RNAPhaSep: a resource of RNAs undergoing phase separation

Haibo Zhu[1,2,†], Hao Fu[1,2,†], Tianyu Cui[3,†], Lin Ning[3,†], Huaguo Shao[1,4], Yehan Guo[5], Yanting Ke[5], Jiayi Zheng[5], Hongyan Lin[5], Xin Wu[5,6], Guanghao Liu[5,6], Jun He[5,6], Xin Han[4], Wenlin Li[2,7], Xiaoyang Zhao[8], Huasong Lu[9], Dong Wang [ID][3,*], Kongfa Hu[2,*] and Xiaopei Shen[1,5,6,*]

[1]Department of Intelligent Medical Engineering, School of Medical Technology and Engineering, Fujian Medical University, Fuzhou 350122, China, [2]School of Artificial Intelligence and Information Technology, Nanjing University of Chinese Medicine, Nanjing Jiangsu 210023, China, [3]Department of Bioinformatics, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, China, [4]Department of Biochemistry and Molecular Biology, School of Medicine & Holistic Integrative Medicine, Jiangsu Collaborative Innovation Center of Chinese Medicinal Resources Industrialization, Nanjing University of Chinese Medicine, Nanjing 210023, China, [5]Department of Bioinformatics, Fujian Key Laboratory of Medical Bioinformatics, School of Medical Technology and Engineering, Fujian Medical University, Fuzhou 350122, China, [6]Key Laboratory of Ministry of Education for Gastrointestinal Cancer, School of Basic Medical Sciences, Fujian Medical University, Fuzhou 350122, China, [7]Jingwen Library, Nanjing University of Chinese Medicine, Nanjing 210023, China, [8]State Key Laboratory of Organ Failure Research, Department of Developmental Biology, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, China and [9]Zhejiang Provincial Key Laboratory for Cancer Molecular Cell Biology, Life Sciences Institute, Zhejiang University, Hangzhou, Zhejiang 310058, China

ABSTRACT

Liquid-liquid phase separation (LLPS) partitions cellular contents, underlies the formation of membraneless organelles and plays essential biological roles. To date, most of the research on LLPS has focused on proteins, especially RNA-binding proteins. However, accumulating evidence has demonstrated that RNAs can also function as 'scaffolds' and play essential roles in seeding or nucleating the formation of granules. To better utilize the knowledge dispersed in published literature, we here introduce RNAPhaSep (http://www.rnaphasep.cn), a manually curated database of RNAs undergoing LLPS. It contains 1113 entries with experimentally validated RNA self-assembly or RNA and protein co-involved phase separation events. RNAPhaSep contains various types of information, including RNA information, protein information, phase separation experiment information and integrated annotation from multiple databases. RNAPhaSep provides a valuable resource for exploring the relationship between RNA properties and phase behaviour, and may further enhance our comprehensive understanding of LLPS in cellular functions and human diseases.

INTRODUCTION

In addition to canonical membrane-bound organelles, eukaryotic cells contain numerous membraneless organelles (MLOs) that concentrate specific collections of proteins and nucleic acids (1,2). Liquid-liquid phase separation (LLPS), a phenomenon that describes the formation of two immiscible fluids from a single homogeneous mixture, has emerged as a general mechanism to interpret how cells can spatiotemporally create MLOs (3–5). To date, a large number of MLOs have been discovered, including but not limited to stress granules, P bodies and even the nucleolus (6,7). These MLOs have been implicated in a wide range of cellular functions, organizing molecules that act in processes ranging from RNA metabolism to signalling to gene regulation (8–10). Moreover, aberrant MLO behaviours have been linked with multiple human diseases, such as neurodegeneration and cancer (11–13).

*To whom correspondence should be addressed. Tel: +86 591 83569250; Fax: +86 591 83569250; Email: xshen@fjmu.edu.cn
Correspondence may also be addressed to Kongfa Hu. Email: kfhu@njucm.edu.cn
Correspondence may also be addressed to Dong Wang Tel: +86 20 61648279; Fax: +86 20 61648279; Email: wangdong79@smu.edu.cn
†The authors wish it to be known that, in their opinion, the first four authors should be regarded as Joint First Authors.

