



OPEN ACCESS

EDITED BY

Ali Rahnavard,
George Washington University, United States

REVIEWED BY

Muhammad Sajjad,
Zhejiang Agriculture and Forestry University,
China
Mary Ann Allen,
University of Colorado Boulder, United States

*CORRESPONDENCE

Bruno César Feltes,
✉ bruno.feltes@ufrgs.br

RECEIVED 09 August 2024

ACCEPTED 27 November 2024

PUBLISHED 16 December 2024

CITATION

Poloni JF, Oliveira FHS and Feltes BC (2024)
Localization is the key to action: regulatory
peculiarities of lncRNAs.
Front. Genet. 15:1478352.
doi: 10.3389/fgene.2024.1478352

COPYRIGHT

© 2024 Poloni, Oliveira and Feltes. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Localization is the key to action: regulatory peculiarities of lncRNAs

Joice de Faria Poloni, Fábio Henrique Schuster de Oliveira and Bruno César Feltes*

Department of Biophysics, Laboratory of DNA Repair and Aging, Institute of Biosciences, Federal University of Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

To understand the transcriptomic profile of an individual cell in a multicellular organism, we must comprehend its surrounding environment and the cellular space where distinct molecular stimuli responses are located. Contradicting the initial perception that RNAs were nonfunctional and that only a few could act in chromatin remodeling, over the last few decades, research has revealed that they are multifaceted, versatile regulators of most cellular processes. Among the various RNAs, long non-coding RNAs (lncRNAs) regulate multiple biological processes and can even impact cell fate. In this sense, the subcellular localization of lncRNAs is the primary determinant of their functions. It affects their behavior by limiting their potential molecular partner and which process it can affect. The fine-tuned activity of lncRNAs is also tissue-specific and modulated by their *cis* and *trans* regulation. Hence, the spatial context of lncRNAs is crucial for understanding the regulatory networks by which they influence and are influenced. Therefore, predicting a lncRNA's correct location is not just a technical challenge but a critical step in understanding the biological meaning of its activity. Hence, examining these peculiarities is crucial to researching and discussing lncRNAs. In this review, we debate the spatial regulation of lncRNAs and their tissue-specific roles and regulatory mechanisms. We also briefly highlight how bioinformatic tools can aid research in the area.

KEYWORDS

long non-coding RNA, lncRNA, subcellular localization, tissue-specific expression, cis- and trans-acting elements, bioinformatics tools and databases

1 Introduction

Eukaryotic transcriptome encompasses a broad diversity of RNA types that differ in functionality, biogenesis, compartment of activity (i.e., nucleus, cytoplasm, etc.), and tissue-specific roles and regulation mode (*cis* or *trans*). Up to 90% of the human genome was found to be transcribed during the eukaryotic life cycle; however, from this total, less than 2% correspond to protein-coding genes (Hatje et al., 2019; Scherrer, 2023). The remaining functional transcriptome comprises ncRNAs categorized as housekeeping or regulatory. Housekeeping ncRNAs are abundant in all cell types, showing constitutive expression, and are associated with regulating primary cellular function (Zhang P. et al., 2019). Among the most known housekeeping ncRNAs are ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), and small nucleolar RNAs (snoRNAs) (Zhang P. et al., 2019). In turn, regulatory ncRNAs act in several biological processes, such as chromatin

remodeling, transcription and translation regulation, and mRNA processing and decay (Yotsukura et al., 2017; Diamantopoulos et al., 2018; Rincón-Riveros et al., 2021) and their biogenesis and activity might be tissue- or compartment-specific. The most known regulatory ncRNAs comprise microRNA (miRNA), endogenous short interfering RNA (endo-siRNA), PIWI-interacting RNA (piRNA), long non-coding RNA (lncRNA), and circular RNA (circRNA) (Loganathan and Doss C, 2023).

lncRNAs are integral to the regulation of biological processes and play significant roles in the etiology of various pathologies, such as cardiovascular, neurodegenerative, and autoimmune diseases, neuroinflammation, and found to be deregulated in several cancer types (Lanzós et al., 2017; Diamantopoulos et al., 2018; Poller et al., 2018; Lodde et al., 2020; Ruffo et al., 2021; Chen et al., 2022; Jiang et al., 2022; Anilkumar et al., 2024).

The pervasive presence of lncRNAs underscores their indispensability in unraveling the intricacies of gene regulation. Despite the detailed characterization of specific lncRNAs, a substantial proportion still needs to be associated with biological processes and their molecular targets. According to GENCODE, the human genome has more than 20,000 lncRNA genes (Frankish et al., 2023), but just a small proportion of lncRNAs has been functionally characterized, about ~500–1,500 (Carlevaro-Fita and Johnson, 2019). Consequently, elucidating direct regulatory targets and linking them to biological functions poses considerable complexity. This gap emphasizes the critical role of bioinformatics tools and the need for additional experimental data to predict and analyze lncRNA structure, function, and other attributes. The advancement of high-throughput techniques has markedly expedited lncRNA research and facilitated the establishment of numerous publicly accessible databases and tools to uncover lncRNA functions and interactions. Bioinformatic tools are indispensable for studying lncRNAs, providing comprehensive methods for their identification, functional prediction, structural analysis, and clinical application. These tools have revolutionized our understanding of ncRNAs, revealing their complexity and significance in gene regulation.

However, to reveal the biological role of lncRNAs, cellular localization remains a crucial factor in addressing their function. lncRNA activities are localization-specific in several ways: they are more tissue-specific than mRNA, show particular subcellular locations, and have an action mechanism that can be performed in *cis*- or *trans*-regulator. Identifying if a lncRNA is expressed within a particular tissue or if the target mature RNA is in a determined subcellular compartment is crucial for future studies involving these molecules. Furthermore, lncRNA localization has practical implications in studies that have proposed using lncRNA as biomarkers or drug targets.

This review highlights the mechanistic spatial insight of lncRNA action considering tissue expression, subcellular location, and mode of action. We also discuss the importance of bioinformatic tools in lncRNA research, demonstrating how this field has revolutionized our understanding of these non-coding elements.

2 Long non-coding RNA

One of the first studies to identify a lncRNA was published in 1990 by Brannan et al., where the authors discovered that the gene

H19 was not translated to protein, even though this gene possesses a small open reading frame (Brannan et al., 1990; Jarroux et al., 2017). H19 has some characteristics compatible with mRNAs, such as gene transcription by RNA Polymerase II, splicing, 3' polyadenylation, and translocation to the cytoplasm (Brannan et al., 1990). Although H19 gene expression is essential to embryonic development, only in 1990 the function of H19 was fully understood when Xist was also characterized, revealing that both lncRNAs are involved in genomic imprinting and dosage compensation in mammals (Jarroux et al., 2017).

lncRNAs have a sequence length above 200 nucleotides and a structure similar to mRNAs. However, they lack the open reading frame (ORF) or show small ORFs (length less than 300 nucleotides), and consequentially, they typically are not able to produce a full-length protein (Gong et al., 2021; Feng et al., 2023). In the case of lncRNA showing sORFs, they may synthesize small and stable micropeptides. However, a significant micropeptide fraction is unstable and degraded after synthesis (Singh, 2024). lncRNAs were considered transcriptional noise because they have low expression and show more specific tissue expression than coding protein genes (Gloss and Dinger, 2016; Feng et al., 2023).

The emergence of lncRNA research can be attributed to their regulatory functions in several biological processes. Like most genes, lncRNA-associated genes are composed of introns and exons, mainly transcribed by RNA polymerase II (Pol II) and less frequently by other polymerases (Statello et al., 2021; Mattick et al., 2023). The transcribed product being is processed by featuring 7-methyl guanosine (m7G) caps at their 5' ends and polyadenylated tails at their 3' ends (Khan et al., 2021; Statello et al., 2021). In addition, lncRNAs are subject to alternative or constitutive splicing after transcription. However, pre-lncRNAs are less efficiently spliced than pre-mRNA of protein-coding genes, possibly due to differences in consensus sequences for the branch point and 5' and 3' splice sites or interaction of specific splicing factors that can decrease alternative usage of particular exons (Khan et al., 2021). Also, lncRNA transcripts are less abundant than mRNA, which could be related to unspliced and low-stability transcripts due to an absence of proximal RNA polymerase II phosphorylation over the lncRNA 5' splice site (Khan et al., 2021).

In contrast to what was first believed, only a minority of lncRNAs are unstable (Bridges et al., 2021). Most are stabilized through polyadenylation, and the non-polyadenylated shows secondary structures that ensure their stability (Bridges et al., 2021). In addition, it was presumed that lncRNAs are transcribed and processed the same way as mRNA; however, distinct cellular fates and functions can be related to specific coordination of lncRNA transcription, processing, exportation, and turnover (Statello et al., 2021). Some lncRNAs are transcribed by RNA polymerase I (Pol I) and RNA polymerase III (Pol III), not being polyadenylated nor receiving m7G caps, or yet resulting from the processing of different precursors, such as introns and repetitive elements (Mattick et al., 2023). In addition, many lncRNAs are successfully spliced and exported to the cytoplasm, while others are inefficiently processed during splicing, and the product is retained in the nucleus (Statello et al., 2021). In this case, a study found that due to their higher transcript complexity (i.e., multiple splice variants per exon, shorter intron length, predominant dinucleotide on splicing sites, and low conservation at 5' and 3' splicing sites),

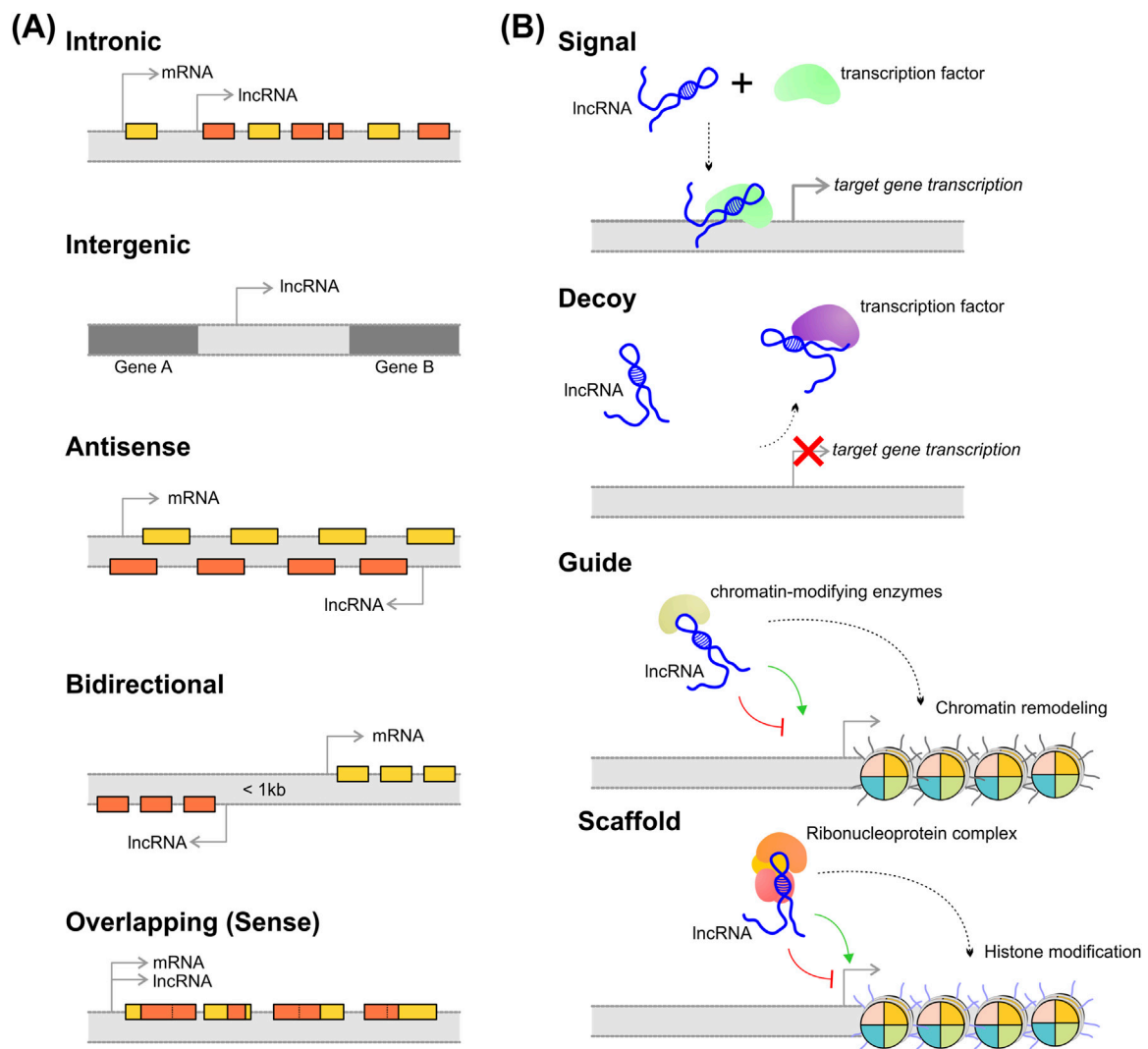


FIGURE 1
 Classification and molecular function of lncRNAs. (A) lncRNAs classification based on their genomic position and orientation relative to nearby protein-coding genes; (B) lncRNAs can execute different molecular functions depending on their mechanism of action. See the main text for the explanation.

