



OPEN

mRNA and miRNA expression profile reveals the role of miR-31 overexpression in neural stem cell

Pengfei Li^{1,3}, Yuantao Gao², Xiao Li³, Feng Tian³, Fei Wang³, Yali Wang³, Bichun Zhao³, Ruxin Zhang³ & Chunfang Wang³✉

A detailed understanding of the character and differentiation mechanism of neural stem cells (NSCs) will help us to effectively utilize their transplantation to treat spinal cord injury. In previous studies, we found that compared with motor neurons (MNs), miR-31 was significantly high-expressed in NSCs and might play an important role in the proliferation of NSCs and the differentiation into MNs. To better understand the role of miR-31, we characterized the mRNA and miRNAs expression profiles in the early stage of spinal cord-derived NSCs after miR-31 overexpression. There were 35 mRNAs and 190 miRNAs differentially expressed between the miR-31 overexpression group and the control group. Compared with the control group, both the up-regulated mRNAs and miRNAs were associated with the stemness maintenance of NSCs and inhibited their differentiation, especially to MNs, whereas the down-regulated had the opposite effect. Further analysis of the inhibition of miR-31 in NSCs showed that interfering with miR-31 could increase the expression of MNs-related genes and produce MNs-like cells. All these indicated that miR-31 is a stemness maintenance gene of NSCs and has a negative regulatory role in the differentiation of NSCs into MNs. This study deepens our understanding of the role of miR-31 in NSCs, provides an effective candidate target for effectively inducing the differentiation of NSCs into MNs, and lays a foundation for the effective application of NSCs in clinic.

Spinal cord injury (SCI) is the structural and functional damage of the spinal cord caused by various reasons, resulting in the impairment below the level of injury of spinal nerve function. The main characteristic of SCI is the death of cholinergic motor neurons (MNs)¹, which is a severely disabling trauma. At the site of spinal cord injury, effective nerve regeneration rarely occurs, and injured neurons seldom repair themselves at the site of injury^{2,3}. Accordingly, finding ways to improve the environment of nerve regeneration at the injured site and promote the recovery of injured MNs has become a research focus in the field of SCI treatment. After the discovery of neural stem cells (NSCs), cell transplantation has become a promising and feasible option for the treatment of SCI⁴. However, in the case of SCI model mice, almost all transplanted NSCs differentiated into glial cells⁵. In addition to the influence of the microenvironment of the injured site, the lack of understanding of the NSCs differentiation mechanism into MNs is also the main reason that the transplanted cells cannot effectively differentiate into the required neurons.

Cell differentiation is a precise process that relies on precise control over the spatial and temporal expression of transcriptional regulators, especially silencing of previously active molecules and activation of new molecular programs⁶, thus establishing clear temporal and spatial boundaries for the expression of corresponding genes⁷, ultimately triggering overall changes in cells. Studies have shown that a kind of small non-coding RNA, called microRNA (miRNA), produced by RNaseII-Dicer, can precisely regulate the expression of target genes by inhibiting the translation of mRNAs and plays an important role in various cellular processes^{8,9}. Although the effect of miRNAs on single target gene inhibition is limited, each miRNA can recognize and inhibit more than hundreds of mRNA targets, and the increase or decrease of their expression may eventually lead to a comprehensive change in the gene expression profile of cells, thus providing a guarantee for the stable transformation of cell fate. The same findings have been found in the study of NSCs, such as let-7b could regulate the proliferation and differentiation of NSCs through the nuclear receptor TLX signal¹⁰, TLX and miR-9 could form a feedback loop to affect the differentiation of NSCs¹¹, and miR-133b plays an important role in regulating maturity and functional aspects of the midbrain dopaminergic neurons¹². Therefore, we believe that the differentiation process of NSCs

¹Translational Medicine Research Center, Shanxi Medical University, Xinjian South Road 56, Taiyuan 030001, Shanxi, People's Republic of China. ²Nanchang University, Nanchang 330000, People's Republic of China. ³Laboratory Animal Center, Shanxi Medical University, Xinjian South Road 56, Taiyuan 030001, Shanxi, People's Republic of China. ✉email: wangchunfang@sxmu.edu.cn

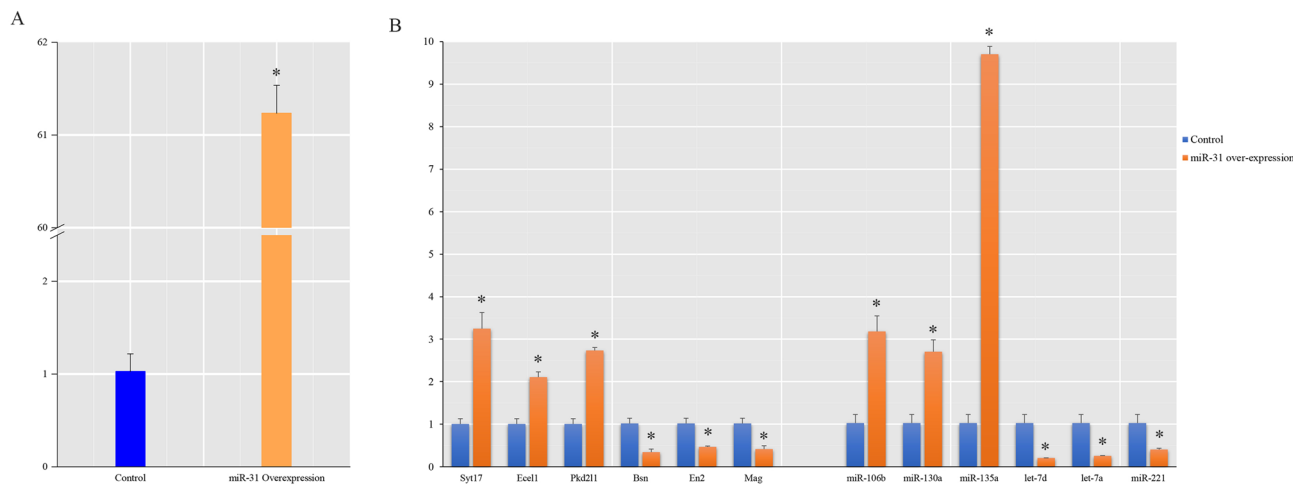


Figure 1. The q-PCR result of miR-31 and some DEGs, DEmiRNAs after miR-31 overexpression. (A) Compared with the control group, the miR-31 overexpression group had significantly higher miR-31 levels. (*Indicates $p \leq 0.05$, compared with its control group.). (B) Experiments to verify the reliability of sequencing results showed that the q-PCR results of randomly selected DEGs and DEmiRNAs had the same trend as their sequencing results. (*Indicates $p \leq 0.05$, compared with its control group).

can be regulated by some key miRNAs. If changing the expression of these key miRNA in the differentiation process, it may improve the differentiation proportion of NSCs to the specific neurons effectively. Based on these understandings, in order to find the miRNAs that play a key role in the differentiation of NSCs into MNs, our group compared the specific expression profiles of spinal cord-derived NSCs and MNs using TaqMan low-density array (TLDA) technology in a previous study, and analyzed the differences between them in the level of regulation of miRNAs¹³. In this study, we found that miR-31 was more than 90 times more expressed in NSCs than MNs, indicating that miR-31 mainly functions in NSCs. Studies have shown that miR-31 can promote the expansion of breast stem cells¹⁴ and intestinal stem cells¹⁵. Taken together, we can speculate that miR-31 is a key stemness maintenance gene of NSCs and may have a negative regulatory role in the differentiation of NSCs into MNs, and the study of the specific role of miR-31 in NSCs can help us further strengthen our understanding of NSCs and their differentiation mechanism into MNs.

For a certain miRNA, regardless of its interference or overexpression, the change of the initial mRNA expression profile of the cell is the most direct function of the miRNA in the cell. Therefore, in this work, we used RNA-seq analysis to determine the differential expression of genes and miRNAs after miR-31 overexpression in spinal cord-derived NSCs, and integrated them to understand the role of miR-31 in NSCs and differentiation.

Result

miR-31 differential expression between overexpression group and control group. We analyzed the miR-31 difference expression between the miR-31 overexpression group and the control group by q-PCR, and found that the expression of miR-31 in the overexpression group was 61.2 times higher than that in the control group (Fig. 1A). This means that our overexpression system can effectively improve the expression of miR-31 in NSCs.

mRNA sequencing data mapping and annotation. A total of 2 cDNA libraries were sequenced from the miR-31 overexpression group and miR-31 overexpression control group. After removing the adaptors and filtering, 35,173,328 clean reads were obtained from the miR-31 overexpression group and 35,353,796 clean reads were obtained from control group. Then we compared the beads with the reference genome sequences by using TopHat2 software¹⁶. We found that 84.25% reads were successfully aligned in miR-31 overexpression group and 84.10% reads were successfully aligned in control group.

miRNA sequencing data mapping and annotation. A total of 2 cDNA libraries were sequenced from the miR-31 overexpression group and miR-31 overexpression control group. After removing reads with low quality, trimming the 3'adapter and discarding the sequences shorter than 18 nt and longer than 30 nt, 15,266,203 clean reads were obtained from the miR-31 overexpression group and 13,923,446 clean reads were obtained from control group. Using Bowtie¹⁷ software, clean Reads were aligned with Silva database, GtRNAdb database, Rfam database and Rfam database respectively. Unannotated reads containing miRNAs were obtained by filtering ncRNAs such as ribosomal RNA (rRNA), transport RNA (tRNA), intranuclear small RNA (snRNA), nucleolar small RNA (snoRNA) and repetitive sequences. The known and new miRNAs were identified by using the software of miRDeep2¹⁸.

