

Post-transcriptional dysregulation in autism, schizophrenia and bipolar disorder

Yuanyuan Wang¹, Liya Liu², Mingyan Lin^{2*}

¹State Key Laboratory of Reproductive Medicine, Nanjing Medical University, Nanjing, Jiangsu 211166, China;

²Department of Neurobiology, School of Basic Medical Sciences, Nanjing Medical University, Nanjing, Jiangsu 211166, China

* Corresponding author:

Mingyan Lin, Department of Neurobiology, School of Basic Medical Sciences, Nanjing Medical University, 101 Longmian Avenue, Nanjing, Jiangsu 211166, China.

Email: linmingyan@njmu.edu.cn

Abstract

Post-transcriptional gene regulation (PTGR) contributes to numerous aspects of RNA metabolism. While multiple regulators of PTGR have been associated with the occurrence and development of psychiatric disorders, a systematic investigation of the role of PTGR in the context of neuropsychiatric disorders is still lacking. In this work, we developed a new transcriptome -based algorithm to estimate PTGR and applied it to an RNA-Seq dataset of 2160 brain samples from individuals with autism spectrum disorder (ASD), schizophrenia (SCZ), bipolar disorder (BD) and controls. The results showed that the contribution of PTGR abnormality to gene differential expression between three common psychiatric disorders and controls was about 30% of that of transcriptional gene regulation (TGR) abnormality. Besides, aberrant PTGR tended to decrease RNA stability in SCZ/BD, while increase RNA stability in ASD, implicating contrasting pathologies among diseases. The abnormal alteration of PTGR in SCZ/BD converged on the inhibition of neurogenesis and neural differentiation, whereas dysregulation of PTGR in ASD induced enhanced activity of apoptosis. This suggested that heterogeneity in disease mechanism and clinical manifestation across different psychiatric disorders may be partially attributed to the diverse role of PTGR. Intriguingly, we identified a promising RBP (RNA bind protein) ELAVL3 (ELAV-Like Protein 3) that have a profound role in all three psychiatric disorders. Our systematic study expands the understanding of the link between PTGR and psychiatric disorders and also open a new avenue for deciphering the pathogenesis of psychiatric disorders.

Keywords: post-transcriptional gene regulation (PTGR), psychiatric disorders, RNA stability, RNA binding protein (RBP), ELAVL3

Introduction

Psychiatric disorders, characterized by brain dysfunction, are caused by a combination of psychological, biological, and environmental factors, including schizophrenia (SCZ), bipolar disorder (BD), depression (Depression) and autism (ASD), and so on. Currently, it is generally postulated that psychiatric disorders are polygenic genetic disease with high genetic heterogeneity and complexity, the pathological mechanisms of psychiatric disorders remain poorly understood¹. In the present study, several studies have revealed that the majority of the psychiatric associated risk loci lies in the gene regulatory region (non-coding region of gene) rather than loci within the gene coding region²⁻⁴, indicating that a greater role for dysregulation of gene expression and alternative splicing compared to dysfunction of the gene in contributing to the development of psychiatric disorders. Gene expression and alternative splicing are jointly determined by the transcriptional regulation mechanism (TGR) and the post-transcriptional regulation mechanism (PTGR). Recognizing the importance of understanding gene expression and splicing, several researches have undertaken large-scale efforts to study the impact of TGR abnormalities on psychiatric disorders^{5,6}, and most of these studies focused on dysregulation of transcription factors and epigenetic patterns^{7,8}. These investigations reported several susceptibility factors including the SCZ high risk *ZNF804A*⁹, 22q11.2 microdeletion¹⁰, ASD high risk *CHD8*¹¹, as well as the fever in maternal immune activation¹². Meanwhile, PTGR, especially RBPs-mediated PTGR, as a major contributor in the control of gene expression and splicing regulation, and participates in the regulation of RNA metabolism^{13,14}. Nevertheless, there are few studies on the PTGR estimation algorithm, leaving large parts of the PTGR, especially RBPs-mediated PTGR, unexplored¹⁵. A more comprehensive analysis

focusing on the PTGR and psychiatric disorders is necessary to accelerate our understanding of the role of RBPs-mediated PTGR in the pathogenesis of psychiatric disorders.

Several studies based on RNA-seq provided evidence that the reads in intronic regions can represent the expression level of pre-mRNAs, which are mainly transcriptionally driven¹⁶⁻¹⁸. The reads in the exonic regions are correlated with the expression level of mature mRNAs (mature mRNAs), which are related to transcriptional in combination with post-transcriptional regulation events¹⁶⁻¹⁸. Therefore, the changes of the expression level of pre-mRNAs and mature mRNAs across different conditions allow us to evaluate the contributions of post-transcriptional regulation to observed changes in steady-state RNA levels and identify abnormal changes of gene post-transcriptional regulatory mechanism. Thus, these analytic methods have defined the difference of the logarithm of fold-change of exonic reads and the logarithm of fold-change of intronic reads ($\Delta_{\text{exon}} - \Delta_{\text{intron}}$) as altered post-transcriptional regulation of mRNAs¹⁶⁻¹⁸. It is known that Δ_{exon} and Δ_{intron} are also affected by the RNA metabolic limitations, measurement errors, and other factors, and the presence of these biases leads to Δ_{intron} much larger than Δ_{exon} . However, these methods do not take into account these biases, which results in the overestimation of Δ_{PTGR} .

To provide a comprehensive view of the abnormal PTGR architecture across three common psychiatric disorders (ASD, SCZ, and BD) at the whole-genome level and measure the contributions of RBPs-mediated PTGR for disease-specific changes in gene expression, we developed a computational method systematically analyzed the PTGR changes of 2160 brain

transcriptome samples data from the PsychENCODE Consortium¹⁹, including 80 ASD samples, 594 SCZ samples, and 253 BD samples, as well as control. In summary, our analyses offer an updated method for PTGR estimation, and highlight that abnormal PTGR as a potential mechanism conferring key aspects of psychiatric disorders specificity and complexity, and find a novel direction for further research into the etiology of psychiatric disorders.

