Non-coding RNA Research 10 (2025) 231–241

Contents lists available at ScienceDirect

Non-coding RNA Research

journal homepage: www.keaipublishing.com/en/journals/non-coding-rna-research

Review Article

Emerging role of endogenous peptides encoded by non-coding RNAs in cancer biology

Anna Lucia Tornesello [a,*], Andrea Cerasuolo [b], Noemy Starita [b], Sara Amiranda [b], Tiziana Pecchillo Cimmino [b], Patrizia Bonelli [b], Franca Maria Tuccillo [b], Franco Maria Buonaguro [b], Luigi Buonaguro [a], Maria Lina Tornesello [b,**]

[a] Innovative Immunological Models Unit, Istituto Nazionale Tumori IRCCS Fondazione G. Pascale, Napoli, Italy
[b] Molecular Biology and Viral Oncology Unit, Istituto Nazionale Tumori IRCCS Fondazione G. Pascale, Napoli, Italy

ARTICLE INFO

Keywords:
Non-coding RNAs
Peptide
Micropeptide
Microprotein
Cancer
Biomarkers
Therapeutic targets

ABSTRACT

Non-coding RNAs have long been recognized for their regulatory roles in various cellular processes, including cancer development and progression. Recent advancements have shed light on a novel aspect of non-coding RNA biology, revealing their ability to encode endogenous peptides also named micropeptides or microprotein through short open reading frames (sORFs). These small proteins play crucial roles in oncogenic processes, acting as either tumour suppressors or tumour promoters, and hold enormous potential as biomarkers for early diagnosis of cancer and as therapeutic targets. This comprehensive review highlights the state of the art on peptides encoded by long non-coding RNAs (lncRNAs), microRNAs (miRNAs), and circular RNAs (circRNAs), elucidating their regulatory functions and implications in different cancer types, including breast cancer, hepatocellular carcinoma and colorectal cancer. The review also discusses challenges and future directions in the exploration of these emerging players in cancer biology, emphasizing the importance of further investigation for their clinical translation in diagnosis and therapy.

1. Introduction

The human genome contains a large number of short open reading frames (sORFs), usually smaller than 100 codons that encode a distinct class of peptides named micropeptides or microproteins [1–5]. Unlike conventional peptides, which frequently arise from post-translational cleavage of longer polypeptides, micropeptides are directly translated from their messenger RNA (mRNA) as mature functional proteins without undergoing proteolytic digestion [6]. Recent studies based on genomic and proteomic analyses identified thousands of microproteins derived from mRNAs lacking the typical features of canonical ORFs, such as the start codon (AUG) and stop codons (UAA, UAG, or UGA). Indeed, the non-canonical ORFs initiate the translation from non-AUG codons or through alternative mechanisms and may encode micropeptides characterized by a relatively small size and key biological functions that have been previously overlooked [7]. Micropeptides typically lack N-terminal signalling sequences and are mainly localized within the cytoplasm [7]. However, emerging studies have shown that micropeptides can also be found in other cellular compartments, such as the nucleus, where they may participate in the transcriptional regulation or chromatin remodelling processes, or in the membrane of cellular organelles such as mitochondria, suggesting a potential role in energy metabolism and signalling pathways [8–10]. Recent studies have identified diverse forms of non-canonical ORFs encoding peptides, which derive from the complex family of non-coding RNAs including long non-coding RNA (lncRNA), microRNA (miRNA) and circular RNA (circRNA) (Fig. 1), [11–13]. Similarly to the conventional protein coding mRNAs, the structural basis of most lncRNAs contains a 5′-m7GpppN cap structure and a polycistronic A tail, which are essential for the translation of lncRNAs [13]. Also miRNAs, which are usually produced from their precursor pri-miRNAs, are subjected to processes of cap addition and polyadenylation [14]. On the other hand, circRNA is a special group of non-coding RNAs originating from linear precursor RNAs undergoing unconventional splicing processes that allow exons to form circles, whose length ranges from less than 100 nucleotides to several thousands. CircRNAs lack the 5′-cap structure, which is essential

* Corresponding author. Innovative Immunological Models Unit Istituto Nazionale Tumori "Fond. G. Pascale", Via Mariano Semmola 80131, Naples, Italy.
** Corresponding author. Molecular Biology and Viral Oncology Unit, Istituto Nazionale Tumori "Fond. G. Pascale", Via Mariano Semmola, 80131, Naples, Italy.
E-mail addresses: a.tornesello@istitutotumori.na.it (A.L. Tornesello), m.tornesello@istitutotumori.na.it (M.L. Tornesello).

https://doi.org/10.1016/j.ncrna.2024.10.006
Received 9 August 2024; Received in revised form 30 September 2024; Accepted 27 October 2024
Available online 28 October 2024

for the recruitment of the ribosomal complex near the 5′ end of coding mRNAs and protein translation. However, there is a growing evidence that circRNAs are involved in a cap-independent protein translation mediated by structures such as the internal ribosome entry site (IRES) and N6-methyladenosine (m6A) [13,15]. Therefore, the translation of sORFs and production of micropeptides is an important aspect of non-coding RNA biology [8].

Short ORFs may express a large number of peptides which are involved in a wide range of physiological and pathological processes [6, 16,17]. Several micropeptides have shown to be involved in multiple oncogenic activities, such as neoplastic transformation, angiogenesis, abnormal signalling pathway transduction and cancer metabolism [18–21]. Micropeptides can act as tumour suppressors [20,22], such as the HOXB-AS3 peptide encoded by the lncRNA *HOX-AS3*, which was shown to inhibit the growth of colorectal cancer cells by regulating tumour energy metabolism [23]. In addition, the FBXW7-185aa peptide encoded by circular *FBXW7*, was shown to inhibit the proliferation and migration of malignant cells derived from triple negative breast cancer via c-Myc degradation and ubiquitination of proliferating cell nuclear antigen (PCNA) [24]. The primary microRNAs (pri-miRNAs) in plants have been found to contain functional sORFs that are translated into small peptides called miPEPs (microRNA-encoded peptides). Initially identified in *Arabidopsis thaliana* and *Medicago truncatula*, miPEPs have been recently found to be encoded by pri-miRNAs in several plants and animals species [25]. Then miPEPs are more common than previously thought and may be a common feature of pri-miRNAs. Few proteins or peptides belong to this group and only a few show activity on tumour cells such as miPEP133, which inhibited the migration and invasion of cancer cells [26].

This comprehensive review describes the current understanding of non-coding RNA-derived peptides across various cancer types. Specifically, it recapitulates the state of the art on the biology of peptides encoded by lncRNAs, miRNAs and circRNAs, elucidating their regulatory functions and implications in cancer pathogenesis.

2. Discovery of peptides encoded by non-coding RNA

Recent advancements in omics-based technologies have revolutionized our knowledge of the coding potential embedded within the human genome. Through these studies, a plethora of unannotated ORFs dispersed throughout the genomes have been unveiled. Contrary to the traditional understanding of protein-coding genes, these non-canonical ORFs possess the capacity to translate novel micropeptides and may also harbour regulatory functions. Despite their genomic abundance, the broader scientific community has only begun to scratch the surface in characterizing these enigmatic elements [1,27].

The limited functional characterization of non-canonical ORFs can be attributed, at least in part, to their under-recognition. Nevertheless, the indispensable roles of a few of them in diverse critical biological processes have been well characterized [6]. More recently, the advancements in high-throughput sequencing technologies, coupled with sophisticated bioinformatics tools, have facilitated the identification and characterization of non-canonical ORFs on a genome-wide scale. These efforts have revealed a myriad of non-canonical ORFs exhibiting diverse characteristics, including various lengths, genomic locations, and regulatory elements [13,28]. Three main methods are used to identify potential sORFs, namely mass spectrometry, ribosome profiling and bioinformatics tools [29]. In addition, machine-learning tools specific for the identification of micropeptides showed clear advantages over existing prediction methods based on coding potential for micropeptides [27,30].

2.1. Mass spectrometry-based proteomics

Mass spectrometry is the gold standard to identify and quantify peptides and proteins as well as to detect post-translational modifications. It has been recently shown that peptidomic assays based on mass spectrometry, compared to ribosome profiling method, are able to directly detect peptides generated by ORFs allowing to validate their

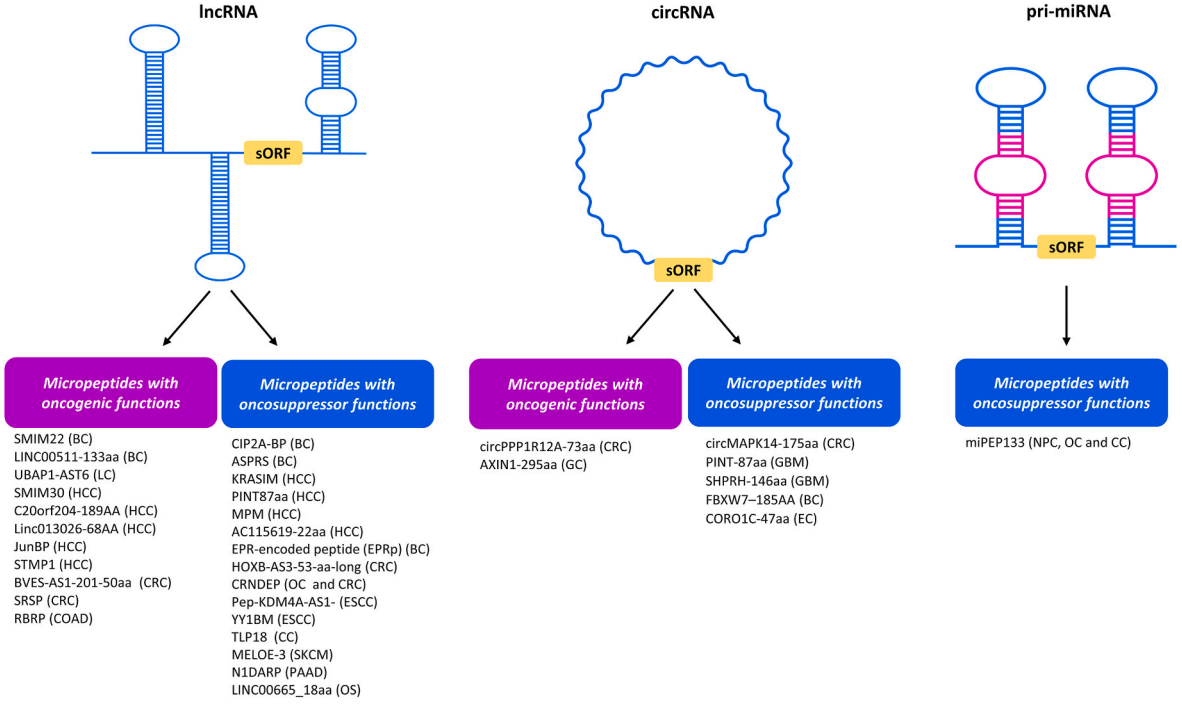

Fig. 1. Schematic representation of micropeptides derived from non-coding RNAs that were shown to have a role in different cancer types. Micropeptides encoded by long non-coding RNA (lncRNA), circular RNA (circRNA) and pri-microRNA (pri-miRNA) are listed and divided into tumour promoters and tumour inhibitors. BC, breast cancer; LC, lung cancer; HCC, hepatocellular carcinoma; CRC, colorectal cancer; COAD, colon adenocarcinoma; OC, ovarian cancer; ESCC, esophageal squamous cells carcinoma; CC, cervical cancer; SKCM, skin cutaneous melanoma; PAAD, pancreatic adenocarcinoma; OS, osteosarcoma; GC, gastric cancer; GBM, glioblastoma multiform; EC, endometrial cancer; NPC, nasopharyngeal carcinoma.

