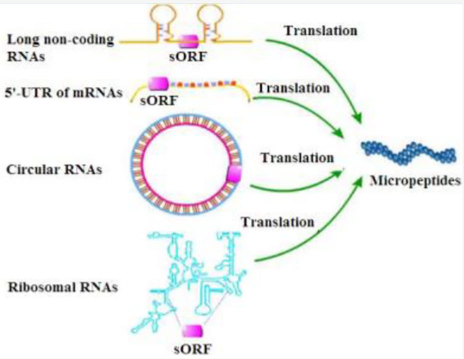
Journal of proteome research
pubs.acs.org/jpr
Reviews

Micropeptides Identified from Human Genomes

Jing Yuanyuan and Yin Xinqiang*

Cite This: J. Proteome Res. 2022, 21, 865−873

ABSTRACT: Advanced analytic techniques, such as ribosome profiling and mass spectrometry, as well as improved bioinformatics technology, have promoted the field of genome annotation forward and have identified thousands of likely coding short open reading frames (sORFs) in the human genome. The discovery of sORFs and their products allows us to realize that the complexity of the human genome is far greater than previously assumed. Here, we provide a review of human micropeptides encoded by various transcripts such as mitochondrial rRNAs, long noncoding RNAs, circular RNAs, upstream of mRNAs, and so on.



KEYWORDS: *sORFs, micropeptides, human genome*

■ INTRODUCTION

Open reading frames (ORFs), hidden in various transcripts, containing a start codon, a line of codons less than 100, and a stop codon, are categorized as short/small open reading frames (sORFs/smORFs).[1−3] Products produced by sORFs/smORFs are named micropeptides or sORFs-encoded peptides (SEPs).[1,3,4] Hundreds of translated, yet nonannotated sORFs derived from long noncoding RNAs (lncRNAs), mRNAs, circular RNAs (circRNAs), pri-miRNAs, and rRNAs have been revealed by using advanced computational methods and large-scale sequencing techniques.[2] Meanwhile, some micropeptides have been identified in a wide range of species from bacteria to humans.[4] Classical bioactive peptides, cleaved from larger precursor proteins, are targeted directly toward the secretory pathway through an N-terminal signaling sequence. However, micropeptides encoded by sORFs are immediately released in the cytoplasm after translation because of the lack of a signaling sequence.[1]

In recent years, strategies for identification of putative translated sORFs have been developed and have disclosed several micropeptides in some organisms. These strategies are based on analyzing sORF sequences by computational methods, ribosome sequencing, and mass spectrometry (MS).[5] The coding potential of sORFs is evaluated by computational methods by analyzing the composition of nucleotides or ORF qualities and conservation.[6−10] Ribosome profiling is a method that sequences the fragments protected by the ribosome and then assesses the coding potential by various computational algorithms. This method can identify the actively translated regions of genes.[11−18] MS peptidomics and proteomics are the third widely used technique which has been used to directly identify micropeptides.[19−21]

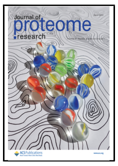
The emerging world of sORFs forms a new part of biological genomes, and the products of these sORFs could have various important functions in cells (Figure 1). Although reviews focused on micropeptides have been published during last several years,[1,3−5,19,22−26] this field is rapidly growing, and regular updates are justified by the pace of advances in this important area. Here, we provide an overview of the micropeptides identified from *Homo sapiens* in recent years (Table 1).

■ MICROPEPTIDES ENCODED BY sORFs DERIVED FROM MITOCHONDRIAL DNA

Human mitochondrial DNA (mtDNA) encodes 37 classically known genes, including 2 rRNAs, 22 tRNAs, and 13 polypeptides which are subunits of the ETC complexes. Recent studies have revealed that the mtDNA contains sORFs that can encode functional micropeptides, named mitochondria-derived peptides (MDPs).[27]

The first MDP, designated as humanin, encoded by a 75bp sORF within the mtDNA, is a 21−24 amino acid peptide that was first identified from an Alzheimer patient.[28,29] There are a

Received: November 20, 2021
Published: March 7, 2022



https://doi.org/10.1021/acs.jproteome.1c00889

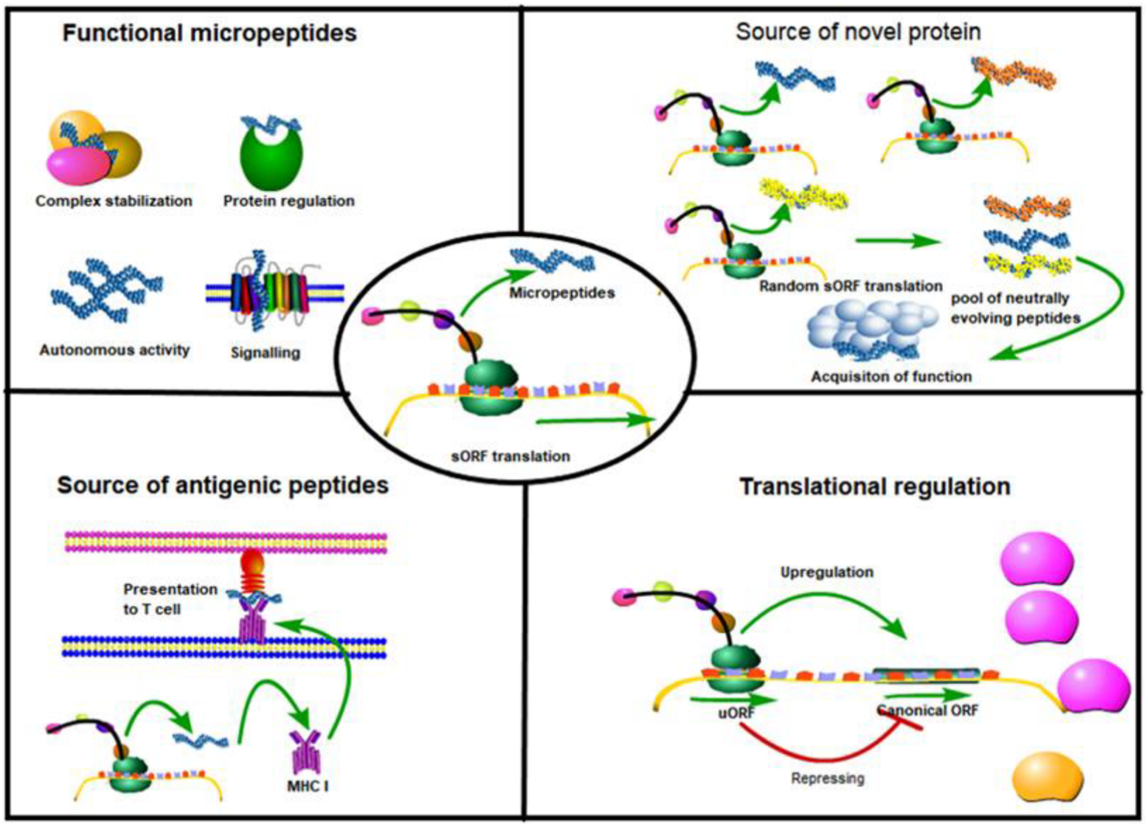


Figure 1. Different functions performed by micropeptides.

positive-charged N-terminal, a central hydrophobic region, and a negatively charged C-terminal in the sequence of humanin.[29,30] The roles of the amino acid residues of the peptide have been identified through the alanine scanning technique. Leu9 to Leu11, and Pro19 to Val20 are responsible for extracellular secretion, and amino acids Pro3−Pro19 are the functional core domain; Pro3, Ser7, Cys8, Leu9, Leu12, Thr13, Ser14, and Pro19 are also key sites for humanin's function.[31] Another research proved that Phe6 and Lys21 are responsible for binding to its interacting partner.[32] In many tissues, including testis, colon, vascular walls, neurons, skeletal muscles, and blood, humanin is widely expressed.[33−35] The wide localization of humanin indicates its important biological function in organisms. To date, many research results have demonstrated that, in different cells, tissues, and animal models, humanin showed versatile properties such as antiaging, neuroprotection, antiapoptosis, anti-inflammatory, and antifibrilogenic.[36−39]

