



Editorial: Emerging Proteins and Polypeptides Expressed by “Non-Coding RNAs”

Wanting Liu¹, Qing-Yu He^{1*} and Marie A. Brunet^{2,3*}

¹MOE Key Laboratory of Tumor Molecular Biology and Key Laboratory of Functional Protein Research of Guangdong Higher Education Institutes, Institute of Life and Health Engineering, Jinan University, Guangzhou, China, ²Department of Pediatrics, Medical Genetics Service, Université de Sherbrooke, Sherbrooke, QC, Canada, ³Centre de Recherche du Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, QC, Canada

Keywords: non-coding RNA (ncRNA), alternative ORFs, small ORFs, MicroProtein, ribosome profiling (RIBO-Seq), mass spectrometry

Editorial on the Research Topic

Emerging Proteins and Polypeptides Expressed by “Non-Coding RNAs”

By definition, non-coding RNAs (ncRNAs) are RNA molecules that do not encode proteins. Yet, emerging evidences, drawn from deep ribosome sequencing and mass spectrometry, show that a subset of ncRNAs including long non-coding RNAs (lncRNAs) and cirRNAs are able to encode functional proteins/polypeptides (Makarewich and Olson, 2017; Orr et al., 2020; Peeters and Menschaert, 2020). Although the function of these novel proteins remains sometimes elusive, some have been demonstrated to play vital functions in human health. The identification and functional characterization of these novel proteins is a new emerging field of biological sciences. Recent studies have shown that these novel proteins are involved in diverse biological functions such as mitochondrial function (Chen et al., 2018; Stein et al., 2018), lipid metabolism (Chibucos et al., 2014; Chen et al., 2018; Polycarpou-Schwarz et al., 2018; Singh et al., 2018; Zhang et al., 2019; Zhang et al., 2020), tumor energy metabolism (Chibucos et al., 2014; Kim et al., 2021), cell development (Kulczynska and Siatecka, 2016; Fazi and Fatica, 2019; Attaway et al., 2021; Kersy et al., 2021; Kim et al., 2021), and DNA repair (Sharma and Misteli, 2013; Slavoff et al., 2014; Zhou et al., 2015; Dianatpour and Ghafouri-Fard, 2017; Thapar, 2018; Attaway et al., 2021; Papaspyropoulos et al., 2021). This research topic in Frontiers in Cell and Developmental Biology focused on recent progress in this emerging field, aiming to better understand “ncRNAs,” and served as a forum to discuss gene annotation and the discovery of novel physiological and pathological molecules.

NON-CODING RNAs: AN OVERLOOKED SOURCE OF FUNCTIONAL PROTEINS

Non-coding RNAs have recently been demonstrated to contain small-open reading frames (sORFs) encoding small proteins. Only a few of these newly discovered proteins have been functionally characterized so far, but they are key players in a variety of cellular processes. In this topic, authors have reviewed or provided new evidence for the overlooked coding potential of some lncRNAs. The collection of article illustrates the diversity of functions of these novel proteins, from glioblastoma biomarkers to neuropeptides and regeneration.

OPEN ACCESS

Edited and reviewed by:

Ana Cuenda,
Spanish National Research Council
(CSIC), Spain

*Correspondence:

Qing-Yu He
tqyhe@email.jnu.edu.cn
Marie A. Brunet
marie.brunet@usherbrooke.ca

Specialty section:

This article was submitted to
Signaling,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 26 January 2022

Accepted: 31 January 2022

Published: 21 February 2022

Citation:

Liu W, He Q-Y and Brunet MA (2022)
Editorial: Emerging Proteins and
Polypeptides Expressed by “Non-
Coding RNAs”.
Front. Cell Dev. Biol. 10:862870.
doi: 10.3389/fcell.2022.862870

In an extensive review, Cardon et al. discuss lncRNAs-encoded proteins as novel biomarkers for glioblastoma (GBM). They review evidence linking these to the patient's survival and bad prognosis. The authors also highlighted the potential functions of these novel proteins in GBM biology by showing their interaction with known proteins in the signaling pathways of cellular mobility and transfer RNA regulation.

Novel proteins originating from lncRNAs have been found in many biological samples, representing a variety of tissues and cell types. To better understand the role of ncRNAs-encoded microproteins in different tissues, Pan et al. profiled the proteomes of five mouse tissues by mass spectrometry with bottom-up, top-down, and *de novo* sequencing strategies. Using the OpenProt database (Brunet et al., 2019; Brunet et al., 2021), they identified 1,074 microproteins, 540 were known and 534 were novel, including 270 from ncRNAs. They performed gene ontology analyses on the 540 already annotated microproteins to highlight tissue-specific functions. For example, the brain contains the largest number of neuropeptides, and the spleen contains the most immune-associated microproteins. Their results expand the mouse proteome and provide insights into the molecular biology of mouse tissues.