Materials and Methods

Data collection

For this study, all processed bulk RNA-seq alignment (.bam) files were collected from the PsychENCODE consortium. All bam files were aligned to the hg19 reference genome and available at <https://www.synapse.org/#!/Synapse:syn6039873>. These include normal controls (n=1233), as well as schizophrenia (SCZ) (n=594), bipolar disorder (BD) (n=253), autism spectrum disorder (ASD) (n=80).

Annotation of exon and intron reads

For reads annotation, we made a custom intron GTF annotation file and a custom composite exon GTF annotation file based on gencode v19 GTF. The following set of rules was applied to generate two custom GTF files supporting a given genomic region as an exonic region or intronic region: intron coordinates of genes were extracted from GENCODE v19 as custom intronic GTF; exon and untranslated region coordinates within any isoform of genes were extracted from GENCODE v19 as custom composite exonic GTF. Gene-level read counts were calculated for intron and exon separately using featureCounts²⁰.

Covariate selection

Exon/intron counts were compiled from gene level imported into R for downstream analyses. Genes were quantified in TPM and further filtered to include those with TPM > 1 in at least 50% of samples.

Bulk RNA-Seq samples clinical data are available at

<https://www.synapse.org/#!/Synapse:syn4587614>. Missing values in clinical data were imputed using the *missMDA* R package²¹. We summarized the top 40 principal components which collectively explained 99% of the total variance. To determine which covariates to include in the final differential post-transcriptional regulation model, we performed multivariate adaptive regression as implemented in the *earth* package in R for intronic and exonic separately. The superset of potential covariates available for all samples included: study, tissue, libraryPrep, strand specificity, platform, individual ID Source, diagnosis, sex, ethnicity, PMI, RIN, ageDeath, along with all 40 seqPCs. For continuous variables, we also included squared terms. These covariates with input into the *earth* model along with count data (limma voom normalized, centered, and scaled). The model was run using linear predictors and otherwise default parameters. As the model fits a maximum of 1000 genes simultaneously, we performed 1000 permutations randomly subsetting 1000 genes at a time. From this, we chose as a set of known covariates those present in a total of the exons and introns resulting pruned models, which consisted of: study, diagnosis, sex, ethnicity, PMI, ageDeath, PC (1-3), PC (5-11), PC (13-18), PC20, PC22, PC23, PC (25-27), PC29, PC31, PC (33-36), PC39, RINS2, ageDeathS2, PC1S2, PC2S2, PC4S2, PC6S2, PC9S2, PC16S2, PC17S2, PC22S2.

Exonic/Intronic count correction

Count level quantifications were normalized for library size using TMM normalization in *edgeR* and were transformed as $\log_2(\text{CPM}+0.5)$. Normalized and transformed exon/intron counts were then calculated using a linear mixed-effects model using the *nlme* package in R. The covariates specified in the

previous section were included as fixed effects in the model. In addition, we included a random effect term for each unique subject to account for subject overlap across sequencing studies. All covariates except for diagnosis and subject were regressed from our exonic/intronic count dataset.

New PTGR estimation algorithm

The intronic counts representing the relative abundance of pre-mRNA, and the exonic counts representing the relative abundance of mature-mRNA. We took the median of all samples within covariates corrected exonic/intronic counts dataset as the reference sample S_{ref} , and then for each gene in each sample S , we subtracted the median value across the samples to obtain Δ_{exon} , Δ_{intron} . Different condition exonic and intronic groups were analyzed separately.

For each of control, SCZ, ASD, and BD set separately, we modeled the ratio of Δ_{exon} and Δ_{intron} for each gene by robust linear regression of Δ_{exon} vs. Δ_{intron} . The control group ratio is recorded as slope A. We used the same method to estimate the ratio of Δ_{exon} and Δ_{intron} in the disease group. The disease group ratio is recorded as slope B. If $B > A$, which means that PTGR is up-regulated in the B(disease) group ($\Delta_{PTGR} > 0$), and representing the increased RNA stability in the disease group. Positive and negative Δ_{PTGR} values correspond to increased RNA stability and reduced RNA stability, respectively. Statistical significance of Δ_{PTGR} was determined using *t-test* shown below:

$$t = \frac{slope(Disease) - slope(Ctrl)}{\sqrt{SE(Disease)^2 + SE(Ctrl)^2}} \quad df = \text{size (Disease)} + \text{size (Ctrl)}$$

Here, *Slope* represents the robust linear regression slope value, and *SE* represents standard error of *Slope*. The *df* is the sum of samples in the control and disease groups.

For each gene, the effect of PTGR vs TGR (transcription gene regulation) on gene expression can be estimated according to the following formula:

$$ratio = \left| \frac{slope(Disease)}{slope(Ctrl)} - 1 \right|$$

The average Δ PTGR of each gene was obtained from the above analysis. To measure the Δ PTGR (Δ PTGR_{in}) of each gene in each individual of the disease, we used the following formula:

$$\Delta PTGR_{in} = \Delta exon(Disease) - \Delta intron(Disease) * slope(Ctrl)$$

Post-transcriptional perturbation dataset analysis

We used RNA-seq data of schizophrenia patient-derived neural progenitor cells with purely post-transcriptional perturbations mediated by microRNAs from ref²². (GEO accession GSE80170) to evaluate the performance of the PTGR algorithm. The data include control samples, samples with the miR-9 knockdown, and samples with the miR-9 overexpression. Reads were mapped to the hg19 assembly of human genome using STAR²³ with default parameters. We used custom intronic GTF and custom composite exonic GTF for reads annotation, and gene-level read counts were calculated for intronic and exonic separately using featureCounts. We identified genes that were inferred to be up-PTGR or down-PTGR in the

knockdown/overexpression group. Because direct targets of miRNAs are enriched in miRNA seed matches in their 3'-UTR, we examined whether genes that were up-PTGR or down-PTGR were enriched for miR-9 RNA seed sites in their 3'UTRs to evaluate the performance of PTGR algorithm.

REMBRA PTGR effect deconvolution algorithm

We closely followed the REMBRA pipeline for PTGR effect deconvolution. We did this to optimally enable comparisons between our PTGR algorithm and REMBRA algorithm. These steps are implemented in a software package available at <https://github.com/csglab/REMBRANDTS>¹⁷.