Identifying the molecules driving or undergoing LLPS is the foundation of understanding the mechanisms of MLOs. Many, but not all, phase-separated biological condensates arise from proteins and RNAs (14). Many nuclear and cytoplasmic condensates are rich in RNAs and RNA-binding proteins (RBPs), which play roles in LLPS (15,16). The roles of proteins in condensation have been well studied (17). However, less attention has been paid to the contribution of RNA to LLPS (18). As an anionic polymer, RNA is an excellent platform for achieving multivalency and accommodating RBPs (19,20). Peptides in condensates usually contain low-complexity domains (LCDs) or intrinsically disordered regions (IDRs) those enabling weak and multivalent interactions to promote liquid-like properties, and a similar role for unstructured sequences in RNAs potentially contribute to RNA-driven phase separation (21,22). Recent findings have confirmed that RNAs can function as 'scaffolds' and play essential roles in seeding or nucleating the formation of MLOs (23,24). Experimental evidence has demonstrated that RNA self-assembly contributes to stress granule formation and defines the stress granule transcriptome (25). Furthermore, apart from driving LLPS, diverse RNA properties, such as composition, length, structure, modification and expression level, can modulate the biophysical features of native condensates, including their size, shape, viscosity, liquidity and surface tension (26–28). The molecular mechanisms under these associations remain unknown (14).

Over the last five years, a large number of publications have reported cases of RNAs being involved in LLPS, and the number is still increasing rapidly. Five databases focused on proteins undergoing LLPS have been created, including LLPSDB, PhaSePro, PhaSepDB, DrLLPS and RNAGranuleDB (29–33). However, centralized resources on newly reported phase separation related RNAs are still lacking. In particular, RNA self-assembly LLPS events have not been recorded in any published phase separation database. To fill this gap, here we introduce RNAPhaSep, a database using RNA as its core and focusing on various properties of RNAs and their roles in the phase separation process. After careful manual collation, a total of 1113 entries of experimentally validated RNA self-assembly or RNA and protein co-involved phase separation events were included in RNAPhaSep. Our RNAPhaSep database provides a convenient interface to help users browse, search and download RNA-related phase separation entries.

DATA COLLECTION AND PROCESSING

RNAPhaSep was constructed based on the curated information derived from published literature and the RNALocate database (34). Literature mining was performed via retrieving from PubMed using the following keywords: ((((phase transition) OR (phase separation) OR (membraneless organelles) OR (biomolecular condensates)) AND RNA) AND cell). A total of 4,804 publications before 30 June 2021, were extracted (Figure 1A). For review articles, we read through each manuscript, extracted sentences describing the RNA-related phase separation events and downloaded the corresponding research

articles for curation. As some of the known MLOs, such as P-bodies, had been extensively studied before the emergence of the LLPS concept, which underlies MLOs formation, the RNALocate database (34), which contains subcellular localization of RNA in MLOs, was used (Figure 1A). For each RNA record related to MLOs in RNALocate (34), an original research article was extracted by PMID. To obtain relevant publications that describe MLOs or phase separation and related RNAs, we manually checked the abstracts or full texts of all these articles. During the curation process, we sought to collect phase separation-associated RNAs and as much helpful information as possible, such as original supporting sentences, RNA sequence, mutation, phase separation experimental conditions, phase diagrams, compositions and corresponding cell lines or tissues used for experimentation (Figure 1B).

RNAPhaSep integrates two types of RNA IDs, including NCBI Gene IDs (35) and RNAcentral identifiers (36). For several rRNAs and snoRNAs, which could not be found in NCBI, an RNAcentral identifier was supplied. The graph of each RNA's structure was generated using the RNAfold server on ViennaRNA web services (37). The LCD and IDR information for each protein were collected from MobiDB (38) or PONDR (39).

The molecular properties of RNAs are essential for understanding their potential phase behaviour. For each natural RNA, the description in NCBI (35), molecular function in Gene Ontology (GO) (40), subcellular localization in RNALocate (34), interaction neighbour in RNAinter (41) and associated disease in DisGeNET (42) and OMIM (43) were all integrated as RNA annotation information (Figure 1B). Designed RNAs were divided into four subclasses based on sequence characteristics. In order of preference, there were: 'poly RNA' if the sequence was just a repetition of a single nucleotide, 'repeat RNA' if the sequence was the duplication of a fragment with at least two different nucleotides, 'nucleotide rich RNA' if the sequence was enriched with two specific types of nucleotides; and the rest are classified as 'irregular RNA'.

It is important to emphasise that RNAPhaSep is concerned with cases where the RNA itself or together with other components (DNA or protein) was experimentally validated *in vitro* or *in vivo* for LLPS. Thus, the systems with only proteins or DNAs as the main components were excluded. Moreover, systems with the mixtures of RNA, such as total mRNA, were included, as RNA was the only component driving LLPS in these systems. After sorting out the records, we noticed no DNA and RNA co-involved reports, which may be due to the limited research in this area. We may include this type of LLPS event in the updated version of RNAPhaSep.