In addition to humanin, other MDPs have also been identified. MOTS-c is a 16 aa peptide located in the 12S rRNA gene. It regulates diabetes, obesity, longevity, and exercise and represents a novel signaling mechanism regulating metabolism.[40] Subsequently, multiple studies demonstrate that MOTS-c facilitates metabolic homeostasis and reduces obesity and insulin resistance,[41] suppresses ovariectomy-induced bone loss,[42] increases survival and decreases bacterial load in MRSA-infected mice,[43] modulates mitochondrial function during cellular senescence,[44] and regulates nuclear gene expression under the condition of metabolic stress via AMPK activation.[45] In addition, we have found that, in the mouse formalin test, MOTS-c exerts antinociceptive and anti-inflammatory effects via activating AMPK pathway[46] and protects the lungs from LPS-induced acute lung injury in mice, probably by activating AMPK and SIRT1 signaling pathways and inhibiting ERK, JNK, p65, and STAT3 signals.[47] Another six small humanin-like peptides (SHLP 1−6) within the same 16S rRNA gene where humanin is located have been identified. They are involved in the activities of mitochondria.[48]

Moreover, a sORF in the C12ORF73 gene encodes a mitochondrial-localized micropeptide termed BRAWNIN. BRAWNIN, identified by analyzing ribosome profiling data combined with a prediction and validation pipeline, is highly expressed in cardiac and skeletal muscle cells and plays key roles in the assembly of respiratory chain complex III (CIII). The AMPK pathway can induce the production of BRAWNIN, and the production of ATP is impaired when this micropeptide is depleted in human cells. These findings suggest that BRAWNIN is necessary for human oxidative phosphorylation.[49]

Most recently, Lee reported a mitochondrial 83-aa micropeptide termed MOCCI.[50] MOCCI is an analogue of the cytochrome C oxidase (Complex IV) NDUFA4 subunit. During inflammation and infection, the expression of the micropeptide is upregulated to promote a host-protective resolution.[50] MOCCI displaces NDUFA4 in Complex IV to decrease the mitochondrial membrane potential and reduce ROS production, resulting in a diminished cytoprotective and immune response during inflammation.[50]

■ MICROPEPTIDES ENCODED BY lncRNA-DERIVED sORFs

Although long noncoding RNAs (lncRNAs) previously have been annotated as noncoding RNAs, several studies have demonstrated that a small part of putative sORFs within lncRNAs are translated.[51]

Table 1. Micropeptides Identified from *Homo sapiens*

micropeptides	length	sequences
humanin	24 aa	MAPRGFSCLLLLTSEMDLPVK
MOTS-c	16 aa	MRWQEMGYIFYPRKLR
BRAWNIN	71 aa	MPAGVPMSTYLKMFAASLLAMCAGAEVVHRYYRPDLTIPEIPPKRGELKTELLGLKERKHKPQVSQQEELK
MOCCI	83 aa	MSFFQLLMKRKELIPLVVFMTVAAGGASSFAVYSLWKTDVILDRKKNPEPWETVDPTVPQKLITINQQWKPIEELQNVQRVTK
SPAR	99 aa	MGAKAPRGPKVAQWAMETAVIGVVVVLFVVTVAITCVLCCFSCDSRAQDPQGGPGRSFTVATFRQEASLFTGPVRHAQPVPSAQDFWTFM
STORM	50 aa	MEGVFPSVKLTTWLPSCEDLTLPGGRGRLTPSIIAVTLPSLVVSLLGSVC
MIAC	51 aa	MERAGVPGFSPRRSSVEAKMQSTSCSVRKSSTVTAWPAVVLLLSWGQRRGG
CIP2A-BP	52 aa	MGRWRVSSVESWPFASTGGKLAIATAGTGTQAAPRPFLIRPPDSWALALALM
YY1BM	21 aa	MLSGQLQPEGRSALPQPGAAL
HOXB-AS3	53 aa	MPVLPGTQRYPHQRRRFQAAGGGAESGKRGSEEAPGVAWSGSESGRDAATPAW
CASIMO1	83 aa	MAVSTEELEATVQEVLGRLKSHQFFQSTWDTVAFIVFLTFMGTVLLLLLLVVARCCCCSSPGPRRESPRKERPKGVDNLALEP
NoBody	68 aa	MGDQPCCASGRSTLPPGNAREAKPPKKRCLLAPRWDYPEGTPNGGSTTLPSAPPPASAGLKSHPPPPEK
KRASIM	99 aa	MVLRRLLAALLHSPQLVERLSESRPIRRAAQLTAFALLQAQLRGQDAARRLQDLAAGPVGSLCRRAERFRDAFTQELRRGLRGRSGPPPGSQRGPGANI
BNLN	48 aa	MVITSENDEDRGGQEKESKEESVLAMLGIIGTILNLIVIIFVYIYTTL
CRNDEP	84 aa	MLAEIHPKAGLQSLQFIMELLYWLLEGGDSEDKEDATGNVEMKNIQPLVFEISCDVFQSRCKEHGKIKVLEWFKYVLGIPVYRL
SMIM30	58 aa	MTSVSTQLSLVLMSLLLVLPVVEAVEAGDAIALLLGVVLSITGICACLGVYARKRNGQ
RBRP	71 aa	MIQQEEIRKLEEEKKQLEGEIIDFYKMKAASEALQTQLSTDTKKDKHPDPYEFLLLRKIKHPGFNEELSPC
ASRPS	60 aa	MTTKMRRLRPSAPSGLGQEQEAEVVEGCFPATETPFAPAYIKKRGGRIWSSDPRSDGEH
PINT87aa	87 aa	MLWLPDRGSCSARSPSGMLRGAPGGWRYGRRCGRRRQSCCCCCCCSHVGAPLSFHREASLVSHDGHDIMKQHCGEESIRGAHGYKNK
CORO1C-47aa	47 aa	MNPRTSTTSTHSAARSLREGWVTCPRGDLMLTNVRLPEKYAGTNCSS
CircPPP1R12A-73aa	73 aa	MKPVLAFKHGNFDSLPQLMEISSIKKIGWVQWLIPVVSQHFGRLRHVNCLSPGVQDQPGQHGETPSLPKVQN
uCHOP	34 aa	MLKMSGWQRQSQNQSWNLRRECSRRKCIFIHHHT
MIEF1-MP	70 aa	MAPWSREAVLSLYRALLRQGRQLRYTDRDFYFASIRREFRKNQKLEDAEARERQLEKGLVFLNGKLGRII
ASDURF	97 aa	MPSRGTRPEDSSVLIPTDNSTPHKEDLSSKIKEQKIVVDELSNLKKNRKVYRQQQNSNIFFLADRTTEMLSESKNILDELKKEYQEIENLDKTKIKK
AGD3	63 aa	MGCGNSTATSAGAGQGPAGAAKDVTEESVTEDDKRRNYGGVYVGLPSEAVNMVSSQTKTVRKN
ELA	32 aa	MFRRLTFAQLLFATVLGIAGGVYIFQPVFEQYAKDQKELKEKMQLVQESEEKKS
MRI-2	69 aa	METLQSETKTRVLPSWLTAQVATKNVAPMKAPKRMRMAAVPVAAARLPATRTVYC
PIGBOS	54 aa	MRFQQFLFAFFIFIMSLLLISGQRPVNLTMRRKLRKHNCLQRRCMPLHSRVPFP

SPAR, identified by proteomics analysis, is a 90 aa peptide encoded by the lncRNAs LINC00961. SPAR inhibits amino acid-induced mTORC1 activation by interacting with the lysosomal v-ATPase, and then it regulates the skeletal muscle regenerative response following injury.[52] A sORF within LINC01420/LOC550643 (updated gene symbol NBDY) RNA encodes a micropeptide termed NoBody (NBDY). mRNA decapping proteins are responsible for promoting 5′-3′ decay via removing the 5′ cap from mRNAs and participating in the turnover of mRNA and nonsense-mediated decay. Endogenous NBDY regulates cellular RNA decapping by directly interacting with the decapping proteins via EDC4 and DCP1A proteins and localizing to P-bodies.[53,54] In addition, in vitro, in the presence of RNA, NBDY undergoes liquid−liquid phase separation.[55] NBDY phosphorylation promotes liquid phase remixing in vitro and macroscopic P-body segregation in cells that undergo growth factor signaling and cell division.[55] Moreover, STORM, a linc00689-derived micropeptide, induced by TNF-α-induced and MST1-mediated eIF4E phosphorylation of eIF4E, exhibits molecular mimicry of SRP19 and thus competes for 7SL RNA.[56]

A group of micropeptides, translated from lncRNAs and serving as oncogenes or tumor suppressors, have been characterized. Several excellent articles have reviewed the micropeptides identified in various cancers.[57−59] Here, we generally list the lncRNA-encoded micropeptides involved in human cancer.

