Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) 2023 24(12):1106-1122
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Review

https://doi.org/10.1631/jzus.B2300128

Micropeptides: origins, identification, and potential role in metabolism-related diseases

Yirui LU, Yutong RAN, Hong LI, Jiao WEN, Xiaodong CUI, Xiaoyun ZHANG, Xiumei GUAN, Min CHENG✉

School of Basic Medicine Sciences, Weifang Medical University, Weifang 261053, China

Abstract: With the development of modern sequencing techniques and bioinformatics, genomes that were once thought to be noncoding have been found to encode abundant functional micropeptides (miPs), a kind of small polypeptides. Although miPs are difficult to analyze and identify, a number of studies have begun to focus on them. More and more miPs have been revealed as essential for energy metabolism homeostasis, immune regulation, and tumor growth and development. Many reports have shown that miPs are especially essential for regulating glucose and lipid metabolism and regulating mitochondrial function. MiPs are also involved in the progression of related diseases. This paper reviews the sources and identification of miPs, as well as the functional significance of miPs for metabolism-related diseases, with the aim of revealing their potential clinical applications.

Key words: Energy metabolism; Micropeptides; Mitochondria; Noncoding RNA (ncRNA); Small open reading frame (sORF)

1 Introduction

Micropeptides (miPs) are small peptides derived from noncoding RNAs (ncRNAs) and are encoded by small open reading frames (sORFs) with a size of less than 100 amino acids. MiPs contain only one protein–protein interaction (PPI) domain and can form non-functional heterodimers to exert adverse regulatory effects. They also generate new protein complexes with the target protein (Staudt and Wenkel, 2011; Bhati et al., 2018). It has been shown that miPs exert many biological functions, with studies revealing that miPs participate in embryonic growth and development. Additionally, miPs play an essential role in immunity and inflammation (Pauli et al., 2014; Bhatta et al., 2020; Niu et al., 2020). Some miPs mediate immune escape and enhance the virulence of pathogens (Zhu et al., 2021), and some reduce inflammation by regulating antigen presentation in dendritic cells (Niu et al., 2020). Major histocompatibility complex class I (MHC-I) immune peptide fragments in cancer are also derived from open reading frames (ORFs) (Ouspenskaia et al., 2022). In terms of muscle function and regeneration, miPs affect muscle function by regulating calcium transport and regulating muscle regeneration by interacting with mammalian target of rapamycin complex 1 (mTORC1) (Magny et al., 2013; Anderson et al., 2015; Matsumoto et al., 2017). It is noteworthy that some miPs, in the process of regulating muscle function, participate in mitochondrial metabolism, affect energy metabolic homeostasis, and regulate fatty acid metabolism (Makarewich et al., 2018; Stein et al., 2018). MiPs also regulate glucose and lipid metabolism processes, which improve insulin resistance, decrease serum cholesterol, and increase fatty acid oxidation (Bal et al., 2012; Tumminia et al., 2018; Friesen et al., 2020). The regulatory function of miPs in glucose metabolism and lipid metabolism may be a new target for the treatment of atherosclerotic cardiovascular disease caused by abnormal lipid metabolism and heart disease caused by abnormal energy metabolism (Ference et al., 2017; Tezze et al., 2017; Kamps et al., 2018; Makarewich et al., 2022). MiPs also regulate cell proliferation and the cell cycle of tumor cells and inhibit tumor progression by regulating glucose metabolism and adenosine

✉ Min CHENG, mincheng@wfmc.edu.cn
Min CHENG, https://orcid.org/0000-0002-2814-5805

Received Feb. 24, 2023; Revision accepted June 6, 2023;
Crosschecked Oct. 10, 2023

triphosphate (ATP) production (Huang et al., 2017; Ge et al., 2021; Polenkowski et al., 2021). Here, we summarize the miP sources, identification techniques, computational analysis methods, and the mechanisms by which miPs regulate diabetes and Alzheimer's disease (AD), amongst others.

MiPs also have rich biological functions in translation, transcription, waste degradation, and DNA repair, and are involved in immune processes and carcinogenesis, which have been elucidated in previous reviews (Hartford and Lal, 2020; Pan et al., 2022).

2 Origins of miPs

NcRNAs can be classified as long noncoding RNAs (lncRNAs), circular RNAs (circRNAs), microRNAs (miRNAs), and small nucleolar RNAs (snoRNAs), all of which have different functions (Adams et al., 2017; Qu et al., 2022; Xu et al., 2022). In recent years, increasing evidence has been found showing that some ncRNAs can encode miPs.

2.1 LncRNAs and untranslated region (UTR)

LncRNAs are "noncoding" RNAs greater than 200 nucleotides (nt) that were once thought to lack coding ability (Rinn and Chang, 2012). With the development of modern sequencing technologies, the Encyclopedia of DNA Elements has shown that lncRNAs have translation potential (The ENCODE Project Consortium, 2007; Wang S et al., 2019). Most lncRNAs are similar to messenger RNAs (mRNAs) at the molecular level, with a 5'-m7GpppN cap structure and a polycistronic A tail (Ulitsky and Bartel, 2013) (Fig. 1a), which is the structural basis for the translation of lncRNAs. The transcriptional and post-transcriptional modification processes of lncRNAs are similar to those of conventional protein-coding genes (Kapusta et al., 2013) (Table 1). There are many functional miPs encoded by lncRNAs; for example, lncRNA *LINC00116* encodes the 56 amino acid (aa)-long microprotein mitoregulin (MTLN), which is localized in mitochondria and regulates the activity of respiratory complex I and lipid metabolism (Chugunova et al., 2019). *LINC0665* is another lncRNA localized in the cytoplasm, and its ORF1 can also be translated into miPs, known as cancerous inhibitor of protein phosphatase 2A (CIP2A)-binding peptides inhibiting triple-negative breast cancer (Guo et al., 2020). ATP synthase-associated peptide (ASAP) is a 94 aa-sized miP encoded by the first sORF at the 5' end of *LINC00467*, which is involved in ATP production and the proliferation of rectal colon cancer cells. ASAP is localized to the inner mitochondrial membrane and regulates mitochondrial ATP production (Ge et al., 2021). In addition, lncRNA-derived miPs not only regulate proliferation and metabolism to participate in the occurrence and development of tumors but also to participate in the process of tumor immunity, which is of great significance for the tumor treatment (Andjus et al., 2021; Xu et al., 2023).

The 5' untranslated region (5' UTR) and 3' UTR are involved in the regulation and expression of miPs (Dangelmaier et al., 2022; Yang et al., 2023). Upstream open reading frames (uORFs) are in the 5' UTR (Ingolia et al., 2014). MiPs derived from uORFs in the 5' UTR show different biological functions (Akimoto et al., 2013; Rathore et al., 2018). Most translation initiation sites are the 5' AUG codons, but there is also non-standard translation UUG or CUG initiation (Ingolia et al., 2014; Starck et al., 2016). Downstream open reading frames (dORFs) are also present in the 3' UTR and these ORFs produce peptides by way of a process similar to that of gene-encoding proteins (Ji et al., 2015; van Heesch et al., 2019).

2.2 CircRNA

CircRNA is a special group of ncRNAs; the precursor RNAs are processed into mature linear mRNAs by conventional splicing, and unconventional splicing allows the exons to form a circle to produce circRNA, whose length ranges from less than 100 nt to several thousand nucleotides (Schindewolf et al., 1996; Wu et al., 2020). CircRNA does not have the 5'-cap structure required for the process of ribosome translation to produce proteins. CircRNA was once thought to be a byproduct of shearing errors; however, there is growing evidence that circRNA is involved in protein translation (Prats et al., 2020). CircRNA initiates translation with structures such as the internal ribosome entry site (IRES) and N^6-methyladenosine (m^6A) (Table 1). The IRES element is a nucleotide sequence that allows the ribosome to initiate translation directly in the middle of the mRNA sequence. At the same time, the m^6A modification binds to the YTH structural domain family protein 3 (YTHD3)