The state of phase separation can vary dynamically in a wide range, from liquid to solid. Four states, including 'solute', 'liquid', 'gel' and 'solid', were used to define the morphological characteristics of phase separation (1). Changes in experimental conditions can lead to a phase transition, such as from liquid to gel, then 'liquid, gel' was recorded as the morphology of this phase separation event. RNA-related phase separation events curated in RNAPhaSep were verified by experiments, including reconstituting LLPS condensation *in vitro* and examining droplet formation
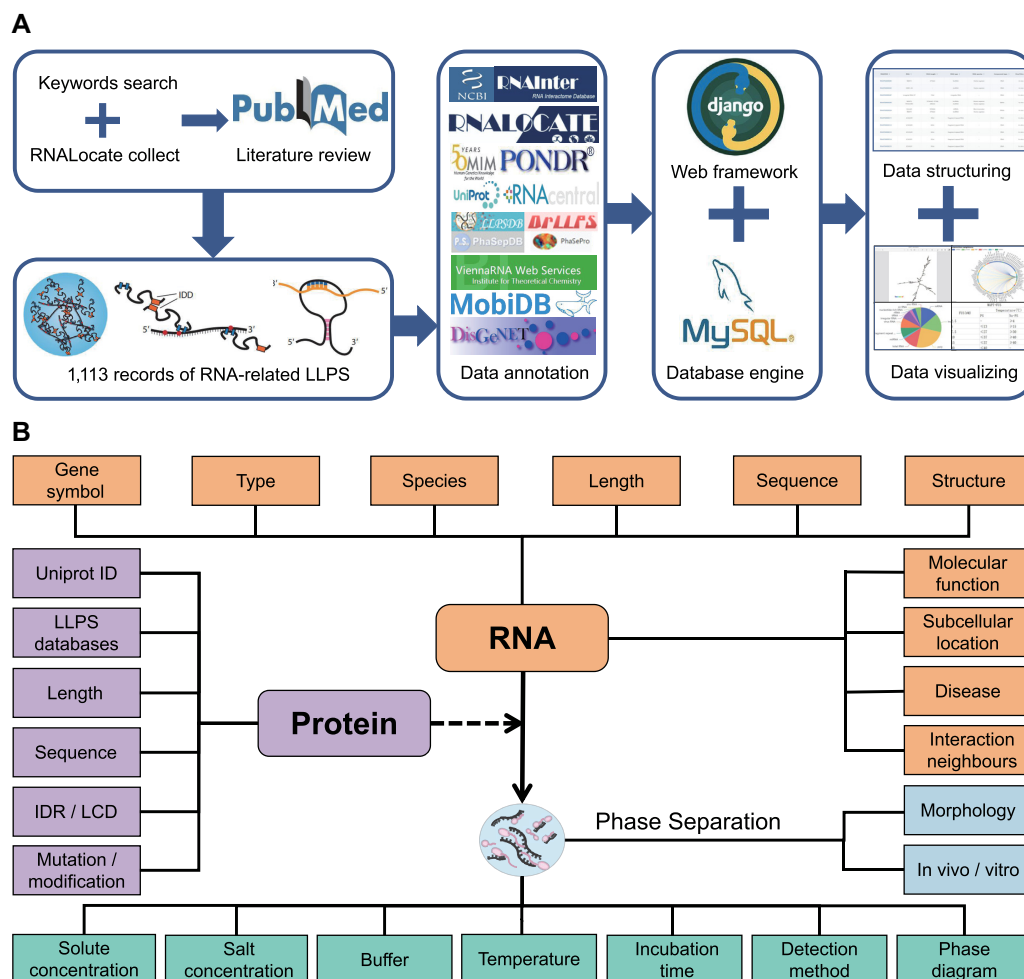
Figure 1. Overview of RNAPhaSep database. (**A**) Workflow of building RNAPhaSep. First, we searched the literature in PubMed by keywords or PMIDs extracted from MLO-related RNA in the RNALocate database, and collected 1113 RNA-protein and RNA self-assembly phase separation events. Then we integrated annotations in 15 public data resources and presented structured and visualized data through web architecture and database servers. (**B**) The data structure of RNAPhaSep. Different types of information are distinguished by colour.

by immunofluorescence *in vivo*. RNAs detected by high-throughput methods were excluded.

DATA CONTENT

As of August 2021, RNAPhaSep included 1113 curated entries about RNA self-assembly or RNA and protein co-involved phase separation events, involving 325 non-redundant RNAs of 22 organisms (Figure 2A). We consolidated the entries with RNA or RNA plus protein names to reduce the data redundancy and assigned all entries to 628 unique RNAPSIDs. RNA properties such as composition, species, classification, sequence, length, structure, subcellular localization, RNA interaction neighbours, related molecular functions and diseases were collected and organized for each RNA.