Working with mouse embryonic stem cells, Senís et al. discovered a conserved microprotein, named pTUNAR, encoded in the *TUNAR* lncRNA. The authors showed that the 48 amino-acid long pTUNAR is expressed in the nervous system using ribosome profiling and a custom antibody. They identified pTUNAR at the membrane of the endoplasmic reticulum, in interaction with SERCA2. Their results validate the previous work of Li et al. (2021) where pTUNAR was independently identified (and named BNLN) and found in interaction with SERCA3. Although further work is needed to understand how pTUNAR regulates calcium dynamics, this work confirmed previous findings and suggest pTUNAR as an important player in neural differentiation and neurite formation.

Another type of non-coding RNA are telomerase RNA. Along with the telomerase reverse transcriptase and regulatory proteins, it makes up the telomerase complex. However, telomerase RNA is expressed in most somatic cells, whereas the telomerase reverse transcriptase is absent. This observation prompted Rubtsova et al. (2018) to investigate the coding potential of human telomerase RNA and discovered the human telomerase RNA protein (hTERP). In this collection, Shliapina et al. further our understanding of hTERP role in autophagy regulation. Using hTERP knock-out and over-expression models, the authors showed that hTERP is involved in the regulation of AMPK and mTORC1 activity. Although more work is needed to fully understand the role of hTERP, it is a pinnacle example of how a deeper characterization of the human proteome is essential to truly decipher cellular and molecular pathways.

DEVELOPING THE NECESSARY TOOLS TO EXPLORE THE DEEP PROTEOME

Ribosome profiling is the major technological advance that revealed pervasive translation throughout the genome in

eukaryotes (Ingolia et al., 2011; Chen et al., 2020). Mass spectrometry quickly followed to demonstrate the existence of protein products from these non-canonical translation sites (Menschaert et al., 2013; Samandi et al., 2017). The development of new technologies and methods is necessary to foster the detection of novel proteins originating from non-coding RNAs.

In this collection, Peeters et al. proposed a proteogenomics workflow combining state-of-the-art mass spectrometer (TimsTOF) and machine learning algorithms to improve the detection of functional peptides in samples. The authors focused on the mouse brain and peptides shorter than 100 amino acids. With an enhanced sensitivity and an optimized search of a large database combining OpenProt (Brunet et al., 2019; Brunet et al., 2021) and the sORFs repository (Olexiouk et al., 2016; Olexiouk et al., 2018), this workflow eases the robust identification of non-canonical peptides.

As the field grows, computational resources have emerged. These include repositories of non-canonical open reading frames (such as OpenProt and sORFs used in studies published in this collection) and browsers of large ribosome profiling data collection, such as GWIPS-viz (Kiniry et al., 2018; Michel et al., 2014) and Trips-Viz (Kiniry et al., 2021; Kiniry et al., 2019). As such, Zaheed et al. present a detailed guide on using GWIPS-Viz and Trips-Viz to explore evidence of translation of allegedly non-coding RNAs. As an example, the authors identify the coding potential of the previously misannotated as lncRNA *LINC00116*. The latter was recently shown to encode the mitoregulin protein and reannotated as the *MTLN* mRNA (Chen et al., 2018) and thus act as a positive control in the method overview from Zaheed et al.

CONCLUDING REMARKS

The field is still young and this collection highlights recent discoveries, novel technologies and avenues for research. All of these are necessary steps to move away from serendipitous discoveries into systematic explorations of the coding potential of eukaryotic “non-coding” RNAs. This unexplored reservoir of functional proteins might hold the key to a better understanding of cellular and molecular mechanisms.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

We also thank the editorial board in the Signaling section of the *Frontiers in Cell and Developmental Biology*. MB holds a Junior 1 Fellowship from the Fonds de Recherche du Québec en Santé (FRQS).

