

Overview of circular RNA databases and other resources

Xianfeng Li & Chunjiang He

 Check for updates

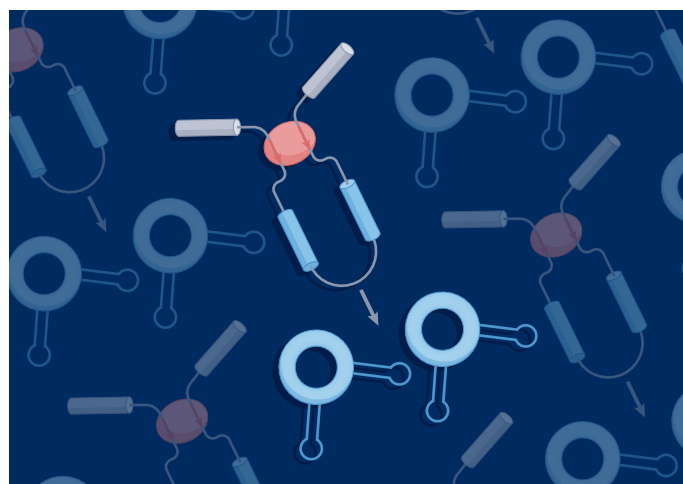
Circular RNAs (circRNAs) form a distinct class of RNA molecules with roles in diverse biological processes. The characterization of large numbers of circRNAs supports extensive database resources. Here we summarize resources dedicated to circRNA tissue and species specificity, processing, interactions, translation and diseases, to facilitate circRNA research and applications.

Circular RNAs (circRNAs) are covalently closed RNA molecules generated from precursor mRNAs via back-splicing. They exert their functions primarily through interactions with DNA, RNA and proteins. In some cases, circRNAs can also function by encoding proteins¹. The widespread use of high-throughput sequencing has enabled the identification of numerous circRNAs from diverse species and tissues, generating large-scale data resources that encompass expression profiles, structural features and functional annotations. In this Comment, we overview circRNA-related resources (Supplementary Table 1) and discuss circRNA functions and therapeutic potential.

Basic versus comprehensive databases

Numerous databases have been constructed to store, annotate and curate circRNA-related information. Basic databases mainly record circRNA genomic coordinates, sequences and expression profiles. These resources cover a broad range of species, including nematodes, fruit flies, *Arabidopsis* spp. and humans, with data scales spanning hundreds to millions of entries.

Early foundational databases, such as circBase², primarily provided genomic localization and origin information. More recent comprehensive platforms, such as circAtlas 3.0 (ref. 3) and CIRCpedia v3 (ref. 4), integrate both short-read and long-read sequencing data and incorporate multidimensional annotations, including expression patterns, coding potential and disease relevance. For example, circAtlas 3.0 includes more than 3 million circRNA sequences from 10 vertebrate species, including humans, and provides full-length sequences, genomic annotations, tissue-specific expression levels, predicted coding potential and disease associations. CIRCpedia v3 integrates over 2,000 next-generation sequencing and third-generation sequencing datasets from 20 species and annotates approximately 2.6 million circRNAs. It provides information on conservation and expression patterns and includes modules for gene-editing validation, therapeutic circRNA optimization, sub-cellular localization prediction and multi-molecular interaction analyses.



Tissue-specific and species-specific databases

Several databases have been developed to characterize the specificity of circRNA expression across tissues, cell types and species. TSCD⁵ is the first large-scale database dedicated to identifying tissue-specific circRNAs in humans and mice, covering dozens of tissues and providing genomic coordinates, tissue distribution and expression profiles. TSCD also offers a tool for comparing circRNA expression patterns across tissues.

PlantcircBase⁶ integrates public transcriptomics and nanopore sequencing data from 21 plant species to provide full-length plant circRNA sequences, conservation information and expression patterns at single-cell resolution. VirusCircBase 2.0 (ref. 7) identifies viral circRNAs from RNA-sequencing data of infected samples and provides information on their sequences, conservation and functional characteristics, covering 43 viruses.

CircRNA processing and modification databases

Like linear RNAs, circRNAs undergo alternative splicing, editing and other post-transcriptional modifications. The parsing and annotation of these processes are essential for exploring circRNA functions. One of databases focusing on internal alternative splicing of circRNAs is circASbase⁸. CircASbase annotates seven types of internal alternative splicing events and distinguishes between events originating from linear or circular transcripts.

