Transcriptome sequencing in a 6-hydroxydopamine rat model of Parkinson’s disease

Jia Li, Yajuan Sun and Jiajun Chen*
Department of Neurology, China–Japan Union Hospital of Jilin University, No. 126 Xiantai Road, Erdao District, Changchun, Jilin 130033, China

(Received 29 June 2018, accepted 26 September 2018; J-STAGE Advance published date: 2 February 2019)

The neurotoxin 6-hydroxydopamine (6-OHDA) has been widely exploited as a tool for modeling Parkinson’s disease (PD) in the rat. This study aimed to provide a comprehensive profile of the mRNAs and long noncoding RNAs (lncRNAs) in rats treated with 6-OHDA as a model of PD. Female SPF Wistar rats were randomly divided into two groups: a PD model group and a control group. The PD model was induced by 6-OHDA injection. RNA-seq analysis was performed on 6-OHDA-treated rats and corresponding controls. Novel lncRNAs were identified. Differentially expressed genes (DEGs) and differentially expressed lncRNAs were identified in the PD group compared with controls. Gene Ontology function and pathway enrichment analyses were conducted on the DEGs, followed by construction of a protein–protein interaction (PPI) network. In addition, prediction of lncRNA target genes and function prediction of lncRNAs were performed. Moreover, microRNAs (miRNAs) that interacted with the DEGs and differentially expressed lncRNAs were predicted to construct a miRNA–lncRNA–mRNA regulatory network. A total of 536 DEGs and 512 differentially expressed lncRNAs (44 up-regulated and 10 down-regulated known lncRNAs; 407 up-regulated and 51 down-regulated novel lncRNAs) were identified in the PD rat model compared with controls. The DEGs and target genes of lncRNAs were mainly associated with the innate immune response, 2′-5′-oligoadenylate synthetase activity, GTPase activity, GTP binding and the RIG-I-like receptor signaling pathway. IRF7 and ISG15 were hub proteins in the PPI network. Many mRNAs and lncRNAs interacted with other molecules in a competing endogenous RNA network, such as MAS1, TMPRSS2, NPTX1, XLOC_016191, XLOC_026924 and XLOC_005439. We conclude that IRF7, ISG15, MAS1, TMPRSS2, NPTX1, XLOC_016191, XLOC_026924 and XLOC_005439 may contribute critical roles in the pathogenesis of PD.

Key words: Parkinson’s disease, 6-hydroxydopamine, lncRNA, functional enrichment

INTRODUCTION

Parkinson’s disease (PD) is a common neurodegenerative disease that impairs voluntary movement in middle-aged and elderly people, affecting 1–2% of the population over the age of 60 (Jungverdorben et al., 2017). It is characterized by dopamine deficiency in areas of the midbrain, resulting in motor dysfunctions such as slowness of movement, akinesia, rigidity, postural instability and resting tremor (Vo et al., 2017). There is still no method that can effectively prevent the occurrence and development of PD. Therefore, identifying the molecular mechanisms of the etiology and progression of the disease is a current research hotspot.

In recent years, an increasing number of studies have identified genes as likely candidates for involvement in PD. For instance, a network-based meta-analysis demonstrated that HNF4A and PTBP1 are longitudinally dynamic biomarkers for PD (Santiago and Potashkin, 2015). Yeo et al. showed that decreased expression of serum- and glucocorticoid-dependent kinase 1 promotes the expression of alpha-synuclein, which is related to dopaminergic cell death in chronic Parkinsonism mice (Yeo et al., 2018). Long noncoding RNAs (lncRNAs) are defined as transcripts longer than 200 nucleotides that are transcribed from DNA but not translated into proteins (Kopp and Mendell, 2018). Soreq et al. used an all-RNA transcriptome sequencing method to screen...
LncRNAs associated with PD and revealed, for example, a significant increase in the expression of lnc-frg1-3, which is associated with muscle rigidity in PD (Soreq et al., 2014). Another study found that the expression of a ubiquitin carboxyl-terminal hydrolase Uchl1 antisense strand lncRNA was regulated by the major transcription factor Nurr1, which is involved in the differentia-
tion and maintenance of dopamine cells (Carriera et al., 2015). However, there are substantially more genes and lncRNAs related to PD remaining to be explored.