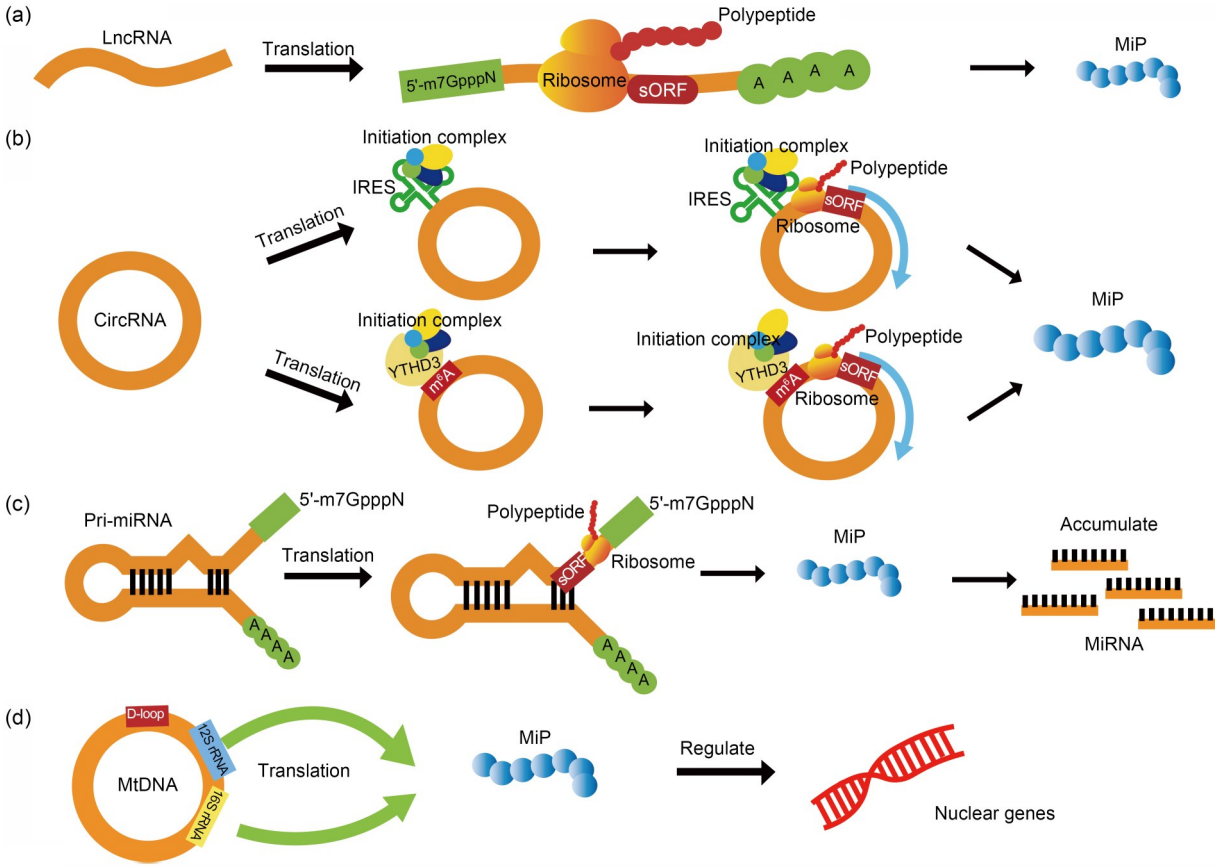

Fig. 1 Origins of micropeptides (miPs). (a) MiPs from long noncoding RNAs (lncRNAs). The translation process of lncRNAs, which have a 5' cap structure and a multimeric A-tail similar to those of messenger RNAs (mRNAs), and the translation initiation site is 5'-m7GpppN. (b) MiPs from circular RNAs (circRNAs). The translation initiation sites of circRNAs are the internal ribosomal entry site (IRES) and N^6-methyladenosine (m^6A)-modified site. IRES binds to the translation initiation complex to initiate translation. The m^6A-modified site first binds to YTH structural domain family protein 3 (YTHD3) and then recruits the translation initiation complex to initiate translation. (c) The translation process of primary microRNA (pri-miRNA), a precursor of miRNA, also has a 5' cap structure and a multimeric A-tail, and the translated miP can promote the accumulation of its cognate miRNAs. (d) Mitochondrial DNA (mtDNA) expresses microproteins. MtDNA replication and transcription start at the mitochondrial displacement loop (D-loop), where 16S ribosomal RNA (rRNA) and 12S rRNA can be translated to produce miPs. These microproteins regulate nuclear genes.

and recruits the translation initiation complex to initiate translation (Sinha et al., 2022). These two translation initiation structures are the structural basis upon which circRNA encodes proteins (Fig. 1b) (Chen et al., 2016). For example, *circ-FBXW7*, mainly located in the cytoplasm, is able to express the 185-aa microprotein FBXW7-185aa, which regulates the cell cycle and reduces the proliferation of glioma cells (Yang et al., 2018).

2.3 MiRNAs

Mature miRNAs are produced by the cleavage of primary miRNAs (pri-miRNAs) by a family of RNA polymerases. RNA polymerase II caps and polyadenylates pri-miRNAs in a series of processes similar to those known for conventional protein coding, which gives pri-miRNAs a structural basis for encoding proteins (Winter et al., 2009) (Table 1). Pri-miRNAs are capable of encoding miPs (Fig. 1c). Plant pri-miR171b and pri-miR165a contain short ORF sequences that encode miP-EP171b and miP-EP165a. These miPs promote the accumulation of their cognate miRNAs by activating transcriptional activators of their corresponding pri-miRNAs. The mechanism by which miRNA-encoding miPs exert their biological functions is closely related to the increased transcription of pri-miRNAs (Lauressergues et al., 2015). During tumorigenesis, miP-EP133, a tumor suppressor microprotein encoded by pri-miRNAs, inhibits tumor progression by regulating mitochondrial energy metabolism (Kang et al., 2020).

Table 1 Origin and characteristics of miPs

Origin	Characteristics	Examples	References
LncRNA	Mostly similar to mRNAs; translation starts from a 5' cap structure with a polyA tail and an exon-to-exon junction linkage.	Pncr003:2L, SPAR, HOXB-AS3, mitoregulin, ASAP	Magny et al., 2013; Matsumoto et al., 2017; Chugunova et al., 2019; Wang JZ et al., 2019; Ge et al., 2021
CircRNA	No 5' cap structure; IRES and m^6A structures required for translation.	FBXW7-185aa	Yang et al., 2018
MiRNA	MiPs are usually produced from their precursor pri-miRNAs and a series of processes, such as cap addition and polyadenylation, are similar to those known for conventional protein coding.	Pri-miR171b, pri-miR165a, miP-EP133	Lauressergues et al., 2015; Kang et al., 2020
MtDNA	MiPs are produced by the two rRNAs they produce and miPs of this origin regulate not only mitochondrial function but also nuclear gene expression.	Humanin, SHLP1–6, OPA1, MOTS-c, SHMOOSE	Tezze et al., 2017; Merry et al., 2020; Miller et al., 2022b
Protein hydrolysis/ cleavage	Usually occurs in vivo, where proteases specifically hydrolyze larger precursor proteins.	Serum response factor cleavage product, cross-reactive DNA binding protein-43 N-terminal cleavage product, insulin, BNP	Wong et al., 2012; Fung et al., 2015; Ramzy and Kieffer, 2022

MiPs: micropeptides; LncRNA: long noncoding RNA; CircRNA: circular RNA; MiRNA: microRNA; MtDNA: mitochondrial DNA; mRNA: messenger RNA; polyA: polycistronic A; IRES: internal ribosome entry site; m^6A: N^6-methyladenosine; Pri-miRNA: primary miRNA; rRNA: ribosomal RNA; SPAR: small regulatory polypeptide of amino acid response; ASAP: adenosine triphosphate (ATP) synthase-associated peptide; SHLP: small humanin-like peptide; OPA1: optic atrophy type 1 protein; MOTS-c: mitochondrial open reading frame of the 12S rRNA-c; SHMOOSE: small human mitochondrial open reading frame over SErine transfer RNA; BNP: brain natriuretic peptide.

2.4 Mitochondrial DNA-derived miPs

Mitochondria, the center of energy metabolism, are complex organelles with their own genome, namely mitochondrial DNA (mtDNA). MtDNA is a double-stranded DNA with 16.6-kb molecule, involved in the oxidative phosphorylation complex, 11 mRNAs, 2 ribosomal RNAs (12S and 16S rRNAs), and 22 transfer RNAs (tRNAs) coding and synthesis (Gustafsson et al., 2016; Boguszewska et al., 2020) (Table 1). Some miPs can be transcriptionally translated from mtDNA. The proportion of miPs in mitochondria is much higher than that in the cytoplasmic matrix (Zhang et al., 2020). The starting point of mtDNA replication and the start of transcription are located in the control region displacement loop (D-loop) (Boguszewska et al., 2020) (Fig. 1d). In addition, sORF in mitochondrial rRNAs can be translated into miPs with protective metabolic functions (Merry et al., 2020). For example, sORF of 16S rRNA in mtDNA encodes humanin (HN) and small humanin-like peptide 1–6 (SHLP1–6), and 12S rRNA in mtDNA encodes mitochondrial ORF of the 12S rRNA-c (MOTS-c) (Merry et al., 2020). Recent study has shown that mitochondria express proteins that can be targeted to the nucleus for reverse transduction, thereby affecting nuclear gene expression (Vizioli et al., 2020) (Table 1).

2.5 Production of miPs by protein hydrolysis/cleavage

In addition to being produced by the transcription and translation of ncRNA and mtDNA, miPs are produced by hydrolytic cleavage of large proteins. The process is the specific hydrolysis of peptide bonds of larger precursor proteins by proteases, producing two or more small peptide fragments. The creation of new amino- and carboxy-termini usually occurs in vivo and is irreversible (Bhati et al., 2018) (Table 1). Those miPs have novel biological functions and often inhibit their precursor proteins (Wong et al., 2012; Fung et al., 2015). Many small molecular peptides, which are essential hormones in the human body, are also produced by protein hydrolysis. For example, the 51 amino acids that make up insulin are derived from its precursor, proinsulin, which is cleaved by prohormone invertase to form a biologically active hormone (Ramzy

and Kieffer, 2022). Brain natriuretic peptide (BNP), a 32-amino acid polypeptide, is important in cardiovascular disease (Kuwahara, 2021). Its production process is produced by the cleavage of pro-hormone peptide of brain natriuretic peptide (ProBNP) by furin (an intracellular serine endopeptidase), which plays an essential biological role in diuretic, natriuretic, vasodilatation, renin-angiotensin-aldosterone system (RAAS) inhibition, and sympathetic inhibition (Nishikimi et al., 2015).

3 Identification of miPs

3.1 Western blot

Western blot (WB) is a routine assay for detecting proteins. However, miPs have a small molecular weight and can target a small selection of antigenic sites. Some miPs contain transmembrane structural domains that mask relatively large portions of their sequences, reducing the regions available for epitope design. In addition, due to the low expression level of miPs, even the use of high-affinity antibodies may not produce a strong detection signal (Makarewich and Olson, 2017). In this case, using clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated gene editing, a marker gene is inserted into an endogenous locus of ORF that encodes a fusion protein, enabling the detection of miPs by means of protein imprinting or mass spectrometry (MS) (Matsumoto et al., 2017) (Table 2). This method of editing Cas9 is easy to customize and can easily target new DNA sequences for efficient editing of target genomes in a variety of cell types and organisms (Ran et al., 2013) (Fig. 2a). While CRISPR can be used to edit individual nucleotides, mtDNA cannot be edited with similar precision due to the lack of efficient translocating guide RNA (gRNA) through the mitochondrial membrane. The introduction of mRNAs into the mitochondria largely depends on tRNAs, resulting in inefficient delivery in the mitochondria (Gammage et al., 2018). Recent studies have shown that the expression of mitochondrial miPs was detected indirectly by targeting nicotinamide adenine dinucleotide (NADH)-ubiquinone oxidoreductase chain 4 (ND4) to guide gRNA attachment to RNA transporter-derived stem-loop element (RP-loop) and expressing Cas9 enzymes with previous mitochondrial localization sequences (Hussain et al., 2021).

3.2 Mass spectrometry

MS is a technique for detecting and quantifying proteins and peptides, and is the gold standard for proteomics research. MS detection of peptides translated by sORF can verify the transcriptome's protein-coding

Table 2 Comparison of miP identification methods

Name	Advantages	Disadvantages	References
WB	Conventional method	Low efficiency of the test alone; must be used in conjunction with gene editing.	Makarewich and Olson, 2017
MS	The gold standard for proteomics research, identifying miPs while understanding their function	The extraction process may lose the target miPs and cannot explore the miPs of unknown coding sequences.	Makarewich and Olson, 2017; Fabre et al., 2021
Ribo-Seq	High resolution and better identification of miPs not initiated by AUG	The false positive rate is high and the requirements for testing samples are high.	Ingolia et al., 2009; Wilson and Masel, 2011; Patraquim et al., 2020
Poly-Ribo-Seq	More accurate identification of translatable sORFs	MiPs with no biological function may be detected, requiring a large sample size.	Aspden et al., 2014; Johnstone et al., 2016; Liang et al., 2018
Computational analysis method	More accurate results; predict the function of miPs	Results vary.	Bazzini et al., 2014; Kustatscher et al., 2019

miP: micropeptide; WB: western blot; MS: mass spectrometry; sORFs: small open reading frames; Ribo-Seq: ribosome sequencing; Poly-Ribo-Seq: polyribosome sequencing technology.

