Altered Expression of a Unique Set of Genes Reveals Complex Etiology of Schizophrenia
Ashutosh Kumar, Vikas Pareek, Himanshu N Singh, Muneeb A Faiq, Ravi K Narayan, Khursheed Raza, Pavan Kumar

To cite this version:

HAL Id: hal-02458882
https://amu.hal.science/hal-02458882
Submitted on 29 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Altered Expression of a Unique Set of Genes Reveals Complex Etiology of Schizophrenia

Ashutosh Kumar¹,²,³*, Vikas Pareek³,⁴, Himanshu N. Singh³,⁵, Muneeb A. Faiq³,⁶, Ravi K. Narayan¹,³, Khursheed Raza²,³ and Pavan Kumar³,⁷

¹ Department of Anatomy, All India Institute of Medical Sciences (AIIMS), Patna, India, ² Department of Anatomy, All India Institute of Medical Sciences (AIIMS), New Delhi, India, ³ Etiologically Elusive Disorders Research Network (EEDRN), New Delhi, India, ⁴ Computational Neuroscience and Neuroimaging Division, National Brain Research Centre (NBRC), Manesar, India, ⁵ TAGC—Theories and Approaches of Genomic Complexity, Aix Marseille University, Inserm U1090, Marseille, France, ⁶ Neuroimaging and Visual Science Laboratory, New York University (NYU) Langone Health Centre, NYU Robert I. Grossman School of Medicine, New York, NY, United States, ⁷ Developmental Neurogenetics Lab, Department of Pediatrics, Medical University of South Carolina, Charleston, SC, United States

Background: The etiology of schizophrenia is extensively debated, and multiple factors have been contended to be involved. A panoramic view of the contributing factors in a genome-wide study can be an effective strategy to provide a comprehensive understanding of its causality.

Materials and Methods: GSE53987 dataset downloaded from GEO-database, which comprised mRNA expression data of post-mortem brain tissue across three regions from control (C) and age-matched subjects (T) of schizophrenia (N = Hippocampus [HIP]: C-15, T-18, Prefrontal cortex [PFC]: C-15, T-19, Associative striatum [STR]: C-18, T-18). Bio-conductor—affy—package used to compute mRNA expression, and further t-test applied to investigate differential gene expression. The analysis of the derived genes performed using the PANTHER Classification System and NCBI database. Further, a protein interactome analysis of the derived gene set was performed using STRING v10 database (https://string-db.org/).

Results: A set of 40 genes showed significantly altered (p < 0.01) expression across all three brain regions. The analyses unraveled genes implicated in biological processes and events, and molecular pathways relating basic neuronal functions.

Conclusions: The aberrant expression of genes maintaining basic cell machinery explains compromised neuronal processing in SCZ.

Keywords: genome-wide expression study, genetic signature, hippocampus, associative striatum, prefrontal cortex

INTRODUCTION

The etiology of schizophrenia (SCZ) is extensively debated (1, 2). Uncertainty of the etiology has greatly impeded the treatment of the disease, and neither of the therapeutic approaches (1) is proving much helpful in halting its progression.

Many candidate genes have been reported (3) for SCZ but none of them got validated in population-based studies for persistent association (4, 5). A disease signature derived from
genome-wide expression patterns in affected brain regions was highly desirable that would help to reach the diagnosis and developing optimal therapeutic approaches for SCZ.

SCZ gives a lifetime risk of ~1% and shows high heritability (~69–81%) (6–8). The SCZ heritability is derived from CNVs, SNPs, de novo mutations, and structural modifications at gene promoter regions without involving gene sequences as have been revealed in the genome-wide studies (9). Expression derangement of the genes also evidenced to arise of the gene-environment interactions during fetal development and in the lifetime of the individuals (10, 11).

