Analysis of long non-coding RNA expression in hemophilia A patients

Niloofer Naderi, Ali Namvar, Nooshin Amani, Nikoo Nasoohi and Azam Bolhassani

Hematology, 2019, Vol. 24, No. 1, 255–262
https://doi.org/10.1080/16078454.2018.1560934

ABSTRACT

Objectives: Hemophilia A (HA) is an inherited and rare X-linked bleeding disorder which is caused by mutations of the factor VIII gene (FVIII). Long non-coding RNAs (lncRNAs) are non-protein coding sequence transcripts containing more than 200 nucleotides and have potential in diagnosis, prevention and treatment of cancers and inherited bleeding disorders like thalassemia. The goal of this study was to determine the relationship between the expression of lncRNAs and the incidence of hemophilia A in Iranian population.

Methods: In this study, after bioinformatics analysis and IncRNA identification, the expression of two lncRNAs including NONHSAT139215 and NONHSAT139219 was investigated among the blood samples of severe hemophilia A and healthy (non-hemophilia) subjects using the relative qRT-PCR technique.

Results: The melting curve analysis confirmed the specificity of primers. Also, the standard curve showed that the efficiency of reactions for β2-microglobulin (B2M), NONHSAT139215 and NONHSAT139219 was 1.91, 1.96 and 2.01, respectively. On the other hand, the statistical analysis using REST software indicated that the expression of NONHSAT139219 and NONHSAT139215 in severe HA patients was significantly down-regulated as compared to control group (p < .05).

Discussion: Our results demonstrated that a decrease in the expression of two lncRNAs located in FVIII gene may play an important role for the development of severe hemophilia A for the first time.

Conclusions: Briefly, the abnormal expression levels of lncRNAs in hemophilia A cases may be correlated with disease intensity. However, further studies of these lncRNAs are required to provide a useful insight into hemophilia biology.

Introduction

Hemophilia A is an X-linked recessive hemorrhagic disorder caused by different mutations throughout large FVIII gene leading to the deficiency in coagulation factor VIII (FVIII) [1,2]. However, heterogeneous bleeding phenotypes were revealed in hemophilia A patients with the same mutation in FVIII gene [3,4]. The FVIII gene (~ 186 kb) consists of 26 exons that encode a precursor polypeptide with 2351 amino acids. The length of exon changes from 69 to 262 nucleotides except for exons 14 and 26 which have 3106 and 1958 nucleotides, respectively [5]. Based on homology analysis, the FVIII sequence was divided into A (1, 2 & 3), C (1 & 2) and B domains (i.e. NH2-A1-A2-B-A3-C1-C2-COOH). FVIII molecules include many glycosylation sites. It was shown that the B-domain has 19 of the 25 Asn glycosylation sites. In general, the FVIII-coding genes only comprise about 1.5% of the human genome [6,7].

Recent studies have shown that noncoding RNAs (ncRNAs), including ribosomal RNA (rRNA), transfer RNA (tRNA), microRNA (miRNA) and long noncoding RNA (lncRNA) play an important role in many biological and pathological processes [8]. LncRNAs are a group of non-coding RNA with more than 200 nucleotides involved in transcriptional and post-transcriptional regulation of gene expression, and in epigenetics studies [9–11]. A large number of IncRNAs were obtained in eukaryotic organisms ranging from nematodes to humans [12]. Generally, the expression levels of IncRNAs were lower than protein-coding genes [13]. Most IncRNAs were expressed during differentiation and development process. The studies indicated that the IncRNAs may have a role in specific cellular functions including cell proliferation, resistance to apoptosis, induction of angiogenesis, promotion of metastasis, and evasion of tumor suppressors [14,15]. Thus, the mutations and dysregulations of IncRNAs could lead to the development of human complex diseases [6]. For instance, the expressed IncRNAs were significantly correlated with human diseases including neurological disorders, cardiovascular diseases, autoimmune diseases, inflammatory diseases, infectious diseases...
diseases [16,17], diabetes [18], leukemia [19], cancer [10] and AIDS [20] indicating the importance of human disease-related lncRNAs in diagnosis, treatment, prognosis and prevention [13]. Moreover, the roles of lncRNAs are important in ribonucleic protein complexes that regulate various stages of gene expression [15].