For different RNA types, we classified RNAs into natural and designed RNA (Figure 2B and C). Although natural RNAs often have diverse annotation information, which could comprehensively describe their molecular functions in the cell, the impact of RNA sequence on phase separa-

tion events is more clearly demonstrated by designed RNAs due to their designability and low complexity.

In vitro experiments are very important as their simplified processes can help researchers clearly identify the conditions involved in phase separation. Researchers can simulate intracellular phase separation events by controlling various experimental details, such as salt concentration, buffer, temperature, and pH. Most importantly, the components involved in these experiments are known, so for *in vitro* experiments, we classified entries as 'RNA(s)', 'RNA + protein(s)' and 'RNA(s) + protein(s)' (Figure 2D). The morphological distribution of phase-separated records was demonstrated in Figure 2E.

The RNA sequences of *in vivo* records, mostly from the annotation in NCBI (35) or RNACentral (36), were not validated in the *in vivo* phase separation experiment. Thus, 184 RNAs from *in vitro* records were used for sequence and length analysis. The sequence analysis demonstrated that LLPS related RNAs' sequences were enriched with adenine and uracil, which together accounted for 58.2%, 58%, 67.5% for all RNA, natural RNA and designed RNA (Figure 2F). Motif analysis of the RNA sequence by STREME (44) was
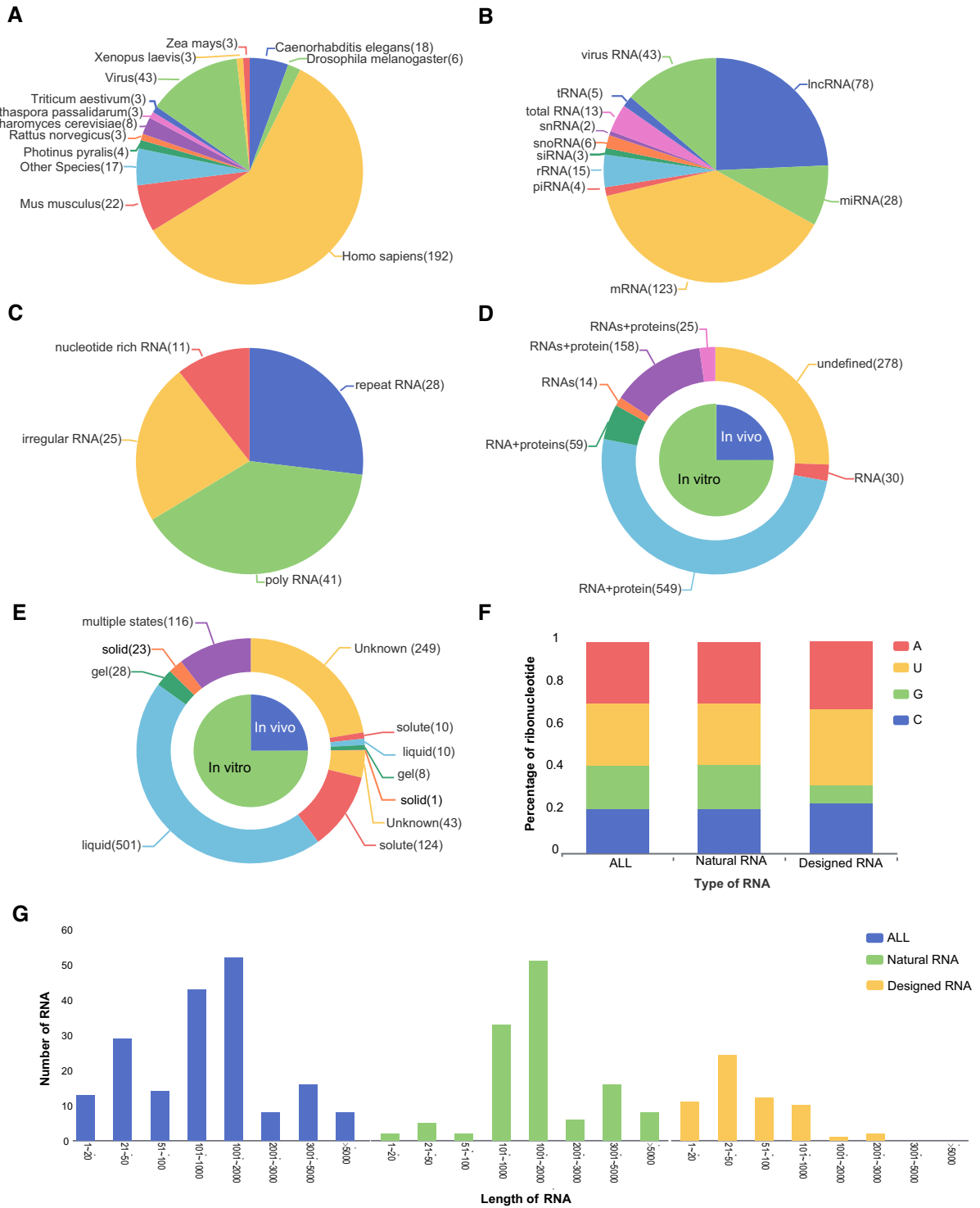
Figure 2. Data statistics on RNAPhaSep. (**A**) Organism distribution and (**B**) category distribution of natural RNAs. (**C**) Category distribution of designed RNAs. (**D**) Distribution of *in vivo* and *in vitro* experiments (inner circle) and composition distribution *in vitro* (outer ring) and (**E**) distribution of morphology for *in vitro* and *in vivo* LLPS events (outer ring). The 'multiple states' means that the phase separation event contains at least two states like 'liquid, gel'. (**F**) Percentage of ribonucleotide and (**G**) distribution of length for RNAs from *in vitro* experiments.