Our team has found that micropeptide MIAC, identified by using computational approaches coupled with translation in vitro and mass spectrometry, interacts with AQP2 and decreases the expression of ITGB4 and SEPT2, which are involved in actin cytoskeleton and modulate cell mobility, and thereby inhibits tumor growth and metastasis of HNSCC.[60] The association of the dysregulation of MIAC with the survival of HNSCC patients and development of other tumors indicates that MIAC may have a wide anticancer spectrum.[60]

BNLN, a highly conserved micropeptide encoded by a lincRNA TUNAR (also known as TUNA, HI-LNC78, or LINC00617) enriched in β and neural cells, was experimentally verified by applying two computational tools in the human pancreas.[61] BNLN regulates Ca^{2+} homeostasis in pancreatic β cells, consequently modulating GSIS.[61] These results suggest that BNLN dysregulation might be associated with the obesity-related impairment of insulin secretion in the pancreas.

SMIM30 is a micropeptide encoded by a sORF within linc00998.[62] In vivo and in vitro, SMIM30, but not the lincRNA, promotes HCC cell proliferation, cell cycle, invasion, and migration. The level of micropeptides is correlated with a poor survival rate in HCC patients.[62] Furthermore, increased SMIM30 transcription by c-Myc activates the MAPK signaling pathway through promoting the expression of the nonreceptor tyrosine kinase SRC/YES1.[62]

RBRP is a 71 aa micropeptide encoded by lncRNA linc00266-1.[63] The micropeptide, but not the lncRNA, exerts its oncogenic functions via binding to IGF2BP1 to enhance m6A recognition of RNAs such as c-Myc mRNA, thus increases mRNA stability and c-Myc expression. In vitro, RBRP promotes proliferation, colony formation, migration, and invasion of cancer cells; in vivo, it drives tumorigenesis and metastasis.[63]

A LINC00665-encoded micropeptide named CIP2A-BP was identified by Guo et al. using both bioinformatics and experimental tools.[64] CIP2A-BP directly binds to CIP2A, a tumor oncogene, and replaces the B56γ subunit of PP2A, thereby releasing PP2A activity and inhibiting the signal pathway of PI3K/AKT, leading to the reduced expression levels of MMP-2, MMP-9, and Snail.[64] They proposed that CIP2A-BP may be a TNBC prognostic biomarker and a therapeutic target.[64] Another micropeptide identified by the same group is YY1BM.[65] YY1BM, encoded by LINC00278, is involved in the progression of esophageal squamous cell carcinoma (ESCC), inhibits the interaction of YY1 and androgen receptor, and then reduces the expression of Eef2k via the AR pathway.[65]

LncRNA linc00908, an ERα-regulated lncRNA, encodes a regulatory micropeptide named ASRPS.[66] This micropeptide, differentially expressed in triple-negative breast cancer (TNBC), was identified by using bioinformatics tools. By directly binding to STAT3 via the coiled coil domain in BC cell lines and mouse models, ASRPS inhibits the phosphorylation of STAT3, reduces the expression of VEGF, and then suppresses tumor angiogenesis.[66] Thus, ASRPS may be both a prognostic biomarker and a novel therapeutic target for TNBC.

KRASIM, a conserved micropeptide encoded by lncRNA NCBP2-AS2, is differentially expressed in normal hepatocytes and HCC cells.[67] It inhibits the growth and proliferation of the HCC cell by interacting and colocalizing with the KRAS protein in the human HuH-7 hepatoma cell cytoplasm.[67] In addition, in HCC cells, overexpression of KRASIM reduces the level of KRAS, resulting in the suppression of ERK signaling activity.[67]

A potential sORF within the noncoding RNA NR_029453 encodes a transmembrane micropeptide, localized in endosomes, termed CASIMO1. CASIMO1 regulates the cytoskeleton and affects cell migration at the transcriptional level.[68] This widely expressed micropeptide impacts proliferation and cell cycle progression through interacting with SQLE and regulates lipid droplet accumulation in breast cancer cells.[68]

LncRNA HOXB-AS3 encodes a conserved 53-aa micropeptide named HOXB-AS3.[69] HOXB-AS3 antagonizes the hnRNP A1-mediated pyruvate kinase M (PKM) splicing by binding to the arginine residues in hnRNP A1 RGG motif to ensure the formation of lower PKM2 and the suppression of glucose metabolism reprogramming, thereby inhibiting colon cancer growth.[69]

CRNDE, overexpressed in human malignancies, previously has been annotated as long noncoding RNA.[70] One of the *CRNDE* transcripts encoded micropeptide, CRNDEP, has been identified by using bioinformatics methods. CRNDEP is mainly located within the nucleus and is increased in the tissues of rapidly proliferating (e. g., intestine and spermatocytes).[70] The translation of CRNDEP is inhibited by the downregulation of *CRNDE* lncRNAs. Overexpression of the fusion protein of the micropeptide with fluorescent tag indicates that the micropeptide appears to induce the formation and localization of the stress granules. The authors speculate that CRNEDP may regulate cell proliferation and oxygen metabolism based on their preliminary experimental results and *in silico* results.[70] But CRNDEP's exact role needs to be further elucidated.

MICROPEPTIDES ENCODED BY sORFs DERIVED FROM circRNAs

Owing to the advanced deep sequencing technique and computational approaches, many circRNAs have been identified, and their roles in regulation, neural development, carcinogenesis, and as “microRNA sponges” have been illustrated.[71] Interestingly, certain synthetic circRNAs can be translated into peptides or proteins.[72,73] In addition, the human circRNAs’ coding potential has been demonstrated by computational methods.[74] In the past few years, several endogenous human circRNAs, circ-LINC-PINT,[75] circ-0000437,[76] and circPPP1R12A,[77] have been translated in vivo, and the roles of their products have been identified and characterized.

A circRNA-encoded micropeptide, named PINT87aa, involved in glioblastoma, was identified by the Zhang group. PINT87aa is translated from circular LINC-PINT and inhibits the proliferation of glioblastoma cells in vitro and in vivo by interacting with the PAF1c complex and suppressing the transcriptional elongation of multiple oncogenes.[75]

CircRNA has-circ-0000437 is remarkably decreased in endometrial cancer compared to matched paracancerous tissue.[76] There is a sORF in the circRNA encoding a functional micropeptide termed CORO1C-47aa. The micropeptide presumably functions via negatively regulating tumor angiogenesis by inhibiting the link between ARNT and TACC3, then suppresses VEGFA expression and secretion, and ultimately inhibits angiogenesis.[76]

CircPPP1R12A-73aa is a 73 aa micropeptide encoded by Has_circ_0000423 (circPPP1R12A). As a transcriptional regulator of the Hippo signaling pathway, Yes-associated protein 1 (YAP1) plays a crucial role in the proliferation and metastasis of tumor cells. The significant inhibiting effect of the YAP1 specific inhibitor Peptide 17 on CircPPP1R12A-73aa-induced tumor growth and metastasis indicates that the promotive effect of circPPP1R12A-73 on cancer cells in vivo and in vitro is through the activation of the Hippo-YAP pathway.[77]

MICROPEPTIDES ENCODED BY uORFs DERIVED FROM THE UPSTREAM OF mRNAs

Over one million uORFs have been identified by using bioinformatic analyses in humans.[78] Recent studies have revealed several micropeptide encoded by uORF. Some of these uORF-encoded micropeptides regulate the translation or transcription of their mRNAs in response to a specific product of metabolic processes via stalling ribosomes.[79]

Chen et al. discovered multiple uORFs encoding functional micropeptides and revealed that these micropetides encoded by uORFs bind to the protein encoded by the downstream of the same mRNA by using a combined strategy.[80] Although they have identified hundreds of previously uncharacterized functional micropeptides, including uORF-encoded micropeptides, in the human genome, the specific roles and mechanisms of action of the most micropeptides remain unclear.

