Dysregulation of Long Intergenic Non-Coding RNA Expression in the Schizophrenia Brain

Изменения экспрессии длинных некодирующих РНК при шизофрении

doi: 10.17816/CP219

Original research

Tuan Nguyen[1], Olga Efimova[1], Artem Tokarchuk[1], Anna Morozova[2,3], Yana Zorkina[2,3], Denis Andreyuk[3], George Kostyuk[3], Philipp Khaitovich[1]

[1] V. Zelman Center for Neurobiology and Brain Restoration, Skolkovo Institute of Science and Technology, Moscow, Russia
[2] V. Serbsky National Medical Research Centre of Psychiatry and Narcology of the Ministry of Health of the Russian Federation, Moscow, Russia
[3] Mental-health Clinic No. 1 named after N.A. Alexeev, Moscow, Russia

Туан Нгуен[1], Ольга Ефимова[1], Артем Токарчук[1], Анна Морозова[2,3], Яна Зоркина[2,3], Денис Андреюк[3], Георгий Костюк[3], Филипп Хайтович[1]

[1] Центр нейробиологии и нейрореабилитации имени Владимира Зельмана, АНОО ВО «Сколковский институт науки и технологий», Москва, Россия
[2] ФГБУ «Национальный медицинский исследовательский центр психиатрии и наркологии им. В.П. Сербского» Минздрава России, Москва, Россия
[3] ГБУЗ «Психиатрическая клиническая больница № 1 им. Н.А. Алексеева Департамента здравоохранения города Москвы», Москва, Россия

ABSTRACT

BACKGROUND: Transcriptomic studies of the brains of schizophrenia (SZ) patients have produced abundant but largely inconsistent findings about the disorder's pathophysiology. These inconsistencies might stem not only from the heterogeneous nature of the disorder, but also from the unbalanced focus on particular cortical regions and protein-coding genes. Compared to protein-coding transcripts, long intergenic non-coding RNA (lincRNA) display substantially greater brain region and disease response specificity, positioning them as prospective indicators of SZ-associated alterations. Further, a growing understanding of the systemic character of the disorder calls for a more systematic screening involving multiple diverse brain regions.

AIM: We aimed to identify and interpret alterations of the lincRNA expression profiles in SZ by examining the transcriptomes of 35 brain regions.

METHODS: We measured the transcriptome of 35 brain regions dissected from eight adult brain specimens, four SZ patients, and four healthy controls, using high-throughput RNA sequencing. Analysis of these data yielded 861 annotated human lincRNAs passing the detection threshold.

RESULTS: Of the 861 detected lincRNA, 135 showed significant region-dependent expression alterations in SZ (two-way ANOVA, BH-adjusted p <0.05) and 37 additionally showed significant differential expression between HC and SZ individuals in at least one region (*post hoc* Tukey test, p <0.05). For these 37 differentially expressed lincRNAs (DELs), 88% of the differences occurred in a cluster of brain regions containing axon-rich brain regions and cerebellum.

Functional annotation of the DEL targets further revealed stark enrichment in neurons and synaptic transmission terms and pathways.

CONCLUSION: Our study highlights the utility of a systematic brain transcriptome analysis relying on the expression profiles measured across multiple brain regions and singles out white matter regions as a prospective target for further SZ research.

АННОТАЦИЯ

ВВЕДЕНИЕ: Исследования транскриптома мозга пациентов с диагнозом шизофрении (ШЗ) не дали однозначной картины механизмов, лежащих в основе этого расстройства. Данная проблема связана не только с возможной гетерогенностью ШЗ, но также с несбалансированным фокусом исследований на определенных областях коры полушарий и экспрессии белок-кодирующих генов. По сравнению с белок-кодирующими генами, длинные некодирующие РНК (дкРНК) демонстрируют значительно большую специфичность и динамику паттернов экспрессии, что позиционирует их как перспективные маркеры молекулярных изменений мозга при ШЗ. Кроме того, растущее понимание системного характера ШЗ требует более систематического анализа экспрессии дкРНК, охватывающего множественные регионы мозга.

ЦЕЛЬ: Идентифицировать и интерпретировать изменения профилей экспрессии дкРНК при ШЗ в 35 регионах мозга.

МЕТОДЫ: Мы провели анализ транскриптома 35 областей мозга четырех пациентов с диагнозом ШЗ и четырех человек из группы контроля, используя высокопроизводительное секвенирование РНК.

РЕЗУЛЬТАТЫ: Из 861 детектированной дкРНК 135 продемонстрировали глобально значимые изменения уровней экспрессии при ШЗ (двусторонний дисперсионный анализ, скорректировано методом Бенджамини–Хохберга p <0,05). Из них 37 дкРНК показали значимые изменения, локализованные в одном или нескольких регионах мозга (тест Тьюки, p <0,05). Из этих изменений 88% произошли в регионах белого вещества мозга и мозжечке. Функциональная аннотация 37 дкРНК выявила значимую корреляцию с генами нейронов и генами, кодирующими элементы синаптической передачи сигнала.

ЗАКЛЮЧЕНИЕ: Наше исследование подчеркивает полезность систематического анализа транскриптома мозга и выделяет области белого вещества в качестве перспективной цели для дальнейших исследований ШЗ.

Keywords: *schizophrenia; long intergenic non-coding RNA; lincRNA; white matter; transcriptome; brain*
Ключевые слова: *шизофрения; длинные некодирующие РНК; дкРНК; белое вещество; транскриптом; головной мозг*

INTRODUCTION

Schizophrenia (SZ) is a neurodevelopmental disorder listed among the top 15 most burdening disabilities worldwide [1]. Despite decades of research, the etiology of the disease remains elusive due to its complexity, heterogeneity, and polygenicity. Genomic abnormalities may fractionally explain the substantial heritability of SZ but show limitations as diagnostic markers and etiology indicators due to a low fraction of explained disorder risk probability, thus suggesting a substantial role of epigenetic and environmental factors [2]. Gene expression analysis can bridge the gap between genomic and environmental risks, making it a promising approach to studying the pathophysiology of the disease.

Multiple regions in the brains of SZ patients display structural and functional abnormalities in neuroimaging studies. Yet, current molecular analysis remains restricted to a few selected brain areas. There is widespread cortical thinning in SZ individuals, with significant volumetric decreases in the frontal, temporal,

and parietal lobes [3, 4]. Several subregions of these lobes, namely the dorsolateral prefrontal cortex and the superior temporal gyrus, are routinely selected for the transcriptomic profiling of psychiatric diseases [5, 6]. Similarly, a large-scale imaging study of subcortical structures revealed smaller hippocampus, amygdala, and thalamus [7]. Consequently, multiple gene expression studies have investigated particular locations within these regions, particularly those with functional relevance to cognitive and emotional functions, but the findings have been surprisingly inconsistent [8]. Furthermore, most of the studies focused on a single region, limiting our capability to unfold the molecular networks underlying such a multiplex disorder like SZ. Furthermore, SZ-associated transcriptome alterations have been found in other "not-as-popular" parts of the brain, such as the parietal lobe [9] and the cerebellum [10, 11]. Yet, these regions are virtually neglected in psychiatric research, leaving their mechanistic involvement in SZ pathology unknown.