Integrated PsychENCODE transcription regulatory results

Summary differential expression genes results from the PsychENCODE were obtained from http://resource.psychencode.org/DER-13_Disorder_DEX_Genes.csv.

Identify key RBPs

We predicted the interaction between DPRGs and 195 human RBPs through the combination of RBPs and phantoms, and then inferred the RBPs that plays a key role through linear models. We limited the analysis of RBP and miRNA binding sites to genes for which all isoforms had the same 3' UTR start coordinates, and the 3' UTR was composed of a single exon, in order to minimize the possibility of confounding effects of alternative splicing. The 3' UTR start coordinates were extracted from GENCODE v19. For analysis of RBP binding sites, we first collected a non-redundant compendium of available sequence preferences for human RBPs. We obtained 217 position

frequency matrices (PFMs) representing 117 human RBPs from beRBP²⁴, including PFMs with direct experimental evidence and those inferred by homology. Then, we calculated all pairwise PFM similarities using MoSBAT²⁵, and then used affinity propagation to cluster the PFMs based on similarity, keeping only the “exemplar” from each cluster. This reduced the total number of PFMs to 154, which we call the non-redundant RBP motif set. Then, we scanned the 3' UTR sequences with each PFM using AffiMx from the MoSBAT package, resulting in a vector of PFM scores that represents the affinity of the corresponding RBP for binding to different 3' UTRs.

To identify RBPs that are associated with disease-specific stability profiles, we used the gene-level stability measures as the response variable in multiple linear regression, with RBP binding site match counts as predictor variables. RBPs and miRNAs whose binding sites were significantly associated with disease-specific stability were identified based on t-test of regression coefficients at *pvalue* < 0.05.

Developmental expression data

Developmental expression data from the BrainSpan: Atlas of the Developing Human Brain (<http://www.brainspan.org/>).

Functional enrichment analysis

We performed gene ontology (GO) enrichment analysis using the R package *clusterProfiler*²⁶ and GO interactome network was performed by Metascape²⁷.

Results

New PTGR evaluation algorithm

For an unbiased estimate of the post-transcriptional regulation contributions to gene expression changes, we devised a computational method and applied it to the standard RNA-seq experiment. The abundance of reads lies in the intron region in RNA-seq data correspond to pre-mature mRNA abundance, changes in the abundance of intronic reads can be used to estimate the change in transcription regulation. While the abundance of reads lies in the exon region in RNA-seq data correspond to mature mRNA abundance, changes in the abundance of exonic reads can be used to estimate the changes in transcription regulation and post-transcription regulation. (Methods and Fig. 1A). We first quantified the abundance of exonic and intronic reads in PsychENCODE RNA-seq datasets. As expected, we found that about 38% of reads were intronic reads, and 62% of reads were exonic reads (Fig. 1A). This observation is in agreement with previous studies^{16,18}, a substantial number of reads in bulk RNA-Seq originate from pre-mature mRNA. For each of control, SCZ, ASD, and BD set separately, we modeled the ratio of Δ_{exon} and Δ_{intron} for each gene by robust linear regression of Δ_{exon} vs Δ_{intron} (Methods and Fig. 1B). The control group ratio is recorded as slope A. We used the same method to estimate the ratio of Δ_{exon} and Δ_{intron} in the disease group. The disease group ratio is recorded as slope B. If $B > A$, which means that the PTGR is up-regulated in the B(disease) group ($\Delta_{\text{PTGR}} > 0$), and representing the increased RNA stability in the disease group. Positive and negative Δ_{PTGR} value corresponds to increased RNA stability and reduced RNA stability, respectively.

To evaluate the performance of our PTGR algorithm, we measured Δ PTGR using RNA-seq data of schizophrenia patient-derived neural progenitor cells with purely post-transcriptional perturbations mediated by microRNAs(*miR-9*) from ref²². The regulation of a particular miRNA on its target messages by binding to the miRNA seed region, and therefore the direct targets 3'UTR of miRNAs are enriched in the miRNA seed region. As expected, our results indeed showed genes that were downregulated at the PTGR level in the *miR-9* overexpression condition were significantly enriched for the *miR-9* direct targets, and genes that were upregulated at the PTGR level in the *miR-9* knockdown condition were significantly enriched for the *miR-9* direct targets (Fig. 1C). Our analysis highlighted that the PTGR algorithm is an accurate method for PTGR estimation.

To deeper assess the accuracy of the PTGR algorithm in big data and compared the accuracy relative to other methods, we next applied the PTGR algorithm to the 2160 brain transcriptome samples across three common psychiatric disorders data from the PsychENCODE Consortium. We used the previous algorithm (REMBRANDTS) and the PTGR algorithm to evaluate the influence of PTGR on changes of gene expression (see Methods section). PTGR algorithm revealed the average effect of abnormal PTGR on changes in gene expression levels was approximately 30% of that of abnormal TGR, but the average influence of PTGR estimated by REMBRANDTS on gene expression was 2 times that of TGR (Fig. 1D). The estimation of REMBRANDTS was contradictory to the theory that the majority of the changes in gene expression are driven by transcriptional regulation. Our results demonstrated that the PTGR algorithm should provide more accurate estimates of mRNA stability compared to REMBRANDTS.

