cncRNAdb: a manually curated resource of experimentally supported RNAs with both protein-coding and noncoding function

Yan Huang[1,†], Jing Wang[2,†], Yue Zhao[3,†], Huafeng Wang[1,†], Tianyuan Liu[2], Yuhe Li[2], Tianyu Cui[2], Weiyi Li[2], Yige Feng[2], Jiaxin Luo[2], Jiaqi Gong[2], Lin Ning[4], Yong Zhang[1,*], Dong Wang[1,2,4,*] and Yang Zhang[1,*]

[1]Shunde Hospital, Southern Medical University (The First People's Hospital of Shunde Foshan), Foshan 528308, China, [2]Department of Bioinformatics, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, China, [3]School of Basic Medical Sciences & Forensic Medicine, Hangzhou Medical College, Hangzhou 310053, China and [4]Dermatology Hospital, Southern Medical University, Guangzhou 510091, China

ABSTRACT

RNA endowed with both protein-coding and noncoding functions is referred to as 'dual-function RNA', 'binary functional RNA (bifunctional RNA)' or 'cncRNA (coding and noncoding RNA)'. Recently, an increasing number of cncRNAs have been identified, including both translated ncRNAs (ncRNAs with coding functions) and untranslated mRNAs (mRNAs with noncoding functions). However, an appropriate database for storing and organizing cncRNAs is still lacking. Here, we developed cncRNAdb, a manually curated database of experimentally supported cncRNAs, which aims to provide a resource for efficient manipulation, browsing and analysis of cncRNAs. The current version of cncRNAdb documents about 2600 manually curated entries of cncRNA functions with experimental evidence, involving more than 2,000 RNAs (including over 1300 translated ncRNAs and over 600 untranslated mRNAs) across over 20 species. In summary, we believe that cncRNAdb will help elucidate the functions and mechanisms of cncRNAs and develop new prediction methods. The database is available at http://www.rna-society.org/cncrnadb/.

INTRODUCTION

Over the past decades, the 'central dogma' of molecular biology has been used to describe messenger RNA (mRNA) as the information-carrying intermediate in protein synthesis (1). Other RNA types, such as transfer RNA (tRNA) and ribosomal RNA (rRNA), are part of the protein-synthesizing machinery as well. However, the rapid development of high-throughput sequencing technologies has revealed the pervasive transcription of eukaryotic genomes (2). Subsequently, various kinds of noncoding RNAs (ncRNAs), such as microRNAs (miRNAs), long noncoding RNAs (lncRNAs) and circular RNAs (circRNAs) were discovered, with functions ranging from epigenetic regulation to transcriptional regulation and signal transduction (3–7). Consequently, the current narrative of RNA involved arbitrarily divided into mRNA and ncRNA according to their protein-coding capacity (8).

However, recent studies have shown that many previously annotated ncRNAs have protein-coding functions (translated ncRNAs) (9,10). For instance, pri-miR171b of *Medicago truncatula* and pri-miR165a of *Arabidopsis thaliana* can produce peptides that enhance the accumulation of their corresponding mature miRNAs (11). The lncRNA LINC00961 harbours a peptide named SPAR that can interact with the lysosomal v-ATPase to negatively regulate mTORC1 activation (12). The circRNA circβ-catenin can activate the Wnt pathway through a 370 aa peptide that it encodes (13). The rapid development of ribosome profiling and mass spectrometry has revealed pervasive translation outside of annotated protein-coding genes (14–18), and some databases have recorded many small open reading frames (smORFs) and micropeptides located in noncoding regions (19–23). All these findings indicate that there is a hidden world of small peptides/proteins produced by ncR-
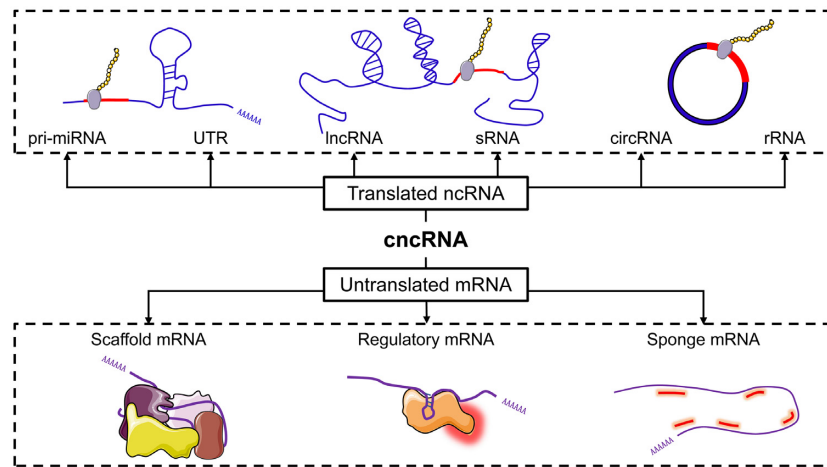
Figure 1. The classification of cncRNAs in cncRNAdb.