lncRNA splicing is more inefficient than mRNA (Basu et al., 2023). Intronic lncRNAs and unspliced lncRNAs are less stable than intergenic, cis-antisense, and spliced lncRNA (Clark et al., 2012). These lncRNAs are submitted to regulators of lncRNA turnover and directed to degradation via different mechanisms, which are believed to be similar to those controlling mRNA turnover, such as decay-promoted by RBPs, microRNAs, decapping and deadenylation followed by exo- and endonucleolytic degradation and translation-associated RNA decay (Yoon et al., 2015). Likewise, despite our knowledge of which factors play a role in lncRNA decay, how this process is regulated and the impact on bioprocess is still to be elucidated (Singh, 2024). According to a study conducted in mouse Neuro-2a cell line, 29% of lncRNAs were unstable (considering a half-life of 2 h or less), and 6% were highly stable (with a half-life superior to 16h) (Clark et al., 2012). Nevertheless, it should be noted that knockout of some lncRNAs does not affect the observable phenotype (Gao et al., 2020), suggesting that due to their higher decay rate, they might exert

no function at all, and only a selected few might truly impact cell biology (Ponting and Haerty, 2022). Despite that, a significant fraction of lncRNAs have nuclear localization signals and remain in the nucleus to exert their roles, which may influence gene expression, nuclear organization, phase separation, compartment formation, and modulating epigenetic regulation (Guh et al., 2020).

The lncRNA classification that is more frequently used was proposed by GENCODE and is based on the genomic context considering their position and orientation relative to protein-coding genes (Antonov et al., 2019). Consequently, lncRNAs may be categorized into six classes (Figure 1A): (i) intronic, referring to lncRNAs that derived from intronic sequences of protein-coding genes; (ii) intergenic, is referred to lncRNA transcribed from regions located between protein-coding genes and do not overlap with known coding sequences; (iii) antisense, the lncRNA is transcribed from the opposite strand relative to protein-coding genes; (iv) bidirectional, when the lncRNA is transcribed from promoters located in the opposite direction of neighboring

protein-coding genes and often share the same promoter region; (v) overlapping (or sense), represents lncRNAs whose gene sequence overlaps another gene on the same strand.

Furthermore, lncRNAs can also be classified according to their action mode, which is closely related to their molecular function. The action mode for lncRNAs can be checked in [Figure 1B](#) and is described as (i) signal mode, which refers to lncRNA that promotes chromatin architecture modifications, silencing their target genes, or could act as an enhancer, promoting the transcriptional machinery recruitment to induce the transcription of target genes in response to different stimuli; (ii) decoy mode, refers to the lncRNA that bind to proteins with regulatory functions, such as transcription factors, and decrease the accessibility of these proteins by “sequestering” them. In this case, decoy lncRNA indirectly acts by activating or inhibiting the transcription of target genes; (iii) guide mode, where the lncRNA directs the assembly of chromatin-modifying enzymes by binding to the molecules and guiding them to specific genomic loci; (iv) scaffold mode, where the lncRNA act as a platform for the assembly and interaction of multiprotein complexes and RNA-binding factors ([Ahmad et al., 2016](#); [Gong et al., 2021](#)).

lncRNAs have been observed to regulate bioprocesses, such as epigenetic, transcriptomic, post-transcriptional, translation, and post-translational modification ([Zhang X. et al., 2019](#)). Although evidence supporting the functionality of most lncRNAs is still missing, there is well-documented evidence that an increasing number of lncRNAs perform critical cellular roles. In part, this difficulty is related to the fact that many lncRNA are reported to be tissue-specific or even cell lineage-specific ([Kashi et al., 2016](#); [Mattick et al., 2023](#)). For example, research found that lncRNAs have tissue-specific expression and distinct subcellular localization patterns and are expressed in narrower time windows than mRNAs along the Zebrafish developmental time course ([Pauli et al., 2012](#)).

Furthermore, different reasons make the elucidations about lncRNAs more complicated, such as the lower expression than mRNAs, making their identification challenging to detect amidst the transcriptome. Another challenge is that most studies use oligo (dT) primers for cDNA synthesis, challenging the investigation of non-polyadenylated lncRNA ([Kashi et al., 2016](#)). lncRNA also shows low overall sequence conservation across different species compared to protein-coding genes, even though their promoters, exon structure, and splice junctions infer selective constraint ([Ponting et al., 2009](#); [Kashi et al., 2016](#); [Mattick et al., 2023](#)). Despite the low overall sequence conservation, many lncRNAs retain preserved functions. While this is, in part, justified by the conservation of a sequence subset, it also indicates that the primary sequence of lncRNAs may not be the determinant of how they exert their function - instead, their functionality could be explained by the formation of secondary structures and specific sequence motifs that remains preserved ([Ponting et al., 2009](#); [Ulitsky and Bartel, 2013](#)).

3 lncRNA landscape

3.1 Subcellular localization

Understanding the subcellular location of lncRNA is the primary determinant that infers its biological function since it

interacts with proteins, DNA, and other RNA types. lncRNA traffic to exact subcellular locations is a process that demands energy; thus, their transport to the correct destination is relevant for the proper execution of their function ([Lee et al., 2019](#)). If this process is halted or abolished, it would impede lncRNAs from finding their interaction partners and regulating their cellular behavior correctly. Otherwise, they could engage in deleterious interactions or activities ([Lee et al., 2019](#)). Due to the importance of lncRNA spatial localization in understanding their partners and participation in different bioprocesses, several databases offering subcellular localization information were developed and explored in [Table 1](#).

The subcellular location of RNAs is not stochastic but a dynamic and regulated process that controls the localization of protein expression, turnover, and subsequent signal regulation in response to homeostatic, stimulated, or cellular stress conditions ([Bridges et al., 2021](#)). The first observation of the asymmetrical distribution of mRNA was made in 1983 ([Jeffery et al., 1983](#)). Subsequent research and methodological advances have shed light on this non-random allocation, proposing that the distribution of RNA in a cell is directly related to local protein concentration, as a non-uniform gradient distribution of RNA may perform a regulatory function ([Carlevaro-Fita and Johnson, 2019](#)).

Several mechanisms regulate the localization of the lncRNA, which are primarily present in the nucleus and cytoplasm; however, many studies have found lncRNA present in organelles and macromolecular structures, such as mitochondria, endoplasmic reticulum, ribosomes, extracellular membrane, exosome, nucleolus, chromatin speckles, and paraspeckles ([Dragomir et al., 2018](#); [Carlevaro-Fita and Johnson, 2019](#); [Bridges et al., 2021](#); [Li C. et al., 2021](#)). The enrichment of lncRNAs in specific locations is orchestrated by different sequence motifs or domains, lncRNA secondary structure, and post-transcriptional modifications ([Carlevaro-Fita and Johnson, 2019](#)). Nuclear retention may be defined by the primary sequence of lncRNA and may include short motifs, such as hexamers, structural elements, transposable element fragments, or longer sequence domains ([Carlevaro-Fita and Johnson, 2019](#)).

Most lncRNA's Post-transcriptional modifications are similar to mRNA's that are crucial for trafficking and nuclear exportation; hence, they share a common exporting pathway ([Carlevaro-Fita and Johnson, 2019](#); [Guo et al., 2020](#)). In this sense, the transcription-export complex (TREX), nuclear RNA export factor 1 (NXF1), and nuclear transport factor 2-like export factor 1 (NXT1) promote the lncRNA transportation, especially by NXF1, which is preferentially used for the exportation of RNA composed by single or few exon or RNA enriched with high A/U content ([Guo et al., 2020](#)). However, lncRNA that have divergent transcription, caused by different phosphorylation pattern Pol II C-terminal domains (CTDs), or do not share the same processing of mRNA transcripts, may be inefficiently exported by the protein complexes mentioned above ([Guo et al., 2020](#); [Khan et al., 2021](#)). This may occur by the fact that RNA processing, co-transcriptional splicing, 3' cleavage, and polyadenylation are strongly related to CTD Ser5 and Ser2 phosphorylation; however, intergenic lncRNA showed to be less selective to CTD profile and Pol II pausing at the transcription start and end sites are generally absent on these lncRNAs ([Schlackow et al., 2017](#)).

TABLE 1 Tools were curated manually mainly through relevant term searches on multiple academic research databases and also through cross-citation in lncRNA-related papers. Only resources updated or released since 2015 and had a working platform at the moment of curation were added to the Table. All data was collected from 17 June 2024, to 26 July 2024.

