doi 10.18699/bgrs2024-11.2-12

Search for upstream open reading frames that are essential for viability of human cell lines

Shepelev N.M.[1, 2]*, Kiniry S.J.[3], Lavrov A.I.[1], Razumova E.A.[1], Suspitsyna A.D.[4], Baranov P.V.[3], Dontsova O.A.[1, 2, 4], Rubtsova M.P.[1, 2]

[1] *Lomonosov Moscow State University, Moscow, Russia*

[2] *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, RAS, Moscow, Russia*

[3] *University College Cork, Cork, Ireland*

[4] *Skolkovo Institute of Science and Technology, Moscow, Russia*

* *nikita.shepelev96@gmail.com*

Key words: small ORFs; sORFs; upstream ORFs; uORFs; CRISPR-Cas9; microproteins

Motivation and Aim: The Human Genome Project concluded over 20 years ago. However, the complete repertoire of coding open reading frames (ORFs) remains uncertain. Among these, the identification of small open reading frames, those consisting of less than 100 codons, poses a significant challenge [1]. Ribosome profiling has made a revolution in our understanding of translated regions of different RNA molecules [2]. Among the plethora of small ORFs, particular attention is drawn to those located in the 5' untranslated regions of mRNAs, commonly known as upstream ORFs. Upstream ORFs can regulate translation *in cis*, encode small functional peptides (microproteins), be a source of antigenic peptides, or serve as the origin for new proteins and their domains during the evolutionary process. It is estimated that nearly half of mammalian mRNAs contain upstream ORFs. Nevertheless, their functions largely remain understudied [3]. Therefore, we performed a functional genomics screening to identify essential upstream ORFs in human cell lines.

Methods and Algorithms: We utilized ribosome profiling data to predict upstream ORFs in human mRNAs, ultimately identifying 977 highly conserved targets. We then designed highly specific guide RNAs to target these upstream ORFs employing the CRISPR-Cas9 system. Our investigation involved an extensive study of selected upstream ORFs using the powerful CRISPR-Cas9 knockout screening technique to assess their influence on cell viability of HAP1, A549, and HEK293T human cell lines. Selected targets were verified by a competition assay using individual guide RNAs. Translation of upstream ORFs was confirmed using a reporter construct or by endogenous tagging. Using a dual-luciferase reporter, we tested the effect of targets on the translation of the main downstream ORFs.

Results: CRISPR-Cas9 knockout screening allowed us to identify numerous upstream ORFs potentially essential for human cell viability. Subsequent functional analysis confirmed the translation of several upstream ORFs by endogenous tagging or exogenous tagging in the reporter construct. Notably, several upstream ORFs were found to play a regulatory role in the translation of the main downstream ORFs. Conservation analysis coupled with peptide motif search indicates that some upstream ORFs may encode functional microproteins.

Conclusion: The CRISPR-Cas9 knockout screening enabled us to identify upstream ORF that play important roles in essential cellular processes. The following analysis

indicates that these selected upstream ORFs have the capability to modulate the translation of main ORFs or potentially encode functional microproteins.
Funding: The study is supported by the Russian Science Foundation (23-14-00058).

References

1. Schlesinger D., Elsässer S.J. Revisiting sORFs: overcoming challenges to identify and characterize functional microproteins. *FEBS J.* 2022;289(1):53-74. doi 10.1111/febs.15769
2. Ingolia N.T., Brar G.A., Stern-Ginossar N. et al. Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep.* 2014;8(5):1365-1379. doi 10.1016/j.celrep.2014.07.045
3. Renz P.F., Valdivia-Francia F., Sendoel A. Some like it translated: small ORFs in the 5'UTR. *Exp Cell Res.* 2020;396(1):112229. doi 10.1016/j.yexcr.2020.112229