Dysregulation of PTGR in psychiatric disorders

To generate a PTGR architecture across three diseases, we used the PTGR algorithm to assess differential PTGR in ASD, SCZ, and BD compared with control. Intriguingly, we noticed PTGR with the largest changes in ASD compared with SCZ/BD (Fig. 1D), and mRNAs appeared to be significantly more stable in ASD; the other two diseases mRNAs stability showed the reverse pattern (Fig. 2A). We identified widespread differential post-transcriptional regulation genes (DPRGs) in ASD, SCZ, and BD [n =1807, 1019, and 979 genes at false discovery rate (FDR) < 0.05, respectively] (Fig. 2B), indicating a prominent role of post-transcriptional programs in disease development. Besides, DPRGs were enriched for up-regulated PTGR genes in ASD, whereas DPRGs were enriched for down-regulated PTGR genes in SCZ as well as BD. Notably, at the Δ PTGR level, there was inconsiderable cross-disorder sharing of this DPRGs signal, suggesting that Δ PTGR confers a substantial portion of disease specificity. Comparison of Δ PTGR of whole genome genes across three diseases suggested there was a significant SCZ/BD cross-disorder correlation in PTGR changes (Fig. 2B), although the majority of DPRGs were disorder-specific in SCZ/BD. This finding implied that SCZ and BD share overlapping genetic etiology at the PTGR level, and ASD is likely to be with disease specificity at the PTGR level (Fig. 2C). To gain better insights into the contributions of abnormal PTGR in disease development, we subsequently performed Gene Ontology (GO) analysis of ASD-, SCZ-, and BD-associated DPRGs, respectively. Regulation of apoptotic process was enriched in ASD up-regulated DPRGs, implying the involvement of apoptotic in the development of ASD (Fig. 2D). SCZ-/BD up-regulated DPRGs showed significant enriched in immune activation related biological processes (Fig. 2D). Moreover, down-regulated DPRGs were associated with neural

development, neuron differentiation, and Synaptic Transmission in SCZ/BD, but no enrichment biological process in ASD (Fig. 2E). Together, this provides direct evidence that the disease specificity and complexity across three disorders partially attributed to the diversity of disease associated mRNA stability, suggesting clear divergence in the pathological mechanisms of psychiatric disorders.

Disease-specific key RBPs that modulate mRNA stability in psychiatric disorders

Since RBPs modulate mRNA stability via the Adenylate-uridylylate (AU)-rich elements (AREs) and regulate the expression of many genes at the post-transcriptional level²⁸, we examined whether DPRGs and differential expressed genes (DEGs) enriched in AREs. We indeed observed ARE enrichment in DPRGs, but not in DEGs (Fig. 2F). This result indicating that RBPs-regulated genes are enriched in DPRGs, and also provides further evidence to support the accuracy of the PTGR algorithm. Additionally, we performed enrichment analysis for transcription factors in DPRGs and DEGs, and observed transcription factors were significantly overrepresented in DPRGs (Fig. 2F), compared with DEGs, suggesting the abnormality of PTGR has a pervasive influence on the whole genome. Although the effect of PTGR on gene expression changes is less than that of TGR, PTGR may contribute as much to the changes of gene expression as TGR by affecting more upstream regulators. This work further highlights PTGR dysregulation as a critical, and relatively underexplored, mechanism linking abnormal gene expression with psychiatric disease pathophysiology.

The above observations prompted us to take a closer look at abnormal mRNAs stability of psychiatric disorders, but it is not clear which factors contribute the most to the abnormal landscape of mRNAs stability of psychiatric disorders. We combined the stability profiles of mRNAs of psychiatric disorders with the binding site predictions of the RBPs, and used multiple linear regression to identify key RBPs that modulate mRNAs stability in psychiatric disorders (see Methods for more details). This analysis identified multiple RBPs that were significantly predictive of disease mRNAs stability (Fig. 3A). Notable examples include FMR1²⁹ and QKI³⁰, which is known ASD risk gene, and SCZ risk gene^{29,31}, respectively. These promising candidate RBPs also including neurodevelopment associated ELAVL3³² and IGF2BP family³³, neuroimmune related gene ZPF36³⁴, classical splicing factor SRSF1³⁵, and so on. Besides, we observed ELAVL3 was upregulated in ASD, ELAVL3 targets were more likely to be bound and more stable in ASD, and ELAVL3 was downregulated in SCZ/BD, ELAVL3 targets were more likely to be bound and less stable in SCZ/BD (Fig. 3A). Therefore, the presence of 3' UTR binding sites for ELAVL3 was significantly associated with increased mRNAs stability in three disorders. In sum, the direction of the changes in stability was consistent with the direction verified experimentally in the literature³⁰.

Intriguingly, ELAVL3, a neural-specific RNA-binding protein which expression centered on early brain development (Figure 3B), displayed alterations of expression in all three psychiatric disorders. We subsequently analyzed 159 ELAVL3-bound genes expression level and found that the average expression of ELAVL3-bound genes also high in brain development (Figure 3E), demonstrating that the involvement of PTGR of ELAVL3 in early

brain development. we further examined the Δ PTGR of 159 ELAVL3-bound mRNAs and observed that these mRNAs were mainly increased stability in ASD and reduced stability in SCZ/BD (Fig. 3C). The GO analysis of the ELAVL3-bound mRNAs showed that these genes were significantly enriched in the biological processes of neurodevelopment, regulation of apoptotic process, neuron differentiation (Fig. 3D). Overall, our study implied the disruption of ELAVL3 may be one of the predominant risk factors that contribute to psychiatry-specific abnormal PTGR and provides a resource for mechanistic insight and therapeutic development.

Discussion

In this study, we present a new algorithm and integrated it into the transcriptome big data of the brain across three major psychiatric disorders to reveal the abnormal pattern and significance of PTGR for the first time.

The PTGR algorithm can be performed with any type of RNA-seq data set, for example, polyA RNA-Seq, single-cell RNA-Seq, whole RNA-Seq, to gain insights into the post-transcriptional regulatory mechanism responsible for the changes of gene expression. Notable, the accuracy of the PTGR estimate depends on the larger sample size, our method is more suitable for large sample data.

RNA stability profiles across three diseases displayed that the aberrant PTGR changes give a new characteristic of each disorder, and PTGR changes have expanded the understanding of the pathogenesis of psychiatric disorders. The similar pattern of RNA stability between SCZ and BD in disease revealed their convergence at the PTGR level, and the reverse pattern on the RNA stability of between ASD and SCZ/BD showed the divergence of PTGR across three disorders. According to the shared and specific PTGR abnormalities of three diseases, we believe that SCZ and BD share the pathogenic mechanism, and the pathogenic mechanism of ASD is specific more than SCZ and BD. The convergence and divergence of mRNA stability across three diseases provide the theoretical basis for investigations targeting shared and specific disease mechanisms to link causal drivers with psychiatric disorders.