origin [31]. However, less than 1 % of lncRNA-encoded peptides have been experimentally characterized using mass spectrometry-based proteomic. Nonetheless, mass spectrometry has a bias toward detecting abundant proteins, limiting the identification of micropeptides. Additionally, small proteins are often lost during sample preparation since standard protein purification protocols tend to exclude proteins smaller than 10 kDa [31]. For this reason, the analysis of micropeptides with mass spectrometry requires several phases of sample preparation, including tryptic fragmentation of proteins, size exclusion methods to enrich peptides, reduction and de-salting. Furthermore, the combination of mass spectroscopy with liquid chromatography has been shown to enhance the separation and analysis of micropeptides significantly [27]. Validation of newly discovered micropeptides is crucial, particularly when only one peptide is detected. One common validation method is based on the use of isotopically labelled standards, which are chemically identical to the native peptides but have a mass shift due to the label. Another approach involves siRNA-mediated silencing of the transcript, followed by targeted mass spectrometry with peptide standards, and confirmation through RT-qPCR [27].

2.2. Ribosome profiling

Ribosome profiling (Ribo-seq) is an innovative technique that maps ribosome positions on all mRNAs by sequencing ribosome-protected mRNA fragments that have the potential to encode peptides [32–34]. Ribo-seq involves sequencing of mRNA fragments approximately 30 nucleotides long that are protected by ribosomes after nuclease digestion. The resulting reads are aligned to the transcriptome, providing a genome-wide view of ribosome occupancy. This technique allows to identify where translation occurs and has revealed unconventional and sORF translation in regions once considered to be non-coding [27]. Ribo-Seq studies in Drosophila, zebrafish, mice, and humans have uncovered widespread translation on lncRNAs, as well as in upstream, downstream, and overlapping coding regions. Since ribosome profiling cannot reliably differentiate between non-coding and coding transcripts, advanced metrics and algorithms like ORF-RATER, PROTEOFORMER, ORFscore, and FLOSS have been developed to enhance ribosome-profiling data analysis [19,35–38]. Important databases collecting ribosome-profiling data include FASTA, RPFdb, GWIPS-viz, and TISdb [39–41]. They have been constructed with computational tools that use distinct approaches to maximize sensitivity and specificity and have significantly expanded the possibility to annotate novel sORFs from ribosome profiling data. However, only few sORFs have been further characterized, unveiling their specific roles in cellular processes. By combining standardization and systematic annotation of sORFs with the development of comprehensive tools to dissect their potential functions, the approaches will likely shed light on which sORFs impact health and disease.

2.3. Bioinformatics tools

Currently, there are few bioinformatics tools specifically designed to predict the coding potential of sORFs. The web server uPEPperoni is designed to identify sORFs in the 5′ untranslated regions (5′-UTR) of mRNAs [42]. It detects conserved sORFs without explicitly predicting their coding potential. Although 5′-UTR sORFs are an important component of the sORF population, many of them are located elsewhere, such as within the coding region of mRNAs, in lncRNAs, etc. The sORF finder was designed to identify sORFs on the basis of nucleotide frequency conditional probabilities of the sequence, however it was developed years ago and the server is no longer accessible [43]. In the past decade, many new micropeptides have been discovered, allowing the construction of larger training datasets, which should improve prediction accuracy. Traditional coding potential prediction tools, like CPC2, CPAT, CNCI, and PhyloCSF, were trained on datasets of normal-sized proteins, making them less effective at predicting sORFs. These tools tend to penalize sORFs and those lacking similarity to known proteins, reducing their ability to accurately predict sORFs. However, some techniques like Ribo-Seq and mass spectrometry have led to the experimental identification and validation of an increasing number of micropeptides [27,30].

Comparative genomic analyses have highlighted the evolutionary conservation of some non-canonical ORFs across mammalian species, which reinforces their importance in cognate biological roles [1]. For instance, some studies have highlighted the involvement of non-canonical ORFs in crucial cellular processes such as transcriptional regulation, alternative splicing, protein homeostasis, and signal transduction as well as in disease pathogenesis, including cancer progression and neurodegenerative disorders [44]. Microproteins identified through bioinformatic tools must be validated experimentally, for example by using sequence-specific antibodies. However, due to their small size, designing effective antigen targets for antibody production is challenging, and transmembrane or protein-associated microproteins may limit antibody accessibility in methods such as immunoprecipitation and immunocytochemistry. Alternatively, CRISPR/Cas9 genome editing can be used to insert a coding sequence for an epitope tag (e.g., HA, FLAG, myc) in-frame with the sORF of interest via homology-directed repair (HDR) [45]. It is recommended to screen clones with the tag inserted both at the N- and C-terminus, as the tags can affect the microprotein localization and function [6].

Recent advances in understanding the diversity and functional relevance of non-coding RNAs have shed new light on the complexity of the mammalian genomes. While much remains to be explored, these findings underscore the importance of further investigating sORFs to unravel their intricate roles in cellular processes and disease pathogenesis. Furthermore, the ability to synthesize and chemically modify endogenous micropeptides opens opportunities for drug development [46]. Modified micropeptides can be radiolabelled or conjugated to clinically effective drugs (doxorubicin, paclitaxel, etc.) to directly target tumours for diagnosis or therapy. Furthermore, they could be modified with cell penetrating peptide (CPP) sequences capable of conveying the modified peptides into various cellular compartments (nucleus, mitochondrion, and cytoplasm) where the micropeptides carry out their antitumour activity. Moreover, different approaches have been considered to improve the binding activity and to prevent degradation including the use of liposomes, micro- and nanoparticles or unnatural amino acids [47].

3. LncRNA-encoded peptides with oncogenic or oncosuppressor functions

Diverse peptides encoded by non-coding RNAs have been shown to play a key role in several oncogenic processes [21]. LncRNAs were initially defined as non-coding RNAs without protein-coding capability. More recently, it has been shown that many lncRNAs contain sORFs translated into functional peptides that play crucial roles either as tumour promoters or tumour suppressors in diverse cancer types. Currently known cancer-promoting and cancer-inhibiting peptides encoded by lncRNAs are listed in Table 1.

3.1. Peptides encoded by lncRNAs in breast cancer

The peptide CASIMO 1 (cancer-associated small integral membrane open reading frame 1) is a 10 kDa molecule mainly overexpressed in hormone-positive breast cancer [48]. CASIMO 1, characterized by a small transmembrane conserved domain, was shown to promote cellular proliferation in three cell lines MCF7, KPL1, and T47D, as well as in non-transformed MCF10a breast cells [48]. The peptide CASIMO 1 was able to interact with squalene epoxidase (SQLE), an enzyme involved in cholesterol synthesis and known to act as an oncogene in breast cancer [48]. The accumulation of SQLE protein promoted by the overexpression of CASIMO1 suggests that these micropeptide plays a crucial

Table 1
Selected peptides encoded by LncRNAs with oncogenic and oncosuppressor functions.