DEGs and DEmiRNAs between miR-31 overexpression group and control group. There were 35 DEGs between the miR-31 overexpression group and the control group. Among these DEGs, 22 (62.9%) genes

COG	GO	KEGG	Swiss-Prot	NR	Total
12	28	16	32	34	34

Table 1. DEGs annotation result statistics.

were down-regulated while 13 (37.1%) genes were up-regulated in the miR-31 overexpression group compared with control group, and 9 (25.7%) genes were the new genes (Supplementary Table S3).

After miRNA-Seq, we obtained 1063 miRNAs, and 980 miRNAs were known miRNAs, others were new predicted miRNAs. Among them, there were 190 miRNAs differentially expressed between overexpression group and its control group. 20 (10.5%) miRNAs were down-regulated while 169 (89.5%) miRNAs were up-regulated in the miR-31 overexpression group compared with control group, and 17 (8.9%) miRNAs may be the new miRNAs (Supplementary Table S4).

The results of q-PCR validated the credibility of sequencing results. To validate the sequencing results, we investigated the relative expression levels by randomly selecting 6 mRNAs (Bsn, Ecel1, En2, Mag, PKD2L1, Syt17) and 6 miRNAs (mmu-let-7a-5p, mmu-let-7d-5p, mmu-miR-106b-5p, mmu-miR-130a-5p, mmu-miR-135a-5p, mmu-miR-221-5p) by q-PCR (Fig. 1B), and the results showed that there was same trend of difference between q-PCR results and sequencing results, which indicated the reliability of the sequencing analysis results.

Functional analysis of DEGs. After functional annotation, the numbers of all DEGs annotated to each database are shown in Table 1.

COG (Cluster of Orthologous Groups of proteins) database is based on the phylogenetic relationship of bacteria, algae and eukaryotes, which can be used for orthologous classification. Among different functional classes, the proportion of genes reflects the metabolic or physiological bias in the corresponding period and environment. Figure 2A showed the up-regulated DEGs were mainly distributed in chromatin structure and dynamics, lipid transport and metabolism, intracellular trafficking, secretion, and vesicular transport of the COG classification. And the down-regulated DEGs were mainly distributed in translation, replication, recombination and repair, signal transduction mechanisms, inorganic ion transport and metabolism of the COG classification.

GO annotated DEGs mainly belonged to the three functional clusters (biological process, BP; cellular component, CC; molecular function, MF). The GO annotation classification statistical graph shows the number of genes annotated to the pathway and their proportion to the total number of genes annotated, and reflects the status of the secondary functions of GO in the context of DEGs as well as all genes. The obvious proportion difference indicates that the proportion trend of DEGs and all genes under this secondary function is different, and the function may be closely related to the expression difference. Figure 2B showed that compared with the whole genetic background, the main differences in the up-regulated DEGs distribution trend were the reproductive process, localization and developmental process and signaling of the BP cluster, the membrane part, cell junction, and synapse part of the CC cluster, and the transporter activity, catalytic activity of the MF cluster. The main differences in the down-regulated DEGs distribution trend were the hormone secretion, rhythmic process and biological adhesion of the BP cluster, the extracellular matrix, synapse part and cell junction of the CC cluster, and the structural molecule activity, nucleic acid binding transcription factor activity and transporter activity of the MF cluster.

The KEGG (Kyoto Encyclopedia of Genes and Genomes) database is the main public database on metabolic pathways. The KEGG classification map shows the number of genes annotated to this pathway and their proportion to the total number of genes annotated. Figure 2C showed the up-regulated DEGs were mainly concentrated in the AMPK signaling pathway (Environmental information processing), and PPAR signaling pathway (Organismal systems). The down-regulated DEGs were mainly concentrated in endocytosis (Cellular processes) and cell adhesion molecules (Environmental information processing).

GO and KEGG enrichment analysis found that only 4 down-regulated DEGs were significantly enriched in cell adhesion molecules (CAMs) in KEGG enrichment analysis, and there were no other significant enrichments.

Functional analysis of DEmiRNA. After functional annotation, the number of all potential target genes (PTGs) of DEmiRNA annotated to each database were shown in Table 2.

Figure 3A showed that in addition to the general function prediction only of COG classification, no matter the up or down regulated DEmiRNAs, distribution of PTGs were mostly in transcription, replication, recombination and repair, signal transduction mechanisms.

Figure 3B showed that compared with the whole genetic background, the main differences in the distribution trend of up-regulated DEmiRNAs PTGs were the biological adhesion, growth and rhythmic process of the BP cluster, the synapse, extracellular matrix part and collagen trimer of the CC cluster, and the nucleic acid binding transcription factor activity, guanyl-nucleotide exchange factor activity, and the protein binding transcription factor activity of the MF cluster. The main differences in the distribution trend of down-regulated DEmiRNAs PTGs were the reproductive process, growth and cell aggregation of the BP cluster, the nucleoid, extracellular matrix part and synapse part of the CC cluster, and the receptor regulator activity, guanyl-nucleotide exchange factor activity, and structural molecule activity of the MF cluster.

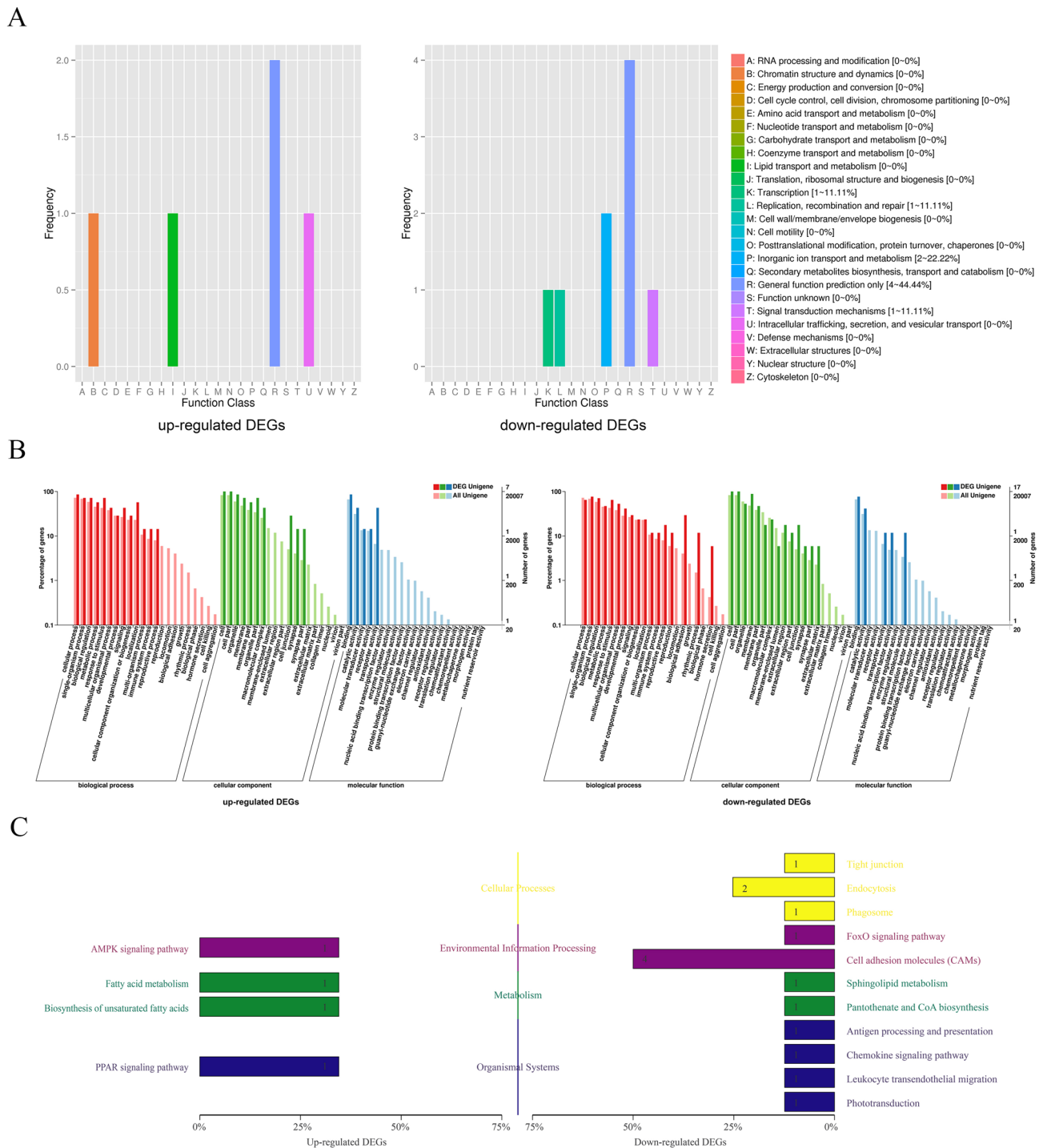


Figure 2. Annotated statistical chart of DEGs. **(A)** Statistical map of COG annotation classification of DEGs showed the up-regulated DEGs were mainly distributed in chromatin structure and dynamics, lipid transport and metabolism, intracellular trafficking, secretion, and vesicular transport. And the down-regulated DEGs were mainly distributed in translation, replication, recombination and repair, signal transduction mechanisms, inorganic ion transport and metabolism. **(B)** Statistical map of GO annotation classification of DEGs showed that the up-regulated DEGs were mainly distributed in the reproductive process, localization and developmental process and signaling of the BP cluster, the membrane part, cell junction, and synapse part of the CC cluster, and the transporter activity, catalytic activity of the MF cluster. The down-regulated DEGs were distributed in the hormone secretion, rhythmic process and biological adhesion of the BP cluster, the extracellular matrix, synapse part and cell junction of the CC cluster, and the structural molecule activity, nucleic acid binding transcription factor activity and transporter activity of the MF cluster. **C.** KEGG classification map of DEGs showed the up-regulated DEGs were mainly concentrated in the AMPK signaling pathway (Environmental information processing), and PPAR signaling pathway (Organismal systems). The down-regulated DEGs were mainly concentrated in endocytosis (Cellular processes) and cell adhesion molecules (Environmental information processing).