Long non-coding RNAs (lncRNAs) play a critical role in gene expression regulation in the brain, with disruption of this regulation implicated in various mental disorders, including SZ [12, 13]. Even though lncRNAs are usually synthesized by RNA polymerase II, similar to the protein-coding transcripts, they vastly exceed the mRNAs in terms of diversity, especially in nervous tissue [14]. Nonetheless, most of the expression analyses of post-mortem brains have focused on protein-coding RNAs, leaving the non-coding component of the transcriptome unexplored. Within the brain, many lncRNAs are specifically expressed in particular regions and at defined developmental stages [14, 15]. Therefore, alterations of lncRNA expression patterns could be affiliated with the disrupted developmental programming postulated by the neurodevelopmental hypothesis of SZ. A growing number of lncRNAs are documented to be regulators at multiple levels of gene expression affecting biological processes encompassing neuronal differentiation and the immune response [15, 16]. Due to the widespread comorbidities and substantial overlap of behavioral symptoms among psychiatric illnesses, many candidate lncRNA regulators could be connected to more than one disease [16].

The largest class of lncRNA is long intergenic non-coding RNA (lincRNAs), which, in addition to the length and non-translated requirement of lncRNA, do not overlap with protein-coding sequences. Compared to mRNAs, lincRNAs are less conserved, less efficiently spliced, and more tissue-specific despite sharing similar biogenesis pathways [17]. Aside from the overlapping features, a few other aspects support the distinction between lincRNAs and the other intragenic lncRNAs [18]. Herein, we identified lincRNAs associated with SZ and annotated their biological functions by comparing the transcriptomes of 35 anatomical regions corresponding to 10 anatomical sections in the post-mortem brains of four healthy and four SZ-diagnosed individuals (Figure 1, Table S1 in the Supplementary).

METHODS

Study design

A postmortem comparative study was conducted jointly by the Mental-health Clinic No. 1 named after N.A. Alexeev and Skolkovo Institute for Science and Technology. The inclusion criterion particular to SZ patients was: paranoid SZ (F20) diagnosed according to ICD-10. The inclusion criteria particular to the healthy control (HC) group were: no history of psychiatric or other brain-related disorders and age and sex range matching that of the SZ patient group. The inclusion criteria shared by both groups were: written informed consent for the collection of postmortem brain sample material for any type of noncommercial biological studies and anonymized processing of socio-demographic, medical, psychometric data; sudden death with no prolonged agony state from causes not directly related to brain function; and tissue collection interval shorter than 24 hours postmortem.

Brain samples

Our study used samples dissected from eight frozen human brains, four from healthy donors (HC), and four from SZ-diagnosed ones (Table S2 in the Supplementary). Post-mortem human brain samples were obtained from the National BioService Russian Biospecimen CRO, St. Petersburg, Russia. Informed consent for the use of human tissues for research was obtained in writing from all donors or next of kin by the tissue provider bank. All HC subjects were defined as healthy with respect to the sampled brain tissue by medical pathologists at the Mental-health Clinic No. 1 named after N.A. Alexeev. From each brain, we dissected 35 regions, listed in Table S1 in the Supplementary, without thawing. The dissected specimens were then preserved at −80°C until RNA extraction.

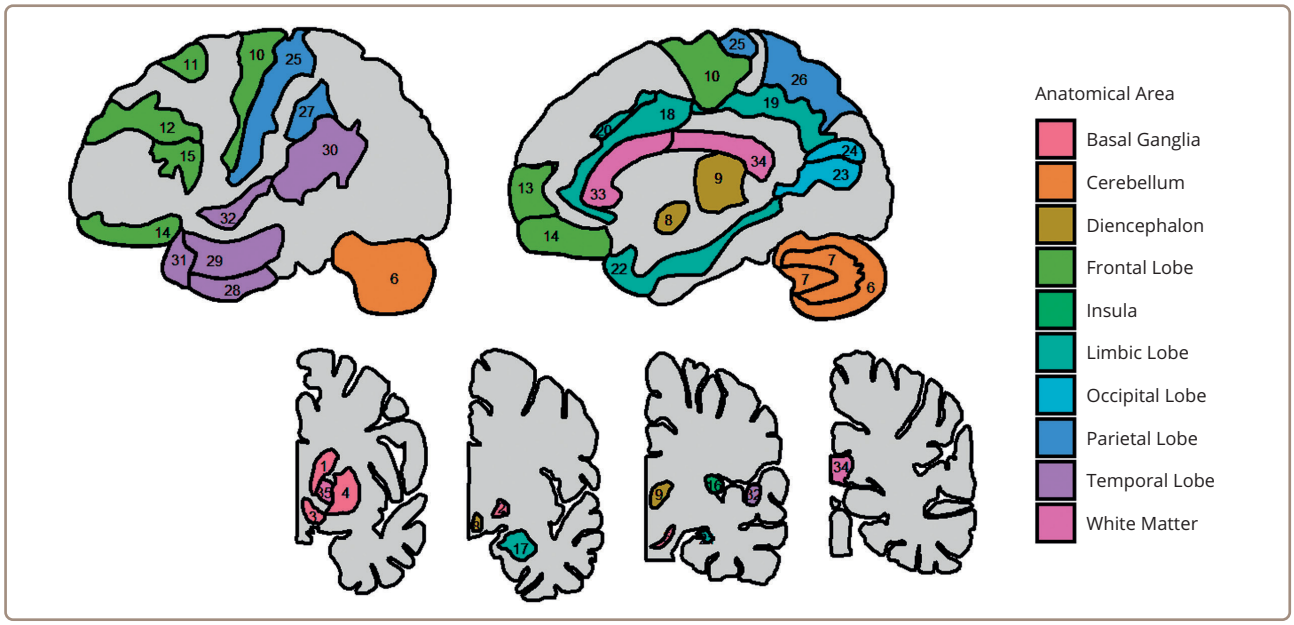

Figure 1. Schematic representation of the 35 examined human brain regions.
Note: Colors indicate the anatomical areas containing the regions. Numbers within the colored regions indicate the corresponding anatomical structures listed in Table S1 in the Supplementary.

RNA library preparation and sequencing data assessment

From each sample, RNA was extracted from an approximately 30-milligram block following the manufacturer's protocol of RNA extraction with QIAzol Lysis Reagent with no modifications [19]. After RNA integrity and concentration were measured, sequencing libraries were prepared following the manufacturer's poly-A selection protocol with no modifications [20]. The libraries were sequenced using the Illumina Hiseq 4000 platform.

FastQC [21] was used to assess the quality of raw reads. We used Trimmomatic [22] to remove low-quality reads and all adapters identified previously or provided by Trimmomatic. We mapped the reads to the human reference genome GRCh38 with HISAT2 [23]. The gene count matrix representing gene expression values was retrieved as transcripts per million (TPM) using Stringtie [24]. Gene annotations were obtained from Ensembl v91 [25].