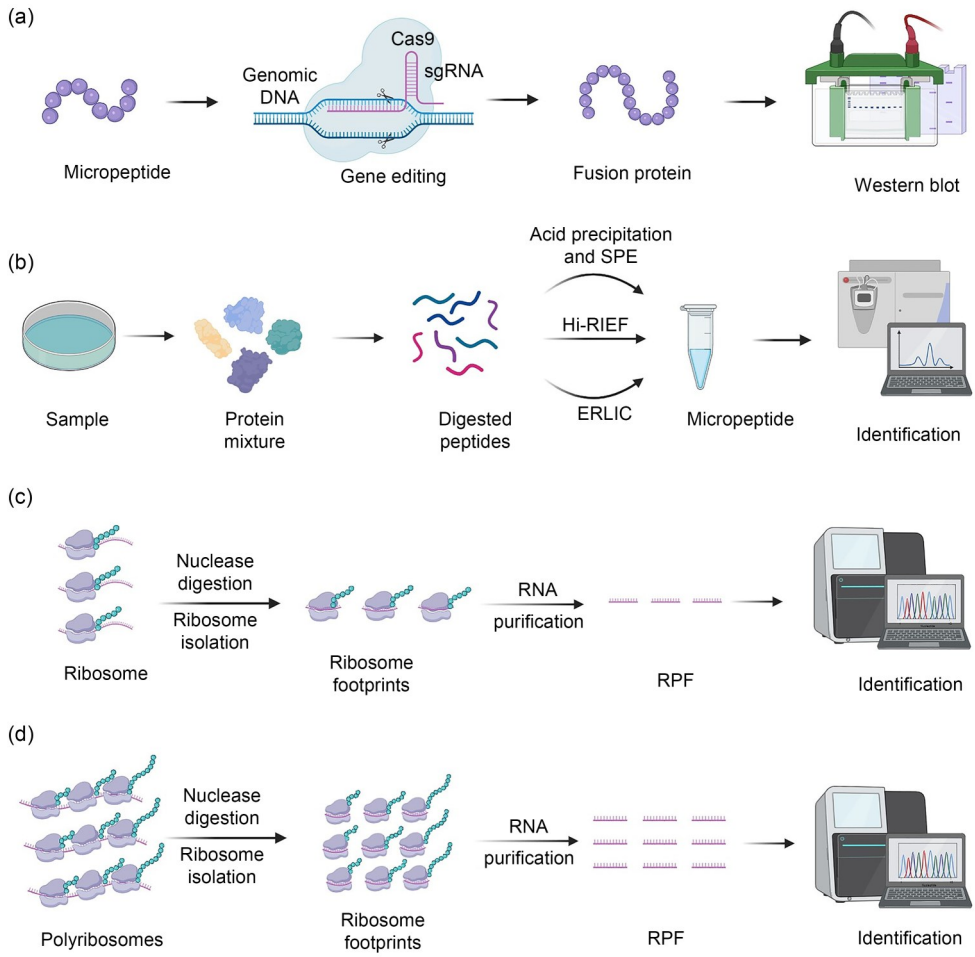

Fig. 2 Identification of micropeptides (miPs). (a) The detection of miPs by western blot (WB) requires gene editing using clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) technology to encode a fusion protein. (b) Identification of miPs by mass spectrometry (MS). Enrichment methods such as acid precipitation and reversed-phase solid phase extraction (SPE), high-resolution isoelectric focusing (Hi-RIEF), and electrostatic force-hydrophilic interaction chromatography (ERLIC) can better detect miPs by MS. (c) Ribosome analysis techniques. The target RNA is digested by nuclease and ribosome-protected fragments (RPFs) are purified and analyzed to identify small open reading frames (sORFs) that translate miPs. (d) Polyribosome sequencing technology. Similar to the ribosome analysis technique, the analysis involves examining multiple ribosome-bounded fragments to determine the sORFs of the translation miPs. sgRNA: single guide RNA.

potential (Makarewich and Olson, 2017). Compared to other methods, MS is used not only for the identification of miPs but also for the identification of their interacting proteins. Most miPs perform their biological functions by directly binding to other proteins (Fabre et al., 2021) (Table 2).

However, miPs have small protein abundance, short lifespan, and tissue specificity, which are susceptible to peptidase degradation during extraction. They need to be enriched and extracted before analysis. During this process, some of the miPs are likely to be lost or destroyed, and the methods of enrichment and extraction need to be adjusted for different localized miPs when performing MS assays (Fabre et al., 2021). Acid precipitation and reversed-phase solid phase extraction (SPE) can better enrich miPs; these two enrichment procedures maximize the recovery of trace proteins (Ma et al., 2016). In the fractionation of miPs, high-resolution isoelectric focusing (Hi-RIEF) pre-classification separation improves the detectability of miPs by MS in a highly reproducible manner (Branca et al., 2014). Electrostatic force-hydrophilic interaction chromatography (ERLIC) allows charge-driven orthogonal fractionation of peptides prior to MS (Khitun and Slavoff, 2019) (Fig. 2b). These protein enrichment and extraction methods enable MS to better detect miPs.

Although miPs can be efficiently enriched and extracted in the process of enzyme digestion, the peptide produced by enzyme digestion may be too small to be detected because of the cleavage of the peptide. Using different enzymes to digest miPs for selecting the appropriate cleavage site can increase the proportion of detectable miPs (Fabre et al., 2021). In addition, many miPs have inherently disordered or hydrophobic regions that may result in the absence of restriction sites as well as non-specific cleavage in MS (Miller et al., 2022a). Database analysis also plays a vital role in MS. The SwePep database for the analysis of small proteins requires peptides with a size of fewer than 10 kDa, which represents a large proportion of miPs that cannot be detected compared to those with molecular weights of less than 100 kDa (Lubec and Afjehi-Sadat, 2007).

3.3 Ribosome profiling

Ribosomes are important organelles involved in the translation process, which produce proteins by recognizing the coding sequence of mRNA molecules. Ribosome sequencing (Ribo-Seq) is a technique that analyzes which mRNA molecules are translated into proteins by detecting the molecular sequences of all mRNAs bound to the ribosome at a given moment. In this technique, active translation is stopped by a translation inhibitor, and the sample is treated with a nuclease to produce ribosome-protected fragments (RPFs) that are protected from digestion by the ribosome. With single nucleotide resolution, Ribo-Seq identifies the precise location of the P site of the ribosome in each sequence and, by using the feature that each codon of the ribosome corresponds to three nucleotide sequences, it is possible to predict the sequence of these RPFs (Ruiz-Orera and Albà, 2019). Then, these RPF sequences are compared to the whole genome to determine which ORFs they correspond to (Ingolia et al., 2014) (Fig. 2c). The advantage of Ribo-Seq for identifying miPs is the ability to simultaneously identify sORFs that are not AUG start codons (Kearse and Wilusz, 2017). It has been shown that more than half of the sORF translations are non-AUG initiated (Cao et al., 2020), so Ribo-Seq can better detect those non-AUG-initiated sORFs. However, the false positive rate of Ribo-Seq results is high because some sORFs bind to ribosomes but do not produce translation (Wilson and Masel, 2011) (Table 2). Based on the Ribo-Seq database, P-site databases have been developed that better enable the identification of translatable sORFs (Chothani et al., 2022). Ribo-Seq, in combination with other analytical methods (e.g., polysome analysis), is also employed to analyze and identify sORF-derived peptides on lncRNAs (Polenkowski et al., 2021).

3.4 Polyribosome sequencing technology

To address the drawback of the high false-positive rate of Ribo-Seq, polyribosome sequencing (Poly-Ribo-Seq) technology was developed to detect ORFs that produce miPs. This technology does not analyze all mRNAs that bind to ribosomes but rather mRNAs that bind multiple ribosomes. In this way, actively translated ribosome-mRNA complexes are distinguished from ribosome-mRNAs that do not produce a translation. Because mRNAs bind to multiple ribosomes during active translation, Poly-Ribo-Seq separates these polyribosomes and determines the sequence of each ribosome-binding mRNA, identifying translatable sORFs precisely (Aspden et al., 2014) (Fig. 2d). It is important to note that not all translation events result in stable functional miPs and the act of translation itself may regulate downstream ORF (Table 2) (Johnstone et al., 2016).

Both Poly-Ribo-Seq and Ribo-Seq are laborious and challenging to apply to studies with larger samples (Liang et al., 2018). Ribo-Seq, which is better at detecting large changes in mRNA translation, is not sensitive enough to analyze small changes (Masvidal et al., 2017). Previous studies have shown that there are differences in the detection efficiency of Ribo-Seq in the translation process of mRNA from whole embryos and cell lines. Ribo-Seq would reveal the changes in gene translation in a cell line but may be drowned in noise in the whole embryo (Patraquim et al., 2020). Poly-Ribo-Seq has lower requirements for detection materials, which can be met by combining with ribose immunoprecipitation (Patraquim et al., 2020). During Poly-Ribo-Seq analysis, the library of effectively translated mRNA was heavily diluted, followed by RNA isolation and purification, and enrichment during RNA precipitation by resuspension. However, for limited tissue cells or small samples, such as primary cells, a large proportion of dilution may lead to sample loss, and more efficient RNA isolation and extraction methods can make up for this deficiency (Liang et al., 2018) (Table 2). Moreover, Poly-Ribo-Seq may prefer to

detect larger sORFs because it is difficult to bind multiple ribosomes on tiny ORF fragments.

3.5 Calculation and analysis methods

MiPs whose existence and physiological function have been confirmed by WB, MS, Ribo-Seq, Poly-Ribo-Seq, and other methods were included in the database FuncPEP. In this database, names of miPs can be retrieved, as well as information on peptide-related ncRNAs, peptide length, size, molecular weight, sequence, and function (Dragomir et al., 2020). For the drawbacks of some miP identification techniques, computational analysis methods are optimized. The ORF determines whether RPFs are uniformly or centrally distributed by quantifying the number of RPFs distributed. The high or low score allows better identification of sORFs in ncRNAs (Bazzini et al., 2014). Among the ORF scoring methods, the basic local alignment search tool (BLAST), PhyloCSF, PhastCons, and other calculation methods can obtain better prediction results when combined with the ORF scoring method (Vitorino et al., 2021) (Table 2). Another integrated calculation method, which takes advantage of the high conservation of the sORF of encoded miPs, provides a database of encodable sORFs (Mackowiak et al., 2015). In terms of the prediction of the function of miPs, the expression of sORF RNAs is estimated by re-annotating microarray probes, which predict the function of miPs (Table 2). A machine learning algorithm, treeClust, is applied to reveal the functional associations between co-regulated human proteins, generating a co-regulatory map of the human proteome. This approach is able to capture the relationships between proteins without physical interactions or co-localization and also to predict the function of miPs that are difficult to study using traditional methods (Kustatscher et al., 2019).