Starck et al. traced translation using T cells to directly detect the translation products of uORFs during the integrated stress response.[79] They verified the translation of uORFs in the upstream of the binding immunoglobulin protein mRNA.[79] The micropeptides encoded by uORFs serve as MHC-I ligands to label cells for adaptive immune system recognition during the integrated stress response.[79]

A conserved uORF within the 5′UTR of CHOP encodes a 31 aa micropeptide. Mutating the 5′ leader region and the sequence of the micropeptide indicate that the peptide suppresses the downstream ORF expression by blocking ribosomal access to downstream initiation sites.[81] MIEF1-MP, a 70 aa micropeptide encoded by the uORF within the 5′ untranslated region of the gene of MIEF1 (mitochondrial dynamics protein MID51), has been identified in several cell lines and intestinal tissues by proteomic analysis.[82] MIEF1-MP facilitates mitochondrial fission through the LYR domain and is involved in regulating mitochondrial translation.[82] ASDURF, a recently discovered uORF derived from the 5′ UTR of ASNSD1 mRNA, encodes a 97 aa micropeptide. ASDURF is the twelfth subunit of the PAQosome, an 11-subunit chaperone involved in the biogenesis of several human protein complexes.[83] The discovery of this uORF-encoded micropeptide provides an example for the micropeptides encoded by the uORFs of eukaryotes. They can function in a more broad range, including as a cis-acting translational regulator of the downstream coding sequence. Therefore, it is important to include the uORF products in proteomic studies.[83]

MICROPEPTIDES ENCODED BY sORFs DERIVED FROM OTHER TRANSCRIPTS

In addition to the micropeptides encoded by mitochondrial rRNAs, lncRNAs, circRNAs, and the upstream of mRNAs, other transcripts (such as transcripts of unknown function (TUFs), transcripts previously annotated as noncoding transcripts) encoding micropeptides have also been identified in human genomes.

Six novel TUFs have been identified by screening for TUFs with controlled expression during the differentiation of pluripotent hMSCs.[84] One of these transcripts, overexpressed in cancer tissues, is AGD3, previously annotated as noncoding RNA.[84] What is interesting is that a 63-amino-acid micropeptide is encoded by AGD3. AGD3 is suppressed by the PKA pathway during adipogenesis.[84] But the mechanism of action of AGD3 in hMSC biology needs to be revealed.

ELABELA (ELA) is a secreted peptide hormone encoded by the AK092578 gene which has been previously annotated as a noncoding RNA.[85] It is essential in vivo for heart development signals via the apelin receptor. ELA is the first hormonal peptide that is involved in the ability of naïve blastomeres to differentiate into one of the three embryonic germ layers.[85]

The C7orf49 (CYREN) gene encodes a 69 aa micropeptide, MRI-2, which was identified by using peptidomics profiling of K562 cells.[21] As one of multiple isoforms produced by the C7orf49 (CYREN) gene, MRI-2 interacts with two subunits of Ku70 and Ku80 proteins of the heterodimeric protein Ku, which is a key effector of the NHEJ pathway for DSBs repair.[86] When a DNA double strand is broken, the Ku heterodimer binds to the break, and this allows the recruitment of the DNA-PK protein and other factors to repair the double strand break.[87] MRI-2 is recruited to the nucleus by theoverexpression of Ku and induction of DSBs and then enhances the rate of NHEJ in vitro.[86] As one member of the NHEJ cofactors, MIR is required for specific DNA end processing and DNA complex stabilizing.[88] Another study reports that MRI is intrinsically disordered, interacting with many DDR proteins and the cNHEJ factors and then forming large multimeric complexes dependent on its N- and C-termini and localized to

DNA DSBs to promote DDR factor retention.[89] They consider that MRI is an adaptor that promotes cNHEJ by increasing the affinity of DDR factors for DSB associated-chromatin through multivalent interactions.[89] It would be challenging to deconvolute the roles of specific short versus longer isoforms of the MRI protein.

PIGBOS is a 54-amino acid micropeptide which is translated from the opposite strand of the PIGB gene and localizes to the outer membrane of mitochondria.[90] PIGBOS regulates the unfolded protein response and endoplasmic reticulum (ER) stress-induced apoptosis in the ER by interacting with the ER protein CLCC1.[90]

■ CONCLUSION AND FUTURE PERSPECTIVES

We have reviewed micropeptides identified recently in the human genome in this review. These micropeptides have been shown to play important roles in many important biological processes. MDPs are important in regulating inflammation, cellular metabolism, apoptosis, and ion homeostasis. The deficiency of MDPs leads to development defects and even cardiovascular or neurodegenerative diseases.[27] The role of micropeptides, especially uORF-encoded peptides, in human pathology has been demonstrated. Recent studies have shown that uORF-encoded peptides play important roles in tumorigenesis. The changes induced by oncogenin alter the landscape of uORF translation, thus allowing a set of mRNAs associated with cancer to remain efficiently translated and evade the overall inhibition of protein synthesis.[78] In addition, the identification of lncRNA-encoded peptides, which play crucial roles in tumorigenesis, inflammation, and metabolism, expand the scope and mode of action of lncRNA.[24] Moreover, the important roles of circRNA-encoded micropeptides have begun to be disclosed. CircRNAs have great application potential in the field of disease diagnosis and treatment. More circRNA-encoded micropeptides and novel functions of these peptides will be identified and characterized in the future.[91]

The sORFs that may play crucial roles are dispersedly distributed in the human genome, a finding that challenges our notion that one gene only encodes one protein and that transcripts without a canonical ORF are noncoding. Likewise, it may be functionally important to use alternative initiation codons, which adds complexity to understanding and annotating genes.[4] Distinguishing these small characteristics from the statistical noise that has so far impeded their annotation is another challenge. Fortunately, advanced technologies and new approaches, such as ribosome profiling, improved bioinformatics strategies, and multiomics identification approaches, are helping us gradually solve these problems.

Exploring the function mode of micropeptides is another challenge. Combining XL-MS with bioinformatics docking predictions provides a promising method for identifying micropeptide−protein interactions at the proteomic scale that can aid in deciphering micropeptide functions. CRISPR-Cas9 combined with single-cell transcriptomics is a conspicuous functional screen strategy.[92] Moreover, loss-of-function in vitro or in vivo is an effective way to decipher the biological roles of micropeptides. This strategy is not suitable when there are no apparent phenotype changes for some micropeptides.