Differential gene expression analysis

Genes identified as long intergenic non-coding RNAs (lincRNAs) and protein-coding (mRNAs) according to Ensembl annotation were chosen for downstream analysis. For both the control and disease samples, we only considered the genes with no more than two zero coverage values among eight samples representing each region (two-zero threshold). The TPM count data were transformed to the logarithmic scale using the package *DESeq2* [26]. We adjusted the expression levels for the sample quality using the linear regression analysis with RIN (RNA Integrity Number) values. Finally, we used donor-centered normalization for each individual by subtracting the means of log-transformed expression values calculated based on expression of the 35 regions of a given brain from the regional expression values of the respective individual. We conducted the Principle Component Analysis to visualize the variation among 280 examined brain samples using the donor-normalized expression levels of 861 detected lincRNAs. Based on the regional means of pooled HC and SZ groups, we identified three clusters of brain regions using the hierarchical clustering method, in which distances were calculated as one minus Pearson correlation coefficients and clusters were defined by the Ward's linkage function.

For the lincRNAs dataset, Levene's test was used to exclude genes with high heteroscedasticity (the threshold for exclusion was p <0.05) yielding 768 detected variation-balanced lincRNAs. Two-way ANOVA including the effect of region and conditions (diagnosis) conducted based on donor-normalized RIN-corrected $\log_{10}$-transformed TPM values was used to identify transcripts with significant differences for the interaction terms (p <0.05 after Benjamini–Hochberg correction). These transcripts were chosen for the *post hoc* Tukey test. Differentially

expressed lincRNAs (DELs) were defined as those with significant difference between the two conditions in the Tukey test in at least one brain region. The fold changes within a region were measured by subtracting the mean transformed expression of the control groups from that of the SZ groups. We defined upregulated DELs as the ones showing higher expression in SZ brains compared to the controls and downregulated DELs as the ones showing an opposite expression behavior.

Enrichment analysis

We calculated the Pearson correlation between the donor-normalized RIN-corrected $\log_{10}$-transformed TPM expression values of DELs and all mRNAs passing the two-zero threshold described above. Protein coding genes with a correlation coefficient $r \geq 0.85$ were defined as potential targets of the respective DELs. DELs with at least 10 targets were used for downstream analyses. For each DEL, we analyzed the enrichment of its targets using all coding genes passing the detection threshold as the background set. Gene Ontology (GO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome pathways analyses for all groups were implemented using the clusterProfiler R package [27]. Terms and pathways with an adjusted p-value <0.05 were considered enriched.

We compiled a set of marker genes for each of the eight main brain cell types based on published gene sets. For excitatory and inhibitory neurons, we chose intersections of the corresponding sets from the studies in [28, 29]. Similarly, other neuron markers, microglia, astrocytes, and oligodendrocytes represented the respective intersections of lists reported in [28, 30, 31]. Markers for oligodendrocyte progenitor cells were reported in [28], and the ones for endothelial cells were extracted from [29]. We tested the overrepresentation of these markers in the list of DEL targets using Fisher's exact test, followed by Benjamini-Hochberg correction.

RESULTS

Region-dependent lincRNA expression differences cluster in white matter regions of the brain

Previously, we published transcriptome data covering 33 regions of a healthy human brain [28]. Here, we analyzed gene expression in the same 33 regions in four individuals diagnosed with SZ, and two additional brain regions, temporopolar cortex (BA38) and secondary auditory anterior cortex (BA21a), in the diagnosed and control groups (Figure 1, Table S1 in the Supplementary). Based on these data, we detected the expression of 861 annotated human lincRNAs that passed the intensity threshold. Visualization of expression variation using the principal component analysis (PCA) showed substantial overlap of HC and SZ samples (Figure 2A), while the segregation of the samples with regard to anatomical regions was more evident (Figure 2B).

Aligning with the PCA results, we identified three clusters of brain regions produced by unsupervised hierarchical clustering based on lincRNA expression levels, aligning with the anatomical subdivision of the brain (Figure

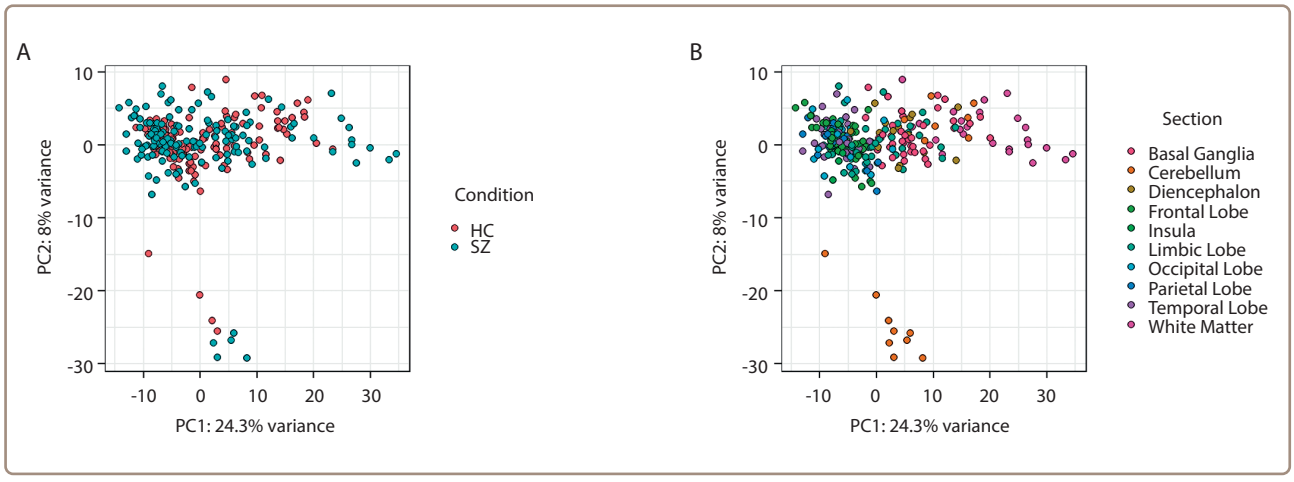

Figure 2. Global patterns of lincRNA expression in the human brain.

Note: (A, B) PCA showing lincRNA expression-based differences among 280 examined brain samples colored by condition (A) or anatomical subdivision (B). Each dot represents a brain sample.

3, Figure S1 in the Supplementary). The first cluster contains mainly the neocortical areas; the second — all the connective nerve tracts and the cerebellum; and the third — the diencephalon and most of the basal ganglia. Substructures of the limbic system belong to both clusters I and III, with the regions spatially related within the clusters.

