 111 | 1.02 × 10^{-10} |
| | Gandal Green | 158 | 2.28 × 10^{-8} |
| Cluster 2 | Fromer M1s | 39 | 2.65 × 10^{-15} |
| | Fromer M2s | 24 | 1.08 × 10^{-9} |
| | Fromer M3s | 17 | 0.002 |
| | Fromer M4s | 18 | 7.90 × 10^{-10} |
| | Gandal Blue | 43 | 6.04 × 10^{-21} |
| | Gandal Pink | 12 | 3.80 × 10^{-7} |
| | Gandal Red | 20 | 5.56 × 10^{-7} |
| | Gandal Turquoise | 32 | 4.77 × 10^{-13} |

Table shows the module, the number of overlapping genes, and the corrected p-value of hypergeometric test statistics.
classically related to schizophrenia psychopathology. The presence of multiple biological processes and signaling pathways in the module, such as the metabolism of neurotransmitters and the regulation of gene expression, and the regulation and transport of proteins in the cell, could reflect complex interplays between these processes in the neurobiology of the disease. This crosstalk between signaling pathways is a critical issue in the underlying polygenic architecture of complex diseases such as schizophrenia. Among these genes, there are several genes that have been widely studied as candidate genes involved in schizophrenia, especially Sirtuin 1 (SIRT1), the ubiquitin-like modifier activating enzyme 3 (UBA3), the neuroblastoma RAS proto-oncogene GTPase (NRAS), the cytoplasmic granule associated RNA binding protein (TIA1), the splicing factor 3b subunit 1 (SF3B1), and the ubiquitin-conjugating enzyme E2 N (UBE2N) were also included in Cluster 1.

Genes belonging to modules included in Cluster 2 showed a significant correlation between higher gene expression and low functioning, higher severity of psychotic symptoms, and poor premorbid adjustment. Hub genes of this cluster, in contrast to cluster 1, were enriched mainly with biological processes related to inflammatory pathways. This is especially relevant considering the immunological and inflammatory hypothesis of schizophrenia and the possible role of inflammatory markers on symptoms severity. Among enriched processes, two signaling pathways have been repeatedly proposed to be involved in the neurobiology of schizophrenia: the Toll-like receptor signaling pathway and the neurotrophin pathway. Both biological processes have been shown to play a vital role in neuronal growth and differentiation, as well as in neural survival, synaptogenesis, and neuroplasticity in the adult brain. Complex crosstalk between these processes has been proposed and could form the neurobiological basis of schizophrenia through the cellular communication between astrocytes and microglia. A notable gene in the network is furin, a paired basic amino acid cleaving enzyme (FURIN), one of the loci that showed a significant overlap between GWAS for schizophrenia and eQTL in the DLPFC reported by Fromer et al. Moreover, in the same study, experimental suppression of this gene had an impact on neuroanatomical and developmental attributes in an experimental model of zebrafish neurodevelopment. The involvement of FURIN in schizophrenia has been extensively studied in recent years. One FURIN substrate is the von Willebrand factor (VWF), a critical gene in the Cluster 2. The VWF is a well-known marker of endothelial cell activation and inflammation that has been associated with brain morphology, cognitive functions, and affective and psychotic symptoms. Other critical genes in the network included the MAP kinase kinase 7 (MAP2K7), a transcription factor with numerous functions in neurodevelopment and adult brain that has been implicated in schizophrenia.

Few studies analyzing gene expression in the peripheral blood have focused specifically on the transcriptome of first-episode psychosis, and its relation with clinical symptomatology, with the notable exception of the study of Leirer et al., although they use individual gene analysis. Some limitations should be considered in the interpretation of our results. First, the sample size may have limited the statistical power of our analysis. Second, although modules of co-expressed genes provide potential insights to understand biological mechanisms, confirmatory evidence requires experimental studies. Last, due to the naturalistic design, drug treatment was not controlled, and the study participants maintained their usual treatment. Despite these limitations, the strength of this study lies in the inclusion of a consistent well-characterized first-episode schizophrenia-patient sample in remission and well-described symptoms and functioning.

