Integrated Analysis of Differentially Expressed Genes and Construction of a Competing Endogenous RNA Network in Human Huntington Neural Progenitor Cells

Xiaoping Tan
Shengjing Hospital of China Medical University https://orcid.org/0000-0002-3256-392X
Shuyan Cong (congsy@sj-hospital.org)
Shengjing Hospital of China Medical University
Yang Liu
Shengjing Hospital of China Medical University
Taiming Zhang
Shengjing Hospital of China Medical University

Research article

Keywords: RNA, lncRNA, Huntington Disease, ceRNA, Bioinformatics

Posted Date: August 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-56137/v1

Version of Record: A version of this preprint was published on February 12th, 2021. See the published version at https://doi.org/10.1186/s12920-021-00894-2.

Integrated analysis of differentially expressed genes and construction of a competing endogenous RNA network in human Huntington neural progenitor cells

Xiaoping Tan[1], Shuyan Cong[1*], Yang Liu[1], Taiming Zhang[1]

Department of Neurology, Shengjing Hospital of China Medical University, Shenyang, Liaoning, PR China.

[*]Corresponding Author:

Shuyan Cong

36 Sanhao Street, Shenyang, Liaoning, 110004, PR China.

Email address: congsy@sj-hospital.org

Abstract

Background

Huntington's disease (HD) is one of the most common polyglutamine disorders, leading to progressive dyskinesia, cognitive impairment, and neuropsychological problems. Besides the dysregulation of many protein-coding genes in HD, previous studies have revealed a variety of non-coding RNAs that are dysregulated in HD, including several long non-coding RNAs (lncRNAs). However, an integrated analysis of differentially expressed (DE) genes based on a competing endogenous RNA (ceRNA) network is still currently lacking.

Results

Here, we have systematically analyzed the gene expression profile data of neural progenitor cells (NPCs) derived from patients with HD and controls (healthy controls and the isogenic controls of HD patient cell lines corrected using CRISPR-Cas9 approach at the *HTT* locus), and we identified 490 DE mRNAs and 94 DE lncRNAs, respectively. Of these, 189 mRNAs and 20 lncRNAs were applied to create a ceRNA network. To learn more about the possible functions of lncRNAs in the ceRNA regulatory network in HD, we conducted a functional analysis of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) and established a protein-protein interaction (PPI) network for mRNAs interacting with these lncRNAs. It is suggested that the function of DE lncRNAs mainly correlated with transcriptional regulation demonstrated by GO analysis. Also, KEGG enrichment analysis showed these lncRNAs were involved in tumor necrosis factor, calcium, Wnt, and NF-kappa B signaling pathways. Interestingly, the PPI network revealed that a variety of transcription factors in the ceRNA network interacted with each other, suggesting such lncRNAs may regulate transcription in HD by controlling the expression of such protein-coding genes, especially transcription factors.

Conclusions

Our research provides new clues for uncovering the mechanism of lncRNAs in HD and can be used as the focus for further investigation.

Key words RNA, lncRNA, Huntington Disease, ceRNA, Bioinformatics

Background

Huntington's disease (HD) is an autosomal dominant-inherited polyglutamine disorder. Typical features include progressive movement disorder, neuropsychiatric problems, and cognitive impairment. Abnormal amplification of cytosine-adenine-guanine (CAG) repeats of the first exon in the huntingtin (*HTT*) gene encoding an abundantly-expressed 3144 amino acid protein, is the root cause of this fatal disease (MacDonald et al., 1993). The DNA of healthy persons contains a region of less than 36 CAG repeats that encodes a polyglutamine (polyQ) tract in the huntingtin gene. An extension in the length of the polyQ tract at N-terminal of the mutant huntingtin (*mutHTT*), encoded by an abnormal *HTT* gene containing the expansion of CAG repeats (>36 in length), changes the conformation of *mutHTT* and leads to intracellular protein aggregates (Chanda et al., 2018). Carriers with longer CAG repeats of *HTT* exhibit different severities of HD performance, depending on the unusual length of CAG repeats (Stine et al., 1993). How this pure mutant of the ubiquitously expressed protein leads to specific neurodegeneration is, however, unclear.

Over the past few decades, along with identifying the *HTT* gene, considerable progress has been made in learning the mechanism of neurodegeneration caused by *mutHTT*, including increased excitotoxicity injury in neurons through the deficiency of wild-type *HTT* (Leavitt et al., 2006), nutritional support disorders from neurons (Zuccato et al., 2001), the appearance of inclusion bodies derived from aberrant protein accumulation and the influence of polyglutamine length on the extent of *mutHTT* aggregates (DiFiglia et al., 1997), mitochondrial impairment (Panov et al., 2002), and defective axonal transport in HD neurons (Orr et al., 2008; Trushina et al., 2004). Except for these cellular dysfunctions, *mutHTT* also caused complex transcriptional and post-transcriptional changes in mouse models and the brain of patients with HD, particularly in the striatum, followed by the cortex and cerebellum, and the severity of transcriptional changes reflected the seriousness of neurodegeneration (Hodges et al., 2006; Kuhn et al., 2007). Dysregulation of the RE1 silencing transcription factor (REST) (Johnson & Buckley, 2009) and its target gene regulatory networks, including microRNA (miRNAs) mir-132(Wayman et al., 2008), mir-9(Packer, Xing, Harper, Jones, & Davidson, 2008), mir-124(Yoo et al., 2011)] and possibly other non-coding RNAs (ncRNAs), were particularly attractive as an explanation of the neurodegeneration caused by *mutHTT* in related studies.

Recently, increasing studies have investigated the role of transcriptional and post-transcriptional dysregulation in the pathogenesis of HD, including a variety of both protein-coding and ncRNAs. In particularly, the dysregulation of protein-coding RNA, REST, in HD has become attractive. As a subtype of ncRNAs, long non-coding RNAs (lncRNAs) are a class of transcripts containing more than 200 bases that do not have a functional open reading frame, and play crucial biological roles in epigenetics, and transcriptional regulation (Guttman et al., 2009). In addition to being involved in normal physiological processes, previous studies have shown that lncRNAs take part in the pathogenesis of HD by regulating the expression of protein-coding genes via both cis- and trans-action pathways (Johnson, 2012). They also interact with varying repressive chromatin regulatory complexes (PRC2, RCOR1, and SMCX), and affect transcription factor (TF) function (Khalil et al., 2009; Tsai et al., 2010). MicroRNAs are endogenous ncRNAs, approximately 22 nucleotides in length, that are involved in the negative regulation of mRNA at the post-transcriptional level by targeting a repressive protein complex (RISC) and pairing to the 3’-untranslated region of mRNA to directly repress post-transcriptional translation (Bartel, 2009). Increasing evidence has illustrated how several neuronal-specific microRNAs were dysregulated in HD disease (Johnson & Buckley, 2009), some of which were targeted by REST. In 2011, the competing endogenous RNA (ceRNA) hypothesis was proposed, In 2011, the competition

endogenous RNA (ceRNA) hypothesis was proposed that emphasized that mRNAs, transcribed pseudogenes, and lncRNAs could interact with each other through competitive binding to microRNA response elements (MREs) (Salmena, Poliseno, Tay, Kats, & Pandolfi, 2011). It is therefore necessary to establish a lncRNA-miRNA-mRNA-ceRNA regulatory network in HD.

Increasing evidence has shown that lncRNAs, as well as ceRNAs, were related to HD (Johnson, 2012). Clarifying the changes of the lncRNA regulatory network in HD neurons may perhaps help to identify and manage this intractable neurodegenerative disease. In this study, we acquired lncRNAs and mRNA expression profile data of neural progenitor cells (NPCs) differentiated from HD (including corrected HD by CRISPR-Cas9 approach) and healthy human-induced pluripotent stem cells (iPSCs) in the Gene Expression Omnibus (GEO) database. We then constructed a ceRNA regulatory network using differentially expressed genes (DEGs) that may help us to shed light on the mechanism of transcriptional changes in HD neurons.

Results

Probe re-annotation

A total of 38,640 probes were re-annotated into corresponding gene symbols, and after deduplication with a mean method and data normalization, a final set of 2,743 lncRNAs and 18,015 mRNAs was retained in the matrix file for subsequent analysis.

Identification of differentially expressed genes

Compared with the control group (CAG33 and HD-C#1,2 cell lines), a total of 94 DE lncRNAs (49 up-regulated, 45 down-regulated) and 490 DE mRNAs (229 up-regulated, 261 down-regulated) were found in the HD cell lines with 180 CAG repeats by differentially expression analysis. The strenghth of differential gene expression was shown in the form of volcano plots, and the top 20 up/down-regulated DE lncRNAs and mRNAs were presented by heatmaps, respectively (Fig.1).

Function analysis of DE mRNAs

We further investigated the biological functions of up- and down-regulated protein-coding mRNAs using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. As shown in Fig. 2, the top 20 biological processes (BPs) and molecular functions (MF) of protein-coding RNAs were exhibited as dot plots. We found that the significant changes in BPs of up-regulated mRNAs included the regulation of GTPase activity, a response to lipopolysaccharide, second messenger-mediated signaling, immune responses, spinal cord, and neuro development and differentiation. For down-regulated mRNAs, BPs were the T cell activation, positive regulation of nervous system development and cell adhesion, the response to BMP, the immune response, T cell cytokine production, and neuron maturation. Similarly, the significant MF changes of up-regulated mRNAs included calmodulin binding, transcription factor activity, RNA polymerase II distal enhancer/proximal promoter sequence-specific DNA binding, iron ion binding, phospholipase C activity, and transmembrane transporter activity. While receptor ligand activity, channel activity, serine-type peptidase/hydrolase activity, and transcriptional activator activity were related to the MF of down-regulated mRNAs.