COG	GO	KEGG	Swiss-Prot	NR	Total
2198	5966	3359	6553	6549	6553

Table 2. Annotation result statistics.

Figure 3C showed the PTGs of the up-regulated DEGs were mainly concentrated in the focal adhesion (Cellular processes), MAPK signaling pathway (Environmental information processing), purine metabolism (Metabolism), axon guidance (Organismal systems) and ubiquitin mediated proteolysis (Genetic information processing). The PTGs of the down-regulated DEGs were mainly concentrated in focal adhesion (Cellular processes) and MAPK signaling pathway (Environmental information processing), purine metabolism (Metabolism), insulin signaling pathway (Organismal systems) and ubiquitin mediated proteolysis (Genetic information processing).

After enrichment analysis, we selected GO categories associated with nervous system to plot Fig. 4A, which showed that PTGs were mainly grouped into the neuron differentiation, axonogenesis and Wnt signaling pathway of the BP cluster, the neuron projection and synapse of the CC cluster, and RNA polymerase II regulatory region DNA binding of the MF cluster. Figure 4B showed the significantly enriched KEGG pathways of PTGs were metabolic pathways, focal adhesion, PI3K-Akt signaling pathway.

The integration analysis of the DEGs and DEmiRNAs. Using DIANA-TarBase V.8 and miRWalk, we obtained VTGs of DEmiRNAs and then cross-analyzed them with DEGs. Since the effect of miRNAs on target genes is mainly to inhibit their expression, we constructed a regulatory network of miRNA-mRNA based on the opposite expression patterns of DEGs, VTGs and DEmiRNAs (Supplementary Table S5, Fig. 5). The results showed that 43 DEmiRNAs and 12 DEGs had opposite expression patterns, suggesting a regulatory relationship between them. In addition, further analysis of PTGs and VTGs revealed that 4 DEGs (Atp10b, Grk1, Ppp1r16b and Slc24a2) were PTGs of miR-31-5p and the miRNA biogenesis gene Dicer1 and inhibition-related genes Ago3 were VTGs of miR-31-5p. Analysis of VTGs of DEmiRNAs with opposite expression patterns to DEGs revealed that about 10 VTGs were MNs differentiation-related genes and 14 VTGs were NSCs differentiation-related genes.

To further understand the PPI network between DEGs and VTGs, we analyzed them using String database (Fig. 6). The results showed that although DEGs had no direct interaction relationship with each other, they established complex interaction networks through some intermediate node genes. Most of these intermediate node genes are not only key genes related to MNs differentiation, NSCs differentiation or stemness maintenance, but also VTGs of DEmiRNAs, especially some of them are VTGs of miR-31.

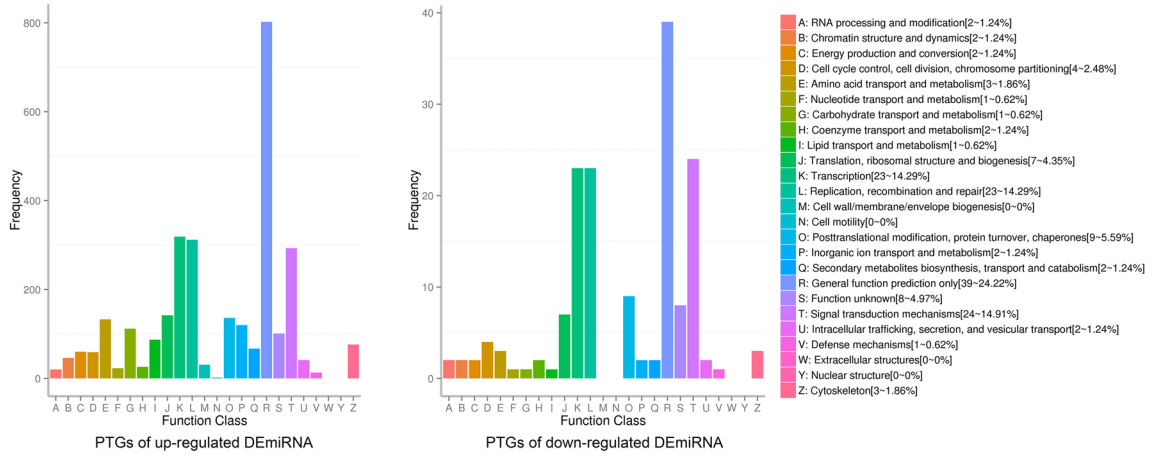
Expression of NSCs and MNs related genes after overexpression or interference with miR-31. To further understand the role of miR-31 in NSCs, we investigated the expression of NSCs and MNs-related genes after overexpression or interference with miR-31. The results (Fig. 7) showed that after interference with the expression of miR-31, the expression of Nestin, a specific marker of NSCs, decreased compared with the control group, while the expression of ChAT, Hb9, Nkx6.1, Nkx6.2, Isl1, Lhx3 and Olig2, which are related to MNs, increased to different degrees. After overexpression of miR-31, the expression of Nestin increased and the expression of MNs-related genes ChAT, Hb9, Nkx6.1, Nkx6.2, Isl1, Lhx3 and Olig2 decreased compared with the control group. (There was a significant difference between the experimental group and the control group, $P \leq 0.05$).

Discussion

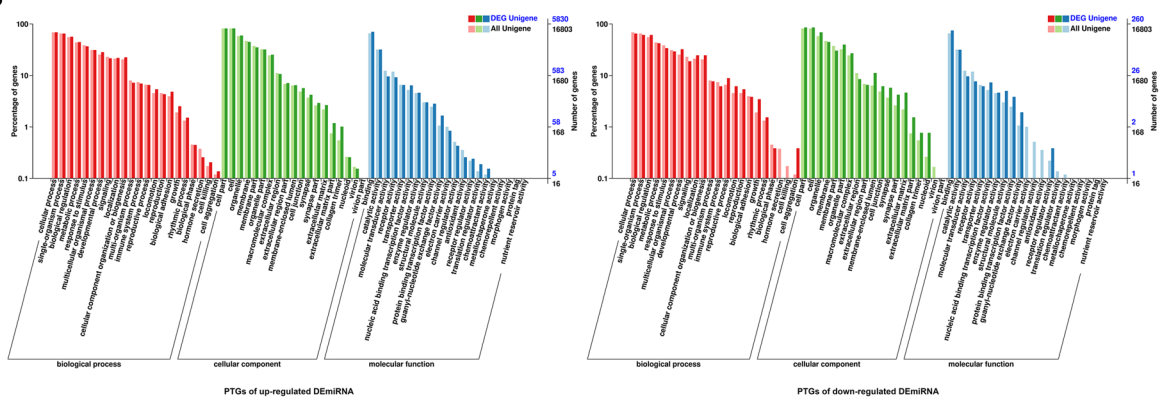
Since a single miRNA can act on hundreds or thousands of target genes, the alternation of its expression not only lead to changes of the corresponding target gene expression level, but also can cause cascade radiation-like changes in the cell through protein-protein interactions, just like the butterfly effect, and ultimately even fundamentally convert the cell fate. In the past few years, many miRNAs have been demonstrated to be involved in the stemness maintenance of NSCs¹⁹ and the differentiation of MNs⁷. But the specific details and cascade of their roles are still unclear. Our previous studies showed that compared with MNs, miR-31 was highly expressed in NSCs, and the difference was very obvious, which suggesting that miR-31 had the opposite regulatory role in NSCs and MNs. It had been confirmed that miR-31 played an important role in the mode of stem cell division^{20,21}. For stem cells, symmetric cell division and asymmetric cell division are their unique renewal and differentiation mechanisms²². Therefore, a detailed study of the role of miR-31 in NSCs will help to understand the stemness maintenance of NSCs and the mechanism of MNs differentiation, so as to better apply NSCs to the treatment of MNs injury. In this study, we conducted some preliminary studies on the role of miR-31 in NSCs by studying the changes of mRNA and miRNA expression profiles in the early stage of miR-31 overexpression.