In conclusion, our study offers a characterization of the cotranscriptome in the peripheral blood of a sample of first-episode schizophrenia patients in remission and its relation to global functioning, clinical symptomatology, and premorbid adjustment. Our results indicate that modules with significant associations with clinical data showed overlap with gene sets identified in differential gene-expression analysis in DLPFC, indicating that peripheral tissues could reveal pathogenic mechanisms. Hub genes involved in these modules revealed multiple signaling pathways, previously related to schizophrenia, which may represent the complex interplay in the pathological mechanisms behind the disease. Overall, Cluster 1 and 2 shared clinical correlates, although in opposite directions. This could be of great clinical interest. Critical genes in the functional networks could represent potential targets for the development of peripheral biomarkers underlying illness traits in clinical remission stages after a first episode of schizophrenia.

METHODS

Study design

This study is part of a naturalistic, multicentre, coordinated, and multimodal project “Clinical and neurobiological determinants of second episodes of schizophrenia. Longitudinal study of first episode of psychosis” (PI11/00325), also known as the 2EPs Project. The project includes multiple sub-studies: general and basic, neuroimaging, adherence, neurocognition, physical health, and biological. Given its main goals, the present study was framed within the general and biological modules.

Subjects

The inclusion criteria for the 2EPs Project were (a) age between 16 and 40 years at the time of first assessment (baseline visit); (b) meeting diagnostic criteria according to DSM-IV for schizophrenia or schizoaffective disorder, (c) being in remission from the first psychotic episode (which should have occurred within the last 5 years), according to Andreasen’s criteria; (d) not having relapsed after the first psychotic episode; (e) speaking Spanish fluently; and (f) providing the signed informed consent form. The exclusion criteria were (a) having experienced a traumatic brain injury with loss of consciousness; (b) presenting intellectual disability understood not only as IQ <70, but also presenting malfunctioning and problems with adaptive processes; and/or (c) presenting somatic pathology with mental repercussion.

From the initial 223 patients recruited in the 2EPs Project, 91 (40.8%) participated in the biological module and provided a biological sample for gene expression analysis at baseline.

The study was approved by the investigation ethics committee of the Hospital Clinic (Barcelona, Spain). Informed consent was obtained from all participants. For children under the age of 18 years old, parents or legal guardians gave written informed consent before beginning their participation in the study, and patients assented to participate. This study was conducted in accordance with the Declaration of Helsinki.

Clinical assessment

Demographic data were collected for all patients through semi-structured interviews. Diagnoses were determined according to the DSM-IV-criteria, with the SCID or the Kiddie-SADS depending on age. In order to obtain global functional outcome information, three different scales were used:

1. The Clinical Global Impression Scale (CGI-S) assesses the severity of global symptomatology;
2. The Functional Assessment Staging Test (FAST) evaluates the patient’s degree of difficulty in autonomy, work functioning, cognitive functioning, finance, interpersonal relationships, and free time functioning;
3. The Global Assessment of Functioning Scale (GAF) measures the severity of symptoms and the level of functioning.

The clinical symptomatology was assessed, separating by different areas as follows:
(1) Psychotic symptoms were assessed using the Spanish validated version of the Positive and Negative Syndrome Scale (PANSS) which comprises three subscales (positive, negative, and general);

(2) Affective symptoms were assessed using the Spanish validated version of the Young Mania Rating Scale (YMRS), designed to assess the severity of manic symptoms and the Spanish validated version of the Montgomery–Asberg Depression Rating Scale (MADRS), to assess the severity of depression.

Finally, premorbid adjustment was assessed using the Premorbid Adjustment Scale (PAS) which explores sociability and withdrawal, peer relationships, school achievement, adaptation to school, and ability to establish socio-affective and sexual relationships.

Sample collection, RNA isolation, and microarray hybridization

Peripheral blood was collected at baseline in PAXgene Blood RNA tubes and total RNA was isolated (PAXgene Blood RNA kit, PreAnalytiX GmbH, Switzerland). The purity and integrity of RNA were assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). The mean RNA integrity number (RIN) was 7.77 ± 0.70, ranging from 6.5 to 8.7. No sample was discarded due to low RIN number. A total of 1 μg of purified RNA from each of the samples was submitted to the Kompetenzzentrum für Fluoreszentle Bioanalytik Mikroarray Technology (KFB, BioPark Regensburg GmbH, Regensburg, Germany) for labeling and hybridization to the Clariom S Human Array (Affymetrix, Santa Clara, CA, USA), following the manufacturer’s protocols. The Clariom S Human Array comprises more than 221,300 probes covering over 337,100 transcripts and variants, which in turn represent 20,800 genes.

