Genomics Regulation and Functional Aspects of Long Non-Coding RNAs

Piero Carninci*
Genomics Research Centre, Fondazione Human Technopole, Milan (Italy) Center for Integrative Medical Sciences, Riken, Yokohama, Japan.
*E-mail: piero.carninci@fht.org

ABSTRACT

Introduction: The sequencing of the human and other mammalian genomes was a landmark achievement, providing the foundation for modern biology. However, genome sequence alone is insufficient to fully understand the repertoire of genes, the RNAs they produce, and the regulatory mechanisms that govern gene expression across all cell types, tissues, and organs of the human body.

Objectives: Our long-term goal has been to systematically identify and characterize mammalian genome elements. We began with the large-scale collection of mRNAs encoding proteins in humans and mouse models, then expanded to mapping gene regulatory elements across diverse cell types, with a particular focus on enhancers as independent layers of regulation. With the discovery of long non-coding RNAs (lncRNAs), we further set out to uncover their functions and their roles in shaping the epigenome.

Methodology: Over the years, my group has developed a suite of complementary technologies. These include large-scale full-length cDNA cloning for gene discovery, followed by Cap Analysis of Gene Expression (CAGE), which enables single-base resolution mapping of promoters and enhancers and quantification of their activity. To study lncRNA function, we created methods for their knockdown and for mapping their interactomes with chromatin, other RNAs, and proteins—for example, through adaptation of PARIS for RNA–RNA interactions, and the use of Hi-C and related approaches for chromatin contacts. More recently, we have applied Oxford Nanopore sequencing for large-scale cDNA sequencing. Alongside these experimental approaches, we have developed bespoke bioinformatics tools, methods, and databases, and are now extending these technologies to single-cell resolution.

Key Findings: We generated the largest collection of mammalian cDNAs in mouse and made substantial contributions to the human cDNA reference set, including alternative splicing isoforms, through the FANTOM project. Unexpectedly, we discovered that mammalian genomes produce a vast number of lncRNAs—one of the most surprising findings of the genome era. While many lncRNAs remain uncharacterized, increasing evidence reveals their structural and regulatory roles. Notably, we identified a new class of antisense lncRNAs, termed SINEUPs, which enhance translation of their target mRNAs and hold promise as therapeutic agents.

In parallel, we systematically mapped mammalian promoters and enhancers, with a strong emphasis on human, uncovering fundamental principles of gene regulation. These maps are now being expanded to single-cell resolution. More recently, through the FANTOM6 project, we have charted RNA–chromatin interactions, revealing a largely unexplored layer of regulation involving lncRNAs, intronic sequences, and RNAs derived from retrotransposons.

Conclusion: Through the FANTOM consortium and in synergy with initiatives such as ENCODE, our work has contributed to detailed maps of gene structures, isoforms, promoters, and enhancers. These efforts have fundamentally advanced our understanding of mammalian gene regulation. The next frontier is to resolve these processes at single-cell resolution across the human body (as envisioned in the Human Cell Atlas) and to elucidate the diverse and expanding roles of lncRNAs in biology and disease.

Keywords: Transcriptome, Non-coding RNAs, Epigenome, Genome Regulation.

REFERENCES

1. Bonetti, A., *et al*. "RADICL-seq Identifies General and Cell Type-Specific Principles of Genome-Wide RNA-Chromatin Interactions." Nature Communications, vol. 11, no. 1, 2020, p. 1018. https://doi.org/10.1038/s41467-020-14337-6.
2. Hon, C. C., *et al*. "An Atlas of Human Long Non-Coding RNAs with Accurate 5′ Ends." Nature, vol. 543, no. 7644, 2017, pp. 199–204. https://doi.org/10.1038/nature21374.
3. Forrest, A. R. R., *et al*. "A Promoter-Level Mammalian Expression Atlas." Nature, vol. 507, 2014, pp. 462–70. https://doi.org/10.1038/nature13182.
4. Andersson, R., *et al*. "An Atlas of Active Enhancers across Human Cell Types and Tissues." Nature, vol. 507, 2014, pp. 455–61. https://doi.org/10.1038/nature12787.
5. Carninci, P., *et al*. "The Transcriptional Landscape of the Mammalian Genome." Science, vol. 309, 2005, pp. 1559–63. https://doi.org/10.1126/science.1112014.