REFERENCES

- Attaway, M., Chwat-Edelstein, T., and Vuong, B. Q. (2021). Regulatory Non-Coding RNAs Modulate Transcriptional Activation During B Cell Development. *Front. Genet.* 12, 678084. doi:10.3389/fgene.2021.678084
- Brunet, M. A., Brunelle, M., Lucier, J. F., Delcourt, V., Levesque, M., Grenier, F., et al. (2019). OpenProt: A More Comprehensive Guide to Explore Eukaryotic Coding Potential and Proteomes. *Nucleic Acids Res.* 47 (D1), D403–D410. doi:10.1093/nar/gky936
- Brunet, M. A., Lucier, J.-F., Levesque, M., Leblanc, S., Jacques, J.-F., Al-Saedi, H. R. H., et al. (2021). OpenProt 2021: Deeper Functional Annotation of the Coding Potential of Eukaryotic Genomes. *Nucleic Acids Res.* 49 (D1), D380–D388. doi:10.1093/nar/gkaa1036
- Chen, C., Cui, Q., Zhang, X., Luo, X., Liu, Y., Zuo, J., et al. (2018). Long Non-Coding RNAs Regulation in Adipogenesis and Lipid Metabolism: Emerging Insights in Obesity. *Cell Signal.* 51, 47–58. doi:10.1016/j.cellsig.2018.07.012
- Chen, J., Brunner, A.-D., Cogan, J. Z., Nuñez, J. K., Fields, A. P., Adamson, B., et al. (2020). Pervasive Functional Translation of Noncanonical Human Open Reading Frames. *Science* 367 (6482), 1140–1146. doi:10.1126/science.aay0262
- Chibucos, M. C., Mungall, C. J., Balakrishnan, R., Christie, K. R., Huntley, R. P., White, O., et al. (2014). *Standardized Description of Scientific Evidence Using the Evidence Ontology (ECO)*. Oxford: Database.
- Dianatpour, A., and Ghafouri-Fard, S. (2017). The Role of Long Non Coding RNAs in the Repair of DNA Double Strand Breaks. *Int. J. Mol. Cell Med.* 6 (1), 1–12.
- Fazi, F., and Fatica, A. (2019). Interplay Between N6-Methyladenosine (m6A) and Non-coding RNAs in Cell Development and Cancer. *Front. Cell Dev. Biol.* 7, 116. doi:10.3389/fcell.2019.00116
- Ingolia, N. T., Lareau, L. F., and Weissman, J. S. (2011). Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell* 147 (4), 789–802. doi:10.1016/j.cell.2011.10.002
- Kersy, O., Salmon-Divon, M., Shpilberg, O., and Hershkovitz-Rokah, O. (2021). Non-Coding RNAs in Normal B-Cell Development and in Mantle Cell Lymphoma: From Molecular Mechanism to Biomarker and Therapeutic Agent Potential. *Int. J. Mol. Sci.* 22 (17). doi:10.3390/ijms22179490
- Kim, S.-H., Lim, K.-H., Yang, S., and Joo, J.-Y. (2021). Long Non-Coding RNAs in Brain Tumors: Roles and Potential as Therapeutic Targets. *J. Hematol. Oncol.* 14 (1), 77. doi:10.1186/s13045-021-01088-0
- Kiniry, S. J., Judge, C. E., Michel, A. M., and Baranov, P. V. (2021). Trips-Viz: An Environment for the Analysis of Public and User-Generated Ribosome Profiling Data. *Nucleic Acids Res.* 49 (W1), W662–W670. doi:10.1093/nar/gkab323
- Kiniry, S. J., Michel, A. M., and Baranov, P. V. (2018). The GWIPS-Viz Browser. *Curr. Protoc. Bioinform.* 62 (1), e50. doi:10.1002/cpbi.50
- Kiniry, S. J., O'Connor, P. B. F., Michel, A. M., and Baranov, P. V. (2019). Trips-Viz: A Transcriptome Browser for Exploring Ribo-Seq Data. *Nucleic Acids Res.* 47 (D1), D847–D852. doi:10.1093/nar/gky842
- Kulczynska, K., and Siatecka, M. (2016). A Regulatory Function of Long Non-Coding RNAs in Red Blood Cell Development. *Acta Biochim. Pol.* 63 (4), 675–680.
- Li, M., Shao, F., Qian, Q., Yu, W., Zhang, Z., Chen, B., et al. (2021). A Putative Long Noncoding RNA-Encoded Micropeptide Maintains Cellular Homeostasis in Pancreatic β Cells. *Mol. Ther. Nucleic Acids* 26, 307–320. doi:10.1016/j.omtn.2021.06.027
- Makarewich, C. A., and Olson, E. N. (2017). Mining for Micropeptides. *Trends Cell Biol.* 27 (9), 685–696. doi:10.1016/j.tcb.2017.04.006
- Menschaert, G., Van Crielinge, W., Notelaers, T., Koch, A., Crappé, J., Gevaert, K., et al. (2013). Deep Proteome Coverage Based on Ribosome Profiling Aids Mass Spectrometry-Based Protein and Peptide Discovery and Provides Evidence of Alternative Translation Products and Near-Cognate Translation Initiation Events. *Mol. Cell Proteom.* 12 (7), 1780–1790. doi:10.1074/mcp.m113.027540