Gene expression and alternative splicing events are determined by both TGR and PTGR. However, there are limited studies to accurately assess the contributions of TGR and PTGR on changes in gene expression in psychiatric disorders. Our results found the effect of PTGR abnormality on gene expression in three common psychiatric disorders was 30% of that of TGR abnormality, suggesting that TGR changes exhibit larger effect sizes in diseased brain gene expression than TGR changes. However, our results found that the upstream regulatory factors were enriched in the DPRG, suggesting that abnormal PTGR in psychiatric disorders leads to abnormal expression of the upstream regulatory factors, which indirectly has a significant impact on downstream gene expression. Therefore, we hypothesized that PTGR and TGR might contribute equally to changes in gene expression of psychiatric disorders, and this finding brought the significance of the PTGR in psychiatric disorders to a higher level.

The current understanding of the relationship between psychiatric disorders and PTGR mainly comes from the research of microRNAs (miRNAs), which regulate the degradation and translation of hundreds of mRNAs by binding to the 3' non-coding region of mRNAs. Numerous well-characterized miRNAs have been associated with psychiatric disorders, including *miR-132*, *miR-195*, *miR-18a*, *miR-137*, and so on. Although the influence of miRNAs and RBPs on PTGR has not been determined, existing studies have shown that RBPs are the all-around regulator of mRNA metabolism, and these findings suggest that RBPs play a non-negligible role in PTGR. Moreover, compared with transcription factors, RBPs regulate the end products of expression more directly, making RBPs a "new favorite" for drug intervention targets. The human genome encodes more than 800 RBPs

involved in the PTGR of mRNAs, yet only a small fraction has been functionally characterized. Encouragingly, our results identified many promising potential psychiatric susceptibility RBPs at the post-transcriptional regulation level. *Ince-Dunn et al.* performed a genome-wide analysis of *nElavl* targets and revealed that one function of *nElavl* is to control excitation-inhibition balance in the brain³². Our results indicated ELAVL3 is a candidate RBP that plays a critical role in three psychiatric disorders, its activity increased in ASD, and in SCZ/BD activity decreased, resulting in cell apoptosis, neural development, and neuron differentiation of related genes mRNA stability aberrance. These observed suggesting that the PTGR mediated by ELAVL3 may be related to the excitation-inhibition imbalance and play an important role in the development of psychiatric disorders. Our results can help clarify the regulatory mechanism of RBP-mediated PTGR.

In summary, we applied a new PTGR estimation method to RNA sequencing data of 2160 brain samples from ASD, SCZ, and BD, as well as controls individuals. We found that SCZ and BD shared a similar abnormal pattern of the PTGR, whereas ASD has a specific abnormality of PTGR, revealing the regularity of PTGR abnormality in three psychiatric disorders. Besides, our results identified RBPs that play a critical role in psychiatric disorders, suggesting that the abnormalities of RBP-mediated PTGR may contribute to the development of psychiatric disorders. These findings bring new strategies for mechanistic insight and disease therapeutic development.

Figures

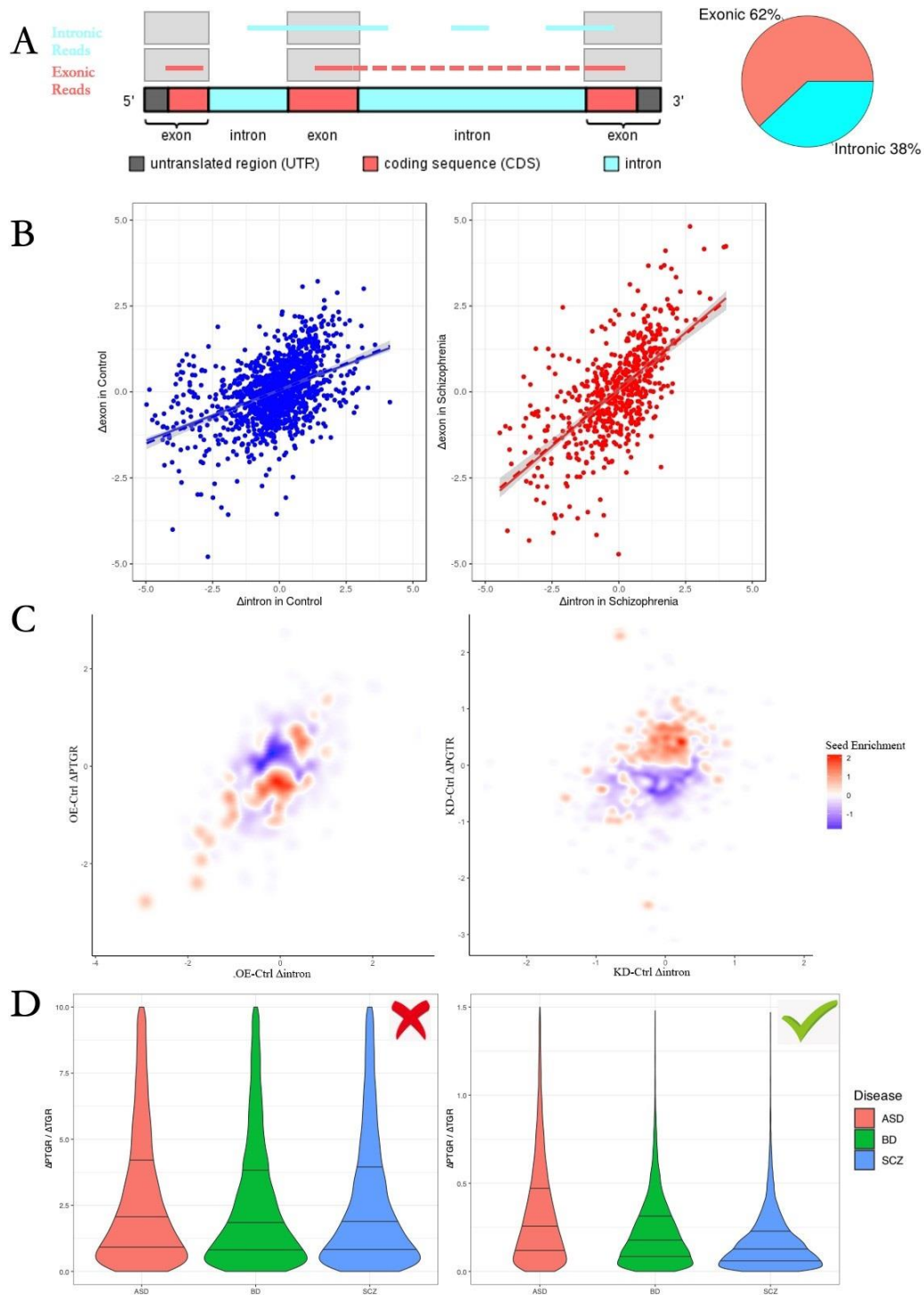


Figure 1. New algorithm for PTGR estimation.

