

EVLncRNAs 2.0: an updated database of manually curated functional long non-coding RNAs validated by low-throughput experiments

Bailing Zhou^{1,†}, Baohua Ji^{1,2,†}, Kui Liu¹, Guodong Hu¹, Fei Wang¹, Qingshuai Chen¹, Ru Yu¹, Pingping Huang¹, Jing Ren¹, Chengang Guo¹, Huiying Zhao³, Hongmei Zhang^{1,4}, Dongbo Zhao¹, Zhiwei Li⁵, Qiangcheng Zeng^{1,4}, Jiafeng Yu¹, Yunqiang Bian¹, Zanxia Cao¹, Shicai Xu¹, Yuedong Yang⁶, Yaoqi Zhou^{①1,7,*} and Jihua Wang^{1,*}

¹Shandong Provincial Key Laboratory of Biophysics, Institute of Biophysics, Dezhou University, Dezhou 253023, China, ²College of Physics and Electronic Information, Dezhou University, Dezhou 253023, China, ³Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou 510120, China, ⁴College of Life Science, Dezhou University, Dezhou 253023, China, ⁵Department of General Surgery, Dezhou Municipal Hospital, Dezhou 253012, China, ⁶School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510275, China and ⁷Institute for Glycomics and School of Information and Communication Technology, Griffith University, Gold Coast, QLD 4222, Australia

Received September 14, 2020; Revised October 21, 2020; Editorial Decision October 21, 2020; Accepted October 22, 2020

ABSTRACT

Long non-coding RNAs (lncRNAs) play important functional roles in many diverse biological processes. However, not all expressed lncRNAs are functional. Thus, it is necessary to manually collect all experimentally validated functional lncRNAs (EVLncRNA) with their sequences, structures, and functions annotated in a central database. The first release of such a database (EVLncRNAs) was made using the literature prior to 1 May 2016. Since then (till 15 May 2020), 19 245 articles related to lncRNAs have been published. In EVLncRNAs 2.0, these articles were manually examined for a major expansion of the data collected. Specifically, the number of annotated EVlncRNAs, associated diseases, lncRNA-disease associations, and interaction records were increased by 260%, 320%, 484% and 537%, respectively. Moreover, the database has added several new categories: 8 lncRNA structures, 33 exosomal lncRNAs, 188 circular RNAs, and 1079 drug-resistant, chemoresistant, and stress-resistant lncRNAs. All records have checked against known retraction and fake articles. This release also comes with a highly interactive visual interaction network that facilitates users to track the underlying relations among lncR-

NA, miRNAs, proteins, genes and other functional elements. Furthermore, it provides links to four new bioinformatics tools with improved data browsing and searching functionality. EVLncRNAs 2.0 is freely available at <https://www.sdklab-biophysics-dzu.net/EVLncRNAs2/>.

INTRODUCTION

The discovery of one order of magnitude more transcripts coded for RNAs (i.e. non-coding RNAs) than proteins provided a paradigm shift in our understanding of genome regulation (1,2). Long non-coding RNAs (lncRNAs), in particular, have emerged as key players in essentially every biological process and associated with many human diseases including cancer, cardiovascular and neurodegenerative diseases (3–6). lncRNAs could regulate gene expression in cis and perform functions in trans (7). Accumulating studies indicated that lncRNAs exert their functions by regulating the expression of target genes (8), while genetic variations, such as single-nucleotide polymorphisms (SNPs), could affect lncRNA expression (9). These RNA molecules are perhaps the final missing pieces of the puzzle to complete our understanding of living organisms on the molecular level. With the availability and development of high-throughput technologies, the number of lncRNAs predicted from high-throughput data (HTLncRNA) has increased exponentially (10), as demonstrated from the his-

*To whom correspondence should be addressed. Tel: +86 534 8985 933; Fax: +86 534 8985 884; Email: jhw25336@126.com

Correspondence may also be addressed to Yaoqi Zhou. Tel: +61 7 55 52 82 28; Fax: +61 7 55 52 90 40; Email: yaoqi.zhou@griffith.edu.au