In the current study, RNA-seq analysis was performed on a 6-hydroxydopamine (6-OHDA) rat model of PD and corresponding control animals. Novel lncRNAs were identified. Differentially expressed mRNAs and lncRNAs were identified in the PD group compared with controls. Gene Ontology (GO) function analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were conducted on differentially expressed genes (DEGs), followed by construction of a protein–protein interaction (PPI) network. In addition, lncRNA target genes and function of lncRNAs were predicted. Moreover, microRNAs (miRNAs) that interacted with the DEGs and differentially expressed lncRNAs were predicted before the construction of a miRNA–lncRNA–mRNA regulatory network. This study aimed to provide a comprehensive profile of the mRNAs and lncRNAs in a 6-OHDA rat model of PD and to identify important new gene targets and lncRNA targets for PD, which should contribute to our understanding of the pathogenesis of PD.

**MATERIALS AND METHODS**

**Animals** Eighteen inbred female SPF Wistar rats weighing 200–220 g, 10 to 12 weeks of age, were used in the present study. Before experiments, the animals were kept in a controlled facility (temperature, 24 ± 2 °C; humidity, 50 ± 5%) with ad libitum access to food and water under a 12/12 h light/dark cycle for one week. All animal procedures were carried out in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

**Experimental design and surgical procedure** The 18 experimental animals were randomly divided into two groups: control group and PD model group (n = 9 for each group). Rats were anesthetized with an intraperitoneal injection of 3% chloral hydrate (10 ml/kg). The animals were fixed into a stereotaxic instrument (Zhongshi, Beijing) secured by ear bars adapted specifically for each rat. The top of the head was shaved. The skull was exposed by an incision in the midline and small holes were drilled in the skull to allow a 4-µl injection of 6-OHDA (Sigma, St. Louis) (dissolved in 0.2% ascorbic acid saline solution) into each of the substantia nigra (antero-posterior: +4.8 ± 0.1 mm, medio-lateral: −1.7 ± 0.1 mm, dorso-ventral: −7.5 ± 0.1 mm below dural surface) and the ventrolateral region of the caudate-putamen (antero-posterior: 4.6 ± 0.1 mm, medio-lateral: 0.9 ± 0.1 mm, V: 7.3 ± 0.1 mm from bregma) according to Paxi-
no's and Watson (1997). Sham-operated rats, which received 0.02% ascorbic acid in sterile saline bilaterally at the same coordinates, served as controls in all experiments. 6-OHDA was injected at a rate of 1 µl/min, and the injection syringe was left in place for a further 10 min after injection to enable full absorption of the solution. Surgical clips were used for closing the incisions before rats were returned to the animal housing facility for recovery. To avoid infections, the rats received a daily injection of penicillin (50,000 units) for five days afterwards.

**Rotational asymmetry analysis** At one week after surgery, behavioral changes and abnormal behavior of the 6-OHDA-lesioned PD rats were observed. Rats were intraperitoneally injected with 0.5 mg/kg of apomorphine (Sigma), and the number of rotations of rats thus induced was recorded manually for a 10-min period. Further inductions were performed at two, four and six weeks after surgery. If the rat turned steadily to the left, and the number of rotations was > 7/min in all of these measurements, it was considered a valid PD model and was used for further experiments (da Conceição et al., 2010).

**RNA isolation and sequencing** Rats were killed by carbon dioxide asphyxiation and the brains were rapidly removed. The right and left striatum were isolated and rinsed with saline, snap-frozen in liquid nitrogen and kept at −80 °C until further use. Total RNA from the rat brain was isolated with RNaio Plus reagent (Takara Bio, Kusatsu) according to the manufacturer’s instruc-
tions. RNA-seq libraries were prepared, followed by Illu-
mina sequencing by Sangon Biotech (Shanghai).