As known, proteins and nucleic acids are the two most important types of biological macromolecules in cells [21]. Protein–RNA and protein–DNA interactions play major roles in many biological processes [22]. The studies indicated that intrinsic nucleic acid nature of lncRNAs confers their dual ability to function as ligands for proteins acting in gene regulation, and to mediate base-pairing interactions directing lncRNA-containing complexes to specific RNA or DNA target sites [15]. Docking methods were widely used to study protein–protein interactions; however, only a few methods were available to develop protein–nucleic acid complexes [22]. Docking is a process of sampling and scoring. Molecular docking is a widely used computer simulation procedure to predict the conformation of a receptor–ligand complex, where the receptor is usually a protein or a nucleic acid molecule and the ligand is either a small molecule or another protein [21,22]. In this study, at first, the molecular docking calculations and sequence-based interaction of lncRNA-protein were performed. According to in silico analysis, the expression levels of the selected lncRNAs specific for FVIII gene was determined in Iranian Hemophilia A patients to identify FVIII activity.

**Methods**

**Bioinformatics procedures**

**LncRNAs sequence retrieval**

A large number of lncRNAs were identified by experimental and bioinformatic approaches with many being recorded in public biological databases. In this study, we used Noncode database [23] to obtain sequence information of lncRNAs stated in coagulation factor 8 loci. After analyzing the database, two lncRNAs were founded in the database (NONHSAT139219.2 and NONHSAT139215.2) as shown in Table 1.

**Prediction of lncRNA 3D structure**

Biological functions of RNA molecules depend on the formation of specific three-dimensional (3D) structures of RNA. However, RNA formation was difficult to determine experimentally leading to the development of many predictive bioinformatics tools. In this case, we used 3dRNA server (Automatic building of non-coding RNA 3D structure) [24] to generate 3D structure of NONHSAT139219.2 and NONHSAT139215.2. Each sequence of lncRNA and their 2D structure (obtained from Noncode database) was used as an input, and 3dRNA with optimization method was selected.

**Prediction of RNA-protein interactions using sequence information**

RNA-protein interactions play significant roles in a variety of cellular processes such as protein synthesis, and transcriptional and post-transcriptional regulation of gene expression [25]. High throughput experiments were developed to identify RNA-protein interaction, but they were expensive and time-consuming. Therefore, various computational methods were used to predict RNA-protein interaction based on sequences. RPIseq server used two different algorithms such as RPISeq-SVM, which applied a Support Vector Machine (SVM) classifier, and RPISeq-RF which utilized a Random Forest classifier to identify RNA-protein interaction by sequence information [26]. In this case, lncRNA and protein sequences of coagulation factor VIII (P00451) were selected to evaluate lncRNA-factor VIII interaction scores.

**RNA-protein molecular docking**

Determining the complex structure of RNA-protein is valuable to identify RNA-protein interaction and binding site, thus molecular docking plays a key role in this connection. HDock server used hybrid algorithms consisting of template-based modeling (TM) and free docking which could run the docking procedure, fast and accurately [21]. In this study, top five models of lncRNAs 3D structure (obtained from 3dRNA server) and coagulation factor VIII (PDB code: 3CDZ) were analyzed separately by HDock server.