3.6 Combined analysis of multiple methods

In screening-translatable ORFs, a combination of techniques is used to predict translatable ORFs. The researchers may look for translatable ORFs in a mass spectrum database. The ORFs are labeled with a detectable V5 tag and these ORFs are detected by detecting the V5 tag. These ORFs are then detected by ribosome analysis, MS, and bioinformatics. Finally, ORFs are translated into proteins by in vitro transcription and translation. The combination of Ribo-Seq, CRISPR technology, MS, and co-immunoprecipitation (Co-IP) enables the systematic discovery and verification of miPs and their roles in different pathways (Chen et al., 2020).

4 Regulation of energy metabolism-related diseases by miPs

4.1 MiPs and diabetes

MOTS-c, an miP encoding 16 amino acids in mitochondrial 12S rRNA, can be transferred to the nucleus and regulates the expression of nuclear genes after metabolic stress (Lee et al., 2015; Khitun et al., 2019). Previous studies have showed that MOTS-c plays a role in cellular glucose and fatty acid metabolism and is closely associated with insulin resistance and obesity (Fig. 3) (Cataldo et al., 2018; Du et al., 2018). MOTS-c can regulate the de novo synthesis of purine nucleotides and the process of folate recycling, improve insulin resistance, and regulate fat metabolism (Lee et al., 2015; Wu et al., 2022) (Table 3). In the regulation process involving MOTS-c, the level of 5-methyltetrahydrofolate decreased and the folate, methionine, and homocysteine levels increased. A decrease in purine biosynthesis accompanies the decrease in 5-methyltetrahydrofolate level and the de novo purine synthesis pathway is negatively regulated by purine, so MOTS-c can accelerate the de novo purine synthesis. During this process, the level of 5-aminimidazole-4-formamide ribonucleotide (AICAR) is increased and then AICAR activates the adenosine 5'-monophosphate (AMP)-activated protein kinase (AMPK) pathway by phosphorylating and inducing the inactivation of acetyl-coenzyme A (CoA) carboxylase, stimulating fat oxidation, reducing the allosteric inhibition of carnitine palmitoyltransferase-1, and increasing fatty acid utilization and cellular glucose uptake (Lee et al., 2015). Meanwhile, the expression level of MOTS-c is negatively correlated with body mass index (BMI), fasting insulin, and glycated hemoglobin levels (Du et al., 2018). MiP SHLP2 can activate the phosphorylation of extracellular signal-regulated kinase (ERK) and signal transducer and activator of transcription-1 (STAT-1), promote cell proliferation, increase cell oxygen consumption rate and ATP level, enhance mitochondrial function, and maintain glucose metabolism homeostasis (Cobb et al., 2016) (Table 3). In vivo, the injection of SHLP2

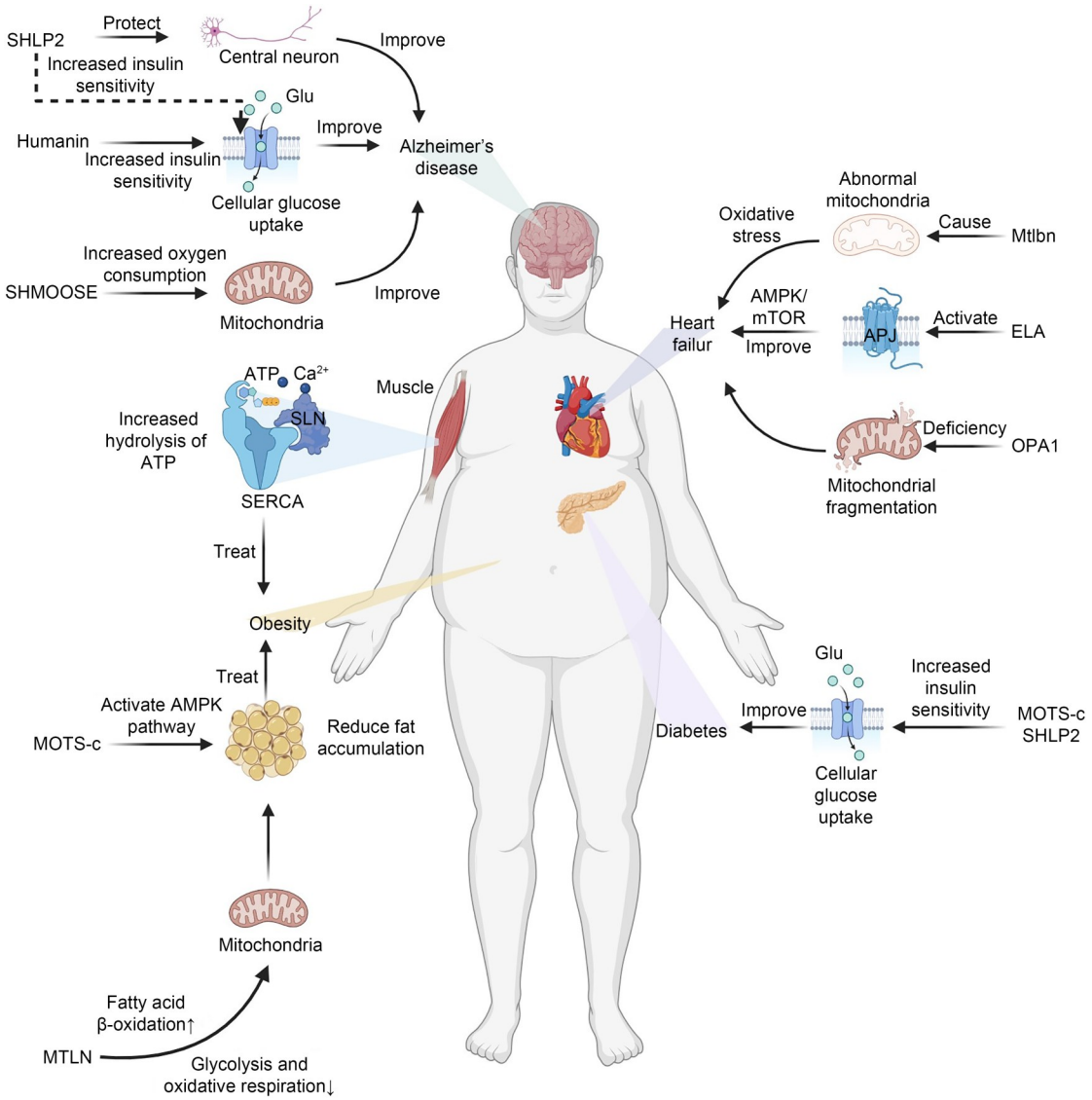

Fig. 3 Metabolic functions of micropeptides (miPs). Small humanin-like peptide 2 (SHLP2) and mitochondrial open reading frame (ORF) of the 12S ribosomal RNA (rRNA)-c (MOTS-c) can improve insulin sensitivity and glucose (Glu) uptake by cells, thus improving diabetes. Mitolamban (Mtlbn) and optic atrophy type 1 protein (OPA1) impair mitochondrial function and lead to heart failure. SHLP2, humanin, and small human mitochondrial ORF over SErine transfer RNA (SHMOOSE) facilitate insulin sensitization, protection of neurons, and enhancement of mitochondrial oxidation function to improve Alzheimer's disease (AD). Sarcolipin (SLN), MOTS-c, and mitoregulin (MTLN) can reduce fat storage by increasing adenosine triphosphate (ATP) hydrolysis and fat oxidation, thus treating obesity. SERCA: sarco/endoplasmic reticulum Ca^{2+}-ATPase; AMPK: adenosine 5'-monophosphate (AMP)-activated protein kinase; mTOR: mammalian target of rapamycin; APJ: apelin-angiotensin receptor-like; ELA: Elabela.

increases glucose uptake and inhibits hepatic glucose production (Cobb et al., 2016) (Fig. 3). HN is a mitochondrial-derived miP, which is a 24-aa peptide encoded by the 16S rRNA of mtDNA (Hashimoto et al., 2001). Potent HN analogs were able to increase glucose transporter 2 (GLUT2) translocation and glucokinase (GCK) activity, which leads to the glucose-stimulated insulin release and glucose-sensitive system in pancreatic β-cells. Moreover, HN analogs enhance islet β-cell metabolism and promote insulin secretion (Kuliawat et al., 2013). HN can also increase glucose uptake and reduce hepatic glucose production by increasing signaling molecules in the insulin-mediated phosphatidylinositol 3 kinase (PI3K)/protein kinase B (AKT) signaling pathway and the fatty acid metabolism signaling pathway (Wu et al., 2022).

Table 3 Mechanisms of action of micropeptides (miPs)

Name	Mechanisms of action	Associated diseases	References
MOTS-c	Inhibits purine nucleotide de novo synthesis and folate cycle, and activates AMPK, thus improving insulin resistance and regulating fat metabolism	Diabetes; obesity	Lee et al., 2015; Wu et al., 2022
SHLP2	Euglycemic agent; prevents neuronal cell death	Diabetes; AD	Cobb et al., 2016
Humanin	Associates with mitochondrial function and increases central insulin sensitivity	AD	Muzumdar et al., 2009; Zárate et al., 2019
SHMOOSE	Increases mitochondrial oxygen consumption and regulates brain gene expression	AD	Miller et al., 2022b
Mitoregulin	Interacts with mitochondrial trifunctional proteins to influence lipid β-oxidation and tricarboxylic acid cycle processes	Obesity	Makarewich et al., 2018; Stein et al., 2018
Sarcolipin	Promotes uncoupling of the membrane pump SERCA, leading to increased ATP hydrolysis and heat production	Obesity	Bal et al., 2012
Elabela/Apela	Regulates cellular function in heart failure by activation of APJ receptors via the AMPK/mTOR signaling pathway	Heart failure	Ma et al., 2021
Mtlbn	Causes altered redox homeostasis, abnormal mitochondrial morphology, and dysfunction	Heart failure	Makarewich et al., 2022
EP133	Leads to mitochondrial energy loss and reduced ATP production, affecting the cell cycle	Nasopharyngeal carcinoma	Kang et al., 2020
ASAP	Interacts with ATP synthase subunits α and γ, enhances ATPase activity and mitochondrial oxygen consumption rate, and promotes cancer cell proliferation	Colorectal cancer	Ge et al., 2021
HOXB-AS3	Inhibits the binding of hnRNP A1 of PKM to its exons, PKM2 formation, glucose metabolic reprogramming, and tumorigenesis	Colon cancer	Huang et al., 2017

MOTS-c: mitochondrial open reading frame (ORF) of the 12S ribosomal RNA (rRNA)-c; AMPK: adenosine 5'-monophosphate-activated protein kinase; SHLP2: small humanin-like peptide 2; AD: Alzheimer's disease; SHMOOSE: small human mitochondrial ORF over SErine transfer RNA; SERCA: sarco/endoplasmic reticulum Ca^{2+}-ATPase; mTOR: mammalian target of rapamycin; APJ: apelin-angiotensin receptor-like; Mtlbn: mitolamban; ASAP: adenosine triphosphate (ATP) synthase-associated peptide; hnRNP A1: heterogeneous nuclear ribonucleoprotein A1; PKM: pyruvate kinase M.