To identify which pathways were involved in the DEGs, all DE mRNAs were matched to pathways in the KEGG database, and 34 significantly enriched pathways were identified ($p < 0.05$; Fig. 3). When

functionally grouped based on shared genes between terms with a kappa score level (⩾0.4), 12 functional groups were contained in the final analysis, such as the calcium signaling pathway, TNF signaling pathway, cell adhesion molecules, and NF-kappa B signaling pathway.

LncRNA-associated ceRNA network

As miRNA sponges, lncRNAs can isolate and bind miRNAs to regulate mRNA expression. By using the predictive databases of starBase, miRcode, miRDB, TargetScanHuman, and miRTarBase, we identified 48 common miRNAs that both targeted by 20 differentially expressed lncRNAs and 189 mRNAs. According to these data, we constructed a lncRNA-miRNA-mRNA competitive regulation network (Fig. 4) and the top 10 genes with the highest connectivity were listed (Table 1).

Construction of PPI network and GO/KEGG analysis of the mRNAs interacting with lncRNAs in the ceRNA network

To better identify critical genes and understand the biological function of mRNAs in the ceRNA network, we established a protein-protein interaction (PPI) network in STRING software (Szklarczyk et al., 2019) with interaction scores of >0.4 (Fig. 5) and performed GO enrichment analysis with the background of a human whole-genome (Table 2 and Table 3). As the PPI network showed, the primary hub nodes were JUN, CXCR4, BCL6, CTGF, NOS3, VEGFC, ITPKB, PTGS2, RAC3, and KAT2B. Through an enrichment analysis of their protein interacting partners, we observed that the most significant biological processes defined by GO were the response to stimulus (GO:0048583), regulation of signaling (GO:0023051) and cell communication (GO:0010646), and signal transduction (GO:0009966) and positive regulation of biological process (GO:0048518). About half of the significant molecular functions were related to transcriptional regulation (GO:0000976, GO:0000977, GO:0003700, GO00140110).

Further, KEGG enrichment analysis based on ClueGO/CluePedia with $p < 0.05$ and a kappa score ⩾0.4 revealed that the top five significant enrichment pathways were axon guidance, and the signaling pathways of tumor necrosis factor, calcium, Wnt, and NF-kappa B (Fig. 6).

Discussion

Emerging evidence over decades has revealed that ncRNAs play a vital role in neuronal development and function by modulating gene expression at different levels of the eukaryotic nucleus and cytoplasm. This includes the biological functions of DNA replication, transcriptional and post-transcriptional regulatory mechanisms, epigenetic functions and regulation of both mRNA and protein stability, and regulation of translation and miRNA activity (Iyengar et al., 2014; Khorkova, Hsiao, & Wahlestedt, 2015; Riva, Ratti, & Venturin, 2016). It has been suggested that lncRNAs are important, unexplored molecules of neurological disease(Mehler & Mattick, 2007), involving Parkinson's Disease (PD), Alzheimer's Disease (AD), fragile X syndrome, spinocerebellar ataxia and HD. For example, the lncRNA, BACE1-AS, was described as positively regulating BACE1 protein positively and promoting Aβ42 synthesis by increasing the stability of BACE1 mRNA and the binding of miR-485-5p in AD (Faghihi et al., 2008; Faghihi et al., 2010). In PD, the lncRNA UCHL1-AS1, plays a role by directly promoting the translation of UCHL1 protein, leading to disturbance of the ubiquitin-proteasome system (Carrieri et al., 2012). Therefore, it is not surprising that lncRNA dysregulation may play a pathogenetic role in HD. In the past few years, several lncRNAs that control the expression of HD-related genes have

been identified, including HTT-AS (Chung, Rudnicki, Yu, & Margolis, 2011), BDNF-AS (Zuccato & Cattaneo, 2007), ABHD11-AS1 (Francelle et al., 2015), neural human accelerated region 1 (HAR1) (Johnson et al., 2010), TUNA (Lin et al., 2014), NEAT1 (Sunwoo et al., 2017), and misregulated lncRNAs such as TUG1, LINC000341, RPS20P22, MEG3, DGCR5 and LINC000342, as revealed by the analysis of microarray data (Johnson, 2012). However, the precise functions of most HD-related lncRNAs have not been elucidated.

LncRNAs can act as signals, decoys, guides, and scaffolds to regulate the expression of protein-coding genes using molecular mechanisms that are probably distinct (Wang & Chang, 2011). Studies have revealed some lncRNAs acting as "miRNA sponges" that can compete with the miRNA target genes to sponge and bind MREs. This alleviates miRNA-mediated target mRNA inhibition and subsequently promotes miRNA degradation through dsRNA formation (Ebert & Sharp, 2010). The ceRNA regulatory network is one of the essential ways for ncRNAs to regulate biological functions by interacting with their downstream target mRNAs. In this study, many DE lncRNAs have been identified that have not fully recognized thus far in HD. We created a ceRNA regulatory network and performed a functional analysis of mRNAs within the network. It was observed that the most significant molecular function of mRNAs in the ceRNA network was transcriptional regulation, which has a role in the pathogenesis of HD. It has been proposed that lncRNAs in the ceRNA network may be involved in the regulation of transcription. Previous studies have revealed that widespread transcriptional dysregulation occurred in HD cells and animal models, and postmortem HD brain tissues (Consortium, 2017; Hodges et al., 2006). Intriguingly, in the present study, transcripts of HAR1, which was previously shown to be down-regulated in HD patients, was also shown to be dysregulated in our study. HAR1 was observed to be down-regulated in the striatum of patients with HD, owing to repression by its direct target, REST, a critical neural gene regulator (Johnson et al., 2010). However, currently little is known about the mechanism or function of HAR1. It is reasonable to expect that HAR1 may regulate downstream target mRNAs by acting as a ceRNA in HD and contribute to the pathogenesis of HD.

The transcription factor, JUN, also known as c-Jun and AP1, is markedly dysregulated in our ceRNA network and is a crucial target of the c-Jun NH2-terminal protein kinase (JNK) pathway. As an essential mediator of apoptosis in different model systems, JNK is implicated in the regulation of biological processes at both transcriptional and post-transcriptional levels (Davis, 2000). JNK is activated in HD so that its inhibition may be beneficial in correcting HD correlated neurotoxicity (Perrin et al., 2009). Other transcription factors, such as KAT2B, EGR2, MAFB, and GTF2A1, which may be dysregulated through ceRNA network interaction, contribute to transcriptional activation and interact with JUN in the PPI network. This shows that transcriptional dysregulation is involved in the pathogenesis of HD, as found in previous research (Hodges et al., 2006; Kuhn et al., 2007). We found that a few DE mRNAs, such as JUN, LGR6, SFRP1, WNT3, NKD1, RAC3 and LEF1, in the ceRNA network were involved in the canonical Wnt/β-catenin signaling pathway (Fig. 6). The Wnt/β-catenin signaling pathway is an important component in the development of many neurodegenerative diseases (Lecarpentier & Vallee, 2016; Vallee, Lecarpentier, Guillevin, & Vallee, 2017). In HD models, *mutHTT* interferes with β-catenin degradation by binding to several components of the β-catenin degradation complex, resulting in an abnormal accumulation of cytoplasmic β-catenin that cannot enter the cellular nucleus to activate the transcription of target genes (Godin, Poizat, Hickey, Maschat, & Humbert, 2010). In addition, as a crucial regulator of the Wnt canonical pathway, alterations of glycogen synthase kinase 3β (GSK-3β) have been shown in HD (Kim & Snider, 2011). The activation of the Wnt/β-catenin pathway inhibits the

function of GSK-3β, and the increase of the activity of β-catenin in the cellular nucleus can activate the transcription of target genes in the Wnt/β-catenin pathway (Libro, Bramanti, & Mazzon, 2016). Inhibition of Wnt/β-catenin signaling in HD is associated with the transcriptional downregulation of pro-survival genes and excessive apoptosis (Vallee, Lecarpentier, Guillevin, & Vallee, 2018). Therefore, further studies are needed to evaluate other molecules, as mentioned above, involved in this critical pathway in HD.

In addition to transcriptional dysregulation, neuroinflammation is also one of the typical features of most neurodegenerative diseases, including HD (Frank-Cannon, Alto, McAlpine, & Tansey, 2009; Moller, 2010). Previous studies have observed elevated levels of various pro-inflammatory cytokines in the blood and brain tissues of mice and patients with HD, suggesting that inflammation may contribute to HD progression; some inflammatory cytokines are related to the TNF signaling pathway (Bjorkqvist et al., 2008; Crocker, Costain, & Robertson, 2006; Silvestroni, Faull, Strand, & Moller, 2009). Moreover, the NF-κB signaling pathway, as a critical regulator of cytokine production, plays a crucial role in inflammation in HD. It has been shown that overexpression of *mutHTT* can activate the NF-κB pathway by directly interacting with the critical regulator of NF-κB, the IkappaB kinase complex; this may contribute to the neurodegeneration (Khoshnan et al., 2004). Additionally, increased activation of NF-κB has been found in the astrocytes of HD patients and mice, while an inflammatory response mediated by NF-κB in astrocytes facilitates the pathogenesis of HD (Hsiao, Chen, Chen, Tu, & Chern, 2013). In our study, we observed several dysregulated mRNAs, such as TRAF1, PTGS2, and MAP3K8, were involved in the inflammatory response by TNF and NF-κB signaling pathway (Fig. 6). It is suggested that neuroinflammation may contribute to the pathogenesis of HD.

A calcium signaling pathway is critical to neuronal function. In neurodegenerative diseases, impaired Ca2+ signal transduction can interfere with mitochondrial function and synaptic plasticity (Pchitskaya, Popugaeva, & Bezprozvanny, 2018). The impairment of mitochondrial function and alteration of intracellular calcium-induced calcium release and blockade mechanisms exacerbate damage in loop function of HD (Raymond, 2017). Reduced Ca2+ levels in the endoplasmic reticulum and enhanced store-operated calcium entry channels, which play an essential signaling function in neurons (Majewski & Kuznicki, 2015), lead to a synaptic decline in HD (Pchitskaya et al., 2018). Dysregulation of GNAS, CXCR4, ORAI2, ITPKB, and CACNA1I genes in the calcium signaling pathway seems related to the pathogenesis of HD and needs to be further elucidated in future. In addition, many of the downstream molecules in the TNF signaling pathway were dysregulated, including BCL3, TRAF1, VEGFC, PTGS2, and CCL20, which can affect a wide range of functions such as apoptosis, cell survival, as well as inflammation and immunity.