To mine for new micropeptides or proteins from genetic variations associated with disease is another direction for this field. The peptides or proteins translated from genetic mutations could give us important clues for understanding the etiology of human diseases, and some of these products may be potential drug candidates for therapy.[4] Micropeptides are particularly suitable as targeted drugs or targets based on their relatively smaller sizes, tissue-specific expression pattern, and low cytotoxicity. In addition, because advanced peptidomic analysis enables the discovery of more biological properties of peptides, mainly including their biological roles and druggability, this approach will certainly facilitate and enhance the use of peptides in the pharmaceutical industry for therapeutic, prognostic, and diagnostic elements.[93,94]

Interestingly, the functions of all trans-acting micropeptides known to date depend on specific interactions with larger proteins.[92] By inducing conformational changes, masking functional and regulatory sites, masking nucleic acid-binding sites and/or cofactor-binding sites, or serving as adaptors that enhance specific interactions, micropeptides modulate their partners' activity.[5] Furthermore, they can also mimic the binding domains of their partners, which are associated with multi-subunits, and modulate their activity via significant negative mechanisms.[5] In the future, the developments of integrated pipelines including information based on the structure of the proteins interacting with each other should help to predict the molecular targets of micropeptides (Figure 1).

To date, only a small part of the putative micropeptides' biological roles have been illustrated. There is still much work to do to prove their existence and clarify their functions.

■ AUTHOR INFORMATION

Corresponding Author

Yin Xinqiang − *School of Basic Medicine and Forensics, North Sichuan Medical College, Nanchong 637000, China;* orcid.org/0000-0002-4163-3600;
Phone: +8618784270139; Email: yinxq06@163.com

Author

Jing Yuanyuan − *School of Public Health, North Sichuan Medical College, Nanchong 637000, China*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jproteome.1c00889

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is supported by the North Sichuan Medical College PhD Research Fund (CBY20-QD01), the Sichuan Province Science and Technology Program (22MZGC0026), and the Sichuan Province Training Program of Innovation for Undergraduates (S20211063006, S20211063069, S20211063158X).

■ REFERENCES

(1) Crappé, J.; Van Criekinge, W.; Menschaert, G. Little things make big things happen: A summary of micropeptide encoding genes. *EuPA Open Proteomics* **2014**, 3, 128−137.

(2) Yin, X.; Hu, J.; Xu, H. Distribution of micropeptide-coding sORFs in transcripts. *Chin. Chem. Lett.* **2018**, *29* (7), 1029−1032.

(3) Sousa, M. E.; Farkas, M. H. Micropeptide. *PLoS Genetics* **2018**, *14* (12), No. e1007764.

(4) Yin, X.; Jing, Y.; Xu, H. Mining for missed sORF-encoded peptides. *Expert review of proteomics* **2019**, *16* (3), 257−266.

(5) Andrews, S. J.; Rothnagel, J. A. Emerging evidence for functional peptides encoded by short open reading frames. *Nature reviews. Genetics* **2014**, *15* (3), 193−204.

(6) Housman, G.; Ulitsky, I. Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs. *Biochimica et biophysica acta* **2016**, *1859* (1), 31−40.

(7) Kong, L.; Zhang, Y.; Ye, Z. Q.; Liu, X. Q.; Zhao, S. Q.; Wei, L.; Gao, G. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Research* **2007**, *35*, W345−W349.

(8) Qin, Q.; Delrio, S.; Wan, J.; Jay Widmer, R.; Cohen, P.; Lerman, L. O.; Lerman, A. Downregulation of circulating MOTS-c levels in patients with coronary endothelial dysfunction. *International Journal of Cardiology* **2018**, *254*, 23−27.

(9) Lin, M. F.; Jungreis, I.; Kellis, M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* **2011**, *27* (13), i275−82.

(10) Miller, W.; Rosenbloom, K.; Hardison, R. C.; Hou, M.; Taylor, J.; Raney, B.; Burhans, R.; King, D. C.; Baertsch, R.; Blankenberg, D.; Kosakovsky Pond, S. L.; Nekrutenko, A.; Giardine, B.; Harris, R. S.; Tyekucheva, S.; Diekhans, M.; Pringle, T. H.; Murphy, W. J.; Lesk, A.; Weinstock, G. M.; Lindblad-Toh, K.; Gibbs, R. A.; Lander, E. S.; Siepel, A.; Haussler, D.; Kent, W. J. 28-way vertebrate alignment and conservation track in the UCSC Genome Browser. *Genome research* **2007**, *17* (12), 1797−808.

(11) Bazzini, A. A.; Johnstone, T. G.; Christiano, R.; Mackowiak, S. D.; Obermayer, B.; Fleming, E. S.; Vejnar, C. E.; Lee, M. T.; Rajewsky, N.; Walther, T. C.; Giraldez, A. J. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *EMBO journal* **2014**, *33* (9), 981−93.

(12) Chew, G. L.; Pauli, A.; Rinn, J. L.; Regev, A.; Schier, A. F.; Valen, E. Ribosome profiling reveals resemblance between long non-coding RNAs and 5′ leaders of coding RNAs. *Development* **2013**, *140* (13), 2828−34.

(13) Erhard, F.; Halenius, A.; Zimmermann, C.; L'Hernault, A.; Kowalewski, D. J.; Weekes, M. P.; Stevanovic, S.; Zimmer, R.; Dolken, L. Improved Ribo-seq enables identification of cryptic translation events. *Nat. Methods* **2018**, *15* (5), 363−366.

(14) Hanada, K.; Akiyama, K.; Sakurai, T.; Toyoda, T.; Shinozaki, K.; Shiu, S. H. sORF finder: a program package to identify small open reading frames with high coding potential. *Bioinformatics* **2010**, *26* (3), 399−400.

(15) Michel, A. M.; Ahern, A. M.; Donohue, C. A.; Baranov, P. V. GWIPS-viz as a tool for exploring ribosome profiling evidence supporting the synthesis of alternative proteoforms. *Proteomics* **2015**, *15* (14), 2410−6.

(16) Michel, A. M.; Fox, G.; Kiran, A. M.; De Bo, C.; O'Connor, P. B.; Heaphy, S. M.; Mullan, J. P.; Donohue, C. A.; Higgins, D. G.; Baranov, P. V. GWIPS-viz: development of a ribo-seq genome browser. *Nucleic acids research* **2014**, *42*, D859−D864.

(17) Mumtaz, M. A.; Couso, J. P. Ribosomal profiling adds new coding sequences to the proteome. *Biochemical Society transactions* **2015**, *43* (6), 1271−6.

(18) Olexiouk, V.; Crappe, J.; Verbruggen, S.; Verhegen, K.; Martens, L.; Menschaert, G. sORFs.org: a repository of small ORFs identified by ribosome profiling. *Nucleic acids research* **2016**, *44* (D1), D324−9.

(19) Chu, Q.; Ma, J.; Saghatelian, A. Identification and characterization of sORF-encoded polypeptides. *Crit. Rev. Biochem. Mol. Biol.* **2015**, *50* (2), 134−41.

(20) Makarewich, C. A.; Olson, E. N. Mining for Micropeptides. *Trends in cell biology* **2017**, *27* (9), 685−696.

(21) Slavoff, S. A.; Mitchell, A. J.; Schwaid, A. G.; Cabili, M. N.; Ma, J.; Levin, J. Z.; Karger, A. D.; Budnik, B. A.; Rinn, J. L.; Saghatelian, A. Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nat. Chem. Biol.* **2013**, *9* (1), 59−64.

(22) Tharakan, R.; Sawa, A. Minireview: Novel Micropeptide Discovery by Proteomics and Deep Sequencing Methods. *Frontiers in genetics* **2021**, *12*, 651485.

(23) Vitorino, R.; Guedes, S.; Amado, F.; Santos, M.; Akimitsu, N. The role of micropeptides in biology. *Cell. Mol. Life Sci.* **2021**, *78* (7), 3285−3298.

(24) Chen, Y.; Ho, L.; Tergaonkar, V. sORF-Encoded Micro-Peptides: New players in inflammation, metabolism, and precision medicine. *Cancer letters* **2021**, *500*, 263−270.

(25) Plaza, S.; Menschaert, G.; Payre, F. In Search of Lost Small Peptides. *Annual Review of Cell and Developmental Biology* **2017**, *33*, 391−416.

(26) Pueyo, J. I.; Magny, E. G.; Couso, J. P. New Peptides Under the s(ORF)ace of the Genome. *Trends in biochemical sciences* **2016**, *41* (8), 665−678.