4.2 MiPs and Alzheimer's disease

The brain is one of the most mitochondria-rich tissues, and the pathogenesis of AD is associated with mitochondrial dysfunction (Yao et al., 2009). HN was first detected in the occipital lobe of AD patients. In the hippocampus of rats, mitochondrial dysfunction induced by ovarian hormone deprivation leads to decreased expression of HN homologous protein (HNr) (Zárate et al., 2019). HN may be a marker of mitochondrial dysfunction in AD. In the serum of humans and mice, HN levels decrease with age. HN can enhance insulin sensitivity (Fig. 3). In addition, in the central system, intravenous administration of novel potent HN derivatives enhanced central insulin sensitivity (Muzumdar et al., 2009) (Table 3). Changes in glucose metabolism are closely related to the development of AD. In the brain, if the insulin-mediated AKT signaling pathway is impaired, the phosphorylation of Ser9 is decreased and the enzymatic activity of glycogen synthase kinase-3β (GSK-3β) is increased, leading to the persistent phosphorylation of Tyr216 (tyrosine site of GSK-3β), which leads to the hyperphosphorylation of tau residues, which are also important pathological features of AD (Zhang et al., 2018). Moreover, miPs can regulate insulin secretion, sensitivity, and mitochondrial function (Muzumdar et al., 2009; Kuliawat et al., 2013; Cobb et al., 2016). These regulatory sites may be potential therapeutic targets for improving cognitive function in AD (Potenza et al., 2021). MiP SHLP2 enhances insulin sensitivity in the periphery. SHLP2 also prevents neuronal cell death caused by lactate dehydrogenase (LDH) leakage in the AD model (Cobb et al., 2016) (Fig. 3, Table 3). Small human

mitochondrial ORF over SErine tRNA (SHMOOSE), a mitochondrial miP, participates in brain energy metabolism by regulating gene expression in mitochondria. The level of the SHMOOSE in CSF is related to the pathological features of AD. Intraventricular administration of SHMOOSE regulates mitochondrial gene expression in the brain, increases mitochondrial reserve capacity during cellular stress, maximizes proton flow through the mitochondrial inner membrane, and increases mitochondrial oxygen consumption (Miller et al., 2022b) (Fig. 3, Table 3).

4.3 MiPs and obesity

MOTS-c reduces lipid accumulation in D-galactose-induced aged mice or ovariectomy-induced obese mice, which might be related to activating the AMPK pathway to improve energy expenditure and insulin sensitivity (Li et al., 2019; Lu et al., 2019). MTLN, another mitochondrial microprotein, regulates the balance of gluconeogenesis and lipid metabolism through affecting fatty acid oxidation and tricarboxylic acid cycle processes. MTLN interacts directly with the β-subunit of mitochondrial trifunctional protein, an enzyme critical in the β-oxidation of long-chain fatty acids (LCFAs) (Stein et al., 2018; Chugunova et al., 2019) (Table 3). MTLN affects the lipolytic phenotype in white adipocytes while inhibiting glycolysis and oxidative respiration processes and promoting fatty acid oxidation by regulating β-oxidation (Fig. 3) (Friesen et al., 2020). Similar to MTLN, the microprotein MOXI, encoded by *LINC00116*, regulates fatty acid metabolism in cardiac and skeletal muscle mitochondria. MOXI specifically interacts with mitochondrial trifunctional proteins to catalyze the oxidation of LCFAs. Loss of MOXI results in decreased β-oxidation of LCFA and increased carbohydrate oxidation (Makarewich et al., 2018). Sarcolipin (SLN) is an miP that regulates membrane pump sarco/endoplasmic reticulum Ca^{2+}-ATPase (SERCA) (Primeau et al., 2018). SLN and Ca^{2+} can combine with SERCA at the same time, promoting the decoupling of the SERCA pump, leading to the increase of ATP hydrolysis, heat production, and energy consumption (Fig. 3), thus becoming a potential target for obesity (Bal et al., 2012) (Table 3).

4.4 MiPs and heart failure

MiPs are involved in heart failure. Elabela (ELA, also known as Apela) plays a key role in the regulation of heart and blood vessel function (Chng et al., 2013; Chen et al., 2021). ELA has anti-apoptotic and stress-induced cell death functions and promotes cell survival and self-renewal in an apelin-angiotensin receptor-like (APJ)-independent manner via the PI3K/AKT pathway. In addition, ELA can stimulate human umbilical vein endothelial cells (HUVECs) to activate APJ receptors to inhibit angiotensin-converting enzyme expression and pathogenic angiotensin II signaling, preventing hypertension and heart failure, which has great potential for the treatment of heart failure (Sato et al., 2017). The ELA–APJ receptor axis has been shown to play an antifibrotic role in the heart by blocking transforming growth factor-β (TGF-β) signaling. The AMPK/mTOR signaling pathway regulates cellular function in heart failure (Ma et al., 2021) (Fig. 3, Table 3). MiP Mtlbn is abundantly expressed in the heart. In Mtlbn transgenic mice, Mtlbn overexpression induces cardiomyopathy that progressively progresses to congestive heart failure (Makarewich et al., 2022). Compared to wild-type mice, *Mtlbn* transgenic mice have reduced total NADH oxidase activity in heart mitochondria, significantly low protein levels of complex I, complex IV, and complex V, increased glutathione peroxidase 4 (GPx4), and a reduced ratio of glutathione (GSH) to oxidized glutathione (GSSG). The hearts of *Mtlbn* transgenic mice likely experience high levels of oxidative stress and thus consume GSH to maintain redox homeostasis. Morphological abnormalities, dysfunction, and altered redox homeostasis of mitochondria caused by overexpression of the microprotein Mtlbn contribute to the development of heart failure in transgenic mice (Makarewich et al., 2022) (Fig. 3, Table 3). Mitochondrial microprotein optic atrophy type 1 protein (OPA1) deficiency leads to mitochondrial fragmentation, which reduces the activity of respiratory chain complexes and the depletion of essential mtDNA (Wai et al., 2015; Tezze et al., 2017) (Fig. 3). MiPs regulate mitochondrial function and protect cardiometabolic function, which might have important clinical significance for treating heart failure.

4.5 MiPs and tumors

MiP-EP133 is a tumor suppressor miP encoded by miR-34a (a tumor suppressor), which affects the cell cycle of tumor cells and regulates apoptosis by influencing mitochondrial energy metabolic processes

(Kang et al., 2020). Overexpression of miP-EP133 induces cancer cell apoptosis, inhibits their migration and invasion, and suppresses tumor growth in vivo (Table 3). When miP-EP133 expression is decreased, a poor prognosis for advanced metastatic nasopharyngeal carcinoma might be predicted. MiP-EP133 prevents heat shock protein family A member 9 (HSPA9) from interacting with HSP60, translocase inner mitochondrial membrane 44 (TIM44), and voltage-dependent anion channel 1 (VDAC1) by binding to HSPA9 (Fig. 4a) (Kang et al., 2020). The 94-aa miP ASAP encoded by lncRNA *LINC00467* has a regulatory role in colorectal cancer. ASAP is involved in mitochondrial metabolism and ATP production. ASAP enhances ATP synthase activity and mitochondrial oxygen consumption rate by interacting with ATP synthase subunits α and γ (ATP5A and ATP5C), thereby promoting colorectal cancer cell proliferation (Fig. 4b). Colorectal cancer patients with high ASAP expression have poor prognosis (Ge et al., 2021) (Table 3). HOXB-AS3 encoded by lncRNA *HOXB-AS3* is associated with multiple cancers (Wang JZ et al., 2019). HOXB-AS3 can inhibit tumor cell growth, migration, invasion, and tumor growth both in vitro and in vivo. HOXB-AS3 increased the formation of pyruvate kinase M1 (PKM1) isoform and decreased the formation of the PKM2 isoform by binding to arginine residues in the RNA-binding RGG box of the splicing factor heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) and then binding to both sides of the exon sequence of *PKM* (Fig. 4c). The expression of PKM2 endows the selective proliferation advantage of tumor cells in vivo. Inhibition