Peptides sequence names	Peptide AA lenght	Cancers	Lnc-RNA symbols	Lnc-RNA name	Cellular compartment	Biological Function	Cell lines	References
Micropeptides with oncogenic functions								
CASIMO1	83	Breast cancer	*SMIM22*	small integral membrane protein 22	cytoplasm	Promotes cell proliferation and migration	MCF7, KPL1, T47D, MCF-10A,	[42]
LINC00511-133aa	133	Breast cancer	*LINC00511*	long intergenic non-protein coding RNA 511	nucleus	Regulates breast cancer cell invasion	MCF7, MDA-MB-231	[43]
UBAP1-AST6	117	Lung cancer	*UBAP1 (MN_033243:1.1)*	ubiquitin associated protein 1	nucleus	Promotes cell proliferation and colony formation	A549	[47]
SMIM30	59	Hepatocellular carcinoma	*SMIM30*	small integral membrane protein 30	cytoplasm, endoplasmic reticulum and mitochondria, membrane	Promotes tumour growth	SK-HEP-1	[50,51]
C20orf204-189AA	189	Hepatocellular carcinoma	*C20orf204*	chromosome 20 open reading frame 204	nucleus	Promotes cell proliferation	HeLa, Huh7, HepG2	[52]
Linc013026-68AA	68	Hepatocellular carcinoma	*LINC013026*		cytoplasm	promotes cell proliferation	HepG2, HeLa	[53]
JunBP	174	Hepatocellular carcinoma	*LINC02551*	long intergenic non-protein coding RNA 2551	cytoplasm	Promotes cell proliferation	97H, Hep3B	[54]
STMP1	47	Hepatocellular carcinoma	*STMP1*	short transmembrane mitochondrial protein 1	mithocondria	Promotes cell proliferation	Hepa1-6, Huh7, Sk-Hep-1	[55]
BVES-AS1-201-50aa	50	Colorectal cancer	*BVES-AS1*	BVES antisense RNA 1	nucleus	Promotes tumour cell viability, migration, and invasion *in vitro*	HCT116, SW480	[60]
SRSP	130	Colorectal cancer	*FLJ20021*	uncharacterized LOC90024	nucleus	Promotes tumorigenesis and progression	HCT-116, SW480, SW620	[61]
RBRP	71	Colon cancer	*SEPTIN14P20*	septin 14 pseudogene 20	nucleus	Promotes proliferation and metastasis	SW620, SW480	[62]
Micropeptides with oncosuppressor functions								
CIP2A-BP	52	Breast cancer	*LINC00665*	long intergenic non-protein coding RNA 665	cytoplasm	inhibited triple-negative breast cancer progression	MDA-MB-231, Hs578T, MCF-10A, BT549	[44]
ASRPS	60	Breast cancer	*LINC00908*	long intergenic non-protein coding RNA 908	cytoplasm	Inhibits triple-negative breast cancer angiogenesis	MDA-MB-231, Hs578T	[45]
EPR-encoded peptide (EPRp)	71	Breast cancer	*SMIM31*	small integral membrane protein 31	cytoplasm	Reduces breast cancer cell proliferation	NMuMG HEK-293	[46]
KRASIM	99	Hepatocellular carcinoma	*NCBP2-AS2*	NCBP2 antisense 2 (head to head)	cytoplasm	Inhibits oncogenic signals, cancer cell growth and proliferation	HuH-7, SK-HEP1	[56]
PINT87aa	87	Hepatocellular carcinoma	*LINC-PINT*	long intergenic non-protein coding RNA, p53 induced transcript	cytoplasm	Induces growth inhibition, cellular senescence, and decreases mitophagy	SMMC-7721, HepG2, HEK293	[57]
MPM	56	Hepatocellular carcinoma	*MTLN*	mitoregulin	mitochondria	Suppresses migration and invasion capabilities of hepatoma cells	Hepa1-6, SK-HEP1 SNU-449, HeLa	[58]
AC115619-22aa	22	Hepatocellular carcinoma	*AC115619*		nucleus	Inhibits growth and proliferation of cancer cells	PLC/PRF/5, HUH7	[59]
HOXB-AS3	53	Colorectal cancer	*HOXB-AS3*	HOXB cluster antisense RNA 3	nucleus	Inhibits growth of cancer cells	HTC-116	[20]
CRNDEP	84	Ovarian cancer and colorectal	*CRNDE*	colorectal neoplasia differentially expressed	nucleus	Regulates carcinogenesis and cell proliferation	HeLa	[63]
KDM4A-AS1	61	Esophageal squamous cell carcinoma	*KDM4A-AS1*	KDM4A antisense RNA 1	mitochondria membrane	Weakenes cancer cell viability and migratory capacity	KYSE150, TE-1	[64]

(*continued on next page*)

Table 1 (*continued*)

Peptides sequence names	Peptide AA lenght	Cancers	Lnc-RNA symbols	Lnc-RNA name	Cellular compartment	Biological Function	Cell lines	References
YY1BM	21	Esophageal squamous cell carcinoma	*LINC00278*	long intergenic non-protein coding RNA 278	cytoplasm	Regulates eEF2K expression and inhibites apoptosis	KYSE-30, Eca-109, TE-1	[65]
TLP18	18	Cervical cancer	*NKILA*	NF-kappaB interacting lncRNA	cytoplasm	Reduces cancer cell proliferation	SiHa, HeLa	[66]
MELOE-3	54	Melanoma	*MGC16025*	Uncharacterized LOC85009	nucleus	Improves T cell targets for immunotherapy	M113, M117	[69]
N1DARP	41	Pancreatic cancer	*LINC00261*	long intergenic non-protein coding RNA 261	cytoplasm	Tumour suppressor	Panc-1, Capan-1	[72]
LINC00665_18aa	18	Osteosarcoma	*LINC00665*	long intergenic non-protein coding RNA 665	nucleus	Suppresses viability, proliferation, and migration of human cells *in vitro* and diminishes tumour growth *in vivo*	MNNG-HOS, U2OS OS	[73]

role in breast cancer development.

The lncRNA *LINC00511*, which encodes a peptide of 133 amino acids (LINC00511), demonstrated to act as an oncogene by promoting proliferation and invasion of breast cancer cells [49]. Indeed, the transfection of LINC00511 into MCF-7 and MDA-MB-231 breast cancer cells was shown to promote increased cell proliferation and migration, whereas apoptosis was inhibited by the expression of Bax, c-myc, and CyclinD1 and localization of β-catenin protein into the nucleus. In addition *LINC00511*-transduced breast cancer cells showed an increased expression of Oct4, Nanog, and SOX2, through the activation of the wnt/β-catenin pathway, which are markers of tumour stem cell-like phenotype.

Among micropeptides acting as oncosuppressors in breast cancer, the 52 amino acids CIP2A-BP, encoded by LINC00665, was identified by bioinformatics analysis [50]. The expression of CIP2A-BP was inhibited by TGF-β in breast cancer cell lines and downregulation of this micropeptide in triple-negative breast cancer (TNBC) patients was found associated with reduced overall survival [50]. Conversely, high levels of CIP2A-BP were found to suppress breast cancer and to inhibit metastasis to the lung.

In addition, the 60 amino acids long ASRPS, an endogenous peptide encoded by LINC00908, was found differentially expressed in TNBC tissues [51]. The ASRPS was shown to reduce angiogenesis by inhibiting the expression of VEGF in a mouse xenograft breast cancer model. Therefore, the micropeptide ASRPS is considered an anti-tumour peptide and proposed as a candidate target for the treatment of TNBC [51].

Moreover, a peptide of 71 amino acids in length originating from the lncRNA *SMIM31* was found to be expressed in epithelia tissues and shown to localize to the cell-junction and to reduce breast cancer cell proliferation [52]. The molecular mechanisms involved in such function are not clear, however, it was hypothesized that such peptide has specific functions related to its cytoskeletal/junctional localization, such as participation in multiprotein complexes involved in the permeability barriers and/or implicated in the establishment of apical-basal polarity as well as in signal transduction [52].

3.2. Peptides encoded by lncRNAs in lung cancer

The study of the coding potential of lncRNAs in lung cancer cell lines, through several experimental techniques including translation studies, mass spectrometry, antibody validation and bioinformatics analysis, allowed to identify 308 small proteins encoded by lncRNAs [53]. Among these five lncRNA encoded microproteins, namely *SNHG17, PAX8-AS1, OTUD6B-AS1, MBNL1-AS1* and *MAPKAPK5*-AS1, were also identified by other groups [54].

In addition, the micropeptide UBAP1-AST6, encoded by a new lncRNA, was also identified in the nucleoli of lung cancer cell lines and observed to have an important role in promoting the proliferation of tumour cells [53]. However, this effect was not observed with the *UBAP1-AST6* RNA form.

3.3. Peptides encoded by lncRNAs in hepatocellular carcinoma

Peptides encoded by sORFs are emerging as important factor that influence various aspects of hepatocellular carcinoma (HCC) development and progression, providing new avenues for research and potential therapeutic strategies [55].

SMIM30 is an endogenous 59 amino acid peptide encoded by the lncRNA *SMIM30*, which modulates HCC cell proliferation and migration [56]. SMIM30 was found in the membranes of endoplasmatic reticulum and mitochondria [57]. High levels of this peptide were shown to increase the proliferation of hepatoma cells *in vitro* and the growth of tumour xenografts and N-nitrosodiethylamine-induced hepatomaby intervening in the control of cell cycle transcription during the G1 and S transition phases via the increased activity of the calcium ATPase of the sarco-endoplasmic reticulum and reducing calcium levels in the cytosol [57].

*C20orf*204 RNA, a splice variant of *Linc00176*, encompasses a 189 amino acid (AA) long ORF known as C20orf204-189AA [58]. This peptide was predominantly expressed in the nuclei of cells in 14 out of 20 primary HCC samples, whereas it was not present in normal liver tissues. Overexpression of C20orf204-189AA was shown to enhance cell proliferation and ribosomal RNA transcription. Additionally, it upregulates the protein level of nucleolin, a multifunctional nucleolar protein involved in various aspects of ribosome biogenesis and cell proliferation. The level of nucleolin and *C20orf*204 mRNA in HCC are correlated with tumour differentiation grade and patient survival, indicating their potential as prognostic markers. Targeting C20orf204-189AA could lead to novel therapeutic strategies aimed at inhibiting tumour progression and improving patient outcomes in HCC [58].

Linc013026 encodes a 68 amino acid peptide whose localization in the perinuclear region implies its potential involvement in critical cellular processes near the nucleus, influencing cell growth and division. Given its specific expression in HCC cells and its role in enhancing cell proliferation, Linc013026-68AA may serve as an HCC-specific target molecule for therapeutic intervention [59].

LINC02551 is transcriptionally activated by TGF-β and encodes a 174 amino acid peptide known as Jun binding peptide (JunBP) [60]. This peptide was identified in both HCC tissues and HCC-derived cell lines. The activation of *LINC02551* by TGF-β and the presence of JunBP in HCC suggest a functional role in the pathogenesis of liver diseases. The interaction of JunBP with cellular components, particularly those involved in the Jun signalling pathway, could influence cell

proliferation and survival, contributing to HCC progression and metastasis. Therefore, targeting the TGF-β/*LINC02551*/JunBP axis might offer new therapeutic opportunities for treating HCC [60].

The *STMP1* (small transmembrane protein 1) encodes a peptide of 47 amino acids [61]. This peptide was found to be up-regulated in various malignancies, including HCC. Importantly, its elevated expression was associated with a shorter recurrence-free survival in HCC patients [61]. The transmembrane nature of STMP1 implies it may have functions related to cell membrane interactions or signalling pathways critical for cancer progression. Its association with poor recurrence-free survival highlights STMP1 as a potential prognostic biomarker in HCC, providing insights into disease progression and potential therapeutic targets for improving patient outcomes. Further research into the molecular mechanisms and biological roles of STMP1 in HCC is warranted to fully understand its implications and therapeutic potential. Mechanistically, STMP1 enhances the mRNA and protein levels of key regulators involved in cell cycle progression and proliferation, namely CCNE2 (cyclin E2), CDK2 (cyclin-dependent kinase 2), and E2F1 (transcription factor E2F1). These proteins play crucial roles in regulating the cell cycle and promoting cell proliferation. Furthermore, STMP1 interacts with mitochondrial complex IV (cytochrome *c* oxidase), a key enzyme involved in the electron transport chain and oxidative phosphorylation. This interaction leads to enhanced activity of mitochondrial complex IV, potentially boosting cellular energy production and supporting increased cell proliferation and tumour growth [61].