Figure 3. An illustration of the RNAPhaSep website. (**A**) Search the database by options (like RNA symbol, protein name) or RNA sequence in the Search/Blast page. (**B**) Dataset can be browsed by two distinct classifications (*In vivo/vitro* or RNA type) in the browse page. RNAPSID is provided in the search result and browse page for jumping to (**C**) phase separation detail page consisting of three main modules (RNA information, protein information and phase separation experiment and description), including visualization of an RNA sequence, structure and a Phase diagram. (**D**) RNA annotation page contains comprehensive information annotated from public databases and visualization of the RNA interaction network.

performed to discover common sequence elements. However, we have not found any significant motifs, which may be due to a relatively small number of RNAs. Designed RNA favours shorter RNA sequences than natural RNA in sequence length distribution (Figure 2G).

A novel coronavirus, SARS-CoV-2, has caused the ongoing worldwide COVID-19 pandemic. Scientists have found that viral genomic RNAs can form phase-separated droplets with nucleocapsid proteins and that these droplets become solid-like structures as the RNA length increases (45,46). Thus, this LLPS morphology is utterly dependent on the length and concentration of the given RNAs. To date, 76 phase-separated entries of SARS-CoV-2 are present in the RNAPhaSep database.

WEB INTERFACE

A user-friendly and fully functional website has been developed for searching, browsing and downloading RNA-related phase separation data. This database includes eight modules, including *Home, Search/Blast, Browse, Submit, Download, Statistics, Help* and *Contact* modules.

For the convenience of the user for searching the data, RNAPhaSep provides two different searching ways on the Search/Blast page, including 'By options' (search by the combination of keywords, component type, species and RNA type) and 'By RNA sequence' (search by inputting RNA sequence) (Figure 3A). For 'Search by options', we have provided three typical examples. By clicking on the example button, the options information is automatically applied, and then by clicking on the 'search' button, the 'Search Result' is presented in a table format. The 'Search by RNA sequence' module enables users to identify the sequence similarity between their target RNA and the LLPS-related RNA stored in the database. The 'Search/Blast' module can help the user quickly screen how their interested RNA contributes to LLPS under available conditions with or without any partners.

The data sources are briefly described on the 'Browse' page, which has a table containing all entries that can be divided into different subsets by *in vivo/vitro* and RNA type (Figure 3B). Users can click on a unique RNAPSID to navigate to the 'Phase Separation Details Page', which includes various descriptions of the involved RNA, protein and experimental condition details (Figure 3C). For RNA, the phase separation-related RNA sequence and structure were demonstrated on the page. The corresponding Uniprot ID, IDR, LCD, mutation, modification, and sequence

information are listed for each involved protein, if available. For each involved protein, in order to allow users to easily obtain a wide range of information on protein-related phase separation, we provided direct links to the protein detail page of known protein-related LLPS databases, including LLPSDB, PhaSePro and DrLLPS (29–30,32). As PhaSepDB did not provide the unique link for each protein, the link to its 'Browse Page' was supplied (31). The same protein involved in different LLPS experiments may have different sequences or structures, such as wild-type, mutant or post-transcriptional modified. These differences potentially affect the LLPS conditions. Thus, the 'State' of the protein is used to discriminate different protein sequences or structures and corresponding experimental conditions (Figure 3C). For each LLPS experiment, a phase diagram is presented in a graphical form, if available. For the natural RNAs, which have extensive information, clicking on the RNA symbol will take the user to the RNA annotation page, which integrates RNA description (35), molecular function (40), subcellular location (34), RNA interaction network (top 100 extracted from RNAinter) (41) and associated diseases (42–43) (Figure 3D). A detailed tutorial for the usage of the database can be found on the 'Help' page.