There is widespread recognition of the vital-pivotal role of BDNF in the development and plasticity of the central nervous system (Chang, Khare, Dani, Nelson, & Jaenisch, 2006; Kline, Ogier, Kunze, & Katz, 2010). BDNF produced in the cortex is required for the proper function of corticostriatal synapse and the survival of the GABAergic as well as medium spiny striatal neurons that die in HD (Zuccato & Cattaneo, 2007). It has been well-studied about reduced BDNF transcription or transport because of the indirect repression by mutHTT. However, the post-transcriptional inhibition of BDNF expression or translation by non-coding RNAs in the ceRNA regulatory network has not yet been elucidated in HD. Regrettably, in this paper, after reanalyzing the original expression profile data, we did not find a significant difference in BDNF expression. This result may be as a result of the differences in the cell type used (the parental cell type of iPSCs/NPCs is derived from human fibroblasts, the altered

expression of which may not be as pronounced as that in striatum neurons of HD patients). Moreover, as the results have shown, the protein-encoding genes included in the ceRNA network are mainly transcription factors but not BDNF, which uncovered a new clue of ceRNA possibly involving gene extensive transcriptional misregulation in HD. In summary, it is unclear whether BDNF expression can be decreased through the ceRNA regulatory network in HD.

Although the functions of most differentially expressed lncRNAs in the study are still unclear, they may serve an essential role in the pathogeneses of HD through control of the expression of coding genes in the ceRNA network. Further studies are needed to systematically clarifiy the mechanisms and pathways of the dysregulated genes in HD referred to above.

The limitations of this study should also be mentioned. First of all, lncRNAs can play a role in different ways by virtue of their nuclear or cytoplasmic localization. Here, we only explored the role of lncRNAs as miRNA sponges in the ceRNA network in NPC cell lines derived from iPSC cell lines. Second, only 12 samples were used in this study, with a comparison made between a CAG33 healthy control cell line and a CAG180 HD cell line. In addition, hub genes predicted by our differential analysis have not been verified in HD models used in the lab. Finally, the upstream regulatory molecules of lncRNAs have not been explored, and the relationship between *mutHTT* and the lncRNAs of this study remains to be elucidated.

Conclusions

In conclusion, by using bioinformatics methods in the present study, we have successfully identified hundreds of DE mRNAs and lncRNAs that were not previously linked to HD. By creating a ceRNA network with differentially expressed coding genes and lncRNAs, and in predicting miRNAs, it was revealed that lncRNAs may be involved in facilitating the expression of corresponding mRNAs through a lncRNA-miRNA-mRNA regulatory mechanism, contributing to the pathological process of HD. Functional and PPI network analyses demonstrated that a variety of transcription factors were dysregulated, the expression of which were controlled by lncRNAs. It is further suggested that the function of lncRNAs may be related to a wide range of transcriptional regulation. This study provides new clues to uncovering the mechanisms of lncRNAs in HD and can be used as the focus for further investigation.

Methods

Data acquisition and pre-processing

The expression dataset used in this study was collected from a previous study of NPC samples generated by HD human iPSCs acquired from GEO (https://www.ncbi.nlm.nih.gov/geo/) on the NCBI website. Series Matrix Files of GEO Series Accession NO.GSE93767 and the corresponding platform file of GPL10558 using an Illumina HumanHT-12 v4 Expression BeadChip (Illumina Inc., San Diego, CA, USA) were downloaded. This experiment was designed to explore transcriptional differences between HD and control cells lines, both of them derived from the HD (CAG180) / healthy (CAG33) human iPSCs and hiPSC-derived NPCs (the parental cell types are derived from human fibroblasts), and to identify genome-wide molecular changes after correction of the mutation of HD cell line using a CRISPR/Cas9 and piggyBac transposon-based approach (Xu et al., 2017). A Cas9 nickase(Cas9n)-mediated cleavage for the CAG180 iPSCs was carried out using a pair of selected sgRNAs at the *HTT* locus to correct the

disease mutation, and a piggyBac transposon selection cassette-based homologous recombination donor was used to establishing isogenic control lines. After genome editing and targeted clones screening, a whole-exome sequencing has been performed and identified a low off-target activity in the comparison of the isogenic control iPSCs and their parental CAG180 line, and then all of these cell lines were efficiently differentiated into forebrain NPCs (Xu et al., 2017). In this paper, two corrected isogenic controls of CAG180 HD NPC cell lines (HD-C#1,2), and the non-isogenic CAG33 healthy NPC cell line, were included in the control group; the non-corrected CAG180 HD NPC cell line was included in the HD group. A total of 12 samples were used for this analysis, with three replicates per cell line per cell type. Expression values in the expression matrix file were the log-converted and adjusted data (Xu et al., 2017).

Due to the probes in the HumanHT-12 v4 Expression BeadChip derived from an earlier version of the National Center for Biotechnology Information (NCBI) Reference Sequence (RefSeq), RNAs Release 38, we re-matched all 48,107 probe sequences to human NCBI RefSeq RNA Release 109 using a Basic Local Alignment Search Tool (BLAST, optimized for highly similar sequences) from NCBI, and 38,640 probes were retained. All probes were then converted to their corresponding gene symbols. Before differential expression analysis, if a gene symbol matched more than one probe, the mean processing of the data was made for all corresponding probes. We then analyzed these expression values using an R-Studio tool (version 1.1.463). All expression values were normalized by quantile normalization in package limma (Smyth & Speed, 2003). To discover which genes were targeting lncRNAs or protein-coding genes, we remapped these genes to the human reference genome annotation file downloaded from NCBI (Homo_sapiens.GRCh38.94.chr.gtf); genes without corresponding annotation information were removed. After the above filtering, 20,897 genes in the final dataset remained.

Differentially expressed genes and construction of lncRNA-associated ceRNA network

Differential expression analysis was performed in R studio software version 1.1.463 (Wettenhall & Smyth, 2004) using a Bioconductor 3.8 package(http://www.bioconductor.org/), with an absolute $\log_2^{\text{fold-change}}$ (|logFC|)>1 and P<0.05 as cutoff values by using an empirical Bayes method for screening DE lncRNAs and mRNAs. A lncRNA-related ceRNA regulatory network was constructed according to the "ceRNA hypothesis", in which lncRNA can regulate the expression of mRNAs by competitively binding miRNA, which contains common MREs. Therefore, the miRNAs targeted to DE lncRNAs were predicted by starBase v3.0 based on Ago CLIP-sequencing (seq) data with a threshold of high stringency (>=3) (Li, Liu, Zhou, Qu, & Yang, 2014) or by miRcode based on highly conserved miRNA targets (Jeggari, Marks, & Larsson, 2012). Next, the DE mRNAs targeted-miRNAs were predicted at least by two or more databases among miRDB (Wong & Wang, 2015), TargetScanHuman 7.2 (Agarwal, Bell, Nam, & Bartel, 2015), and the experimentally validated miRNA-target interactions database-miRTarBase 7.0 (Chou et al., 2018). Based on the acquired DE lncRNAs, DE mRNAs, and their co-targeted miRNAs, a ceRNA regulatory network was established and depicted by Cytoscape v3.7.0 software (Shannon et al., 2003). A flow diagram (Fig. 7) clearly shows how we undertook data analysis.

PPI network analysis

To better identify critical genes and clarify the potential relationships of mRNAs in the ceRNA network, protein-protein (PPI) network analysis was carried out using STRING software (score >0.4) (Szklarczyk et al., 2019).

GO, KEGG functional enrichment analysis

To clarify which biological processes and pathways of NPCs induce significant changes in HD for all aberrantly expressed protein-coding RNAs and those included in the ceRNA network, we performed functional enrichment analysis of GO using the R software (Yu, Wang, Han, & He, 2012), and created and visualized a functional enrichment network of KEGG pathways using Cytoscape plug-in ClueGO v2.5.4 (Bindea et al., 2009) and CluePedia v1.5.4 (Bindea, Galon, & Mlecnik, 2013) with the human genome as a background. The ClueGO network was built using kappa statistics to reflect the relationship between terms according to the resemblance of related genes. Functionally grouped networks with terms as nodes were connected based on their kappa score level ($\geqslant$0.4). Pathways showing *p*-value <0.05 were regarded as remarkably enriched for both GO and KEGG analyses.

Abbreviations

HD: Huntington's disease

BP(s): Biological processe(s)

CAG: Cytosine-adenine-guanine

ceRNA: Competing endogenous RNA

DE/DEG(s): Differentially expressed/differentially expressed gene(s)

GO: Gene Ontology

HTT: Huntingtin

iPSC(s): Induced pluripotent stem cell(s)

KEGG: Kyoto Encyclopedia of Genes and Genomes

lncRNA(s): Long non-coding RNA(s)

MF: Molecular function

mutHTT: Mutant huntingtin

NPC(s): Neural progenitor cell(s)

polyQ: Polyglutamine

PPI: Protein-protein interaction

REST: Repressor element 1 (RE1) silencing transcription factor

RISC: Repressive protein complex

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The data analyzed in this study available from pre-existing datasets from Gene Expression Omnibus (NO.GSE93767).

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the National Natural Science Foundation of China (Grant No 81371271) and was also sponsored by the “Liaoning BaiQianWan Talents Program” and 345 Talent Project of Shengjing Hospital of China Medical University.