Although the number of genes with obvious differences in the early stage of NSCs after miR-31 overexpressing was small and there was no obvious trend of GO and KEGG pathway clustering, the distribution of these DEGs in COG, GO and KEGG showed that the up-regulated DEGs were mainly attributed to localization (GO) and metabolism (KEGG), while the down-regulated DEGs were mainly attributed to inorganic ion transport and metabolism (COG), biological adhesion (GO) and cell adhesion molecules (KEGG). These indicate that the direct effect of overexpressing miR-31 on NSCs is also associated with these GO categories and KEGG pathways. Since the role of miRNAs is mainly to inhibit the expression of target genes, down-regulated DEGs were the main objects of our analysis. Studies have shown that Bsn, namely Bassoon, was a presynaptic marker²³; En2

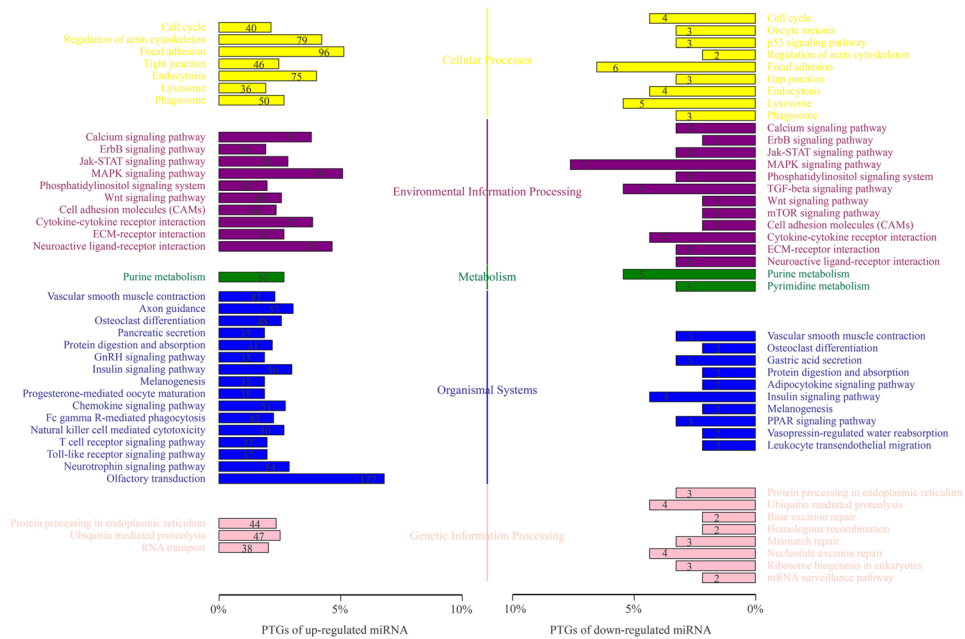
A



B



C



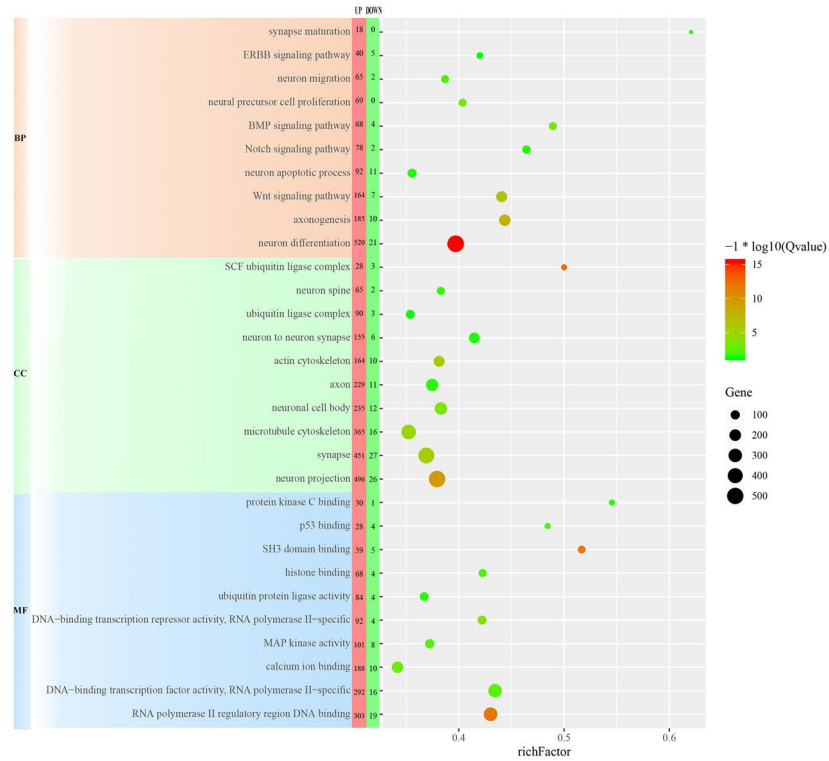
◀ **Figure 3.** Annotated statistical chart of PTGs. **(A)** Statistical map of COG annotation classification of PTGs showed that in addition to the general function prediction only of COG classification, no matter the up or down regulated DEmiRNAs, distribution of PTGs were mostly in transcription, replication, recombination and repair, signal transduction mechanisms. **(B)** Statistical map of GO annotation classification of PTGs showed that the main differences in the up-regulated DEmiRNAs PTGs were mainly distributed in the biological adhesion, growth and rhythmic process of the BP cluster, the synapse, extracellular matrix part and collagen trimer of the CC cluster, and the nucleic acid binding transcription factor activity, guanyl-nucleotide exchange factor activity, and the protein binding transcription factor activity of the MF cluster. The down-regulated DEmiRNAs PTGs were mainly distributed in the reproductive process, growth and cell aggregation of the BP cluster, the nucleoid, extracellular matrix part and synapse part of the CC cluster, and the receptor regulator activity, guanyl-nucleotide exchange factor activity, and structural molecule activity of the MF cluster. **(C)** KEGG Classification Map of PTGs showed the PTGs of the up-regulated DEGs were mainly concentrated in the focal adhesion (Cellular processes), MAPK signaling pathway (Environmental information processing), purine metabolism (Metabolism), axon guidance (Organismal systems) and ubiquitin mediated proteolysis (Genetic information processing). The PTGs of the down-regulated DEGs were mainly concentrated in focal adhesion (Cellular processes) and MAPK signaling pathway (Environmental information processing), purine metabolism (Metabolism), insulin signaling pathway (Organismal systems) and ubiquitin mediated proteolysis (Genetic information processing).

gene encodes a transcription factor containing homeobox, which participated in the development of embryonic midbrain-hindbrain and could promote the differentiation of NSCs into GABAergic neurons²⁴; Mag played an important role in neurite outgrowth²⁵; and *smpd3* was mainly expressed in neurons of the central nervous system²⁶. These genes were closely related to the differentiation of NSCs, and their expression declines after miR-31 overexpression. These suggested that miR-31 overexpression could inhibit the differentiation of NSCs. PKD2L1, which belong to the up-regulated DEGs, was a transient receptor potential channel, mainly expressed in spinal cerebrospinal fluid-contacting neurons²⁷. The NSCs we selected in this experiment were obtained from the embryonic spinal cord. In the spinal cord, NSCs resided in the ependymal region around the central canal²⁸, which were in close contact with circulating cerebrospinal fluid, and migrated out of this region only when differentiated. The increased expression of PKD2L1 suggested that miR-31 overexpression could further maintain the location of NSCs in the spinal cord to adapt to their niche, and also hinted that miR-31 had a certain effect on the stemness maintenance of NSCs.

Our study found that the number of DEmiRNAs in early stage of NSCs after miR-31 overexpressing was much more than DEGs, which suggested that the direct role of miR-31 in NSCs was mainly achieved by changing the expression of numerous miRNAs, which in turn played a regulatory role against their own target genes, thereby expanding the regulatory effect of miR-31. GO and KEGG cluster analysis for PTGs showed that they were mainly enriched in neuron differentiation and RNA polymerase II regulatory region DNA binding (GO), metabolic pathways and PI3K-Akt signaling pathway (KEGG). These GO categories and KEGG pathways are closely related to the proliferation and differentiation of NSCs. Studies have shown that miR-106b could promote the renewal of NSCs and inhibit their differentiation²⁹; miR-130a and miR-138 could inhibit axon growth and regeneration^{30,31}; miR-20b overexpression could downregulate the expressions of Map2 and Tubb3 (well-known neuronal markers)³². In this study, we observed that these miRNAs were upregulated in the early stage of NSCs after miR-31 overexpression. It has also been confirmed that let-7a was mainly involved in neuronal differentiation³³; let-7d could reduce the proliferation of NSCs and promote neuronal differentiation and migration when overexpressed in vivo³⁴; and overexpression of miR-221 could induce neuronal differentiation of PC12 cells³⁵. These miRNAs were all among the down-regulated DEmiRNAs in our study. Taken together, these findings suggest that in the early stage of NSCs after miR-31 overexpression, the expression of miRNA that can promote its proliferation and renewal and inhibit differentiation is up-regulated in NSCs, while the expression of miRNA that have the opposite effect is down-regulated, further suggesting that the main role of miR-31 on NSCs is to maintain stemness and inhibit their differentiation.

