Recent Research Progress of Long Non-coding RNA

Wang Guofeng, Tang Haibao

Center for Genomic and Biotechnology, College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou, Fujian, 350002, China

Corresponding author email: tanghaibao@gmail.com

Genomics and Applied Biology, 2018, Vol.9, No.10  doi: 10.5376/gab.2018.09.0010

Received: 02 Nov., 2018
Accepted: 25 Dec., 2018
Published: 30 Dec., 2018

Copyright © 2018 Wang and Tang. This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:
Wang G.F., and Tang H.B., 2018, Recent research progress of long non-coding RNA, Genomics and Applied Biology, 9(10): 62-71 (doi: 10.5376/gab.2018.09.0010)

Abstract In general, long non-coding RNA (lncRNA) refers to the transcripts that are greater than 200 nucleotides in length which cannot encode a protein but have influential functions. Recently, lncRNA has gradually become a hot spot in the study of biology. This review summarized the advances in understanding the roles of lncRNA in human disease, plant biology among many other directions and how these results would affect us. An increasing number of lncRNA software has been developed to predict lncRNA with the increasing attention paid to lncRNA. We mainly introduced several related software for lncRNA predictions, including their sorting algorithms, and different advantages and disadvantages in their performance and accuracy. With the improvement of sequencing technologies, the approaches for predicting lncRNA have also changed accordingly. In this review, the lncRNA prediction and quantitative processes of different sequencing technologies were introduced in detail. Finally, we summarized some of the challenges in lncRNA research as well as future prospects of lncRNA research, hoping to provide some help for future lncRNA research.

Keywords LncRNA; Bioinformatics; Sequencing; Gene regulation; LncRNA prediction

Background

Gene expression regulation plays an important role in growth and development of creature, and it can be divided into regulation before transcription, regulation during transcription, regulation after transcription, regulation during gene translation and after translation (Wang et al., 2015). LncRNA plays a vital role in gene expression regulation of eucaryon (Hu et al., 2016). It is regarded that many LncRNA is not only correlated with mRNA expression regulation, but also with growth and development of creature as well as biological diseases (Guo and Liu, 2014).

As more and more researchers are involved in the study of lncRNA, the main mechanism by which lncRNA performs biological functions has become increasingly clear. LncRNA mainly perform their biological functions through the following biological mechanisms. For example, cell cycle regulation, mRNA degradation and chromatin remodeling, gene imprinting, increasing the stability of mRNA, serving as the skeleton molecule of histone modified complex and regulating the phosphorylation of serine and arginine splicing factors, and so on (Xia et al., 2013).

With the in-depth study and development of lncRNA, it has been found that lncRNA not only performs its biological functions in organisms with various mechanisms, but also is related to the occurrence of certain diseases. For example, latest studies have shown that lncRNA molecules can significantly accelerate the metastasis and regulate the epithelial interstitial transformation that controls colon cancer (Chen et al., 2017). Although with the progress of sequencing technology, a large number of lncRNA have been identified. Our understanding of the functions and actions of these identified lncRNA is still quite limited, and we still need to explore the functions and actions of these lncRNA in vivo through new research methods.

LncRNA research is at the forefront of biological research and is one of the hot spots at present. The study of lncRNA helps us to deep understand the structure and functions of genes from a new perspective. With the continuous improvement of research technology and level, some biological regulatory processes related to
lncRNA will be more clearly displayed to us, thus allowing us to further understand the functions of different transcripts. Some lncRNA have been found to have abnormal expression in tumor tissues, and these related lncRNA can be used as biomarkers for the prediction of tumor or cancer (Yarmishyn and Kurochkin, 2015). Up to now, a large number of lncRNA have been identified as biomarkers for cancer (Iyer et al., 2015). For example, in a study of lncRNA in prostate cancer cells, researchers discovered an lncRNA called SChLAP1 (Zhao, 2014). The expression of these lncRNA in prostate cancer cells was significantly higher than that in prophase prostate cancer cells. This characteristic is helpful for judgment of cancer and postoperative observation, thus increasing means of studying the pathogenesis of cancer of researchers. In plant studies, researchers have found that rice specific lncRNA-LDMAR could cause programmed cell death in the anther tapetal layer of rice, and result in photosensitive male sterility (Ding et al., 2012). This result might be of great significance to the research of light regulation pathway of rice.

