

## REVIEW ARTICLE OPEN



## Coding, or non-coding, that is the question

Laura Polisenò <sup>1,2,6</sup>, Martina Lanza <sup>1,2,3,6</sup> and Pier Paolo Pandolfi <sup>4,5</sup>

© The Author(s) 2024

The advent of high-throughput sequencing uncovered that our genome is pervasively transcribed into RNAs that are seemingly not translated into proteins. It was also found that non-coding RNA transcripts outnumber canonical protein-coding genes. This mindboggling discovery prompted a surge in non-coding RNA research that started unraveling the functional relevance of these new genetic units, shaking the classic definition of “gene”. While the non-coding RNA revolution was still taking place, polysome/ribosome profiling and mass spectrometry analyses revealed that peptides can be translated from non-canonical open reading frames. Therefore, it is becoming evident that the coding vs non-coding dichotomy is way blurrier than anticipated. In this review, we focus on several examples in which the binary classification of coding vs non-coding genes is outdated, since the same bifunctional gene expresses both coding and non-coding products. We discuss the implications of this intricate usage of transcripts in terms of molecular mechanisms of gene expression and biological outputs, which are often concordant, but can also surprisingly be discordant. Finally, we discuss the methodological caveats that are associated with the study of bifunctional genes, and we highlight the opportunities and challenges of therapeutic exploitation of this intricacy towards the development of anticancer therapies.

*Cell Research* (2024) 0:1–21; <https://doi.org/10.1038/s41422-024-00975-8>

## INTRODUCTION

In his article published in *Nature* in 1970, Francis Crick stated the central dogma of molecular biology: information is passed from DNA (gene) to protein, through messenger RNA (mRNA).<sup>1</sup> Approximately 50 years later, the central dogma still stands. However, it has become just one among many mechanisms through which functional molecules are expressed from our genome.<sup>2</sup>

Over the years it has become more and more evident that our genomic DNA is pervasively transcribed. The portion transcribed into mRNAs, which in turn are translated into ~20,000 proteins, is minimal (2%–5%). Far from being “junk DNA”, most of the rest (75%–90%) is rather transcribed into hundreds of thousands of non-coding RNAs (ncRNAs)<sup>3</sup> that are emerging as sophisticated regulators of gene expression. Together with proteins, they govern embryonic development,<sup>4,5</sup> maintain the physiological state,<sup>6</sup> define organism complexity,<sup>3,7</sup> and are causally linked with hereditary and non-hereditary diseases,<sup>8,9</sup> including cancer.<sup>5,10</sup>

According to their function, ncRNAs are classified into housekeeping ncRNAs and regulatory ncRNAs. Housekeeping ncRNAs (ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs), and small nucleolar RNAs (snoRNAs)) were the first ncRNAs to be identified and studied. They are up to ~4.5 kb long and ubiquitously expressed, as they are involved in essential processes for cell survival. Specifically, rRNAs and tRNAs reside in the cytoplasm and are core components of the mechanism of protein translation. rRNAs represent the RNA component of ribosomes, while tRNAs act as adaptors between the mRNA and amino acids. snRNAs and snoRNAs are instead located in the

nucleus. snRNAs represent the RNA component of spliceosomes. snoRNAs represent the RNA component of small nucleolar ribonucleoprotein particles (snoRNPs), which are responsible for maturation of pre-rRNAs through nucleoside modifications (mainly methylation and pseudouridylation).<sup>11,12</sup> Interestingly, examples are accumulating of rRNAs, tRNAs, snRNAs, and snoRNAs with an oncogenic or tumor-suppressive role in human cancer, precisely because they regulate core mechanisms of gene expression. They are also explored as diagnostic and prognostic biomarkers.<sup>13–15</sup>

Regulatory ncRNAs are further divided according to their lengths. Among short ncRNAs (< 200 nt) there are PIWI-interacting RNAs<sup>16</sup> and microRNAs (miRNAs). Long ncRNAs (lncRNAs, ≥ 500 nt) include long intergenic non-coding RNAs (lincRNAs), pseudogenetic RNAs (PGs), Natural Antisense Transcripts (NATs), and circular RNAs (circRNAs).<sup>5</sup> The expression of regulatory ncRNAs is tightly controlled, they perform a range of functions that is as wide as that performed by proteins, and their dysregulation has been linked to many pathological conditions, including cancer.