SCZ has been noted to cause significant architectural changes in many brain regions, the hippocampus, prefrontal cortex, and basal nuclei (more specifically associative or dorsal striatum) have been chief among them (12–14). Neuronal gene expressions in the affected brain regions are known to alter in SCZ (15, 16). A comprehensive study of the gene expression dynamics in SCZ patients, which is shared by all three brain regions (hippocampus, prefrontal cortex, and associative striatum) may plausibly give a glimpse of the disease etiology.

How the neural architectural changes are instructed by the changes in the neural genes has also been shown by some recent studies. Piskorowski et al. (17) have shown in the mouse model that deletion of 22q11 locus may involve the genes making synaptic proteins and that may produce SCZ like symptoms (17). Fromer and colleagues (18) have identified over 100 of genetic loci harboring SCZ associated variants which together involve scores of genes, and altering the expression or knock down of some of such genes in animal or human stem cell models has shown to compromise neural functions effectively (18). Plausibly, the dysregulation of the neuronal genes, especially which are involved in maintaining basic cell architecture and machinery, may compromise the information processing in neurons in affected brain regions in SCZ (19, 20).

In this study, we hypothesized that the contributory factors involved in the etiogenesis of SCZ may get reflected in the altered expression of neuronal genes; hence an ontological analysis of these genes from the affected brain regions may unravel the components of the complex etiology of SCZ.

MATERIALS AND METHODS

Data Resources

The mRNA expression data were retrieved from the GEO (Genome Expression Omnibus, GSE53987) (http://www.ncbi.nlm.nih.gov/geo/), a public repository for high-throughput microarray. The RNA was originally isolated from post-mortem brain tissue across three specific regions (hippocampus [HIP], prefrontal cortex [PFC]; Brodmann Area 46 [dorsolateral part of PFC], and Associative [or dorsal] striatum [STR]) of control (N = 18 [HIP], 19 [PFC], 18 [STR]) and age-matched subjects with schizophrenia (N = 15 [HIP], 15 [PFC], 18 [STR]). For the original data, designated tissue samples were acquired from the curated collection of brain samples from the University of Pittsburgh which was permitted by the institute ethics committee (21, 22). Equal numbers of male and female (except for odd number samples) diagnosed SCZ cases and controls of adult age were chosen for this purpose (Table S1). The controls were matched for the age and sex with cases, and were free of any neurological or psychiatric illness during their life course. Tissue was collected from same hemisphere of the brain using same anatomical landmarks in all individuals. The post-mortem interval (PMI) and pH of the brain tissue, storage timing (at ~80 degree Celsius), and RNA integrity number (RIN) (to confirm quality of processed RNA) for the test and controls were maintained to the set standard as were declared in the published records related to the original data source (21, 22). Also, history of taking tobacco products, centrally acting drugs, and any other medications and manner of death for the tests as well control samples were noted from the records (Table S1) (21, 22).

Data Retrieval and Analysis

The RNA was isolated from HIP, PFC (Brodmann Area 46), and associative STR and hybridized to U133_Plus2 Affymetrix chips for m-RNA expression study (21, 22). Expression analysis of mRNA was done by using “affy” package (http://www.bioconductor.org/packages/release/bioc/html/affy.html), which was deposited at Bioconductor and developed in R statistical software program and scripting language. It used three steps to calculate the expression intensities: (i) background correction; (ii) normalization (data were normalized by RMA, subjected to pair wise comparison followed by Benjamini and Hochberg False Discovery rate correction [FDR]), and (iii) expression calculation. After calculation of mRNA expression intensity, a simple unpaired two tailed t-test (significance set at p ≤ 0.01) was applied to the data to filter out the set of genes expressed significantly in all three brain regions.

To categorize the derived significantly altered genes on the basis of their involvement in molecular functions, molecular pathways, and biological events, PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System (http://www.pantherdb.org/) and NCBI gene database (http://www.ncbi.nlm.nih.gov/gene/) were exploited. To construct a protein interactome network of the derived gene set STRING v10 database was used (https://string-db.org/). The pathway enrichment analysis results were extracted using Reactome Pathways of the STRING utility.