**Sequence read processing and genome mapping** The raw sequencing data were first filtered to generate clean data using SeqPrep (Schmieder and Edwards, 2011) (https://github.com/jstjohn/SeqPrep). Reads containing adaptors were removed. Bases having a quality score of less than 20 at the 3' end of the read were trimmed; if the remaining sequence still had a quality score of less than 10, the entire read was removed. Reads containing more than 10% of unknown nucleotides and reads with a minimum length of less than 30 bp were discarded. Clean reads were mapped to the rat reference genome (Rnor_6.0.88, http://www.ensembl.org/index.html) using tophat2 (Schmieder and Edwards, 2011) software (v2.1.0, http://ccb.jhu.edu/software/tophat/index.shtml) using the following parameters: read-mismatches 2, read-
edit-dist 2, max-intron-length 500000, num-threads 8,
RNA sequence analysis and identification of differentially expressed transcripts  Transcript assembly was performed using Cufflinks (Mortazavi et al., 2008) software (v2.2.1, http://cufflinks.cbcb.umd.edu/) and the number of read-counts of the corresponding genes was then obtained using the HTSeq tool (Mortazavi et al., 2008) (V0.6.1p2) based on the rat genome annotation information. The results were normalized by the CPM (counts per million) method, and genes having a CPM value of less than 0.1 in five or more samples were defined as genes with low expression abundance. The obtained results were classified into IncRNA and mRNA according to the gene type in the annotation information.

A transcript of class code type iu in the assembled transcript (gene) is considered as an unknown transcript (i indicates that the predicted transcript is located entirely within the intron range, u indicates an unknown intergenic transcript, and x indicates exonic overlap with a reference on the opposite strand). Three software packages were used to predict the coding ability of transcript sequences, CPC (Kong et al., 2007) (v0.9-r2, http://cpc2.cbi.pku.edu.cn/), CNCI (Sun et al., 2013) (https://github.com/www-bioinfo-org/CNCI) and PfamScan (Mistry et al., 2007) (https://www.ebi.ac.uk/Tools/pfa/pfamscan/), with the following threshold settings: CPC_threshold = 0 (assembled transcripts > 0 was mRNA, or assembled transcripts < 0 was lncRNA), CNCI_threshold = 0 (assembled transcripts > 0 was mRNA, or assembled transcripts < 0 was lncRNA) and Pfam as a protein database. The results obtained from the intersection of t for all three judgment methods were used for subsequent analysis.

The edgeR (Lun et al., 2016) software package in R was used for differential expression analysis. Differentially expressed lncRNAs and mRNAs were identified in the PD model group compared with controls. Prior to analysis, genes with lower expression abundance were deleted and thresholds were set as |log2 fold change (FC)| > 0.585 and P < 0.05. In addition, a two-dimensional clustering heatmap was generated.

GO and KEGG enrichment analysis  GO functions and KEGG pathways of DEGs were analyzed using the DAVID (Huang et al., 2008) (https://david.ncifcrf.gov/) tool. GO terms and pathways with gene count ≥ 2 were enriched. P values < 0.05 based on the hypergeometric test were considered significant.

PPI prediction  PPI analysis was performed on DEGs using the STRING (Huang et al., 2008) (v10.5, https://string-db.org/) online tool. Medium Confidence (combined score) > 0.4 was selected as the threshold. Cytoscape (Shannon et al., 2003) (v3.5.1, http://www.cytoscape.org/) was used for the construction of a PPI network. The topology of the network was analyzed, the connectivity of network nodes was ranked, and nodes with higher connectivity were considered as the important nodes.

Prediction of IncRNA target genes and function prediction of lncRNAs  For the resulting DEGs and differentially expressed lncRNAs, the correlation between them was calculated using Pearson’s correlation coefficient (Benesty et al., 2009), and the P value was corrected by the Benjamini and Hochberg method (Benjamini and Hochberg, 1995). A correlation coefficient greater than 0.9 and corrected P < 0.01 were chosen as the threshold. mRNAs that had a correlation with lncRNAs were potential target genes for lncRNAs. The IncRNA–mRNA network was built using cytoscape software (Shannon et al., 2003).