miRNAs are 19–25-nt short ncRNAs that inhibit gene expression at post-transcriptional level. The human genome contains hundreds of miRNAs (<https://mirbase.org/>). Primary miRNAs (pri-miRNAs) are transcribed and processed (first in the nucleus, then in the cytoplasm) to precursor miRNAs (pre-miRNAs) and finally to single-strand mature miRNAs, which recruit the Argonaute (AGO) protein to form the miRNA-induced silencing complex (miRISC).<sup>10</sup> miRISC binds target RNAs in correspondence of specific sequences termed miRNA Recognition Elements (MREs). The result is a decrease in target expression through RNA degradation or

<sup>1</sup>Oncogenomics Unit, Core Research Laboratory, ISPRO, Pisa, Italy. <sup>2</sup>Institute of Clinical Physiology, CNR, Pisa, Italy. <sup>3</sup>University of Siena, Siena, Italy. <sup>4</sup>Department of Molecular Biotechnology and Health Sciences, Molecular Biotechnology Center, University of Turin, Torino, Italy. <sup>5</sup>Renown Institute for Cancer, Nevada System of Higher Education, Reno, NV, USA. <sup>6</sup>These authors contributed equally: Laura Polisenò, Martina Lanza. ✉email: [laura.polisenò@cnr.it](mailto:laura.polisenò@cnr.it); [pierpaolo.pandolfi@unito.it](mailto:pierpaolo.pandolfi@unito.it)

Received: 3 January 2024 Accepted: 30 April 2024

Published online: 25 July 2024

translational repression. Each miRNA recognizes hundreds of targets, and each target can be recognized by multiple miRNAs. Therefore, most miRNAs are soft and pervasive tuners of gene expression.<sup>17–19</sup> In addition, as described in “Translation of pri-miRNAs” chapter below, examples exist of miRNA precursors that are translated into functional peptides.

In cancer, miRNA expression and sequence are altered,<sup>20</sup> and, depending on the repressed targets, they act as oncogenes (oncomiRs) or as tumor suppressors.<sup>21–23</sup> Interestingly, there are cases in which the same miRNA behaves as oncogene in some cancer types and tumor suppressor in other cancer types. This attests the pervasive nature of miRNA-mediated regulation of gene expression, with net results depending on the balance among all targets at play.<sup>24</sup> miRNAs are also extensively used as diagnostic, prognostic, and predictive biomarkers. Since they can be detected in body fluids such as blood, saliva and urine, circulating miRNAs represent in fact a much less invasive alternative to the classic biopsy.<sup>25–27</sup> Finally, they are exploited therapeutically (inhibition of oncomiRs and replacement of tumor-suppressive miRNAs).<sup>28</sup>

Relevant to the topic of this review, we discovered a new mechanism for the regulation of miRNA activity: miRNAs are negative regulators of target RNA expression, but in turn target RNAs are negative regulators of miRNA function. This is because they compete for binding to common miRNAs, i.e., they act as competing endogenous RNAs (ceRNAs). ceRNA partners that share MRE(s) for the same miRNA(s) dilute the miRNA(s), blunting their efficacy and at the same time sustaining their own expression.<sup>29</sup> ceRNA-based regulation can occur only if some conditions are satisfied: ceRNA partners and the shared miRNA(s) need to be present in the same subcellular compartment and to be expressed at similar level. Furthermore, the sponging effect gets stronger at the increase of the number and the affinity of shared MREs.<sup>30–38</sup> Experimental evidence indicate that ceRNA-based regulation is common among coding (see “The 3’UTR exerts non-coding functions” chapter below) and all classes of non-coding RNAs. Furthermore, it can produce extended networks, which get heavily disrupted in cancer.<sup>39–47</sup>

The human genome contains more than 30,000 lncRNA genes, which are expressed into more than 100 thousand transcripts ([https://www.genecodegenes.org/human/stats\\_45.html](https://www.genecodegenes.org/human/stats_45.html))<sup>5,48</sup>. Although the mechanism of transcription initiation and termination can be different from those of mRNAs,<sup>49</sup> most lncRNAs are transcribed by Pol II. Nevertheless, examples exist of lncRNAs transcribed by Pol I or Pol III. Furthermore, many lncRNAs undergo splicing, but they can be transcribed from single-exon genes as well.<sup>50</sup> Analogously, they may or may not undergo 5’ capping and 3’ polyadenylation.<sup>5</sup> Besides transcription mechanisms, heterogeneity extends to many other features of lncRNAs. According to the position of their genes, they are grouped into lincRNAs, intronic lncRNAs or pseudogenes, while according to the direction of their transcription they are divided into sense transcripts and NATs. Furthermore, all lncRNAs are linear transcripts, except for circRNAs, while subcellular localization defines nuclear vs cytoplasmatic lncRNAs.<sup>5,51</sup>