RESULTS

A set of 40 genes (protein coding-38; RNA-gene-2) was identified showing statistically significant (p ≤ 0.01) altered mRNA expression in schizophrenic patients in the all three brain regions studied (Table 1). Interestingly, it was observed that most of the genes were down-regulated in all three brain regions (32/40). Also, the same genes in all three brain regions have shown the similar direction of expression changes.

These genes were classified into six categories on the basis of their molecular functions (Figure 1). Further, the genes were
classified into fourteen categories on the basis of involvement in biological processes and events (Table 2). However, some genes belong to more than one category. The protein interactome network analysis of the gene-set revealed weak to strong interaction between only limited genes (Figure S1). Two cluster hubs were identified where more than two genes were observed to be interacting with each other. The salient genes which showed cross-talk in the interactome were ATP5D, UQRC1 and ANAPC5, PSMC3 which are mainly involved in the basic cellular functions like energy production mechanisms and cellular growth, cell cycle regulation, and enzyme activity (Table 2).

Furthermore, in pathway linkage analysis, the gene set was found to link with 36 molecular pathways (Figure 2) that broadly could be placed in seven categories based on their commonality (Table S2).

**DISCUSSION**

The structural and functional brain abnormalities have been repeatedly reported in patients with SCZ (23). The brain regions chosen for this study (hippocampus, prefrontal cortex, and...
Altered Gene-Expressions Reveal Etiology of Schizophrenia

Kumar et al.

The genes selected for the analysis were found significantly altered in all three studied brain regions, these expression changes may be reflecting the actual pathophysiological changes occurring in neuronal functions in SCZ. None of the genes revealed in this study had been reported earlier as a candidate gene for SCZ, hence the new set appealed for a fresh attention to understand the disease etiology.

Involvement of the Gene Set in Molecular Functions

The functional analysis of the gene-set (Figure 1) elucidated the genes being involved in regulation of basic machinery and housekeeping functions of the neurons viz. receptor-ligand binding, catalysis, enzymatic regulation, nucleic acid binding transcription factor activity, structural molecule activity, and transport activities. It is well evident that dysregulation of these basic functions in neurons may manifest in compromised information processing in brain which has been a hallmark of the progressed SCZ (29). The molecular function analysis also showed hierarchy of the functions that may be compromised in SCZ (Figure 1) (the catalysis and receptor-ligand binding being most affected functions).
Involvement of the Gene Set in Biological Processes and Events

The comprehensive influence of the dysregulation of these genes in the pathogenesis of SCZ gets further clarified in the analysis for the involvement in the biological processes and cellular events (Table 2). The implication of genes involved in ubiquitination (Table 2-a), enzyme activity (Table 2-b), and energy production mechanisms (Table 2-c) may point towards a failure of the basic functions in neurons; as ubiquitination is known to regulate the diverse spectrum of cellular functions (30) and the same should be true for the genes encoding enzymes, especially those necessary for mitochondrial functions (ATP5D, PDK4) (31), regulating specific signaling pathways (MAPK9) (32) and involved in phosphorylation (ATP5D, PDK4) or dephosphorylation (PPM1E) (25, 33). The dysregulation of genes involved in energy production (Table 2-c) supports prevailed view in the literature that energy production mechanisms get compromized in SCZ (34, 35).

Furthermore, down-regulation of genes which function as regulator of the cell growth mechanisms (Table 2-d) provides a possible explanation for reduced neuronal cell sizes, synaptic connection and brain volume in specific brain regions noted in schizophrenia (36). The significant upregulation of genes involved in the programmed death (Table 2-e) may indicate pro-apoptotic mechanisms prevailing in particular brain regions in schizophrenia which gets support from some earlier studies (37, 38). Though, the pro-apoptotic mechanisms may not be a generalized feature in SCZ, as we also noted contrary evidence that an anti-apoptotic gene MCL-1 was found significantly upregulated in all three brain regions. (Contrastingly, mMCL-1 regulates cell cycle negatively hence limiting the mitosis (39)).