Among micropeptides characterized by oncosuppressor features, KRASIM is a conserved 99-amino acid micropeptide, differentially expressed in normal hepatocytes and HCC cells, which is encoded by *NCBP2-AS2* [62]. It is localized in the cytoplasm where it interacts with the KRAS protein carrying out a regulatory activity of oncogenic signalling. High levels of KRASIM corresponded to a decrease of KRAS protein level, causing the inhibition of ERK signalling activity in HCC cells [62]. The peptide PINT encoded by *LINK-PINT* in the cytoplasm played an important role in cellular senescence of HCC, induced growth inhibition and reduced elimination of dysfunctional mitochondria (mitophagy) both *in vivo* and *in vitro* [63]. It was also shown that shorter peptides from 2 to 39 amino acids performed the same activity of PINT, demonstrating that it can be considered a new therapeutic target for HCC [63].

Micropeptide in mitochondria (MPM) regulated myogenic differentiation and muscle development, but its role in cancer development is unknown. In HCC tissues, MPM was significantly downregulated, correlating with increased metastasis potential and recurrence. Studies showed that restoring MPM expression inhibited, while silencing MPM enhanced, hepatoma cell migration, invasion and liver metastasis to the lung. MPM interacted with NDUFA7, inhibiting the activity of mitochondrial complex I. This inhibition was reversed by siMPM, and the effect of siMPM was reduced by knocking down NDUFA7. MPM caused the reduction of NAD+/NADH ratio, while siMPM determined an increase. The NAD + precursor nicotinamide eliminated the inhibitory effect of MPM on hepatoma cell migration [64].

The micropeptide AC115619-22aa inhibited HCC progression by binding to WTAP and by disrupting the assembly of the N6-methyladenosine (m6A) methyltransferase complex, which controls the expression of tumour-associated genes like *SOCS2* and *ATG14*. AC115619 was co-transcribed with the upstream gene APOB, and their transcription was repressed under hypoxia via HIF1A/HDAC3 and HNF4A signalling. In animal and patient-derived models, AC115619-22aa reduced the global m6A levels and suppressed tumour growth, suggesting that AC115619 and its encoded peptide could be potential prognostic markers and therapeutic targets for HCC patients [65].

3.4. Peptides encoded by lncRNAs in colorectal cancer and ovarian cancer

Bioinformatic analyses revealed that BVES-AS1-201is a peptide encoded by *BVES-AS1* involved in promoting the growth, viability and migration of colorectal cancer (CRC) cells *in vitro* [66]. Splicing Regulatory Small Protein (SRSP) is a small protein of 130 amino acids in length encoded by the uncharacterized lncRNA *LOC90024*. Overexpression of SRSP correlated with malignant phenotypes and poor prognosis in CRC patients [67].

The 71 amino acids peptide, called RNA-binding regulatory peptide (RBRP) for its interaction with RNA-binding proteins, is encoded by *SEPTIN14P20* [68]. RBRP was observed to be upregulated in CRC tissues. In particular, this peptide binded to IGF2BP1 and improved the recognition of N6-methyladenosine, which increased the stability of mRNAs and c-Myc expression, thus exerting an oncogenic function [68]. Elevated levels of RBRP are associated with poor prognosis in CRC patients [68].

The lncRNAs acting as oncosuppressors in colorectal cancer include the LncRNA *HOX-AS3*, encoding a 53 amino acids peptide, which is able to suppress CRC cell growth. The HOXB-AS3 peptide competitively binds to the arginine residues in the RGG motif of hnRNP A1, thereby antagonizing hnRNP A1-mediated regulation of the alternative splicing of pyruvate kinase M (PKM). This interaction blocked the binding of hnRNP A1 to the sequences flanking PKM exon 9, resulting in the reduced formation of PKM2 and suppression of glucose metabolism reprogramming [23].

The endogenous CRNDEP peptide (84 amino acids in length) exhibited strong nuclear expression in cancer cells and normal tissues characterized by a high proliferation rate, promoted the formation of stress granules and regulated cell proliferation [69]. The upregulation of CRNDEP in various malignancies, including colorectal, blood, brain, and ovarian cancers, underscores its significance in cancer biology. The identification of CRNDEP as a marker of poor prognosis in ovarian cancer adds to its potential utility in clinical practice, both for prognostic assessment and as a target for novel therapeutic approaches. Further research is needed to fully understand the mechanisms by which CRNDEP influences cancer progression and to explore its potential in cancer management [69].

3.5. Peptides encoded by lncRNAs in esophageal squamous cell carcinoma

Several lncRNA-coding peptides were identified in esophageal squamous cell carcinoma (ESCC) by using weighted gene co-expression network analysis (WGCNA), ribosome sequencing analyses, ORF prediction, MS dataset identification, and western blotting validation [70]. In particular, six lncRNAs, namely LINC01116, *KDM4A-AS1, KMT2E-AS1,* LINC00839, *UBL7-AS1*, and *DLX6-AS1*, were shown to encode peptides and among them the 61 amino acids long sequence identified as Pep-KDM4A-AS1 reduced cell viability and inhibited cell migration in ESCC by influencing fatty acid metabolism and the redox process [70].

LINC00278 and its encoded peptide YY1BM play crucial roles in the progression of male esophageal squamous cell carcinoma by modulating the interaction between YY1 and the androgen receptor, thereby affecting eEF2K expression [71]. Downregulation of YY1BM, influenced by cigarette smoking through decreased m6A modification, leads to increased eEF2K expression, reduced apoptosis, and enhanced cell survival under nutrient deprivation. These findings offer new insights into the molecular mechanisms of ESCC progression and suggest potential therapeutic targets and biomarkers for this malignancy [71].

3.6. Peptides encoded by lncRNAs in cervical cancer

The peptide TLP18 (FKCRRWQWRMKKLGAPSI) derived from lactoferricin, an antimicrobial peptide of 25 amino acids, which was isolated from cow's milk and exhibited *in vitro* cytotoxic activity against many different types of mouse and human cancer cell lines, including leukemia cells, fibrosarcoma cells, various carcinomas, and neuroblastoma cells [72]. In cervical cancer, TLP-18, was found to regulate SiHa

and HeLa cells proliferation through the lncRNA-NKILA/NF-κB feedback cancer This mechanism highlights the potential of TLP18 as a therapeutic agent and the lncRNA-NKILA/NF-κB pathway as a target for cervical cancer treatment [73,74].

3.7. Peptides encoded by lncRNAs in melanoma

MELOE-1 and MELOE-2, are highly specific melanoma antigens that play a role in T cell immunosurveillance but are poorly expressed in melanoma cells. These antigens present potential targets for immunotherapy and could serve as biomarkers for melanoma. Understanding and manipulating their expression levels may enhance immune system recognition against melanoma cells, offering new opportunities for treatment and diagnosis of this aggressive cancer [75,76]. A new peptide, MELOE-3 (54 amino acids long) encoded by MGC16025 (uncharacterized LOC85009), was found highly expressed in melanoma cells but showed low immunogenicity [76]. In addition, the stimulation ex vivo of peripheral blood monocyte cells from healthy donors with MELOE-1 or MELOE-3 overlapping peptides showed lower specificity of MELOE-3 T cell repertoire as compared to that against MELOE-1. The weak immunogenicity of MELOE-3 and its expression in melanocytes is consistent with the immune tolerance towards a physiologically expressed protein.

3.8. Peptides encoded by lncRNAs in pancreatic cancer

The Notch receptor signalling pathway is implicated in malignant transformation, playing a critical role during the development of pancreatic ductal adenocarcinoma [77,78]. Notch1 peptide degradation-associated regulatory polypeptide (N1DARP) encoded by *LINC00261* was found to be a tumour suppressor [79]. Low levels of N1DARP were shown to increase tumour progression and to enhance stem cell properties in pancreatic cancer organoids and LSL-Kras, LSL-Trp53, and Pdx1-Cre (KPC) mice [79]. The therapeutic potential of N1DARP was evaluated using the cell-penetrating stapled peptide, SAH-mAH2-5, which is structurally similar to N1DARP [79]. SAH-mAH2-5, a peptide that promotes the proteasome-mediated degradation of N1ICD, demonstrated significant therapeutic benefits in Notch1-activated pancreatic cancer models. The treatment is associated with limited off-target and systemic adverse effects, making it a promising candidate for targeted therapy in pancreatic cancer [79].

3.9. Peptides encoded by lncRNAs in osteosarcoma

A short peptide, encoded by the lncRNA *LINC00665*, was recently identified that suppresses the proliferation and migration of human osteosarcoma cells [80]. The *LINC00665*_18aa peptide was identified by bioinformatics analysis. LINC00665_18aa was observed to interfere with the transcriptional activity, nuclear localization, and phosphorylation of cAMP response element binding protein 1 (CREB1), a key nuclear transcriptional factor. Furthermore, the short peptide was shown to disrupt the interaction between CREB1 and the ribosomal protein S6 kinase A3 (RPS6KA3, RSK2) while increasing the CREB1 expression, reversing its inhibitory effects on osteosarcoma cell proliferation and migration [80].

4. Peptides encoded by circRNAs

CircRNAs, previously supposed to be untranslatable, were found to encode peptides that play crucial roles in cancer biology (Table 2). These findings challenge the traditional view of circRNAs and highlight their potential importance in tumorigenesis. The discovery of circRNA-encoded proteins opens new research directions for understanding cancer mechanisms and developing targeted therapies [81,82].

4.1. Peptides encoded by circRNAs in glioblastoma

circRNA *LINC-PINT* encodes a peptide of 87 amino acids in length that suppresses the proliferation of glioblastoma cells *in vitro* and *in vivo* by interacting with the polymerase-associated factor complex (PAFc)

Table 2
Selected peptides encoded by circRNAs and miRNAs with oncogenic and oncosuppressor functions.