NAs waiting to be explored (24). On the other hand, some mRNAs have also been proven to harbour noncoding functions independent of the proteins they encode (25). For example, the mRNA of P53 can interact with MDM2, blocking E3 ligase activity and stimulating translation of the P53 mRNA (26,27). The pre-mRNA of H2B can function as a scaffold in the formation of major nuclear bodies (28). Several mRNAs (CNOT6L, SERINC1 and VAPA) have been shown to regulate PTEN by competing for common miRNAs (29). These clues suggest that some mRNAs in certain situations may be able to fulfil regulatory functions similar to those of ncRNAs (30).

Given the above observations, the class of RNAs endowed with both protein-coding and noncoding functions has been referred to as 'dual-function RNA' (31–33), 'binary functional RNA (bifunctional RNA)' (8,34) or 'cncRNA (this term was proposed as RNA with both coding and noncoding function in 2015)' (35,36). In a sense, the discovery of cncRNA blurred the boundary between 'coding' and 'noncoding' RNAs and led researchers to reconsider the function, evolution and understanding of RNAs (8,25,37,38). Recently, an increasing number of cncRNAs have been identified, including translated ncRNAs and untranslated mRNAs (18,39–41), and some studies began to identify cncRNAs by combining different high-throughput experimental technologies, such as mass spectrometry, ribosome profiling and CRISPR-based screening (18,24,42,43). Here, we developed cncRNAdb (http://www.rna-society.org/cncrnadb/), a manually curated database of experimentally supported cncRNAs, which aims to provide a resource for efficient manipulation, browsing and analysis of cncRNAs. The current version of cncRNAdb documents about 2600 manually curated entries of cncRNA functions with experimental evidence, involving >2000 RNAs (including over 1,300 translated ncRNAs and over 600 untranslated mRNAs) across over 20 species. In summary, cncRNAdb provides a user-friendly interface to query and browse detailed information about these cncRNAs and will be of help in integrating, analysing and predicting cncRNAs, enabling faster structural and functional research of RNA.

DATA COLLECTION AND ORGANIZATION

The cncRNAs in cncRNAdb were curated manually from peer-reviewed literature (before July 2020). First, we retrieved the literature from the PubMed database using the following keywords: 'cncRNA', 'bifunctional RNA', 'dual function RNA', 'translated ncRNA', 'untranslated mRNA', 'regulatory mRNA', 'scaffold mRNA', 'sponge mRNA', 'competitive endogenous mRNA' and 'ncRNA encode peptide'. Then, all retrieved literature was preliminarily reviewed by expert curators to remove irrelevant articles. Meanwhile, to avoid omitting any new cncRNAs, we further checked all the articles that cited at least one of the remaining relevant articles. Only the experimentally supported data were collected in the cncRNAdb. Besides, cncRNAdb collected 504 translated lncRNAs and 53 translated circRNAs from the SmProt (19) and circRNADb (44) databases, respectively.