The functions of lncRNAs were predicted by performing GO function and KEGG pathway enrichment analysis on lncRNA target genes. We used the clusterProfiler (Yu et al., 2012) (http://www.bioconductor.org/) package to compare the KEGG pathways of IncRNA target genes. The screening threshold was set as P < 0.05.

miRNA prediction and ceRNA network construction  Competing endogenous RNA (ceRNA) regulatory networks are more elaborate and complex, involving more RNA molecules, including mRNA, pseudogene transcripts, lncRNAs and miRNAs (Yuan et al., 2017). We downloaded conserved miRNA-regulated mRNA data from miRanda (http://34.236.212.39/microrna/home.do), and the miRanda tool (Betel et al., 2010) was also applied to predict binding sites between miRNAs and lncRNAs (parameter: -sc -160, -en -20). miRNA–IncRNA, miRNA–mRNA and lncRNA–mRNA and lncRNA–IncRNA regulatory relationships were integrated to obtain the miRNA–IncRNA–mRNA regulatory network.

RESULTS

Overview of RNA-seq and mRNA and lncRNA identification  After the rotational behavior test, six PD rats were valid and were used for the subsequent sequencing. The six PD rats in the model group were named MW1-L, MW1-R, MW2-L, MW2-R, MW3-L and MW3-R; six rats in the control group were chosen for sequence analysis and named CW7-L, CW7-R, CW8-L, CW8-R, CW9-L and CW9-R. After sequence data preprocessing, more than 91% reads in each sample were aligned to the reference genome (Supplementary Table S1). The reads were mapped to the reference genome, resulting in 40,752 assembled transcripts, of which 13,613 genes had low expression abundance.

Under the threshold conditions (|log2 FC| > 0.585 and P < 0.05), a total of 536 differentially expressed mRNAs (418 up-regulated and 118 down-regulated) and 512 dif-
ferentially expressed lncRNAs (44 up-regulated known lncRNAs, 10 down-regulated known lncRNAs, 407 up-regulated novel lncRNAs and 51 down-regulated novel lncRNAs) were identified. Detailed information about differentially expressed mRNAs and known lncRNAs is presented in Supplementary Table S2. Figure 1 shows two-dimensional heatmaps of differentially expressed mRNAs and lncRNAs.

**GO and KEGG pathway enrichment analysis of DEGs** A total of 141 GO terms and 29 KEGG pathways were identified for DEGs. In the GO analysis, the top five enriched terms in the biological process (BP), cellular component (CC) and molecular function (MF) categories are shown in Fig. 1C. We found that the DEGs were mainly associated with the following BP terms: defense response to virus, cellular response to interferon-beta, innate immune response and negative regulation of viral genome replication, whereas with regard to the CC terms the DEGs appeared to be active in both the extracellular region and the extracellular matrix (ECM) (Fig. 1C). The main MF terms were 2′-5′-oligoadenylate synthetase activity, double-stranded RNA binding, CXCR3 chemokine receptor binding, GTPase activity and GTP binding. Additionally, the DEGs were mainly involved in pathways such as protein digestion and absorption, herpes simplex and influenza A infection, RIG-I-like receptor signaling pathway and ECM–receptor interaction.

**PPI analysis** A PPI network for DEGs was constructed (Fig. 2), comprising 149 nodes and 434 interaction pairs. The key nodes in the network were identified. Table 1 lists the top 20 node genes with the highest connectivity in the network, such as interferon regulatory factor 7 (Irf7, degrees = 18), ISG15 ubiquitin-like modifier (Isg15, degrees = 13), ubiquitin-like modifier activating enzyme 7 (Uba7, degrees = 13), 2′-5′-oligoadenylate synthetase 1 (Oas1a, degrees = 12) and C-X-C motif chemokine ligand 10 (Cxcl10, degrees = 11).

![Fig. 1. Heatmaps and functional enrichment analysis. (A) Heatmap of differentially expressed genes (DEGs). Red indicates up-regulated genes and blue indicates down-regulated genes. (B) Heatmap of differentially expressed lncRNAs. Red indicates up-regulated lncRNAs and blue indicates down-regulated lncRNAs. (C) The top five GO terms and pathways enriched in DEGs. Number of genes: the number of differentially expressed genes enriched in the indicated terms. The gray trend lines represent the –log₁₀ (P values), and the values (0–30) on the horizontal axis also apply to these data. BP, biological process; CC, cellular component; MF, molecular function.](image)
Prediction of target genes for differentially expressed lncRNAs and functional analysis. Using the CPC, CNCI and PfamScan software to obtain intersection results, 7,506 novel lncRNAs were obtained (Fig. 3A), 458 of which were differentially expressed. Target genes for differentially expressed lncRNAs were predicted using Pearson’s correlation coefficient. lncRNAs with more than five target genes were analyzed. A lncRNA–mRNA interaction network was constructed (Fig. 3B). The top 20 lncRNAs with the highest connectivity are shown in Table 2, including XLOC_026924, XLOC_029397, XLOC_004631, XLOC_005439 and XLOC_028318.