Peptide sequence names	Peptide length	Cancers	circRNA and miRNA symbols	circRNA and miRNA names	Cellular compartment	Biological function	Cell lines	References
PINT	87	Glioblastoma	*LINC-PINT*	long intergenic non-protein coding RNA, p53 induced transcript	cytoplasm	Suppresses cancer cell proliferation	SW1783, Hs683, U251, 293T	[76] [83]
SHPRH	146	Glioblastoma	*SHPRH*	SNF2 histone linker PHD RING helicase	cytoplasm	Suppresses tumour growth	U251, U373	[77]
FBXW7	185	Breast cancer	*FBXW7-AS1*	FBXW7 antisense RNA 1	cytoplasm	Inhibits cancer cell proliferation and migration	BT549, 4T1	[21]
AXIN1	295	Gastric cancer	*AXIN1*	axin 1	cytoplasm	Promotes cell proliferation and migration	AGS	[78]
circPPP1R12A	73	Colorectal cancer	*PPP1R12A*	protein phosphatase 1 regulatory subunit 12A	cytoplasm	Promotes proliferation and metastasis	HT-29, HCT-116, SW480, SW620, LoVo, SW48, DLD-1, Caco2, HCT-15, NCM460	[79]
circMAPK14	175	Colorectal cancer	*MAPK14*	mitogen-activated protein kinase 14	cytoplasm	Blocks the proliferation and metastasis in cancer cells	SW480, DLD-1, LoVo, HT29, HCT116, Caco2, NCM460	[80]
CORO1C	47	Endometrium tumour	*CORO1C*	coronin 1C	nucleus	Suppresses cell proliferation and migration	Ishikawa, HEC-1-B	[81]
miPEP133	133	Nasopharyngeal carcinoma, ovarian cancer and cervical cancer	*miR-34a*	microRNA 34a	mitochondria	Induces apoptosis and inhibits migration and invasion of cancer cells	HNE3, C666–1, CNE2, CNE1, 5–8F, TWO3, NP69	[23]

causing inhibition of the transcriptional elongation of multiple oncogenes [84]. In addition, *SHPRH* encoding the peptide SHPRH is a new anti-tumour protein that blocks the proliferation and the tumorigenicity of glioblastoma cells. The mechanism involved in the oncosuppressor activity relies on the ability of SHPRHto interact with with denticleless protein homolog (DTL), thus promoting the ubiquitantion of the PCNA [83,85].

4.2. Peptides encoded by circRNAs in breast cancer

FBXW7AS1, already considered a tumour suppressor in glioma, was shown to inhibit the proliferation and migration of TNBCcells [24]. In particular, the peptide FBXW7 encoded by F*BXW7AS1* was shown to inhibit the proliferation and migration capacity of TNBC cells by increasing the abundance of FBXW7 and inducing the degradation of c-Myc [24].

4.3. Peptides encoded by circRNAs in gastric cancer

AXIN1 encoding the novel peptide, AXIN1, was shown to promote the proliferation, migration, invasion, and metastasis of gastric cancer cells *in vitro* and *in vivo* [86]. It was found that AXIN1 competitively interacts with antigen presenting cells (APC) leading to dysfunction of the "destruction complex" of the Wnt pathway. Released β-catenin was shown to translocate to the nucleus and to bind the TCF consensus site on the promoter, inducing downstream gene expression [86].

4.4. Peptides encoded by circRNAs in colorectal cancer

The expression pattern of *PPP1R12A* mainly localized in the cytoplasm of colorectal cancer cells was confirmed by real-time quantitative PCR and tissue microarray [87]. It was demonstrated that *PPP1R12A* encoding the peptide, circPPP1R12A, promoted the growth and metastasis of colorectal cancer via activating Hippo-YAP (Yes-associated protein) signalling pathway. Furthermore, Pep 17, a specific inhibitor of YAP, was shown to alleviate the promoting effect of circPPP1R12A on colorectal cancer cells [87]. Liquid chromatography and mass analysis combined with western blotting demonstrated the presence in the cytoplasm and relative expression of the circMAPK14 micropeptide with the property of inhibiting the expression level of colorectal cancer cells and control tissue samples [88].

4.5. Peptides encoded by circRNAs in endometrial tumour

CORO1C, encoding the peptide CORO1C, was demonstrated to be significantly downregulated in human endometrial cancer (EC) compared with matched para-cancerous tissues [89]. The overexpression of CORCO1Cinhibited the angiogenesis in the early phases by suppressing endothelial cell proliferation, migration, and differentiation by competing with the transcription factor TACC3 on the binding to ARNT and by suppressing VEGF expression [89].

5. Oncosuppressors peptides encoded by miRNAs

MiRNAs are small RNAs approximately 18–25 nucleotides in length that regulate gene expression by explicitly cleaving or inhibiting the translation of target miRNAs [90]. The biogenesis of miRNAs undergoes two main steps: 1) transcription into primary miRNA (pri-miRNAs) in the nucleus, 2) processing to form a pre-miRNA for translocation to the cytoplasm, and finally cleavage by the Dicer enzyme [90]. Furthermore, pre-miRNAs initiate a new phase of genetic regulation with peptides encoded by miRNAs (miPeps) [91]. Very few peptides encoded by miRNA (precursor of miR-34a) are reported in the literature. A micropeptide of 133 amino acids in length, miPEP133, encoded by *miR-34a* precursor miRNA, was shown to act as tumour-suppressor in a xenograft nasopharyngeal carcinoma (NPC) model [26]. The overexpression of miPEP133 in cancer cell lines induced apoptosis and inhibited their migration and invasion. It can be considered a prognostic marker associated with advanced metastatic NPC [26].

6. Immunogenic micropeptides

In recent years, new cancer treatments have emerged, such as immunotherapy and targeted therapy, which associated with conventional treatments, including surgery, radiotherapy and chemotherapy, have greatly improved the prognosis of cancer patients. Particularly, neo-epitopes identified by screening technologies have attracted much attention as targets for immunotherapy against cancer and for development of personalized vaccines [92]. A novel source of neo-epitopes derived from cancer-specific lncRNAs and circRNAs, which have greatly expanded the pool for neo-antigen discovery [82,93]. The automated bioinformatics pipeline are able to predict circRNA-derived neo-epitopes on the basis of RNA sequencing data and their binding to leukocyte antigen types (HLA I and HLA II). Functional peptides encoded by circRNA showed important roles in the development of glioblastoma multiforme (GBM), which is the most malignant and prevalent subtype of glioma. Particularly, it was shown that *circ-AKT3* encodes a novel tumour suppressor micropeptide, Akt3-174AA, which can be a source of neo-epitopes [82,94]. *FBXW7* already described in this article as a glioma suppressor and as an inhibitor of tumour cells in breast cancer can also be considered a source of neo-antigens encoding the peptide FBXW7-185AA, which contains eight overlapping peptides with high HLA I binding potential and 24 overlapping peptides with high binding potential to HLA II (Table 3) [82].

Furthermore, it was reported that the protein arginine methyltransferase (PRMT) 5, which is over-expressed in many cancer types, and the master regulator E2F1 have a striking effect on the expression of lncRNAs. The selective pharmacological inhibition of PRMT5 with the small molecule active site inhibitor T1-44 qualitative alteration of lncRNA-derived antigens displayed by tumour cells [96]. In addition, lncRNA-derived peptides triggered a potent antigen-specific CD8 T lymphocyte response, which caused a significant delay in tumour growth. Thus, lncRNA genes encode immunogenic peptides that can be deployed as a cancer vaccine [96].

Two peptides (HEDTGNPGL and RLQEGLAAV) encoded by pseudogene RHOXF1P3 and by lncRNA lnc-SERPIND1-41:10 (DGCR9 intron) were found to be expressed at high levels in tumours compared to normal tissues [97]. These peptides, considered neo-antigens and called "dark antigens", possess potent immunogenic activity and could be considered new candidates for cancer vaccines [97]. Furthermore, the circRNA FAM53B, which derives from the circularization of exon 2 of the human FAM53B gene, was shown to be specifically expressed in multiple human breast cancer cell lines, but not in normal breast epithelial MCF-10A cells [95]. In addition, circRNA FAM53B has been found expressed in the tumour tissue of 198 out of 212 (93.4 %) cases of invasive ductal carcinoma of the breast but was not detected in 29 normal breast tissues by in situ hybridization. The circRNA FAM53B encodes a 219-amino-acid peptide, termed circFAM53B-219, which showed strong binding affinity to both HLA-I and HLA-II molecules. *In vivo* experiments demonstrated that vaccines made of tumour-specific circRNA or its encoded peptides enhanced the infiltration of tumour-antigen specific cytotoxic T cells, leading to effective tumour control in mice with breast cancer tumours or melanoma [95]. Peptide candidates circFAM53B (192–200) or circFAM53B(210–218) showed strong binding affinity to HLA-I and where present in HLAI peptidomes of 50 % of breast cancer patients tested. Additionally, peptides circFAM53B(190–204), circFAM53B(191–204) and circFAM53B(192–205), from an HLA-II peptide library displayed high affinity scores to HLA-DRB1. Both the full-length circFAM53B(181–219) and truncated peptides circFAM53B(190–204) induced strong CD4 T cell responses.

Table 3
Immunogenic peptides encoded by circRNAs.

Micropeptides sequence name	Cancers	circRNA symbol	circRNA name	Neoantigens sequences binding HLA I	Neoantigens sequences binding HLA II	References
FBXW7-185AA	Glioma and triple negative breast cancer	*FBXW7*	F-box and WD repeat domain containing 7	SPFYTKTTKDY FYTKTTKDY TKTTKDYFL KTTKDYFLRI TTKDYFLRI YFLRIDCQKW LRIDCQKWSY RIDCQKWSY	DYFLRIDCQKWSY DYFLRIDCQKWSYW DYFLRIDCQKWSYWV DYFLRIDCQKWSYWVR KDYFLRIDCQKWSY KDYFLRIDCQKWSYW KDYFLRIDCQKWSYWV YFLRIDCQKWSYWVR LSSPFYTKTTKDYF LSSPFYTKTTKDYFL LSSPFYTKTTKDYFL LSSPFYTKTTKDYFLR LSSPFYTKTTKDYFLR LSSPFYTKTTKDYFLRI LSSPFYTKTTKDYFLRID SPFYTKTTKDYFL SPFYTKTTKDYFLR SPFYTKTTKDYFLRI SPFYTKTTKDYFLRID SSPFYTKTTKDYF SSPFYTKTTKDYFL SSPFYTKTTKDYFL SSPFYTKTTKDYFLR SSPFYTKTTKDYFLRI	[75]
circFAM53B-219	Breast cancer	*FAM53B*	family with sequence similarity 53 member B	ALFRLTNRA RTAHYGTGR	GNALFRLTNRAPASG KMTDGETWTGNALFRLTNRAPASG-NACLKRTAHYGTGRQ	[95]

7. Conclusion

Non-coding RNA-derived peptides, containing less than or slightly more than 100 amino acids, play a crucial role in the regulation of energy metabolism, proliferation and migration of tumour cells. These molecules are considered novel candidate targets for drugs and immunotherapeutic to inhibit tumour growth as well as novel diagnostic and prognostic biomarkers in cancer patients. Recent advances in omics technologies and bioinformatics have led to the identification of several sORFs from RNAs annotated as noncoding on a large genomic scale. Despite the presence of databases (https://metamorf.hb.univ-amu.fr/, https://github.com/MindAI/MiPepid, http://www.jianglab.cn/ncEP/, https://gwips.ucc.ie/cgi-bin/hgGateway) and technological support, to date few micropeptides have been identified. Therefore, novel methods and more sophisticated techniques are needed to better characterize and validate these molecules which have been shown to have key roles in various types of tumours.