(A) Theoretical basis schematic diagram of PTGR estimation algorithm (left). The fraction of exonic reads and intronic reads in PsychENCODE bulk RNA-Seq dataset is shown in the pie chart. (right). The intronic reads represent the expression of pre-mRNAs, and the exonic reads represent the expression of mature- mRNAs. **(B)** Schematic diagrams of PTGR estimation method. Scatter plots comparing PTGR of disease (left) and control(right). Y axis represents the changes of exonic reads. X axis represents the changes of intronic reads. Each dot represents a sample, and the solid line represents fitted line. **(C)** Density plot for miR-9 seed sequence enrichments in genes of miR-9 knockdown (left) and miR-9 overexpression (right). Y axis represents the changes of exonic reads. X axis represents the changes of intronic reads. Each dot represents a gene, and the color represents the degree of enrichment of genes with miR-9 seed sequence in the 3'UTR, with red representing enrichment and blue representing depletion. **(D)** Comparison of REMBRANDTS (left) and new PTGR algorithm (right). The violin plot shows the ratio of change of PTGR in disease vs. control to change of TGR in disease vs. control. The three lines within the violin plot represent the 75%, median, and 25% percentiles from up to bottom. ASD, BD and SCZ are respectively displayed in red, green and blue.

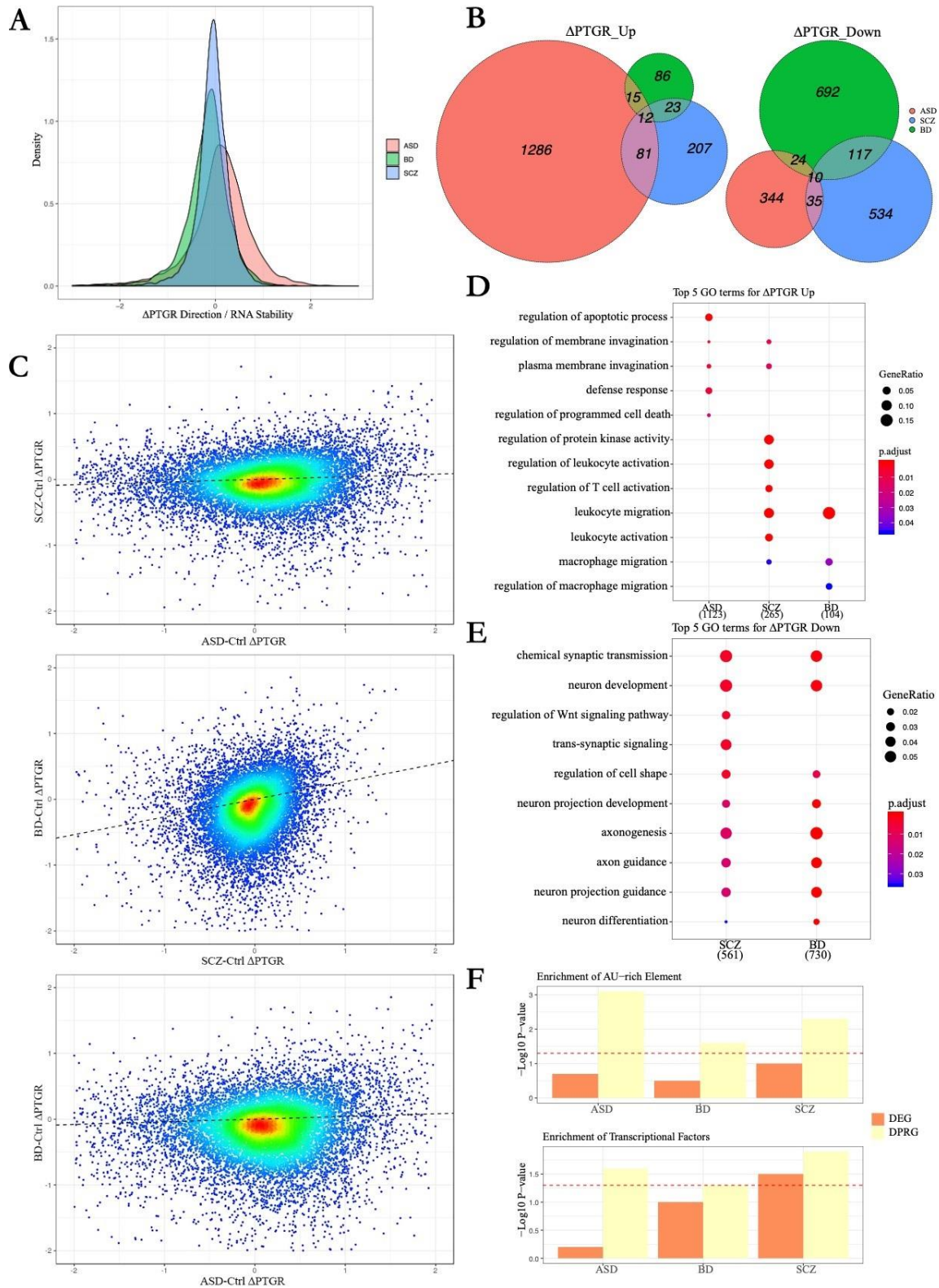


Figure 2. Dysregulation of PTGR in psychiatric disorders.