Authors' contributions

SC conceived and designed the study; XT and YL contributed to data analyses; TZ created the diagrams; XP drafted the manuscript. All authors reviewed the manuscript, made critical revisions and approved the final draft.

Acknowledgements

The authors are grateful to Xiaohong Xu, Yilin Tay, Bernice Sim et al. for providing the original expression dataset GSE93767 on GEO.

Author information

Affiliations

Department of Neurology, Shengjing Hospital of China Medical University, Shenyang, 110004, China

Xiaoping Tan, Shuyan Cong, Yang Liu & Taiming Zhang

Corresponding authors

Correspondence to Shuyan Cong; congsy@sj-hospital.org

References

Agarwal, V., Bell, G. W., Nam, J. W., & Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *Elife, 4*. doi:10.7554/eLife.05005

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell, 136*(2), 215-233. doi:10.1016/j.cell.2009.01.002

Bindea, G., Galon, J., & Mlecnik, B. (2013). CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics, 29*(5), 661-663. doi:10.1093/bioinformatics/btt019

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., . . . Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics, 25*(8), 1091-1093. doi:10.1093/bioinformatics/btp101

Bjorkqvist, M., Wild, E. J., Thiele, J., Silvestroni, A., Andre, R., Lahiri, N., . . . Tabrizi, S. J. (2008). A novel pathogenic pathway of immune activation detectable before clinical onset in Huntington's disease. *J Exp Med, 205*(8), 1869-1877. doi:10.1084/jem.20080178

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., . . . Gustincich, S. (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature, 491*(7424), 454-457. doi:10.1038/nature11508

Chanda, K., Das, S., Chakraborty, J., Bucha, S., Maitra, A., Chatterjee, R., . . . Bhattacharyya, N. P. (2018). Altered Levels of Long NcRNAs Meg3 and Neat1 in Cell And Animal Models Of Huntington's Disease. *RNA Biol, 15*(10), 1348-1363. doi:10.1080/15476286.2018.1534524

Chang, Q., Khare, G., Dani, V., Nelson, S., & Jaenisch, R. (2006). The disease progression of Mecp2 mutant mice is affected by the level of BDNF expression. *Neuron, 49*(3), 341-348. doi:10.1016/j.neuron.2005.12.027

Chou, C. H., Shrestha, S., Yang, C. D., Chang, N. W., Lin, Y. L., Liao, K. W., . . . Huang, H. D. (2018). miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res, 46*(D1), D296-D302. doi:10.1093/nar/gkx1067

Chung, D. W., Rudnicki, D. D., Yu, L., & Margolis, R. L. (2011). A natural antisense transcript at the Huntington's disease repeat locus regulates HTT expression. *Hum Mol Genet, 20*(17), 3467-3477. doi:10.1093/hmg/ddr263

Consortium, H. D. i. (2017). Developmental alterations in Huntington's disease neural cells and pharmacological rescue in cells and mice. *Nat Neurosci, 20*(5), 648-660. doi:10.1038/nn.4532

Crocker, S. F., Costain, W. J., & Robertson, H. A. (2006). DNA microarray analysis of striatal gene expression in symptomatic transgenic Huntington's mice (R6/2) reveals neuroinflammation and insulin associations. *Brain Res, 1088*(1), 176-186. doi:10.1016/j.brainres.2006.02.102

Davis, R. J. (2000). Signal transduction by the JNK group of MAP kinases. *Cell, 103*(2), 239-252. doi:10.1016/s0092-8674(00)00116-1

DiFiglia, M., Sapp, E., Chase, K. O., Davies, S. W., Bates, G. P., Vonsattel, J. P., & Aronin, N. (1997). Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science, 277*(5334), 1990-1993. doi:10.1126/science.277.5334.1990

Ebert, M. S., & Sharp, P. A. (2010). Emerging roles for natural microRNA sponges. *Curr Biol, 20*(19), R858-861. doi:10.1016/j.cub.2010.08.052

Faghihi, M. A., Modarresi, F., Khalil, A. M., Wood, D. E., Sahagan, B. G., Morgan, T. E., . . . Wahlestedt, C. (2008). Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med, 14*(7), 723-730. doi:10.1038/nm1784

Faghihi, M. A., Zhang, M., Huang, J., Modarresi, F., Van der Brug, M. P., Nalls, M. A., . . . Wahlestedt, C. (2010). Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol, 11*(5), R56. doi:10.1186/gb-2010-11-5-r56

Francelle, L., Galvan, L., Gaillard, M. C., Petit, F., Bernay, B., Guillermier, M., . . . Brouillet, E. (2015). Striatal long noncoding RNA Abhd11os is neuroprotective against an N-terminal fragment of

mutant huntingtin in vivo. *Neurobiol Aging, 36*(3), 1601 e1607-1616. doi:10.1016/j.neurobiolaging.2014.11.014

Frank-Cannon, T. C., Alto, L. T., McAlpine, F. E., & Tansey, M. G. (2009). Does neuroinflammation fan the flame in neurodegenerative diseases? *Mol Neurodegener, 4*, 47. doi:10.1186/1750-1326-4-47

Godin, J. D., Poizat, G., Hickey, M. A., Maschat, F., & Humbert, S. (2010). Mutant huntingtin-impaired degradation of beta-catenin causes neurotoxicity in Huntington's disease. *EMBO J, 29*(14), 2433-2445. doi:10.1038/emboj.2010.117

Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., . . . Lander, E. S. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature, 458*(7235), 223-227. doi:10.1038/nature07672

Hodges, A., Strand, A. D., Aragaki, A. K., Kuhn, A., Sengstag, T., Hughes, G., . . . Luthi-Carter, R. (2006). Regional and cellular gene expression changes in human Huntington's disease brain. *Hum Mol Genet, 15*(6), 965-977. doi:10.1093/hmg/ddl013

Hsiao, H. Y., Chen, Y. C., Chen, H. M., Tu, P. H., & Chern, Y. (2013). A critical role of astrocyte-mediated nuclear factor-kappaB-dependent inflammation in Huntington's disease. *Hum Mol Genet, 22*(9), 1826-1842. doi:10.1093/hmg/ddt036

Iyengar, B. R., Choudhary, A., Sarangdhar, M. A., Venkatesh, K. V., Gadgil, C. J., & Pillai, B. (2014). Non-coding RNA interact to regulate neuronal development and function. *Front Cell Neurosci, 8*, 47. doi:10.3389/fncel.2014.00047

Jeggari, A., Marks, D. S., & Larsson, E. (2012). miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics, 28*(15), 2062-2063. doi:10.1093/bioinformatics/bts344

Johnson, R. (2012). Long non-coding RNAs in Huntington's disease neurodegeneration. *Neurobiol Dis, 46*(2), 245-254. doi:10.1016/j.nbd.2011.12.006

Johnson, R., & Buckley, N. J. (2009). Gene dysregulation in Huntington's disease: REST, microRNAs and beyond. *Neuromolecular Med, 11*(3), 183-199. doi:10.1007/s12017-009-8063-4

Johnson, R., Richter, N., Jauch, R., Gaughwin, P. M., Zuccato, C., Cattaneo, E., & Stanton, L. W. (2010). Human accelerated region 1 noncoding RNA is repressed by REST in Huntington's disease. *Physiol Genomics, 41*(3), 269-274. doi:10.1152/physiolgenomics.00019.2010

Khalil, A. M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., . . . Rinn, J. L. (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A, 106*(28), 11667-11672. doi:10.1073/pnas.0904715106

Khorkova, O., Hsiao, J., & Wahlestedt, C. (2015). Basic biology and therapeutic implications of lncRNA. *Adv Drug Deliv Rev, 87*, 15-24. doi:10.1016/j.addr.2015.05.012

Khoshnan, A., Ko, J., Watkin, E. E., Paige, L. A., Reinhart, P. H., & Patterson, P. H. (2004). Activation of the IkappaB kinase complex and nuclear factor-kappaB contributes to mutant huntingtin neurotoxicity. *J Neurosci, 24*(37), 7999-8008. doi:10.1523/JNEUROSCI.2675-04.2004

Kim, W. Y., & Snider, W. D. (2011). Functions of GSK-3 Signaling in Development of the Nervous System. *Front Mol Neurosci, 4*, 44. doi:10.3389/fnmol.2011.00044

Kline, D. D., Ogier, M., Kunze, D. L., & Katz, D. M. (2010). Exogenous brain-derived neurotrophic factor rescues synaptic dysfunction in Mecp2-null mice. *J Neurosci, 30*(15), 5303-5310. doi:10.1523/JNEUROSCI.5503-09.2010

Kuhn, A., Goldstein, D. R., Hodges, A., Strand, A. D., Sengstag, T., Kooperberg, C., . . . Luthi-Carter, R. (2007). Mutant huntingtin's effects on striatal gene expression in mice recapitulate changes observed in human Huntington's disease brain and do not differ with mutant huntingtin length or wild-type huntingtin dosage. *Hum Mol Genet, 16*(15), 1845-1861. doi:10.1093/hmg/ddm133

Leavitt, B. R., van Raamsdonk, J. M., Shehadeh, J., Fernandes, H., Murphy, Z., Graham, R. K., . . . Hayden, M. R. (2006). Wild-type huntingtin protects neurons from excitotoxicity. *J Neurochem, 96*(4), 1121-1129. doi:10.1111/j.1471-4159.2005.03605.x

Lecarpentier, Y., & Vallee, A. (2016). Opposite Interplay between PPAR Gamma and Canonical Wnt/Beta-Catenin Pathway in Amyotrophic Lateral Sclerosis. *Front Neurol, 7*, 100. doi:10.3389/fneur.2016.00100

Li, J. H., Liu, S., Zhou, H., Qu, L. H., & Yang, J. H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res, 42*(Database issue), D92-97. doi:10.1093/nar/gkt1248

Libro, R., Bramanti, P., & Mazzon, E. (2016). The role of the Wnt canonical signaling in neurodegenerative diseases. *Life Sci, 158*, 78-88. doi:10.1016/j.lfs.2016.06.024

Lin, N., Chang, K. Y., Li, Z., Gates, K., Rana, Z. A., Dang, J., . . . Rana, T. M. (2014). An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. *Mol Cell, 53*(6), 1005-1019. doi:10.1016/j.molcel.2014.01.021

MacDonald, M. E., Ambrose, C. M., Duyao, M. P., Myers, R. H., Lin, C., Srinidhi, L., . . . Taylor, S. A. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell, 72*(6), 971-983.