Resource	Evidence	Spatial Information	Last update	Supported species	Website link	References
Annocript	Prediction-based	-	2018	-	https://github.com/frankMusacchia/Annocript	Musacchia et al. (2015)
AnnoLnc2	Experiment and prediction-based	Tissue specific and subcellular localization	2020	Human, mouse	http://annolnc.gao-lab.org/	Ke et al. (2020)
CANTATAdb	Experiment-based	-	2024	108 plant species	http://yeti.amu.edu.pl/CANTATA/	Szcześniak and Wanowska (2024)
CPC2	Prediction-based	-	2017	-	https://cpc2.gao-lab.org/	Kang et al. (2017)
CPPred	Experiment-based	-	2019	-	http://www.mabinding.com/ CPPred	Tong and Liu (2019)
DeepLGP	Experiment-based	-	2019	-	https://github.com/zty2009/LncRNA-target-gene	Zhao et al. (2020)
DeepLncLoc	Prediction-based	Subcellular localization	2022	-	http://bioinformatics.csu.edu.cn/DeepLncLoc/	Zeng et al. (2022)
DIANA-LncBase*	Experiment and prediction-based	Tissue specific and Subcellular localization	2019	Human, mouse	www.microrna.gr/LncBase	Karagkouni et al. (2020)
ENCORI/starBase*	Experiment-based	-	2023	23 species, including human	https://rnasyu.com/encori	Li et al. (2014)
exoRBase 2.0*	Experiment-based	Tissue specific	2021	Human	http://www.exoRBase.org	Lai et al. (2022)
GENCODE*	Experiment and prediction-based	-	2023	Human, mouse	https://www.genecodegenes.org	Frankish et al. (2019)
GTEEx*	Experiment-based	Tissue specific	2023	Human	https://gtexportal.org/	Lonsdale et al. (2013)
iLoc-lncRNA	Experiment and prediction-based	Subcellular localization	2018	-	http://lin-group.cn/server/iLoc-LncRNA	Su et al. (2018)
iLoc-lncRNA (2.0)	Prediction-based	Subcellular localization	2022	-	http://lin-group.cn/server/iLoc-LncRNA(2.0)/	Zhang et al. (2022)
IntaRNA	Prediction-based	-	2017	-	http://rna.informatik.uni-freiburg.de/IntaRNA	Mann et al. (2017)
Lnc2Cancer 3.0*	Experiment and prediction-based	Subcellular localization	2020	Human	http://bio-bigdata.hrbmu.edu.cn/Lnc2cancer	Gao et al. (2021)
LncACTdb*	Experiment and prediction-based	Subcellular localization	2022	25 species, including human	http://bio-bigdata.hrbmu.edu.cn/LncACTdb/	Wang et al. (2022)
LncADeep	Prediction-based	-	2018	-	https://cqb.pku.edu.cn/zhuLab/info/1006/1160.htm	Yang et al. (2018)
LnCaNet	Experiment-based	-	2016	Human	http://lncanet.bioinfo-minzhao.org/	Liu and Zhao (2016)
LncATLAS	Experiment-based	Subcellular localization	2017	Human	https://lncatlas.crg.eu/	Mas-ponte et al. (2017)
LNCBook	Experiment and prediction-based	-	2022	Human	https://ngdc.cnbc.ac.cn/lncbook	Li et al. (2023)
LncCeRBase*	Experiment-based	-	2018	Human	http://www.insect-genome.com/LncCeRBase	Pian et al. (2018)
LncExpDB	Experiment and prediction-based	Tissue specific and subcellular localization	2021	Human	https://bigd.big.ac.cn/lncexpdb	Li Z. et al. (2021)
LncFunTK	Experiment-based	-	2018	-	https://github.com/zhoujj2013/lncfuntk	Zhou et al. (2018)
LNCipedia	Experiment-based	-	2018	Human	https://lncipedia.org	Volders et al. (2019)
LncLocation	Prediction-based	Subcellular localization	2020	-	https://github.com/FengSY-JLU/Core-LncLocation/	Feng et al. (2020)

(Continued on following page)

TABLE 1 (Continued) Tools were curated manually mainly through relevant term searches on multiple academic research databases and also through cross-citation in lncRNA-related papers. Only resources updated or released since 2015 and had a working platform at the moment of curation were added to the Table. All data was collected from 17 June 2024, to 26 July 2024.

Resource	Evidence	Spatial Information	Last update	Supported species	Website link	References
lncLocator 2.0	Prediction-based	Subcellular localization	2021	-	www.csbio.sjtu.edu.cn/bioinf/lncLocator2	Lin et al. (2021)
lncRNA2Target*	Experiment-based	-	2021	Human, mouse	http://bio-computing.hrbmu.edu.cn/lncrna2target/	Cheng et al. (2019)
lncRNAdb	Experiment-based	Subcellular localization	2015	10 species, including human	https://rnacentral.org/expert-database/lncrnadb	Amaral et al. (2011)
lncRNADisease v3.0*	Experiment and prediction-based	-	2023	4 species, including human	http://www.rnaut.net/lncrnadisease	Lin et al. (2024)
lncRNASNP2*	Experiment and prediction-based	-	2018	Human, mouse	https://guolab.wchscu.cn/lncRNASNP/#/	Miao et al. (2018)
lncRNAWiki 2.0	Experiment-based	Subcellular localization	2021	Human	https://ngdc.cncb.ac.cn/lncrnawiki	Liu et al. (2022)
lncRRISearch	Experiment and prediction-based	Tissue specific and subcellular localization	2019	Human, mouse	http://rtools.cbrc.jp/lncRRISearch/	Fukunaga et al. (2019)
lncSEA	Experiment-based	-	2021	Human	http://bio.liclab.net/lncSEA/index.php	Chen et al. (2021)
lncSpA	Experiment-based	Tissue specific	2019	Human	http://bio-bigdata.hrbmu.edu.cn/lncSpA	Lv et al. (2020)
lncTarD 2.0	Experiment and prediction-based	-	2022	Human	http://bio-bigdata.hrbmu.edu.cn/lncTarD	Zhao et al. (2023)
Locate-R	Prediction-based	Subcellular localization	2020	-	http://locate-r.azurewebsites.net/	Ahmad et al. (2020)
ncFANs v2.0*	Experiment-based	Tissue specific	2021	Human, mouse	http://ncfans.gene.ac/	Zhang et al. (2021)
NONCODEV5*	Experiment and prediction-based	-	2020	17 species, including human	http://www.noncode.org/	Fang et al. (2018)
NPInter v5.0*	Experiment and prediction-based	-	2022	60 species, including human	http://bigdata.ibp.ac.cn/npinter5/	Zheng et al. (2023)
RefLnc	Experiment-based	-	2017	Human	https://reflnc.gao-lab.org/	Jiang et al. (2019)
Rlblast*	Prediction-based	-	2020	-	https://github.com/fukunagatsu/Rlblast	Fukunaga and Hamada (2017)
RISE*	Experiment-based	-	2017	3 species, including human	http://rise.life.tsinghua.edu.cn	Gong et al. (2018)
RNA Disease v4.0*	Experiment and prediction-based	-	2022	117 species, including human	http://www.rnadisease.org/	Chen et al. (2023)
RNAenrich*	Experiment-based	-	2023	Human, mouse	http://idrblab.cn/rnaenrich/	Zhang et al. (2023)
RNAInter v4.0*	Experiment and prediction-based	-	2021	156 species, including human	http://www.rnainter.org	Kang et al. (2022)
RNAlocate v3.0*	Experiment and prediction-based	Tissue specific and subcellular localization	2024	242 species, including human	http://www.rmalocate.org/	Wu et al. (2022)
SEEKR	Prediction-based	-	2024	-	https://github.com/CalabreseLab/seekr	Kirk et al. (2018)
SFPEL-lpi	Prediction-based	-	2017	-	http://www.bioinfotech.cn/SFPEL-LPI/	Zhang et al. (2018)
TANRIC	Experiment-based	-	2022	Human	https://www.tanric.org	Li et al. (2015)
TransCistor	Prediction-based	-	2024	Human and mouse	https://github.com/gold-lab/TransCistor https://transcistor.unibe.ch/	Dhaka et al. (2024)

Table legends: *supports other types of RNA besides lncRNA.

Furthermore, chromatin marks and Pol II promoter-proximal pausing are also related to nuclear exportation, possibly allowing the association of lncRNA with exporting proteins (Zuckerman and Ulitsky, 2019). Altogether, splicing efficiency is substantially increased in lncRNA transcript found in the cytoplasm compared to those retained in the nucleus (Zuckerman and Ulitsky, 2019).

Alternatively, unusual stabilization may occur by forming structures at their ends that confer stability and ensure RNA stability and the exportation of the RNA from the nucleus (Wilusz et al., 2012). Different structures at the end of lncRNA are formed by distinct motifs and described to confer stability and direct the lncRNA to proper subcellular localization. The nascent transcript is processed at the 3' end during transcription, ensuring the mature RNA's functionality. Most RNA Pol II transcripts receive a poly(A) tail conceived from cleavage by endonucleases and the addition of adenosine (A) residues. This poly(A) tail confers stability and allows the mature RNA exportation from the nucleus. However, some Pol II transcripts, instead of having poly(A) tail, show a different 3' RNA conformation, such as the triple-helical structure found in the lncRNA MALAT1, formed by highly conserved A- and U-rich motifs at 3' end and promote nuclear retention of MALAT1 and its enrichment in nuclear speckles (Wilusz et al., 2012). Another example is SLERT, a lncRNA specially located in the nucleolus, and its localization and biogenesis are mediated by a box H/ACA at both ends (Zhang Y. et al., 2017). Different lncRNA sequence motifs may interact with several partners, such as hnRNPs, RNA helicases, and RNA processing factors, to restrict the lncRNA to the nucleus (Guo et al., 2020).

In general, cytoplasmic lncRNAs are more stable than their nuclear counterparts. In the nucleus, lncRNA regulates transcription by interacting with and remodeling chromatin and establishing the spatial organization of nuclear compartments through their scaffold function (Bridges et al., 2021). Additionally, lncRNA may interact with splicing factors and regulate splicing events (Ouyang et al., 2022). In the cytoplasm, lncRNAs play a crucial role in the post-transcriptional regulation of gene expression and influence several pathways by, for instance, acting as sponges for miRNAs, thereby regulating the expression of miRNA target genes and interacting with RBPs to modulate their activity and influence mRNA stability and translation (Noh et al., 2018; Aillaud and Schulte, 2020). However, many lncRNAs exhibit dynamic subcellular localization in response to pathological stimuli. In these contexts, the altered localization of lncRNAs can profoundly affect cellular function, gene expression, and disease outcomes. In the literature, we can find plenty of lncRNA that change location in certain diseases, but cancer is the most common. Under normal conditions, MALAT1 is primarily located in the nucleus, retained within nuclear speckles (Miyagawa et al., 2012). However, the translocation of MALAT1 from the nucleus to the cytoplasm can occur under various cellular conditions, particularly in response to cellular stress, such as oxidative stress, or during certain pathological conditions like cancer (He et al., 2018). MALAT1 has been shown to interact with microRNAs (miRNAs) in the cytoplasm, which suggests that it may function as a molecular sponge for these miRNAs. This interaction can influence the availability and activity of the miRNAs, thereby affecting gene expression and cellular processes (He et al., 2018). Nevertheless, the exact trigger for this translocation is not clear.

Another case is the lncRNA Taurine Upregulated Gene 1 (TUG1), an oncogenic lncRNA whose splicing process determines TUG1 localization. Intron retention promotes nuclear compartmentalization, while fully spliced transcripts are found in the cytoplasm (Dumbović et al., 2021). Although intron retention in nuclear TUG1 was increased in HeLa and U-2 OS cells, the functional consequences remain elusive.

Comprehending the subcellular localization of lncRNAs is not just a scientific pursuit but a potential pathbreaker in molecular biology, genetics, and bioinformatics. This knowledge is crucial for unraveling the diverse functions of lncRNAs in cellular physiology and pathology. However, this path has only begun to be traveled, and many challenges still need to be overcome due to the dynamic nature of the lncRNA and the complexity of their interactions associated with low gene expression. Also, it cannot be ignored that some of these molecules may have localization context-dependent, varying with cellular conditions, developmental stages, or disease states.

3.2 Lineage and tissue-specific expression

The regulatory action of lncRNA is directly related to their spatial expression across different tissues, being more tissue-specific than protein-coding gene expression (Lv et al., 2020). In this sense, many lncRNA expressions are confined or more expressed in a specific tissue while, at the same time, showing lower levels in other tissues. In this scenario, a comprehensive picture of the lncRNA landscape is essential once a dysfunction of its activity could be related to a tissue-specific pathology. While lncRNA expression levels are typically lower than mRNA, they show stronger tissue-specific spatial expression patterns, suggesting that they may have a differential role in specific cell types (Lv et al., 2020; Bridges et al., 2021). As mentioned above, lncRNA genes show less overall sequence conservation than protein-coding genes; however, their promoter sequence and transcription factor binding sites are conserved, indicating that a conserved regulatory mechanism governs lncRNA transcription (Mattioli et al., 2019).