- Michel, A. M., Fox, G., M Kiran, A., O'Connor, P. B., Heaphy, S. M., Mullan, J. P., et al. (2014). GWIPS-viz: Development of a Ribo-Seq Genome Browser. *Nucleic Acids Res.* 42 (Database issue), D859–D864. doi:10.1093/nar/gkt1035
- Olexiouk, V., Crappé, J., Verbruggen, S., Verhegen, K., Martens, L., and Menschaert, G. (2016). sORFs.org: A Repository of Small ORFs Identified by Ribosome Profiling. *Nucleic Acids Res.* 44 (D1), D324–D329. doi:10.1093/nar/gkv1175
- Olexiouk, V., Van Crielinge, W., and Menschaert, G. (2018). An Update on sORFs.org: A Repository of Small ORFs Identified by Ribosome Profiling. *Nucleic Acids Res.* 46 (D1), D497–D502. doi:10.1093/nar/gkx1130
- Orr, M. W., Mao, Y., Storz, G., and Qian, S.-B. (2020). Alternative ORFs and Small ORFs: Shedding Light on the Dark Proteome. *Nucleic Acids Res.* 48 (3), 1029–1042. doi:10.1093/nar/gkz734
- Papaspapopoulos, A., Lagopati, N., Mourkioti, I., Angelopoulou, A., Kyriazis, S., Lontos, M., et al. (2021). Regulatory and Functional Involvement of Long Non-Coding RNAs in DNA Double-Strand Break Repair Mechanisms. *Cells* 10 (6). doi:10.3390/cells10061506
- Peeters, M. K. R., and Menschaert, G. (2020). The Hunt for sORFs: A Multidisciplinary Strategy. *Exp. Cell Res.* 391 (1), 111923. doi:10.1016/j.yexcr.2020.111923
- Polycarpou-Schwarz, M., Gross, M., Mestdagh, P., Schott, J., Grund, S. E., Hildenbrand, C., et al. (2018). The Cancer-Associated Microprotein CASIMO1 Controls Cell Proliferation and Interacts with Squalene Epoxidase Modulating Lipid Droplet Formation. *Oncogene* 37 (34), 4750–4768. doi:10.1038/s41388-018-0281-5
- Rubtsova, M., Naraykina, Y., Vasilkova, D., Meerson, M., Zvereva, M., Prassolov, V., et al. (2018). Protein Encoded in Human Telomerase RNA is Involved in Cell Protective Pathways. *Nucleic Acids Res.* 46 (17), 8966–8977. doi:10.1093/nar/gky705
- Samandi, S., Roy, A. V., Delcourt, V., Lucier, J. F., Gagnon, J., Beaudoin, M. C., et al. (2017). Deep Transcriptome Annotation Enables the Discovery and Functional Characterization of Cryptic Small Proteins. *Elife* 6. doi:10.7554/eLife.27860
- Sharma, V., and Misteli, T. (2013). Non-Coding RNAs in DNA Damage and Repair. *FEBS Lett.* 587 (13), 1832–1839. doi:10.1016/j.febslet.2013.05.006
- Singh, A. K., Aryal, B., Zhang, X., Fan, Y., Price, N. L., Suárez, Y., et al. (2018). Posttranscriptional Regulation of Lipid Metabolism by Non-Coding RNAs and RNA Binding Proteins. *Semin. Cell Dev. Biol.* 81, 129–140. doi:10.1016/j.semcdb.2017.11.026
- Slavoff, S. A., Heo, J., Budnik, B. A., Hanakahi, L. A., and Saghatelian, A. (2014). A Human Short Open Reading Frame (sORF)-Encoded Polypeptide that Stimulates DNA End Joining. *J. Biol. Chem.* 289 (16), 10950–10957. doi:10.1074/jbc.c113.533968
- Stein, C. S., Jadiya, P., Zhang, X., McLendon, J. M., Abouassaly, G. M., Witmer, N. H., et al. (2018). Mitoregulin: A lncRNA-Encoded Microprotein that Supports Mitochondrial Supercomplexes and Respiratory Efficiency. *Cell Rep.* 23 (13), 3710–3720. doi:10.1016/j.celrep.2018.06.002
- Thapar, R. (2018). Regulation of DNA Double-Strand Break Repair by Non-Coding RNAs. *Molecules* 23 (11). doi:10.3390/molecules23112789
- Zhang, T.-N., Wang, W., Yang, N., Huang, X.-M., and Liu, C.-F. (2020). Regulation of Glucose and Lipid Metabolism by Long Non-Coding RNAs: Facts and Research Progress. *Front. Endocrinol.* 11, 457. doi:10.3389/fendo.2020.00457
- Zhang, X., Price, N. L., and Fernández-Hernando, C. (2019). Non-Coding RNAs in Lipid Metabolism. *Vasc. Pharmacol.* 114, 93–102. doi:10.1016/j.vph.2018.06.011
- Zhou, Z., Liu, H., Wang, C., Lu, Q., Huang, Q., Zheng, C., et al. (2015). Long Non-Coding RNAs as Novel Expression Signatures Modulate DNA Damage and Repair in Cadmium Toxicology. *Sci. Rep.* 5, 15293. doi:10.1038/srep15293

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.

Copyright © 2022 Liu, He and Brunet. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.