(27) Kim, S. J.; Xiao, J.; Wan, J.; Cohen, P.; Yen, K. Mitochondrially derived peptides as novel regulators of metabolism. *Journal of physiology* **2017**, *595* (21), 6613−6621.

(28) Hashimoto, Y.; Ito, Y.; Niikura, T.; Shao, Z.; Hata, M.; Oyama, F.; Nishimoto, I. Mechanisms of neuroprotection by a novel rescue factor humanin from Swedish mutant amyloid precursor protein. *Biochemical and biophysical research communications* **2001**, *283* (2), 460−8.

(29) Hashimoto, Y.; Niikura, T.; Tajima, H.; Yasukawa, T.; Sudo, H.; Ito, Y.; Kita, Y.; Kawasumi, M.; Kouyama, K.; Doyu, M.; Sobue, G.; Koide, T.; Tsuji, S.; Lang, J.; Kurokawa, K.; Nishimoto, I. A rescue factor abolishing neuronal cell death by a wide spectrum of familial Alzheimer's disease genes and Abeta. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (11), 6336−41.

(30) Zapała, B.; Kaczyński, Ł.; Kieć-Wilk, B.; Staszel, T.; Knapp, A.; Hege Thoresen, G.; Wybrańska, I.; Dembińska-Kieć, A. Humanins, the neuroprotective and cytoprotective peptides with antiapoptotic and anti-inflammatory properties. *Pharmacological Reports* **2010**, *62* (5), 767−777.

(31) Yamagishi, Y.; Hashimoto, Y.; Niikura, T.; Nishimoto, I. Identification of essential amino acids in Humanin, a neuroprotective factor against Alzheimer's disease-relevant insults. *Peptides* **2003**, *24* (4), 585−595.

(32) Ikonen, M.; Liu, B.; Hashimoto, Y.; Ma, L.; Lee, K. W.; Niikura, T.; Nishimoto, I.; Cohen, P. Interaction between the Alzheimer's survival peptide humanin and insulin-like growth factor-binding protein 3 regulates cell survival and apoptosis. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (22), 13042−7.

(33) Bachar, A. R.; Scheffer, L.; Schroeder, A. S.; Nakamura, H. K.; Cobb, L. J.; Oh, Y. K.; Lerman, L. O.; Pagano, R. E.; Cohen, P.; Lerman, A. Humanin is expressed in human vascular walls and has a cytoprotective effect against oxidized LDL-induced oxidative stress. *Cardiovascular research* **2010**, *88* (2), 360−6.

(34) Kin, T.; Sugie, K.; Hirano, M.; Goto, Y.; Nishino, I.; Ueno, S. Humanin expression in skeletal muscles of patients with chronic progressive external ophthalmoplegia. *Journal of human genetics* **2006**, *51* (6), 555−8.

(35) Zacharias, D. G.; Kim, S. G.; Massat, A. E.; Bachar, A. R.; Oh, Y. K.; Herrmann, J.; Rodriguez-Porcel, M.; Cohen, P.; Lerman, L. O.; Lerman, A. Humanin, a cytoprotective peptide, is expressed in carotid atherosclerotic [corrected] plaques in humans. *PloS one* **2012**, *7* (2), No. e31065.

(36) Minasyan, L.; Sreekumar, P. G.; Hinton, D. R.; Kannan, R. Protective Mechanisms of the Mitochondrial-Derived Peptide Humanin in Oxidative and Endoplasmic Reticulum Stress in RPE Cells. *Oxidative medicine and cellular longevity* **2017**, *2017*, 1675230.

(37) Matsuoka, M. Protective effects of Humanin and calmodulin-like skin protein in Alzheimer's disease and broad range of abnormalities. *Molecular neurobiology* **2015**, *51* (3), 1232−9.

(38) Charununtakorn, S. T.; Shinlapawittayatorn, K.; Chattipakorn, S. C.; Chattipakorn, N. Potential Roles of Humanin on Apoptosis in the Heart. *Cardiovascular therapeutics* **2016**, *34* (2), 107−14.

(39) Xiao, J.; Kim, S. J.; Cohen, P.; Yen, K. Humanin: Functional Interfaces with IGF-I. *Growth hormone & IGF research: official journal*

of the Growth Hormone Research Society and the International IGF Research Society **2016**, *29*, 21−27.

(40) Lee, C.; Kim, K. H.; Cohen, P. MOTS-c: A novel mitochondrial-derived peptide regulating muscle and fat metabolism. *Free radical biology & medicine* **2016**, *100*, 182−187.

(41) Lee, C.; Zeng, J.; Drew, B. G.; Sallam, T.; Martin-Montalvo, A.; Wan, J.; Kim, S. J.; Mehta, H.; Hevener, A. L.; de Cabo, R.; Cohen, P. The mitochondrial-derived peptide MOTS-c promotes metabolic homeostasis and reduces obesity and insulin resistance. *Cell metabolism* **2015**, *21* (3), 443−54.

(42) Ming, W.; Lu, G.; Xin, S.; Huanyu, L.; Yinghao, J.; Xiaoying, L.; Chengming, X.; Banjun, R.; Li, W.; Zifan, L. Mitochondria related peptide MOTS-c suppresses ovariectomy-induced bone loss via AMPK activation. *Biochemical and biophysical research communications* **2016**, *476* (4), 412−419.

(43) Zhai, D.; Ye, Z.; Jiang, Y.; Xu, C.; Ruan, B.; Yang, Y.; Lei, X.; Xiang, A.; Lu, H.; Zhu, Z.; Yan, Z.; Wei, D.; Li, Q.; Wang, L.; Lu, Z. MOTS-c peptide increases survival and decreases bacterial load in mice infected with MRSA. *Molecular immunology* **2017**, *92*, 151−160.

(44) Li, N.; Han, Z.-L.; Xu, B.; Zhang, M.-N.; Zhang, T.; Shi, X.-R.; Zhao, W.-D.; Guo, Y.-Y.; Zhang, Q.-Q.; Fang, Q. Systemic administration of the bifunctional opioid/neuropeptide FF receptors agonist BN-9 produced peripheral antinociception in preclinical mouse models of pain. *Eur. J. Pharmacol.* **2018**, *837*, 53−63.

(45) Kim, K. H.; Son, J. M.; Benayoun, B. A.; Lee, C. The Mitochondrial-Encoded Peptide MOTS-c Translocates to the Nucleus to Regulate Nuclear Gene Expression in Response to Metabolic Stress. *Cell metabolism* **2018**, *28* (3), 516−524.

(46) Yin, X.; Jing, Y.; Chen, Q.; Abbas, A. B.; Hu, J.; Xu, H. The intraperitoneal administration of MOTS-c produces antinociceptive and anti-inflammatory effects through the activation of AMPK pathway in the mouse formalin test. *European journal of pharmacology* **2020**, *870*, 172909.

(47) Xinqiang, Y.; Quan, C.; Yuanyuan, J.; Hanmei, X. Protective effect of MOTS-c on acute lung injury induced by lipopolysaccharide in mice. *International immunopharmacology* **2020**, *80*, 106174.

(48) Wu, H.; Zhao, G.; Jiang, K.; Chen, X.; Zhu, Z.; Qiu, C.; Li, C.; Deng, G. Plantamajoside ameliorates lipopolysaccharide-induced acute lung injury via suppressing NF-*κ*B and MAPK activation. *International Immunopharmacology* **2016**, *35*, 315−322.

(49) Zhang, S.; Reljic, B.; Liang, C.; Kerouanton, B.; Francisco, J. C.; Peh, J. H.; Mary, C.; Jagannathan, N. S.; Olexiouk, V.; Tang, C.; Fidelito, G.; Nama, S.; Cheng, R. K.; Wee, C. L.; Wang, L. C.; Duek Roggli, P.; Sampath, P.; Lane, L.; Petretto, E.; Sobota, R. M.; Jesuthasan, S.; Tucker-Kellogg, L.; Reversade, B.; Menschaert, G.; Sun, L.; Stroud, D. A.; Ho, L. Mitochondrial peptide BRAWNIN is essential for vertebrate respiratory complex III assembly. *Nat. Commun.* **2020**, *11* (1), 1312.