Despite this tissue-specific expression preference, some lncRNAs are ubiquitously expressed in almost all tissues, indicating a universal housekeeping function (Jiang et al., 2016; Xu et al., 2023) showed that the testis, brain, and kidney expressed a significant fraction of lncRNAs, but only a small number were among the highly expressed genes (Xu et al., 2023). In contrast, the liver and muscle expressed fewer lncRNAs (Xu et al., 2023). More than 95% of lncRNA expressed in the muscle were among the top thousand highly expressed genes; in the testis, however, less than 63% were among the top expressed genes (Xu et al., 2023). The authors suggested that the difference observed between these tissues could be explained by the more specialized nature, and the greater the cell types number, the greater the lncRNAs proportion present in abundance (Xu et al., 2023). In addition, a comparative analysis across 94 samples and 20 tissue types showed that 1,184 lncRNAs expressed in all evaluated tissues have a higher expression level pattern than lncRNA exclusively expressed in only one tissue, suggesting that ubiquitous lncRNA may have higher expression levels than tissue-specific lncRNA (Jiang et al., 2016). Furthermore, lncRNA ubiquitously expressed tends to have the highest sequence

conservation levels, fewer exons, and fewer isoform transcripts than tissue-specific expressed lncRNA (Jiang et al., 2016). This result is consistent with the observed for housekeeping protein-coding genes. Combined with these ubiquitous RNA's higher expression levels, it suggests a selective pressure for energy conservation, resulting in a more compacted gene structure that minimizes transcription and processing energetic costs (Jiang et al., 2016). An alternative hypothesis suggests that tissue-specific protein-coding genes have a longer gene architecture due to regulatory and functional complexity because these genes show more functional domains than housekeeping genes (Vinogradov, 2004). Also, tissue-specific protein-coding genes could be subjected to optimized chromatin suppression and more complex regulation of its expression (Vinogradov, 2004). This hypothesis is still a matter of discussion, especially considering lncRNAs, where the differences between housekeeping and tissue-specific lncRNA are still under exploration.

Jiang et al. found that 76.5% of tissue-specific lncRNA genes are located in intergenic regions and are targeted by fewer transcription factors and regulatory miRNA than ubiquitously expressed lncRNA (Jiang et al., 2016). This result suggests that lncRNA ubiquitously expressed are under the strictest regulation than tissue-specific lncRNA (Jiang et al., 2016).

Mattioli et al. showed that the core promoter sequence is a determinant that explains tissue specificity. They proposed that highly abundant genes have complex and promiscuous transcription factor binding sites (Mattioli et al., 2019). The authors showed that the overlapping transcription factor binding sites are more frequent in ubiquitously expressed genes, contrasting tissue-specific genes showing fewer overlapping motifs (Mattioli et al., 2019). In addition, it was found that the intergenic lncRNA has less complex transcription factor motifs at the core promoter and fewer overlapping motifs, corroborating with the overall lower expression and higher tissue specificity described for most lncRNAs (Mattioli et al., 2019).

Another candidate that could contribute to this tissue specificity regulation observed in lncRNA is the transposable elements (TE) that could influence the lncRNA regulatory network (Kelley and Rinn, 2012; Chishima et al., 2018). One example is the long terminal repeats (LTR) of endogenous retroviruses (ERV) that could act as alternative promoters (Kelley and Rinn, 2012). LTR/ERV elements generally are silenced in most human tissues except for a subset of family members that become active in some tissue or cell types (Seifarth et al., 2005; Kelley and Rinn, 2012; Tokuyama et al., 2018). It was observed that lncRNA containing ERV1 element was more likely expressed in the testis than lncRNA containing Alu element, which was less likely expressed in this tissue (Chishima et al., 2018). Interestingly, the LIPA2 element was found to act as lncRNA promoter, driving the tissue-specific lncRNA expression specifically in the placenta (Chishima et al., 2018).

Furthermore, it is known that TE are prevalent in lncRNAs exons relative to protein-coding gene exons, which are directly related to their composition and diversification (Chishima et al., 2018). Johnson and Guigó proposed that TE plays a fundamental role in composing functional domains in lncRNA and drives regulatory lncRNA evolution (Johnson and Guigó, 2014). In fact, several pieces of evidence have proposed TE as a functional element of the lncRNA (Kapusta et al., 2013; Chishima et al., 2018; Fort et al.,

2021). A comparative analysis between primate species considering overall genes showed that genome regions associated with transcription and development are devoided of TE, possibly due to a stronger selective pressure that avoids TE inserted in genes with essential functions and their vicinity (Mortada et al., 2010). Corroborating with this matter, TE-free genes and their surrounding regions are highly conserved across primates compared to TE-rich genes. However, TE-derived lncRNA exons showed higher purifying selection than non-TE-derived sequences, suggesting that at least a subset of lncRNA sequences are under purifying selection constraints (Kapusta et al., 2013).

lncRNA exhibits spatiotemporal pattern expression and has emerged as an essential tissue physiology regulator. A study considering 1340 post-mortem human brain samples of different brain regions during distinct developmental stages showed that temporal changes in lncRNA expression across fetal development are higher than those observed in the postnatal period through late adulthood (Zhang X. Q. et al., 2017). However, fewer differences in lncRNA expression were observed during the prenatal period across different brain regions, contrasting with those found in postnatal and adult stages (Zhang X. Q. et al., 2017). In these last stages, the cerebellar cortex revealed the most distinguishable lncRNA expression compared to the mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus, and neocortex regions (Zhang X. Q. et al., 2017). Considering samples of the neurodegenerative Alzheimer's and Parkinson's disease, the authors showed that many lncRNA involved in these disorders were related to fetal brain development (Zhang X. Q. et al., 2017). These results prove that lncRNAs are regulated spatiotemporally, deeply influencing neurodevelopment (Zhang X. Q. et al., 2017). In fact, several lncRNA genes are upregulated in Alzheimer's disease, such as BACE1-AS, 51A, and LRP1-AS, and the lncRNA 51A has been proposed as a biomarker to evaluate cognitive decline (Srinivas et al., 2023; Fang et al., 2024).

Exploring the landscape of ubiquitously expressed and tissue-specific lncRNA is essential to better understanding these molecules in different pathologies. More importantly, it highlights their potential as diagnostic biomarkers and therapeutic targets in various diseases.

3.3 *Cis* and *trans*: mechanism of action

To execute their regulatory role, lncRNAs interact with specific DNA target sequences, other RNAs, or RNA-binding proteins (RBPs), which could work in *cis* or *trans*-acting mode (Tao et al., 2023). In the *cis*-acting lncRNA mode, they can execute their function at the site of transcription, affecting neighboring genes, while the lncRNA that function in the *trans*-acting mode can affect their target at distant locations of its own transcription (Tao et al., 2023). The *cis*- and *trans*-regulatory roles of lncRNAs are fundamental to understanding the magnitude with which these molecules affect cellular biological processes and how they can interfere with pathological conditions.

Cis-acting lncRNA executes its regulatory function by targeting genes at the same loci as the lncRNA gene. *Cis*-acting lncRNA can act as repressors and activators in targets as distant as one hundred base pairs to hundreds of kilobases (Dhaka et al., 2024). This may

occur when *cis*-regulatory lncRNAs modulate gene expression by the recruitment and/or influencing the function of regulatory factors, such as preinitiation complex formation and transcription factors (Ma et al., 2013; Zhao et al., 2024). Yet, *cis*-acting lncRNA may affect the expression of neighbor genes by different mechanisms. One way to perform their activity directly relates to regulating chromatin states by using chromatin looping or local domains created by preferential chromatin interactions called “topologically associating domains” (Dhaka et al., 2024). Also, it was proposed that *cis*-acting lncRNA promotes chromatin remodeling by recruiting chromatin modification complexes, such as the polycomb repressive complex (PRC) (Ma et al., 2013). Another mechanism involves *cis*-acting lncRNA transcribed from enhancer regions (Gil and Ulitsky, 2020). Enhancer lncRNA (e-lncRNA) have been found overlapped to enhancer regions, and it is proposed that these transcripts contribute to modulating the enhancer activity, which has been shown that enhancer-producing lncRNA has a stronger activity compared to enhancers that do not (Gil and Ulitsky, 2020). It is essential to distinguish e-lncRNA from enhancer RNA (e-RNA), which can be accomplished by their size and stability once e-RNA are shorter and unstable (Gil and Ulitsky, 2020).

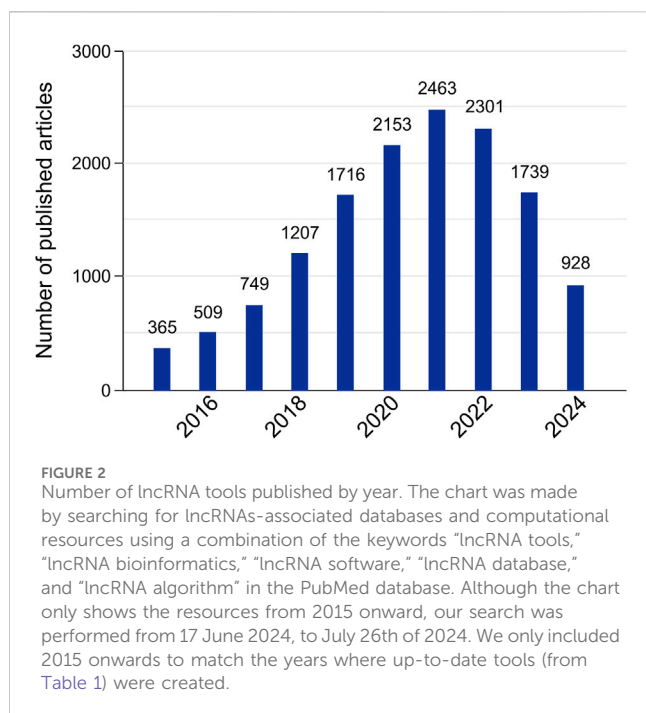
In contrast, *trans*-acting lncRNAs demonstrate a versatile regulatory function, with no apparent preference for regulating targets in specific distances or positions. Similar to mRNAs, after transcription and processing, *trans*-acting lncRNAs exert their function in a destination independent of their transcription site (Gil and Ulitsky, 2020). They regulate chromatin state and gene expression, interact and regulate different biomolecules in different cellular locations, and influence nuclear structure and organization (Zhao et al., 2024). The localization of *trans*-acting lncRNAs in the cytoplasm, organelles, other cellular structures, and/or nucleoplasm supports their identification as *trans*-regulatory lncRNAs. (Zibitt et al., 2021).

The origin of new lncRNA could be the result of different mechanisms responsible for the origin and diversity of lncRNA loci, such as sequence duplication, loss of coding potential of protein-coding genes, formation of the new transcriptional unit due to the integration of TE, mutations that alter the splicing process, and mutation that contribute to the creation of a favorable combination of sequences, such as promoters, polyadenylation elements, and splice sites (Ulitsky, 2016). All these mechanisms represent possibilities for the emergence of a new lncRNA from a locus previously transcriptionally silent (Gil and Ulitsky, 2020). In addition, it has been suggested that enhancer and promoter regions could be favorable sources for novel transcripts (Gil and Ulitsky, 2020; Palazzo and Koonin, 2020). It is hypothesized that lncRNA derived from protein-coding genes are more likely to act in *trans* than enhancer-derived lncRNA, which is more likely to act in *cis* (Gil and Ulitsky, 2020). Interestingly, 30%–60% of transcribed lncRNA are derived from regions with enhancer characteristics, and their transcription start sites tend to overlap these regions (Gil and Ulitsky, 2020). However, besides these lncRNAs initially showing a *cis*-activity by promoting chromatin remodeling and locally coordinating gene expression, it is proposed that progressively over time, *cis*-acting lncRNA become *trans*-regulators or, in some cases, carry both functions (Gil and Ulitsky, 2020; Palazzo and Koonin, 2020).