In the most up to date version of cncRNAdb, all the cncRNAs were divided into two categories: translated ncRNAs, representing ncRNAs that harbour protein-coding functions, and untranslated mRNAs, representing mRNAs that possess noncoding functions that are independent of the proteins they encode (see Figure 1). Then, the translated ncRNAs was further divided into 6 subcategories: lncRNAs, pri-miRNAs, circRNAs, sRNAs, rRNAs and untranslated regions (UTRs). Meanwhile, after careful consideration common perspectives from multiple review articles (25,30,36,45) and the characteristics of the data, the untranslated mRNAs were further divided into three subcategories: regulatory mRNAs, scaffold mRNAs and sponge mRNAs. Regulatory mRNAs refer to mRNAs that play regulatory roles in various biological processes, such as protein binding, action, inhibition and chromatin remodelling (30,36). Scaffold mRNAs refer to the mRNAs that can serve as a scaffold to help form ribonucleoprotein complexes (30,36,45). Sponge mRNA refers to the mRNAs that can act as a competitive endogenous RNAs by competing with other RNAs (30,36,45). In addition, many studies have found some ncRNAs play new regulatory roles independent of their primary function (such as tRNA) (46). It should be
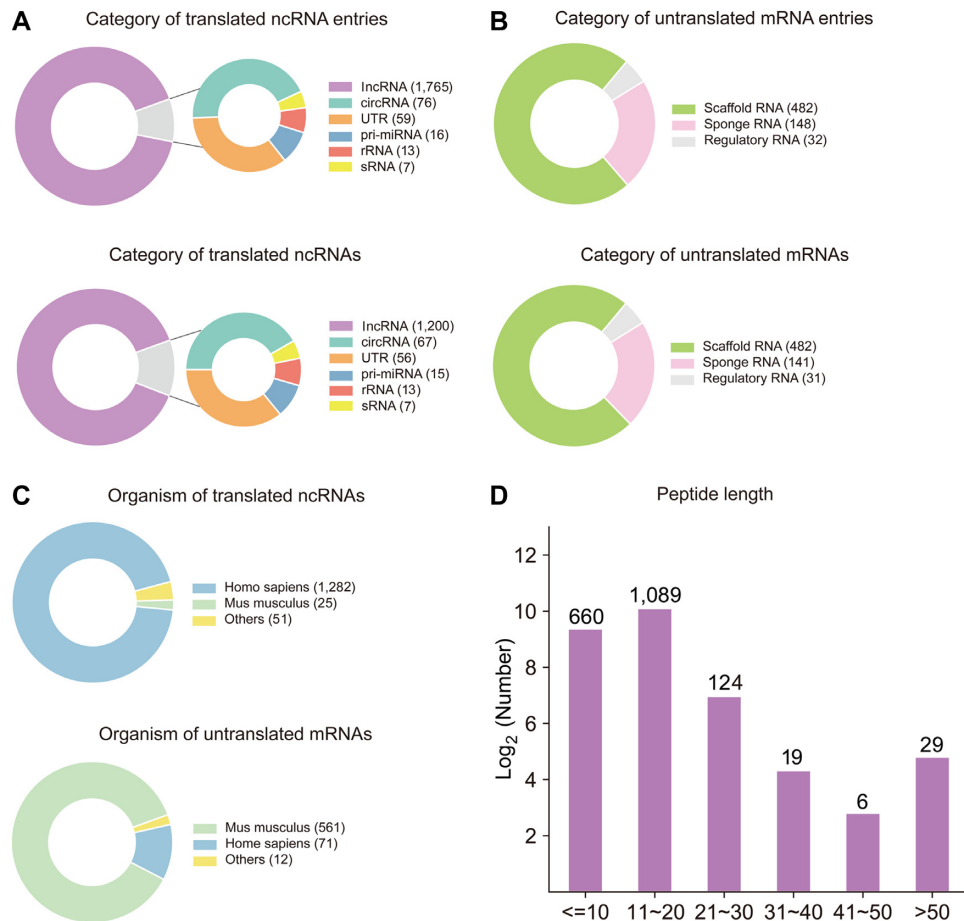
Figure 2. Data statistics of cncRNAdb. (**A**) Subcategory distribution of translated ncRNA entries (upper) and ncRNAs (down). (**B**) Subcategory distribution of untranslated mRNA entries (upper) and mRNAs (down). (**C**) Organismal distribution of translated ncRNAs (upper) and untranslated mRNAs (down). (**D**) Sequence length distribution of peptide encoded by the translated ncRNAs.

noted that these RNAs were not included in cncRNAdb because of lacking evidence for coding capacity.

To unify the cncRNAs into authoritative reference databases, the translated ncRNAs were mapped to the Ensemble database (Ensemble Gene ID) (47), NCBI gene database (Entrez ID), RNAcentral database (RNAcentral ID) (48) and miRbase (miRbase Accession) (49). The untranslated mRNAs were mapped to the NCBI gene database (Entrez ID). The orthologous genes of cncRNA were obtained from Ensemble database (Ensemble Gene ID) (47).