Through the study of interaction analysis between DEGs, DEmiRNAs and their VTGs and PTGs, we observed that there were complex regulatory and interaction networks between them. Among these interacting genes, a large number of DEmiRNAs showed negative regulatory effect on more than half of the DEGs. The up-regulated DEGs were mainly related to the stemness maintenance genes of NSCs, among them, Sox2 and Nup153 could maintain NSCs stemness^{36,37}; Bmi-1 and Pou5f1 played a role in NSCs self-renewal and proliferation^{38,39}; Mcm2 was one of NSCs markers⁴⁰. The down-regulated DEGs mainly interacts with NSCs differentiation related genes and MNS differentiation related genes, in which GFAP was the marker of glial cells, Tubb3 was the marker of neurons, Sox10 could direct NSCs to oligodendrocyte lines⁴¹; Mecp2 was mainly expressed in neurons⁴²; Shh, Ptc1, CXCR4, Nkx2-2, Olig1, Irx3, Crebbp, Tle1, Pou3f4, Foxg1, Ep300, Gdnf, Sv2a, Olig2, Slc18a3, Calca and Rbfox3 was related to motor neuron differentiation⁴³. Above all, these genes interacting with DEGs were closely related to the proliferation and differentiation of NSCs, and most of these genes were also VTGs of DEmiRNAs. In addition, the further study also found that Dicer1, an important gene of miRNA biogenesis, and ago3, a related gene that played an inhibitory role, were the VTGs of miR-31-5p, suggesting that the change of miRNA expression profile in early stage of NSCs after miR-31 overexpressed might be related to this. At the same time, in addition to the fact that some DEGs were VTGs or PTGs of miR-31-5p, there were interaction networks between DEGs and some VTGs of miR-31. Among them, Stat3 was an important transcription factor regulating the expression of Gfap⁴⁴; Sod2 played an important role in the regulation of cell cycle⁴⁵; and Gsk3b could be used as a signal "node" to coordinate multiple key signaling pathways in NSCs^{46,47}. Based on the above analysis, we conclude that the action mechanism of miR-31 in early stage of NSCs after overexpression was not

A



B

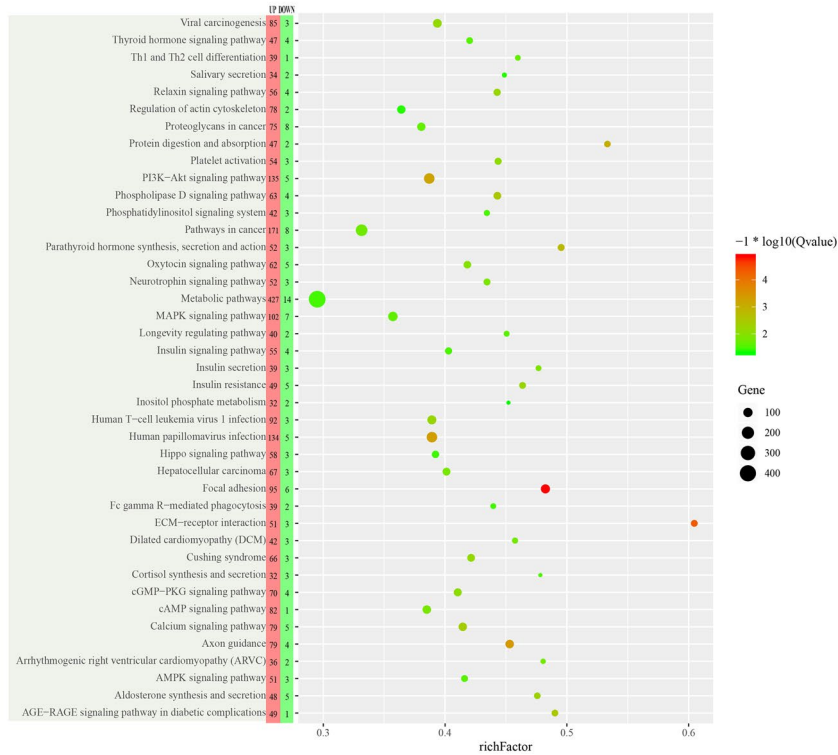


Figure 4. The results of GO and KEGG enrichment analysis of PTGs. (A) The GO categories associated with nervous system from GO enrichment analysis of PTGs showed that PTGs were mainly grouped into the neuron differentiation, axonogenesis and Wnt signaling pathway of the BP cluster, the neuron projection and synapse of the CC cluster, and RNA polymerase II regulatory region DNA binding of the MF cluster. (B) The results of KEGG enrichment analysis of PTGs showed the significantly enriched KEGG pathways of PTGs were metabolic pathways, focal adhesion, PI3K-Akt signaling pathway.

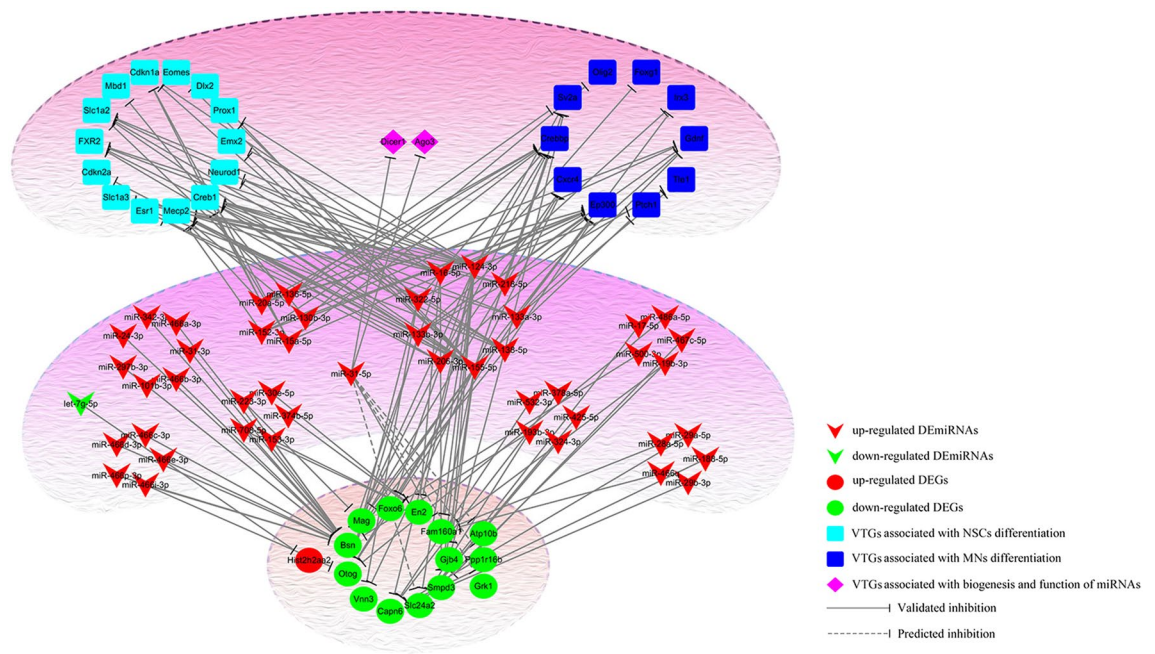


Figure 5. The regulatory relationship between DEmiRNAs, DEGs and VTGs showed that there were complex regulatory relationships between them. Among them, 43 DEmiRNAs and 12 DEGs had opposite expression patterns, suggesting a regulatory relationship between them. VTGs of DEmiRNAs with opposite expression patterns to DEGs revealed that about 10 VTGs were MNs differentiation-related genes and 14 VTGs were NSCs differentiation-related genes.

only to inhibit its target gene, but also to expand its action scope by regulating the expression of other miRNAs, ultimately increasing the expression of genes or miRNAs related to NSCs stemness maintenance, and inhibiting NSCs, especially MNs, differentiation-related genes, through protein interaction and DEmiRNAs, to maintain the undifferentiated state of NSCs. All these indicated that miR-31 is a stemness maintenance gene of NSCs and has a negative regulatory role in the differentiation of NSCs into MNs. After interfering with the expression of miR-31 in spinal cord-derived NSCs, we further found that the expression levels of ChAT, Hb9, Nkx6.1, Nkx6.2, Isl1, Lhx3, and Olig2 genes increased in different degrees. At present, previous studies have shown that Olig2, Nkx6.1 and Nkx6.2 are downstream class II transcription factors of Shh. The mutual inhibition between Olig2, Nkx6.1, Nkx6.2 and Irx3, Dbx2, Dbx1 can ultimately determine the boundary between pMN and other domains in the spinal cord^{43,48,49}. Hb9 (Mnx1)⁵⁰, Isl1⁵¹ and ChAT⁵² are markers of MNs; Lhx3 can bind to Isl1 to form a complex, which plays an important role in the generation of specific motor neurons⁵³. Meanwhile, when miR-31 was overexpressed in NSCs, the expression of these important MNs related genes decreased, while the expression of Nestin (a NSCs marker) increased, which was opposite to that when interfering with miR-31. These suggest that interference with miR-31 expression can induce NSCs to produce MNs like cells, while overexpression of miR-31 can maintain NSCs stemness. This further confirms our above conjecture about the role of miR-31 in NSCs from the above sequencing results. This study deepens our understanding of the role of miR-31 in NSCs, provides an effective candidate target for effectively inducing the differentiation of NSCs into MNs, and lays a foundation for us to effectively apply NSCs to the clinical treatment of motor neuron diseases.

Methods

Ethics statement. All animal procedures were performed according to guidelines developed by the China Council on Animal Care and protocols were approved by the Animal Care and Use Committee of Shanxi Province, China. The permit numbers are SCXK2009-0001.

