SCIENCE CHINA PRESS

SCIENCE CHINA Life Sciences
life.scichina.com link.springer.com
https://doi.org/10.1007/s11427-024-2603-5

RESEARCH HIGHLIGHT

Illuminating functions of evolutionarily conserved long non-coding RNAs

Ruiqiao He[1,2] & Fangqing Zhao[1,2*]

[1]Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China;
[2]Key Laboratory of Systems Health Science of Zhejiang Province, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China

*Corresponding author (email: zhfq@biols.ac.cn)

Received 07 April 2024; Accepted 30 April 2024; Published online 24 May 2024

Long non-coding RNAs (lncRNAs) serve as crucial regulators of gene expression, influencing various physiological processes and disease states (Cabili et al., 2011; Ulitsky and Bartel, 2013; Yao et al., 2019). Dysregulation of certain notable lncRNAs, such as HOTAIR and MALAT1, has been implicated in a wide range of pathological processes, including tumorigenesis in various human cancers (Gupta et al., 2010; Kim et al., 2018). However, the functions of the majority of lncRNAs remain elusive. Traditional sequence comparison methods are limited in their ability to identify homologous lncRNAs across species due to the poor sequence conservation (Derrien et al., 2012; Hezroni et al., 2015). For instance, among tens of thousands of lncRNA genes in zebrafish and humans, only a very small fraction exhibit sequence-conserved homologous (Ulitsky et al., 2011).

In response to this challenge, Huang et al. (2024) introduced a novel computational approach, known as lncHOME (Figure 1A), designed to identify conserved homologous lncRNAs across eight vertebrate species, including human, mouse, and zebrafish. This method focused on a distinct class of lncRNAs characterized by conserved genomic positions and patterns of RNA-binding protein (RBP) binding site, referred to as coPARSE-lncRNAs. Additionally, the authors used a CRISPR-based knockout and rescue screening system to experimentally validate the conserved functions of the identified homologous coPARSE-lncRNAs in different species. This experimental evidence supports the notion that distantly related species retain similar regulatory grammar.