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

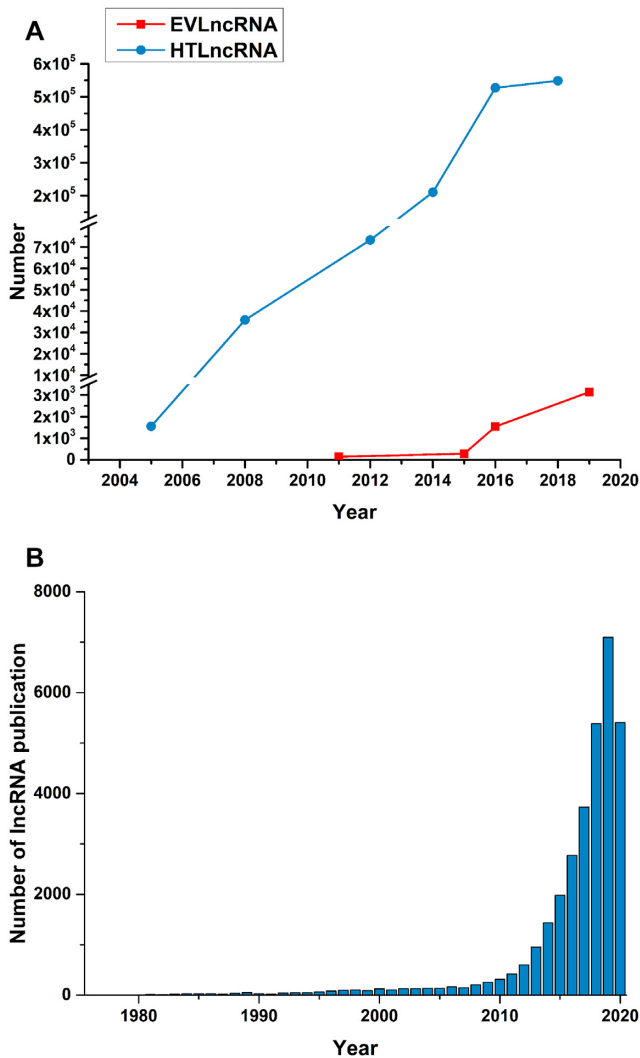


Figure 1. (A) The number of EVLncRNA and HTLncRNA as a function of year (Drawn from the history data of NONCODE (for HTLncRNA), lncRNADB and EVLncRNAs (for EVLncRNA) databases). (B) The number of lncRNA publications per year from 1990 to August 2020 (data from PubMed).

torical data in NONCODE (10–15) (Figure 1A). By comparison, a significant increase in the number of functional lncRNAs (EVLncRNA) validated by low-throughput experimental techniques did not occur until after 2016 (Figure 1A) (16–18). The vast difference between the number of HTLncRNA and the number of EVLncRNA demands a separate and yet comprehensive database for the sequences, structures, and functions along with associated phenotypes for EVLncRNA, similar to what the UniProt database has done for proteins (19). The EVLncRNAs database was the first such database to take up the challenge (18).

Since its publication, the EVLncRNAs database remains the only comprehensive database for experimentally validated functional lncRNAs. With an increasing interest in lncRNAs, the number of associated studies and publications are increasing exponentially, especially in recent years (Figure 1B, from 1990 to August 2020), and novel func-

tions and mechanisms are constantly discovered (20). For example, LINC-PINT is a circular lncRNA encoding a peptide, which suppresses oncogenic transcriptional elongation in glioblastoma (21). Expression signatures of exosomal lncRNAs MALAT1, PCAT-1 and SPRY4-IT1 in urine could serve as non-invasive biomarkers for diagnosing and predicting bladder-cancer recurrence (22). Here, we report a major update of EVLncRNAs to version 2.0 by manually extracting the information from nearly 19 000 articles of the last four years (1 May 2016 to 15 May 2020). The new version now contains 4010 EVLncRNAs in 124 species more than doubled from the previous version. The major increase of data, however, is in functional and disease associations along with new types of lncRNAs. The database is freely available at <https://www.sdcklab-biophysics-dzu.net/EVLncRNAs2/>.

DATA COLLECTION

As in the previous version of EVLncRNAs (18), we only included those functional lncRNAs confirmed by low-throughput experiments such as qRT-PCR, knockdown, Northern blot, and luciferase reporter assays. We also employed the same literature-search procedure (18). Briefly, we first searched the PubMed (20) with keywords matching ‘lncRNA’, ‘long non coding RNA’, ‘long ncRNA’, or ‘long non-code RNA’ along with their plural forms. We found 19 245 publications from 1 May 2016 to 15 May 2020. We manually went over these new articles and extracted necessary information on new as well as existing experimentally validated lncRNAs, including their names, species, function annotations, disease associations, interaction information, experimental methods and samples, expression patterns, and brief descriptions.