**Experimental procedures**

**Patient samples**

A 5 ml of venous blood was collected from 50 severe Hemophilia A male patients (the FVIII level < 1%) with the mean age (±SD) of 25.66 (±8.95) and 50 healthy male donors with the mean age (±SD) of 27.42 (±5.51). This study was performed at Iranian comprehensive Hemophilia Care Center, Tehran according to 1975 deceleration of Helsinki and relevant local regulation. All included patients gave written informed consent before entering the study and the Iranian

---

**Table 1. Sequence information of lncRNAs stated in coagulation factor VIII.**

| Name               | Chromosome | Site              | Strand | Length | Exon number | Type   |
|--------------------|------------|------------------|--------|--------|-------------|--------|
| NONHSAT139219.2    | chrX       | 154897560–154898092 | -      | 532    | 1           | Intronic |
| NONHSAT139215.2    | chrX       | 154817098–154817401 | -      | 303    | 1           | Intronic |
Comprehensive Hemophilia Care Center Ethics Committee cleared and approved the study protocol.

**RNA extraction**
Total RNA was extracted from peripheral blood using Qiagen RNeasy Mini Kit (Qiagen, Hilden, Germany) according to manufacturer’s protocol. Total RNA from each sample was quantified using NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific). RNA integrity was assessed for 100 samples via standard denaturing agarose gel electrophoresis.

**Gene expression analysis of lncRNAs by quantitative real-time PCR (qRT-PCR)**
The cDNA was synthesized from total RNA by reverse transcription method using Takara PrimeScript RT Reagent Kit (TaKaRa, Dalian, China) with gDNA Eraser (Perfect Real Time) according to manufacturer’s instructions. The expression of the selected lncRNAs (NONHSAT139215 and NONHSAT13219) was analyzed using qRT-PCR with a SYBR Green PCR kit (TaKaRa) and Real-time PCR system (Roche, Lightcycler 96 system). The sequences of specific primers were listed in Table 2. The levels of lncRNAs were normalized by β2-microglobulin (B2M) expression levels as an internal control.

**Statistical analysis**
The Pfaff method was used to calculate relative gene expression, while accounting for differences in primer efficiency [27]. To perform the Pfaff formula, it requires primer efficiency (E) for gene of interest and the housekeeping gene as well as cycle threshold (Ct) values for all samples. The Pfaff formula is [27]:

\[
\text{ratio} = \frac{(E_{\text{target}})\Delta\text{CP target (control − sample)}}{(E_{\text{ref}})\Delta\text{CP ref (control − sample)}}
\]

Based on this formula, the gene expression ratios for the lncRNAs were analyzed by the REST-2009 software. The results of the molecular assay were reported as the mean ± standard deviation (SD) using Excel or SPSS software.

**Results**

**Prediction of lncRNAs 3D structures**
Three-dimensional (3D) structures of lncRNAs were generated and optimized by 3dRNA server. In each lncRNA, top five models with the highest prediction scores were selected as a model for future molecular docking analysis (Figure 1). For NONHSAT139215.2 and NONHSAT139219.2, free energy scores of the selected 3D models were obtained from 25.26338 to 25.8191 and from 25.1951 to 25.2874, respectively. The results demonstrated high efficiency and accurate 3D modeling of lncRNAs.

**Sequence-based prediction of RNA-protein interactions**
Selected lncRNAs (NONHSAT139215.2 and NONHSAT13219.2) and coagulation factor VIII were analyzed by RPIseq server to predict RNA-protein interaction scores. RF classifier and SVM classifier scores of both lncRNAs were above cutoff (0.5) which showed a high quality of interaction between lncRNA and coagulation factor VIII (Table 3). For evaluation of the experiments, the predictions with probabilities > 0.5 were considered ‘positive,’ indicating that the corresponding RNA and protein were likely to interact. Using this threshold, the accuracy of the classifiers ranged from 87% to 90% in cross-validation experiments on benchmark datasets.