On the 'Submission' page, we supplied three choices for different data sources, including published, preprint and unpublished data. The submission of published or preprint data will be routinely reviewed manually and formatted into database form with PMID or tagged as 'preprint'. For the unpublished data, authors may not want to share the original experiment evidence, and we will tag the data by 'unpublished'. We added the reminder on the phase separation page that 'Entries with 'unpublished' tag have relatively lower reliability'.

Database implementation

The database is stored in the MySQL v5.7 (https://www.mysql.com) database engine. The web framework was constructed on Django v3.2 (https://www.djangoproject.com) and ran on a CentOS Linux operating system server. We have tested it on Google Chrome, Mozilla Firefox, Microsoft Edge and Apple Safari browsers. The RNAPhaSep database is freely available to the research community online at http://www.rnaphasep.cn.

CONCLUSION

Here, we present a novel resource on RNA-related phase separation, RNAPhaSep, generated from information obtained from the literature and RNALocate database. It contains 1113 experimentally validated RNA self-assembly or RNA and protein co-involved phase separation events, helping and guiding researchers to perform further studies related to LLPS. RNAPhaSep was designed explicitly for RNA-related phase separation, and we believe it will be a handy tool for researchers in this field. From recently published perspectives and review articles, we noticed that the role of RNA in driving phase separation has attracted increasing researchers' attention and interest. The establishment of RNAPhaSep is only the first step, and we will continue to expand and improve it to satisfy more requirements

in this field. We are now collecting a reliable LLPS corpus for developing a text mining system, which can automatically extract LLPS information from biomedical literature in PubMed. To ensure that all data from literature has consistently high reliability, these automatically extracted records will still need to be manually curated. After applying this system, we can update RNAPhaSep more frequently.

REFERENCES

1. Boeynaems,S., Alberti,S., Fawzi,N.L., Mittag,T., Polymenidou,M., Rousseau,F., Schymkowitz,J., Shorter,J., Wolozin,B., Van Den Bosch,L. *et al.* (2018) Protein phaseseparation: a new phase in cell biology. *Trends Cell Biol.*, **28**, 420–435.
2. Bentley,E.P., Frey,B.B. and Deniz,A.A. (2019) Physical chemistry of cellular liquid-phase separation. *Chemistry*, **25**, 5600–5610.
3. Hyman,A.A., Weber,C.A. and Jülicher,F. (2014) Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.*, **30**, 39–58.
4. Smith,J., Calidas,D., Schmidt,H., Lu,T., Rasoloson,D. and Seydoux,G. (2016) Spatial patterning of P granules by RNA-induced phase separation of the intrinsically-disordered protein MEG-3. *Elife*, **5**, e21337.
5. Alberti,S., Gladfelter,A. and Mittag,T. (2019) Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell*, **176**, 419–434.
6. Gomes,E. and Shorter,J. (2019) The molecular language of membraneless organelles. *J. Biol. Chem.*, **294**, 7115–7127.
7. Banani,S.F., Lee,H.O., Hyman,A.A. and Rosen,M.K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.*, **18**, 285–298.
8. Yang,Y., Hsu,P.J., Chen,Y.-S. and Yang,Y.-G. (2018) Dynamic transcriptomic m(6)A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell Res.*, **28**, 616–624.
9. Taliaferro,J.M., Lambert,N.J., Sudmant,P.H., Dominguez,D., Merkin,J.J., Alexis,M.S., Bazile,C. and Burge,C.B. (2016) RNA sequence context effects measured in vitro predict in vivo protein binding and regulation. *Mol. Cell*, **64**, 294–306.
10. Boehning,M., Dugast-Darzacq,C., Rankovic,M., Hansen,A.S., Yu,T., Marie-Nelly,H., McSwiggen,D.T., Kokic,G., Dailey,G.M., Cramer,P. *et al.* (2018) RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.*, **25**, 833–840.
11. Mackenzie,I.R., Nicholson,A.M., Sarkar,M., Messing,J., Purice,M.D., Pottier,C., Annu,K., Baker,M., Perkerson,R.B., Kurti,A. *et al.* (2017) TIA1 mutations in amyotrophic lateral sclerosis and frontotemporal dementia promote phase separation and alter stress granule dynamics. *Neuron*, **95**, 808–816.
12. Boija,A., Klein,I.A. and Young,R.A. (2021) Biomolecular condensates and cancer. *Cancer Cell*, **39**, 174–192.
13. Alexander,E.J., Ghanbari Niaki,A., Zhang,T., Sarkar,J., Liu,Y., Nirujogi,R.S., Pandey,A., Myong,S. and Wang,J. (2018) Ubiquilin 2 modulates ALS/FTD-linked FUS–RNA complex dynamics and stress granule formation. *Proc. Natl Acad. Sci. U.S.A.*, **115**, E11485–E11494.
14. Roden,C. and Gladfelter,A.S. (2021) RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.*, **22**, 183–195.