An important subset of lncRNAs is the non-coding natural antisense transcripts (NATs) that can be defined as transcripts originating from the antisense DNA strand of coding or non-coding sense transcripts (Balbin et al., 2015; Werner et al., 2024). These ncNATs are known to act as regulatory components of gene expression by either *cis*-acting, which refers to ncNATs that regulate their sense counterparts, or *trans*-acting, which originate from genomic regions that differ from their regulatory targets' origins. The dysregulation of ncNATs can be observed in multiple human diseases such as breast cancer (Iranpour et al., 2016), Angelman's disease (Meng et al., 2012), and Alzheimer's (Faghihi et al., 2008), indicating that these regulatory molecules could play an essential role in distinct diseases. On the subject of spatial information concerning ncNATs, it is essential to note that, as observed in other types of lncRNA, ncNATs' transcription levels can fluctuate significantly across different tissues and cell types (Werner et al., 2007; Balbin et al., 2015). Antisense RNAs can be found both in the nucleus and cytoplasm but tend to be concentrated in the nucleus (Lee et al., 2014). Finally, general information on ncNATs is still limited, given that the field of study on these molecules is still highly recent. Yet, the current advances in this area hint at optimistic perspectives on developing ncNATS comprehensive knowledge and its applications in science. We recommend the following reviews for more detailed information on ncNATS (Wight and Werner, 2013; Krappinger et al., 2021; Werner et al., 2024), as it is not the main focus of this review.

4 Investigating the lncRNA world with bioinformatics tools

The post-genomic era is marked by the constant development of techniques and increased analytic complexity. This term described the myriads of omics science that emerged after human genome sequencing. Altogether, the advancement of large-scale technologies has enabled the growing generation of data that allows advances in molecular characterization under different conditions and pathologies. This colossal generation of data is attractive to the advancement of research and valuable material for studying different molecules, which could have been initially ignored by the original research, such as lncRNAs. In this sense, bioinformatics has been representing an area of paramount importance, playing an essential interdisciplinary role in the development of methodological approaches to decipher and understand the massive amount of data generated, integrating and organizing different biological evidence. However, although many lncRNAs were well-characterized, the function of most lncRNAs remains unknown or poorly understood. In addition, the interest in lncRNA emerged only in the last decade and has been a source of debate. Different approaches have emerged with the fast advance of next-generation sequencing to solve this question. lncRNA-centric methods are used to investigate the interactions between known lncRNA and their molecular partners, such as RNA pull-down and ChIRP (chromatin isolation by RNA purification) (Tao et al., 2023). In this situation, preliminary experimentation is necessary to determine a specific lncRNA of interest and, consequently, if there is a limited detection capacity. For this reason, more methods focused on throughput approaches have



been explored to comprehend better lncRNA and their interactions with RNA, DNA, and proteins (Tao et al., 2023).

In the last few years, there has been a great increase in the research interest regarding lncRNA, as observed in the graph in Figure 2. With that, numerous bioinformatics resources dedicated to studying lncRNA have been developed. In this work, we compiled some of these tools that have been updated/released since 2015. The resources are collected in Table 1.

In general, most of the investigated tools were developed after 2015, and only a small portion of the ones released in earlier years have not been updated after 2015. This corroborates with the crescent research interest in lncRNA and the consequent increased demand for computational resources that may assist such studies. Additionally, it is worth mentioning that a significant number of the tools, especially databases, have received regular updates since their launch.

We also found that approximately 40% of the resources are not specific to lncRNAs and support other types of ncRNA. A considerable portion of these are tools developed for the general area of ncRNAs and, therefore, contain some information on or support for lncRNA.

When it comes to the species contemplated in each of these computational resources, we observed that most of them support few species, and, in most cases, these few species include humans. This sort of “bias” most likely results from the known role of lncRNAs in the pathogenesis of numerous human diseases and their medical importance as biomarkers and therapeutic targets. Furthermore, it is also notable that a few resources support plants despite the recently unveiled importance of lncRNA in various regulatory processes in the organisms of this taxa.

AnnoLnc2, LncExpDB, and RNALocate (Table 1) represent metasearch resources with multi-functionalities that include tissue-specific expression and subcellular location (Ke et al., 2020; Li Z. et al., 2021; Wu et al., 2024). In addition, AnnoLnc2 extends the

information and offers data regarding regulation and partner interaction, secondary structure and genomic location, genetic association, and sequence conservation (Ke et al., 2020). Otherwise, despite Lnc2Cancer3.0 and DIANA-LncBaseV3 (Table 1) being repositories focused on human cancer and miRNA targets, respectively, both databases show relevant and interesting data about lncRNA that could complement or support information about the expression profile of lncRNA in different cell types and tissues (Karagkouni et al., 2020; Gao et al., 2021). In addition, DIANA-Lnc2Cancer3.0 offers user survival analysis, correlation analysis, and transcription factor motif prediction that could be very helpful depending on the user project context (Gao et al., 2021).

When we consider prediction tools or databases to unravel the lncRNA subcellular localization, we found numerous resources that could help investigate this matter (Table 1). Several experimental methods can be used to map lncRNA to their cell compartments, such as fluorescence *in situ* hybridization (FISH), fluorescent *in situ* sequencing (FISSEQ), and subcellular RNA sequencing (subcRNaseq), which only the least is capable of potentially mapping whole transcriptome (Mas-ponte et al., 2017). LncAtlas (Table 1) is a comprehensive database that provides information on lncRNA localization that is expressed in units of Relative Concentration Index (RCI), which represents the comparison between two cellular compartments (Mas-ponte et al., 2017). It provides valuable information based on experimental data prevented from RNA-seq datasets from 15 cell lines containing information about subcellular localization. In this sense, these databases are limited by the lncRNA captured in previous experiments. In contrast, prediction tools such as LncLocator, DeepLncLoc, and iLoc-lncRNA (Table 1) use the primary sequence of lncRNA to infer the subcellular localization using a benchmark dataset (Su et al., 2018; Lin et al., 2021; Zeng et al., 2022). These tools use machine learning approaches, such as SVM + Random Forest, SVM, and Deep Neural Network, respectively, inputted by k-mers features to achieve a decent performance in predicting subcellular localization of lncRNAs (Zeng et al., 2022).

Predictive analysis of lncRNA localization involves complex algorithms that estimate the compartment distribution based on sequence features, interactions, and expression patterns. However, these models often face limitations in their ability to accurately predict localization, particularly for novel or poorly characterized lncRNAs. Additionally, the development and optimization of computational tools must balance computational efficiency and the ability to provide detailed and accurate localization predictions. The quality of input data can influence the accuracy of predictions, the quality of benchmarks used during model training, and the sophistication of the underlying algorithms.

Investigating whether a lncRNA acts in a *cis*- or *trans*-mode involves utilizing various bioinformatic tools and approaches to elucidate its functional mechanisms. To determine the mode of action of lncRNAs, researchers employ a combination of computational tools and databases to analyze expression patterns, chromatin interactions, and genetic perturbations. However, in most cases, several studies use methods to infer *cis*- or *trans*-regulation based on indirect relations. For instance, bioinformatic tools like LncRNA2Target (Table 1) facilitate the prediction of lncRNA-target interactions (Cheng et al., 2019). LncRNA2Target

is a curated database to investigate the lncRNA-target partnership using mainly lncRNA knockdown and overexpression experiments, luciferase reporter assays, immunoprecipitation assays, and RNA pull-down assays as a source of information (Cheng et al., 2019). Dhaka et al. developed the framework TransCistor (Table 1) that proposed elucidating *cis*-regulatory relationship by using loss-of-function perturbations from FANTOM6 repositories to identify target genes of lncRNA (Dhaka et al., 2024). The author considers a molecule as a lncRNA target when the steady-state level of this molecule is significantly altered in response to the loss-of-function of the lncRNA (Dhaka et al., 2024). Another approach used is to evaluate the gene co-expression analysis to compare the gene expression between the lncRNA and their neighbor to infer if the target expression could be influenced by lncRNA (Li et al., 2017; Wu et al., 2019). Thus, this approach can help identify potential *cis*-acting targets by examining whether lncRNAs and their target genes are co-expressed and physically proximal in the genome.

Additionally, Chromosome Conformation Capture coupled with High Throughput Sequencing (Hi-C), Hi-C coupled with Chromatin Isolation by RNA Purification (Hi-ChIRP), and Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET) are examples of methods that provide insights into the three-dimensional interactions between lncRNAs and genomic regions, revealing potential *cis*- or *trans*-acting relationships (Saxena and Carninci, 2011; Ariel et al., 2021). These approaches help determine whether lncRNAs interact with nearby or distant genomic regions.

Despite these advancements, predicting lncRNA localization remains complex due to the dynamic nature of lncRNA interactions and cellular environments. Future developments are expected to focus on integrating multi-omic data and improving model accuracy. The complexity of the transcriptome, including the presence of multiple isoforms across tissues and overlapping regions between lncRNAs and protein-coding genes, increases the difficulty of finding a solution.

Furthermore, public databases and repositories may have incomplete or inconsistent data coverage across different tissues. Available datasets often represent only a limited subset of tissues, which may only partially capture the tissue-specific expression profiles of lncRNAs. This limitation can hinder identifying and characterizing lncRNAs with unique or restricted expression patterns. Variability in data quality due to differences in sample preparation, sequencing platforms, and experimental protocols can impact the reliability of expression measurements and subsequent analyses.

Investigating lncRNA localization or tissue-specific expression typically involves integrating various types of omics data, such as transcriptomic, proteomic, and imaging data. The differences in data formats, resolutions, and scales complicate the integration of all this information. Bioinformatics tools must accurately differentiate between isoforms accounting for their specific localizations and effectively combine these disparate data types to provide a coherent scenario of lncRNA. The challenges lie in developing algorithms that integrate multi-dimensional data while minimizing artifacts and maintaining biological relevance.

5 Conclusion

Given their low overall expression and tissue/cell-line specificity, the understanding of lncRNAs remains complex. Localizing lncRNAs

within the subcellular compartments is a crucial determinant of their molecular functions, playing a significant role in cellular processes. As lncRNA can interact with a wide range of molecules, their localization and local partner interaction are crucial to understanding and predicting their function (Bridges et al., 2021). Based on the experimental information in public repositories associated with *in silico* prediction techniques, computational tools can be a differential approach to guide experiments to evaluate a hypothesis, saving time and financial resources. However, ideally, different experimental validation approaches are necessary to prove the findings, such as fluorescent in-situ hybridization (FISH)-based methods and MS2-system-based techniques (Choudhury et al., 2024). Ultimately, further studies focused on specific cases will be needed to expand our current, albeit incomplete, knowledge of lncRNA biology. Fortunately, new-generation technologies and bioinformatics tools have advanced our understanding of these regulatory molecules. To further expand our knowledge, future research should focus on elucidating the specific mechanisms by which lncRNAs regulate gene expression and cellular processes context-dependently. This will involve investigating the intricate interplay between lncRNAs and other cellular components, such as proteins, DNA, and other RNA species, and how these complex interactions contribute to lncRNA-mediated regulation of cellular homeostasis and disease pathogenesis.

Author contributions

JP: Conceptualization, Data curation, Formal Analysis, Investigation, Validation, Visualization, Writing–original draft, Writing–review and editing. FO: Data curation, Resources, Validation, Writing–original draft, Writing–review and editing. BF: Conceptualization, Investigation, Project administration, Supervision, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant (151491/2024-0), by the Fundação de Amparo a Pesquisa do Estado do Rio Grande do Sul (FAPERGS) grant (24/2551-0001277-0), and partially funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), finance code 001.

