Table S1. Comparison of lncRNA bioinformatic analysis pipelines

	LncPipe [1]	LncEvo [2]	CALINCA [3]	lncRNA Detector [4]	lncRNA-screen [5]	ICAnnoLncRNA [6]
Quality control and assembly of the transcriptome	+	+	+	–	+	–
Prediction model training for a non-model organism	–	–	–	–	+	+
Possibility of using data for plants	–	–	–	+	+	+
Prediction of lncRNA using specialized programs	+	+	–	+	+	+
Classification of lncRNAs according to their location relative to protein-coding genes	+	+	+	–	+	+
Removal of transcriptional noise and assembly redundancy	–	–	–	–	–	+
Deletion of sequences that are potential mobile elements	–	–	–	–	–	+
Differential expression analysis	+	–	+	–	+	-
Analysis of lncRNA conservation between different organisms	–	+	+	–	+	+
Analysis of lncRNA distribution in tissues based on its expression	+	–	+	–	+	+

Table S2. Databases for lncRNAs in plants

Title	Description	Data source	Reference
PLncDBv2.0	Plant database with more than 1,246,372 lncRNAs predicted for 80 organisms from chlorophytes to embryophytes. Presents information on expression, tissue specificity, mutations, and developmental stages of lncRNA transcripts in 13834 datasets from different organisms. Allows the user to study the relationship between lincRNAs and epigenetic markers	Methods for lncRNA prediction	[7]
PNRD	Includes from 166 plant species, over 25,739 lncRNAs of 16 types. Offers several analytical tools and includes a proprietary genomic search engine, protein-coding sequence prediction methods and a microRNA prediction method	Literature and databases	[8]
NONCODE v6	The database contains 16 animal species and 549,813 lncRNAs belonging to them, and 23 plant species and 94,697 lncRNA transcripts	Literature and databases	[9]
GREENC v2	Includes more than 495,000 lncRNA transcripts for 94 plant and algal species	Methods for lncRNA prediction	[10]
CANTATAdb v2	The database contains more than 239,631 lncRNAs from 36 plant species and 3 algal species	Methods for lncRNA prediction	[11]
EVLncRNAs v2	Includes 4010 lncRNAs for 124 species of organisms. Provides the user with a network of lncRNA interactions with microRNAs, proteins, genes, and other functional elements. Also contains several lncRNA analysis and prediction programs	Literature	[12]

REFERENCES

1. *Zhao Q. et al.* LncPipe: A Nextflow-based pipeline for identification and analysis of long non-coding RNAs from RNA-Seq data // J Genet Genomics. 2018. Vol. 45, № 7. P. 399–401.
2. *Bryzghalov O., Makałowska I., Szcześniak M.W.* lncEvo: Automated identification and conservation study of long noncoding RNAs // BMC Bioinformatics. 2021. Vol. 22, № 1. P. 59.
3. *Talyan S. et al.* CALINCA—A Novel Pipeline for the Identification of lncRNAs in Podocyte Disease // Cells. MDPI, 2021. Vol. 10, № 3. P. 692.
4. *Shukla B. et al.* lncRNADetector: A bioinformatics pipeline for long non-coding RNA identification and MAPslnc: a repository of medicinal and aromatic plant lncRNAs // RNA Biol. 2021. Vol. 18, № 12. P. 2290–2295.
5. *Gong Y. et al.* lncRNA-screen: An interactive platform for computationally screening long non-coding RNAs in large genomics datasets // BMC Genomics. 2017. Vol. 18, № 1. P. 434.
6. *Pronozin A.Y., Afonnikov D.A.* ICAnnoLncRNA: A Snakemake Pipeline for a Long Non-Coding-RNA Search and Annotation in Transcriptomic Sequences // Genes. MDPI, 2023. Vol. 14, № 7. P. 1331.
7. *Jin J. et al.* PLncDB V2. 0: A comprehensive encyclopedia of plant long noncoding RNAs // Nucleic Acids Res. Oxford University Press, 2021. Vol. 49, № D1. P. D1489–D1495.
8. *Yi X. et al.* PNRD: A plant non-coding RNA database // Nucleic Acids Res. Oxford University Press, 2015. Vol. 43, № D1. P. D982–D989.
9. *Zhao Y. et al.* NONCODE 2016: An informative and valuable data source of long non-coding RNAs // Nucleic Acids Res. Oxford University Press, 2016. Vol. 44, № D1. P. D203–D208.
10. *Gallart A.P. et al.* GREENC: A Wiki-based database of plant lncRNAs // Nucleic Acids Res. Oxford University Press, 2016. Vol. 44, № Database issue. P. D1161.
11. *Szcześniak M.W. et al.* CANTATAdb 2.0: Expanding the Collection of Plant Long Noncoding RNAs // Plant Long Non-Coding RNAs / ed. Chekanova J.A., Wang H.-L.V. New York, NY: Springer New York, 2019. Vol. 1933. P. 415–429.
12. *Zhou B. et al.* EVLncRNAs 2.0: An updated database of manually curated functional long non-coding RNAs validated by low-throughput experiments // Nucleic Acids Res. Oxford University Press, 2021. Vol. 49, № D1. P. D86–D91.

The Role of Long Non-Coding Rnas in Plants

A. Y. Pronozin [1, 2, *], D. A. Afonnikov [1, 2, 3]

[1] *Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090 Russia*
[2] *Kurchatov Genomic Center, Institute of Cytology and Genetics, Siberian Branch,*
Russian Academy of Sciences, Novosibirsk, 630090 Russia
[3] *Novosibirsk National Research State University, Novosibirsk, 630090 Russia*
**e-mail: pronozinartem95@gmail.com*

Long non-coding RNAs (lncRNAs) are a class of linear or circular RNA molecules longer than 200 nucleotides without open reading frames. Experimental studies have shown the involvement of lncRNAs in the regulation of resistance to cold, salt, and heat stress, and in fruit, root, and leaf development. However, experimental methods are labor-intensive and costly approaches and cannot yet be used for genome-wide mass studies of lncRNAs. For this purpose, bioinformatic approaches that aim at large-scale recognition of lncRNA sequences in genomes and transcriptomes have been applied. However, despite the growing number of studies devoted to the structural and functional analysis of lncRNAs, this type of molecule remains poorly understood. This is due to the many factors that need to be considered when identifying lncRNAs. The use of pan-genomes and pan-transcriptomes will improve the efficiency of the study and the total number of predicted lncRNAs compared to using the genome of a single species representative. This review focuses on describing the molecular and biological functions of lncRNAs, experimental and bioinformatic methods of identification, patterns of evolution, detection and analysis of lncRNAs at the scale of pan-genomes and pan-transcriptomes.

Keywords: long non-coding RNAs, microRNAs, transcription regulation, pan-genome.