In total, lncHOME identified 570 hu-

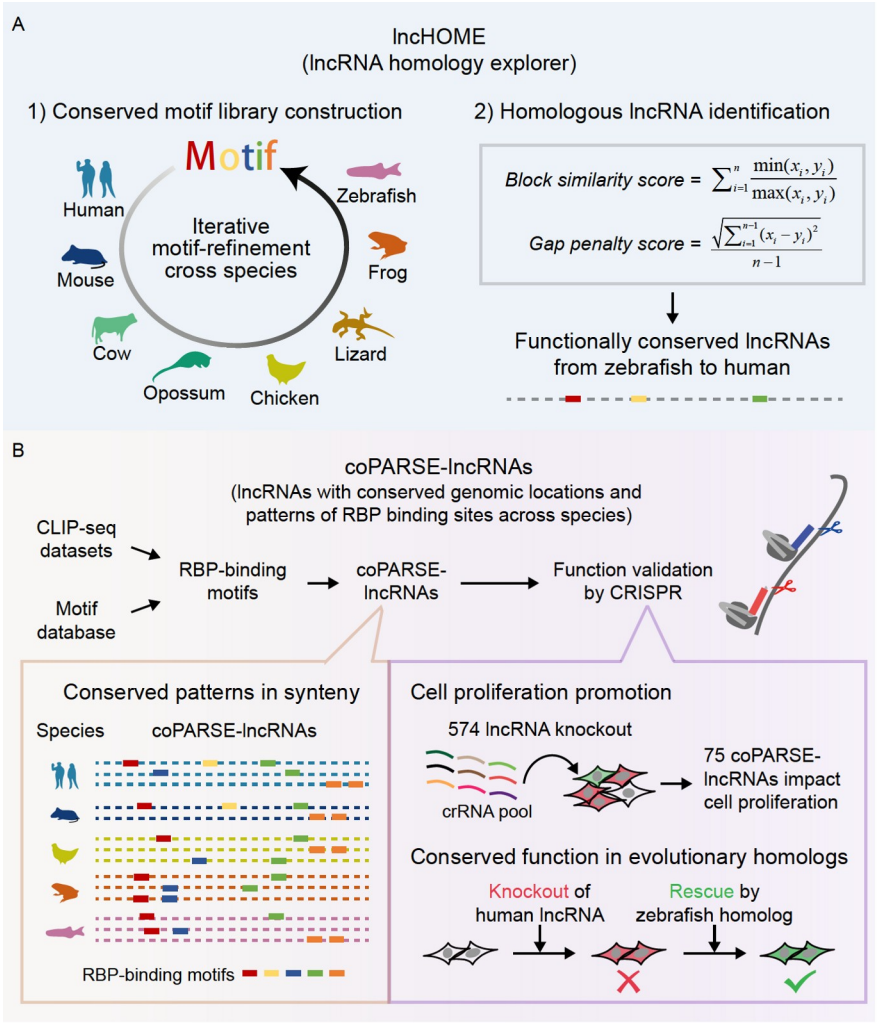

Figure 1. The workflow and features of lncHOME. A, Workflow for lncHOME analysis of functional conserved lncRNAs in vertebrates. lncHOME delineates conserved lncRNAs by combining the identification of functional conserved motif across species (left) and the analysis of similar motif distribution patterns (right). B, The evolutionary features and conserved functions of coPARSE-lncRNAs. lncHOME recognizes coPARSE-lncRNAs as a distinct class of lncRNAs characterized by conserved genomic locations and patterns of RBP binding sites. Using a knockout screen based on Cas12a with paired crRNAs revealed 75 coPARSE-lncRNAs that modulate cancer cell proliferation. Moreover, several prioritized human coPARSE-lncRNAs, along with their zebrafish homologs, were validated to exert shared functions across evolutionarily distant species.

man coPARSE-lncRNAs bearing homologous RBP-regulatory signatures with zebrafish lncRNAs, representing a substantial expansion compared with the mere 17 identifiable homologous lncRNAs detected through sequence alignment alone. These newly classified coPARSE-lncRNAs exhibit enrichment for disease-associated mutations and are more frequently dysregulated in cancer tissues compared with non-homologous fragments. These findings suggest that coPARSE-lncRNAs may have significant physiological functions.

The authors delved deeper into the functional conservation of the identified coPARSE-lncRNAs. They established a CRISPR-Cas12a-mediated knockdown screening system and pinpointed 75 coPARSE-lncRNAs that promote cancer cell proliferation (Figure 1B), with 37 exhibiting pivotal roles in HeLa cells. Subsequently, they devised a single-step system based on CRISPR-Cas12a for knockout and rescue. Through this system, they observed that zebrafish-derived coPARSE-lncRNA homologs could alleviate the proliferation defects induced by the knockdown of human coPARSE-lncRNA (see Figure 1B). Similarly, human homologous coPARSE-lncRNAs were able to rescue the severe developmental delays in zebrafish embryos resulting from the knockout of four zebrafish coPARSE-lncRNAs. These findings underscore the potential functional conservation of coPARSE-lncRNAs across evolutionarily related species.

The authors postulate that coPARSE-lncRNAs exhibit similar RBP binding profiles across homologs due to their identification based on a conserved RBP binding site pattern. This hypothesis was validated through RNA pull-down and mass spectrometry experiments of coPARSE-lncRNAs. Notably, the homologous lncRNAs, which could rescue cell proliferation or embryonic developmental defects, shared a large number of RBPs in their interactomes among human, mouse and zebrafish. In addition, the conserved rescuing effect was compromised upon mutating specific RBP binding sites, such as NONO and IGF2BP2. These findings robustly support the notion that coPARSE-lncRNA homologs maintain functional conservation through interactions with specific RBPs.

This study presents a computational analysis method for identifying potentially homologous lncRNAs across species and experimentally verifying their functional conservation. Despite losing sequence conservation during evolution, these lncRNAs retained conserved RBP binding patterns. Looking ahead, this work endeavors to unravel the common regulatory grammar preserved in evolutionarily related species. It is anticipated that lncHOME will have a wide range of applications, including the recognition of lncRNA homologs with diverse functions and the expansion of RNA homologs with conserved functions to other types of non-coding RNAs, such as circular RNAs.

Compliance and ethics

The author(s) declare that they have no conflict of interest.

References

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 25, 1915–1927.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Res 22, 1775–1789.

Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H. M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 464, 1071–1076.

Hezroni, H., Koppstein, D., Schwartz, M.G., Avrutin, A., Bartel, D.P., and Ulitsky, I. (2015). Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. Cell Rep 11, 1110–1122.

Huang, W., Xiong, T., Zhao, Y., Heng, J., Han, G., Wang, P., Zhao, Z., Shi, M., Li, J., Wang, J., et al. (2024). Computational prediction and experimental validation identify functionally conserved lncRNAs from zebrafish to human. Nat Genet 56, 124–135.

Kim, J., Piao, H.L., Kim, B.J., Yao, F., Han, Z., Wang, Y., Xiao, Z., Siverly, A.N., Lawhon, S.E., Ton, B.N., et al. (2018). Long noncoding RNA MALAT1 suppresses breast cancer metastasis. Nat Genet 50, 1705–1715.

Ulitsky, I., and Bartel, D.P. (2013). lincRNAs: genomics, evolution, and mechanisms. Cell 154, 26–46.

Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011). Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 147, 1537–1550.

Yao, R.W., Wang, Y., and Chen, L.L. (2019). Cellular functions of long noncoding RNAs. Nat Cell Biol 21, 542–551.