N⁶-methyladenosine (m⁶A) is a vital post-transcriptional modification and one of the mechanisms of cap-independent translation of circRNAs. A specific database, m6A2Circ⁹, collects sequencing data and related experimental evidence, systematically identifies and characterizes m⁶A modifications on circRNAs in humans and mice, and provides corresponding functional and disease-associated annotations.

CircRNA interactions databases

A major function of circRNAs is to bind and regulate other macromolecules, especially acting as sponges of microRNAs or RNA-binding proteins. Several databases have been developed to characterize interactions between circRNAs and other RNAs or proteins, by integrating high-throughput experimental data from techniques such as RIP-seq (RNA immunoprecipitation sequencing), CLIP-seq (cross-linking immunoprecipitation sequencing), SPLASH (sequencing of psoralen crosslinked, labelled, and selected hybrids) and RIC-seq (RNA in situ conformation sequencing). These datasets are primarily derived from human and mouse samples, and some databases provide secondary-structure prediction, analysis of disease-associated variants, and functional annotation. For example, CircTarget¹⁰ integrates several datasets obtained from sequencing technologies to identify around 130,000 high-confidence circRNA–RNA interactions from human and mouse transcriptomes. In addition to basic information, CircTarget provides detailed annotations on interaction sites, double-stranded structures and disease-associated variant loci, offering a reliable resource for investigating circRNA regulation mechanisms.

CircRNA translation databases

An increasing number of circRNAs have been demonstrated to possess translation potential, prompting the construction of several circRNA translation databases. These databases collect circRNAs with translation evidence, such as ribosome association, translation initiation sites, mass spectrometry data, internal ribosome entry sites (IRESs), m⁶A sites and open-reading frames, to systematically evaluate the coding potential of circRNAs. For example, TransCirc¹¹ contains over 320,000 human circRNA sequences and incorporates 7 types of translation evidence, thereby providing a comprehensive resource for assessing circRNA coding capacity. Some databases, such as riboCIRC¹², further include analyses of the physicochemical properties, structural features and functional annotations of circRNA-encoded peptides.

Disease-associated circRNA databases

CircRNAs are closely associated with the development and progression of various diseases. Several databases focusing on disease-related circRNAs have been constructed. These databases integrate high-throughput sequencing data, experimental evidence and literature mining, and cover a wide spectrum of physiological and pathological conditions, including cancer, immune disorders, metabolic diseases, inflammation and ageing. Most of these datasets are derived from human samples, although some also include data from model organisms. Some databases further provide molecular interaction networks, expression analyses and primer design tools. For instance, circRNADisease v2.0 (ref. 13) contains more than 4,000 experimentally validated disease-associated circRNAs, offering valuable support for disease-related research. circAge¹⁴ compiles ageing-related circRNAs from diverse tissues and age groups, and provides expression profiles, regulatory networks and functional predictions to facilitate investigations into the roles of circRNAs in ageing.

Database limitations and future perspective

CircRNA databases cover several research dimensions, including basic information, translation potential, molecular interactions and disease associations (Supplementary Table 1). These databases provide diverse resources for researchers in different fields and are crucial for deciphering circRNA functions in vivo. Nevertheless, current circRNA databases have several limitations. First, a standardized nomenclature for circRNAs is lacking, and in most databases circRNAs are named with custom IDs, host gene symbols or genomic coordinates. Such naming methods cannot intuitively distinguish between circRNA isoforms derived from the same gene and hinder cross-database comparison. Only a few databases, such as circAtlas 3.0 and CIRCpedia v3, have adopted standardized naming criteria.

Second, dataset heterogeneity and widespread discrepancies in algorithms exist across platforms. Most databases integrate data from various tissues, pathophysiological conditions and library construction strategies, with inconsistent bioinformatics tools, analytical pipelines and screening thresholds. These inconsistencies result in poor data reproducibility and compatibility among different databases. Accordingly, researchers are recommended to combine multiple databases to obtain comprehensive and reliable circRNA information.

Third, there is a general lack of solid experimental evidence. Most circRNA identification and functional annotations in existing databases rely on computational prediction, and only a small proportion have been experimentally validated. Users need to carefully evaluate the reliability of evidence. For example, circRNA-microRNA interactions should be validated by CLIP-seq data, and circRNA translation should be supported by mass spectrometry.