**RNA-protein molecular docking**
RNA-protein interactions could play a key role in different biological pathways such as RNA transcription, RNA splicing, protein synthesis and degradation of nucleic acids. At first, available FVIII pdb file was obtained from RCSB Server. Then, the predicted 3D structure of the selected lncRNA (5 models for each) and FVIII pdb files were submitted to the server separately. After finalizing of the docking process, top models with the lowest free energy scores (the highest docking ranks) were selected for each model of lncRNAs and FVIII (Figures 2 and 3). For NONHSAT139215.2, Model 2 with C1 and C2 binding sites on FVIII and Model 1 with A3 and C1 binding sites on FVIII had the lowest free energy scores. Among the predicted complex for NONHSAT139219.2, Model 2 with A2 binding site on FVIII and Model 4 with A2 and A3 binding sites on FVIII had the lowest energy scores.

**Quantitative RT–PCR validation**
According to the bioinformatics data analysis, two lncRNAs (NONHSAT139215.2 and NONHSAT139219.2) were selected for gene expression analysis by qRT-PCR based on their features. Quantitative values were obtained as threshold fluorescent signal of PCR

Table 2. The designed primers for qRT-PCR validation of candidate lncRNAs and B2M.

| Name           | Forward primer    | Reverse primer    | Product length (bp) |
|---------------|-------------------|-------------------|--------------------|
| NONHSAT139215.2 | ATGCCCAAGGAAAGTCAGTCCACCA | TGCTCTCCTTCCTCCCTTTGTGCA | 174                |
| NONHSAT139219.2 | AAGGCTTATGTGAGCTGCTTCTCA | AGAGTCTCTTCTGGAGACCCAGTGGATT | 89                 |
| B2M           | GGCTATCCAGGCTACTCCAAA | CGGCGACGATGCTACTCTTTT | 246                |
product indicating exponential amplification. Target IncRNA level was normalized to that of the housekeeping gene in the same sample. Each sample was measured in triplicate for each experiment. Moreover, the melting and amplification curves for each PCR product were analyzed to ensure the specificity of the amplification product. The comparative cycle threshold (Ct) method was used to compute the relative quantification of IncRNAs in hemophilia blood samples by comparing them with the healthy samples. A comparison between the hemophilia and non-hemophilia groups showed that the mean of transcript levels of
two lncRNAs was significantly lower in hemophilia A samples compared to normal samples (5.52 in control & 1.25 in case, and 4.86 in control & 2.14 in case for NONHSAT139215 and NONHSAT139219, respectively) ($p < .05$; Figure 4(A,B)). Considering that the expression of lncRNAs was low in the case versus the control, thus the selected lncRNAs could be related to FVIII activity in Iranian hemophilia A patients.

**Discussion**

Hemophilia is one of the most prevalent hereditary diseases and also a public health problem in the world [28]. The goal of this study was to determine the relationship between the expression of lncRNAs and the incidence of severe hemophilia A in the Iranian population for the first time. In this study, the expression of two lncRNAs including NONHSAT139215 and NONHSAT139219 was investigated among the blood samples of severe hemophilia A and healthy (non-hemophilia) subjects. At first, the NONCODE database was used to find the lncRNAs in the sense and anti-sense regions of the FVIII gene on the X chromosome (i.e. Xq28 region). Using in silico analysis, the lncRNAs were selected by the lowest free energy scores on FVIII binding. The studies showed that several structural models are available for analysis of FVIII domains. For instance, a hypothetical model of the FVIII ‘A’ domains was generated from the crystal structure of ceruloplasmin [28]. Nine separate mutations that cause severe disease symptoms (less than 1% of normal circulating FVIII activity) were found on the surface of the A1, C1, and C2 domains involved in binding interactions with membranes, VWF or FIXa [28]. Our data showed that for NONHSAT139215.2, the IncRNA could interact with the C1 and C2 domains, and the A3 and C1 domains on FVIII in models 2 and 1, respectively with the lowest free energy scores. Moreover, for NONHSAT139219.2, the IncRNA could interact with the A2 domain, and the A2 and A3 domains on FVIII in models 2 and 4, respectively with the lowest free energy scores. Then, the expression of these lncRNAs was experimentally evaluated by the relative qRT-PCR technique. The β2-Microglobulin (B2M) was used as an appropriate reference gene. The melting and standard curves were analyzed to determine the specificity and efficiency of the reactions. The melting curve analysis confirmed the specificity of the primers. Also, the standard curve showed that the efficiency of the reactions for B2M, NONHSAT139215 and NONHSAT139219 was 1.91, 1.96, and 2.01, respectively. On the other hand, the statistical analysis using REST software indicated that the expression of NONHSAT139219 and NONHSAT139215 in severe HA patients was significantly down-regulated as compared to control group ($p < .05$).