CRediT authorship contribution statement

Anna Lucia Tornesello: Writing – review & editing, Writing – original draft, Data curation. **Andrea Cerasuolo:** Writing – original draft. **Noemy Starita:** Writing – original draft. **Sara Amiranda:** Writing – original draft. **Tiziana Pecchillo Cimmino:** Writing – original draft. **Patrizia Bonelli:** Writing – original draft. **Franca Maria Tuccillo:** Writing – original draft. **Franco Maria Buonaguro:** Supervision. **Luigi Buonaguro:** Supervision. **Maria Lina Tornesello:** Writing – review & editing, Writing – original draft, Validation, Supervision, Funding acquisition, Data curation, Conceptualization.

Data availability statement

Data and material are available at 10.5281/zenodo.13152073

Funding

This work was funded by grants from the Italian Ministry of Health Ricerca Corrente L1/10.

Declaration of competing interest

The authors declare no conflict of interest.

References

[1] B.W. Wright, Z. Yi, J.S. Weissman, J. Chen, The dark proteome: translation from noncanonical open reading frames, Trends Cell Biol. 32 (2022) 243–258.
[2] A.Z. Leong, P.Y. Lee, M.A. Mohtar, S.E. Syafruddin, Y.F. Pung, T.Y. Low, Short open reading frames (sORFs) and microproteins: an update on their identification and validation measures, J. Biomed. Sci. 29 (2022) 19.
[3] C.L. Sandmann, J.F. Schulz, J. Ruiz-Orera, M. Kirchner, M. Ziehm, E. Adami, M. Marczenke, A. Christ, N. Liebe, J. Greiner, A. Schoenenberger, M.B. Muecke, N. Liang, R.L. Moritz, Z. Sun, E.W. Deutsch, M. Gotthardt, J.M. Mudge, J. R. Prensner, T.E. Willnow, P. Mertins, S. van Heesch, N. Hubner, Evolutionary origins and interactomes of human, young microproteins and small peptides translated from short open reading frames, Mol. Cell. 83 (2023) 994–1011 e18.
[4] A. Sahgal, V. Uversky, V. Dave, Microproteins transitioning into a new Phase: defining the undefined, Methods 220 (2023) 38–54.
[5] J. Pan, R. Wang, F. Shang, R. Ma, Y. Rong, Y. Zhang, Functional micropeptides encoded by long non-coding RNAs: a comprehensive review, Front. Mol. Biosci. 9 (2022) 817517.
[6] K.R. Hassel, O. Brito-Estrada, C.A. Makarewich, Microproteins: overlooked regulators of physiology and disease, iScience 26 (2023) 106781.
[7] M.E. Sousa, M.H. Farkas, Micropeptide, PLoS Genet. 14 (2018) e1007764.
[8] S. Wang, C. Mao, S. Liu, Peptides encoded by noncoding genes: challenges and perspectives, Signal Transduct. Targeted Ther. 4 (2019) 57.
[9] X. Zheng, M. Xiang, Mitochondrion-located peptides and their pleiotropic physiological functions, FEBS J. 289 (2022) 6919–6935.
[10] J.P. Couso, P. Patraquim, Classification and function of small open reading frames, Nat. Rev. Mol. Cell Biol. 18 (2017) 575–589.
[11] M.L. Tornesello, R. Faraonio, L. Buonaguro, C. Annunziata, N. Starita, A. Cerasuolo, F. Pezzuto, A.L. Tornesello, F.M. Buonaguro, The role of microRNAs, long non-coding RNAs, and circular RNAs in cervical cancer, Front. Oncol. 10 (2020) 150.
[12] P. Kapranov, J. Cheng, S. Dike, D.A. Nix, R. Duttagupta, A.T. Willingham, P. F. Stadler, J. Hertel, J. Hackermuller, I.L. Hofacker, I. Bell, E. Cheung, J. Drenkow, E. Dumais, S. Patel, G. Helt, M. Ganesh, S. Ghosh, A. Piccolboni, V. Sementchenko, H. Tammana, T.R. Gingeras, RNA maps reveal new RNA classes and a possible function for pervasive transcription, Science 316 (2007) 1484–1488.
[13] Y. Lu, Y. Ran, H. Li, J. Wen, X. Cui, X. Zhang, X. Guan, M. Cheng, Micropeptides: origins, identification, and potential role in metabolism-related diseases, J. Zhejiang Univ. - Sci. B 24 (2023) 1106–1122.
[14] G.C. Shukla, J. Singh, S. Barik, MicroRNAs: processing, maturation, target recognition and regulatory functions, Mol. Cell. Pharmacol. 3 (2011) 83–92.