1 Action Mode of LncRNA

LncRNA does not encode proteins, but it can perform a lot of important functions in vivo through the regulation of genes. LncRNA can perform their functions in the following ways (Figure 1): 1. LncRNA can be transcribed in the upstream region of genes, which can affect gene expression; 2. LncRNA can affect gene expression by inhibiting certain polymerases; 3. LncRNA acts by interfering with mRNA splicing; 4. Cooperate with Dicer enzyme to regulate gene expression; 5. Directly regulate the activity of related proteins; 6. Formation of nucleic acid protein complex; 7. Change the localization of cytoplasm; 8. Interaction with related small molecules (Jeremy et al., 2008).

The regulation of lncRNA on gene expression can be classified into the following categories (Figure 2): 1. lncRNA will regulate genes at the level of epigenetics; 2. LncRNA will regulate during RNA transcription; 3. It will regulate after RNA transcription completed. Relevant studies have shown that lncRNA can activate nearby gene expression and also regulate the activity of transcription factors and RNA polymerase so as to regulate gene expression as a transcription activator (Ørom et al., 2010).

2 Research Progress of LncRNA in Human Diseases

We now know that lncRNA can bind to and interact with large molecules in cells, such as DNA, RNA and proteins, to control some important cancer phenotypes related to transcripts. Numerous studies have shown that HOTAIR lncRNA is closely related to human cancer (Wu et al., 2014). HOTAIR was found for the first time in...
human breast cancer cells. If this expression level increases in the tumor, it means that the tumor will deteriorate to the advanced stage of cancer and cannot be treated, so it can be used as an important tumor marker for diagnostic detection (Gupta et al., 2010). LncRNA also regulates the genes that encode proteins that are self-related. If the regulation of these lncRNA is wrong, it may also lead to the occurrence of diseases. Relevant studies have shown that p15, a gene that inhibits cancer, can be transcribed to generate an antisense lncRNA, which is able to induce DNA methylation, leading to leukocytosis (Yu et al., 2008). In addition, there are other related diseases, such as liver cancer related lncRNA:ZFAS1. If the expression of this lncRNA increases in mice, the metastasis of liver cancer cells will be promoted (Li et al., 2015).

Figure 2 Expression regulation on genes by lncRNA at 3 levels

Another disease is also closely related to lncRNA. Alzheimer’s disease is also known as senile dementia. Some recent studies have found that lncRNA is also an important factor in the formation of Alzheimer’s disease. The main cause of Alzheimer’s disease is the amyloid produced by a secretase, which will induce Alzheimer’s disease if it accumulates in the body (Burns, 2009). The antisense chain of the secretase encoding gene BACE1 can be transcribed to generate lncRNA:BACE1AS (Tan et al., 2013). This lncRNA prevents mRNA produced by the coding genes of secretase from being degraded, leading to the continuous increase of amyloid protein, while the accumulation of amyloid in turn leads to the expression of secretase encoding genes. This kind of positive feedback mechanism makes Alzheimer’s disease or senile dementia worse. According to this mechanism, silencing or inactivating lncRNA might be a way to treat or alleviate Alzheimer’s disease.

3 Research Progress of LncRNA in Plant

Compared with lncRNA in animals, researches of lncRNA in plant are far from enough. Relatively few species are studied on lncRNA at present. Studies on lncRNA in mammals are relatively abundant, compared with those in plants. Among them, human and mouse lncRNA have been studied relatively deeply in mammals (Sun et al., 2012; Shi et al., 2013), while the types of plants studied are more limited. These lncRNA in plants have surprisingly strong tissue expression specificity. A specific example is the identification of rice lncRNA at the genome-wide level through the use of high-throughput sequencing technology (Zhang et al., 2014). Analysis of the identified rice lncRNA revealed that these lncRNA had high tissue specific expression.