(50) Lee, C. Q. E.; Kerouanton, B.; Chothani, S.; Zhang, S.; Chen, Y.; Mantri, C. K.; Hock, D. H.; Lim, R.; Nadkarni, R.; Huynh, V. T.; Lim, D.; Chew, W. L.; Zhong, F. L.; Stroud, D. A.; Schafer, S.; Tergaonkar, V.; St John, A. L.; Rackham, O. J. L.; Ho, L. Coding and non-coding roles of MOCCI (C15ORF48) coordinate to regulate host inflammation and immunity. *Nat. Commun.* **2021**, *12* (1), 2130.

(51) Martinez, T. F.; Chu, Q.; Donaldson, C.; Tan, D.; Shokhirev, M. N.; Saghatelian, A. Accurate annotation of human protein-coding small open reading frames. *Nat. Chem. Biol.* **2020**, *16* (4), 458−468.

(52) Matsumoto, A.; Pasut, A.; Matsumoto, M.; Yamashita, R.; Fung, J.; Monteleone, E.; Saghatelian, A.; Nakayama, K. I.; Clohessy, J. G.; Pandolfi, P. P. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature* **2017**, *541* (7636), 228−232.

(53) D'Lima, N. G.; Ma, J.; Winkler, L.; Chu, Q.; Loh, K. H.; Corpuz, E. O.; Budnik, B. A.; Lykke-Andersen, J.; Saghatelian, A.; Slavoff, S. A. A human microprotein that interacts with the mRNA decapping complex. *Nat. Chem. Biol.* **2017**, *13* (2), 174−180.

(54) Na, Z.; Luo, Y.; Schofield, J. A.; Smelyansky, S.; Khitun, A.; Muthukumar, S.; Valkov, E.; Simon, M. D.; Slavoff, S. A. The NBDY Microprotein Regulates Cellular RNA Decapping. *Biochemistry* **2020**, *59* (42), 4131−4142.

(55) Na, Z.; Luo, Y.; Cui, D. S.; Khitun, A.; Smelyansky, S.; Loria, J. P.; Slavoff, S. A. Phosphorylation of a Human Microprotein Promotes Dissociation of Biomolecular Condensates. *J. Am. Chem. Soc.* **2021**, *143* (32), 12675−12687.

(56) Min, K. W.; Davila, S.; Zealy, R. W.; Lloyd, L. T.; Lee, I. Y.; Lee, R.; Roh, K. H.; Jung, A.; Jemielity, J.; Choi, E. J.; Chang, J. H.; Yoon, J. H. eIF4E phosphorylation by MST1 reduces translation of a subset of mRNAs, but increases lncRNA translation. *Biochimica et biophysica acta. Gene regulatory mechanisms* **2017**, *1860* (7), 761−772.

(57) Bakhti, Z. S.; Latifi-Navid, S. Non-coding RNA-Encoded Peptides/Proteins in Human Cancer: The Future for Cancer Therapy. *Curr. Med. Chem.* **2021**, *28*, 1−1.

(58) Ye, M.; Zhang, J.; Wei, M.; Liu, B.; Dong, K. Emerging role of long noncoding RNA-encoded micropeptides in cancer. *Cancer cell international* **2020**, *20*, 506.

(59) Merino-Valverde, I.; Greco, E.; Abad, M. The microproteome of cancer: From invisibility to relevance. *Experimental cell research* **2020**, *392* (1), 111997.

(60) Li, M.; Li, X.; Zhang, Y.; Wu, H.; Zhou, H.; Ding, X.; Zhang, X.; Jin, X.; Wang, Y.; Yin, X.; Li, C.; Yang, P.; Xu, H. A Micropeptide MIAC Inhibits HNSCC Progression by Interacting with Aquaporin 2. *J. Am. Chem. Soc.* **2020**, *142* (14), 6708−6716.

(61) Li, M.; Shao, F.; Qian, Q.; Yu, W.; Zhang, Z.; Chen, B.; Su, D.; Guo, Y.; Phan, A. V.; Song, L. S.; Stephens, S. B.; Sebag, J.; Imai, Y.; Yang, L.; Cao, H. A putative long noncoding RNA-encoded micropeptide maintains cellular homeostasis in pancreatic beta cells. *Molecular therapy. Nucleic acids* **2021**, *26*, 307−320.

(62) Pang, Y.; Liu, Z.; Han, H.; Wang, B.; Li, W.; Mao, C.; Liu, S. Peptide SMIM30 promotes HCC development by inducing SRC/YES1 membrane anchoring and MAPK pathway activation. *Journal of Hepatology* **2020**, *73* (5), 1155−1169.

(63) Zhu, S.; Wang, J. Z.; Chen; He, Y. T.; Meng, N.; Chen, M.; Lu, R. X.; Chen, X. H.; Zhang, X. L.; Yan, G. R. An oncopeptide regulates m(6)A recognition by the m(6)A reader IGF2BP1 and tumorigenesis. *Nat. Commun.* **2020**, *11* (1), 1685.

(64) Guo, B.; Wu, S.; Zhu, X.; Zhang, L.; Deng, J.; Li, F.; Wang, Y.; Zhang, S.; Wu, R.; Lu, J.; Zhou, Y. Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression. *EMBO journal* **2020**, *39* (1), No. e102190.

(65) Wu, S.; Zhang, L.; Deng, J.; Guo, B.; Li, F.; Wang, Y.; Wu, R.; Zhang, S.; Lu, J.; Zhou, Y. A Novel Micropeptide Encoded by Y-Linked LINC00278 Links Cigarette Smoking and AR Signaling in Male Esophageal Squamous Cell Carcinoma. *Cancer research* **2020**, *80* (13), 2790−2803.

(66) Wang, Y.; Wu, S.; Zhu, X.; Zhang, L.; Deng, J.; Li, F.; Guo, B.; Zhang, S.; Wu, R.; Zhang, Z.; Wang, K.; Lu, J.; Zhou, Y. LncRNA-encoded polypeptide ASRPS inhibits triple-negative breast cancer angiogenesis. *Journal of Experimental Medicine* **2020**, *217* (3), e20190950.

(67) Xu, W.; Deng, B.; Lin, P.; Liu, C.; Li, B.; Huang, Q.; Zhou, H.; Yang, J.; Qu, L. Ribosome profiling analysis identified a KRAS-interacting microprotein that represses oncogenic signaling in hepatocellular carcinoma cells. *Science China. Life sciences* **2020**, *63* (4), 529−542.

(68) Polycarpou-Schwarz, M.; Gross, M.; Mestdagh, P.; Schott, J.; Grund, S. E.; Hildenbrand, C.; Rom, J.; Aulmann, S.; Sinn, H. P.; Vandesompele, J.; Diederichs, S. The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation. *Oncogene* **2018**, *37* (34), 4750−4768.

(69) Huang, J. Z.; Chen, M.; Chen; Gao, X. C.; Zhu, S.; Huang, H.; Hu, M.; Zhu, H.; Yan, G. R. A Peptide Encoded by a Putative lncRNA HOXB-AS3 Suppresses Colon Cancer Growth. *Molecular cell* **2017**, *68* (1), 171−184.

(70) Szafron, L. M.; Balcerak, A.; Grzybowska, E. A.; Pienkowska-Grela, B.; Felisiak-Golabek, A.; Podgorska, A.; Kulesza, M.; Nowak, N.; Pomorski, P.; Wysocki, J.; Rubel, T.; Dansonka-Mieszkowska, A.;

Konopka, B.; Lukasik, M.; Kupryjanczyk, J. The Novel Gene CRNDE Encodes a Nuclear Peptide (CRNDEP) Which Is Overexpressed in Highly Proliferating Tissues. *PloS one* **2015**, *10* (5), No. e0127475.