DATABASE CONTENT

This current version of cncRNAdb documents 2598 manually curated cncRNA-associated function entries with experimental evidence (including 1,936 translated ncRNA entries and 662 untranslated mRNA entries) involving 2002 cncRNAs (including 1358 translated ncRNAs and 644 untranslated mRNAs) across 21 species. The distribution of the translated ncRNA entries and ncRNAs in categories is shown in Figure 2A, over 85% of translated ncRNAs are lncRNAs (1200/1358). In addition, we found 144 previously annotated translated ncRNAs that have been reannotated as protein-coding genes in the NCBI and/or Ensemble databases. The distribution of the untranslated mRNA entries and ncRNAs are shown in Figure 2B, there are 32 regulatory mRNA entries (31 mRNAs), 482 scaffold mRNA entries (482 mRNAs) and 148 sponge mRNA entries (141 mRNAs). The organismal distribution of translated ncRNAs and untranslated mRNAs is shown in Figure 2C, and over 95% of cncRNAs are from humans and mice. Moreover, the length distribution of the ncRNA-coding peptides is shown in Figure 2D, over 90% of the ncRNA-coding peptides contain fewer than 20 aa (1749/1927).

DATA QUERYING, SEARCHING AND BROWSING

To make it convenient for users to query and browse data, cncRNAdb provided three different search methods on the 'Search' page (see Figure 3), including 'By Keyword search' (search by inputting keywords of RNA name or disease name, fuzzy search is supported), 'By Locus search' (search by inputting a locus of RNA with the associated option of organism) and 'By Blast search' (search by inputting an RNA or peptide sequence).

A brief summary of the search results is presented in a table. Detailed descriptions, such as peptide sequence and

Figure 3. The 'Search' page and 'Detail' page of cncRNAdb. cncRNAdb provides three different search methods on the 'Search' page, including 'By Keyword search', 'By Locus search' and 'By Blast search'. The 'Detail' page presents more associated information on the cncRNAs, such as basic cncRNA information, peptide information (only for translated ncRNA), supporting experimental evidence, orthologues of the cncRNA, PubMed ID and description of the reference. Moreover, the disease, subcellular localization and interaction information (from OMIM database (50), MNDR v3.0 (51), RNALocate (52) and RNAInter (53)) of the cncRNAs were also provided.

description of the reference, are shown in 'Detail' page by clicking 'more'. The 'Detail' page presents more information associated with the cncRNAs (see Figure 3), such as basic information for the cncRNA, peptide sequence and length (only for translated ncRNA), supporting experimental evidence, orthologous genes (Human, Chimpanzee, Mouse, Drosophila melanogaster and Zebrafish) of the cncRNA, PubMed ID and description of the reference. Moreover, the disease, subcellular localization and interaction information (from OMIM database (50), MNDR v3.0 (51), RNALocate (52) and RNAInter (53)) of the cncRNAs were also provided.

Furthermore, cncRNAdb provides the 'Browse' page to help quickly browse data. Users can browse translated ncRNA and untranslated mRNA in three different ways, 'Browse by RNA type', 'Browse by Organism' and 'Browse by Method'. The related cncRNA information is presented by clicking each entry. Moreover, users can also quickly browse cncRNAs by clicking the schematic plots presented on the 'Home' page.

DISCUSSION AND FUTURE PROSPECTS

At present, many bifunctional RNAs have been identified by multiple experiments (*in vitro* and *in vivo* assay) (12,40). However, there is still a large part of bifunctional

RNAs screened by high-through methods need further experimental validation (24). Systematically collecting, organizing, and grading these RNAs by diverse experimental evidence would be valuable for study on bifunctional RNA. Therefore, we developed cncRNAdb, a manually curated database of experimentally supported cncRNAs containing both translated ncRNAs and untranslated mRNAs. The current version of cncRNAdb documents about 2,600 manually curated entries on cncRNA function and provides stratified and detailed experimental evidence. On the one hand, the various unconventional functions of RNAs stored in cncRNAdb might provide new clues for the exploration and functional characterization of RNA. On the other hand, the continued identification and collection of cncRNAs will definitely extend our former understanding of the role of RNAs in cellular processes and organismal evolution, leading us to reconsider the function, evolution and understanding of RNAs (8,25,37,38). In addition, given the growing evidence that ncRNA translation products play important roles in various biological processes related to diseases (34), cncRNAdb might provide a useful resource for the identification and exploration of new disease biomarkers and/or drug targets. In summary, cncRNAdb provides a user-friendly interface for the query and browsing of detailed information about cncRNAs, and will help elucidate their functions and mechanisms of action for the

development of new prediction methods. In the future, we will continue to collect cncRNAs from both translated ncR-NAs and untranslated mRNAs and update cncRNAdb.