15. Hofweber,M., Hutten,S., Bourgeois,B., Spreitzer,E., Niedner-Boblenz,A., Schifferer,M., Ruepp,M.-D., Simons,M., Niessing,D., Madl,T. *et al.* (2018) Phase separation of FUS is suppressed by its nuclear import receptor and arginine methylation. *Cell*, **173**, 706–719.

16. Murakami,T, Qamar,S., Lin,J.Q., Schierle,G.S.K., Rees,E., Miyashita,A., Costa,A.R., Dodd,R.B., Chan,F.T.S., Michel,C.H. *et al.* (2015) ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. *Neuron*, **88**, 678–690.

17. Lin,Y., Protter,D.S.W., Rosen,M.K. and Parker,R. (2015) Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell*, **60**, 208–219.

18. Kroschwald,S., Maharana,S., Mateju,D., Malinovska,L., Nüske,E., Poser,I., Richter,D. and Alberti,S. (2015) Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *Elife*, **4**, e06807.

19. Guo,L., Kim,H.J., Wang,H., Monaghan,J., Freyermuth,F., Sung,J.C., O'Donovan,K., Fare,C.M., Diaz,Z., Singh,N. *et al.* (2018) Nuclear-import receptors reverse aberrant phase transitions of RNA-binding proteins with prion-like domains. *Cell*, **173**, 677–692.

20. Yoshizawa,T., Ali,R., Jiou,J., Fung,H.Y.J., Burke,K.A., Kim,S.J., Lin,Y., Peeples,W.B., Saltzberg,D., Soniat,M. *et al.* (2018) Nuclear import receptor inhibits phase separation of FUS through binding to multiple sites. *Cell*, **173**, 693–705.

21. Langdon,E.M., Qiu,Y., Ghanbari Niaki,A., McLaughlin,G.A., Weidmann,C.A., Gerbich,T.M., Smith,J.A., Crutchley,J.M., Termini,C.M., Weeks,K.M. *et al.* (2018) mRNA structure determines specificity of a polyQ-driven phase separation. *Science*, **360**, 922–927.

22. Ries,R.J., Zaccara,S., Klein,P., Olarerin-George,A., Namkoong,S., Pickering,B.F., Patil,D.P., Kwak,H., Lee,J.H. and Jaffrey,S.R. (2019) m(6)A enhances the phase separation potential of mRNA. *Nature*, **571**, 424–428.

23. Jain,A. and Vale,R.D. (2017) RNA phase transitions in repeat expansion disorders. *Nature*, **546**, 243–247.

24. Bounedjah,O., Desforges,B., Wu,T.-D., Pioche-Durieu,C., Marco,S., Hamon,L., Curmi,P.A., Guerquin-Kern,J.-L., Piétrement,O. and Pastré,D. (2014) Free mRNA in excess upon polysome dissociation is a scaffold for protein multimerization to form stress granules. *Nucleic Acids Res.*, **42**, 8678–8691.

25. Van Treeck,B., Protter,D.S.W., Matheny,T., Khong,A., Link,C.D. and Parker,R. (2018) RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 2734–2739.

26. Elbaum-Garfinkle,S., Kim,Y., Szczepaniak,K., Chen,C.C.-H., Eckmann,C.R., Myong,S. and Brangwynne,C.P. (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 7189–7194.

27. Zhang,H., Elbaum-Garfinkle,S., Langdon,E.M., Taylor,N., Occhipinti,P., Bridges,A.A., Brangwynne,C.P. and Gladfelter,A.S. (2015) RNA controls PolyQ protein phase transitions. *Mol. Cell*, **60**, 220–230.

28. Giustiniani,A., Drenckhan,W. and Poulard,C. (2017) Interfacial tension of reactive, liquid interfaces and its consequences. *Adv. Colloid Interface Sci.*, **247**, 185–197.

29. Li,Q., Peng,X., Li,Y., Tang,W., Zhu,J., Huang,J., Qi,Y. and Zhang,Z. (2020) LLPSDB: a database of proteins undergoing liquid-liquid phase separation in vitro. *Nucleic. Acids. Res.*, **48**, D320–D327.