Out of 861 detected lincRNAs, the expression of 135 depended significantly on both conditions and brain regions (two-way ANOVA, Benjamini–Hochberg (BH)-adjusted p <0.05 for the interaction term, excluding 93 genes with unequal variance). Among these lincRNAs, we identified 37 DELs showing differences between HC and SZ individuals significant in at least one individual brain region (Tukey test, p <0.05; Figure 4A, Table S3 in the Supplementary). Of them, four DELs were dysregulated in two regions, while the rest were dysregulated in one. Further, 31 of the 37 DELs showed a two-fold or greater expression level difference between the two conditions (Table S3 in the Supplementary). Notably, these significant differences were not distributed uniformly within the brain but were associated with 10 out of the 35 examined regions. Most of the associations were found in the regions containing connective axonal tracks: cerebellar white matter contained two down- with 14 up-regulated lincRNAs and three regions of the cerebral white matter contained 13 down- with seven up-regulated lincRNAs (Figure 4B).

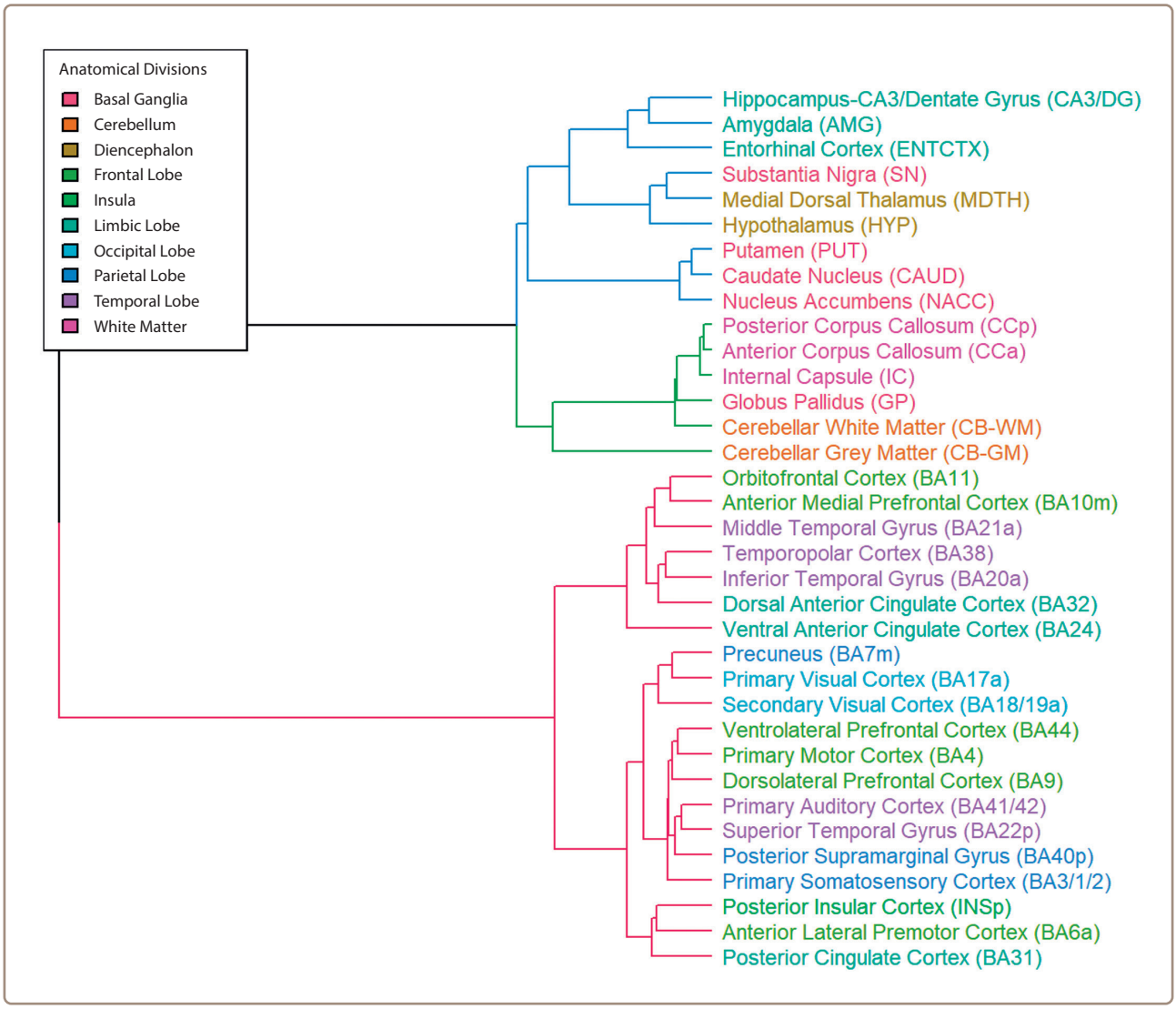

Figure 3. Unsupervised hierarchical clustering of brain regions based on the donor-normalized expression profiles of 861 lincRNAs averaged across HC and SZ samples.