**Table 3.** Sequence-based prediction of lncRNA-protein interactions.

| Name           | RF Classifier Score | SVM Classifier Score |
|----------------|---------------------|----------------------|
| NONHSAT139215.2| 0.65                | 0.75                 |
| NONHSAT139219.2| 0.7                 | 0.99                 |

*Predictions with probabilities > 0.5 were considered ‘positive’.

![Figure 2](image_url)

**Figure 2.** (A) Top five models of molecular docking between NONHSAT139215.2 IncRNA and FVIII; (B) Docking free energy scores and binding sites of NONHSTA139215.2 IncRNA on FVIII.
Recent advances in sequencing technology revealed that the genome is widely transcribed, generating a large range of non-coding RNAs such as lncRNAs with major roles in different biological processes [29]. Like proteins, the function of lncRNAs depends on their subcellular localization. Many lncRNAs were recognized as important modulators for nuclear functions and exhibited distinct nuclear localization patterns. Others were exported to the cytoplasm to perform their regulatory roles [30]. Up to now, only a small percentage of lncRNAs was indicated to be biologically functional [11,31,32], thus the identification of potential disease-related lncRNAs has been important for understanding their critical roles in development and progression of complex human diseases which may provide new methods in diagnosis, treatment and prevention of diseases (e.g. biomarker or drug design) [33]. In particular, methods for predicting the 3D structures of RNA molecules and protein-RNA complexes were relatively rare. There were simple methods for docking rigid ligands with rigid receptors and flexible ligands with rigid receptors, but general methods of docking for conformationally flexible ligands and receptors were problematic [34].

| Model  | Docking Energy Score | Binding Site on FVIII |
|--------|----------------------|-----------------------|
| Model 1 | -249.86              | A2 & C2 Domain        |
| Model 2 | -284.57              | A2 Domain             |
| Model 3 | -261.20              | A1 Domain             |
| Model 4 | -283.33              | A2 & A3 Domain        |
| Model 5 | -247.56              | A2 Domain             |

**Figure 3.** (A) Top five models of molecular docking between NONHSAT139219.2 lncRNA and FVIII; (B) Docking free energy scores and binding sites of NONHSATA139219.2 lncRNA on FVIII.

**Figure 4.** Relative expression of NONHSAT139215 (A) and NONHSAT139219 (B) lncRNAs in cumulus cells from hemophilia A patients (case) and healthy individuals (control). The blue bars showed the relative lncRNA expression measured using qRT-PCR. * is p-value < 0.05.
no previous report on expression profiles of IncRNAs and the association between IncRNA expression and clinical characteristics in hemophilia A. Up to now only a correlation study between IncRNA expression and β-thalassemia was reported among blood disorders [15]. The researchers showed that some IncRNAs were up- or down-regulated in β-thalassemia patients compared to control (normal) groups. Moreover, abnormal expression levels of IncRNAs in β-thalassemia cases (β-thalassemia major and β-thalassemia minor) may be related to their different clinical phenotypes [15]. In this study, we generated a 3D structure of IncRNAs and then predicted IncRNA-protein interaction by two methods (i.e. sequence based and molecular docking methods). We performed the four-step procedure adopted to exhibit the molecular docking such as target selection, ligand selection and preparation, docking and evaluating its results, and docking software description. Sequence-based and molecular docking prediction scores suggested a high probability of interaction between the selected IncRNAs (NONHSAT139215 and NONHSAT139219) and FVIII. The results also determined the binding site of IncRNAs on FVIII. Based on the predictions, A2, A3 and C1 domains were considered as a candidate for interaction with IncRNAs. Finally, the expression levels of these IncRNAs were quantitatively represented and C1 domains were considered as a candidate for interaction with lncRNAs. Finally, the expression levels of these IncRNAs were quantitatively represented remarkable differential expression between hemophilia and healthy samples. These findings could provide information for the relationship between IncRNA expression profile and FVIII activity. The current study showed that the altered expression of IncRNAs could contribute to hemophilia development.