REFERENCES

1. Crick,F. (1970) Central dogma of molecular biology. *Nature* **227**, 561–563.
2. Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C., Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C. *et al.* (2005) The transcriptional landscape of the mammalian genome. *Science*, **309**, 1559–1563.
3. Dunham,I., Kundaje,A., Aldred,S.F., Collins,P.J., Davis,C.A., Doyle,F., Epstein,C.B., Frietze,S., Harrow,J., Kaul,R. *et al.* (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.
4. Hon,C.-C., Ramilowski,J.A., Harshbarger,J., Bertin,N., Rackham,O.J.L., Gough,J., Denisenko,E., Schmeier,S., Poulsen,T.M., Severin,J. *et al.* (2017) An atlas of human long non-coding RNAs with accurate 5′ ends. *Nature*, **543**, 199–204.
5. Zhang,Y., Liu,T., Chen,L., Yang,J., Yin,J., Zhang,Y., Yun,Z., Xu,H., Ning,L., Guo,F. *et al.* (2019) RIscoper: a tool for RNA-RNA interaction extraction from the literature. *Bioinformatics*, **35**, 3199–3202.
6. Li,Y., Wang,C., Miao,Z., Bi,X., Wu,D., Jin,N., Wang,L., Wu,H., Qian,K., Li,C. *et al.* (2015) ViRBase: a resource for virus-host ncRNA-associated interactions. *Nucleic Acids Res.*, **43**, D578–D582.
7. Liu,T.Y., Zhang,Y.C., Lin,Y.Q., Hu,Y.F., Zhang,Y., Wang,D., Wang,Y. and Ning,L. (2020) Exploration of invasive mechanisms via global ncRNA-associated virus-host crosstalk. *Genomics*, **112**, 1643–1650.
8. Hubé,F. and Francastel,C. (2018) Coding and non-coding RNAs, the frontier has never been so blurred. *Front. Genet.*, **9**, 140.
9. Andrews,S.J. and Rothnagel,J.A. (2014) Emerging evidence for functional peptides encoded by short open reading frames. *Nat. Rev. Genet.*, **15**, 193–204.
10. Choi,S.W., Kim,H.W. and Nam,J.W. (2019) The small peptide world in long noncoding RNAs. *Brief Bioinform*, **20**, 1853–1864.
11. Lauressergues,D., Couzigou,J.M., Clemente,H.S., Martinez,Y., Dunand,C., Becard,G. and Combier,J.P. (2015) Primary transcripts of microRNAs encode regulatory peptides. *Nature*, **520**, 90–93.
12. Matsumoto,A., Pasut,A., Matsumoto,M., Yamashita,R., Fung,J., Monteleone,E., Saghatelian,A., Nakayama,K.I., Clohessy,J.G. and Pandolfi,P.P. (2017) mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature*, **541**, 228–232.
13. Liang,W.C., Wong,C.W., Liang,P.P., Shi,M., Cao,Y., Rao,S.T., Tsui,S.K., Waye,M.M., Zhang,Q., Fu,W.M. *et al.* (2019) Translation of the circular RNA circbeta-catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome Biol*, **20**, 84.
14. Ingolia,N.T., Brar,G.A., Stern-Ginossar,N., Harris,M.S., Talhouarne,G.J., Jackson,S.E., Wills,M.R. and Weissman,J.S. (2014) Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep.*, **8**, 1365–1379.
15. van Heesch,S., Witte,F., Schneider-Lunitz,V., Schulz,J.F., Adami,E., Faber,A.B., Kirchner,M., Maatz,H., Blachut,S., Sandmann,C.L.
et al. (2019) The translational landscape of the human heart. *Cell*, **178**, 242–260.
16. Wilhelm,M., Schlegl,J., Hahne,H., Gholami,A.M., Lieberenz,M., Savitski,M.M., Ziegler,E., Butzmann,L., Gessulat,S., Marx,H. *et al.* (2014) Mass-spectrometry-based draft of the human proteome. *Nature*, **509**, 582–587.
17. Friedman,R.C., Kalkhof,S., Doppelt-Azeroual,O., Mueller,S.A., Chovancova,M., von Bergen,M. and Schwikowski,B. (2017) Common and phylogenetically widespread coding for peptides by bacterial small RNAs. *BMC Genomics*, **18**, 553.
18. Chen,J., Brunner,A.D., Cogan,J.Z., Nunez,J.K., Fields,A.P., Adamson,B., Itzhak,D.N., Li,J.Y., Mann,M., Leonetti,M.D. *et al.* (2020) Pervasive functional translation of noncanonical human open reading frames. *Science*, **367**, 1140–1146.