To avoid inclusion of the papers that may have been exposed as fraudulent or retracted, we searched PubMed with keywords ‘lncRNA retracted’, combining with the papermill list (23) and the search result of the PubPeer website (for post-publication peer review) (24), and removed 267 papers along with 470 associated entries in the database. This retraction removal will be included as a standard operation in all future updates of the database.

In the interaction category, we collected not only direct physical contacts between biomolecules, but also associations by co-expression and regulation (similar to NPInter database (25)). We used ‘co-expression’ to indicate positively or negatively correlated expression of lncRNA with other molecules in the same physiological process. The ‘regulation’ shows that lncRNA can regulate the expression of other biomolecules in the same physiological process. The ‘binding’ indicates that lncRNA has direct physical contact with other biomolecules. State-of-the-art methods such as ChIRP (26), CHART (27), RAP (28), HyPR (29), CRUIS (30), CLIP (31,32) or CLASH (33) make screening of the direct physical binding interactions between lncRNAs and proteins, RNAs or DNAs. However, whether or not these interactions have any biological significance require further low-throughput experimental validations. Because our database only collects functional lncRNAs, we only include functional interactions demonstrated by subsequent exper-

Table 1. Data collected in EVLncRNAs V2.0 in comparison to that in V1.0 according to the numbers of experimentally validated lncRNAs, species, lncRNA-associated diseases, lncRNA-disease associations and annotated functions and interactions, along with newly included lncRNA structures, exosomal RNA, circular RNAs and drug-resistant, chemoresistant and stress-resistant lncRNAs

Version	#lncRNA	#Species	#Diseases	#Disease association	#Function	#Interaction	#peptide	#Structure	#Exosomal RNA	#CircRNA	#Resist-RNA
1.0	1543	77	338	2324	793	1163	14	0	0	0	0
2.0	4010	124	1082	11 257	1665	6244	37	8	33	188	1079
Inc%	260%	161%	320%	484%	210%	537%	264%				

iments that confirmed functional roles of the binding in diseases or biological processes.

The new additions contain lncRNAs with experimentally solved structures, experimentally validated circular lncRNAs, exosomal lncRNAs, peptide-encoded lncRNAs, and drug-resistant, chemoresistant, and stress-resistant lncRNAs. The structure information of lncRNAs were manually collected from PDB database (34) and presented with the structural views and PDB links. For peptide-encoded lncRNAs, we extracted the names, sequences and lengths of the encoded peptides. Each entry was curated and checked by at least two specialists and proofread non-standard names of lncRNAs according to the HUGO Gene Nomenclature Committee (HGNC) (35). Sequence and position information of lncRNAs were annotated according to NCBI with links pointing to accessions in the NCBI (20) and Ensembl (36) provided if known.

DATABASE CONSTRUCTION

All data in EVLncRNAs were organized using MySQL. The web interface was designed by PHP. The web services were built on Apache Tomcat. Echarts4 (37) was employed to construct interactive interaction networks.