Majewski, L., & Kuznicki, J. (2015). SOCE in neurons: Signaling or just refilling? *Biochim Biophys Acta, 1853*(9), 1940-1952. doi:10.1016/j.bbamcr.2015.01.019

Mehler, M. F., & Mattick, J. S. (2007). Noncoding RNAs and RNA editing in brain development, functional diversification, and neurological disease. *Physiol Rev, 87*(3), 799-823. doi:10.1152/physrev.00036.2006

Moller, T. (2010). Neuroinflammation in Huntington's disease. *J Neural Transm (Vienna), 117*(8), 1001-1008. doi:10.1007/s00702-010-0430-7

Orr, A. L., Li, S., Wang, C. E., Li, H., Wang, J., Rong, J., . . . Li, X. J. (2008). N-terminal mutant huntingtin associates with mitochondria and impairs mitochondrial trafficking. *J Neurosci, 28*(11), 2783-2792. doi:10.1523/JNEUROSCI.0106-08.2008

Packer, A. N., Xing, Y., Harper, S. Q., Jones, L., & Davidson, B. L. (2008). The bifunctional microRNA miR-9/miR-9* regulates REST and CoREST and is downregulated in Huntington's disease. *J Neurosci, 28*(53), 14341-14346. doi:10.1523/JNEUROSCI.2390-08.2008

Panov, A. V., Gutekunst, C. A., Leavitt, B. R., Hayden, M. R., Burke, J. R., Strittmatter, W. J., & Greenamyre, J. T. (2002). Early mitochondrial calcium defects in Huntington's disease are a direct effect of polyglutamines. *Nat Neurosci, 5*(8), 731-736. doi:10.1038/nn884

Pchitskaya, E., Popugaeva, E., & Bezprozvanny, I. (2018). Calcium signaling and molecular mechanisms underlying neurodegenerative diseases. *Cell Calcium, 70*, 87-94. doi:10.1016/j.ceca.2017.06.008

Perrin, V., Dufour, N., Raoul, C., Hassig, R., Brouillet, E., Aebischer, P., . . . Deglon, N. (2009). Implication of the JNK pathway in a rat model of Huntington's disease. *Exp Neurol, 215*(1), 191-200. doi:10.1016/j.expneurol.2008.10.008

Raymond, L. A. (2017). Striatal synaptic dysfunction and altered calcium regulation in Huntington disease. *Biochem Biophys Res Commun, 483*(4), 1051-1062. doi:10.1016/j.bbrc.2016.07.058

Riva, P., Ratti, A., & Venturin, M. (2016). The Long Non-Coding RNAs in Neurodegenerative Diseases: Novel Mechanisms of Pathogenesis. *Curr Alzheimer Res, 13*(11), 1219-1231.

Salmena, L., Poliseno, L., Tay, Y., Kats, L., & Pandolfi, P. P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell, 146*(3), 353-358. doi:10.1016/j.cell.2011.07.014

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res, 13*(11), 2498-2504. doi:10.1101/gr.1239303

Silvestroni, A., Faull, R. L., Strand, A. D., & Moller, T. (2009). Distinct neuroinflammatory profile in post-mortem human Huntington's disease. *Neuroreport, 20*(12), 1098-1103. doi:10.1097/WNR.0b013e32832e34ee

Smyth, G. K., & Speed, T. (2003). Normalization of cDNA microarray data. *Methods, 31*(4), 265-273.

Stine, O. C., Pleasant, N., Franz, M. L., Abbott, M. H., Folstein, S. E., & Ross, C. A. (1993). Correlation between the onset age of Huntington's disease and length of the trinucleotide repeat in IT-15. *Hum Mol Genet, 2*(10), 1547-1549.

Sunwoo, J. S., Lee, S. T., Im, W., Lee, M., Byun, J. I., Jung, K. H., . . . Kim, M. (2017). Altered Expression of the Long Noncoding RNA NEAT1 in Huntington's Disease. *Mol Neurobiol, 54*(2), 1577-1586. doi:10.1007/s12035-016-9928-9

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., . . . Mering, C. V. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res, 47*(D1), D607-D613. doi:10.1093/nar/gky1131

Trushina, E., Dyer, R. B., Badger, J. D., 2nd, Ure, D., Eide, L., Tran, D. D., . . . McMurray, C. T. (2004). Mutant huntingtin impairs axonal trafficking in mammalian neurons in vivo and in vitro. *Mol Cell Biol, 24*(18), 8195-8209. doi:10.1128/MCB.24.18.8195-8209.2004

Tsai, M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., . . . Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science, 329*(5992), 689-693. doi:10.1126/science.1192002

Vallee, A., Lecarpentier, Y., Guillevin, R., & Vallee, J. N. (2017). Effects of cannabidiol interactions with Wnt/beta-catenin pathway and PPARgamma on oxidative stress and neuroinflammation in Alzheimer's disease. *Acta Biochim Biophys Sin (Shanghai), 49*(10), 853-866. doi:10.1093/abbs/gmx073

Vallee, A., Lecarpentier, Y., Guillevin, R., & Vallee, J. N. (2018). Aerobic glycolysis in amyotrophic lateral sclerosis and Huntington's disease. *Rev Neurosci, 29*(5), 547-555. doi:10.1515/revneuro-2017-0075

Wang, K. C., & Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAs. *Mol Cell, 43*(6), 904-914. doi:10.1016/j.molcel.2011.08.018

Wayman, G. A., Davare, M., Ando, H., Fortin, D., Varlamova, O., Cheng, H. Y., . . . Impey, S. (2008). An activity-regulated microRNA controls dendritic plasticity by down-regulating p250GAP. *Proc Natl Acad Sci U S A, 105*(26), 9093-9098. doi:10.1073/pnas.0803072105

Wettenhall, J. M., & Smyth, G. K. (2004). limmaGUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics, 20*(18), 3705-3706. doi:10.1093/bioinformatics/bth449

Wong, N., & Wang, X. (2015). miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res, 43*(Database issue), D146-152. doi:10.1093/nar/gku1104

Xu, X., Tay, Y., Sim, B., Yoon, S. I., Huang, Y., Ooi, J., . . . Pouladi, M. A. (2017). Reversal of Phenotypic Abnormalities by CRISPR/Cas9-Mediated Gene Correction in Huntington Disease Patient-Derived Induced Pluripotent Stem Cells. *Stem Cell Reports, 8*(3), 619-633. doi:10.1016/j.stemcr.2017.01.022

Yoo, A. S., Sun, A. X., Li, L., Shcheglovitov, A., Portmann, T., Li, Y., . . . Crabtree, G. R. (2011). MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature, 476*(7359), 228-231. doi:10.1038/nature10323

Yu, G., Wang, L. G., Han, Y., & He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS, 16*(5), 284-287. doi:10.1089/omi.2011.0118

Zuccato, C., & Cattaneo, E. (2007). Role of brain-derived neurotrophic factor in Huntington's disease. *Prog Neurobiol, 81*(5-6), 294-330. doi:10.1016/j.pneurobio.2007.01.003

Zuccato, C., Ciammola, A., Rigamonti, D., Leavitt, B. R., Goffredo, D., Conti, L., . . . Cattaneo, E. (2001). Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease. *Science, 293*(5529), 493-498. doi:10.1126/science.1059581

Figure Legends

Figure 1. DEGs analysis and visualization.

(A) and (C). Heat map of the top 20 DE long non-coding RNAs (lncRNAs) and mRNAs. (B) and (D). Volcano plots of lncRNAs and mRNAs. Red and green represented up- and down- regulated differentially expressed genes (DEGs), respectively.

Figure 2. Top 20 significantly enriched GO categories of DE mRNAs.

Biological processes (A and B) and molecular functions (C and D) of aberrantly expressed protein-coding RNAs are shown in the pattern of dot-plots. The size of the solid dots represents gene counts, and the colors from red to blue correspond to p-values from small to large. GO, Gene Ontology; DE, differentially expressed.

Figure 3. KEGG Pathway Enrichment Analysis of DE mRNAs.

The results of Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis acquired by ClueGO/CluePedia software were visualized by a Cytoscape platform, with $p < 0.05$ as the cutoff value. Terms grouped with a kappa score level ($\geqslant 0.4$) were shown as nodes of different colors. The size of each node represented significances where only the labels of the most critical terms per group were shown in colors. Edges linked small nodes representing genes related to the corresponding terms. DE, differentially expressed.

Figure 4. A view of the ceRNA network.

Long non-coding RNA (lncRNA), mRNA, and micro(mi)RNA were defined by red circles, blue squares and green rhombuses, respectively. A total of 20 lncRNAs, 48 miRNAs, 189 mRNAs, and 687

edges were included in the competing endogenous RNA (ceRNA) network. The size of the node represented the number of edges connected to it. The transparency of the color represented the significance of the differences; the more in- depth the color, the more significant of the differences.

Figure 5. The protein-protein interaction network.

This PPI network comprised of 123 mRNAs with an interaction score >0.4 in STRING software and was optimized with Cytoscape software. The mRNAs are shown in different sizes and gradient colors of nodes based on the degree of edges. Green, yellow, or red correspond to low, medium, or high degrees, respectively. The edge thickness indicated interaction scores.

Figure 6. A view of KEGG Pathway Enrichment Analysis of mRNAs in ceRNA.