Also, the significantly altered expression of the genes involved in cytoplasmic vesicular transport and exocytosis (Table 2-f), dynamic regulation of actin and tubulin cytoskeleton (Table 2-g), and ion channel homeostasis (Table 2-h), lipid-binding (PITPNA) and synthesis (CADPS) (Table 2-i) hint of compromised neuronal information processing in SCZ.

Indications From Interactome Analysis of Gene Set

The protein-protein interaction analysis of the gene set revealed cross-talk between ATP5D (Table 2-b, c), UQRC1 (Table 2-c),
and ANAPC5 (Table 2-d, n), PSMC3 (Table 2-b) (Figure S1) which further strengthen the view that basic (neuronal) cell functionary of the selective brain regions could be the focus of the pathogenesis in SCZ. We detected two central hubs showing strong interaction between the genes involved in ubiquitination (ANAPC5), enzyme activity (ATP5D, PSMC3), energy production mechanism or mitochondrial functions (ATP5D), and cell cycle regulation (ANAPC5) (Figure S1). Existing literature suggests that a dysregulation of the noted biological processes and events could be at the core of SCZ pathogenesis (40–42).

**Involvement of the Gene Set in Molecular Pathways**

In pathway linkage analysis (Figure 2, Table S2), the category involving largest number of molecular pathways has been that of neurotransmitters/modulators and neurohormones (Table S2-a) which fits with clinical manifestations of the disease and also gets support from existing theories that the etiology of SCZ majorly may be based on dysregulation of this category of molecules (43).

Various neurotransmitters-based hypotheses have been proposed for the etiology of SCZ (44) but none of them are primarily explaining causality of the diseases. The result of this study (Figure 2, Table S2) indicates that disease etiology is not implicating any single transmitter but many of them together (45).

The linkage of the immune cell/chemokine mediated pathways (Table S2-b) is strongly supported by literature (46, 47). An immunogenic basis of SCZ etiogenesis had also been brought forward (48) although counter to this hypothesis has also been placed which limits the role of immune function related genes as a solo or major factor in SCZ etiology (49).

In a recent study (21), Lanz et al. who performed pathway analysis of transcriptional profiles in post-mortem samples of HPC, PFC, and associative STR from SCZ patients (using the same dataset which we used in our study) found enrichment of the transcripts involved in inflammatory pathways. Enrichment of the inflammatory pathways was also reported by Scarr et al., who examined transcriptional profiles from PFC of post-mortem SCZ patients (50).

Also, the involvement of growth, differentiation, and survival of neurons in the specific brain regions (Table S2-c) (51, 52) (also discussed in subsection Involvement of the Gene Set in Biological Processes and Events) and pathways related to apoptosis (Table S2-d) (37, 38), and related to protein synthesis (Table S2-e) and degradation (Table S2-f) (53) has been well documented in the literature (also discussed in subsection Involvement of the Gene Set in Biological Processes and Events). The linkage of FGF signaling pathway (Table S2-c) under neuronal growth, differentiation, and survival to SCZ etiology has been corroborated by a freshly published study by Narla et al. (54) who regarded it as a central pathway commanding all other pathways in developing brain strengthening the view that SCZ has a neurodevelopmental etiology (54). In contrast, the linking of the pathways involved in the pathogenesis of major neurodegenerative diseases (Table S2-g) such as Alzheimer, Parkinson, and Huntington’s disease indicates neurodegenerative nature of SCZ.

**Non-Protein-Coding Genes: Unknown Functions**

The neuronal functions associated with two non-coding genes (LOC100507534, LOC100507534) couldn’t be ascertained from the literature but it’s interesting to find significant alterations of these long non-coding RNAs in SCZ which have never been reported before. There are now strong indications that non-coding genes are implicated in SCZ pathology (55).

**LIMITATIONS OF THIS STUDY AND FURTHER RESEARCH**

The confounding factors such as the history of addiction (including alcohol intake) or substance abuse, and drug intake might have some impact on the transcriptional data. A separate analysis of the impacts of these confounding factors (especially that for addiction and substance abuse) might have given additional insights to this study. Additionally, a separate analysis of the impacts of the sex of the subject on the resultant data could have provided important insights regarding sex-specific pathogenesis in SCZ. We couldn’t analyze these factors separately either due to the lack of related data or limitation of the sample sizes.