Sample preparation and experimental grouping. According to our previous study¹³, we cultured spinal cord-derived NSCs obtained from Balb/c mice embryos on days 16. Spinal cords were mechanically dissected using sterile technique under a dissecting microscope. Discarded the pia mater spinalis, and triturated spinal cord gently with a pipette to dissociate cells. Centrifuged dispersed cells at 337 × g for 5 min to get the cells' pellet. Resuspended the pellet in the medium consisting of DMEM/F12, 2% B27 supplement with 20 ng/ml basic fibroblast growth factor (bFGF, R&D Systems) and 20 ng/ml epidermal growth factor (EGF, R&D Systems) at 37 °C in 5% CO₂.

After cultured 14 days, NSCs were plated in 6-well culture plates, and then divided into two groups, one was miR-31 overexpression group which treated with miR-31 mimics (Thermo Fisher, MC10653); another group was miR-31 overexpression control group which treated with negative control of miRNA mimic (Thermo Fisher, 4464058). Each group contained six samples. The experimental process referred to the protocol of products.

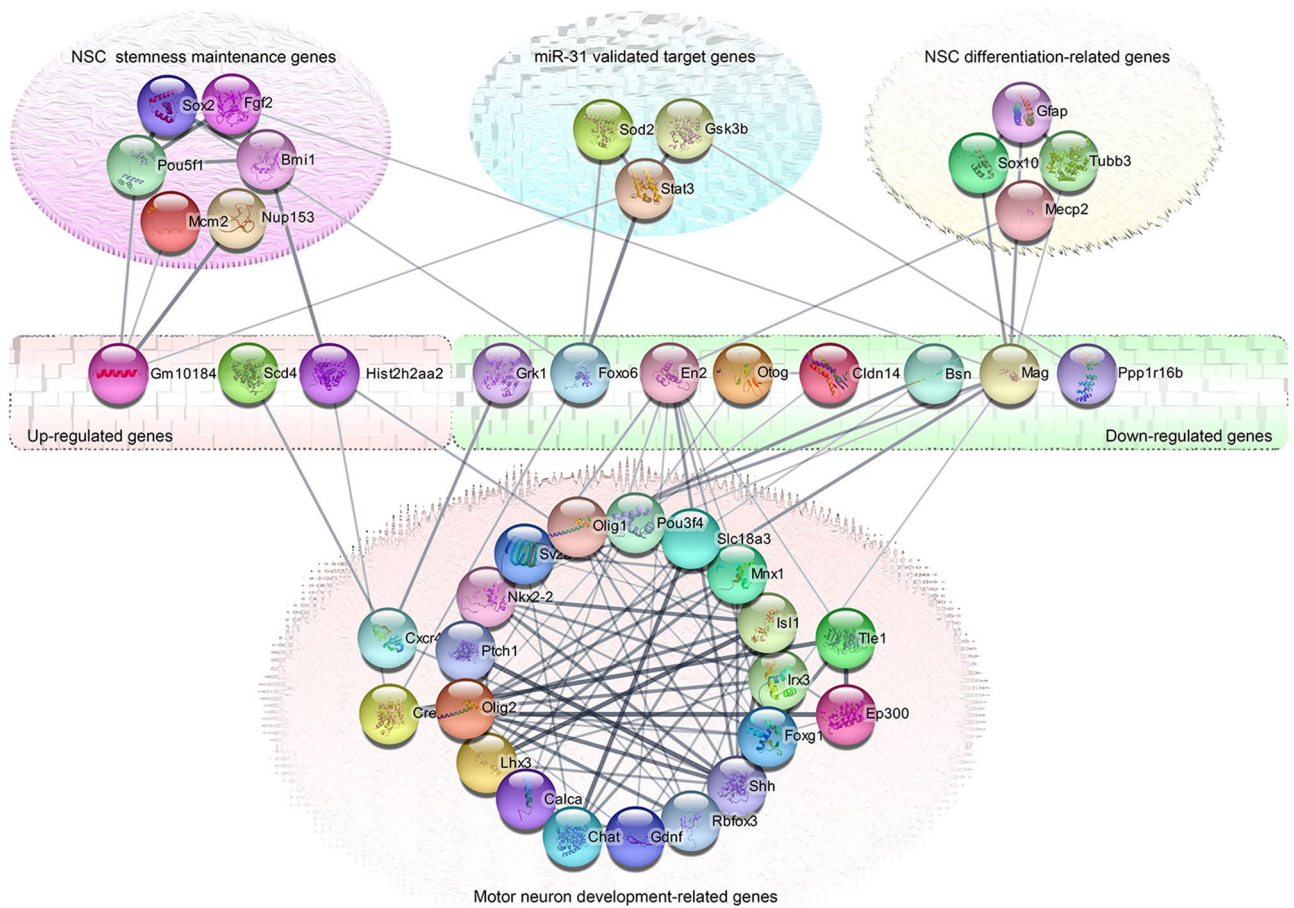


Figure 6. The PPI network between DEGs and VTGs showed that the DEGs could established complex interaction networks through some intermediate node genes. Most of these intermediate node genes were not only key genes related to MNs differentiation, NSCs differentiation or stemness maintenance, but also VTGs of DE miRNAs, especially some of them were the VTGs of miR-31.

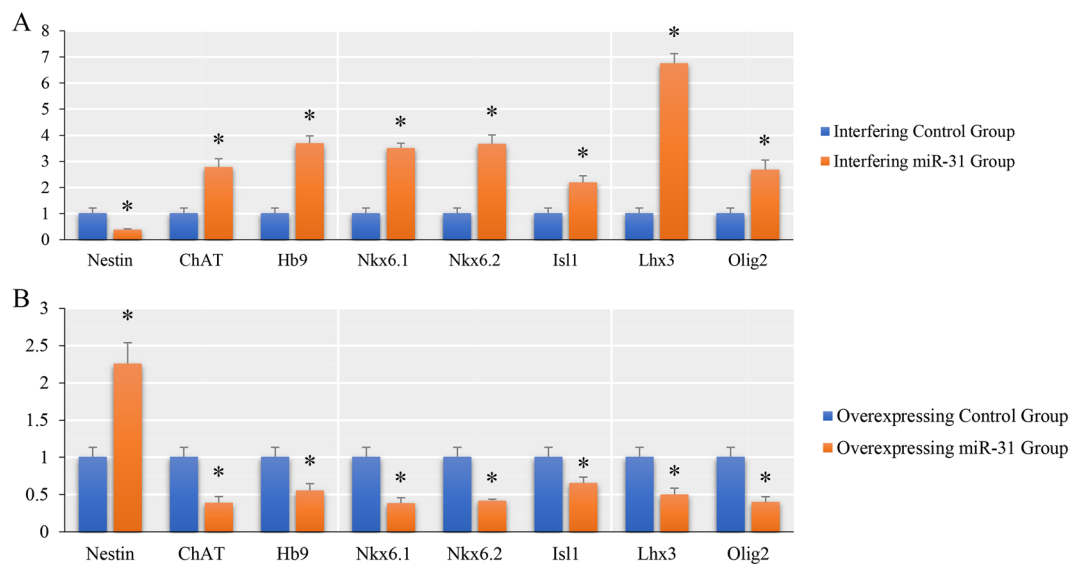


Figure 7. The q-PCR results of NSCs and MNs related genes after overexpression or interference with miR-31. (A) The q-PCR results of NSCs and MNs related genes after miR-31 overexpression (*indicates $p \leq 0.05$, compared with its control group.). (B) The q-PCR results of NSCs and MNs related genes after inhibition of miR-31 (*indicates $p \leq 0.05$, compared with its control group).

RNA extraction and Illumina sequencing. After 3 days of overexpression, the total RNA of miR-31 overexpression group and control group was extracted with Trizol reagent (Thermo Fisher), purified with anhydrous ethanol and treated with DNase. Randomly selected three samples in each group, mixed the total RNA of three samples in the same group, and obtained the final miR-31 overexpression group and its control group, which were finally used for sequencing analysis. Using Agilent Bioanalyzer 2100 (Agilent Technologies) to evaluate the integrity/quality of the two groups of total RNA.

Following the methods provided by TruSeq RNA Sample Preparation V2 Guideline (Illumina) and TruSeq Small RNA Sample Preparation Kit Version 2 (Illumina), constructed the mRNA and miRNA sequencing libraries, respectively. Sequencing analysis was performed respectively by using a HiSeq 2500 (Illumina) at Beijing Biomarker Technologies CO., LTD (Beijing, China).

Analyses of RNA-Seq data and miRNA-Seq data. For the sequencing results of mRNA, FPKM (Fragments Per Kilobase of transcript per Million fragments mapped)⁵⁴ was used as a measure of gene expression level, and EBSeq⁵⁵ was used for differential expression analysis. For the results of miRNA sequencing, TPM algorithm⁵⁶ was used to calculate the amount of miRNA expression in the samples, and IDEG6⁵⁷ was used to analyze the differential expression. Both of them used Benjamini–Hochberg correction method to correct the significant p-value obtained from the original hypothesis test, and finally used False Discovery Rate (FDR) and Fold Change (FC), the ratio of expression between the two groups, as the key indicators for screening differentially expressed genes and miRNAs. Differentially expressed genes (DEGs) and differentially expressed miRNAs (DEmiRNAs) between the two samples were obtained using $\log_2(\text{FC}) \geq 1$ and $\text{FDR} \leq 0.01$ as screening criteria, respectively.

q-PCR validation of DEGs and DEmiRNAs. Analyzed the relative expression levels of DEGs and DEmiRNAs in the remaining samples of the miR-31 overexpression group and its control group by real-time quantitative PCR. The primer information of the analyzed genes and miRNAs are listed in Supplementary Tables S1 and S2. The $2^{-\Delta\Delta C_t}$ method was used to analyze the relative gene expression level, and the Student's t-test was used to analyze the expression difference between the two groups. For both mRNAs and miRNAs, a $p \leq 0.05$ after the Student's t-test was considered statistically significant.