[15] A.C. Prats, F. David, L.H. Diallo, E. Roussel, F. Tatin, B. Garmy-Susini, E. Lacazette, Circular RNA, the key for translation, Int. J. Mol. Sci. 21 (2020).
[16] B.Y. Wang, Q. Gao, Y. Sun, X.B. Qiu, Biochemical targets of the micropeptides encoded by lncRNAs, Noncoding RNA Res 9 (2024) 964–969.
[17] J. Pan, X. Meng, N. Jiang, X. Jin, C. Zhou, D. Xu, Z. Gong, Insights into the noncoding RNA-encoded peptides, Protein Pept. Lett. 25 (2018) 720–727.
[18] H. Zhou, Y. Wu, J. Cai, D. Zhang, D. Lan, X. Dai, S. Liu, T. Song, X. Wang, Q. Kong, Z. He, J. Tan, J. Zhang, Micropeptides: potential treatment strategies for cancer, Cancer Cell Int. 24 (2024) 134.
[19] O. Boix, M. Martinez, S. Vidal, M. Gimenez-Alejandre, L. Palenzuela, L. Lorenzo-Sanz, L. Quevedo, O. Moscoso, J. Ruiz-Orera, P. Ximenez-Embun, N. Ciriaco, P. Nuciforo, C. Stephan-Otto Attolini, M.M. Alba, J. Munoz, T.V. Tian, I. Varela, A. Vivancos, Y.C.S. Ramon, P. Munoz, C. Rivas, M. Abad, pTINCR microprotein promotes epithelial differentiation and suppresses tumor growth through CDC42 SUMOylation and activation, Nat. Commun. 13 (2022) 6840.
[20] M. Ye, J. Zhang, M. Wei, B. Liu, K. Dong, Emerging role of long noncoding RNA-encoded micropeptides in cancer, Cancer Cell Int. 20 (2020) 506.
[21] S. Setrerrahmane, M. Li, A. Zoghbi, X. Lv, S. Zhang, W. Zhao, J. Lu, D.J. Craik, H. Xu, Cancer-related micropeptides encoded by ncRNAs: promising drug targets and prognostic biomarkers, Cancer Lett. 547 (2022) 215723.
[22] M. Li, G. Liu, X. Jin, H. Guo, S. Setrerrahmane, X. Xu, T. Li, Y. Lin, H. Xu, Micropeptide MIAC inhibits the tumor progression by interacting with AQP2 and inhibiting EREG/EGFR signaling in renal cell carcinoma, Mol. Cancer 21 (2022) 181.
[23] J.Z. Huang, M. Chen, D. Chen, X.C. Gao, S. Zhu, H. Huang, M. Hu, H. Zhu, G. R. Yan, A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth, Mol. Cell. 68 (2017) 171–184 e6.
[24] F. Ye, G. Gao, Y. Zou, S. Zheng, L. Zhang, X. Ou, X. Xie, H. Tang, circFBXW7 inhibits malignant progression by sponging miR-197-3p and encoding a 185-aa protein in triple-negative breast cancer, Mol. Ther. Nucleic Acids 18 (2019) 88–98.
[25] M. Ormancey, P. Thuleau, J.P. Combier, S. Plaza, The essentials on microRNA-encoded peptides from plants to animals, Biomolecules 13 (2023).
[26] M. Kang, B. Tang, J. Li, Z. Zhou, K. Liu, R. Wang, Z. Jiang, F. Bi, D. Patrick, D. Kim, A.K. Mitra, Y. Yang-Hartwich, Identification of miPEP133 as a novel tumor-suppressor microprotein encoded by miR-34a pri-miRNA, Mol. Cancer 19 (2020) 143.
[27] F. Valdivia-Francia, A. Sendoel, No country for old methods: new tools for studying microproteins, iScience 27 (2024) 108972.
[28] K. Grillone, C. Riillo, F. Scionti, R. Rocca, G. Tradigo, P.H. Guzzi, S. Alcaro, M.T. Di Martino, P. Tagliaferri, P. Tassone, Non-coding RNAs in cancer: platforms and strategies for investigating the genomic "dark matter", J. Exp. Clin. Cancer Res. 39 (2020) 117.
[29] D. Schlesinger, S.J. Elsasser, Revisiting sORFs: overcoming challenges to identify and characterize functional microproteins, FEBS J. 289 (2022) 53–74.
[30] M. Zhu, M. Gribskov, MiPepid: MicroPeptide identification tool using machine learning, BMC Bioinf. 20 (2019) 559.
[31] R. Vitorino, S. Guedes, F. Amado, M. Santos, N. Akimitsu, The role of micropeptides in biology, Cell. Mol. Life Sci. 78 (2021) 3285–3298.
[32] J. Ma, J.K. Diedrich, I. Jungreis, C. Donaldson, J. Vaughan, M. Kellis, J. R. Yates 3rd, A. Saghatelian, Improved identification and analysis of small open reading frame encoded polypeptides, Anal. Chem. 88 (2016) 3967–3975.
[33] C.M. Laumont, K. Vincent, L. Hesnard, E. Audemard, E. Bonneil, J.P. Laverdure, P. Gendron, M. Courcelles, M.P. Hardy, C. Cote, C. Durette, C. St-Pierre, M. Benhammadi, J. Lanoix, S. Vobecky, E. Haddad, S. Lemieux, P. Thibault, C. Perreault, Noncoding regions are the main source of targetable tumor-specific antigens, Sci. Transl. Med. 10 (2018).
[34] G. Loughran, A.V. Zhdanov, M.S. Mikhaylova, F.N. Rozov, P.N. Datskevich, S. I. Kovalchuk, M.V. Serebryakova, S.J. Kiniry, A.M. Michel, P.B.F. O'Connor, D. B. Papkovsky, J.F. Atkins, P.V. Baranov, I.N. Shatsky, D.E. Andreev, Unusually efficient CUG initiation of an overlapping reading frame in POLG mRNA yields novel protein POLGARF, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 24936–24946.
[35] J.L. Aspden, Y.C. Eyre-Walker, R.J. Phillips, U. Amin, M.A. Mumtaz, M. Brocard, J. P. Couso, Extensive translation of small open reading frames revealed by poly-ribo-seq, Elife 3 (2014) e03528.
[36] P. Bi, A. Ramirez-Martinez, H. Li, J. Cannavino, J.R. McAnally, J.M. Shelton, E. Sanchez-Ortiz, R. Bassel-Duby, E.N. Olson, Control of muscle formation by the fusogenic micropeptide myomixer, Science 356 (2017) 323–327.
[37] D.M. Anderson, K.M. Anderson, C.L. Chang, C.A. Makarewich, B.R. Nelson, J. R. McAnally, P. Kasaragod, J.M. Shelton, J. Liou, R. Bassel-Duby, E.N. Olson, A micropeptide encoded by a putative long noncoding RNA regulates muscle performance, Cell 160 (2015) 595–606.
[38] S.A. Slavoff, A.J. Mitchell, A.G. Schwaid, M.N. Cabili, J. Ma, J.Z. Levin, A. D. Karger, B.A. Budnik, J.L. Rinn, A. Saghatelian, Peptidomic discovery of short open reading frame-encoded peptides in human cells, Nat. Chem. Biol. 9 (2013) 59–64.
[39] A.M. Michel, G. Fox, A.M. A, C. De Bo, P.B. O'Connor, S.M. Heaphy, J.P. Mullan, C. A. Donohue, D.G. Higgins, P.V. Baranov, GWIPS-viz: development of a ribo-seq genome browser, Nucleic Acids Res. 42 (2014) D859–D864.
[40] J. Wan, S.B. Qian, TISdb: a database for alternative translation initiation in mammalian cells, Nucleic Acids Res. 42 (2014) D845–D850.
[41] H. Wang, L. Yang, Y. Wang, L. Chen, H. Li, Z. Xie, RPFdb v2.0: an updated database for genome-wide information of translated mRNA generated from ribosome profiling, Nucleic Acids Res. 47 (2019) D230–D234.
[42] A. Skarshewski, M. Stanton-Cook, T. Huber, S. Al Mansoori, R. Smith, S.A. Beatson, J.A. Rothnagel, uPEPperoni: an online tool for upstream open reading frame location and analysis of transcript conservation, BMC Bioinf. 15 (2014) 36.
[43] K. Hanada, K. Akiyama, T. Sakurai, T. Toyoda, K. Shinozaki, S.H. Shiu, sORF finder: a program package to identify small open reading frames with high coding potential, Bioinformatics 26 (2010) 399–400.
[44] C.A. Makarewich, E.N. Olson, Mining for micropeptides, Trends Cell Biol. 27 (2017) 685–696.
[45] Z. Zhang, Generation of epitope tag knock-in mice with CRISPR-Cas9 to study the function of endogenous proteins, STAR Protoc 4 (2023) 102518.
[46] A.L. Tornesello, M. Sanseverino, F.M. Buonaguro, Solid phase formylation of N-terminus peptides, Molecules 21 (2016).
[47] P.A. Wender, D.J. Mitchell, K. Pattabiraman, E.T. Pelkey, L. Steinman, J. B. Rothbard, The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 13003–13008.
[48] M. Polycarpou-Schwarz, M. Gross, P. Mestdagh, J. Schott, S.E. Grund, C. Hildenbrand, J. Rom, S. Aulmann, H.P. Sinn, J. Vandesompele, S. Diederichs, The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation, Oncogene 37 (2018) 4750–4768.
[49] Z. Tan, L. Zhao, S. Huang, Q. Jiang, Y. Wei, J.L. Wu, Z. Zhang, Y. Li, Small peptide LINC00511-133aa encoded by LINC00511 regulates breast cancer cell invasion and stemness through the Wnt/beta-catenin pathway, Mol. Cell. Probes 69 (2023) 101913.
[50] B. Guo, S. Wu, X. Zhu, L. Zhang, J. Deng, F. Li, Y. Wang, S. Zhang, R. Wu, J. Lu, Y. Zhou, Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression, EMBO J. 39 (2020) e102190.
[51] Y. Wang, S. Wu, X. Zhu, L. Zhang, J. Deng, F. Li, B. Guo, S. Zhang, R. Wu, Z. Zhang, K. Wang, J. Lu, Y. Zhou, LncRNA-encoded polypeptide ASRPS inhibits triple-negative breast cancer angiogenesis, J. Exp. Med. 217 (2020).
[52] M. Rossi, G. Bucci, D. Rizzotto, D. Bordo, M.J. Marzi, M. Puppo, A. Flinois, D. Spadaro, S. Citi, L. Emionite, M. Cilli, F. Nicassio, A. Inga, P. Briata, R. Gherzi, LncRNA EPR controls epithelial proliferation by coordinating Cdkn1a transcription and mRNA decay response to TGF-beta, Nat. Commun. 10 (2019) 1969.
[53] S. Lu, J. Zhang, X. Lian, L. Sun, K. Meng, Y. Chen, Z. Sun, X. Yin, Y. Li, J. Zhao, T. Wang, G. Zhang, Q.Y. He, A hidden human proteome encoded by 'non-coding' genes, Nucleic Acids Res. 47 (2019) 8111–8125.
[54] S. van Heesch, F. Witte, V. Schneider-Lunitz, J.F. Schulz, E. Adami, A.B. Faber, M. Kirchner, H. Maatz, S. Blachut, C.L. Sandmann, M. Kanda, C.L. Worth, S. Schafer, L. Calviello, R. Merriott, G. Patone, O. Hummel, E. Wyler, B. Obermayer, M.B. Mucke, E.L. Lindberg, F. Trnka, S. Memczak, M. Schilling, L. E. Felkin, P.J.R. Barton, N.M. Quaife, K. Vanezis, S. Diecke, M. Mukai, N. Mah, S. J. Oh, A. Kurtz, C. Schramm, D. Schwinge, M. Sebode, M. Harakalova, F. W. Asselbergs, A. Vink, R.A. de Weger, S. Viswanathan, A.A. Widjaja, A. Gartner-Rommel, H. Milting, C. Dos Remedios, C. Knosalla, P. Mertins, M. Landthaler, M. Vingron, W.A. Linke, J.G. Seidman, C.E. Seidman, N. Rajewsky, U. Ohler, S. A. Cook, N. Hubner, The translational landscape of the human heart, Cell 178 (2019) 242–260 e29.
[55] Q. Zhang, L. Liu, Novel insights into small open reading frame-encoded micropeptides in hepatocellular carcinoma: a potential breakthrough, Cancer Lett. 587 (2024) 216691.
[56] Y. Pang, Z. Liu, H. Han, B. Wang, W. Li, C. Mao, S. Liu, Peptide SMIM30 promotes HCC development by inducing SRC/YES1 membrane anchoring and MAPK pathway activation, J. Hepatol. 73 (2020) 1155–1169.
[57] J.E. Yang, W.J. Zhong, J.F. Li, Y.Y. Lin, F.T. Liu, H. Tian, Y.J. Chen, X.Y. Luo, S. M. Zhuang, LINC00998-encoded micropeptide SMIM30 promotes the G1/S transition of cell cycle by regulating cytosolic calcium level, Mol. Oncol. 17 (2023) 901–916.
[58] S. Burbano De Lara, D.D.H. Tran, A.B. Allister, M. Polenkowski, B. Nashan, M. Koch, T. Tamura, C20orf204, a hepatocellular carcinoma-specific protein interacts with nucleolin and promotes cell proliferation, Oncogenesis 10 (2021) 31.
[59] M. Polenkowski, S. Burbano de Lara, A.B. Allister, T.N.Q. Nguyen, T. Tamura, D.D. H. Tran, Identification of novel micropeptides derived from hepatocellular carcinoma-specific long noncoding RNA, Int. J. Mol. Sci. 23 (2021).
[60] H. Zhang, Z. Liao, W. Wang, Y. Liu, H. Zhu, H. Liang, B. Zhang, X. Chen, A micropeptide JunBP regulated by TGF-beta promotes hepatocellular carcinoma metastasis, Oncogene 42 (2023) 113–123.
[61] Y. Sang, J.Y. Liu, F.Y. Wang, X.Y. Luo, Z.Q. Chen, S.M. Zhuang, Y. Zhu, Mitochondrial micropeptide STMP1 promotes G1/S transition by enhancing mitochondrial complex IV activity, Mol. Ther. 30 (2022) 2844–2855.
[62] W. Xu, B. Deng, P. Lin, C. Liu, B. Li, Q. Huang, H. Zhou, J. Yang, L. Qu, Ribosome profiling analysis identified a KRAS-interacting microprotein that represses oncogenic signaling in hepatocellular carcinoma cells, Sci. China Life Sci. 63 (2020) 529–542.
[63] X. Xiang, Y. Fu, K. Zhao, R. Miao, X. Zhang, X. Ma, C. Liu, N. Zhang, K. Qu, Cellular senescence in hepatocellular carcinoma induced by a long non-coding RNA-encoded peptide PINT87aa by blocking FOXM1-mediated PHB2, Theranostics 11 (2021) 4929–4944.
[64] M.H. Xiao, Y.F. Lin, P.P. Xie, H.X. Chen, J.W. Deng, W. Zhang, N. Zhao, C. Xie, Y. Meng, X. Liu, S.M. Zhuang, Y. Zhu, J.H. Fang, Downregulation of a mitochondrial micropeptide, MPM, promotes hepatoma metastasis by enhancing mitochondrial complex I activity, Mol. Ther. 30 (2022) 714–725.
[65] Q. Zhang, T. Wei, L. Yan, S. Zhu, W. Jin, Y. Bai, Y. Zeng, X. Zhang, Z. Yin, J. Yang, W. Zhang, M. Wu, Y. Zhang, L. Liu, Hypoxia-responsive lncRNA AC115619 encodes a micropeptide that suppresses m6A modifications and hepatocellular carcinoma progression, Cancer Res. 83 (2023) 2496–2512.
[66] W. Zheng, Y. Guo, G. Zhang, J. Bai, Y. Song, X. Song, Q. Zhu, X. Bao, G. Wu, C. Zhang, Peptide encoded by lncRNA BVES-AS1 promotes cell viability, migration,