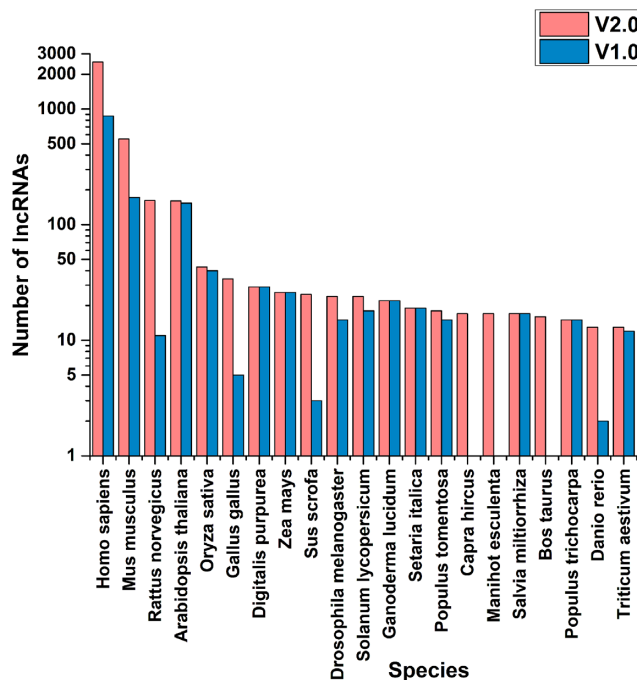
CONTENTS

Data expansion over previous records

Table 1 summarizes the changes of the contents since the last release. More specifically, the number of annotated EVLncRNAs was increased by 260%, (4010 in 124 species), the number of lncRNA-associated diseases by 320%, the number of lncRNA-disease associations by 484%, the number of functions annotated (excluding interactions) by 210%, and the number of interaction records by 537%. The database also includes 37 peptide-encoded lncRNAs, 264% increase from 14 in the previous release. Figure 2 compares the top 20 species for the number of EVLncRNAs in the two versions of the database. Human and mouse remain the top 2 species with the largest number of lncRNAs validated. *Capra hircus* (goat), *Manihot esculenta* (cassava) and *Bos taurus* (cattle) are the newly collected species in top 20 species, indicating the increasing interest in animal and plant lncRNAs.

New contents

In addition to the expansion of the original records, several new contents were added, reflecting new discoveries in the field of lncRNAs. Exosomal lncRNAs were found in exosomes along with their potential roles in drug resistance

**Figure 2.** Top 20 species for the number of experimentally validated lncRNAs in the two versions of the database.

and possible use as diagnostic biomarkers (38,39) (33 collected in the database). lncRNAs were shown circularizable with covalently closed-loop structures without a 5' cap or a 3' Poly A tail (40) (188 collected). Moreover, many lncRNAs were discovered in response to or induced by the drugs, chemicals or stress treatments (1079 added). What is more significant is that the protein databank now contains eight structures solved for lncRNAs.

Interaction network of lncRNAs

As shown in Table 1, the interaction records have the largest expansion with more than five times annotations. To facilitate the navigation of such highly complex interaction network, we implemented Echarts4 to visualize interactions associated with lncRNAs (37). As an example, Figure 3D1 shows the interaction network of lncRNAs with proteins, RNA, DNA and transcription factors as well as their related diseases (if available). When an item is selected, its associated factors and diseases will be highlighted. Figure 3D shows a few examples: illustration of liver cancer-related lncRNAs (Figure 3D2), display of lncRNAs interacting with a microRNA miR-124 (Figure 3D3), and visualization of the diseases and interactive biomolecules related