To further understand the characteristics of these lncRNAs, functional enrichment analysis was performed.
on the target genes involved in the network. A total of 46, 15 and seven terms were enriched for the GO categories BP, MF and CC, respectively, and 10 KEGG pathways were enriched simultaneously. The top five most significant terms of BP, MF and CC as well as pathways are represented in Fig. 3C. The lncRNAs were associated with the defense response to virus, the cellular response to interferon-beta, the innate immune response, 2′-5′-oligoadenylate synthetase activity, GTPase activity, GTP binding, calcium channel activity and the RIG-I-like receptor signaling pathway.

miRNA prediction and ceRNA construction We downloaded the conserved miRNAs for rat mRNA regulation from the miranda website and screened for miRNAs that regulate the target genes of differentially expressed lncRNAs. In addition, miranda software was used to predict binding sites between miRNAs and lncRNAs. Thus, the top 10 miRNAs that had binding sites with the differentially expressed lncRNAs are listed in Table 3, and are rno-miR-466b, rno-miR-349, rno-miR-34b, rno-miR-200b, rno-miR-483, rno-miR-27a, rno-miR-532-5p, rno-miR-34c, rno-miR-532-5p and rno-miR-466b. To integrate the miRNA–lncRNA, miRNA–mRNA and lncRNA–mRNA interactions, a miRNA–lncRNA–mRNA regulatory network was constructed (Fig. 4). Several clusters were extracted from the ceRNA network (Fig. 5). From the results, we found that DEGs such as neuronal pentraxin 1 (Nptx1), transient receptor potential cation channel subfamily V member 5 (Trpv5), MAS1 proto-oncogene, G protein-coupled receptor (Mgs1), MX dynamin like GTPase 1 (Mx1), calcium-binding protein 7 (Cabp7), transmembrane protease, serine 2 (Tmprss2), and radical S-adenosyl methionine domain containing 2 (Rsad2) could be regulated by many lncRNAs and miRNAs.

DISCUSSION

In the current study, 536 DEGs (418 up-regulated and 118 down-regulated) and 512 differentially expressed lncRNAs (44 up-regulated known lncRNAs, 10 down-

Table 3. The top 10 miRNAs that had binding sites with differentially expressed lncRNAs

<table>
<thead>
<tr>
<th>miRNA</th>
<th>lncRNA</th>
<th>Max. score</th>
<th>Max. energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>rno-miR-466b</td>
<td>XLOC_022926</td>
<td>183.00</td>
<td>−26.10</td>
</tr>
<tr>
<td>rno-miR-349</td>
<td>XLOC_004631</td>
<td>183.00</td>
<td>−33.06</td>
</tr>
<tr>
<td>rno-miR-34b</td>
<td>Cd93</td>
<td>181.00</td>
<td>−26.90</td>
</tr>
<tr>
<td>rno-miR-200b</td>
<td>XLOC_036577</td>
<td>181.00</td>
<td>−20.54</td>
</tr>
<tr>
<td>rno-miR-483</td>
<td>XLOC_029657</td>
<td>178.00</td>
<td>−32.47</td>
</tr>
<tr>
<td>rno-miR-27a</td>
<td>Cd93</td>
<td>178.00</td>
<td>−24.35</td>
</tr>
<tr>
<td>rno-miR-532-5p</td>
<td>XLOC_029397</td>
<td>177.00</td>
<td>−28.28</td>
</tr>
<tr>
<td>rno-miR-34c</td>
<td>Cd93</td>
<td>177.00</td>
<td>−28.30</td>
</tr>
<tr>
<td>rno-miR-532-5p</td>
<td>XLOC_004631</td>
<td>176.00</td>
<td>−29.62</td>
</tr>
<tr>
<td>rno-miR-466b</td>
<td>AABR07006685.1</td>
<td>176.00</td>
<td>−29.69</td>
</tr>
</tbody>
</table>