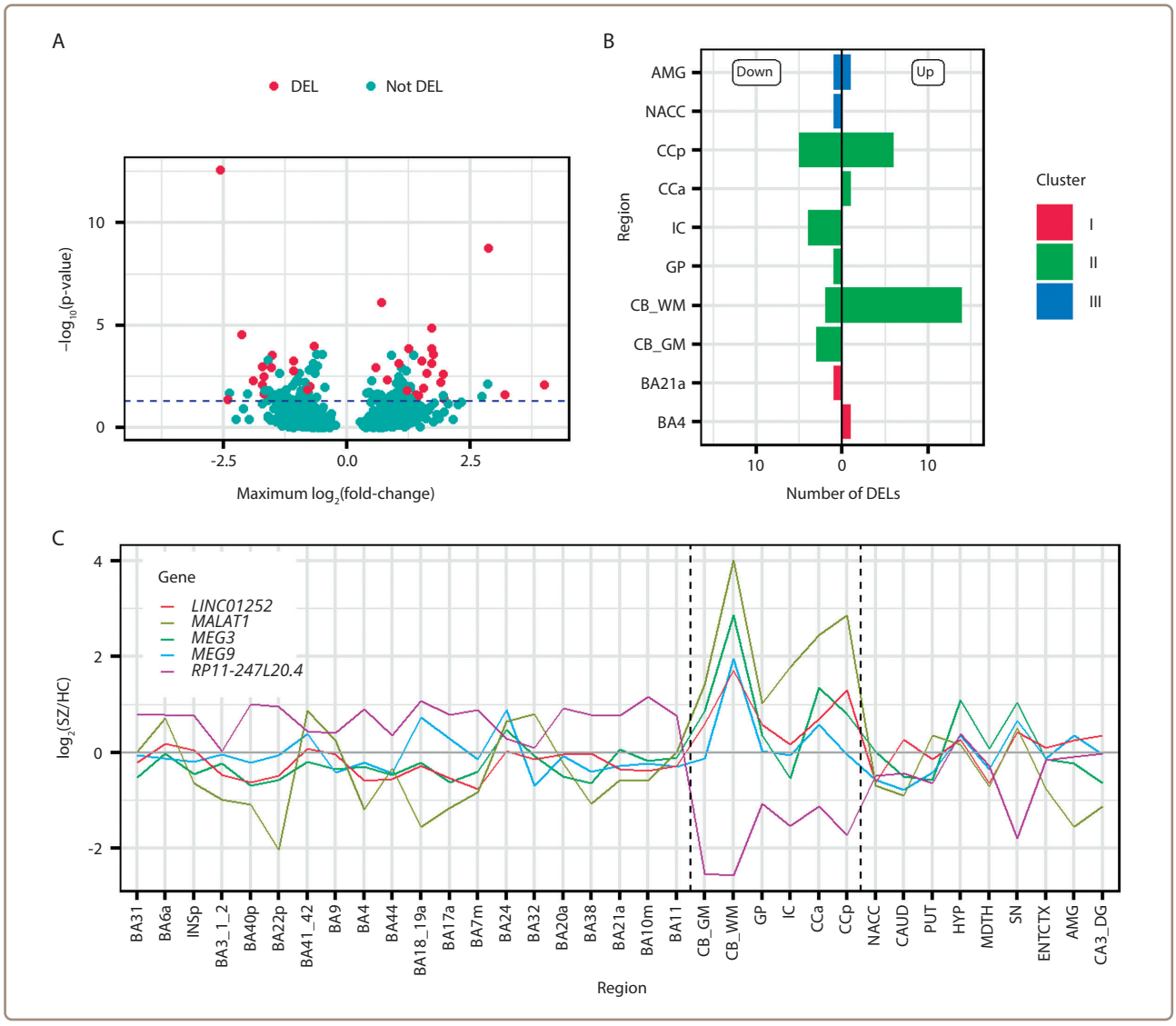

Figure 4. Differentially expressed lincRNAs (DELs) and the corresponding brain regions.
Note: (A) A volcano plot showing 768 lincRNAs used in ANOVA with 37 DELs marked. The vertical axis shows the ANOVA BH-corrected p-value for the region-diagnosis interaction term. The dashed blue line indicates the threshold p=0.05. The plotted fold-change values represent the maximum fold-change among 35 regions for each lincRNA. (B) The number of down- and up-regulated DELs identified in each of the 10 brain regions containing at least one significant expression difference. Colors represent the brain region clusters shown in Figure 2C. (C) The profiles of the difference between the SZ and HC expression levels drawn for five DELs having the lowest SZ/HC comparison p-value, previously reported in the literature as SZ-associated, or both. The dashed black lines delineate the brain region cluster containing white-matter-rich regions.

The DELs showing the most significant difference between SZ and HC expression in our study included the following transcripts: *MEG3*, *RP11-247L20.4*, and *LINC01252*. These lincRNAs were all dysregulated in the white matter of the cerebellum, and *RP11-247L20.4* was also significantly downregulated in the cerebellar gray matter. We illustrated the expression profiles of these five genes in Figure 4C. Notably, *MALAT1* was also the gene with the biggest difference amplitude among DELs, showing a 16-fold increase in the cerebellum of SZ patients compared to HC individuals.

Re-analysis of published lincRNA data sets revealed a positive and significant correlation of SZ-associated fold changes between our data and published lincRNA ones in the amygdala (two studies, Spearman correlation test, p >0.45, p <0.03) [32, 33] but not in the dorsolateral prefrontal cortex (one study, Spearman correlation test, p=0.15, p=0.39) [34] (Table S4 in the Supplementary).

Functional annotation of DELs links them to neuroplasticity and neurotransmission

While human protein-coding genes tend to have substantial functional annotation, this is not the case for the vast majority of lincRNA. Nonetheless, the co-expression of mRNA and lincRNA transcripts could be an indication of the functional roles of non-coding counterparts. To perform such an annotation, we defined mRNAs strongly correlated with a DEL expression difference profile as potential targets of the respective DEL. Whilst most DELs had few or no targets, three DELs correlated with outstanding numbers of mRNAs, which altogether constituted 218 out of the 231 identified DEL-mRNA correlations. These three DELs included *RP11-74E22.8* lincRNA upregulated in the cerebellar white matter and *LINC01963* and *RP11-416I2.1,* both downregulated in the internal capsule, as well as a white matter region. Furthermore, the profiles of *RP11-74E22.8* and *LINC01963* were strongly positively correlated (Pearson r=0.76) and they shared 17 common mRNA targets out of 71 and 105 targets, respectively.

The potential mRNA targets of the three DELs were significantly associated with the neuronal activity terms listed in the GO database (hypergeometric test, BH-adjusted p-values <0.05). Specifically, mRNA targets of *RP11-416I2.1* were associated with voltage-gated channels and neuroplasticity, while targets of the other two DELs were linked to synaptic transmission and signaling terms (Figure 5A, Figures S2A, S2B in the Supplementary). Analysis of mRNA target enrichment using another functional annotation database, the KEGG [35], linked four pathways, including the "synaptic vesicle cycle" and "metabolism of alanine, aspartate, and glutamate", with targets of *LINC01963* (Figure S2B in the Supplementary). Functional annotation using Reactome Knowledgebase [36] yielded similar results:

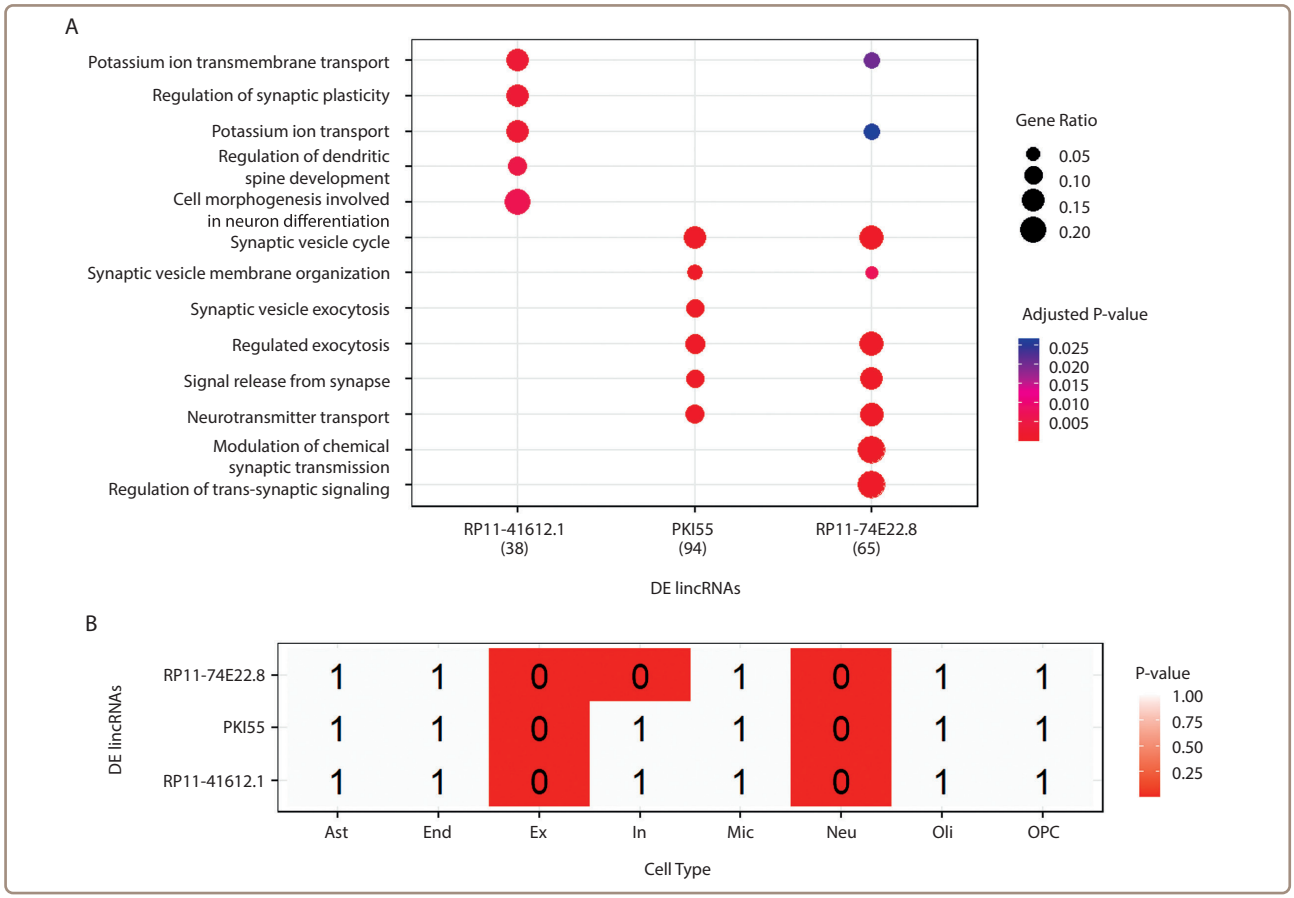

Figure 5. Annotation of three lincRNAs with the highest number of correlated mRNAs.
Note: (A) Top GO biological process terms enriched in each target group. (B) Brain cell type markers enriched in each target group. Ast: Astrocyte, End: Endothelium, Ex: Excitatory neuron, In: Inhibitory neuron, Mic: Microglia, Neu: Neuron, Oli: Oligodendrocyte, OPC: Oligodendrocyte progenitor cell.

the targets of *RP11-416I2.1* were enriched in potassium channels and G-protein-coupled receptors, while the other two target groups shared mutual entries related to the neurotransmitter release cycle (Figure S2D in the Supplementary).

We further investigated the association of DEL target mRNAs with eight main brain cell types by testing them using a customized list of marker genes extracted from publications. The analysis revealed an evident and significant association between the targets of all three target-rich DELs and general neuronal markers, as well as markers of excitatory neurons (Figure 5B). This result aligns with the functional annotation outcome dominated by terms related to neuronal functionality. In addition, the targets of *RP11-74E22.8* overlapped significantly with inhibitory neuron markers, which might be related to their enrichment in the GABA, dopaminergic, and norepinephrine pathways. These results suggest that the three DELs could modulate a network of genes expressed in neurons and involved in synaptic signal transduction.

DISCUSSION

Our analysis of lincRNA expression in the SZ patients' brains revealed few alterations in the cerebral cortex and basal ganglia regions commonly thought to be associated with the disorder. Instead, it shows substantial lincRNA dysregulation in the cortical white matter regions and the cerebellum. Our study defers from most of the previous SZ brain expression analyses in two substantial aspects. First, by measuring gene expression in multiple regions of the same brain, we based our analysis on the expression profiles of the transcripts within the brain, thus minimizing any interindividual variation. This approach allowed us to focus on the expression differences particular to specific brain regions, including the ones neglected by previous studies. Interindividual variation poses a serious problem in human studies due to uncontrollable and diverse genetic and environmental factor effects, resulting in the loss of biologically meaningful differences with marginal to modest effect sizes [37]. Most existing gene expression studies of the SZ brain focus on either a single or a few regions of the cerebrum [8]. However, sporadic omics screening of the neglected brain regions, such as the transcriptome and proteome assessment of the cerebellum [10, 11], have reported meaningful expression alterations. In this study, focusing on the expression alteration patterns recorded across the 35 brain regions, we show that the white matter and the cerebellum might warrant more attention in future SZ studies.

The main components of the white matter are myelinated axons extending from the neuronal cell bodies. Thus, it might seem unusual that the observed lincRNA expression alterations were not accompanied by changes in the corresponding gray matter. One hypothesis is that the changes in the gene expression of these regions arose mainly from local glial cells. This notion, however, is unlikely to hold for the cerebellum gray matter, given that the non-neuronal cells account for less than one-fifth of the total cell population in this structure [38]. Thus, gene dysregulations in glial cells have to be substantial in order to explain the observed differences. Another possible explanation of the observed differences is the redistribution of transcripts leading to the accumulation of the DELs in axons, possibly due to molecular transport impairment. This explanation aligns with the white matter pathology of SZ [39] and could be linked to a transcriptomic study of the cerebellum reporting dysregulation of the genes involved in the Golgi function and presynaptic vesicular transport in SZ [11]. Alternatively, the disconnected patterns of white matter regions and the cerebellum from the rest of the brain could be a consequence of cell-type-specific expression. There are many nuclei in the cerebellar and cerebral cortices, including all of the cerebellar granular layer neurons, that do not project into the other brain regions. Thus, the transcriptomes of such non-projecting neurons might not be reflected in the white matter transcriptome. Future experiments should investigate the cellular locations of these molecules to uncover the underlying mechanism and clarify these speculations.

The second particular aspect of our study is its focus on non-coding RNA expression. Unlike mRNA, lincRNA expression displays more pronounced tissue and brain region specificity, as well as greater response amplitudes, making them better perspective markers of disorder-related alterations [40]. On the other hand, the evident drawback of the lincRNA research is an almost complete lack of functional annotation, hindering results interpretation. In our study, however, we were able to largely overcome this limitation by taking advantage of a lincRNA-mRNA co-expression analysis relying on the transcript profiles measured across the 35 brain regions. The reliance on these profiles, instead of the variation-prone inter-individual

comparisons, allowed us to unambiguously assign co-expressed lincRNA targets to neurons and neuron-specific functions, such as synaptic signal transduction. It is also noteworthy that a substantial fraction of the DELs identified in our study overlapped with lincRNAs previously reported by the few corresponding analyses. At the level of individual lincRNA, *MEG3* has been previously reported to be differentially expressed in the hippocampus [41], superior temporal gyrus [42], and amygdala [32, 43] of SZ patients. Similarly, *LINC01252* has been reported to be upregulated in SZ in the amygdala [33]. Two other DELs that have appeared in the related literature included *MEG9* dysregulation in the amygdala [32] and *MALAT1* — in the dorsolateral prefrontal cortex [34]. The direction of effect reported for these lincRNAs coincided with the one found in our analysis, with the sole exception of the *MEG3* expression difference in the hippocampus, where the difference was not statistically significant in our study. Besides particular lincRNAs, three published studies contained lincRNA datasets: two from amygdala and one from prefrontal cortex [32–34]. While the agreement of our results with the reported amygdala differences was significant, in the cortex we only detected a positive correlation trend. The absence of a statistically significant agreement in the prefrontal cortex could be due to insufficient power of the comparison and, more importantly, lack of substantiation of lincRNA expression alterations in this region. Our general analysis, as well as the expression profiles of the five selected DELs, shows the concentration of large-amplitude SZ-associated expression differences in white matter regions and cerebellar gray matter, with some significant differences also found in the amygdala, but none in the prefrontal cortex. These preliminary findings support further confirmative research focusing on disease-associated non-coding RNA.