Though, we have included only those genes for further analysis which showed significant change of expression in all three studied brain regions, the fold changes for the genes are not very large to derive strong conclusions. Validation of the analyzed data with more than one gene expression analysis methods could have been necessary, and could have further augmented the value of this study. The genes which were significantly altered in only one or two and not in all three brain regions selected for the study have not been included in analysis to keep the study design robust, but they might carry some value in disease etiology. A neural circuit specific analysis of the changes in gene expressions targeted to the individual neurocognitive domains may further enhance the etiological clarity on SCZ.

Testing validity of the proposed gene set as a SCZ genetic signature is a remaining task which needs to be studied further. A rigorous search of the SCZ gene expression databases linked to the noted brain regions will make this clear. Also, looking for the similar gene expression changes in the blood cells and/or skin fibroblasts (though there is limited evidence for this in literature for now) (56, 57), will be greatly informative for assigning any prognostic (or diagnostic) value to the gene set.

**CONCLUSIONS**

Molecular characterization of the gene set unraveled in this study gives a glimpse of the complex etiogenetic mechanisms involved in SCZ, an understanding of which may have useful implications in the therapeutic management of the disease. Most
of the genes in the set participate in the maintenance of basic cell machinery which explains why their aberrant expression may cause compromised neuronal processing in SCZ.

DATA AVAILABILITY STATEMENT

The dataset used for this study can be found in the NCBI Genome Expression Omnibus, GEO accession: GSE53987 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53987).

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

AK conceived and designed the study. HS, VP, and AK analyzed the data. AK wrote the first draft. AK, VP, HS, MF, RN, KR, and PK edited the first draft. AK, VP, and RN prepared final draft of the paper.

REFERENCES


ACNOWLEDGMENTS

We are thankful to Thomas A. Lanz, contributor of the GSE53987 dataset at Gene Expression Omnibus, NCBI.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00906/full#supplementary-material

TABLE S1 | Human subject demographics and sample metadata.

TABLE S2 | Involvement of the gene set in molecular pathways.

FIGURE S1 | Protein Interactome Network Analysis of the gene set. Certain of the genes shown in the figure, such as JADE-1 and ENSG00000251349 or MSANTD3-TMEFF1, PSMC6, PSMC2, ATP5F1, are linker genes generated by the database and are not part of the original gene set.

Supplementary Material

Evaluating historical candidate genes for schizophrenia.

What we know in 2008. 2. Epidemiology and etiology.

ACKNOWLEDGMENTS

We are thankful to Thomas A. Lanz, contributor of the GSE53987 dataset at Gene Expression Omnibus, NCBI.

Supplementary Material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00906/full#supplementary-material

TABLE S1 | Human subject demographics and sample metadata.

TABLE S2 | Involvement of the gene set in molecular pathways.

FIGURE S1 | Protein Interactome Network Analysis of the gene set. Certain of the genes shown in the figure, such as JADE-1 and ENSG00000251349 or MSANTD3-TMEFF1, PSMC6, PSMC2, ATP5F1, are linker genes generated by the database and are not part of the original gene set.

Supplementary Material

Evaluating historical candidate genes for schizophrenia.

What we know in 2008. 2. Epidemiology and etiology.

ACKNOWLEDGMENTS

We are thankful to Thomas A. Lanz, contributor of the GSE53987 dataset at Gene Expression Omnibus, NCBI.


43. Scarr U, Udawela M, Dean B. Changed frontal pole gene expression suggest altered interplay between neurotransmitter, developmental, and inflammatory pathways in schizophrenia. NPJ Schizophrenia (2018) 4:4. doi: 10.1038/s41537-018-0044-x


Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Kumar, Pareek, Singh, Faiz, Narayan, Raza and Kumar. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.