Genome-wide expression analysis and the WGCNA procedure

Microarray data preprocessing was performed using the Oligo R package. The data were standardized using robust multichip analysis. Multiple probes mapping to the same gene were merged using the average as the summary of the hybridization values. Normalized expression data from all samples were used to identify and remove any source of unwanted variation before the construction of co-expression networks. We applied the sva R package that contains functions for identifying and building surrogate variables for high-dimensional data. First, the normalized data matrix was converted in an ExpressionSet object. Second, we applied the sva function using the n.sv argument and the num.sv function to estimate the number of surrogate variables. After applying this approach, the sva functions returned 0 arguments, assuming that no-sources of unwanted variation are present in our expression matrix. Cell counts were measured using the Gene Expression Deconvolution Interactive Tool (GEDIT). The following cells were counted: CD8+ T cells, CD4+ T cells, natural killer (NK) cells, B cells, and monocytes, using the Human Body atlas as reference matrix. Co-expression modules were identified using the R software package for weighted gene co-expression network analysis (WGCNA).

First, in order to remove outlier samples, distance-based adjacency matrices of samples were estimated, and sample network connectivity according to the distances was standardized. Samples with connectivity less than −5 were considered as outliers and were excluded (Supplementary Fig. 1A). The co-expression analysis involved constructing a matrix of pairwise correlations between all pairs of genes across all selected samples. Next, the matrix was raised to a soft-thresholding power (β = 6 in this study) to obtain an adjacency matrix. To identify modules of co-expressed genes, we constructed the topological overlap-based dissimilarity, which was then used as input to average linkage hierarchical clustering. This step resulted in a clustering tree (dendrogram), branches of which were identified for cutting based on their shape, using the dynamic tree-cutting algorithm (Supplementary Fig. 1C). The above steps were performed using the automatic network construction and module detection function (blockwiseModules in WGCNA) with the following parameters: minModuleSize of 30, reassignThreshold of 0, and mergeCutHeight of 0.25.

The modules were then tested for their associations with clinical variables by correlating module eigengenes (MEs, defined as the first principal component of each module) with scale values using Pearson partial correlation. For each significant module, the correlation between the gene significance (GS, the absolute value of the Pearson correlation between each gene expression and scale score) and its module membership (MM, the correlation between gene expression and the module eigengene) was calculated adjusted by sex, age, age at onset, ethnicity, and cell count. Bootstrapping was used as a cross-validation approach and the confidence intervals for the correlation coefficient were estimated using 1000 random samples. Multiple testing corrections were applied using the false discovery rate (FDR) as described by Benjamini & Hochberg. According to this method, p-value < 0.007 was considered statistically significant.

Modules showing significant correlations with clinical data after multiple testing corrections were clustered according to the computed correlation matrix. We first calculated the distance matrix and then performed hierarchical clustering of modules as implemented in the ComplexHeatmap R package.

Validation of the gene co-expression network

To assess whether the resulting co-expression modules were robustly defined in our cohort, we performed a subsampling analysis (Supplementary Fig. 2). This analysis consisted in the network construction and module identification using the previous parameters, with 50 iterations including randomly drawn individuals, as implemented in the sampleBlockwiseModules function in the WGCNA R package. For each gene, the consistency was calculated as the percentage of iterations in which it was assigned to the original module. Finally, the stability of each module was defined as the average gene consistency of all genes constituting the given module.

An external validation was performed assessing the replication of the identified co-expression modules in the general blood architecture (of an independent sample). To this end, we test the degree of overlap, using the hypergeometric test implemented in the userListEnrichment function from the WGCNA package, between our modules and the modules described in peripheral blood in a large cohort of healthy human samples (N = 2127).