It was found in the studies on rice lncRNA that some of these lncRNA were highly expressed during sexual reproduction of rice, suggesting that the function of these lncRNA was related to sexual reproduction in rice. Further studies showed that one lncRNA (XLOC_057324) in rice could affect the development of rice spikes.
Researchers found the insertion mutation of a lncRNA, XLOC_057324, in the rice mutation database. Mutations at the lncRNA insertion site resulted in a significant decrease in isoforms content of such lncRNA, leading to phenotypic changes. Researchers planted the lncRNA mutant and wild type of ZH11 rice at the same time, and found that the mutant rice bloomed earlier than the wild type, but the ears of the mutant rice were not significantly as full as the wild type. The findings of this study fully proved that lncRNA (XLOC_057324) was involved in the regulation of rice spike growth and development and was closely related to sexual reproduction of rice.

Latest researches have indicated that lncRNA also plays an important regulatory role in the growth and development of soybeans. In a study on lncRNA in soybean, researchers found that lncRNA could be involved in stress response, signal transduction and development process through co-expression analysis of protein-coding genes and lncRNA. In addition, the researchers also observed the expression of lncRNA in centromere regions, especially in active meristems, suggesting that lncRNA might be involved in the regulation of cell division.

4 Related Databases and Its Classification of lncRNA

4.1 LncRNA databases
With the wide spreading of lncRNA researches, a number of different organizations and academic research institutes have established lncRNA databases in order to facilitate relevant researches and academic exchanges (Amaral et al., 2011). The data of these databases are from a wide range of sources, of which some are from published literature or directly from experiments, and are classified according to different species or different research objects (Table 1).

| Database name     | Characteristics                                                                 | Website                                |
|-------------------|---------------------------------------------------------------------------------|----------------------------------------|
| Noncode           | Established by Institute of Computing, Chinese Academy of Sciences, it is by far the most complete database of non-coding RNA (Liu, 2004). When using this database, there is a place that needs to be paid special attention. Since this database uses a set of lncRNA naming system specified by itself, if the general naming system is used to search in this database, the desired lncRNA cannot be found. Here is a way to find the lncRNA we want in this database. Blast function is provided in this database. If the user has the nucleic acid sequence of lncRNA, the corresponding number of lncRNA in this database can be found according to blast results, so as to further find the annotation information that the user wants. | http://www.noncode.org                  |
| lncRNA Disease    | The database collects more than 160 disease-related lncRNA and provides annotations of disease-related lncRNA reported in literature (Chen et al., 2012). It is worth mentioning that this database provides the function of browsing lncRNA annotation information on the website, which makes it much easier for us to annotate lncRNA. | http://cmbi.bjmu.edu.cn/lncrnadisease  |
| CHIPBase          | This database is significantly rich in content, so you can find a lot of lncRNA comments here. This database also provides the loci that bind lncRNA and cn/chipbase/ expression map to transcription factors identified by RNA-seq (Yang et al., 2013). | http://deepbase.sysu.edu.cn/            |
| lncRNAdb          | The database collects eukaryotes lncRNA information that has been reported in the literature. Each entry in the database contains reference information about RNA, the expression of lncRNA, subcellular localization and conservative, functional evidence and other relevant information. We can use the lncRNAdb database to continuously proofread the literature related to lncRNA and other genomic elements (Amaral et al., 2011). | http://www.lncrnadb.org/               |

4.2 LncRNA classification
There are many classification methods for lncRNA. lncRNA can be divided into the following categories according to the relative positions of lncRNA in the genomes and exons (Figure 3). Sense lncRNA refers to lncRNA that overlaps with one or more exons of the protein-coding gene of the chain in the same position.
Antisense lncRNA refers to lncRNA that overlaps with one or more exons of antisense chain gene. Intergenic lncRNA refers to the lncRNA generated by two protein-coding genes. Intronic lncRNA refers to the lncRNA between two exons from the protein-coding gene. At present, the functions of different types of lncRNA have not been deeply studied. Finding common features related to the functions of different types of lncRNA will be of great help to further understand the mechanism of lncRNA in the body.