(A) Density plot of Δ PTGR distribution within the whole genome genes in three diseases. The Δ PTGR > 0 represents the increase of RNA stability (disease vs. control), whereas Δ PTGR < 0 represents the decrease of RNA stability (disease vs. control). (B) Venn diagram showing cross-disorder overlap for genes with significantly differential PTGR (FDR < 0.05). ASD, BD and SCZ are displayed in red, green and blue, respectively. (C) The scatter plot shows the degree of correlation of Δ PTGR between diseases. Each dot represents a gene. The color represents the localized density of genes. Red represents high gene density, and the blue represents low gene density. The dotted line represents fitted line. (D) Gene ontology enrichments are shown for up-regulated gene of PTGR in each disorder. Numbers in parentheses represent the count of genes in each disorder. (E) Gene ontology enrichments are shown for down regulated gene of PTGR in each disorder. Numbers in parentheses represent the count of genes in each disorder. (F) The bar plot shows the enrichment degree of regulatory elements AREs(up) and transcription factors(bottom) in DPRGs and DEGs in three diseases.

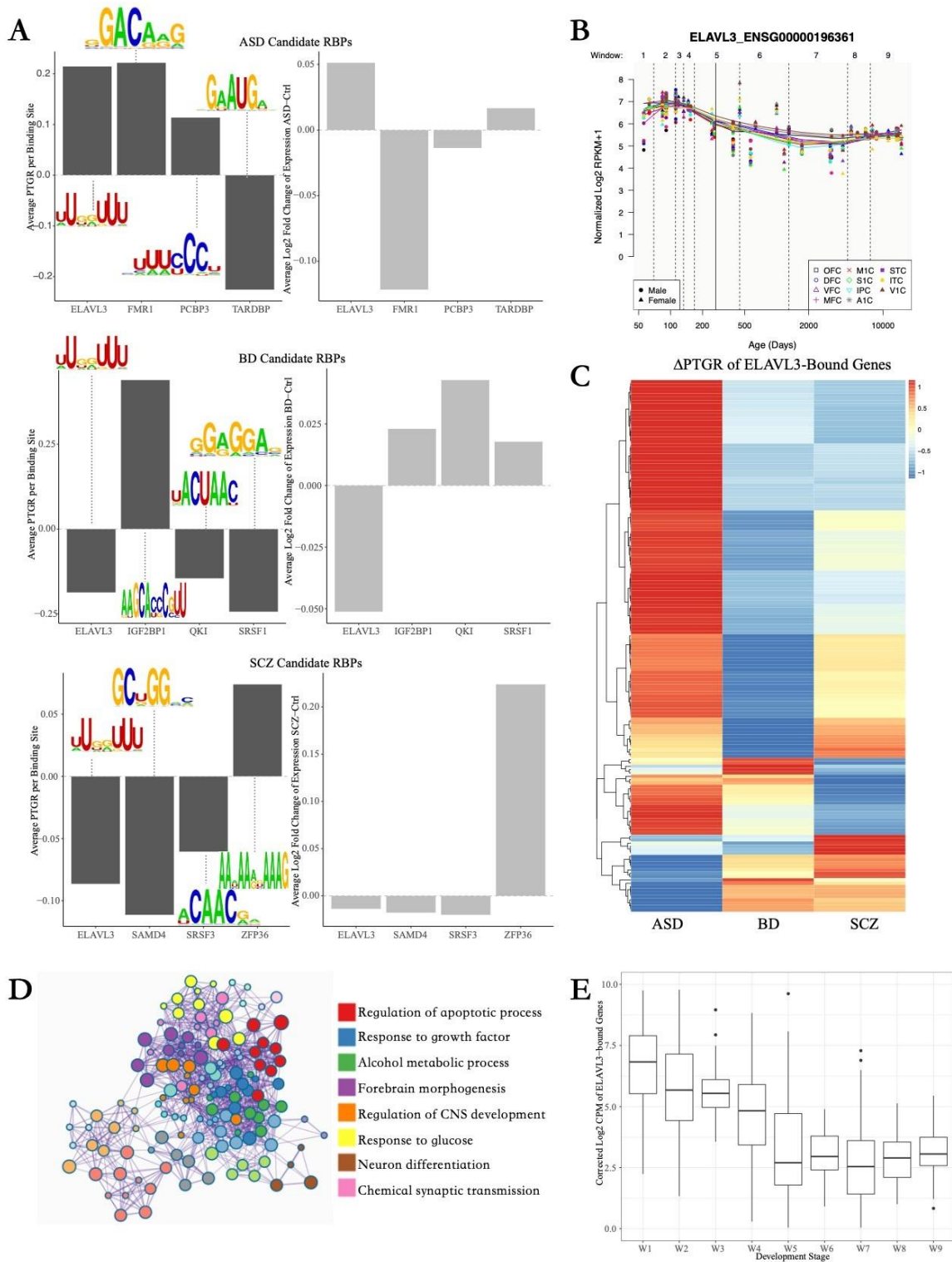


Figure 3. Disease-specific top candidate RBPs that modulate RNA stability in psychiatric disorders.

(A) RBPs that are significantly associated with disease-specific mRNA stability are shown (FDR < 0.05, *t-test* of regression coefficients). From up to bottom, there are identified RBPs in ASD, BD and SCZ. The left bar plot shows the predicted key RBPs, the corresponding RBP binding motif. Y axis within the left bar plot represents the average PTGR of predicted RBPs-bound genes in disease vs. control. The right bar chart shows the average log₂fold-change of the expression of RBP in disease vs. control. **(B)** Plot of spatiotemporal expression of ELAVL3 in human brain. **(C)** Heatmap shows the relative RNA stability of 159 predicted ELAVL3-bound genes in three diseases. The Δ PTGR represents the changes of PTGR of 159 predicted ELAVL3-bound genes in disease vs. control. **(D)** Enrichment analysis of ELAVL3-bound genes. Each color represents a kind of biological process category. **(E)** The box plot shows the average expression of 159 ELAVL3-bound genes at different developmental stages. The stages windows of W1-9 correspond to the developmental time windows in Figure 3 (B).

Author contributions

The study was designed by M.L.. M.L. and L.L. collected data. M.L., L.L. and Y.W. did the bioinformatic analyses. Y.W., M.L., and L.L. wrote the manuscript.

Acknowledgements

We are grateful to all laboratory members for comments on the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

Funding

This study was supported by research grants from the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (17KJB180009) to M.L., the Natural Science Foundation of Jiangsu Province (BK20171062) to M.L., and the National Natural Science Foundation of China (81701320) to M.L..