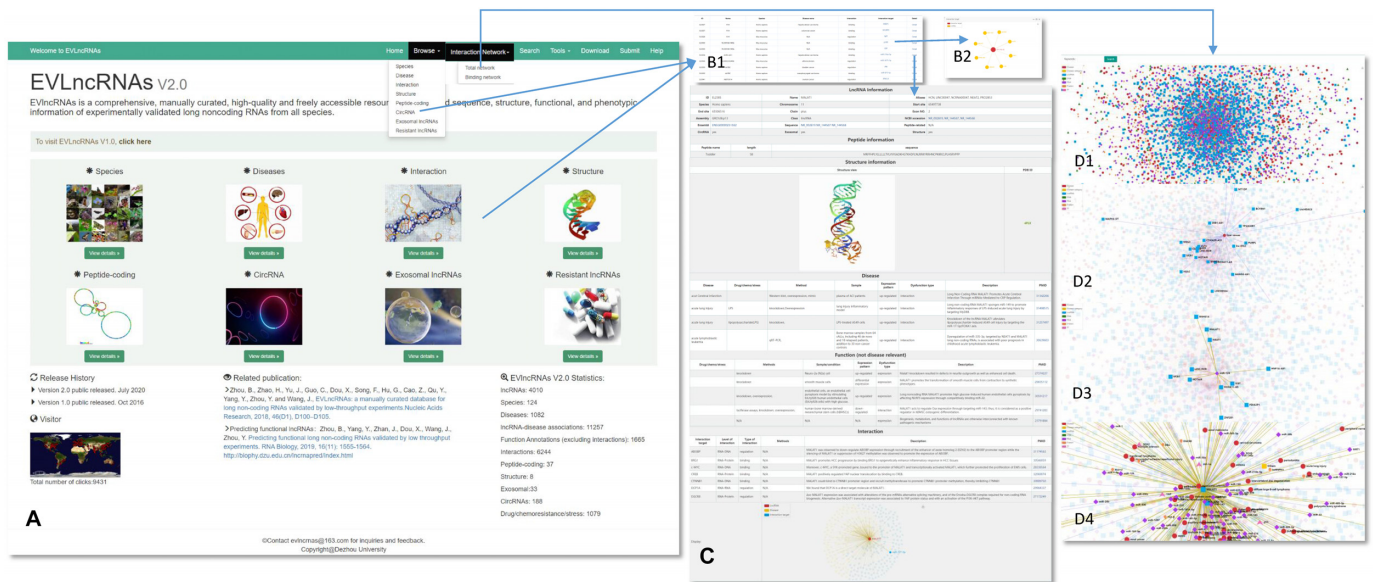


Figure 3. The user interface of EVLncRNAs2. (A) The homepage. (B1) The browse page. (B2) The interaction network graph after clicking on the ‘interaction target’ button. (C) The detailed information after clicking on the ‘Detail’ button. (D1) The highly interactive visual interaction network of lncRNAs with biomolecules and related diseases. (D2) An illustration of liver cancer-related lncRNAs. (D3) A display of lncRNAs interacting with a microRNA miR-124. (D4) Visualization of the diseases and interactive biomolecules related to MALAT1.

to MALAT1 (Figure 3D4). The relations between lncRNAs and diseases and between lncRNAs and biomolecules are distinguished in purple and yellow colors, respectively. Using the search box in the upper left corner, user can search for the lncRNA or interaction target of interest, and examine the small interaction network for the keyword searched. Such a network map allows users to locate overlapped interactions between lncRNAs and biomolecules in different diseases and different biological processes and to discover previously unknown relations between lncRNAs, miRNAs, proteins, genes and other functional elements. The types of interaction networks that can be displayed are binding, regulation and co-expression networks (see Total network in the pull-down menu, Figure 3A). Binding networks are those lncRNAs having direct contact interactions with other functional elements. A detailed help page is included to facilitate the usage.

User interface

To make EVLncRNAs more accessible and user friendly, we improved the web interface (Figure 3). For example, in the ‘Browse’–‘Species’, the records are now classified according to species first, followed by lncRNA names. In the ‘Browse’ page, the information on the same lncRNA function from different articles are now combined to display as a single item for simplifying the view. The ‘Browse’ page now contains ‘interaction target’ and ‘Detail’ buttons (Figure 3B). The ‘interaction target’ button will lead to an interaction network graph which shows involved interaction partners (Figure 3B2). More specific details will be shown after further clicking on the ‘Detail’ button. The ‘Detail’ contains the tables of lncRNA information, peptide information, structure information, disease information, function information, interaction information and a

mini-network of lncRNAs along with their associated interactive biomolecules and diseases (Figures 3C). Only the tables with information available will be displayed. The database ‘Search’ function was also improved substantially. Now users can search for a specific lncRNA, disease or lncRNA-interacting molecule within a specific drop-down option (lncRNAs, Diseases, Interactions) in addition to the previously provided keyword-based search. The keyword search now includes the first round of screening experimental methods such as ChIRP (26), CHART (27), RAP (28), HyPR (29), CRUIS (30), CLIP (31,32) and CLASH (33) in addition to low-throughput techniques. The ‘Tools’ page now includes EVLncRNA-pred for predicting the likelihood of a given lncRNA from high-throughput expression experiments to be functional (41), Context Fold (42) and SPOT-RNA (43) for RNA secondary structure prediction, and RNAsnap2 for RNA solvent accessibility prediction (44).

DISCUSSION AND CONCLUSION

The field of lncRNAs has expanded significantly in the last few years. This is not only because they were found in more species but more significantly in previously unknown locations (exosomes) and structures (in circular forms). They are also induced by various environmental stress factors. The much-needed update of EVLncRNAs reflects a largely unexplored world that is emerging with periodic new discoveries. EVLncRNAs 2.0 attempts to capture this exciting dynamic by manually collecting all recent literature relevant to lncRNAs. This timely update should provide the scientific community with a must-have platform to further explore what seems to be unlimited possibilities of lncRNAs in every aspect of biology.