19. Hao,Y., Zhang,L., Niu,Y., Cai,T., Luo,J., He,S., Zhang,B., Zhang,D., Qin,Y., Yang,F. *et al.* (2018) SmProt: a database of small proteins encoded by annotated coding and non-coding RNA loci. *Brief Bioinform*, **19**, 636–643.
20. Liu,W., Xiang,L., Zheng,T., Jin,J. and Zhang,G. (2018) TranslatomeDB: a comprehensive database and cloud-based analysis platform for translatome sequencing data. *Nucleic Acids Res.*, **46**, D206–D212.
21. Wang,H., Yang,L., Wang,Y., Chen,L., Li,H. and Xie,Z. (2019) RPFdb v2.0: an updated database for genome-wide information of translated mRNA generated from ribosome profiling. *Nucleic Acids Res.*, **47**, D230–D234.
22. Olexiouk,V., Van Criekinge,W. and Menschaert,G. (2017) An update on sORFs.org: a repository of small ORFs identified by ribosome profiling. *Nucleic Acids Res.*, **46**, D497–D502.
23. Liu,H., Zhou,X., Yuan,M., Zhou,S., Huang,Y.E., Hou,F., Song,X., Wang,L. and Jiang,W. (2020) ncEP: a manually curated database for experimentally validated ncRNA-encoded proteins or peptides. *J. Mol. Biol.*, **432**, 3364–3368.
24. Lu,S., Zhang,J., Lian,X., Sun,L., Meng,K., Chen,Y., Sun,Z., Yin,X., Li,Y., Zhao,J. *et al.* (2019) A hidden human proteome encoded by 'non-coding' genes. *Nucleic Acids Res.*, **47**, 8111–8125.
25. Li,J. and Liu,C. (2019) Coding or noncoding, the converging concepts of RNAs. *Front. Genet.*, **10**, 496.
26. Candeias,M.M., Malbert-Colas,L., Powell,D.J., Daskalogianni,C., Maslon,M.M., Naski,N., Bourougaa,K., Calvo,F. and Fåhraeus,R. (2008) P53 mRNA controls p53 activity by managing Mdm2 functions. *Nat. Cell Biol.*, **10**, 1098–1105.
27. Naski,N., Gajjar,M., Bourougaa,K., Malbert-Colas,L., Fåhraeus,R. and Candeias,M.M. (2009) The p53 mRNA-Mdm2 interaction. *Cell Cycle*, **8**, 31–34.
28. Shevtsov,S.P. and Dundr,M. (2011) Nucleation of nuclear bodies by RNA. *Nat. Cell Biol.*, **13**, 167–173.
29. Tay,Y., Kats,L., Salmena,L., Weiss,D., Tan,S.M., Ala,U., Karreth,F., Poliseno,L., Provero,P., Di Cunto,F. *et al.* (2011) Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell*, **147**, 344–357.
30. Kloc,M., Foreman,V. and Reddy,S.A. (2011) Binary function of mRNA. *Biochimie*, **93**, 1955–1961.
31. Ulveling,D., Francastel,C. and Hubé,F. (2011) When one is better than two: RNA with dual functions. *Biochimie*, **93**, 633–644.
32. Raina,M., King,A., Bianco,C. and Vanderpool,C.K. (2018) Dual-function RNAs. *Microbiol. Spectrum*, **6**, RWR-0032-2018.
33. Gimpel,M. and Brantl,S. (2017) Dual-function small regulatory RNAs in bacteria. *Mol. Microbiol.*, **103**, 387–397.
34. Zhu,S., Wang,J., He,Y., Meng,N. and Yan,G.R. (2018) Peptides/proteins encoded by non-coding RNA: a novel resource bank for drug targets and biomarkers. *Front. Pharmacol.*, **9**, 1295.
35. Kumari,P. and Sampath,K. (2015) cncRNAs: bi-functional RNAs with protein coding and non-coding functions. *Semin. Cell Dev. Biol.*, **47-48**, 40–51.
36. Sampath,K. and Ephrussi,A. (2016) CncRNAs: RNAs with both coding and non-coding roles in development. *Development*, **143**, 1234–1241.
37. Dinger,M.E., Gascoigne,D.K. and Mattick,J.S. (2011) The evolution of RNAs with multiple functions. *Biochimie*, **93**, 2013–2018.
38. Scheidler,C.M., Kick,L.M. and Schneider,S. (2019) Ribosomal peptides and small proteins on the rise. *Chembiochem*, **20**, 1479–1486.
39. Crerar,H., Scott-Solomon,E., Bodkin-Clarke,C., Andreassi,C., Hazbon,M., Logie,E., Cano-Jaimez,M., Gaspari,M., Kuruvilla,R.