References

1. Horwitz T, Lam K, Chen Y, Xia Y, Liu C. A decade in psychiatric GWAS research. *Mol Psychiatry* 2019; **24**(3): 378-89.
2. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**(7510): 421-7.
3. Robinson EB, St Pourcain B, Anttila V, et al. Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nature Genetics* 2016; **48**(5): 552-5.
4. Stahl EA-O, Breen GA-O, Forstner AJ, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. (1546-1718 (Electronic)).
5. Petronis A, Labrie V. The crossroads of psychiatric epigenomics. *World Psychiatry* 2019; **18**(3): 353-4.
6. Huo Y, Li S, Liu J, Li X, Luo X-J. Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nature Communications* 2019; **10**.
7. Willsey AJ, Sanders Stephan J, Li M, et al. Coexpression Networks Implicate Human Midfetal Deep Cortical Projection Neurons in the Pathogenesis of Autism. *Cell* 2013; **155**(5): 997-1007.
8. Chen CA-O, Meng QA-OX, Xia YA-O, et al. The transcription factor POU3F2 regulates a gene coexpression network in brain tissue from patients with psychiatric disorders. LID - 10.1126/scitranslmed.aat8178 [doi] LID - eaat8178. (1946-6242 (Electronic)).
9. Chen J, Lin M, Hrabovsky A, et al. ZNF804A Transcriptional Networks in Differentiating Neurons Derived from Induced Pluripotent Stem Cells of Human Origin. *PLoS One* 2015; **10**(4): e0124597-e.
10. Lin M, Pedrosa E, Hrabovsky A, et al. Integrative transcriptome network analysis of iPSC-derived neurons from schizophrenia and schizoaffective disorder patients with 22q11.2 deletion. (1752-0509 (Electronic)).
11. Wang P, Mokhtari R, Pedrosa E, et al. CRISPR/Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in cerebral organoids derived from iPS cells. (2040-2392 (Electronic)).
12. Lin M, Zhao D, Hrabovsky A, Pedrosa E, Zheng D, Lachman HM. Heat shock alters the expression of schizophrenia and autism candidate genes in an induced pluripotent stem cell model of the human telencephalon. (1932-6203 (Electronic)).
13. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding

- proteins. (1471-0064 (Electronic)).
14. Keene JD. RNA regulons: coordination of post-transcriptional events. (1471-0056 (Print)).
 15. Zhou Y, Dong F, Mao Y. Control of CNS functions by RNA-binding proteins in neurological diseases. *Curr Pharmacol Rep* 2018; **4**(4): 301-13.
 16. Gaidatzis D, Burger L, Florescu M, Stadler MA-O. Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. (1546-1696 (Electronic)).
 17. Alkallas R, Fish L, Goodarzi H, Najafabadi HS. Inference of RNA decay rate from transcriptional profiling highlights the regulatory programs of Alzheimer's disease. *Nat Commun* 2017; **8**(1): 909.
 18. La Manno G, Soldatov R, Zeisel A, et al. RNA velocity of single cells. (1476-4687 (Electronic)).
 19. Akbarian S, Liu C, Knowles JA, et al. The PsychENCODE project. (1546-1726 (Electronic)).
 20. Liao Y, Smyth Gk Fau - Shi W, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. (1367-4811 (Electronic)).
 21. Josse J, Husson F. missMDA : A Package for Handling Missing Values in Multivariate Data Analysis. *Journal of Statistical Software* 2016; **70**.
 22. Topol A, Zhu S, Hartley BJ, et al. Dysregulation of miRNA-9 in a Subset of Schizophrenia Patient-Derived Neural Progenitor Cells. *Cell Rep* 2016; **15**(5): 1024-36.
 23. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; **29**(1): 15-21.
 24. Yu H, Wang J, Sheng Q, Liu Q, Shyr Y. beRBP: binding estimation for human RNA-binding proteins. *Nucleic Acids Research* 2019; **47**(5): e26-e.
 25. Lambert SA, Albu M, Hughes TR, Najafabadi HS. Motif comparison based on similarity of binding affinity profiles. *Bioinformatics* 2016; **32**(22): 3504-6.
 26. Yu G, Wang Lg Fau - Han Y, Han Y Fau - He Q-Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. (1557-8100 (Electronic)).
 27. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications* 2019; **10**(1): 1523-.
 28. García-Mauriño SM, Rivero-Rodríguez F, Velázquez-Cruz A, et al. RNA Binding Protein Regulation and Cross-Talk in the Control of AU-rich mRNA Fate. (2296-889X (Print)).
 29. El Fatimy R, Davidovic L, Tremblay S, et al. Tracking the Fragile X

Mental Retardation Protein in a Highly Ordered Neuronal RiboNucleoParticles Population: A Link between Stalled Polyribosomes and RNA Granules. (1553-7404 (Electronic)).

30. Ray D, Kazan H, Cook KB, et al. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 2013; **499**(7457): 172-7.

31. Aberg K, Saetre P Fau - Lindholm E, Lindholm E Fau - Ekholm B, et al. Human QKI, a new candidate gene for schizophrenia involved in myelination. (1552-4841 (Print)).

32. Ince-Dunn G, Okano Hirotsuka J, Jensen KB, et al. Neuronal Elav-like (Hu) Proteins Regulate RNA Splicing and Abundance to Control Glutamate Levels and Neuronal Excitability. *Neuron* 2012; **75**(6): 1067-80.

33. Samuels TA-O, Järvelin AA-O, Ish-Horowicz DA-O, Davis IA-O. Imp/IGF2BP levels modulate individual neural stem cell growth and division through myc mRNA stability. LID - 10.7554/eLife.51529 [doi] LID - e51529. (2050-084X (Electronic)).

34. Brooks SA, Blackshear PJ. Tristetraprolin (TTP): Interactions with mRNA and proteins, and current thoughts on mechanisms of action. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 2013; **1829**(6): 666-79.

35. Han J, Ding J-H, Byeon CW, et al. SR proteins induce alternative exon skipping through their activities on the flanking constitutive exons. *Mol Cell Biol* 2011; **31**(4): 793-802.