One new addition in EVLncRNAs 2.0 is the visualization capability of interaction networks. Such a visualiza-

tion tool should be useful for discovering the hubs and islands of the interactions. The binding interaction networks of lncRNAs in various cancers such as hepatocellular carcinoma, gastric cancer, and colorectal cancer have the most extensive data with a network of more than 100 interacting molecules. We found that MALAT1 (Figure 3D4), NEAT1, H19, MEG3, XIST, PVT1 and UCA1 are the lncRNAs with the most interactions with other biomolecules (all >70). MiR-21, miR-34a, miR-206, miR-145, miR-214, miR-129-5p, miR-124 and miR-124-3p are the miRNAs with most known interactions with lncRNAs (all >20). The protein EZH2, which binds with the most lncRNAs, can directly interact with 170 lncRNAs. This is likely due to its associations with many types of cancer (45). lncRNAs can behave differently in different cancers. For example, MALAT1 binds with 13 biomolecules, including 10 miRNAs and 3 proteins (EZH2, MIT, BRG1) in hepatocellular carcinoma but 7 biomolecules, including 6 miRNAs and 1 protein (EZH2) in gastric cancer. Such a difference is interesting and requires further studies for better understanding.

In conclusion, the new release of EVLncRNAs reveals fast accumulations of our knowledge in lncRNAs. We fully expect that such a fast expansion will remain for some times in the near future because what we have explored is only a tip of the iceberg. In future, we will update it more frequently due to unexpected sheer volume of the literature. This will allow the scientific community to stay current with the evolving field. We also encourage all scientists to submit and enhance the impacts of their discoveries by using the submission page.

ACKNOWLEDGEMENTS

We thank Hucheng Tang for his work on building the database. We thank Hui Wang, Tingting Sun, Huan Liu, Hanyu Liu and many other students in the Shandong Key Laboratory of Biophysics of Dezhou University for contributing data extraction.

FUNDING

National Natural Science Foundation of China [62071085, 61801081, 61671107]; Taishan Scholars Program of Shandong Province of China [Tshw201502045]; Talent Introduction Project of Dezhou University of China [320111 to B.Z.]. Funding for open access charge: Natural Science Foundation of China.

Conflict of interest statement. None declared.