External validation of relevant modules

To characterize modules significantly associated with clinical symptomatology, the hub genes in each module were used to define a gene set. Hub genes were defined according to their gene significance (GS) with the clinical variable of interest (GS > 0.3) and their module membership (MM > 0.8). These gene sets were tested for overlap with the gene composition and module labels from gene expression studies in dorsolateral prefrontal cortex (DLPFC) of large cohorts of schizophrenia subjects, including studies from Fromer et al. (N = 159 schizophrenia patients) and Gandal et al. (N = 258 schizophrenia patients). We also compared our list of hub genes with genes associated with schizophrenia using gene expression imputation (Transcriptome Wide Analysis, TWAS) across multiple brain regions in 40,299 schizophrenia cases and 65,264 matched controls. To test the degree of overlap, we used the hypergeometric test implemented in the userListEnrichment function from the WGCNA package.

The gene sets of selected hub genes were imported to ClueGO v2.17 to perform a gene set enrichment analysis. The Biological Processes of the Gene Ontology databases were selected for the enrichment analysis. Genes involved in each network were mapped to their enriched Biological Processes based on the hypergeometric test (two-sided), with the p-value being corrected using the Benjamini–Hochberg method (adjusted P < 0.05 was considered significant).

DATA AVAILABILITY

Gene expression data that support the findings of this study have been deposited in Gene Expression Omnibus (GEO) Database (https://www.ncbi.nlm.nih.gov/geo/) with the accession code GSE193818. The clinical data that support the findings of this study are not openly available due to contain human data and are available from the corresponding author upon reasonable request.

Received: 14 June 2021; Accepted: 1 February 2022;
Published online: 27 April 2022

REFERENCES

1. Guloksuz, S. et al. Examining the independent and joint effects of molecular genetic liability and environmental exposures in schizophrenia: results from the EUGEI study. World Psychiatry 18, 173–182 (2019).

2. Mas, S. et al. Examining gene-environment interactions using aggregate scores in a first-episode psychosis cohort. Schizophr. Bull. 46, 1019–1025 (2020).

3. Pardinas, A. F. et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat. Genet. 50, 381–389 (2018).
4. Sullivan, P. F. & Geschwind, D. H. Defining the genetic, genomic, cellular, and diagnostic architectures of psychiatric disorders. Cell 177, 162–183 (2019).

5. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat. Neurosci. 18, 199–209 (2015).

6. Gassó, P. et al. Microarray gene-expression study in fibroblast and lymphoblastoid cell lines from antipsychotic-naive first-episode schizophrenia patients. J. Psychiatr. Res. 95, 91–101 (2017).

7. Fromer, M. et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat. Neurosci. 19, 1442–1453 (2016).

8. Gandal, M. J. et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. Science 359, 693–697 (2018).

9. Gandal, M. J., Leppa, V., Won, H., Parikshak, N. N. & Geschwind, D. H. The road to precision psychiatry: translating genetics into disease mechanisms. Nat. Neurosci. 19, 1397–1407 (2016).

10. Gudmundsdottir, V. et al. Whole blood co-expression modules associate with metabolic traits and type 2 diabetes: an IMI-DIRECT study. Genome Med. 12, 109 (2020).

11. Horvath, S. & Dong, J. Geometric interpretation of gene coexpression network analysis. PLoS Comput. Biol. 4, e1000117 (2008).

12. Horvath, S. Weighted Network Analysis. Applications in Genomics and Systems Biology (Springer Book, 2011).

13. Langfelder, P., Mischel, P. S. & Horvath, S. When is hub gene selection better than standard meta-analysis? PLoS ONE 8, e61505 (2013).

14. Yan, Z. & Rein, B. Mechanisms of synaptic transmission dysregulation in the prefrontal cortex: pathophysiological implications. Mol. Psychiatry https://doi.org/10.1038/s41380-021-01092-3 (2021).

15. Radulescu, E. et al. Identification and prioritization of gene sets associated with schizophrenia risk by co-expression network analysis in human brain. Mol. Psychiatry 25, 791–804 (2020).

16. Wang, D. et al. A comprehensive analysis of the effect of SIRT1 variation on the risk of schizophrenia and depressive symptoms. Front. Genet. 11, 832 (2020).