30. Mészáros,B., Erdős,G., Szabó,B., Schád,É., Tantos,Á., Abukhairan,R., Horváth,T., Murvai,N., Kovács,O.P., Kovács,M.

et al. (2020) PhaSePro: the database of proteins driving liquid-liquid phase separation. *Nucleic Acids Res.*, **48**, D360–D367.

31. You,K., Huang,Q., Yu,C., Shen,B., Sevilla,C., Shi,M., Hermjakob,H., Chen,Y. and Li,T. (2020) PhaSepDB: a database of liquid-liquid phase separation related proteins. *Nucleic Acids Res.*, **48**, D354–D359.

32. Ning,W., Guo,Y., Lin,S., Mei,B., Wu,Y., Jiang,P., Tan,X., Zhang,W., Chen,G., Peng,D. *et al.* (2020) DrLLPS: a data resource of liquid-liquid phase separation in eukaryotes. *Nucleic Acids Res.*, **48**, D288–D295.

33. Youn,J.-Y., Dyakov,B.J.A., Zhang,J., Knight,J.D.R., Vernon,R.M., Forman-Kay,J.D. and Gingras,A.-C. (2019) Properties of stress granule and P-body proteomes. *Mol. Cell*, **76**, 286–294.

34. Cui,T., Dou,Y., Tan,P., Ni,Z., Liu,T., Wang,D., Huang,Y., Cai,K., Zhao,X., Xu,D. *et al.* (2021) RNALocate v2.0: an updated resource for RNA subcellular localization with increased coverage and annotation. *Nucleic Acids Res.*, https://doi.org/10.1093/nar/gkab825.

35. Resource Coordinators,NCBI, Agarwala,R., Barrett,T., Beck,J., Benson,D.A., Bollin,C., Bolton,E., Bourexis,D., Brister,J.R., Bryant,S.H. *et al.* (2018) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **46**, D8–D13.

36. Consortium,The RNAcentral, Sweeney,B.A., Petrov,A.I., Burkov,B., Finn,R.D., Bateman,A., Szymanski,M., Karlowski,W.M., Gorodkin,J., Seemann,S.E. *et al.* (2019) RNAcentral: a hub of information for non-coding RNA sequences. *Nucleic Acids Res.*, **47**, D221–D229.

37. Gruber,A.R., Bernhart,S.H. and Lorenz,R. (2015) The ViennaRNA web services. *Methods Mol. Biol.*, **1269**, 307–326.

38. Piovesan,D., Necci,M., Escobedo,N., Monzon,A.M., Hatos,A., Mičetić,I., Quaglia,F., Paladin,L., Ramasamy,P., Dosztányi,Z. *et al.* (2021) MobiDB: intrinsically disordered proteins in 2021. *Nucleic Acids Res.*, **49**, D361–D367.

39. Xue,B., Dunbrack,R.L., Williams,R.W., Dunker,A.K. and Uversky,V.N. (2010) PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta*, **1804**, 996–1010.

40. The Gene Ontology Consortium, Carbon,S., Douglass,E., Good,B.M., Unni,D.R., Harris,N.L., Mungall,C.J., Basu,S., Chisholm,R.L., Dodson,R.J. *et al.* (2021) The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.*, **49**, D325–D334.

41. Kang,J., Tang,Q., He,J., Li,L., Yang,N., Yu,S., Wang,M., Zhang,Y., Lin,J., Cui,T. *et al.* (2021) RNAInter v4.0: RNA interactome repository with redefined confidence scoring system and improved accessibility. *Nucleic Acids Res.*, https://doi.org/10.1093/nar/gkab997.

42. Piñero,J., Ramírez-Anguita,J.M., Saüch-Pitarch,J., Ronzano,F., Centeno,E., Sanz,F. and Furlong,L.I. (2020) The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.*, **48**, D845–D855.

43. 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.

44. Bailey,T.L. (2021) STREME: accurate and versatile sequence motif discovery. *Bioinformatics*, **37**, 2834–2284.

45. Iserman,C., Roden,C.A., Boerneke,M.A., Sealfon,R.S.G., McLaughlin,G.A., Jungreis,I., Fritch,E.J., Hou,Y.J., Ekena,J., Weidmann,C.A. *et al.* (2020) Genomic RNA elements drive phase separation of the SARS-CoV-2 nucleocapsid. *Mol. Cell*, **80**, 1078–1091.

46. Carlson,C.R., Asfaha,J.B., Ghent,C.M., Howard,C.J., Hartooni,N., Safari,M., Frankel,A.D. and Morgan,D.O. (2020) Phosphoregulation of phase separation by the SARS-CoV-2 N protein suggests a biophysical basis for its dual functions. *Mol. Cell*, **80**, 1092–1103.