Identification of PTGs and experimentally validated target genes (VTGs) of DEmiRNAs. The PTGs of DEmiRNAs were predicted by miRanda⁵⁸ and RNAhybrid⁵⁹, and the VTGs of DEmiRNAs were identified by DIANA-TarBase v.8⁶⁰ and miRWalk⁶¹.

Functional annotation and enrichment analysis. The DEGs and PTGs were compared with NR⁶², Swiss-Prot⁶³, COG⁶⁴, GO^{65,66}, KEGG^{67–69} databases to obtain annotation information of target genes, and cluster analysis was performed.

miRNA-mRNA regulatory network and protein–protein interactions (PPI) network. The miRNA-mRNA regulatory network is mainly constructed based on the opposite expression patterns of DEmiRNAs and DEGs, and PPI between DEGs and VTGs are constructed using STRING database (<https://string-db.org/>).

Effects of overexpression or interference with miR-31 on the expression of NSCs and MNs related genes. NSCs were divided into four groups: interfering miR-31 group, interfering control group, overexpressing miR-31 group and overexpressing control group. mmu-miR-31-5p mimic, mmu-miR-31-5p mimics Negative Control, mmu-miR-31-5p inhibitor and mmu-miR-31-5p inhibitor Negative Control (Supplementary Table S6) were transfected by Lipofectamine RNAiMAX (Thermo Fisher) according to group. After incubation at 37 °C and 5% CO₂ for 7 days, the expression of Nestin, ChAT, Hb9, Nkx6.1, Nkx6.2, Isl1, Lhx3 and Olig2 in each group was detected by q-PCR (the information of primers is listed in Supplementary Table S7).

Data availability

The datasets generated during and/or analyzed during the current study are available by request.

Received: 29 March 2020; Accepted: 28 August 2020

Published online: 16 October 2020

References

- Grumbles, R. M. & Thomas, C. K. Motoneuron death after human spinal cord injury. *J. Neurotrauma* **34**, 581–590. <https://doi.org/10.1089/neu.2015.4374> (2017).
- Ahuja, C. S. *et al.* Traumatic spinal cord injury. *Nat. Rev. Dis. Primers* **3**, 17018. <https://doi.org/10.1038/nrdp.2017.18> (2017).
- Koprivica, V. *et al.* EGFR activation mediates inhibition of axon regeneration by myelin and chondroitin sulfate proteoglycans. *Science* **310**, 106–110. <https://doi.org/10.1126/science.1115462> (2005).
- Zheng, Y., Mao, Y. R., Yuan, T. F., Xu, D. S. & Cheng, L. M. Multimodal treatment for spinal cord injury: A sword of neuroregeneration upon neuromodulation. *Neural. Regen. Res.* **15**, 1437–1450. <https://doi.org/10.4103/1673-5374.274332> (2020).
- Kameda, T., Imamura, T. & Nakashima, K. Epigenetic regulation of neural stem cell differentiation towards spinal cord regeneration. *Cell Tissue Res.* **371**, 189–199. <https://doi.org/10.1007/s00441-017-2656-2> (2018).
- Cheng, L. C., Tavazoie, M. & Doetsch, F. Stem cells: From epigenetics to microRNAs. *Neuron* **46**, 363–367. <https://doi.org/10.1016/j.neuron.2005.04.027> (2005).

7. Chen, T. H. & Chen, J. A. Multifaceted roles of microRNAs: From motor neuron generation in embryos to degeneration in spinal muscular atrophy. *Elife*. <https://doi.org/10.7554/eLife.50848> (2019).
8. Hammond, S. M. An overview of microRNAs. *Adv. Drug Deliv. Rev.* **87**, 3–14. <https://doi.org/10.1016/j.addr.2015.05.001> (2015).
9. Mens, M. M. J. & Ghanbari, M. Cell cycle regulation of stem cells by microRNAs. *Stem Cell Rev. Rep.* **14**, 309–322. <https://doi.org/10.1007/s12015-018-9808-y> (2018).
10. Zhao, C. *et al.* MicroRNA let-7b regulates neural stem cell proliferation and differentiation by targeting nuclear receptor TLX signaling. *Proc. Natl. Acad. Sci. USA* **107**, 1876–1881. <https://doi.org/10.1073/pnas.0908750107> (2010).
11. Zhao, C., Sun, G., Li, S. & Shi, Y. A feedback regulatory loop involving microRNA-9 and nuclear receptor TLX in neural stem cell fate determination. *Nat. Struct. Mol. Biol.* **16**, 365–371. <https://doi.org/10.1038/nsmb.1576> (2009).
12. Kim, J. *et al.* A MicroRNA feedback circuit in midbrain dopamine neurons. *Science* **317**, 1220–1224. <https://doi.org/10.1126/science.1140481> (2007).
13. Wei, H. *et al.* Comparative profiling of microRNA expression between neural stem cells and motor neurons in embryonic spinal cord in rat. *Int. J. Dev. Neurosci.* **28**, 545–551. <https://doi.org/10.1016/j.ijdevneu.2010.04.007> (2010).
14. Lv, C. *et al.* MiR-31 promotes mammary stem cell expansion and breast tumorigenesis by suppressing Wnt signaling antagonists. *Nat. Commun.* **8**, 1036. <https://doi.org/10.1038/s41467-017-01059-5> (2017).
15. Tian, Y. *et al.* Stress responsive miR-31 is a major modulator of mouse intestinal stem cells during regeneration and tumorigenesis. *Elife*. <https://doi.org/10.7554/eLife.29538> (2017).
16. Kim, D. *et al.* TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36. <https://doi.org/10.1186/gb-2013-14-4-r36> (2013).
17. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25. <https://doi.org/10.1186/gb-2009-10-3-r25> (2009).
18. Friedlander, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **40**, 37–52. <https://doi.org/10.1093/nar/gkr688> (2012).
19. Diana, A., Gaido, G. & Murtas, D. MicroRNA signature in human normal and tumoral neural stem cells. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms20174123> (2019).
20. Su, Y. *et al.* Fate decision of satellite cell differentiation and self-renewal by miR-31-IL34 axis. *Cell Death Differ.* **27**, 949–965. <https://doi.org/10.1038/s41418-019-0390-x> (2020).
21. Wang, J. *et al.* Zeb1 regulates the symmetric division of mouse lewis lung carcinoma stem cells through numb mediated by miR-31. *Int. J. Biol. Sci.* **14**, 1399–1410. <https://doi.org/10.7150/ijbs.27446> (2018).
22. Morrison, S. J. & Kimble, J. Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature* **441**, 1068–1074. <https://doi.org/10.1038/nature04956> (2006).
23. Haverkamp, S., Muller, U., Zeilhofer, H. U., Harvey, R. J. & Wässle, H. Diversity of glycine receptors in the mouse retina: Localization of the alpha2 subunit. *J. Comp. Neurol.* **477**, 399–411. <https://doi.org/10.1002/cne.20267> (2004).
24. Boschian, C. *et al.* Impaired neuronal differentiation of neural stem cells lacking the engrailed-2 gene. *Neuroscience* **386**, 137–149. <https://doi.org/10.1016/j.neuroscience.2018.06.032> (2018).
25. Zearfoss, N. R., Farley, B. M. & Ryder, S. P. Post-transcriptional regulation of myelin formation. *Biochim. Biophys. Acta* **1779**, 486–494. <https://doi.org/10.1016/j.bbagr.2008.06.003> (2008).
26. Stoffel, W., Jenke, B., Block, B., Zumbansen, M. & Koebke, J. Neutral sphingomyelinase 2 (smpd3) in the control of postnatal growth and development. *Proc. Natl. Acad. Sci. USA* **102**, 4554–4559. <https://doi.org/10.1073/pnas.0406380102> (2005).
27. Djenoune, L. *et al.* Investigation of spinal cerebrospinal fluid-contacting neurons expressing PKD2L1: Evidence for a conserved system from fish to primates. *Front. Neuroanat.* **8**, 26. <https://doi.org/10.3389/fnana.2014.00026> (2014).
28. Martens, D. J., Seaberg, R. M. & van der Kooy, D. In vivo infusions of exogenous growth factors into the fourth ventricle of the adult mouse brain increase the proliferation of neural progenitors around the fourth ventricle and the central canal of the spinal cord. *Eur. J. Neurosci.* **16**, 1045–1057. <https://doi.org/10.1046/j.1460-9568.2002.02181.x> (2002).
29. Xia, X. *et al.* miR-106b regulates the proliferation and differentiation of neural stem/progenitor cells through Tp53inp1-Tp53-Cdkn1a axis. *Stem Cell Res. Ther.* **10**, 282. <https://doi.org/10.1186/s13287-019-1387-6> (2019).
30. Zhang, Y. *et al.* MiR-130a regulates neurite outgrowth and dendritic spine density by targeting MeCP2. *Protein Cell* **7**, 489–500. <https://doi.org/10.1007/s13238-016-0272-7> (2016).
31. Liu, C. M. *et al.* MicroRNA-138 and SIRT1 form a mutual negative feedback loop to regulate mammalian axon regeneration. *Genes Dev.* **27**, 1473–1483. <https://doi.org/10.1101/gad.209619.112> (2013).
32. Wang, H., Ban, W., Wang, T., Li, Z. & Dang, X. miR-20b/106a modulate Ngn2 gene expression during neural differentiation of human umbilical cord mesenchymal stem cells. *NeuroReport* **28**, 1225–1231. <https://doi.org/10.1097/WNR.0000000000000900> (2017).
33. Song, J., Cho, K. J., Oh, Y. & Lee, J. E. Let7a involves in neural stem cell differentiation relating with TLX level. *Biochem. Biophys. Res. Commun.* **462**, 396–401. <https://doi.org/10.1016/j.bbrc.2015.05.004> (2015).
34. Zhao, C., Sun, G., Ye, P., Li, S. & Shi, Y. MicroRNA let-7d regulates the TLX/microRNA-9 cascade to control neural cell fate and neurogenesis. *Sci. Rep.* **3**, 1329. <https://doi.org/10.1038/srep01329> (2013).
35. Hamada, N. *et al.* MicroRNA expression profiling of NGF-treated PC12 cells revealed a critical role for miR-221 in neuronal differentiation. *Neurochem. Int.* **60**, 743–750. <https://doi.org/10.1016/j.neuint.2012.03.010> (2012).
36. Graham, V., Khudyakov, J., Ellis, P. & Pevny, L. SOX2 functions to maintain neural progenitor identity. *Neuron* **39**, 749–765. [https://doi.org/10.1016/s0896-6273\(03\)00497-5](https://doi.org/10.1016/s0896-6273(03)00497-5) (2003).
37. Toda, T. *et al.* Nup153 interacts with Sox2 to enable bimodal gene regulation and maintenance of neural progenitor cells. *Cell Stem Cell* **21**, 618–634.e617. <https://doi.org/10.1016/j.stem.2017.08.012> (2017).
38. Yadirgi, G. *et al.* Conditional activation of Bmi1 expression regulates self-renewal, apoptosis, and differentiation of neural stem/progenitor cells in vitro and in vivo. *Stem Cells* **29**, 700–712. <https://doi.org/10.1002/stem.614> (2011).
39. Reeve, R. L., Yammine, S. Z., Morshead, C. M. & van der Kooy, D. Quiescent Oct4(+) neural stem cells (NSCs) repopulate ablated glial fibrillary acidic protein(+) NSCs in the adult mouse brain. *Stem Cells* **35**, 2071–2082. <https://doi.org/10.1002/stem.2662> (2017).
40. Mohan, A., Kandalam, M., Ramkumar, H. L., Gopal, L. & Krishnakumar, S. Stem cell markers: ABCG2 and MCM2 expression in retinoblastoma. *Br. J. Ophthalmol.* **90**, 889–893. <https://doi.org/10.1136/bjo.2005.089219> (2006).
41. Pozniak, C. D. *et al.* Sox10 directs neural stem cells toward the oligodendrocyte lineage by decreasing Suppressor of Fused expression. *Proc. Natl. Acad. Sci. USA* **107**, 21795–21800. <https://doi.org/10.1073/pnas.1016485107> (2010).
42. Jung, B. P., Zhang, G., Ho, W., Francis, J. & Eubanks, J. H. Transient forebrain ischemia alters the mRNA expression of methyl DNA-binding factors in the adult rat hippocampus. *Neuroscience* **115**, 515–524. [https://doi.org/10.1016/s0306-4522\(02\)00383-4](https://doi.org/10.1016/s0306-4522(02)00383-4) (2002).
43. Stifani, N. Motor neurons and the generation of spinal motor neuron diversity. *Front. Cell Neurosci.* **8**, 293. <https://doi.org/10.3389/fncel.2014.00293> (2014).
44. Park, S. Y. & Han, J. S. Phospholipase D1 signaling: Essential roles in neural stem cell differentiation. *J. Mol. Neurosci.* **64**, 333–340. <https://doi.org/10.1007/s12031-018-1042-1> (2018).
45. Sarsour, E. H., Kalen, A. L. & Goswami, P. C. Manganese superoxide dismutase regulates a redox cycle within the cell cycle. *Antioxid. Redox Signal* **20**, 1618–1627. <https://doi.org/10.1089/ars.2013.5303> (2014).