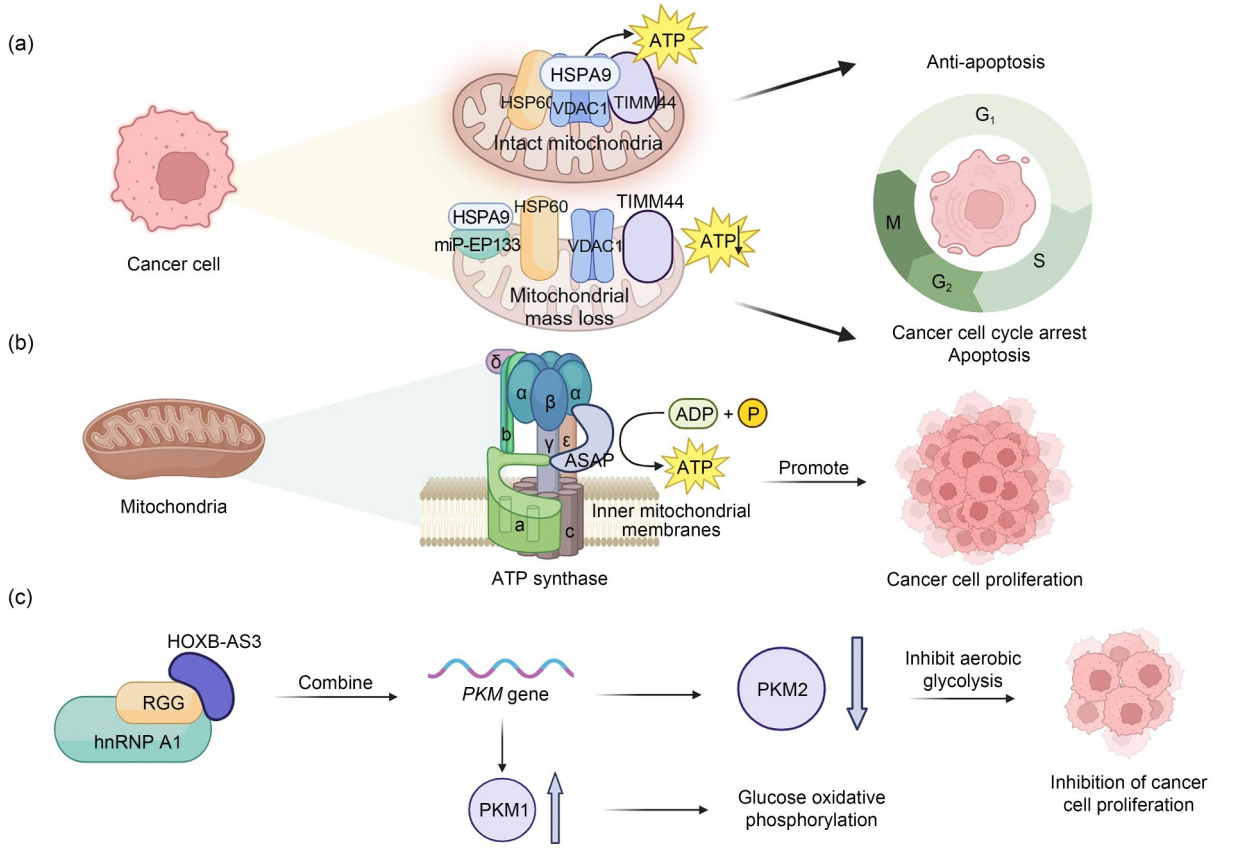

Fig. 4 Micropeptides (miPs) in tumor progression. (a) The combination of miP-EP133 with heat shock protein family A member 9 (HSPA9) resulted in the failure of HSPA9 to combine with HSP60, voltage-dependent anion channel 1 (VDAC1), and translocase of inner mitochondrial membrane 44 (TIMM44), which reduced mitochondrial mass and adenosine triphosphate (ATP) production and then led to cell cycle arrest and apoptosis of cancer cells. (b) MiP ATP synthase-associated peptide (ASAP) binds the α and γ subunits of ATP synthase, which increases the production of ATP and accelerates the proliferation of cancer cells. (c) MiP HOXB-AS3 binds to RGG of heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1). It acts on the pyruvate kinase M (*PKM*) gene, resulting in PKM2 downregulation and PKM1 upregulation. This transformation shifts glucose metabolism from aerobic glycolysis to glucose oxidative phosphorylation, thus inhibiting the proliferation of cancer cells. ADP: adenosine diphosphate.

of PKM2 expression can inhibit the reprogramming of glucose metabolism and the occurrence of tumors. In colon cancer patients, the low level of HOXB-AS3 expression may be a sign of poor prognosis (Huang et al., 2017) (Table 3).

5 Future outlook

An increasing number of studies on miPs have demonstrated that miPs are closely related to mitochondrial function and may be critical regulatory targets for treating diabetes, obesity, heart failure, AD, tumors, and other highly prevalent diseases. Although with the development of technology there are many assays for miPs, the current miP research still faces difficulties (Tharakan and Sawa, 2021). Regarding new miP exploration, some methods cannot identify miPs that are not annotated, which hinders the discovery and exploration of new miPs. However, it has been shown that some miPs are expressed from lncRNAs that have regulatory functions (van Heesch et al., 2019). Intron transcripts (ITs) in protein-coding genes that form lncRNAs may encode miPs, and ITs are closely related to the occurrence and development of diseases (Rothzerg et al., 2022). MiPs are inextricably linked to mitochondria, which suggests that the enrichment of miPs in mitochondria is a low-energy strategy that mitochondria have evolved. Mitochondria prefer to transport proteins with small molecular weights. This strategy improves transport efficiency and reduces mitochondrial energy consumption, which can maintain the homeostasis of cellular energy balance (van Heesch et al., 2019). In addition, in miPs-related diseases, many miPs target and regulate the function of mitochondria, for example, affecting the production and hydrolysis of ATP and participating in the processes of glucose metabolism and lipid metabolism. Thus, further exploration of miPs might be a new direction for the study of diseases related to energy metabolism.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 81870237). All the figures were created with Biorender.com.

Author contributions

Yirui LU and Yutong RAN were responsible for writing – original draft. Hong LI and Jiao WEN were responsible for supervision. Xiaodong CUI and Xiaoyun ZHANG were responsible for conceptualization. Xiumei GUAN was responsible for writing – review & editing. Min CHENG was responsible for supervision, writing – review & editing, and funding acquisition. All authors have read and approved the final manuscript.

Compliance with ethics guidelines

Yirui LU, Yutong RAN, Hong LI, Jiao WEN, Xiaodong CUI, Xiaoyun ZHANG, Xiumei GUAN, and Min CHENG declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Adams BD, Parsons C, Walker L, et al., 2017. Targeting noncoding RNAs in disease. *J Clin Invest*, 127(3):761-771. https://doi.org/10.1172/JCI84424

Akimoto C, Sakashita E, Kasashima K, et al., 2013. Translational repression of the *McKusick-Kaufman syndrome* transcript by unique upstream open reading frames encoding mitochondrial proteins with alternative polyadenylation sites. *Biochim Biophys Acta*, 1830(3):2728-2738. https://doi.org/10.1016/j.bbagen.2012.12.010

Anderson DM, Anderson KM, Chang CL, et al., 2015. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell*, 160(4):595-606. https://doi.org/10.1016/j.cell.2015.01.009

Andjus S, Morillon A, Wery M, 2021. From yeast to mammals, the nonsense-mediated mRNA decay as a master regulator of long non-coding RNAs functional trajectory. *Non-Coding RNA*, 7(3):44. https://doi.org/10.3390/ncrna7030044

Aspden JL, Eyre-Walker YC, Phillips RJ, et al., 2014. Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq. *Elife*, 3:e03528. https://doi.org/10.7554/eLife.03528

Bal NC, Maurya SK, Sopariwala DH, et al., 2012. Sarcolipin is a newly identified regulator of muscle-based thermogenesis in mammals. *Nat Med*, 18(10):1575-1579. https://doi.org/10.1038/nm.2897

Bazzini AA, Johnstone TG, Christiano R, et al., 2014. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *EMBO J*, 33(9): 981-993. https://doi.org/10.1002/embj.201488411

Bhati KK, Blaakmeer A, Paredes EB, et al., 2018. Approaches to identify and characterize microproteins and their potential uses in biotechnology. *Cell Mol Life Sci*, 75(14):2529-2536. https://doi.org/10.1007/s00018-018-2818-8

Bhatta A, Atianand M, Jiang ZZ, et al., 2020. A mitochondrial micropeptide is required for activation of the Nlrp3 inflammasome. *J Immunol*, 204(2):428-437. https://doi.org/10.4049/jimmunol.1900791

Boguszewska K, Szewczuk M, Kaźmierczak-Barańska J, et al., 2020. The similarities between human mitochondria and

bacteria in the context of structure, genome, and base excision repair system. *Molecules*, 25(12):2857. https://doi.org/10.3390/molecules25122857

Branca RMM, Orre LM, Johansson HJ, et al., 2014. HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics. *Nat Methods*, 11(1):59-62. https://doi.org/10.1038/nmeth.2732

Cao XW, Khitun A, Na Z, et al., 2020. Comparative proteomic profiling of unannotated microproteins and alternative proteins in human cell lines. *J Proteome Res*, 19(8):3418-3426. https://doi.org/10.1021/acs.jproteome.0c00254

Cataldo LR, Fernández-Verdejo R, Santos JL, et al., 2018. Plasma MOTS-c levels are associated with insulin sensitivity in lean but not in obese individuals. *J Investig Med*, 66(6):1019-1022. https://doi.org/10.1136/jim-2017-000681

Chen J, Brunner AD, Cogan JZ, et al., 2020. Pervasive functional translation of noncanonical human open reading frames. *Science*, 367(6482):1140-1146. https://doi.org/10.1126/science.aay0262

Chen XP, Han P, Zhou T, et al., 2016. CircRNADb: a comprehensive database for human circular RNAs with protein-coding annotations. *Sci Rep*, 6:34985. https://doi.org/10.1038/srep34985

Chen Y, Ho L, Tergaonkar V, 2021. sORF-Encoded MicroPeptides: new players in inflammation, metabolism, and precision medicine. *Cancer Lett*, 500:263-270. https://doi.org/10.1016/j.canlet.2020.10.038

Chng SC, Ho L, Tian J, et al., 2013. ELABELA: a hormone essential for heart development signals via the apelin receptor. *Dev Cell*, 27(6):672-680. https://doi.org/10.1016/j.devcel.2013.11.002

Chothani SP, Adami E, Widjaja AA, et al., 2022. A high-resolution map of human RNA translation. *Mol Cell*, 82(15):2885-2899.e8. https://doi.org/10.1016/j.molcel.2022.06.023

Chugunova A, Loseva E, Mazin P, et al., 2019. *LINC00116* codes for a mitochondrial peptide linking respiration and lipid metabolism. *Proc Natl Acad Sci USA*, 116(11):4940-4945. https://doi.org/10.1073/pnas.1809105116

Cobb LJ, Lee C, Xiao JL, et al., 2016. Naturally occurring mitochondrial-derived peptides are age-dependent regulators of apoptosis, insulin sensitivity, and inflammatory markers. *Aging*, 8(4):796-809. https://doi.org/10.18632/aging.100943

Dangelmaier EA, Li XL, Hartford CCR, et al., 2022. An evolutionarily conserved AU-rich element in the 3' untranslated region of a transcript misannotated as a long noncoding RNA regulates RNA stability. *Mol Cell Biol*, 42(4):e00505-21. https://doi.org/10.1128/mcb.00505-21

Dragomir MP, Manyam GC, Ott LF, et al., 2020. FuncPEP: a database of functional peptides encoded by non-coding RNAs. *Non-Coding RNA*, 6(4):41. https://doi.org/10.3390/ncrna6040041