17. Fang, X., Chen, Y., Wang, Y., Ren, J. & Zhang, C. Depressive symptoms in schizotypy. Schizophr. Res. 215, 222–228 (2015).

18. Hope, S. et al. Similar immune profile in bipolar disorder and schizophrenia: selective increase in soluble tumour necrosis factor receptor 1 and von Willebrand factor. Bipolar Disord. 11, 726–734 (2009).

19. Lee, Y. S. et al. von Willebrand’s disease and psychotic disorders: co-segregation and genetic associations. Bipolar Disord. 16, 150–155 (2004).

20. Openshaw, R. L. et al. Map2k7 haploinsufficiency induces brain imaging endophenotypes and behavioral phenotypes relevant to schizophrenia. Schizophr. Bull. 46, 211–223 (2020).

21. Openshaw, R. L. et al. JNK signalling mediates aspects of maternal immune activation: importance of maternal genotype in relation to schizophrenia risk. J. Neuroinflammation 16, 18 (2019).

22. Openshaw, R. L., Thomson, D. M., Penningger, J. M., Pratt, J. A. & Morris, B. J. Mice haploinsufficient for Map2k7, a gene involved in neurodevelopment and risk for schizophrenia, show impaired attention, a vigilance decrement deficit and unstable cognitive processing in an attentional task: impact of minocycline. Psychopharmacology 234, 293–305 (2017).

23. Winchester, C. L. et al. Converging evidence that sequence variations in the novel candidate gene MAP2K7 (MKK7) are functionally associated with schizophrenia. Hum. Mol. Genet. 21, 4910–4921 (2012).

24. Smeland, O. B. et al. Genetic overlap between schizophrenia and volumes of hippocampus, putamen, and intracranial volume indicates shared molecular genetic mechanisms. Schizophr. Bull. 44, 854–864 (2018).

25. Leirer, D. J. et al. Differential gene expression analysis in blood of first episode psychosis patients. Schizophr. Res. 209, 88–97 (2019).

26. Bernardo, M., Amoretti, S., Cuesta, M. J., Parellada, M. & Mezquida, G. 2EPs Group. The prevention of relapses in first episodes of schizophrenia: The 2EPs Project, background, rationale and study design. Rev. Psiquiatr. Salud Ment. https://doi.org/10.1016/j.rpsm.2020.07.004 (2020).

27. American Psychiatric Association. DSM-5: Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994).

28. Andreasen, N. C. et al. Remission in schizophrenia: proposed criteria and rationale for consensus. Am. J. Psychiatry 162, 441–449 (2005).

29. Williams, J. B. et al. The Structured Clinical Interview for DSM-III-R (SCID). II. Multisite test–retest reliability. Arch. Gen. Psychiatry 49, 630–636 (1992).

30. Kaufman, J. et al. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. J. Am. Acad. Child Adolesc. Psychiatry 36, 980–988 (1997).

31. guy, W. ECDEU Assessment Manual for Psychopharmacology-revised. (Department of Health, Education and Welfare, 1976).

32. Rosa, A. R. et al. Validity and reliability of the functioning assessment short test (FAST) in bipolar disorder. Clin. Proct. Epidemiol. Ment. Health 3, 5 (2007).
ACKNOWLEDGEMENTS

This study was supported by the Carlos III Healthcare Institute, the Spanish Ministry of Science, Innovation and Universities, the European Regional Development Fund (ERDF/FEDER) (PI08/0208, PI11/00325, and PI14/00612), Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), CERCA Program, Catalan Government, the Secretariat of Universities and Research of the Department of Enterprise and Knowledge (2017SGR1562 and 2017SGR1355), and Institut de Neurociències, Universitat de Barcelona. The authors thank the Language Advisory Service at the University of Barcelona for the paper revision. The authors also thank all subjects and their families for the time and effort spent on this study, as well as Ana Meseguer for sample-collection assistance.