Fig. 4. miRNA–lncRNA–mRNA regulatory network. Red represents up-regulation; triangles are lncRNAs, diamonds are miRNAs and circles are mRNAs.
regulated known lncRNAs, 407 up-regulated novel lncRNAs and 51 down-regulated novel lncRNAs) were identified in a PD rat model compared with controls. The DEGs and target genes of lncRNAs were mainly associated with the innate immune response, 2′-5′-oligoadenylate synthetase activity, GTPase activity, GTP binding and the RIG-I-like receptor signaling pathway. IRF7 and ISG15 were hub proteins in the PPI network. Many mRNAs and lncRNAs interacted with other molecules in the ceRNA network such as Mas1, Tmprss2, Nptx1, XLOC_016191, XLOC_026924 and XLOC_005439.

Chronic neuroinflammation processes in PD have an important role that leads to protracted neuronal degeneration (Lee et al., 2009; Tansey and Goldberg, 2010). Evidence indicates that 2′-5′-oligoadenylate synthetase-like protein can be a useful biomarker for neurodegenerative disease diagnosis (Goldknopf et al., 2006). Several studies demonstrated the important roles of GTPase activity and GTP binding in neuronal toxicity and the development of PD (Guo et al., 2007; Ito et al., 2007; West et al., 2007). In addition, one study reported that cytosolic RIG-I-like helicases could be negative regulators of inflammation in the central nervous system (Dann et al., 2011). These studies provide an explanation for the results obtained by functional enrichment analysis, suggesting that our analysis is reliable.

Evidence demonstrates that type I interferon (IFN-) dependent immune responses are under the control of IRF7 (Honda et al., 2005). Recently, type I IFNs were reported to contribute to the early neuroinflammatory response and neuronal cell death in the mouse model of PD (Main et al., 2016). In addition, a previous genome-wide association study indicated that IRF7 was associated with systemic lupus erythematosus, which suggests a mechanistic relationship between systemic lupus erythematosus and PD via IFN signaling (Liu et al., 2015). The ubiquitin-like protein ISG15 is one of several hundred proteins induced by IFN-α/β (Zhao et al., 2005). ISG15-mediated impairment of protein degradation was shown to contribute to progressive neurodegeneration (Wood et al., 2011). In this study, IRF7 and ISG15 were distinct hub proteins in the PPI network. Taken together, our results suggest that IRF7 and ISG15 play significant roles in the pathogenesis of PD, which warrants further experimental validation in future.

Neuronal pentraxin II (NPTX2) is highly up-regulated in PD (Moran et al., 2008). Qu and D’Mello showed that NPTX1 can be a target of histone deacetylase-3-mediated neurodegeneration (Qu and D’Mello, 2018). In this study, Nptx1 was an important player in the ceRNA network. Additionally, Mas1 and Tmprss2 were also shown to be critical in the ceRNA network. The appearance of Mx1 protein in reactive microglia may contribute to Alzheimer’s disease pathology (Wennström et al.,
Lahiry et al. showed a mutation in the serine protease TMPRSS4 in a novel neurodegenerative disorder (Lahiry et al., 2013). However, no study has reported the roles of XLOC_016191, XLOC_026924 or XLOC_005439. Nevertheless, we suggest that MAS1, TMPRSS2, NPTX1, XLOC_016191, XLOC_026924 and XLOC_005439 may contribute critical roles in the development of PD. Future studies are required to address their functional significance.

CONCLUSIONS

In conclusion, the present study has provided a comprehensive analysis of the DEGs and differentially expressed lncRNAs in a rat PD model by RNA-seq and bioinformatics analysis. This study indicated significant functions for the immune response, 2′-5′-oligoadenylate synthetase activity, GTPase activity, GTP binding, and the RIG-I-like receptor signaling pathway. Moreover, IRF7, ISG15, MAS1, TMPRSS2, NPTX1, XLOC_016191, XLOC_026924 and XLOC_005439 may contribute critical roles in the pathogenesis of PD, which provides a foundation for further research into the molecular mechanisms of PD and may aid the identification of sensitive biomarkers and therapeutic targets for this disease.

This work was supported by the Jilin Province Neurological Diseases Precision Medicine Science and Technology Innovation Center (No. 20170623006TC).

REFERENCES