As known, HA intensity can be classified as mild, moderate and severe forms based on the amount of FVIII. This study showed that low expression levels of IncRNAs in severe HA cases are correlated with the reduction of FVIII levels (< 1 IU/dL). It is possible that the transcription of IncRNAs leads to gene silencing or activation, i.e. the regulatory effects. The results of the present study indicated that down-regulation of two IncRNAs in patients with severe HA may be a strong biomarker with potential clinical applications in prognosis as well as a therapeutic target. Nowadays, IncRNA identification becomes easier due to improved sequencing and bioinformatics tools. However, the number of annotated IncRNA transcripts are raising quickly that should be further studied for HA patients with different levels of coagulation factors in future. It is possible for up-regulation of some IncRNAs in HA disease.

Conclusions

Although the number of identified IncRNAs is growing quickly, further study will be needed to explore their molecular and biological functions in a variety of human diseases. Our results demonstrated that a decrease in the expression of two IncRNAs located in the FVIII gene may play an important role for the development of severe hemophilia A. However, further studies of these IncRNAs are required to provide a useful insight into hemophilia biology. It is possible that the differential expression of IncRNAs facilitates the classification of hemophilia disease into distinct subtypes and the development of IncRNA-based therapies. Thus, IncRNA functional analysis should be performed in future studies.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

None

References

[1] Elmahmoudi H, Khodjet-el-khil H, Wigren E, et al. First report of molecular diagnosis of Tunisian hemophiliacs A: identification of 8 novel causative mutations. Diagn Pathol. 2012;7:93.
[2] Pezeshkpoor B, Theophilus BD, Guilliat AM, et al. Novel characterization of a breakpoint in F8: an individualized approach to gene analysis when PCR and MLPA results contradict. Haemophilia. 2015;21:392–397.
[3] Trampuš Bakija A, Debeljak M, Preložnik Zupan I, et al. Specific and global coagulation tests in patients with mild haemophilia A with a double mutation (Glu113Asp, Arg593Cys). Blood Transfus. 2015;13:622–630.
[4] Rydz N, Leggo J, Tinlin S, et al. The Canadian “national program for hemophilia mutation testing” database: A ten-year review. Am J Hematol. 2013;88:1030–1034.
[5] Antonarakis SE. Molecular genetics of coagulation factor VIII gene and hemophilia A. Thromb Haemost. 1995;74:322–328.
[6] Chen X, Yan CC, Zhang X, et al. Long non-coding RNAs and complex diseases: from experimental results to computational models. Brief Bioinform. 2017;18:558–576.
[7] Chen X, Yan CC, Luo C, et al. Constructing IncRNA functional similarity network based on IncRNA–disease associations and disease semantic similarity. Sci Rep. 2015;5:11338.
[8] Chen D, Sun Q, Cheng X, et al. Genome-wide analysis of long noncoding RNA (IncRNA) expression in colorectal cancer tissues from patients with liver metastasis. Cancer Med. 2016;5:1629–1639.
[9] Dong R, Jia D, Xue P, et al. Genome-wide analysis of long noncoding RNA (IncRNA) expression in hepatoblastoma tissues. PLoS One. 2014;9:e85599.
[10] Zhu H, Wang Q, Yao Y, et al. Microarray analysis of long non-coding RNA expression profiles in human gastric cells and tissues with Helicobacter pylori Infection. BMC Med Genomics. 2015;8:84.
[11] Ding L, Wang M, Sun D, et al. TPGLDA: Novel prediction of associations between IncRNAs and diseases via...
[12] Gu C, Liao B, Li X, et al. Global network random walk for predicting potential human lncRNA-disease associations. Sci Rep. 2017;7:12442.