The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment graph was visualized by Cytoscape software based on ClueGO/CluePedia KEGG analysis with $p < 0.05$ and a kappa score $\geqslant$ 0.4. See Fig. 3 for the meaning of nodes and edges. ceRNA, competing endogenous RNA; VEGF, vascular endothelial growth factor; TNF, tumor necrosis factor.

Figure 7. Flow diagram of the data analysis.

Figures

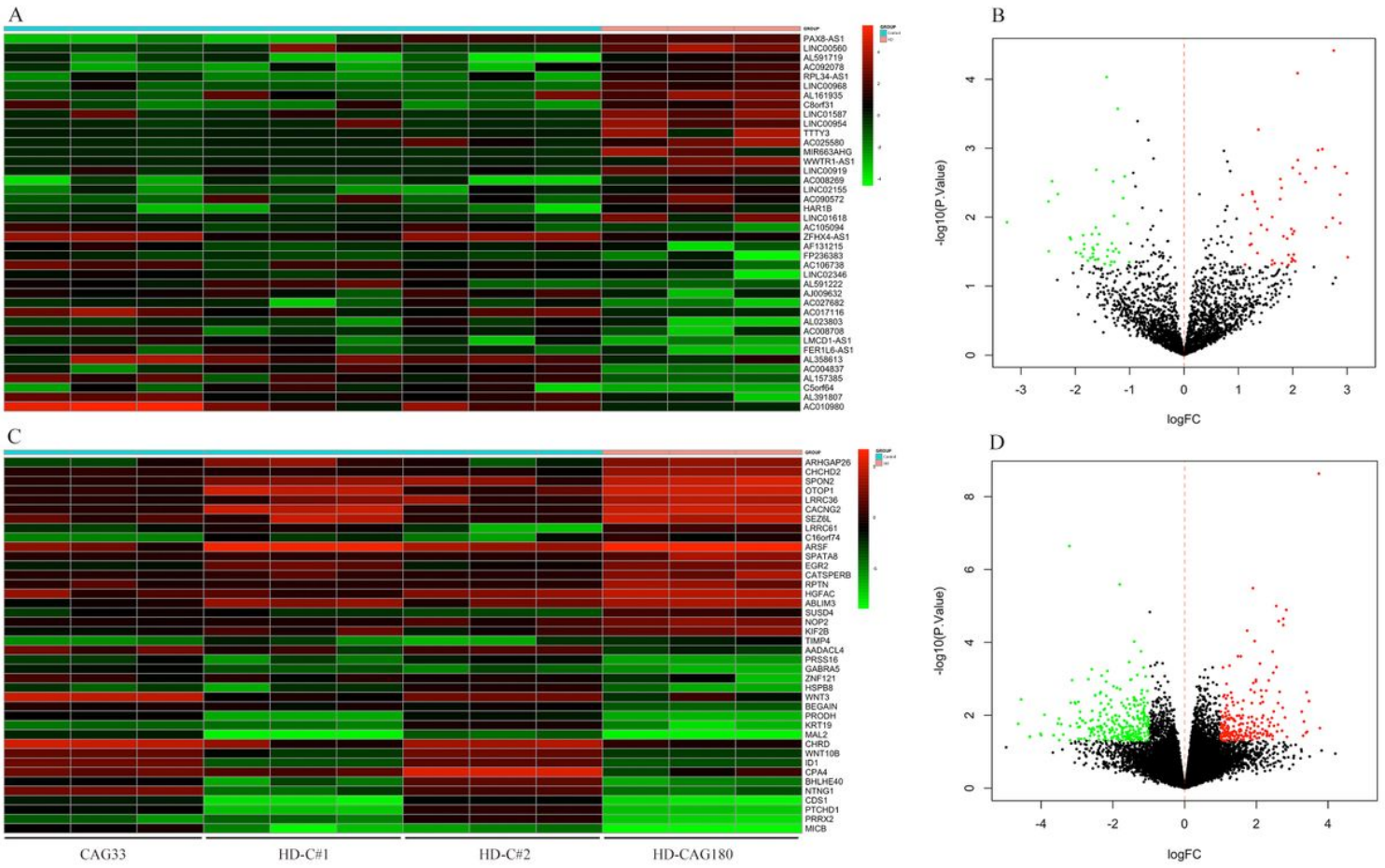

Figure 1

DEGs analysis and visualization. (A) and (C). Heat map of the top 20 DE long non-coding RNAs (lncRNAs) and mRNAs. (B) and (D). Volcano plots of lncRNAs and mRNAs. Red and green represented up- and down- regulated differentially expressed genes (DEGs), respectively.

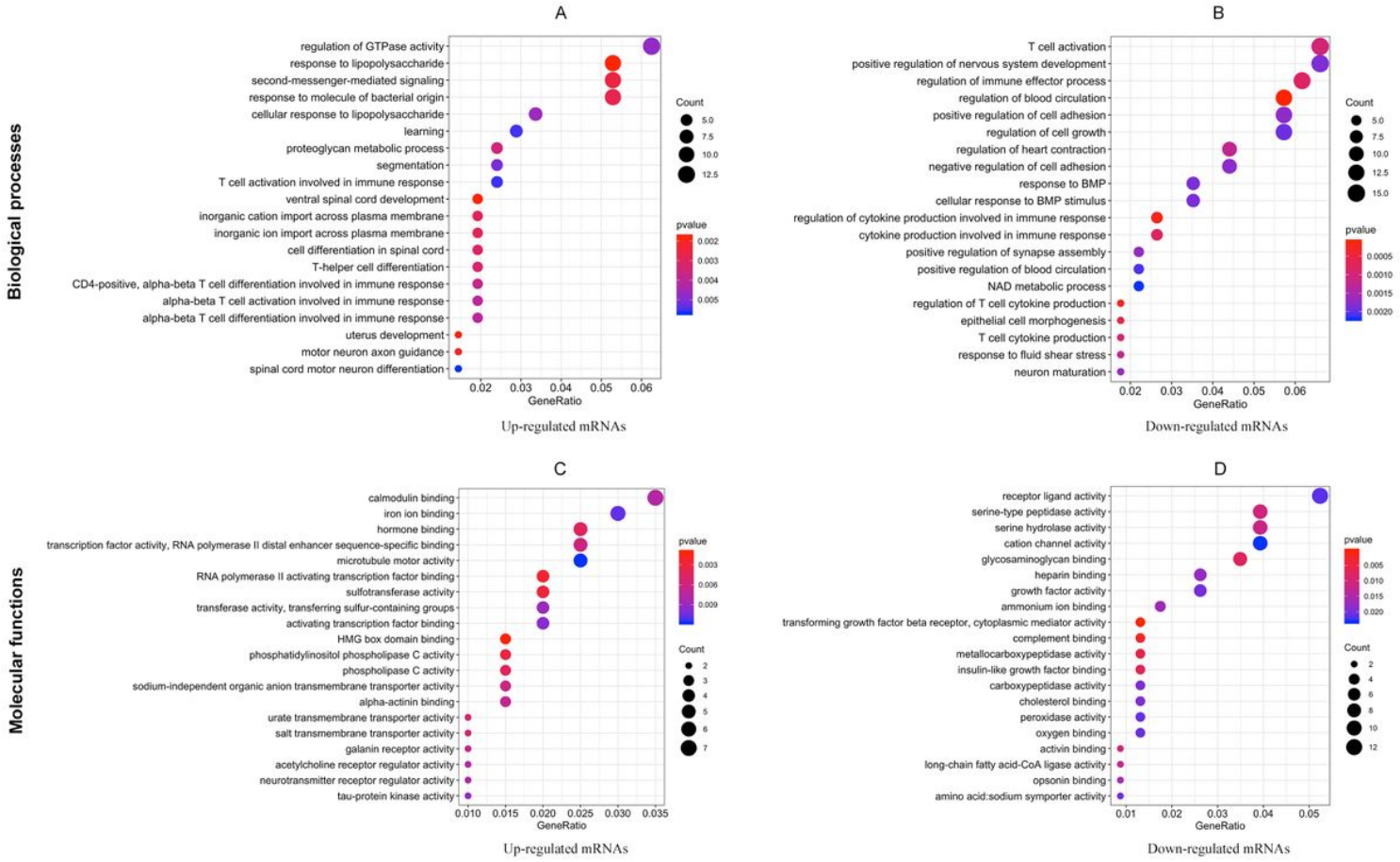

Figure 2

Top 20 significantly enriched GO categories of DE mRNAs. Biological processes (A and B) and molecular functions (C and D) of aberrantly expressed proteincoding RNAs are shown in the pattern of dot-plots. The size of the solid dots represents gene counts, and the colors from red to blue correspond to p-values from small to large. GO, Gene Ontology; DE, differentially expressed.