AUTHOR CONTRIBUTIONS

The results presented here are part of a broader project, the 2EPS study. M.Be. is the coordinator of the 2EPS study. N.F. and P.G. performed the statistical analysis and wrote the first draft of the manuscript, and both authors contributed equally to this work. A.M.P. performed the sample isolation and preparation and participated in the statistical analysis. A.G.S. participated in the sample preparation. G.M. and M.Bi. participated in the coordination of the sample shipment, the maintenance of the 2EPS database, and in the recruitment and assessment of the sample. E.B. is the coordinator of the Biological module of the 2EPS study. L.M.-J., J.G.P., I.Z., M.M., R.R.J., I.C., S.S., A.I., A.B., F.C., M.J.C., M.P., and A.G.P. participated in the recruitment and assessment of the sample. S.M. designed, supervised, and performed the statistical analysis, performed the interpretation of the results, and wrote the first draft of the manuscript. All the authors, including the 2EPSs group authors listed in the acronym, contributed to the final draft of the manuscript.

COMPETING INTERESTS

A. Ibáñez has received research support from or served as speaker or advisor for Janssen-Cilag, Lundbeck, and Otsuka. A. Gonzalez-Pinto has received grants and served as consultant, advisor, or CME speaker for the following entities: Janssen-Cilag, Lundbeck, Otsuka, Pfizer, Sanofi-Aventis, Exeltis, the Spanish Ministry of Science and Innovation (CIBERSAM), the Ministry of Science (Carlos III Institute), and the Basque Government. J. Saiz-Ruiz has been as speaker for and on the advisory boards of Adamed, Lundbeck, Servier, Medtronic, Casen Recordati, Neurofarmagen, Otsuka, Indivior, Lilly, Schwabe, Janssen, and Pfizer, outside the submitted work. M. Bernardo has been a consultant for, received grant/research support and honoraria from, and been on the speakers/advisory board of ABBiotics, Adamed, Angelini, Casen Recordati, Janssen-Cilag, Menarini, Rovi, and Takeda. M. Bioque has been a consultant for, received grant/research support and honoraria from, and been on the speakers/ advisory board of, has received honoraria from talks and/or consultancy of Adamed, Angelini, Ferrer, Janssen-Cilag, Lundbeck, Neuraxpharm, Otsuka, Pfizer, and Sanofi, and grants from Spanish Ministry of Health, Instituto de Salud Carlos III. M. Gutiérrez Fraile has been on the speakers/advisory board of Janssen-Cilag. Pilar A. Saiz has been a consultant to and/or has received honoraria or grants from Adamed, CIBERSAM, European Commission, Government of the Principality of Asturias, Instituto de Salud Carlos III, Janssen-Cilag, Lundbeck, Otsuka, Pfizer, Plan Nacional Sobre Drogas, and Servier. R. Rodríguez-Jimenez has been a consultant for, spoken in activities of, or received grants from Instituto de Salud Carlos III, Fondo de Investigación Sanitaria (FIS), Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid Regional Government (S2010/BMD-2422 AGES, S2017/BMD-3790), Janssen-Cilag, Lundbeck, Otsuka, Pfizer, Ferrer, Juste, Takeda, Exeltis, Casen-Recordati, and Angelini.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41537-022-00215-1.

Correspondence and requests for materials should be addressed to Miquel Bernardo or Sergi Mas.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2022

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022

2EPS GROUP

Silvia Amoretti1,2,4, Constanza Moreno28,29, Carol Stella7,8, Xaquin Gurrriaran7,8, Anna Alonso-Solis8,16,17, Eva Grasa8,16,17, Jessica Fernandez8,9,10,11, Itxaso Gonzalez-Ortega8,9,10,11, Francesc Casanovas30, Antoni Bulbuena31, Ágatha Núñez-Doyle32, Miguel Gutiérrez33,34, Olga Jiménez-Rodríguez32, Edith Pomarol-Clotet33,34, Isabel López-Torres39, Juan Nacher40,41,42, Concepción De-la-Cámara34,44,45, Ana María Sánchez-Torres5,6, Jeronimo Saiz-Ruiz8,20, Isabel Feria-Raposo35, Judith Usall36,37, Daniel Muñoz-Samons38, Jose L. Ilundain5,6,7,1.

© The Author(s) 2022

Schizophrenia (2022) 45

Published in partnership with the Schizophrenia International Research Society