and Riccio,A. (2019) Regulation of NGF signaling by an axonal untranslated mRNA. *Neuron*, **102**, 553–563.

40. Chugunova,A., Loseva,E., Mazin,P., Mitina,A., Navalayeu,T., Bilan,D., Vishnyakova,P., Marey,M., Golovina,A., Serebryakova,M. *et al.* (2019) LINC00116 codes for a mitochondrial peptide linking respiration and lipid metabolism. *PNAS*, **116**, 4940–4945.

41. Pamudurti,N.R., Bartok,O., Jens,M., Ashwal-Fluss,R., Stottmeister,C., Ruhe,L., Hanan,M., Wyler,E., Perez-Hernandez,D., Ramberger,E. *et al.* (2017) Translation of CircRNAs. *Mol. Cell*, **66**, 9–21.

42. Cao,X., Khitun,A., Na,Z., Dumitrescu,D.G., Kubica,M., Olatunji,E. and Slavoff,S.A. (2020) Comparative proteomic profiling of unannotated microproteins and alternative proteins in human cell lines. *J. Proteome Res.*, **19**, 3418–3426.

43. Saghatelian,A. and Couso,J.P. (2015) Discovery and characterization of smORF-encoded bioactive polypeptides. *Nat. Chem. Biol.*, **11**, 909–916.

44. Chen,X., Han,P., Zhou,T., Guo,X., Song,X. and Li,Y. (2016) circRNADb: a comprehensive database for human circular RNAs with protein-coding annotations. *Sci. Rep.*, **6**, 34985.

45. Karapetyan,A.R., Buiting,C., Kuiper,R.A. and Coolen,M.W. (2013) Regulatory roles for long ncRNA and mRNA. *Cancers*, **5**, 462–490.

46. Geslain,R. and Pan,T. (2011) tRNA: vast reservoir of RNA molecules with unexpected regulatory function. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 16489–16490.

47. 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.

48. (2019) RNAcentral: a hub of information for non-coding RNA sequences. *Nucleic Acids Res.*, **47**, D1250–D1251.

49. Kozomara,A., Birgaoanu,M. and Griffiths-Jones,S. (2018) miRBase: from microRNA sequences to function. *Nucleic Acids Res.*, **47**, D155–D162.

50. Amberger,J.S., Bocchini,C.A., Scott,A.F. and Hamosh,A. (2019) OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res.*, **47**, D1038–D1043.

51. Ning,L., Cui,T., Zheng,B., Wang,N., Luo,J., Yang,B., Du,M., Cheng,J., Dou,Y. and Wang,D. (2020) MNDR v3.0: mammal ncRNA–disease repository with increased coverage and annotation. *Nucleic Acids Res.*, doi:10.1093/nar/gkaa707.

52. Zhang,T., Tan,P., Wang,L., Jin,N., Li,Y., Zhang,L., Yang,H., Hu,Z., Zhang,L., Hu,C. *et al.* (2017) RNALocate: a resource for RNA subcellular localizations. *Nucleic Acids Res.*, **45**, D135–D138.

53. Lin,Y., Liu,T., Cui,T., Wang,Z., Zhang,Y., Tan,P., Huang,Y., Yu,J. and Wang,D. (2020) RNAInter in 2020: RNA interactome repository with increased coverage and annotation. *Nucleic Acids Res.*, **48**, D189–D197.