Figure 3 Classification of lncRNA based on its location on genome

5 Software Related to LncRNA Prediction

The key problem to predict lncRNA is how to distinguish mRNA from ncRNA. At present, the main method to distinguish mRNA from ncRNA is to establish a classifier, which is mainly based on the sequence characteristics of lncRNA, such as the modification site of histones, the arrangement of bases and the conservatism of sequences.

One of the representatives of the software developed based on the characteristics of base sequence is PLEK (predictor of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme) (Li et al., 2014). This software classifies transcripts based on SVM (support vector machine) algorithm. It calculates the k-mer frequency of transcripts, and can divide transcripts into protein-coding transcripts and non-protein-coding transcripts. The classification of PLEK does not depend on sequence alignment or genomic information. In addition, one of the advantages of the software is that it runs faster. When forecasting the same set of data, the prediction speed of PLEK software is 8 times that of CNCI and 244 times that of CPC which is the most popular prediction software. The accuracy of PLEK may fluctuate when it predicts different species. For example, when using the known mRNA of mice for testing, PLEK wrongly judged the most lncRNA, but if the data of mice were replaced to the data of corn, PLEK was satisfactory for its higher accuracy performance. In general, it is a stable and reliable prediction software.

Another classic lncRNA prediction software was CNCI (Coding-non-coding Index), which was developed by the team of Zhao Yi from the Institute of Computing, Chinese Academy of Sciences (Sun et al., 2013). The software
classifies sequences based on their characteristics. This software calculates the replacement frequency of codons adjacent to the transcripts of coding protein and non-coding protein, which is later used to construct score matrix. The most similar coding sequence (CDS) are selected, and SVM classifier is constructed by combining single nucleotide frequency. This software works for incomplete sequences, so it is more suitable for partial EST sequences or transcripts spliced from scratch.

At present, there are many other prediction software in addition to the two kinds of prediction software described above. The core ideas of these software are the same, but the specific implementation methods are different. According to the different cases of specific sequence, the predicted results of these software have their own advantages and disadvantages, and the intersection of the predicted results of several software can be used as a reliable result.

6 Bioinformatical Identification Procedure of LncRNA

Now the analysis and identification process of LncRNA can be roughly divided into two categories according to their data sources (Figure 4). The data generated by Illumina sequencing technology can be identified and analyzed in one category. Another category is to identify and analyze the data obtained by single-molecule sequencing with third-generation sequencing technology.

Illumina sequencing platform is widely used. The data generated by Illumina sequencing platform need to be analyzed and converted into original sequence data (Raw Data). The next thing is to filter and clean the data, because the raw offline data generally contains some joint contamination, and may also contain some low-quality reading lengths. The general requirements of filtration operations are to remove reads with sequencing connectors
or selectively remove bases that are not accurately measured (Iyer et al., 2015). Clean filtered data after filtration is assembled into transcripts by software. There are two ways of assembling transcripts. One is the assembly based on the reference sequence, which used Tophat to compare the sequencing data to the reference genome (Kim, 2014), and later used cufflinks for stitching (Trapnell, 2014). This method has high sensitivity and requires less memory for calculation. The other is to obtain transcripts directly from overlap assembly between sequencing reads without reference genome assembly. This method does not rely on the comparison software and the existing reference genome, but requires large memory resources as well as higher sequencing depth.