REFERENCES

- Djebali,S., Davis,C.A., Merkel,A., Dobin,A., Lassmann,T., Mortazavi,A., Tanzer,A., Lagarde,J., Lin,W., Schlesinger,F. *et al.* (2012) Landscape of transcription in human cells. *Nature*, **489**, 101–108.
- Davis,C.A., Hitz,B.C., Sloan,C.A., Chan,E.T., Davidson,J.M., Gabdank,I., Hilton,J.A., Jain,K., Baymuradov,U.K., Narayanan,A.K. *et al.* (2018) The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.*, **46**, D794–D801.
- Uchida,S. and Dimmeler,S. (2015) Long noncoding RNAs in cardiovascular diseases. *Circ. Res.*, **116**, 737–750.
- Kung,J.T.Y., Colognori,D. and Lee,J.T. (2013) Long noncoding RNAs: past, present, and future. *Genetics*, **193**, 651–669.
- Dong,D., Bao,Z., Zhou,Y., Huang,Z., Cui,Q. and Yang,Z. (2019) LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Res.*, **47**, D1034–D1037.
- Gao,Y., Wang,P., Wang,Y., Ma,X., Zhi,H., Zhou,D., Li,X., Fang,Y., Shen,W., Xu,Y. *et al.* (2019) Lnc2Cancer v2.0: updated database of experimentally supported long non-coding RNAs in human cancers. *Nucleic Acids Res.*, **47**, D1028–D1033.
- Kopp,F. and Mendell,J.T. (2018) Functional classification and experimental dissection of long noncoding RNAs. *Cell*, **172**, 393–407.
- Cheng,L., Wang,P., Tian,R., Wang,S., Guo,Q., Luo,M., Zhou,W., Liu,G., Jiang,H. and Jiang,Q. (2019) LncRNA2Target v2.0: a comprehensive database for target genes of lncRNAs in human and mouse. *Nucleic Acids Res.*, **47**, D140–D144.
- Li,J., Xue,Y., Amin,M.T., Yang,Y., Yang,J., Zhang,W., Yang,W., Niu,X., Zhang,H.-Y. and Gong,J. (2020) ncRNA-eQTL: a database to systematically evaluate the effects of SNPs on non-coding RNA expression across cancer types. *Nucleic Acids Res.*, **48**, D956–D963.
- Fang,S., Zhang,L., Guo,J., Niu,Y., Wu,Y., Li,H., Zhao,L., Li,X., Teng,X., Sun,X. *et al.* (2018) NONCODEV5: a comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Res.*, **46**, D308–D314.
- Liu,C., Bai,B., Skogerbø,G., Cai,L., Deng,W., Zhang,Y., Bu,D., Zhao,Y. and Chen,R. (2005) NONCODE: an integrated knowledge database of non-coding RNAs. *Nucleic Acids Res.*, **33**, D112–D115.
- He,S., Liu,C., Skogerbø,G., Zhao,H., Wang,J., Liu,T., Bai,B., Zhao,Y. and Chen,R. (2008) NONCODE v2.0: decoding the non-coding. *Nucleic Acids Res.*, **36**, D170–D172.
- Bu,D., Yu,K., Sun,S., Xie,C., Skogerbø,G., Miao,R., Xiao,H., Liao,Q., Luo,H., Zhao,G. *et al.* (2012) NONCODE v3.0: integrative annotation of long noncoding RNAs. *Nucleic Acids Res.*, **40**, D210–D215.
- Xie,C., Yuan,J., Li,H., Li,M., Zhao,G., Bu,D., Zhu,W., Wu,W., Chen,R. and Zhao,Y. (2014) NONCODEv4: exploring the world of long non-coding RNA genes. *Nucleic Acids Res.*, **42**, D98–D103.
- Zhao,Y., Li,H., Fang,S., Kang,Y., Wu,W., Hao,Y., Li,Z., Bu,D., Sun,N., Zhang,M.Q. *et al.* (2016) NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Res.*, **44**, D203–D208.
- 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 Res.*, **39**, D146–D151.
- Quek,X.C., Thomson,D.W., Maag,Jesper L.V., Bartonicek,N., Signal,B., Clark,M.B., Gloss,B.S. and Dinger,M.E. (2015) lncRNAdb v2.0: expanding the reference database for functional long noncoding RNAs. *Nucleic Acids Res.*, **43**, D168–D173.
- Zhou,B., Zhao,H., Yu,J., Guo,C., Dou,X., Song,F., Hu,G., Cao,Z., Qu,Y., Yang,Y. *et al.* (2018) EVLncRNAs: a manually curated database for long non-coding RNAs validated by low-throughput experiments. *Nucleic Acids Res.*, **46**, D100–D105.
- Consortium,U. (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.*, **47**, D506–D515.
- Sayers,E.W., Beck,J., Brister,J.R., Bolton,E.E., Canese,K., Comeau,D.C., Funk,K., Ketter,A., Kim,S., Kimchi,A. *et al.* (2020) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **48**, D9–D16.
- Zhang,M., Zhao,K., Xu,X., Yang,Y., Yan,S., Wei,P., Liu,H., Xu,J., Xiao,F., Zhou,H. *et al.* (2018) A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nat. Commun.*, **9**, 4475.
- Zhan,Y., Du,L., Wang,L., Jiang,X., Zhang,S., Li,J., Yan,K., Duan,W., Zhao,Y., Wang,L. *et al.* (2018) Expression signatures of exosomal long non-coding RNAs in urine serve as novel non-invasive biomarkers for diagnosis and recurrence prediction of bladder cancer. *Mol. Cancer*, **17**, 142.
- Papermill Productions (2020) <https://docs.google.com/spreadsheets/d/1KXqTAY4j-jVorFPMD2XRpr76LcIKJ0CVyIvRj0exYQ/edit#gid=0>.
- PubPeer (2020) <https://pubpeer.com/>.
- Teng,X., Chen,X., Xue,H., Tang,Y., Zhang,P., Kang,Q., Hao,Y., Chen,R., Zhao,Y. and He,S. (2020) NPInter v4.0: an integrated database of ncRNA interactions. *Nucleic Acids Res.*, **48**, D160–D165.
- Chu,C., Qu,K., Zhong, Franklin L., Artandi, Steven E. and Chang, Howard Y. (2011) Genomic maps of long noncoding RNA