The main limitation of our work is the low count of investigated brains per group. Although our study included 35 regions from each individual and the number of samples per group was balanced, we had only four biological replicates in each group. This number is certainly low for such a heterogeneous disease as SZ. The clinical presentation of SZ is highly diverse, so whether the subjective, behavior-based diagnosis of the disease agrees with the pattern of molecular alterations remains a topic of controversy [44]. However, this limitation is universal and represents a problem largely unresolved in most post-mortem brain studies [2, 8]. Unlike other studies, our analysis minimizes inter-individual variation by using the average expression level of each transcript within a given brain as an internal control. As a result, despite the limited sample size, we identified numerous DELs whose involvements in SZ could be supported by previous SZ studies, as well as their evident association with the mRNA transcripts involved in neuronal-specific functionality, suggesting biologically meaningful signals.

CONCLUSION

Our analysis of long, non-coding RNA expression patterns across 35 diverse brain regions reveals the clustering of SZ-associated expression alterations in brain structures routinely neglected by transcriptome studies: white matter and cerebellar brain regions. Further, the identified lincRNA expression alterations were associated with mRNA preferentially expressed in neurons and involved in neuron-specific functions, such as synaptic transmission. These results strongly indicate that further studies of SZ molecular mechanisms should involve a broad selection of brain structures, including the white matter regions and cerebellum.

Article history:
Submitted: 09.10.2022
Accepted: 07.11.2022
Published Online: 07.12.2022

Acknowledgments: We thank all authors for their contributions to the research.

Authors' contribution: A.Yu. Morozova, Ya.A. Zorkina, D.S. Andreyuk — performed clinical diagnostic assessment of the SZ patients and curated the brain sample collection; O.I. Efimova — carried out brain sample dissection; T. Nguyen, A.V. Tokarchuk — performed formal analysis; T. Nguyen, Ph.E. Khaitovich — wrote the manuscript; G.P. Kostyuk, Ph.E. Khaitovich — designed the study and supervised the work. All authors read and approved the final manuscript.

Funding: This work was supported by the Russian Science Foundation under grant No. 22-15-00474.

Conflict of interest: The authors declare no conflicts of interest.

Supplementary data

Supplementary material related to this article can be found in the online version at doi: 10.17816/CP219

For citation:

Nguyen T, Efimova OI, Tokarchuk AV, Morozova AYu, Zorkina YaA, Andreyuk DS, Kostyuk GP, Khaitovich PhE. Dysregulation of long intergenic non-coding RNA expression in the schizophrenia brain. Consortium Psychiatricum 2023;4(1):CP219. doi: 10.17816/CP219

Information about the authors

Tuan Nguyen, MS, V. Zelman Center for Neurobiology and Brain Restoration Skolkovo Institute of Science and Technology;
ORCID: https://orcid.org/0000-0002-1703-0404

Olga Igorevna Efimova, MS, V. Zelman Center for Neurobiology and Brain Restoration Skolkovo Institute of Science and Technology;
ORCID: https://orcid.org/0000-0003-0842-3203

Artem Valerievich Tokarchuk, MS, V. Zelman Center for Neurobiology and Brain Restoration Skolkovo Institute of Science and Technology;
ORCID: https://orcid.org/0000-0002-9234-5828

Anna Yurievna Morozova, Senior Researcher, Department of Basic and Applied Neurobiology, V. Serbsky National Medical Research Centre of Psychiatry and Narcology of the Ministry of Health of the Russian Federation; The officer-in-charge, Mental-health Clinic No. 1 named after N.A. Alexeev; ORCID: https://orcid.org/0000-0002-8681-5299

Yana Alexandrovna Zorkina, Cand. Sci. (Biology), Senior Researcher, V. Serbsky National Medical Research Centre of Psychiatry and Narcology of the Ministry of Health of the Russian Federation;
Mental-health Clinic No. 1 named after N.A. Alexeev;
ORCID: https://orcid.org/0000-0003-0247-2717;
e-Library SPIN-code: 3017-3328, Researcher ID: H-2424-2013

Denis Sergeevich Andreyuk, Cand. Sci. (Biology), Senior Fellow at the Education Center, Mental-health Clinic No. 1 named after N.A. Alexeev;
ORCID: https://orcid.org/0000-0002-3349-539

George Petrovich Kostyuk, Dr. Sci (Med.), Professor, Director of Mental-health Clinic No. 1 named after N.A. Alexeev, Editor-in-Chief, Consortium Psychiatricum; ORCID: https://orcid.org/0000-0002-3073-6305; Scopus Author ID: 57200081884, RSCI: 3424-4544

***Philipp Efimovich Khaitovich,** Ph.D. , Professor, V. Zelman Center for Neurobiology and Brain Restoration Skolkovo Institute of Science and Technology; ORCID: https://orcid.org/0000-0002-4305-0054
E-mail: p.khaitovich@skoltech.ru