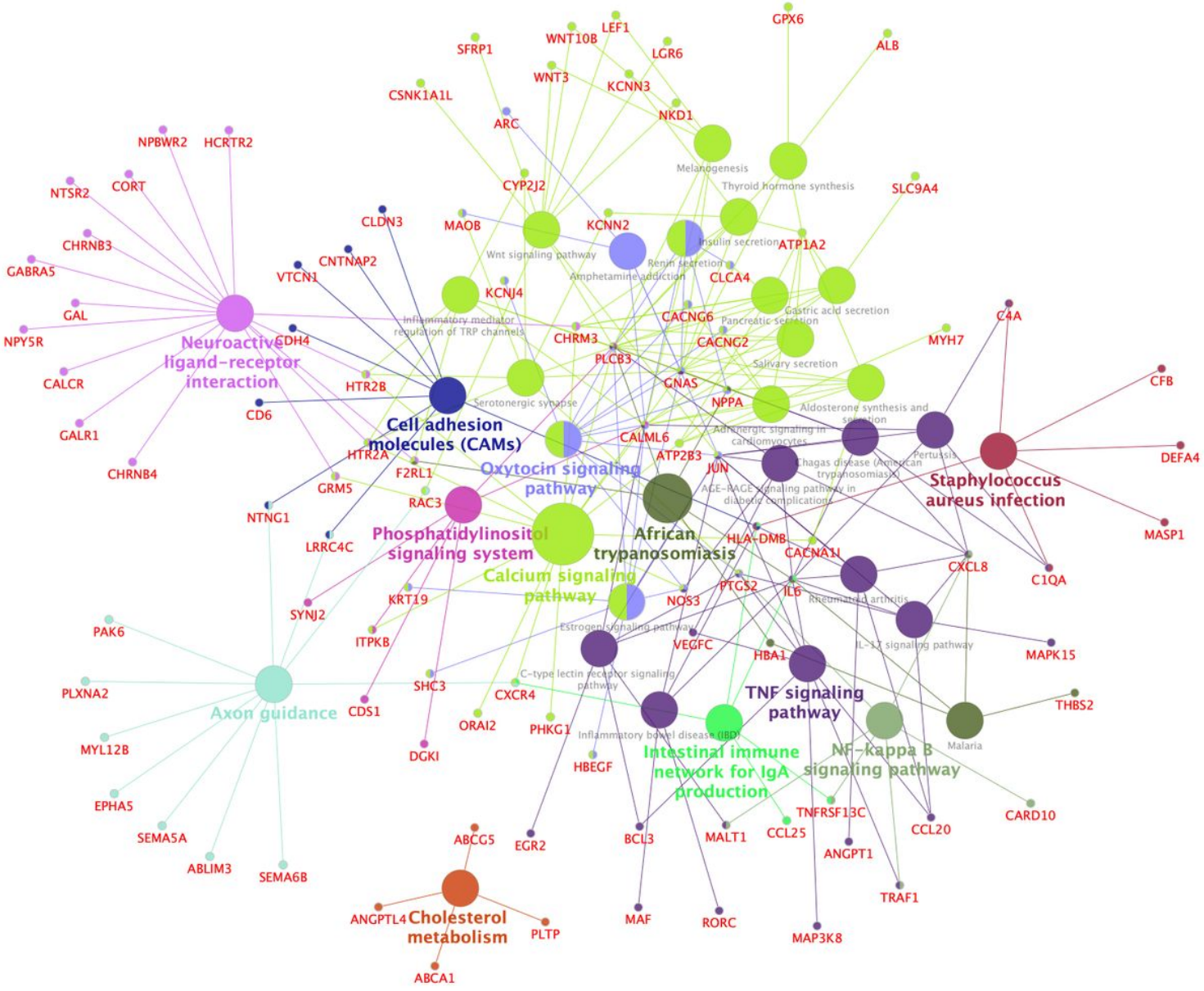

Figure 3

KEGG Pathway Enrichment Analysis of DE mRNAs. The results of Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis acquired by ClueGO/CluePedia software were visualized by a Cytoscape platform, with $p < 0.05$ as the cutoff value. Terms grouped with a kappa score level (≥0.4) were shown as nodes of different colors. The size of each node represented significances where only the labels of the most critical terms per group were shown in colors. Edges linked small nodes representing genes related to the corresponding terms. DE, differentially expressed.

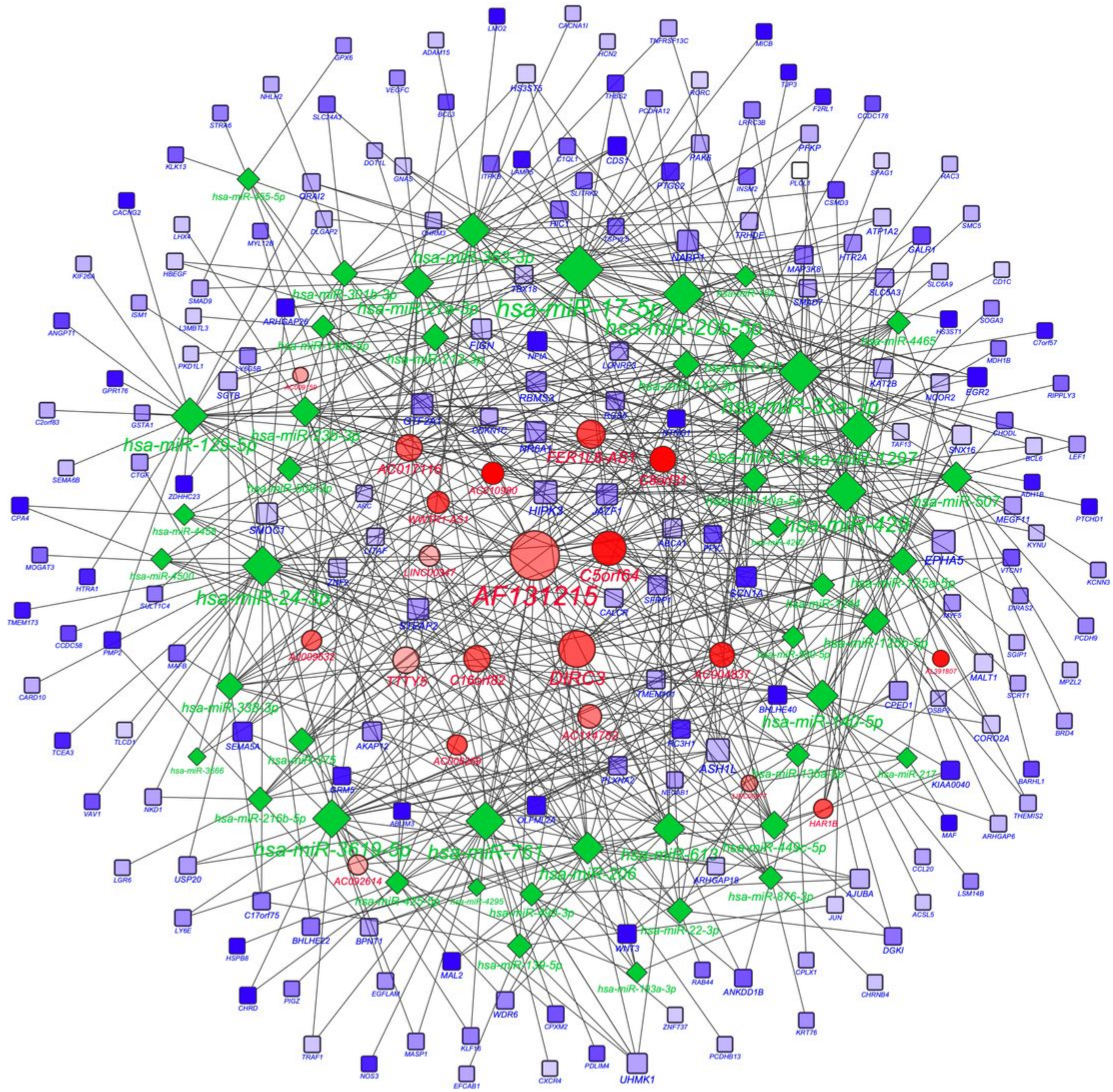

Figure 4

A view of the ceRNA network. Long non-coding RNA (lncRNA), mRNA, and micro(mi)RNA were defined by red circles, blue squares and green rhombuses, respectively. A total of 20 lncRNAs, 48 miRNAs, 189 mRNAs, and 687 edges were included in the competing endogenous RNA (ceRNA) network. The size of the node represented the number of edges connected to it. The transparency of the color represented the significance of the differences; the more in- depth the color, the more significant of the differences.