After the assembly of transcripts, the next step is to screen lncRNA. There are several basic screening methods for lncRNA: one is length, the other is the exon number. For instance, we usually select the length longer than or equal to 200 bp from the assembled transcript according to the common distribution of lncRNA (Wilusz et al., 2009). The resulting transcripts were then functionally screened to predict their potential ability to encode proteins. The analysis methods commonly used now are as follows: CPC software analysis, CNCI software analysis, pfam protein domain analysis and CPAT software analysis. In general, the intersection of these software types will be taken to reduce the potential false positives probably caused by a single software.

Once lncRNA is obtained, downstream analysis can be carried out. For instance, lncRNA family classification, lncRNA expression analysis, lncRNA difference analysis, lncRNA-mRNA co-expression analysis, Pathway enrichment analysis, and so on. Finally, functional analysis and sample properties of lncRNA will be combined to discuss relevant biological problems.

With the continuous advance of sequencing technology, the third-generation of high-throughput sequencing technology has been recognized and accepted by more and more researchers, and the latest technology has been timely applied to scientific research practice, so we have a powerful tool on the way to explore life science. In particular, the third-generation high-throughput sequencing technology represented by the Single Molecule real-time Sequencing (SMRT) technology of Pacific Biosciences is widely applied in scientific research. One of the reasons for the popularity of third-generation sequencing technology is that its ultra-long read-length data is extremely convenient for research. For instance, the average reading length of PucBio single-molecule real-time sequencing technology can reach about 10 kb, which enables us to obtain a complete transcript sequence without relying on splicing. Due to the short reading length of second-generation sequencing technology, full-length transcript sequences cannot be obtained directly. In the process of research, a step of splicing transcript must be added. However, splicing is inevitably leading to errors, which limits our research on transcript. Benefit from the advantage of the ultra-long read-length of third-generation high-throughput sequencing technology, we have more and more flexible methods for the study of transcripts. The main processes of lncRNA identification using the third-generation high-throughput sequencing technology can be summarized as the following parts.

Firstly, the RNA of the sequencing object was extracted, and later was reversely transcribed into cDNA. Then, libraries of different lengths and sizes were established according to specific requirements, and finally the libraries were sequenced. After getting the original sequencing data, we still need to filter the data. Insertion of high quality was classified into full-length transcripts and non-full-length transcripts. Full-length transcripts were clustered and non-full-length transcripts were used to correct the clustered full-length transcripts (Gordon et al., 2016) to obtain high-quality transcripts.

The following procedures were similar to the analysis processes after next generation sequencing. Obtained transcripts of high quality were classified by PLEK software, and the transcript sequences of encoded protein and noncoding protein were obtained. Then, sequences with a length greater than or equal to 200 bp were selected from the transcript sequences of non-coding proteins. Then, EMBOSS filter was used to remove the transcript sequences encoded by ORFs with more than 100 amino acids. Finally, BLAST was used to compare the remaining transcripts to the NR protein database to further filter out the gene sequences of encoding proteins and highly reliable lncRNA transcript sequences were obtained.
7 Expectation

With more and more researchers paying attention to lncRNA and the continuous development of sequencing technology, more and more lncRNA have been identified, of which some play a very wide role in the organism. lncRNA is also very important in the generation and development of cancer, which shows a variety of biological functions in cancer, such as epigenetic regulation, inhibition or activation of gene expression. It has been fully recognized that lncRNA plays an important role in organisms. However, compared with other researches, researches on lncRNA are still at the beginning stage so far, because there are still many problems that need to be solved urgently about lncRNA.

There is still no accurate method for the identification of lncRNA. Some transcripts have open reading frames but cannot encode proteins, while some lncRNA have the function of translating small peptides because of the definition of lncRNA. In these two cases, if we take whether proteins can be encoded as the basis to determine whether transcripts are lncRNA, errors will be caused.

Due to the complexity and diversity of lncRNA function and structure, the identification of lncRNA should not only be through several existing methods. We need to develop more effective methods to identify lncRNA and conduct more extensive and systematic studies.