and invasion in colorectal cancer cells via the SRC/mTOR signaling pathway, PLoS One 18 (2023) e0287133.
[67] N. Meng, M. Chen, D. Chen, X.H. Chen, J.Z. Wang, S. Zhu, Y.T. He, X.L. Zhang, R. X. Lu, G.R. Yan, Small protein hidden in lncRNA LOC90024 promotes "cancerous" RNA splicing and tumorigenesis, Adv. Sci. 7 (2020) 1903233.
[68] S. Zhu, J.Z. Wang, D. Chen, Y.T. He, N. Meng, M. Chen, R.X. Lu, X.H. Chen, X. L. Zhang, G.R. Yan, An oncopeptide regulates m(6)A recognition by the m(6)A reader IGF2BP1 and tumorigenesis, Nat. Commun. 11 (2020) 1685.
[69] L.M. Szafron, A. Balcerak, E.A. Grzybowska, B. Pienkowska-Grela, A. Felisiak-Golabek, A. Podgorska, M. Kulesza, N. Nowak, P. Pomorski, J. Wysocki, T. Rubel, A. Dansonka-Mieszkowska, B. Konopka, M. Lukasik, J. Kupryjanczyk, The novel gene CRNDE encodes a nuclear peptide (CRNDEP) which is overexpressed in highly proliferating tissues, PLoS One 10 (2015) e0127475.
[70] B. Zhou, Y. Wu, P. Cheng, C. Wu, Long noncoding RNAs with peptide-encoding potential identified in esophageal squamous cell carcinoma: KDM4A-AS1-encoded peptide weakens cancer cell viability and migratory capacity, Mol. Oncol. 17 (2023) 1419–1436.
[71] S. Wu, L. Zhang, J. Deng, B. Guo, F. Li, Y. Wang, R. Wu, S. Zhang, J. Lu, Y. Zhou, A novel micropeptide encoded by Y-linked LINC00278 links cigarette smoking and AR signaling in male esophageal squamous cell carcinoma, Cancer Res. 80 (2020) 2790–2803.
[72] D.W. Hoskin, A. Ramamoorthy, Studies on anticancer activities of antimicrobial peptides, Biochim. Biophys. Acta 1778 (2008) 357–375.
[73] Y. Pan, Y. Jiang, Y. Cui, J. Zhu, Y. Yu, Truncated lactoferricin peptide controls cervical cancer cell proliferation via lncRNA-NKILA/NF-kappaB feedback loop, Protein Pept. Lett. 29 (2022) 268–280.
[74] D. Huang, J. Chen, L. Yang, Q. Ouyang, J. Li, L. Lao, J. Zhao, J. Liu, Y. Lu, Y. Xing, F. Chen, F. Su, H. Yao, Q. Liu, S. Su, E. Song, NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death, Nat. Immunol. 19 (2018) 1112–1125.
[75] D. Carbonnelle, V. Vignard, D. Sehedic, A. Moreau-Aubry, L. Florenceau, M. Charpentier, W. Mikulits, N. Labarriere, F. Lang, The melanoma antigens MELOE-1 and MELOE-2 are translated from a bona fide polycistronic mRNA containing functional IRES sequences, PLoS One 8 (2013) e75233.
[76] M. Charpentier, M. Croyal, D. Carbonnelle, A. Fortun, L. Florenceau, C. Rabu, M. Krempf, N. Labarriere, F. Lang, IRES-dependent translation of the long non coding RNA meloe in melanoma cells produces the most immunogenic MELOE antigens, Oncotarget 7 (2016) 59704–59713.
[77] S.J. Bray, Notch signalling: a simple pathway becomes complex, Nat. Rev. Mol. Cell Biol. 7 (2006) 678–689.
[78] J.L. Avila, J.L. Kissil, Notch signaling in pancreatic cancer: oncogene or tumor suppressor? Trends Mol. Med. 19 (2013) 320–327.
[79] S. Zhai, J. Lin, Y. Ji, R. Zhang, Z. Zhang, Y. Cao, Y. Liu, X. Tang, J. Liu, P. Liu, J. Lin, F. Li, H. Li, Y. Shi, D. Fu, X. Deng, B. Shen, A microprotein N1DARP encoded by LINC00261 promotes Notch1 intracellular domain (N1ICD) degradation via disrupting USP10-N1ICD interaction to inhibit chemoresistance in Notch1-hyperactivated pancreatic cancer, Cell Discov. 9 (2023) 95.
[80] J. Pan, M. Liu, X. Duan, D. Wang, A short peptide LINC00665_18aa encoded by lncRNA LINC00665 suppresses the proliferation and migration of osteosarcoma cells through the regulation of the CREB1/RPS6KA3 interaction, PLoS One 18 (2023) e0286422.
[81] X. Shi, S. Liao, Z. Bi, J. Liu, H. Li, C. Feng, Newly discovered circRNAs encoding proteins: recent progress, Front. Genet. 14 (2023) 1264606.
[82] J. Xia, S. Li, B. Ren, P. Zhang, Circular RNAs as a potential source of neoepitopes in cancer, Front. Oncol. 13 (2023) 1098523.
[83] M. Zhang, N. Huang, X. Yang, J. Luo, S. Yan, F. Xiao, W. Chen, X. Gao, K. Zhao, H. Zhou, Z. Li, L. Ming, B. Xie, N. Zhang, A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis, Oncogene 37 (2018) 1805–1814.
[84] M. Zhang, K. Zhao, X. Xu, Y. Yang, S. Yan, P. Wei, H. Liu, J. Xu, F. Xiao, H. Zhou, X. Yang, N. Huang, J. Liu, K. He, K. Xie, G. Zhang, S. Huang, N. Zhang, A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma, Nat. Commun. 9 (2018) 4475.
[85] A. Biswas, N. Chowdhury, A. Bagchi, Structural characterization of the hidden peptide SHPRH-146aa encoded by non-coding circ-SHPRH to act as tumor suppressor, Appl. Biochem. Biotechnol. 193 (2021) 2076–2086.
[86] Y. Peng, Y. Xu, X. Zhang, S. Deng, Y. Yuan, X. Luo, M.T. Hossain, X. Zhu, K. Du, F. Hu, Y. Chen, S. Chang, X. Feng, X. Fan, H. Ashktorab, D. Smoot, S.J. Meltzer, G. Hou, Y. Wei, S. Li, Y. Qin, Z. Jin, A novel protein AXIN1-295aa encoded by circAXIN1 activates the Wnt/beta-catenin signaling pathway to promote gastric cancer progression, Mol. Cancer 20 (2021) 158.
[87] X. Zheng, L. Chen, Y. Zhou, Q. Wang, Z. Zheng, B. Xu, C. Wu, Q. Zhou, W. Hu, C. Wu, J. Jiang, A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling, Mol. Cancer 18 (2019) 47.
[88] L. Wang, J. Zhou, C. Zhang, R. Chen, Q. Sun, P. Yang, C. Peng, Y. Tan, C. Jin, T. Wang, J. Ji, Y. Sun, A novel tumour suppressor protein encoded by circMAPK14 inhibits progression and metastasis of colorectal cancer by competitively binding to MKK6, Clin. Transl. Med. 11 (2021) e613.
[89] F. Li, Y. Cai, S. Deng, L. Yang, N. Liu, X. Chang, L. Jing, Y. Zhou, H. Li, A peptide CORO1C-47aa encoded by the circular noncoding RNA circ-0000437 functions as a negative regulator in endometrium tumor angiogenesis, J. Biol. Chem. 297 (2021) 101182.
[90] S.Y. Ying, D.C. Chang, S.L. Lin, The microRNA (miRNA): overview of the RNA genes that modulate gene function, Mol. Biotechnol. 38 (2008) 257–268.
[91] J. Yang, M. Liu, X. Fang, H. Zhang, Q. Ren, Y. Zheng, Y. Wang, Y. Zhou, Advances in peptides encoded by non-coding RNAs: a cargo in exosome, Front. Oncol. 12 (2022) 1081997.
[92] N. Xie, G. Shen, W. Gao, Z. Huang, C. Huang, L. Fu, Neoantigens: promising targets for cancer therapy, Signal Transduct. Targeted Ther. 8 (2023) 9.
[93] M.E. Camarena, P. Theunissen, M. Ruiz, J. Ruiz-Orera, B. Calvo-Serra, R. Castelo, C. Castro, P. Sarobe, P. Fortes, J. Perera-Bel, M.M. Alba, Microproteins encoded by noncanonical ORFs are a major source of tumor-specific antigens in a liver cancer patient meta-cohort, Sci. Adv. 10 (2024) eadn3628.
[94] X. Xia, X. Li, F. Li, X. Wu, M. Zhang, H. Zhou, N. Huang, X. Yang, F. Xiao, D. Liu, L. Yang, N. Zhang, A novel tumor suppressor protein encoded by circular AKT3 RNA inhibits glioblastoma tumorigenicity by competing with active phosphoinositide-dependent Kinase-1, Mol. Cancer 18 (2019) 131.
[95] D. Huang, X. Zhu, S. Ye, J. Zhang, J. Liao, N. Zhang, X. Zeng, J. Wang, B. Yang, Y. Zhang, L. Lao, J. Chen, M. Xin, Y. Nie, P.E. Saw, S. Su, E. Song, Tumour circular RNAs elicit anti-tumour immunity by encoding cryptic peptides, Nature 625 (2024) 593–602.
[96] W. Barczak, S.M. Carr, G. Liu, S. Munro, A. Nicastri, L.N. Lee, C. Hutchings, N. Ternette, P. Klenerman, A. Kanapin, A. Samsonova, N.B. La Thangue, Long non-coding RNA-derived peptides are immunogenic and drive a potent anti-tumour response, Nat. Commun. 14 (2023) 1078.
[97] R. Xiang, L. Ma, M. Yang, Z. Zheng, X. Chen, F. Jia, F. Xie, Y. Zhou, F. Li, K. Wu, Y. Zhu, Increased expression of peptides from non-coding genes in cancer proteomics datasets suggests potential tumor neoantigens, Commun. Biol. 4 (2021) 496.