Acknowledgments

We want to thank all the funding agencies that provided financial aid for this project.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Ahmad, A., Lin, H., and Shatabda, S. (2020). Locate-R: subcellular localization of long non-coding RNAs using nucleotide compositions. *Genomics* 112, 2583–2589. doi:10.1016/j.ygeno.2020.02.011
- Ahmad, S., Mudasir, S., Tajamul, P., Ahad, A., Ahmad, M., Urwat, U., et al. (2016). Long non-coding RNAs: mechanism of action and functional utility. *Non-coding RNA Res.* 1, 43–50. doi:10.1016/j.ncrna.2016.11.002
- Aillaud, M., and Schulte, L. N. (2020). Emerging roles of long noncoding RNAs in the cytoplasmic milieu. *Noncoding RNA* 6, 44. doi:10.3390/ncrna6040044
- 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, 146–151. doi:10.1093/nar/gkq1138
- Anilkumar, A. K., Vij, P., Lopez, S., Leslie, S. M., Doxtater, K., Khan, M. M., et al. (2024). Long non-coding RNAs: new insights in neurodegenerative diseases. *Int. J. Mol. Sci.* 25, 2268. doi:10.3390/ijms25042268
- Antonov, I. V., Mazurov, E., Borodovsky, M., and Medvedeva, Y. A. (2019). Prediction of lncRNAs and their interactions with nucleic acids: benchmarking bioinformatics tools. *Brief. Bioinform.* 20, 551–564. doi:10.1093/bib/bby032
- Ariel, F., Fonouni-farde, C., and Ariel, F. (2021). Functional classification of plant long noncoding RNAs: a transcript is known by the company it keeps. *New Phytol.* 229, 1251–1260. doi:10.1111/nph.16903
- Balbin, O. A., Malik, R., Dhanasekaran, S. M., Prensner, J. R., Cao, X., Wu, Y.-M., et al. (2015). The landscape of antisense gene expression in human cancers. *Genome Res.* 25, 1068–1079. doi:10.1101/gr.180596.114
- Basu, K., Dey, A., and Kiran, M. (2023). Inefficient splicing of long non-coding RNAs is associated with higher transcript complexity in human and mouse. *RNA Biol.* 20, 563–572. doi:10.1080/15476286.2023.2242649
- Brannan, C. I., Dees, E. C., Ingram, R. S., and Tilghman, S. M. (1990). The product of the H19 gene may function as an RNA. *Mol. Cell Biol.* 10, 28–36. doi:10.1128/mcb.10.1.28-36.1990
- Bridges, M. C., Daulagala, A. C., and Kourtidis, A. (2021). LNCcation: lncRNA localization and function. *J. Cell Biol.* 220, e202009045. doi:10.1083/jcb.202009045
- Carlevaro-Fita, J., and Johnson, R. (2019). Global positioning system: understanding long noncoding RNAs through subcellular localization. *Mol. Cell* 73, 869–883. doi:10.1016/j.molcel.2019.02.008
- Chen, B., Dragomir, M. P., Yang, C., Li, Q., Horst, D., and Calin, G. A. (2022). Targeting non-coding RNAs to overcome cancer therapy resistance. *Signal Transduct. Target Ther.* 7, 121. doi:10.1038/s41392-022-00975-3
- Chen, J., Lin, J., Hu, Y., Ye, M., Yao, L., Wu, L., et al. (2023). RNADisease v4.0: an updated resource of RNA-associated diseases, providing RNA-disease analysis, enrichment and prediction. *Nucleic Acids Res.* 51, 1397–1404. doi:10.1093/nar/gkac814
- Chen, J., Zhang, J., Gao, Y., Li, Y., Feng, C., Song, C., et al. (2021). LncSEA: a platform for long non-coding RNA related sets and enrichment analysis. *Nucleic Acids Res.* 49, 969–980. doi:10.1093/nar/gkaa806
- Cheng, L., Wang, P., Tian, R., Wang, S., Guo, Q., Luo, M., et al. (2019). lncRNA2Target v2.0: a comprehensive database for target genes of lncRNAs in human and mouse. *Nucleic Acids Res.* 47, D140–D144. doi:10.1093/nar/gky1051
- Chishima, T., Iwakiri, J., and Hamada, M. (2018). Identification of transposable elements contributing to tissue-specific expression of long non-coding RNAs. *Genes (Basel)* 9, 23. doi:10.3390/genes9010023
- Choudhury, S., Rathore, A. S., and Raghava, G. P. S. (2024). Compilation of resources on subcellular localization of lncRNA. *Front. RNA Res.* 2. doi:10.3389/frnar.2024.1419979
- Clark, M. B., Johnston, R. L., Inostroza-Ponta, M., Fox, A. H., Fortini, E., Moscatto, P., et al. (2012). Genome-wide analysis of long noncoding RNA stability. *Genome Res.* 22, 885–898. doi:10.1101/gr.131037.111
- Dhaka, B., Zimmerli, M., Hanhart, D., Moser, M. B., Guillen-Ramirez, H., Mishra, S., et al. (2024). Functional identification of cis-regulatory long noncoding RNAs at controlled false discovery rates. *Nucleic Acids Res.* 52, 2821–2835. doi:10.1093/nar/gkac075
- Diamantopoulos, M. A., Tsiakanikas, P., and Scorilas, A. (2018). Non-coding RNAs: the riddle of the transcriptome and their perspectives in cancer. *Ann. Transl. Med.* 6, 241. doi:10.21037/atm.2018.06.10
- Dragomir, M., Chen, B., and Calin, G. A. (2018). Exosomal lncRNAs as new players in cell-to-cell communication. *Transl. Cancer Res.* 7, 243–252. doi:10.21037/tcr.2017.10.46
- Dumbović, G., Braunschweig, U., Langner, H. K., Smallegan, M., Biayna, J., Hass, E. P., et al. (2021). Nuclear compartmentalization of TERT mRNA and TUG1 lncRNA is driven by intron retention. *Nat. Commun.* 12, 3308. doi:10.1038/s41467-021-23221-w
- Faghihi, M. A., Modarresi, F., Khalil, A. M., Wood, D. E., Sahagan, B. G., Morgan, T. E., et al. (2008). Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat. Med.* 14, 723–730. doi:10.1038/nm1784
- Fang, H., Li, J., Zhang, L., Li, B., Song, J., Lu, X., et al. (2024). lncRNA 51A: a promising diagnostic biomarker for assessing cognitive decline in occupationally exposed aluminum workers. *Environ. Toxicol. Pharmacol.* 111, 104548. doi:10.1016/j.etap.2024.104548
- Fang, S., Zhang, L., Guo, J., Niu, Y., Wu, Y., Li, H., et al. (2018). NONCODEV5: a comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Res.* 46, 308–314. doi:10.1093/nar/gkx1107
- Feng, H., Wang, S., Wang, Y., Ni, X., Yang, Z., Hu, X., et al. (2023). LncCat: an ORF attention model to identify lncRNA based on ensemble learning strategy and fused sequence information. *Comput. Struct. Biotechnol. J.* 21, 1433–1447. doi:10.1016/j.csbj.2023.02.012
- Feng, S., Liang, Y., Du, W., Lv, W., and Li, Y. (2020). LncLocation: Efficient subcellular location prediction of long non-coding RNA-based multi-source heterogeneous feature fusion. *Int. J. Mol. Sci.* 21, 7271. doi:10.3390/ijms21197271
- Fort, V., Khelifi, G., and Hussein, S. M. I. (2021). Long non-coding RNAs and transposable elements: a functional relationship. *Biochim. Biophys. Acta - Mol. Cell Res.* 1868, 118837. doi:10.1016/j.bbamcr.2020.118837
- Frankish, A., Carbonell-Sala, S., Diekhans, M., Jungreis, I., Loveland, J. E., Mudge, J. M., et al. (2023). GENCODE: reference annotation for the human and mouse genomes in 2023. *Nucleic Acids Res.* 51, D942–D949. doi:10.1093/nar/gkac1071
- Frankish, A., Diekhans, M., Ferreira, A., Johnson, R., Jungreis, I., Loveland, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* 47, 766–773. doi:10.1093/nar/gky955
- Fukunaga, T., and Hamada, M. (2017). RIBlast: an ultrafast RNA–RNA interaction prediction system based on a seed-and-extension approach. *Bioinformatics* 33, 2666–2674. doi:10.1093/bioinformatics/btx287
- Fukunaga, T., Iwakiri, J., Ono, Y., and Hamada, M. (2019). LncRRISearch: a web server for lncRNA-RNA interaction prediction integrated with tissue-specific expression and subcellular localization data. *Front. Genet.* 10, 462. doi:10.3389/fgene.2019.00462
- Gao, F., Cai, Y., Kapranov, P., and Xu, D. (2020). Reverse-genetics studies of lncRNAs—what we have learnt and paths forward. *Genome Biol.* 21, 93. doi:10.1186/s13059-020-01994-5
- Gao, Y., Shang, S., Guo, S., Li, X., Zhou, H., Liu, H., et al. (2021). Lnc2Cancer 3.0: an updated resource for experimentally supported lncRNA/circRNA cancer associations and web tools based on RNA-seq and scRNA-seq data. *Nucleic Acids Res.* 49, 1251–1258. doi:10.1093/nar/gkaa1006
- Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. *Nat. Rev. Genet.* 21, 102–117. doi:10.1038/s41576-019-0184-5
- Gloss, B. S., and Dinger, M. E. (2016). The specificity of long noncoding RNA expression. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1859, 16–22. doi:10.1016/j.bbagr.2015.08.005
- Gong, J., Shao, D., Xu, K., Lu, Z., Lu, Z. J., Yang, Y. T., et al. (2018). RISE: a database of RNA interactome from sequencing experiments. *Nucleic Acids Res.* 46, 194–201. doi:10.1093/nar/gkx864
- Gong, Y., Zhu, W., Sun, M., and Shi, L. (2021). Bioinformatics analysis of long non-coding RNA and related diseases: an overview. *Front. Genet.* 12, 813873–813879. doi:10.3389/fgene.2021.813873
- Guh, C.-Y., Hsieh, Y.-H., and Chu, H.-P. (2020). Functions and properties of nuclear lncRNAs—from systematically mapping the interactomes of lncRNAs. *J. Biomed. Sci.* 27, 44. doi:10.1186/s12929-020-00640-3
- Guo, C., Xu, G., and Chen, L. (2020). Mechanisms of long noncoding RNA nuclear retention. *Trends biochem. Sci.* 45, 947–960. doi:10.1016/j.tics.2020.07.001
- Hatje, K., Mühlhausen, S., Simm, D., and Kollmar, M. (2019). The protein-coding human genome: annotating high-hanging fruits. *BioEssays* 41, e1900066. doi:10.1002/bies.201900066
- He, X., Yan, Q., Kuang, G., Wang, Y., Cao, P., and Ou, C. (2018). Metastasis-associated lung adenocarcinoma transcript 1 regulates tumor progression: old wine in a new bottle. *J. Thorac. Dis.* 10, S1088–S1091. doi:10.21037/jtd.2018.04.13