46. Grimes, C. A. & Jope, R. S. The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling. *Prog. Neurobiol.* **65**, 391–426. [https://doi.org/10.1016/s0301-0082\(01\)00011-9](https://doi.org/10.1016/s0301-0082(01)00011-9) (2001).
47. Kim, W. Y. *et al.* GSK-3 is a master regulator of neural progenitor homeostasis. *Nat. Neurosci.* **12**, 1390–1397. <https://doi.org/10.1038/nn.2408> (2009).
48. Novitsch, B. G., Chen, A. I. & Jessell, T. M. Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. *Neuron* **31**, 773–789. [https://doi.org/10.1016/s0896-6273\(01\)00407-x](https://doi.org/10.1016/s0896-6273(01)00407-x) (2001).
49. Sugimori, M. *et al.* Combinatorial actions of patterning and HLH transcription factors in the spatiotemporal control of neurogenesis and gliogenesis in the developing spinal cord. *Development* **134**, 1617–1629. <https://doi.org/10.1242/dev.001255> (2007).
50. Arber, S. *et al.* Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. *Neuron* **23**, 659–674. [https://doi.org/10.1016/s0896-6273\(01\)80026-x](https://doi.org/10.1016/s0896-6273(01)80026-x) (1999).
51. Ericson, J., Thor, S., Edlund, T., Jessell, T. M. & Yamada, T. Early stages of motor neuron differentiation revealed by expression of homeobox gene Islet-1. *Science* **256**, 1555–1560. <https://doi.org/10.1126/science.1350865> (1992).
52. Bagher, Z. *et al.* Differentiation of neural crest stem cells from nasal mucosa into motor neuron-like cells. *J. Chem. Neuroanat.* **92**, 35–40. <https://doi.org/10.1016/j.jchemneu.2018.05.003> (2018).
53. Thaler, J. P., Lee, S. K., Jurata, L. W., Gill, G. N. & Pfaff, S. L. LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein–protein interactions. *Cell* **110**, 237–249. [https://doi.org/10.1016/s0092-8674\(02\)00823-1](https://doi.org/10.1016/s0092-8674(02)00823-1) (2002).
54. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515. <https://doi.org/10.1038/nbt.1621> (2010).
55. Leng, N. *et al.* EBSeq: An empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **29**, 1035–1043. <https://doi.org/10.1093/bioinformatics/btt087> (2013).
56. Fahlgren, N. *et al.* High-throughput sequencing of Arabidopsis microRNAs: Evidence for frequent birth and death of MIRNA genes. *PLoS ONE* **2**, e219. <https://doi.org/10.1371/journal.pone.0000219> (2007).
57. Romualdi, C., Bortoluzzi, S., D'Alessi, F. & Danieli, G. A. IDEG6: A web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol. Genomics* **12**, 159–162. <https://doi.org/10.1152/physiolgenomics.00096.2002> (2003).
58. Betel, D., Koppal, A., Agius, P., Sander, C. & Leslie, C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.* **11**, R90. <https://doi.org/10.1186/gb-2010-11-8-r90> (2010).
59. Kruger, J. & Rehmsmeier, M. RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res.* **34**, W451–454. <https://doi.org/10.1093/nar/gkl243> (2006).
60. Karagkouni, D. *et al.* DIANA-TarBase v8: A decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res.* **46**, D239–D245. <https://doi.org/10.1093/nar/gkx1141> (2018).
61. Sticht, C., De La Torre, C., Parveen, A. & Gretz, N. miRWalk: An online resource for prediction of microRNA binding sites. *PLoS ONE* **13**, e0206239. <https://doi.org/10.1371/journal.pone.0206239> (2018).
62. Deng, Y. Y., Li, J. Q., Wu, S. F., Zhu, Y. P. & He, F. C. Integrated nr database in protein annotation system and its localization. *Comput. Eng.* **32**, 71–72 (2006).
63. Apweiler, R. *et al.* UniProt: The Universal Protein knowledgebase. *Nucleic Acids Res.* **32**, D115–119. <https://doi.org/10.1093/nar/gkh131> (2004).
64. Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36. <https://doi.org/10.1093/nar/28.1.33> (2000).
65. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29. <https://doi.org/10.1038/75556> (2000).
66. The Gene Ontology, C. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338. <https://doi.org/10.1093/nar/gky1055> (2019).
67. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. <https://doi.org/10.1093/nar/28.1.27> (2000).
68. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **47**, D590–D595. <https://doi.org/10.1093/nar/gky962> (2019).
69. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951. <https://doi.org/10.1002/pro.3715> (2019).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant no. 81371384), Key Research and Development Projects of Shanxi Province (Grant no. 201803D31068) and Basic Application Research of Shanxi Province (Grant no. 201801D121212).

Author contributions

P.L. and C.W. conceived, designed and supervised the study. P.L., Y.G., X.L., F.T., F.W., Y.W., B.Z. and R.Z. conducted the experiments. P.L. and C.W. analyzed the data and wrote the paper. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-74541-8>.

Correspondence and requests for materials should be addressed to C.W.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2020