*corresponding author

References

1. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390(10100):1211-59. doi: 10.1016/S0140-6736(17)32154-2.
2. Khavari B, Cairns MJ. Epigenomic dysregulation in schizophrenia: in search of disease etiology and biomarkers. Cells. 2020;9(8):1837. doi: 10.3390/cells9081837.
3. Van Erp TG, Walton E, Hibar DP, et al. Cortical brain abnormalities in 4474 individuals with schizophrenia and 5098 control subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. Biol Psychiatry. 2018;84(9):644-54. doi: 10.1016/j.biopsych.2018.04.023.
4. Vita A, de Peri L, Deste G, Sacchetti E. Progressive loss of cortical gray matter in schizophrenia: a meta-analysis and meta-regression of longitudinal MRI studies. Translational Psychiatry. 2012;2(11):e190. doi: 10.1038/tp.2012.116.
5. Smucny J, Dienel SJ, Lewis DA, Carter CS. Mechanisms underlying dorsolateral prefrontal cortex contributions to cognitive dysfunction in schizophrenia. Neuropsychopharmacology. 2022;47(1):292-308. doi: 10.1038/s41386-021-01089-0.
6. Bobilev AM, Perez JM, Tamminga CA. Molecular alterations in the medial temporal lobe in schizophrenia. Schizophr Res. 2020;217:71-85. doi: 10.1016/j.schres.2019.06.001.
7. Van Erp TG, Hibar DP, Rasmussen JM, et al. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. Mol Psychiatry. 2016;21(4):547-53. doi: 10.1038/mp.2015.63.
8. Merikangas AK, Shelly M, Knighton A, et al. What genes are differentially expressed in individuals with schizophrenia? A systematic review. Molecular Psychiatry. 2022;27(3):1373-83. doi: 10.1038/s41380-021-01420-7.
9. Yildiz M, Borgwardt SJ, Berger GE. Parietal lobes in schizophrenia: do they matter? Schizophr Res Treatment. 2011:581686. doi: 10.1155/2011/581686.
10. Vidal-Domènech F, Riquelme G, Pinacho R, et al. Calcium-binding proteins are altered in the cerebellum in schizophrenia. PLoS One. 2020;15(7):e0230400. doi: 10.1371/journal.pone.0230400.
11. Mudge J, Miller NA, Khrebtukova I, et al. Genomic convergence analysis of schizoprenia: mRNA sequencing reveals altered synaptic vesicular transport in post-mortem cerebellum. PLoS One. 2008;3(11):e3625. doi: 10.1371/journal.pone.0003625.
12. Aliperti V, Skonieczna J, Cerase A. Long non-coding RNA (lncRNA) roles in cell biology, neurodevelopment and neurological disorders. Noncoding RNA. 2021;7(2):36. doi: 10.3390/ncrna7020036.
13. Gibbons A, Udawela M, Dean B. Non-Coding RNA as novel players in the pathophysiology of schizophrenia. Noncoding RNA. 2018;4(2):11. doi: 10.3390/ncrna4020011.
14. Salvatori B, Biscarini S, Morlando M. Non-coding RNAs in nervous system development and disease. Front Cell Dev Biol. 2020;8:273. doi: 10.3389/fcell.2020.00273.
15. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol. 2021;22(2):96-118. doi: 10.1038/s41580-020-00315-9.
16. Rusconi F, Battaglioli E, Venturin M. Psychiatric disorders and lncRNAs: a synaptic match. Int J Mol Sci. 2020;21(9):3030. doi: 10.3390/ijms21093030.
17. Melé M, Mattioli K, Mallard W, et al. Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. Genome Res. 2017;27 (1):27-37. doi: 10.1101/gr.214205.116.
18. Ransohoff JD, Wei Y, Khavari PA. The functions and unique features of long intergenic non-coding RNA. Nat Rev Mol Cell Biol. 2018;19(3):143-57. doi: 10.1038/nrm.2017.104.

19. Quick-Start Protocols — QIAzol Lysis Reagent 2011 [Internet]. [cited 2022 Oct 5]. Available from: https://www.qiagen.com/us/resources/download.aspx?id=6c452080-142a-44a7-a902-9177dea57d7c&lang=en.
20. TruSeq RNA Sample Prep Guide (15008136 A) [Internet]. 2010. [cited 2022 Oct 5]. Available from: https://support.illumina.com/downloads/truseq_rna_sample_preparation_guide_15008136.html.
21. Andrews S. FastQC: A quality control tool for high throughput sequence data [Internet]. 2010. [cited 2022 Oct 5]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 5 Oct 2022.
22. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20. doi: 10.1093/bioinformatics/btu170.
23. Kim D, Paggi JM, et al. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907-15. doi: 10.1038/s41587-019-0201-4.
24. Pertea M, Pertea GM, Antonescu CM, et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290-5. doi: 10.1038/nbt.3122.
25. Cunningham F, Allen JE, Allen J, et al. Ensembl 2022. Nucleic Acids Res. 2022;50(D1):D988-95. doi: 10.1093/nar/gkab1049.
26. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8.
27. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284-7. doi: 10.1089/omi.2011.0118.
28. Khrameeva E, Kurochkin I, Han D, et al. Single-cell-resolution transcriptome map of human, chimpanzee, bonobo, and macaque brains. Genome Res. 2020;30(5):776-89. doi: 10.1101/gr.256958.119.
29. Lake BB, Chen S, Sos BC, et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. Nat Biotechnol. 2018;36(1):70-80. doi: 10.1038/nbt.4038.
30. Darmanis S, Sloan SA, Zhang Y, et al. A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci U S A. 2015;112(23):7285-90. doi: 10.1073/pnas.1507125112.
31. Zhang Y, Sloan SA, Clarke LE, et al. Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. Neuron. 2016;89(1):37-53. doi: 10.1016/j.neuron.2015.11.013.
32. Liu Y, Chang X, Hahn CG, et al. Noncoding RNA dysregulation in the amygdala region of schizophrenia patients contributes to the pathogenesis of the disease. Transl Psychiatry. 2018;8(1):44. doi: 10.1038/s41398-017-0030-5.
33. Tian T, Wei Z, Chang X, et al. The long noncoding RNA landscape in amygdala tissues from schizophrenia patients. EBioMedicine. 2018;34:171-81. doi: 10.1016/j.ebiom.2018.07.022.
34. Hauberg ME, Fullard JF, Zhu L, et al.; CommonMind Consortium. Differential activity of transcribed enhancers in the prefrontal cortex of 537 cases with schizophrenia and controls. Mol Psychiatry. 2019;24(11):1685-95. doi: 10.1038/s41380-018-0059-8.
35. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27. doi: 10.1093/nar/28.1.27.
36. Gillespie M, Jassal B, Stephan R, et al. The reactome pathway knowledgebase 2022. Nucleic Acids Res. 2022;50(D1):D687-92. doi: 10.1093/nar/gkab1028.
37. Huang G, Osorio D, Guan J, et al. Overdispersed gene expression in schizophrenia. NPJ Schizophr. 2020;6(1):9. doi: 10.1038/s41537-020-0097-5.
38. Azevedo FA, Carvalho LR, Grinberg LT, et al. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J Comp Neurol. 2009;513(5):532-41. doi: 10.1002/cne.21974.
39. Najjar S, Pearlman DM. Neuroinflammation and white matter pathology in schizophrenia: systematic review. Schizophrenia Research. 2015;161(1):102-12. doi: 10.1016/j.schres.2014.04.041.
40. Sánchez Y, Huarte M. Long non-coding RNAs: challenges for diagnosis and therapies. Nucleic Acid Ther. 2013;23(1):15-20. doi: 10.1089/nat.2012.0414.
41. Hwang Y, Kim J, Shin JY, et al. Gene expression profiling by mRNA sequencing reveals increased expression of immune/inflammation-related genes in the hippocampus of individuals with schizophrenia. Transl Psychiatry. 2013;3(10):e321. doi: 10.1038/tp.2013.94.
42. Wu JQ, Wang X, Beveridge NJ, et al. Transcriptome sequencing revealed significant alteration of cortical promoter usage and splicing in schizophrenia. PLoS One. 2012;7(4):e36351. doi: 10.1371/journal.pone.0036351.
43. Chang X, Liu Y, Hahn CG, et al. RNA-seq analysis of amygdala tissue reveals characteristic expression profiles in schizophrenia. Transl Psychiatry. 2017;7(8):e1203. doi: 10.1038/tp.2017.154.
44. Barch DM, Bustillo J, Gaebel W, et al. Logic and justification for dimensional assessment of symptoms and related clinical phenomena in psychosis: relevance to DSM-5. Schizophr Res. 2013;150(1):15-20. doi: 10.1016/j.schres.2013.04.027.