- Iranpour, M., Soudyab, M., Geranpayeh, L., Mirfakhraie, R., Azargashb, E., Movafagh, A., et al. (2016). Expression analysis of four long noncoding RNAs in breast cancer. *Tumor Biol.* 37, 2933–2940. doi:10.1007/s13277-015-4135-2
- Jarroux, J., Morillon, A., and Pinskaya, M. (2017). History, discovery, and classification of lncRNAs. *Adv. Exp. Med. Biol.* 1008, 1–46. doi:10.1007/978-981-10-5203-3_1
- Jeffery, W. R., Tomlinson, C. R., and Brodeur, R. D. (1983). Localization of actin messenger RNA during early ascidian development. *Dev. Biol.* 99, 408–417. doi:10.1016/0012-1606(83)90290-7
- Jiang, C., Li, Y., Zhao, Z., Lu, J., Chen, H., Ding, N., et al. (2016). Identifying and functionally characterizing tissue-specific and ubiquitously expressed human lncRNAs. *Oncotarget* 7, 7120–7133. doi:10.18632/oncotarget.6859
- Jiang, H., Zhang, Y., Yue, J., Shi, Y., Xiao, B., Xiao, W., et al. (2022). Non-coding RNAs: the neuroinflammatory regulators in neurodegenerative diseases. *Front. Neurol.* 13, 929290. doi:10.3389/fneur.2022.929290
- Jiang, S., Cheng, S., Ren, L., Wang, Q., Kang, Y., Ding, Y., et al. (2019). An expanded landscape of human long noncoding RNA. *Nucleic Acids Res.* 47, 7842–7856. doi:10.1093/nar/gkz621
- Johnson, R., and Guigó, R. (2014). The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *Rna* 20, 959–976. doi:10.1261/rna.044560.114
- Kang, J., Tang, Q., He, J., Li, L., Yang, N., Yu, S., et al. (2022). RNAInter v4.0: RNA interactome repository with redefined confidence scoring system and improved accessibility. *Nucleic Acids Res.* 50, 326–332. doi:10.1093/nar/gkab997
- Kang, Y., Yang, D., Kong, L., Hou, M., Meng, Y., Wei, L., et al. (2017). CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res.* 45, 12–16. doi:10.1093/nar/gkx428
- Kapusta, A., Kronenberg, Z., Lynch, V. J., Zhuo, X., Ramsay, L. A., Bourque, G., et al. (2013). Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet.* 9, e1003470. doi:10.1371/journal.pgen.1003470
- Karakouni, D., Paraskevopoulou, M. D., Tastsoglou, S., Skoufos, G., Karavangeli, A., Pterros, V., et al. (2020). DIANA-LncBase v3: indexing experimentally supported miRNA targets on non-coding transcripts. *Nucleic Acids Res.* 48, 101–110. doi:10.1093/nar/gkz1036
- Kashi, K., Henderson, L., Bonetti, A., and Carninci, P. (2016). Discovery and functional analysis of lncRNAs: methodologies to investigate an uncharacterized transcriptome. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1859, 3–15. doi:10.1016/j.bbagr.2015.10.010
- Ke, L., Yang, D., Wang, Y., Ding, Y., and Gao, G. (2020). AnnoLnc2: the one-stop portal to systematically annotate novel lncRNAs for human and mouse. *Nucleic Acids Res.* 48, 230–238. doi:10.1093/nar/gkaa368
- Kelley, D., and Rinn, J. (2012). Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* 13, R107. doi:10.1186/gb-2012-13-11-r107
- Khan, M. R., Wellinger, R. J., and Laurent, B. (2021). Exploring the alternative splicing of long noncoding RNAs. *Trends Genet.* 37, 695–698. doi:10.1016/j.tig.2021.03.010
- Kirk, J. M., Kim, S. O., Inoue, K., Smola, M. J., Lee, D. M., Schertz, M. D., et al. (2018). Functional classification of long non-coding RNAs by k-mer content. *Nat. Genet.* 50, 1474–1482. doi:10.1038/s41588-018-0207-8
- Krappinger, J. C., Bonstingl, L., Pansy, K., Sallinger, K., Wreglesworth, N. I., Grinninger, L., et al. (2021). Non-coding natural antisense transcripts: analysis and application. *J. Biotechnol.* 340, 75–101. doi:10.1016/j.jbiotec.2021.08.005
- Lai, H., Li, Y., Zhang, H., Hu, J., Liao, J., Su, Y., et al. (2022). exoRBase 2.0: an atlas of mRNA, lncRNA and circRNA in extracellular vesicles from human biofluids. *Nucleic Acids Res.* 50, 118–128. doi:10.1093/nar/gkab1085
- Lanzós, A., Carlevaro-fita, J., Mularoni, L., Reverter, F., Palumbo, E., Guigó, R., et al. (2017). Discovery of cancer driver long noncoding RNAs across 1112 tumour genomes: new candidates and distinguishing features. *Nat. Publ. Gr.* 7, 41544. doi:10.1038/srep41544
- Lee, H., Zhang, Z., and Krause, H. M. (2019). Long noncoding RNAs and repetitive elements: junk or intimate evolutionary partners? *Trends Genet.* 35, 892–902. doi:10.1016/j.tig.2019.09.006
- Lee, J. H., Daugharthy, E. R., Scheiman, J., Kalhor, R., Yang, J. L., Ferrante, T. C., et al. (2014). Highly multiplexed subcellular RNA sequencing *in situ*. *Sci. (80-)* 343, 1360–1363. doi:10.1126/science.1250212
- Li, C., Ni, Y., Xu, H., Xiang, Q., Zhao, Y., Zhan, J., et al. (2021). Roles and mechanisms of exosomal non-coding RNAs in human health and diseases. *Signal Transduct. Target Ther.* 6, 383. doi:10.1038/s41392-021-00779-x
- Li, J., Han, L., Roebuck, P., Diao, L., Liu, L., Yuan, Y., et al. (2015). TANRIC: an interactive open platform to explore the function of lncRNAs in cancer. *Cancer Res.* 75, 3728–3737. doi:10.1158/0008-5472.CAN-15-0273
- Li, J., Liu, S., Zhou, H., Qu, L., and Yang, J. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42, 92–97. doi:10.1093/nar/gkt1248
- Li, Z., Liu, L., Feng, C., Qin, Y., Xiao, J., Zhang, Z., et al. (2023). LncBook 2.0: integrating human long non-coding RNAs with multi-omics annotations. *Nucleic Acids Res.* 51, 186–191. doi:10.1093/nar/gkac999
- Li, Z., Liu, L., Jiang, S., Li, Q., Feng, C., Du, Q., et al. (2021). LncExpDB: an expression database of human long non-coding RNAs. *Nucleic Acids Res.* 49, D962–D968. doi:10.1093/nar/gkaa850
- Li, Z., Ouyang, H., Zheng, M., Cai, B., Han, P., Abdalla, B. A., et al. (2017). Integrated analysis of long non-coding RNAs (lncRNAs) and mRNA expression profiles reveals the potential role of lncRNAs in skeletal muscle development of the chicken. *Front. Physiol.* 7, 687. doi:10.3389/fphys.2016.00687
- Lin, X., Lu, Y., Zhang, C., Cui, Q., Cui, C., Ji, X., et al. (2024). LncRNA Disease v3.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Res.* 52, 1365–1369. doi:10.1093/nar/gkad828
- Lin, Y., Pan, X., and Shen, H.-B. (2021). lncLocator 2.0: a cell-line-specific subcellular localization predictor for long non-coding RNAs with interpretable deep learning. *Bioinformatics* 37, 2308–2316. doi:10.1093/bioinformatics/btab127
- Liu, L., Li, Z., Liu, C., Zou, D., Li, Q., Feng, C., et al. (2022). LncRNA Wiki 2.0: a knowledgebase of human long non-coding RNAs with enhanced curation model and database system. *Nucleic Acids Res.* 50, 190–195. doi:10.1093/nar/gkab998
- Liu, Y., and Zhao, M. (2016). InCaNet: pan-cancer co-expression network for human lncRNA and cancer genes. *Bioinformatics* 32, 1595–1597. doi:10.1093/bioinformatics/btw017
- Lodde, V., Murgia, G., Simula, E. R., Steri, M., Floris, M., and Idda, M. L. (2020). Long noncoding RNAs and circular RNAs in autoimmune diseases. *Biomolecules* 10, 1044. doi:10.3390/biom10071044
- Loganathan, T., and Doss, C. G. P. (2023). *Non-coding RNAs in human health and disease: potential function as biomarkers and therapeutic targets*. Springer Berlin Heidelberg. doi:10.1007/s10142-022-00947-4
- Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., et al. (2013). The genotype-tissue expression (GTEx) project. *Nat. Genet.* 45, 580–585. doi:10.1038/ng.2653
- Lv, D., Xu, K., Jin, X., Li, J., Shi, Y., Zhang, M., et al. (2020). LncSpA: LncRNA spatial atlas of expression across normal and cancer tissues. *Cancer Res.* 80, 2067–2071. doi:10.1158/0008-5472.CAN-19-2687
- Ma, L., Bajic, V. B., and Zhang, Z. (2013). On the classification of long non-coding RNAs. *RNA Biol.* 10, 925–933. doi:10.4161/rna.24604
- Mann, M., Wright, P. R., and Backofen, R. (2017). IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. *Nucleic Acids Res.* 45, 435–439. doi:10.1093/nar/gkx279
- Mas-ponte, D., Carlevaro-fita, J., Palumbo, E., Pulido, T. H., Guigo, R., and Johnson, R. (2017). LncATLAS database for subcellular localization of long noncoding RNAs. *RNA* 23, 1080–1087. doi:10.1261/rna.060814.117
- Mattick, J. S., Amaral, P. P., Carninci, P., Carpenter, S., Chang, H. Y., Chen, L. L., et al. (2023). Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* 24, 430–447. doi:10.1038/s41580-022-00566-8
- Mattioli, K., Volders, P., Gerhardinger, C., Lee, J. C., Maass, P. G., Melé, M., et al. (2019). High-throughput functional analysis of lncRNA core promoters elucidates rules governing tissue specificity. *Genome Res.* 29, 344–355. doi:10.1101/gr.242222.118
- Meng, L., Person, R. E., and Beaudet, A. L. (2012). Ube3a-ATS is an atypical RNA polymerase II transcript that represses the paternal expression of Ube3a. *Hum. Mol. Genet.* 21, 3001–3012. doi:10.1093/hmg/ddsl30
- Miao, Y., Liu, W., Zhang, Q., and Guo, A. (2018). lncRNASNP2: an updated database of functional SNPs and mutations in human and mouse lncRNAs. *Nucleic Acids Res.* 46, 276–280. doi:10.1093/nar/gkx1004
- Miyagawa, R., Tano, K., Mizuno, R., Nakamura, Y., Ijiri, K., Rakwal, R., et al. (2012). Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. *RNA* 18, 738–751. doi:10.1261/rna.028639.111
- Mortada, H., Vieira, C., and Lerat, E. (2010). Genes devoid of full-length transposable element insertions are involved in development and in the regulation of transcription in human and closely related species. *J. Mol. Evol.* 71, 180–191. doi:10.1007/s00239-010-9376-5
- Musacchia, F., Basu, S., Petrosino, G., and Sanges, R. (2015). Annotcript: a flexible pipeline for the annotation of transcriptomes able to identify putative long noncoding RNAs. *Bioinformatics* 31, 2199–2201. doi:10.1093/bioinformatics/btv106
- Noh, J. H., Kim, K. M., McClusky, W., Abdelmohsen, K., and Gorospe, M. (2018). Cytoplasmic functions of long noncoding RNAs. *Wiley Interdiscip. Rev. RNA* 9, e1471. doi:10.1002/wrna.1471
- Ouyang, J., Zhong, Y., Zhang, Y., Yang, L., Wu, P., Hou, X., et al. (2022). Long non-coding RNAs are involved in alternative splicing and promote cancer progression. *Br. J. Cancer* 126, 1113–1124. doi:10.1038/s41416-021-01600-w
- Palazzo, A. F., and Koonin, E. V. (2020). Functional long non-coding RNAs evolve from junk transcripts. *Cell* 183, 1151–1161. doi:10.1016/j.cell.2020.09.047
- Pauli, A., Valen, E., Lin, M. F., Garber, M., Vastenhout, N. L., Levin, J. Z., et al. (2012). Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Res.* 22, 577–591. doi:10.1101/gr.133009.111