Authors’ contributions

WGF carried out the analysis and drafted the manuscript. THB supervised the study and critically revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This research is supported by “Seven Crops” key projects of National Key Research and Development Program (2016YFD0100305), Fujian ‘Hundred Talents’ Program and National Natural Science Foundation for Young Scholars.

References

Amaral P.P., Clark M.B., Gascoigne D.K., Dinger M.E., and Mattick J.S., 2011, lncRNAdb: a reference database for long noncoding RNAs, Nucleic Acids Research, 39(suppl_1): D146-D151

Burns A.I.S., 2009, Alzheimer’s disease, BMJ: British Medical Journal, 338(7692): 467-471

Chen D.L., Chen L.Z., Lu Y.X., Zhang D.S., Zeng Z.L., Pan Z.Z., Huang P., Wang F.H., Li Y.H., Ju H.Q., and Xu R.H., 2017, Long noncoding RNA XIST expedites metastasis and modulates epithelial-mesenchymal transition in colorectal cancer, Cell Death and Disease, 8(8): e3011

https://doi.org/10.1038/cddis.2017.421

PMid:28837144 PMCid:PMC5596599

Chen G., Wang Z.Y., Wang D.Q., Qiu C.X., Liu M.X., Chen X., Zhang Q.P., Yan G.Y., and Cai Q.H., 2012, LncRNA Disease: a database for long-non-coding RNA-associated diseases, Nucleic Acids Research, 41(D1): D983-D986

https://doi.org/10.1093/nar/gks1099

PMid:23175614 PMCid:PMC3551173

Ding J.H., Lu Q., Ouyang Y.D., Mao H.L., Zhang P.B., Yao J.L., Xu C.G., Li X.H., Xiao J.H., and Zhang Q.F., 2012, A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice, Proc. Natl. Acad. Sci. USA., 109(7): 2654-2659

https://doi.org/10.1073/pnas.1121374109

PMid:22308482 PMCid:PMC3289353

Gordon D., Huddleston J., Chaisson M.J.P., et al., 2016, Long-read sequence assembly of the gorilla genome, Science, 352(6281): e344

https://doi.org/10.1126/science.aag0344

PMid:27034376 PMCid:PMC4920363

Guo C.X., and Liu H., 2014, Structure research and progress of LncRNA, Hebei Yiyao (Hebei Medical Journal), 36(17): 2666-2671

Gupta R.A., Shah N., Wang K.C., et al., 2010, Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis, Nature, 464(7291): 1071-1076

https://doi.org/10.1038/nature08975

PMid:20393566 PMCid:PMC3049919

Hu G., Gong A., Wang Y., Ma S., Chen X., Chen J., Su C., Shibata A., Strauss-Soukup J., Drescher K., and Chen X., 2016, LncRNA-Cox2 promotes late inflammatory gene transcription in macrophages through modulating SW I/SNF-Mediated chromatin remodeling, J. Immunol., 196(6): 2799-2808

https://doi.org/10.4049/jimmunol.1502146

PMid:26880762 PMCid:PMC4779692
Iyer M.K., Niknafs Y.S., Malik R., et al., 2015, The landscape of long noncoding RNAs in the human transcriptome, Nature Genetics, 47(3): 199-208
https://doi.org/10.1038/ng.3192
PMID:25599403 PMCID:PMC4417758

Jeremy E., and Wilusz S.M.F.A., 2008, 3' End Processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA, Cell, 135(5)

Kim D., Pertege A., Trapnell C., Pimentel H., Kelley R., and Salzberg S.L., 2014, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions, Genome Biology, 14(4): R36
https://doi.org/10.1186/gb-2013-14-4-r36
PMID:23641808 PMCID:PMC4053844

Li A., Zhang J., and Zhou Z., 2014, PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme, BMC Bioinformatics. 15: 311
https://doi.org/10.1186/1471-2105-15-311
PMID:25239089 PMCID:PMC4177586