- occupancy reveal principles of RNA-Chromatin interactions. *Mol. Cell*, **44**, 667–678.
27. Simon, M.D. (2013) Capture hybridization analysis of RNA targets (CHART). *Curr. Protoc. Mol. Biol.*, **101**, 21.25.21–21.25.16.
 28. Munschauer, M., Nguyen, C.T., Sirokman, K., Hartigan, C.R., Hogstrom, L., Engreitz, J.M., Ulirsch, J.C., Fulco, C.P., Subramanian, V., Chen, J. *et al.* (2018) The NORAD lncRNA assembles a topoisomerase complex critical for genome stability. *Nature*, **561**, 132–136.
 29. Spiniello, M., Knoener, R.A., Steinbrink, M.I., Yang, B., Cesnik, A.J., Buxton, K.E., Scalf, M., Jarrard, D.F. and Smith, L.M. (2018) HyPR-MS for multiplexed discovery of MALAT1, NEAT1, and NORAD lncRNA protein interactomes. *J. Proteome Res.*, **17**, 3022–3038.
 30. Zhang, Z., Sun, W., Shi, T., Lu, P., Zhuang, M. and Liu, J.-L. (2020) Capturing RNA–protein interaction via CRUIS. *Nucleic Acids Res.*, **48**, e52.
 31. Grosswendt, S., Filipchyk, A., Manzano, M., Klironomos, F., Schilling, M., Herzog, M., Gottwein, E. and Rajewsky, N. (2014) Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Mol. Cell*, **54**, 1042–1054.
 32. Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.-Y., Cody, N.A.L., Dominguez, D. *et al.* (2020) A large-scale binding and functional map of human RNA-binding proteins. *Nature*, **583**, 711–719.
 33. Helwak, A. and Tollervey, D. (2016) In: Dassi, E. (ed). *Post-Transcriptional Gene Regulation*. Springer, NY, pp. 229–251.
 34. Burley, S.K., Berman, H.M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., Christie, C., Dalenberg, K., Duarte, J.M., Dutta, S. *et al.* (2019) RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.*, **47**, D464–D474.
 35. Bruford, E.A., Braschi, B., Denny, P., Jones, T.E.M., Seal, R.L. and Tweedie, S. (2020) Guidelines for human gene nomenclature. *Nat. Genet.*, **52**, 754–758.
 36. Yates, A.D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R. *et al.* (2020) Ensembl 2020. *Nucleic Acids Res.*, **48**, D682–D688.
 37. Li, D., Mei, H., Shen, Y., Su, S., Zhang, W., Wang, J., Zu, M. and Chen, W. (2018) ECharts: a declarative framework for rapid construction of web-based visualization. *Visual Informatics*, **2**, 136–146.
 38. Naderi-Meshkin, H., Lai, X., Amirkhah, R., Vera, J., Rasko, J.E.J. and Schmitz, U. (2018) Exosomal lncRNAs and cancer: connecting the missing links. *Bioinformatics*, **35**, 352–360.
 39. Li, Y., Yin, Z., Fan, J., Zhang, S. and Yang, W. (2019) The roles of exosomal miRNAs and lncRNAs in lung diseases. *Signal Transduct. Targeted Ther.*, **4**, 47.
 40. Meng, S., Zhou, H., Feng, Z., Xu, Z., Tang, Y., Li, P. and Wu, M. (2017) CircRNA: functions and properties of a novel potential biomarker for cancer. *Mol. Cancer*, **16**, 94.
 41. Zhou, B., Yang, Y., Zhan, J., Dou, X., Wang, J. and Zhou, Y. (2019) Predicting functional long non-coding RNAs validated by low throughput experiments. *RNA Biol.*, **16**, 1555–1564.
 42. Zakov, S., Goldberg, Y., Elhadad, M. and Ziv-ukelson, M. (2011) Rich parameterization improves RNA structure prediction. *J. Comput. Biol.*, **18**, 1525–1542.
 43. Singh, J., Hanson, J., Paliwal, K. and Zhou, Y. (2019) RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning. *Nat. Commun.*, **10**, 5407.
 44. Kumar, A., Singh, J., Paliwal, K., Singh, J. and Zhou, Y. (2020) Single-sequence and profile-based prediction of RNA solvent accessibility using dilated convolution neural network. *Bioinformatics*. btaa652.
 45. Li, B. and Chng, W.-J. (2019) EZH2 abnormalities in lymphoid malignancies: underlying mechanisms and therapeutic implications. *J. Hematol. Oncol.*, **12**, 118.