Du CQ, Zhang C, Wu W, et al., 2018. Circulating MOTS-c levels are decreased in obese male children and adolescents and associated with insulin resistance. *Pediatr Diabetes*, 19(6):1058-1064. https://doi.org/10.1111/pedi.12685

The ENCODE Project Consortium, 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447(7146):799-816. https://doi.org/10.1038/nature05874

Fabre B, Combier JP, Plaza S, 2021. Recent advances in mass spectrometry-based peptidomics workflows to identify short-open-reading-frame-encoded peptides and explore their functions. *Curr Opin Chem Biol*, 60:122-130. https://doi.org/10.1016/j.cbpa.2020.12.002

Ference BA, Ginsberg HN, Graham I, et al., 2017. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur Heart J*, 38(32):2459-2472. https://doi.org/10.1093/eurheartj/ehx144

Friesen M, Warren CR, Yu HJ, et al., 2020. Mitoregulin controls β-oxidation in human and mouse adipocytes. *Stem Cell Rep*, 14(4):590-602. https://doi.org/10.1016/j.stemcr.2020.03.002

Fung G, Shi J, Deng H, et al., 2015. Cytoplasmic translocation, aggregation, and cleavage of TDP-43 by enteroviral proteases modulate viral pathogenesis. *Cell Death Differ*, 22(12):2087-2097. https://doi.org/10.1038/cdd.2015.58

Gammage PA, Moraes CT, Minczuk M, 2018. Mitochondrial genome engineering: the revolution may not be CRISPR-ized. *Trends Genet*, 34(2):101-110. https://doi.org/10.1016/j.tig.2017.11.001

Ge QW, Jia DJC, Cen D, et al., 2021. Micropeptide ASAP encoded by LINC00467 promotes colorectal cancer progression by directly modulating ATP synthase activity. *J Clin Invest*, 131(22):e152911. https://doi.org/10.1172/JCI152911

Guo BB, Wu SQ, Zhu X, et al., 2020. Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression. *EMBO J*, 39(1):e102190. https://doi.org/10.15252/embj.2019102190

Gustafsson CM, Falkenberg M, Larsson NG, 2016. Maintenance and expression of mammalian mitochondrial DNA. *Annu Rev Biochem*, 85:133-160. https://doi.org/10.1146/annurev-biochem-060815-014402

Hartford CCR, Lal A, 2020. When long noncoding becomes protein coding. *Mol Cell Biol*, 40(6):e00528-19. https://doi.org/10.1128/MCB.00528-19

Hashimoto Y, Niikura T, Tajima H, et al., 2001. A rescue factor abolishing neuronal cell death by a wide spectrum of familial Alzheimer's disease genes and Aβ. *Proc Natl Acad Sci USA*, 98(11):6336-6341. https://doi.org/10.1073/pnas.101133498

Huang JZ, Chen M, Chen D, et al., 2017. A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth. *Mol Cell*, 68(1):171-184.e6. https://doi.org/10.1016/j.molcel.2017.09.015

Hussain SRA, Yalvac ME, Khoo B, et al., 2021. Adapting CRISPR/Cas9 system for targeting mitochondrial genome. *Front Genet*, 12:627050. https://doi.org/10.3389/fgene.2021.627050

Ingolia NT, Ghaemmaghami S, Newman JRS, et al., 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, 324(5924):218-223. https://doi.org/10.1126/science.1168978

Ingolia NT, Brar GA, Stern-Ginossar N, et al., 2014. Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep*, 8(5):1365-1379. https://doi.org/10.1016/j.celrep.2014.07.045

Ji Z, Song RS, Regev A, et al., 2015. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *Elife*, 4:e08890. https://doi.org/10.7554/eLife.08890

Johnstone TG, Bazzini AA, Giraldez AJ, 2016. Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J*, 35(7):706-723. https://doi.org/10.15252/embj.201592759

Kamps R, Szklarczyk R, Theunissen TE, et al., 2018. Genetic defects in mtDNA-encoded protein translation cause pediatric, mitochondrial cardiomyopathy with early-onset brain disease. *Eur J Hum Genet*, 26(4):537-551. https://doi.org/10.1038/s41431-017-0058-2

Kang M, Tang B, Li JX, et al., 2020. Identification of miPEP133 as a novel tumor-suppressor microprotein encoded by miR-34a pri-miRNA. *Mol Cancer*, 19:143. https://doi.org/10.1186/s12943-020-01248-9

Kapusta A, Kronenberg Z, Lynch VJ, et al., 2013. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet*, 9(4):e1003470. https://doi.org/10.1371/journal.pgen.1003470

Kearse MG, Wilusz JE, 2017. Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev*, 31(17):1717-1731. https://doi.org/10.1101/gad.305250.117

Khitun A, Slavoff SA, 2019. Proteomic detection and validation of translated small open reading frames. *Curr Protoc Chem Biol*, 11(4):e77. https://doi.org/10.1002/cpch.77

Khitun A, Ness TJ, Slavoff SA, 2019. Small open reading frames and cellular stress responses. *Mol Omics*, 15(2):108-116. https://doi.org/10.1039/c8mo00283e

Kuliawat R, Klein L, Gong ZW, et al., 2013. Potent humanin analog increases glucose-stimulated insulin secretion through enhanced metabolism in the β cell. *FASEB J*, 27(12):4890-4898. https://doi.org/10.1096/fj.13-231092

Kustatscher G, Grabowski P, Schrader TA, et al., 2019. Co-regulation map of the human proteome enables identification of protein functions. *Nat Biotechnol*, 37(11):1361-1371. https://doi.org/10.1038/s41587-019-0298-5

Kuwahara K, 2021. The natriuretic peptide system in heart failure: diagnostic and therapeutic implications. *Pharmacol Ther*, 227:107863. https://doi.org/10.1016/j.pharmthera.2021.107863

Lauressergues D, Couzigou JM, Clemente HS, et al., 2015. Primary transcripts of microRNAs encode regulatory peptides. *Nature*, 520(7545):90-93. https://doi.org/10.1038/nature14346

Lee C, Zeng J, Drew BG, et al., 2015. The mitochondrial-derived peptide MOTS-c promotes metabolic homeostasis and reduces obesity and insulin resistance. *Cell Metab*, 21(3):443-454. https://doi.org/10.1016/j.cmet.2015.02.009

Li QY, Lu HY, Hu GY, et al., 2019. Earlier changes in mice after D-galactose treatment were improved by mitochondria derived small peptide MOTS-c. *Biochem Biophys Res Commun*, 513(2):439-445. https://doi.org/10.1016/j.bbrc.2019.03.194

Liang S, Bellato HM, Lorent J, et al., 2018. Polysome-profiling in small tissue samples. *Nucleic Acids Res*, 46(1):e3. https://doi.org/10.1093/nar/gkx940

Lu HY, Wei M, Zhai Y, et al., 2019. MOTS-c peptide regulates adipose homeostasis to prevent ovariectomy-induced metabolic dysfunction. *J Mol Med*, 97(4):473-485. https://doi.org/10.1007/s00109-018-01738-w

Lubec G, Afjehi-Sadat L, 2007. Limitations and pitfalls in protein identification by mass spectrometry. *Chem Rev*, 107(8):3568-3584. https://doi.org/10.1021/cr068213f

Ma J, Diedrich JK, Jungreis I, et al., 2016. Improved identification and analysis of small open reading frame encoded polypeptides. *Anal Chem*, 88(7):3967-3975. https://doi.org/10.1021/acs.analchem.6b00191

Ma Z, Song JJ, Martin S, et al., 2021. The Elabela-APJ axis: a promising therapeutic target for heart failure. *Heart Fail Rev*, 26(5):1249-1258. https://doi.org/10.1007/s10741-020-09957-5

Mackowiak SD, Zauber H, Bielow C, et al., 2015. Extensive identification and analysis of conserved small ORFs in animals. *Genome Biol*, 16:179. https://doi.org/10.1186/s13059-015-0742-x

Magny EG, Pueyo JI, Pearl FMG, et al., 2013. Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. *Science*, 341(6150):1116-1120. https://doi.org/10.1126/science.1238802

Makarewich CA, Olson EN, 2017. Mining for micropeptides. *Trends Cell Biol*, 27(9):685-696. https://doi.org/10.1016/j.tcb.2017.04.006

Makarewich CA, Baskin KK, Munir AZ, et al., 2018. MOXI is a mitochondrial micropeptide that enhances fatty acid β-oxidation. *Cell Rep*, 23(13):3701-3709. https://doi.org/10.1016/j.celrep.2018.05.058

Makarewich CA, Munir AZ, Bezprozvannaya S, et al., 2022. The cardiac-enriched microprotein mitolamban regulates mitochondrial respiratory complex assembly and function in mice. *Proc Natl Acad Sci USA*, 119(6):e2120476119. https://doi.org/10.1073/pnas.2120476119

Masvidal L, Hulea L, Furic L, et al., 2017. mTOR-sensitive translation: cleared fog reveals more trees. *RNA Biol*, 14(10):1299-1305.

https://doi.org/10.1080/15476286.2017.1290041

Matsumoto A, Pasut A, Matsumoto M, et al., 2017. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature*, 541(7636):228-232. https://doi.org/10.1038/nature21034

Merry TL, Chan A, Woodhead JST, et al., 2020. Mitochondrial-derived peptides in energy metabolism. *Am J Physiol Endocrinol Metab*, 319(4):E659-E666. https://doi.org/10.1152/ajpendo.00249.2020

Miller B, Kim SJ, Kumagai H, et al., 2022a. Mitochondria-derived peptides in aging and healthspan. *J Clin Invest*, 132(9):e158449. https://doi.org/10.1172/JCI158449

Miller B, Kim SJ, Mehta HH, et al., 2022b. Mitochondrial DNA variation in Alzheimer's disease reveals a unique microprotein called SHMOOSE. *Mol Psychiatry*, 28(4):1813-1826. https://doi.org/10.1038/s41380-022-01769-3

Muzumdar RH, Huffman DM, Atzmon G, et al., 2009. Humanin: a novel central regulator of peripheral insulin action. *PLoS ONE*, 4(7):e6334. https://doi.org/10.1371/journal.pone.0006334

Nishikimi T, Nakagawa Y, Minamino N, et al., 2015. Pro-B-type natriuretic peptide is cleaved intracellularly: impact of distance between *O*-glycosylation and cleavage sites. *Am J Physiol Regul Integr Comp Physiol*, 309(6):R639-R649. https://doi.org/10.1152/ajpregu.00074.2015

