DOI 10.18699/BGRSSB-2018-003

Proteomic analysis of affinity captured LINE-1 macromolecular complexes

I. Altukhov[1]*, M.S. Taylor[2], K.R. Molloy[3], P. Mita[4], H. Jiang[5], E.M. Adney[4, 6], A. Wudzinska[6], S. Badri[7], D. Ischenko[1], G. Eng[2], K.H. Burns[5, 8], D. Fenyö[4], B.T. Chait[3], D. Alexeev[9], M.P. Rout[5], J.D. Boeke[4], J. LaCava[4, 5]

[1] *Moscow Institute of Physics and Technology, Dolgoprudny, Russia*
[2] *Department of Pathology, Massachusetts General Hospital, Boston, United States*
[3] *Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, United States*
[4] *Department of Biochemistry and Molecular Pharmacology, Institute for Systems Genetics, NYU Langone Health, New York, United States*
[5] *Laboratory of Cellular and Structural Biology, The Rockefeller University, New York, United States*
[6] *McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, United States*
[7] *Department of Pathology, NYU Langone Health, New York, United States*
[8] *Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, United States*
[9] *Novosibirsk State University, Novosibirsk, Russia*
** e-mail: ilya.altukhov@gmail.com*

Key words: protein-protein interactions, protein complex, mass-spectrometry, LINE-1, retrotransposon, data analysis

Motivation and Aim: Long Interspersed Nuclear Element-1 (LINE-1, L1) is a mobile genetic element active in human genomes. L1-encoded ORF1 and ORF2 proteins bind L1 RNAs, forming ribonucleoproteins (RNPs). These RNPs interact with diverse host proteins, some repressive and others required for the L1 lifecycle. In a never-ending battle, our cells have been fighting to keep LINE-1 and its ancestors from replicating, and so evolved various defense mechanisms. Yet, LINE-1 has learned to circumvent these barriers, and continues to replicate and cause disease. Our understanding of these defenses and of how LINE-1 evades them is limited.

Methods and Algorithms: Using differential affinity purifications, quantitative mass spectrometry, and next generation RNA sequencing, we have characterized the proteins and nucleic acids associated with distinctive, enzymatically active L1 macromolecular complexes.

Results: We described a cytoplasmic intermediate that we hypothesize to be the canonical ORF1p/ORF2p/L1-RNA-containing RNP, and we described a nuclear population containing ORF2p, but lacking ORF1p, which likely contains host factors participating in target-primed reverse transcription. We additionally explored proteomes associated with catalytically-inactivated ORF2p point mutants and monitored the rates of protein exchange from L1 macromolecules in vitro. Taken together, our data support the existence of a variety of putative L1-related protein complexes.

Conclusion: Custom computer code written in the R programming language was used in the analysis of mass spectrometry and RNA sequencing data; it has been published at https://github.com/elifesciences-publications/altukhov-line-1.