To date, just a small fraction of all lncRNAs have been studied, and what we already know might just be the tip of the iceberg. In any case, they appear as flexible molecules that organize in thermodynamically-stable secondary or even higher-order structures. Acting as signals, guides, decoys, or scaffolds, they regulate virtually each step of gene expression: nuclear organization and genome integrity, chromatin remodeling by epigenetic modification, transcription, RNA splicing and processing, mRNA stability and translation, protein post-translational modification (e.g., phosphorylation), subcellular localization, and activity.<sup>5,52–55</sup>

lncRNAs are altered in cancer and, depending on their mechanisms of action and on the effectors involved, they can act either as oncogenes or as tumor suppressors. Furthermore, in

the last few years lncRNAs have taken center stage as diagnostic, prognostic and predictive biomarkers. Crucially, oncogenic lncRNAs are currently explored as therapeutic targets, using approaches based on antisense oligonucleotides, RNA interference, or CRISPR/Cas9 technology.<sup>56,57</sup> Furthermore, lncRNAs encoding tumor-specific antigens can be exploited as anticancer vaccines.<sup>58</sup>

Among linear lncRNAs, lincRNAs are the most conspicuous group and include some of the most well-studied examples, in development and in cancer.<sup>59</sup> lincRNA *X-inactive specific transcript* (*Xist*), which is responsible for X chromosome inactivation and gene dosage compensation, was discovered in the early 90s and is the first long non-coding RNA to be ever studied.<sup>60</sup> Fifteen years later, *HOX antisense intergenic RNA* (*HOTAIR*) was discovered. This lincRNA is transcribed from an independent promoter located in antisense strand within the *HOXC* locus, and it causes epigenetic silencing of the *HOXD* locus.<sup>61</sup> In cancer, examples of well-studied lincRNAs include: *BRAF-Activated Non-protein Coding RNA* (*BANCR*)<sup>62</sup>, *linc-Regulator Of Reprogramming* (*linc-RoR*)<sup>63</sup>, *Metastasis-Associated Lung Adenocarcinoma Transcript 1* (*MALAT-1*)<sup>50</sup>, and *Survival Associated Mitochondrial Melanoma Specific Oncogenic Non-coding RNA* (*SAMMSON*)<sup>64</sup>. These are pleiotropic lincRNAs that exert oncogenic functions in multiple cancer types by affecting key cellular processes (motility, stemness, chemosensitivity, and mitochondrial metabolism, respectively). Mechanistically, lincRNAs have the activities listed above for lncRNA. Interestingly, they can also be translated into functional peptides, as we describe in “Translation of lincRNAs” chapter below.

A group of linear lncRNAs with peculiar features is represented by NATs. The development of high-throughput sequencing has shed light on this class of RNA molecules that are transcribed from the opposite DNA strand. They originate from bidirectional promoters shared with the corresponding sense transcripts, from independent antisense promoters, or from latent antisense promoters located within sense transcriptional units. According to the degree of overlap with sense transcripts, NATs are defined as head-to-head (the overlap is in the 5’ region), tail-to-tail (the overlap is in the 3’ region) or embedded (the overlap is complete). Finally, *cis*-NATs exert their function on their own genomic locus, while *trans*-NATs act on other genomic loci. In cancer, several NATs are known to be involved in the regulation of gene expression at multiple levels: in the nucleus, epigenetic modification, transcription, RNA splicing and processing; in the cytoplasm, mRNA stability and translation, post-translational modification of proteins. As such, they play oncogenic or tumor-suppressive roles and they are highly valued as diagnostic, prognostic, and predictive biomarkers.<sup>65,66</sup>

Another group of linear lncRNAs with peculiar features is represented by pseudogenes. With the name “pseudogene” we refer to a region of the genome that contains a defective copy of a parental protein-coding gene. Indeed, pseudogenes are characterized by the presence of mutations, deletions or insertions that lead to frameshifts and prevent translation of parental protein products. The human genome contains ~14,000 pseudogenes, 10% of which are transcribed ([https://www.genecodegenes.org/human/stats\\_45.html](https://www.genecodegenes.org/human/stats_45.html))<sup>67</sup>.