[13] Chen X, Yan GY. Novel human lncRNA-disease association inference based on lncRNA expression profiles. Bioinformatics. 2013;29:2617–2624.

[14] Fang Y, Fullwood MJ. Roles, functions and mechanisms of long non-coding RNAs in cancer. Genomics Proteomics Bioinformatics. 2016;14:42–54.

[15] Ma J, Liu F, Du X, et al. Changes in lncRNAs and related genes in β-thalassemia minor and β-thalassemia major. Front Med. 2017;11:74–86.

[16] Shi X, Sun M, Liu H, et al. Long non-coding RNAs: a new frontier in the study of human diseases. Cancer Lett. 2013;339:159–166.

[17] Maass PG, Luft FC, Bahringer S. Long non-coding RNA in health and disease. J Mol Med. 2014;92:337–346.

[18] Alvarez ML, Di Stefano JK. Functional characterization of the plasmacytoma variant translocation 1 gene (PVT1) in diabetic nephropathy. PLoS One. 2011;6:e18671.

[19] Calin GA, Liu CG, Ferracin M, et al. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell. 2007;12:215–229.

[20] Zhang Q, Chen CY, Yedavalli VS, et al. NEAT1 long non-coding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. M Bio. 2013;4:e00596–12.

[21] Yan Y, Zhang D, Zhou P, et al. HDock: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. Nucleic Acids Res. 2017;45:W365–W373.

[22] Tuszynska I, Magnus M, Jonak K, et al. NPDock: a web server for protein-nucleic acid docking. Nucleic Acids Res. 2015;43:425–430.

[23] Zhao Y, Li H, Fang S, et al. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic Acids Res. 2016;44:D203–D208.

[24] Zhao Y, Huang Y, Gong Z, et al. Automated and fast building of three-dimensional RNA structures. Sci Rep. 2012;2:734.

[25] Yi Y, Zhao Y, Huang Y, et al. A brief review of RNA-protein interaction database resources. Noncoding RNA. 2017;3:e6.

[26] Muppirala UK, Honavar VG, Dobbs D. Predicting RNA-protein interactions using only sequence information. BMC Bioinformatics. 2011;12:489.

[27] Pfaffl MW. A new mathematical model for relative quantification in real time RT-PCR. Nucleic Acids Res. 2001;29:e45.

[28] Shen BW, Spiegel PC, Chang CH, et al. The tertiary structure and domain organization of coagulation factor VIII. Blood. 2008;111:1240–1247.

[29] Wang X, Hao C, Bao H, et al. Aberrant expression of long noncoding RNAs in cumulus cells isolated from PCOS patients. J Assist Reprod Genet. 2016;33:111–121.

[30] Chen LL. Linking long noncoding RNA localization and function. Trends Biochem Sci. 2016;41:761–772.

[31] Dinger ME, Amaral PP, Mercer TR, et al. Pervasive transcription of the eukaryotic genome: functional indices and conceptual implications. Brief Funct Genomic Proteomic. 2009;8:407–423.

[32] Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009;10:155–159.

[33] Sun J, Shi H, Wang Z, et al. Inferring novel lncRNA-disease associations based on a random walk model of an lncRNA functional similarity network. Mol Biosyst. 2014;10:2074–2081.

[34] Hernández-Santoyo A, Tenorio-Barajas AY, Altuzar V. Protein-protein and protein-ligand docking. Intech Open. 2013;3:63–81.