- Pian, C., Zhang, G., Tu, T., Ma, X., and Li, F. (2018). LncCerBase: a database of experimentally validated human competing endogenous long non-coding RNAs. *Database* 2018, bay061. doi:10.1093/database/bay061
- Poller, W., Dimmeler, S., Heymans, S., Zeller, T., Haas, J., Karakas, M., et al. (2018). Non-coding RNAs in cardiovascular diseases: diagnostic and therapeutic perspectives. *Eur. Hear. J.* 39, 2704–2716. doi:10.1093/eurheartj/ehx165
- Ponting, C. P., and Haerty, W. (2022). Genome-wide analysis of human long noncoding RNAs: a provocative review. *Annu. Rev. Genomics Hum. Genet.* 23, 153–172. doi:10.1146/annurev-genom-112921-123710
- Ponting, C. P., Oliver, P. L., and Reik, W. (2009). Evolution and functions of long noncoding RNAs. *Cell* 136, 629–641. doi:10.1016/j.cell.2009.02.006
- Rincón-Riveros, A., Morales, D., Rodríguez, J. A., Villegas, V. E., and López-Kleine, L. (2021). Bioinformatic tools for the analysis and prediction of ncRNA interactions. *Int. J. Mol. Sci.* 22, 11397. doi:10.3390/ijms222111397
- Ruffo, P., Straffella, C., Cascella, R., Caputo, V., Conforti, F. L., Andò, S., et al. (2021). Deregulation of ncRNA in neurodegenerative disease: focus on circRNA, lncRNA and miRNA in amyotrophic lateral sclerosis. *Front. Genet.* 12, 784996. doi:10.3389/fgene.2021.784996
- Saxena, A., and Carninci, P. (2011). Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs. *Bioessays* 33, 830–839. doi:10.1002/bies.201100084
- Scherrer, K. (2023). Is RNA the working genome in eukaryotes? The 60 year evolution of a conceptual challenge. *Exp. Cell Res.* 424, 113493. doi:10.1016/j.yexcr.2023.113493
- Schlackow, M., Nojima, T., Gomes, T., Dhir, A., Carmo-Fonseca, M., and Proudfoot, N. J. (2017). Distinctive patterns of transcription and RNA processing for human lincRNAs. *Mol. Cell* 65, 25–38. doi:10.1016/j.molcel.2016.11.029
- Seifarth, W., Frank, O., Zeilfelder, U., Spiess, B., Greenwood, A. D., Hehlmann, R., et al. (2005). Comprehensive analysis of human endogenous retrovirus transcriptional activity in human tissues with a retrovirus-specific microarray. *J. Virol.* 79, 341–352. doi:10.1128/jvi.79.1.341-352.2005
- Singh, A. K. (2024). Rules and impacts of nonsense-mediated mRNA decay in the degradation of long noncoding RNAs. *Wiley Interdiscip. Rev. RNA* 15, 1–14. doi:10.1002/wrna.1853
- Srinivas, T., Mathias, C., Oliveira-Mateos, C., and Guil, S. (2023). Roles of lncRNAs in brain development and pathogenesis: emerging therapeutic opportunities. *Mol. Ther.* 31, 1550–1561. doi:10.1016/j.yymthe.2023.02.008
- Stattell, L., Guo, C. J., Chen, L. L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* 22, 96–118. doi:10.1038/s41580-020-00315-9
- Su, Z., Huang, Y., Zhang, Z., Zhao, Y., Wang, D., Chen, W., et al. (2018). iLoc-lncRNA: predict the subcellular location of lncRNAs by incorporating octamer composition into general PseKNC. *Bioinformatics* 34, 4196–4204. doi:10.1093/bioinformatics/bty508
- Szcześniak, M. W., and Wanowska, E. (2024). CANTATAdb 3.0: an updated repository of plant long noncoding RNAs. *Plant Cell Physiol.* 17, 1486–1493. doi:10.1093/pcp/pcae081
- Tao, X., Li, S., Chen, G., Wang, J., and Xu, S. (2023). Approaches for modes of action study of long non-coding RNAs: from single verification to genome-wide determination. *Int. J. Mol. Sci.* 24, 5562. doi:10.3390/ijms24065562
- Tokuyama, M., Kong, Y., Song, E., Jayewickreme, T., Kang, I., and Iwasaki, A. (2018). ERVmap analysis reveals genome-wide transcription of human endogenous retroviruses. *Proc. Natl. Acad. Sci. U. S. A.* 115, 12565–12572. doi:10.1073/pnas.1814589115
- Tong, X., and Liu, S. (2019). CPPred: coding potential prediction based on the global description of RNA sequence. *Nucleic Acids Res.* 47, e43. doi:10.1093/nar/gkz087
- Ulitsky, I. (2016). Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. *Nat. Publ. Gr.* 17, 601–614. doi:10.1038/nrg.2016.85
- Ulitsky, I., and Bartel, D. P. (2013). lincRNAs: genomics, evolution, and mechanisms. *Cell* 154, 26–46. doi:10.1016/j.cell.2013.06.020
- Vinogradov, A. E. (2004). Compactness of human housekeeping genes: selection for economy or genomic design? *Trends Genet.* 20, 248–253. doi:10.1016/j.tig.2004.03.006
- Volders, P., Anckaert, J., Verheggen, K., Nuytens, J., Martens, L., Mestdagh, P., et al. (2019). LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Res.* 47, 135–139. doi:10.1093/nar/gky1031
- Wang, P., Guo, Q., Qi, Y., Hao, Y., Gao, Y., Zhi, H., et al. (2022). LncACTdb 3.0: an updated database of experimentally supported ceRNA interactions and personalized networks contributing to precision medicine. *Nucleic Acids Res.* 50, 183–189. doi:10.1093/nar/gkab1092
- Werner, A., Kanhere, A., Wahlestedt, C., and Mattick, J. S. (2024). Natural antisense transcripts as versatile regulators of gene expression. *Nat. Rev. Genet.* 25, 730–744. doi:10.1038/s41576-024-00723-z
- Werner, A., Schmutzler, G., Carlile, M., Miles, C. G., and Peters, H. (2007). Expression profiling of antisense transcripts on DNA arrays. *Physiol. Genomics* 28, 294–300. doi:10.1152/physiolgenomics.00127.2006
- Wight, M., and Werner, A. (2013). The functions of natural antisense transcripts. *Essays Biochem.* 54, 91–101. doi:10.1042/bse0540091
- Wilusz, J. E., Jnbaptiste, C. K., Lu, L. Y., Kuhn, C., Joshua-tor, L., and Sharp, P. A. (2012). A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly (A) tails. *Genes Dev.* 26, 2392–2407. doi:10.1101/gad.204438.112
- Wu, Y., Guo, J., Wang, T., Cao, F., and Wang, G. (2019). Transcriptional profiling of long noncoding RNAs associated with leaf-color mutation in Ginkgo biloba L. *BMC Plant Biol.* 19, 527. doi:10.1186/s12870-019-2141-z
- Wu, L., Wang, L., Hu, S., Tang, G., Chen, J., Yi, Y., et al. (2024). RNALocate v3.0: advancing the repository of rna subcellular localization with dynamic analysis and prediction. *Nucleic Acids Res.* gkae872. doi:10.1093/nar/gkae872
- Xu, K., Jin, X., Luo, Y., Zou, H., Lv, D., Wang, L., et al. (2023). Spatial transcriptome analysis of long non-coding RNAs reveals tissue specificity and functional roles in cancer. *J. Zhejiang Univ. Sci. B* 24, 15–31. doi:10.1631/jzus.B2200206
- Yang, C., Yang, L., Zhou, M., Xie, H., Zhang, C., Wang, M. D., et al. (2018). LncADeep: an *ab initio* lncRNA identification and functional annotation tool based on deep learning. *Bioinformatics* 34, 3825–3834. doi:10.1093/bioinformatics/bty428
- Yoon, J.-H., Kim, J., and Gorospe, M. (2015). Long noncoding RNA turnover. *Biochimie* 117, 15–21. doi:10.1016/j.biochi.2015.03.001
- Yotsukura, S., du Verle, D., Hancock, T., Natsume-Kitatani, Y., and Mamitsuka, H. (2017). Computational recognition for long non-coding RNA (lncRNA): software and databases. *Brief. Bioinform.* 18, 9–27. doi:10.1093/bib/bbv114
- Zeng, M., Wu, Y., Lu, C., Zhang, F., Wu, F., and Li, M. (2022). DeepLncLoc: a deep learning framework for long non-coding RNA subcellular localization prediction based on subsequence embedding. *Br. Bioinform* 23, bbab360. doi:10.1093/bib/bbab360
- Zhang, P., Wu, W., Chen, Q., and Chen, M. (2019). Non-coding RNAs and their integrated networks. *J. Integr. Bioinform* 16, 20190027. doi:10.1515/jib-2019-0027
- Zhang, S., Amahong, K., Zhang, Y., Hu, X., Huang, S., Lu, M., et al. (2023). RNAenrich: a web server for non-coding RNA enrichment. *Bioinformatics* 39, btad421. doi:10.1093/bioinformatics/btad421
- Zhang, W., Yue, X., Tang, G., Wu, W., Huang, F., and Id, X. Z. (2018). SFPEL-LPI: sequence-based feature projection ensemble learning for predicting lncRNA-protein interactions. *PLoS Comput. Biol.* 14, e1006616. doi:10.1371/journal.pcbi.1006616
- Zhang, X., Wang, W., Zhu, W., Dong, J., Cheng, Y., Yin, Z., et al. (2019). Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. *Int. J. Mol. Sci.* 20, 5573. doi:10.3390/ijms20225573
- Zhang, X. Q., Wang, Z. L., Poon, M. W., and Yang, J. H. (2017). Spatial-temporal transcriptional dynamics of long non-coding RNAs in human brain. *Hum. Mol. Genet.* 26, 3202–3211. doi:10.1093/hmg/ddx203
- Zhang, Y., Bu, D., Huo, P., Wang, Z., Rong, H., Li, Y., et al. (2021). ncFANs v2.0: an integrative platform for functional annotation of non-coding RNAs. *Nucleic Acids Res.* 49, 459–468. doi:10.1093/nar/gkab435
- Zhang, Y., Y., Dong, R., Yang, L., Chen, L., Xing, Y., Yao, R., et al. (2017). SLERT regulates DDX21 rings associated with Pol I transcription. *Cell* 169, 664–678.e16. doi:10.1016/j.cell.2017.04.011
- Zhang, Z., Sun, Z., Yang, Y., and Lin, H. (2022). Towards a better prediction of subcellular location of long non-coding RNA. *Front. Comp. Sci.* 16, 165903. doi:10.1007/s11704-021-1015-3
- Zhao, H., Yin, X., Xu, H., Liu, K., Liu, W., Wang, L., et al. (2023). LncTarD 2.0: an updated comprehensive database for experimentally-supported functional lncRNA – target regulations in human diseases. *Nucleic Acids Res.* 51, 199–207. doi:10.1093/nar/gkac984
- Zhao, T., Hu, Y., Peng, J., and Cheng, L. (2020). DeepLGP: a novel deep learning method for prioritizing lncRNA target genes. *Bioinformatics* 36, 4466–4472. doi:10.1093/bioinformatics/btaa428
- Zhao, X., Li, F., Ali, M., Li, X., Fu, X., and Zhang, X. (2024). Emerging roles and mechanisms of lncRNAs in fruit and vegetables. *Hortic. Res.* 44, uhae046. doi:10.1093/hr/uhae046
- Zheng, Y., Luo, H., Teng, X., Hao, X., Yan, X., Tang, Y., et al. (2023). NPInter v5.0: ncRNA interaction database in a new era. *Nucleic Acids Res.* 51, 232–239. doi:10.1093/nar/gkac1002
- Zhou, J., Huang, Y., Ding, Y., Yuan, J., Wang, H., and Sun, H. (2018). lncFunTK: a toolkit for functional annotation of long noncoding RNAs. *Bioinformatics* 34, 3415–3416. doi:10.1093/bioinformatics/bty339
- Zibitt, M. S., Hartford, C. C. R., and Lal, A. (2021). Interrogating lncRNA functions via CRISPR/Cas systems. *RNA Biol.* 18, 2097–2106. doi:10.1080/15476286.2021.1899500
- Zuckerman, B., and Ulitsky, I. (2019). Predictive models of subcellular localization of long RNAs. *RNA* 25, 557–572. doi:10.1261/rna.068288.118