(71) Pamudurti, N. R.; Bartok, O.; Jens, M.; Ashwal-Fluss, R.; Stottmeister, C.; Ruhe, L.; Hanan, M.; Wyler, E.; Perez-Hernandez, D.; Ramberger, E.; Shenzis, S.; Samson, M.; Dittmar, G.; Landthaler, M.; Chekulaeva, M.; Rajewsky, N.; Kadener, S. Translation of CircRNAs. *Molecular cell* **2017**, *66* (1), 9−21.

(72) Chen, C.-y.; Sarnow, P. Initiation of Protein Synthesis by the Eukaryotic Translational Apparatus on Circular RNAs. *Science* **1995**, *268* (5209), 415−417.

(73) Wang, Y.; Wang, Z. Efficient backsplicing produces translatable circular mRNAs. *Rna* **2015**, *21* (2), 172−9.

(74) Chen, X.; Han, P.; Zhou, T.; Guo, X.; Song, X.; Li, Y. circRNADb: A comprehensive database for human circular RNAs with protein-coding annotations. *Sci. Rep.* **2016**, *6*, 34985.

(75) Zhang, M.; Zhao, K.; Xu, X.; Yang, Y.; Yan, S.; Wei, P.; Liu, H.; Xu, J.; Xiao, F.; Zhou, H.; Yang, X.; Huang, N.; Liu, J.; He, K.; Xie, K.; Zhang, G.; Huang, S.; Zhang, N. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nat. Commun.* **2018**, *9* (1), 4475.

(76) Li, F.; Cai, Y.; Deng, S.; Yang, L.; Liu, N.; Chang, X.; Jing, L.; Zhou, Y.; Li, H. A peptide CORO1C-47aa encoded by the circular noncoding RNA circ-0000437 functions as a negative regulator in endometrium tumor angiogenesis. *J. Biol. Chem.* **2021**, *297* (5), 101182.

(77) Zheng, X.; Chen, L.; Zhou, Y.; Wang, Q.; Zheng, Z.; Xu, B.; Wu, C.; Zhou, Q.; Hu, W.; Wu, C.; Jiang, J. A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling. *Molecular cancer* **2019**, *18* (1), 47.

(78) Dever, T. E.; Ivanov, I. P.; Sachs, M. S. Conserved Upstream Open Reading Frame Nascent Peptides That Control Translation. *Annual review of genetics* **2020**, *54*, 237−264.

(79) Starck, S. R.; Tsai, J. C.; Chen, K.; Shodiya, M.; Wang, L.; Yahiro, K.; Martins-Green, M.; Shastri, N.; Walter, P. Translation from the 5′ untranslated region shapes the integrated stress response. *Science* **2016**, *351* (6272), aad3867.

(80) Chen, J.; Brunner, A.-D.; Cogan, J. Z.; Nunez, J. K.; Fields, A. P.; Adamson, B.; Itzhak, D. N.; Li, J. Y.; Mann, M.; Leonetti, M. D.; Weissman, J. S. Pervasive functional translation of noncanonical human open reading frames. *Science* **2020**, *367* (6482), 1140−1146.

(81) Jousse, C.; Bruhat, A.; Carraro, V.; Urano, F.; Ferrara, M.; Ron, D.; Fafournoux, P. Inhibition of CHOP translation by a peptide encoded by an open reading frame localized in the chop 5′UTR. *Nucleic Acids Research* **2001**, *29* (21), 4341−4351.

(82) Rathore, A.; Chu, Q.; Tan, D.; Martinez, T. F.; Donaldson, C. J.; Diedrich, J. K.; Yates, J. R., 3rd; Saghatelian, A. MIEF1Microprotein Regulates Mitochondrial Translation. *Biochemistry* **2018**, *57* (38), 5564−5575.

(83) Cloutier, P.; Poitras, C.; Faubert, D.; Bouchard, A.; Blanchette, M.; Gauthier, M. S.; Coulombe, B. Upstream ORF-Encoded ASDURF Is a Novel Prefoldin-like Subunit of the PAQosome. *J. Proteome Res.* **2020**, *19* (1), 18−27.

(84) Kikuchi, K.; Fukuda, M.; Ito, T.; Inoue, M.; Yokoi, T.; Chiku, S.; Mitsuyama, T.; Asai, K.; Hirose, T.; Aizawa, Y. Transcripts of unknown function in multiple-signaling pathways involved in human stem cell differentiation. *Nucleic acids research* **2009**, *37* (15), 4987−5000.

(85) Chng, S. C.; Ho, L.; Tian, J.; Reversade, B. ELABELA: a hormone essential for heart development signals via the apelin receptor. *Developmental cell* **2013**, *27* (6), 672−80.

(86) Slavoff, S. A.; Heo, J.; Budnik, B. A.; Hanakahi, L. A.; Saghatelian, A. A human short open reading frame (sORF)-encoded polypeptide that stimulates DNA end joining. *J. Biol. Chem.* **2014**, *289* (16), 10950−7.

(87) Grundy, G. J.; Rulten, S. L.; Arribas-Bosacoma, R.; Davidson, K.; Kozik, Z.; Oliver, A. W.; Pearl, L. H.; Caldecott, K. W. The Ku-binding motif is a conserved module for recruitment and stimulation of non-homologous end-joining proteins. *Nat. Commun.* **2016**, *7*, 11242.

(88) Castaneda-Zegarra, S.; Huse, C.; Rosand, O.; Sarno, A.; Xing, M.; Gago-Fuentes, R.; Zhang, Q.; Alirezaylavasani, A.; Werner, J.; Ji, P.; Liabakk, N. B.; Wang, W.; Bjoras, M.; Oksenych, V. Generation of a Mouse Model Lacking the Non-Homologous End-Joining Factor Mri/Cyren. *Biomolecules* **2019**, *9* (12), 798.

(89) Hung, P. J.; Johnson, B.; Chen, B.-R.; Byrum, A. K.; Bredemeyer, A. L.; Yewdell, W. T.; Johnson, T. E.; Lee, B. J.; Deivasigamani, S.; Hindi, I.; Amatya, P.; Gross, M. L.; Paull, T. T.; Pisapia, D. J.; Chaudhuri, J.; Petrini, J. J. H.; Mosammaparast, N.; Amarasinghe, G. K.; Zha, S.; Tyler, J. K.; Sleckman, B. P. MRI Is a DNA Damage Response Adaptor during Classical Non-homologous End Joining. *Molecular cell* **2018**, *71* (2), 332−342.

(90) Chu, Q.; Martinez, T. F.; Novak, S. W.; Donaldson, C. J.; Tan, D.; Vaughan, J. M.; Chang, T.; Diedrich, J. K.; Andrade, L.; Kim, A.; Zhang, T.; Manor, U.; Saghatelian, A. Regulation of the ER stress response by a mitochondrial microprotein. *Nat. Commun.* **2019**, *10* (1), 4883.

(91) Lu, Y.; Li, Z.; Lin, C.; Zhang, J.; Shen, Z. Translation role of circRNAs in cancers. *Journal of Clinical Laboratory Analysis* **2021**, *35* (7), No. e23866.

(92) Schlesinger, D.; Elsasser, S. J. Revisiting sORFs: overcoming challenges to identify and characterize functional microproteins. *FEBS journal* **2022**, *289* (1), 53−74.

(93) Wang, B.; Hao, J.; Pan, N.; Wang, Z.; Chen, Y.; Wan, C. Identification and analysis of small proteins and short open reading frame encoded peptides in Hep3B cell. *Journal of proteomics* **2021**, *230*, 103965.

(94) Zhang, Q.; Wu, E.; Tang, Y.; Cai, T.; Zhang, L.; Wang, J.; Hao, Y.; Zhang, B.; Zhou, Y.; Guo, X.; Luo, J.; Chen, R.; Yang, F. Deeply Mining a Universe of Peptides Encoded by Long Noncoding RNAs. *Molecular & cellular proteomics: MCP* **2021**, *20*, 100109.