Li T., Xie J., Shen C., et al., 2015, Amplification of long noncoding RNA ZFAS1 promotes metastasis in hepatocellular carcinoma, Cancer Research, 75(15)
https://doi.org/10.1158/0008-5472.CAN-14-3723

Liu C., 2004, NONCODE: an integrated knowledge database of non-coding RNAs, Nucleic Acids Research, 33(Database issue): D112-D115
https://doi.org/10.1093/nar/gki041
PMID:15068158 PMCID:PMC539995

Shi X.F., Sun M., Liu H.B., Yao Y.W., and Song Y., 2013, Long non-coding RNAs: A new frontier in the study of human diseases, Cancer Letters, 339(2): 159-166
https://doi.org/10.1016/j.canlet.2013.06.013
PMID:23791884

Sun L., Luo H.T., Bu D.C., Zhao G.G., Yu K.T., Zhang C.H., Liu Y.N., Chen R.S., and Zhao Y., 2013, Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts, 41(17): e166
Sun L., Zhang Z.H., Bailey T.L., Perkins A.C., Tallack M.R., Xu Z., and Liu H., 2012, Prediction of novel long non-coding RNAs based on RNA-Seq data of mouse Klf1 knockout study, BMC Bioinformatics, 13(1): 331
https://doi.org/10.1186/1471-2105-13-331
PMID:23237380 PMCID:PMC3577497

Tan L., Yu J.T., He N., and Tan L., 2013, Non-coding RNAs in Alzheimer's disease, Molecular Neurobiology, 47(1): 382-393
https://doi.org/10.1007/s12035-013-8359-5
PMID:23054683

Trapnell C., Roberts A., Goff L., et al., 2014, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks, Nature Protocols, 7(3): 562-578
https://doi.org/10.1038/nprot.2012.016
PMID:22830306 PMCID:PMC3334321

Wang K.Y., Ling Y.H., and Zhang X.R., 2015, Progress in association of gene expression and long noncoding RNA, Xumu Shouyi Xuebao (Acta Geneticae et Genomics), 6: 145
https://doi.org/10.3389/fgene.2015.00145
PMID:25954300 PMCID:PMC4407501

Yang J.H., Li J.H., Jiang S., Zhou H., and Qu L.H., 2013, ChIPBase: a database for decoding the transcriptional regulation of long non-coding RNA and microRNA genes from ChIP-Seq data, Nucleic Acids Research, 41(D1): D177-D187
https://doi.org/10.1093/nar/gks1060
PMID:23161675

Yarmishyn A.A., and Kurochkin I.V., 2015, Long noncoding RNAs: a potential novel class of cancer biomarkers, Frontiers in Genetics, 6: 145
https://doi.org/10.3389/fgene.2015.00145
PMID:25954300 PMCID:PMC4407501

Yu W.Q., Gius D., Onyango P., Muldoon-Jacobs K., Karp J., Feinberg A.P., and Cui H.M., 2008, Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA, Nature, 451(7175): 202-206
https://doi.org/10.1038/nature06468
PMID:18185590 PMCID:PMC2743558

70
Zhang Y.C., Liao J.Y., Li Z.Y., Yu Y., Zhang J.P., Li Q.F., Qu L.H., Shu W.S., and Chen Y.Q., 2014, Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice, Genome Biol., 15(12): 512
https://doi.org/10.1186/s13059-014-0512-1
PMid:25517485 PMCid:PMC4253996

Zhao S., Prensner J.R., Erho N., et al., 2014, Identification and validation of the long noncoding RNA SChLAP1 as a prognostic biomarker in prostate cancer, International Journal of Radiation Oncology, Biology, Physics, 90(1): S1-S2
https://doi.org/10.1016/j.ijrobp.2014.06.020

Ørom U.A., Derrien T., Beringer M., et al., 2010, Long noncoding RNAs with enhancer-like function in human cells, Cell., 143(1): 46-58
https://doi.org/10.1016/j.cell.2010.09.001
PMid:20887892 PMCid:PMC4108080