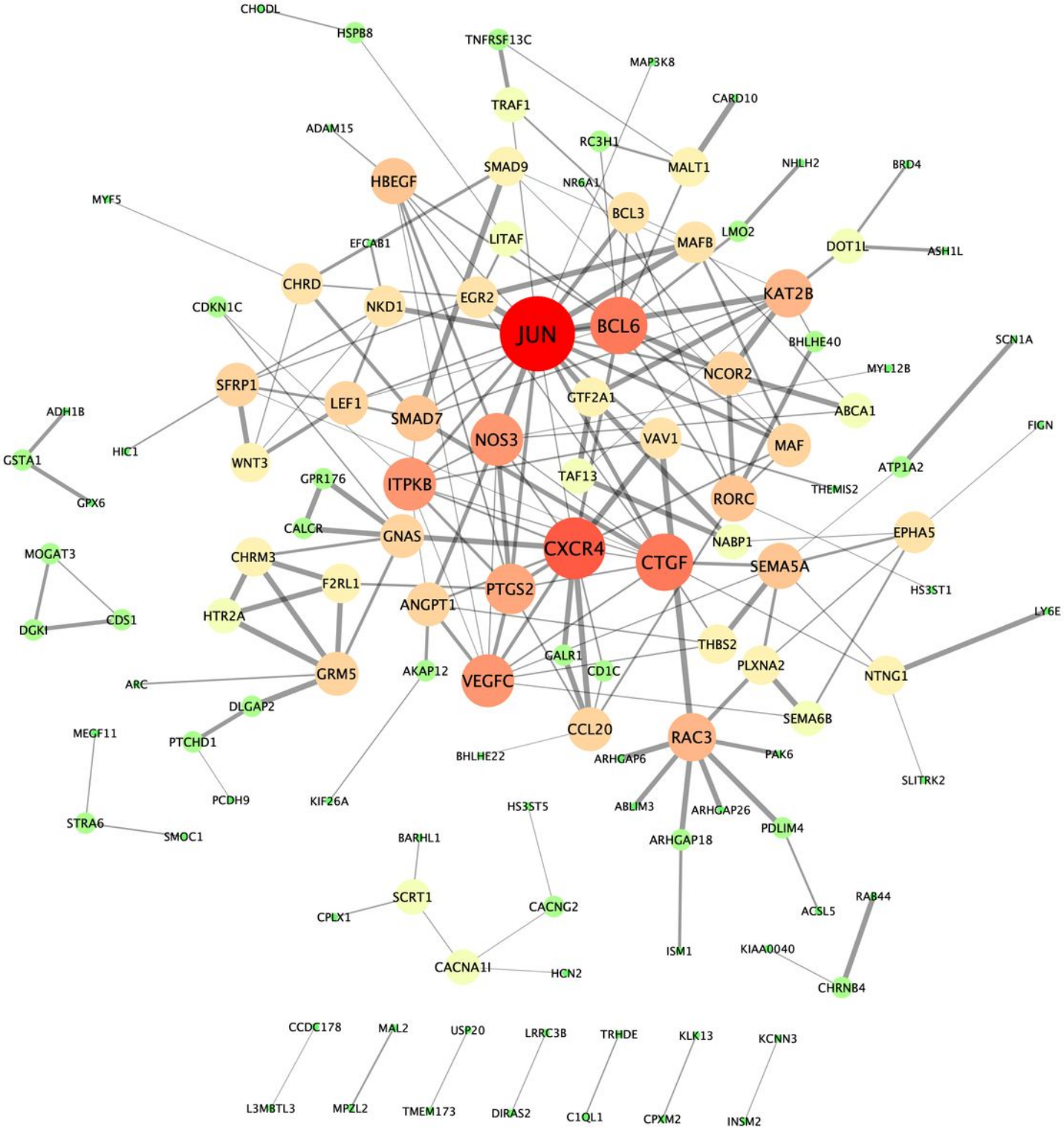

Figure 5

The protein-protein interaction network. This PPI network comprised of 123 mRNAs with an interaction score >0.4 in STRING software and was optimized with Cytoscape software. The mRNAs are shown in different sizes and gradient colors of nodes based on the degree of edges. Green, yellow, or red correspond to low, medium, or high degrees, respectively. The edge thickness indicated interaction scores.

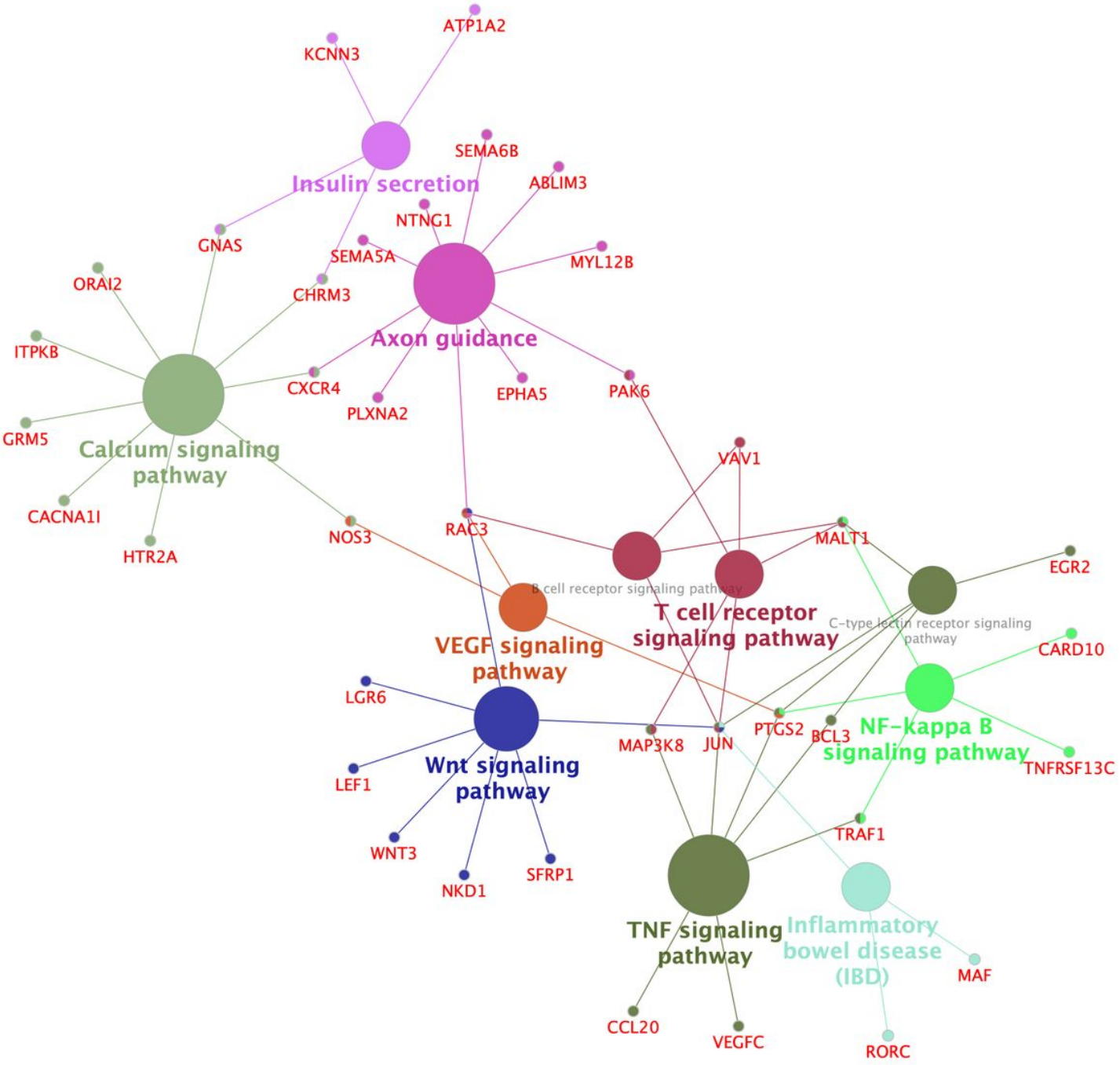

Figure 6

A view of KEGG Pathway Enrichment Analysis of mRNAs in ceRNA. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment graph was visualized by Cytoscape software based on ClueGO/CluePedia KEGG analysis with $p < 0.05$ and a kappa score≥ 0.4. See Fig. 3 for the meaning of nodes and edges. ceRNA, competing endogenous RNA; VEGF, vascular endothelial growth factor; TNF, tumor necrosis factor.

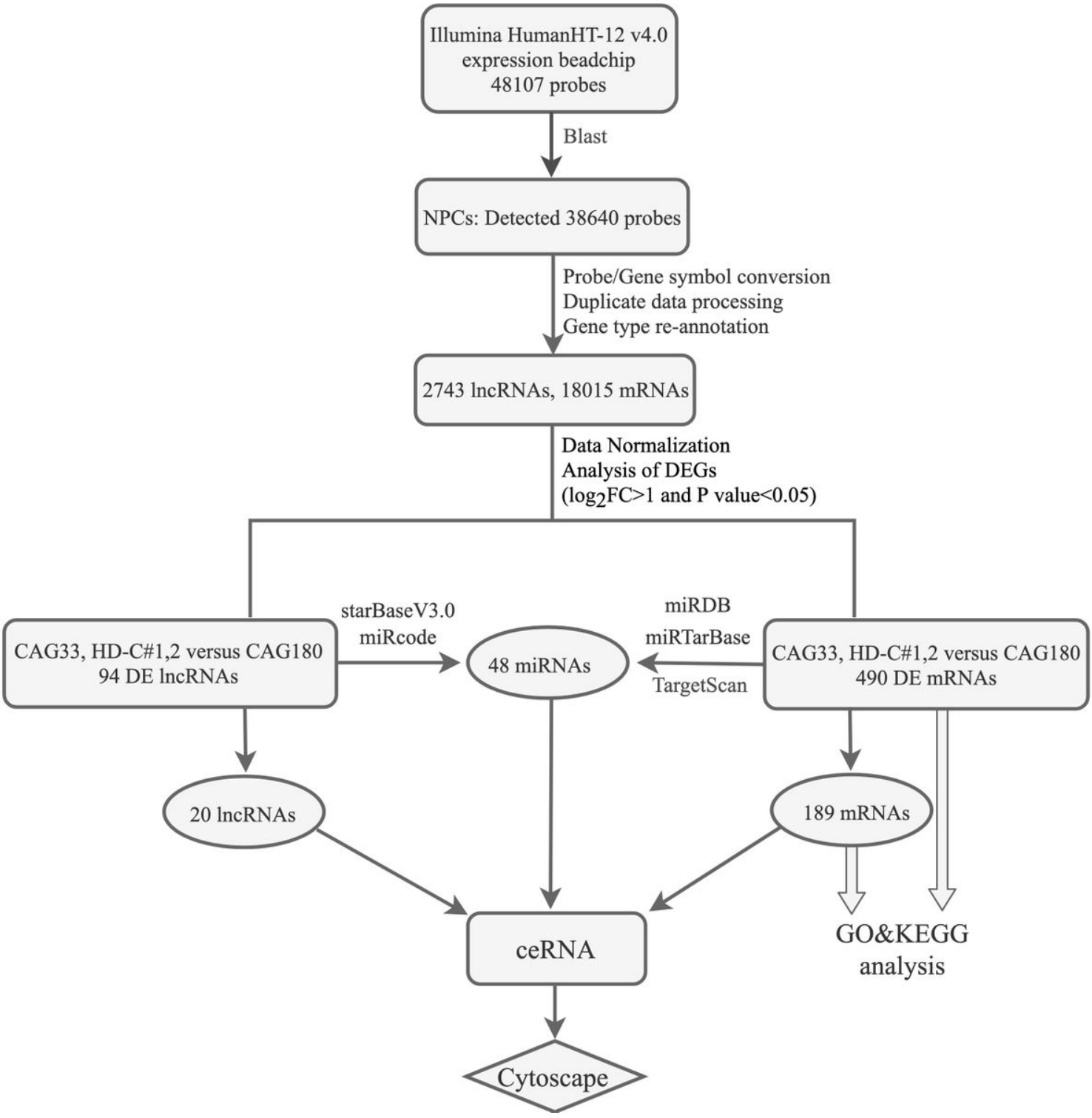

Figure 7

Flow diagram of the data analysis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table.pdf