Niu LM, Lou FZ, Sun Y, et al., 2020. A micropeptide encoded by lncRNA MIR155HG suppresses autoimmune inflammation via modulating antigen presentation. *Sci Adv*, 6(21): eaaz2059. https://doi.org/10.1126/sciadv.aaz2059

Ouspenskaia T, Law T, Clauser KR, et al., 2022. Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer. *Nat Biotechnol*, 40(2):209-217. https://doi.org/10.1038/s41587-021-01021-3

Pan JF, Wang RJ, Shang FZ, et al., 2022. Functional micropeptides encoded by long non-coding RNAs: a comprehensive review. *Front Mol Biosci*, 9:817517. https://doi.org/10.3389/fmolb.2022.817517

Patraquim P, Mumtaz MAS, Pueyo JI, et al., 2020. Developmental regulation of canonical and small ORF translation from mRNAs. *Genome Biol*, 21:128. https://doi.org/10.1186/s13059-020-02011-5

Pauli A, Norris ML, Valen E, et al., 2014. Toddler: an embryonic signal that promotes cell movement via Apelin receptors. *Science*, 343(6172):1248636. https://doi.org/10.1126/science.1248636

Polenkowski M, Burbano de Lara S, Allister AB, et al., 2021. Identification of novel micropeptides derived from hepatocellular carcinoma-specific long noncoding RNA. *Int J Mol Sci*, 23(1):58. https://doi.org/10.3390/ijms23010058

Potenza MA, Sgarra L, Desantis V, et al., 2021. Diabetes and Alzheimer's disease: might mitochondrial dysfunction help deciphering the common path? *Antioxidants*, 10(8):1257. https://doi.org/10.3390/antiox10081257

Prats AC, David F, Diallo LH, et al., 2020. Circular RNA, the key for translation. *Int J Mol Sci*, 21(22):8591. https://doi.org/10.3390/ijms21228591

Primeau JO, Armanious GP, Fisher ME, et al., 2018. The Sarco-Endoplasmic reticulum calcium ATPase. *In*: Harris JR, Boekema EJ (Eds.), Membrane Protein Complexes: Structure and Function. Springer, Singapore, p.229-258. https://doi.org/10.1007/978-981-10-7757-9_8

Qu L, He XY, Tang Q, et al., 2022. Iron metabolism, ferroptosis, and lncRNA in cancer: knowns and unknowns. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 23(10):844-862. https://doi.org/10.1631/jzus.B2200194

Ramzy A, Kieffer TJ, 2022. Altered islet prohormone processing: a cause or consequence of diabetes? *Physiol Rev*, 102(1):155-208. https://doi.org/10.1152/physrev.00008.2021

Ran FA, Hsu PD, Wright J, et al., 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*, 8(11):2281-2308. https://doi.org/10.1038/nprot.2013.143

Rathore A, Chu Q, Tan D, et al., 2018. MIEF1 microprotein regulates mitochondrial translation. *Biochemistry*, 57(38): 5564-5575. https://doi.org/10.1021/acs.biochem.8b00726

Rinn JL, Chang HY, 2012. Genome regulation by long noncoding RNAs. *Annu Rev Biochem*, 81:145-166. https://doi.org/10.1146/annurev-biochem-051410-092902

Rothzerg E, Xu JK, Wood D, 2022. Identification of differentially expressed intronic transcripts in osteosarcoma. *Non-Coding RNA*, 8(6):73. https://doi.org/10.3390/ncrna8060073

Ruiz-Orera J, Albà MM, 2019. Translation of small open reading frames: roles in regulation and evolutionary innovation. *Trends Genet*, 35(3):186-198. https://doi.org/10.1016/j.tig.2018.12.003

Sato T, Sato C, Kadowaki A, et al., 2017. ELABELA-APJ axis protects from pressure overload heart failure and angiotensin II-induced cardiac damage. *Cardiovasc Res*, 113(7): 760-769. https://doi.org/10.1093/cvr/cvx061

Schindewolf C, Braun S, Domdey H, 1996. *In vitro* generation of a circular exon from a linear pre-mRNA transcript. *Nucleic Acids Res*, 24(7):1260-1266. https://doi.org/10.1093/nar/24.7.1260

Sinha T, Panigrahi C, Das D, et al., 2022. Circular RNA translation, a path to hidden proteome. *Wiley Interdiscip Rev RNA*, 13(1):e1685. https://doi.org/10.1002/wrna.1685

Starck SR, Tsai JC, Chen KL, et al., 2016. Translation from the 5' untranslated region shapes the integrated stress response. *Science*, 351(6272):aad3867. https://doi.org/10.1126/science.aad3867

Staudt AC, Wenkel S, 2011. Regulation of protein function by 'microProteins'. *EMBO Rep*, 12(1):35-42. https://doi.org/10.1038/embor.2010.196

Stein CS, Jadiya P, Zhang XM, et al., 2018. Mitoregulin: a lncRNA-encoded microprotein that supports mitochondrial

supercomplexes and respiratory efficiency. *Cell Rep*, 23(13): 3710-3720.e8. https://doi.org/10.1016/j.celrep.2018.06.002

Tezze C, Romanello V, Desbats MA, et al., 2017. Age-associated loss of OPA1 in muscle impacts muscle mass, metabolic homeostasis, systemic inflammation, and epithelial senescence. *Cell Metab*, 25(6):1374-1389.e6. https://doi.org/10.1016/j.cmet.2017.04.021

Tharakan R, Sawa A, 2021. Minireview: novel micropeptide discovery by proteomics and deep sequencing methods. *Front Genet*, 12:651485. https://doi.org/10.3389/fgene.2021.651485

Tumminia A, Vinciguerra F, Parisi M, et al., 2018. Type 2 diabetes mellitus and Alzheimer's disease: role of insulin signalling and therapeutic implications. *Int J Mol Sci*, 19(11):3306. https://doi.org/10.3390/ijms19113306

Ulitsky I, Bartel DP, 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell*, 154(1):26-46. https://doi.org/10.1016/j.cell.2013.06.020

van Heesch S, Witte F, Schneider-Lunitz V, et al., 2019. The translational landscape of the human heart. *Cell*, 178(1): 242-260.e29. https://doi.org/10.1016/j.cell.2019.05.010

Vitorino R, Guedes S, Amado F, et al., 2021. The role of micropeptides in biology. *Cell Mol Life Sci*, 78(7):3285-3298. https://doi.org/10.1007/s00018-020-03740-3

Vizioli MG, Liu TH, Miller KN, et al., 2020. Mitochondria-to-nucleus retrograde signaling drives formation of cytoplasmic chromatin and inflammation in senescence. *Genes Dev*, 34(5-6):428-445. https://doi.org/10.1101/gad.331272.119

Wai T, García-Prieto J, Baker MJ, et al., 2015. Imbalanced OPA1 processing and mitochondrial fragmentation cause heart failure in mice. *Science*, 350(6265):aad0116. https://doi.org/10.1126/science.aad0116

Wang JZ, Zhu S, Meng N, et al., 2019. ncRNA-encoded peptides or proteins and cancer. *Mol Ther*, 27(10):1718-1725. https://doi.org/10.1016/j.ymthe.2019.09.001

Wang S, Mao CB, Liu SR, 2019. Peptides encoded by non-coding genes: challenges and perspectives. *Signal Transduct Target Ther*, 4:57. https://doi.org/10.1038/s41392-019-0092-3

Wilson BA, Masel J, 2011. Putatively noncoding transcripts show extensive association with ribosomes. *Genome Biol Evol*, 3:1245-1252. https://doi.org/10.1093/gbe/evr099

Winter J, Jung S, Keller S, et al., 2009. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol*, 11(3):228-234. https://doi.org/10.1038/ncb0309-228

Wong J, Zhang JC, Yanagawa B, et al., 2012. Cleavage of serum response factor mediated by enteroviral protease 2A contributes to impaired cardiac function. *Cell Res*, 22(2): 360-371. https://doi.org/10.1038/cr.2011.114

Wu WY, Ji PF, Zhao FQ, 2020. CircAtlas: an integrated resource of one million highly accurate circular RNAs from 1070 vertebrate transcriptomes. *Genome Biol*, 21:101. https://doi.org/10.1186/s13059-020-02018-y

Wu Y, Sun LK, Zhuang ZD, et al., 2022. Mitochondrial-derived peptides in diabetes and its complications. *Front Endocrinol*, 12:808120. https://doi.org/10.3389/fendo.2021.808120

Xu K, Jin XY, Luo Y, et al., 2023. Spatial transcriptome analysis of long non-coding RNAs reveals tissue specificity and functional roles in cancer. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 24(1):15-31. https://doi.org/10.1631/jzus.B2200206

Xu YT, Zhang L, Ocansey DKW, et al., 2022. HucMSC-Ex alleviates inflammatory bowel disease via the lnc78583-mediated miR3202/HOXB13 pathway. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 23(5):423-431. https://doi.org/10.1631/jzus.B2100793

Yang JE, Zhong WJ, Li JF, et al., 2023. LINC00998-encoded micropeptide SMIM30 promotes the G1/S transition of cell cycle by regulating cytosolic calcium level. *Mol Oncol*, 17(5):901-916. https://doi.org/10.1002/1878-0261.13358

Yang YB, Gao XY, Zhang ML, et al., 2018. Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis. *J Natl Cancer Inst*, 110(3):304-315. https://doi.org/10.1093/jnci/djx166

Yao J, Irwin RW, Zhao LQ, et al., 2009. Mitochondrial bioenergetic deficit precedes Alzheimer's pathology in female mouse model of Alzheimer's disease. *Proc Natl Acad Sci USA*, 106(34):14670-14675. https://doi.org/10.1073/pnas.0903563106

Zárate SC, Traetta ME, Codagnone MG, et al., 2019. Humanin, a mitochondrial-derived peptide released by astrocytes, prevents synapse loss in hippocampal neurons. *Front Aging Neurosci*, 11:123. https://doi.org/10.3389/fnagi.2019.00123

Zhang S, Reljić B, Liang C, et al., 2020. Mitochondrial peptide BRAWNIN is essential for vertebrate respiratory complex III assembly. *Nat Commun*, 11:1312. https://doi.org/10.1038/s41467-020-14999-2

Zhang Y, Huang NQ, Yan F, et al., 2018. Diabetes mellitus and Alzheimer's disease: GSK-3β as a potential link. *Behav Brain Res*, 339:57-65. https://doi.org/10.1016/j.bbr.2017.11.015

Zhu LF, Xu L, Wang CG, et al., 2021. T6SS translocates a micropeptide to suppress STING-mediated innate immunity by sequestering manganese. *Proc Natl Acad Sci USA*, 118(42):e2103526118. https://doi.org/10.1073/pnas.2103526118