Fourth, insufficient maintenance and irregular updates represent another major challenge. Several databases are not updated after their publication, or suffer from unstable accessibility, which may result in outdated information. By contrast, continuously maintained databases such as circAtlas 3.0 and CIRCpedia v3 offer higher research reference value.

Notably, circRNAs are attracting increasing attention as therapeutic molecules and protein expression vectors. Compared with linear RNAs, circRNAs are highly stable owing to their covalently closed structure, and purified circRNAs exhibit low immunogenicity. Furthermore, efficient protein production can be achieved by optimizing IRES design and the sequence of untranslated regions. These advantages make circRNAs particularly promising therapeutically relative to other RNA types¹⁵. At present, dedicated databases for the rational design of therapeutic circRNAs are still scarce. Researchers have to collate scattered data from several independent resources, which greatly limits research efficiency. Therefore, the construction of specialized circRNA design databases that integrate information including aptamer sequences, IRES elements, coding and untranslated region sequences and secondary structures will be a highly valuable direction for future research.

Xianfeng Li & Chunjiang He✉

College of Biomedicine and Health, Hubei Jiangxia Laboratory, Huazhong Agricultural University, Wuhan, China.

✉ e-mail: che@whu.edu.cn

Published online: 29 May 2026

References

1. Liu, C. X. & Chen, L. L. Circular RNAs: Characterization, cellular roles, and applications. *Cell* **185**, 2016–2034 (2022).
2. Glazar, P., Papavasiliou, P. & Rajewsky, N. circBase: a database for circular RNAs. *RNA* **20**, 1666–1670 (2014).
3. Wu, W., Zhao, F. & Zhang, J. circAtlas 3.0: a gateway to 3 million curated vertebrate circular RNAs based on a standardized nomenclature scheme. *Nucleic Acids Res.* **52**, D52–D60 (2024).
4. Zhai, S. N. et al. CIRCpedia v3: an interactive database for circular RNA characterization and functional exploration. *Nucleic Acids Res.* **54**, D78–D88 (2026).
5. Xia, S. et al. Comprehensive characterization of tissue-specific circular RNAs in the human and mouse genomes. *Brief Bioinform.* **18**, 984–992 (2017).
6. Xu, X. et al. PlantcircBase 7.0: Full-length transcripts and conservation of plant circRNAs. *Plant Commun.* **3**, 100343 (2022).
7. Fu, P., Cai, Z., Zhang, Z., Meng, X. & Peng, Y. An updated database of virus circular RNAs provides new insights into the biogenesis mechanism of the molecule. *Emerg. Microbes Infect.* **12**, 2261558 (2023).
8. Zou, L. et al. circASbase: a comprehensive database of alternative splicing events in circRNAs. *Genom. Proteomics Bioinformatics* <https://doi.org/10.1093/gpbjnl/qzaf121> (2025).
9. Li, Y. et al. m6A2Circ: a comprehensive database for decoding the regulatory relationship between m6A modification and circular RNA. *Comput. Struct. Biotechnol. J.* **27**, 813–820 (2025).
10. Zhang, H. et al. CircTarget: a comprehensive database of circRNA-target RNA interactions across multiple cell types. *Nucleic Acids Res.* **54**, D89–D97 (2026).
11. Huang, W. et al. TransCirc: an interactive database for translatable circular RNAs based on multi-omics evidence. *Nucleic Acids Res.* **49**, D236–D242 (2021).
12. Xu, Y., Yang, L., Tang, Y., Wang, Y. & Wang, H. riboCIRC v2.0: an expanded resource for translatable circRNAs. *Nucleic Acids Res.* **54**, D273–D278 (2026).

-
13. Sun, Z. Y. et al. circRNADisease v2.0: an updated resource for high-quality experimentally supported circRNA-disease associations. *Nucleic Acids Res.* **52**, D1193–D1200 (2024).
 14. Dong, X. et al. CircAge: a comprehensive resource for aging-associated circular rnas across species and tissues. *Genom. Proteomics Bioinformatics* **23**, qzaf044 (2025).
 15. O’Leary, E., Jiang, Y., Kristensen, L. S., Hansen, T. B. & Kjems, J. The therapeutic potential of circular RNAs. *Nat. Rev. Genet.* **26**, 230–244 (2025).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41580-026-00985-x>.

Peer review information *Nature Reviews Molecular Cell Biology* thanks Trees-Juen Chuang and Li Yang for their contribution to the peer review of this work.