According to their origin, pseudogenes are classified as follows. Processed pseudogenes, which represent the most abundant class, derive from a retrotransposition event. They do not contain introns, are located on different chromosomes compared to parental genes, and are subjected to a distinct regulation of gene expression. They accumulate alterations because retrotranscription is error prone. Nonprocessed pseudogenes derive from gene duplication. They are located on the same chromosome as parental genes and retain introns, as well as regulatory regions (promoter). They accumulate alterations because the sequence of the parental copy is the only one preserved under selective pressure. Finally, unitary pseudogenes arise from the

accumulation of alterations in an ancestral protein-coding gene that has no other copy in the genome.<sup>68,69</sup>

Even though they have been considered functionless for a long time, in the recent years we and others have contributed to discovering that pseudogenes exert a wide range of parental gene-related as well as parental gene-unrelated functions. Specifically, PGs are involved in chromatin remodeling, sponging of miRNAs and RNA Binding Proteins (RBPs), and mRNA degradation through endosRNAs.<sup>68,70–74</sup> As described in “Translation of PGs” chapter below, they are also translated into functional peptides and proteins.

The contribution of pseudogenes to cancer initiation and progression is gaining momentum. There are in fact several examples of oncogenic and tumor-suppressive pseudogenes.<sup>69</sup> Interestingly, with their random “landing” upon retrotranscription, processed pseudogenes acquired somatically can potentially disrupt otherwise functional genetic units and therefore can be considered mutagenic factors.<sup>75</sup> PGs are also highly valuable as diagnostic and prognostic biomarkers.<sup>76</sup> Oncogenic PGs are explored as therapeutic targets,<sup>77,78</sup> while PGs that act as sponges of oncogenic miRNAs are envisioned as drugs.<sup>73</sup> Finally, PGs that encode immunogenic peptides could reveal effective as anticancer vaccines.<sup>73,79</sup>

The ~25,000 circRNAs constitute the group of non-linear lncRNAs.<sup>80</sup> circRNAs are single-stranded, covalently-closed RNA molecules characterized by high stability since they are immune from exonucleases activity. circRNAs are mostly generated by back-splicing of pre-mRNAs. Contrary to canonical splicing in which an upstream 5′ splice donor is joined with a downstream 3′ splice acceptor, back-splicing is an unconventional splicing event in which an upstream 3′ splice acceptor is joined with a 5′ downstream splice donor, leading to the formation of a circular-shaped structure. As a result of this peculiar splicing mechanism, circRNAs are classified as follows. Exonic circRNAs (EcircRNAs) are composed entirely by exons. They are the largest subclass of circRNAs (they account for ~85% of all circRNAs) and are mainly located in the cytoplasm. Conversely, Exonic-Intronic circRNAs (ElcircRNAs) are composed both by exons and by introns and are mainly retained in the nucleus. In alternative to back-splicing, circRNAs can also originate from lariat introns that fail to undergo debranching and are subsequently subjected to trimming of the lariat tail. Because of such biogenesis, circular intronic RNAs (ciRNAs) are composed only by introns and are mainly retained in the nucleus.<sup>81</sup>

The molecular functions attributed to circRNAs are in line with their subcellular localization. Nuclear ElcircRNAs and ciRNAs mainly act *in cis* as regulators of transcription or splicing of their own gene. Conversely, cytoplasmic EcircRNAs mainly act *in trans*: as sponges for miRNAs and for RBPs, they play a crucial role in the post-transcriptional regulation of gene expression; they form circRNPs that modulate signaling pathways; as described in “Translation of circRNAs” chapter below, they can also be translated into functional peptides.<sup>82,83</sup>

There are several examples of oncogenic or tumor-suppressive circRNAs in cancer.<sup>84</sup> They have gained attention in the context of resistance to traditional chemotherapeutic drugs, as well as targeted and immunotherapy approaches.<sup>85,86</sup> Extensively studied as diagnostic, prognostic, and predictive biomarkers,<sup>83,87</sup> circRNAs are the focus of many therapeutic strategies as well. Oncogenic circRNAs are targeted through antisense oligonucleotides, RNA interference, or CRISPR/Cas9 technology. In addition, their exceedingly high stability has prompted testing of circRNAs as drugs: synthetically engineered circRNAs are currently under evaluation as sponges for oncogenic miRNAs,<sup>83,87,88</sup> while circRNAs encoding tumor-specific antigens can be envisioned as anticancer vaccines.<sup>89,90</sup>

Still overwhelmed by the discovery of pervasive genome transcription, in the most recent years we have been hit by a

counter-wave of pervasive translation. In the last decades, the ~20,000 proteins encoded by our genome ([https://www.genecodegenes.org/human/stats\\_45.html](https://www.genecodegenes.org/human/stats_45.html)) have been the undivided focus of cancer research. The most potent oncogenes and tumor suppressors known so far (e.g., BRAF and c-MYC vs p53 and PTEN) are in fact proteins. They have almost completely monopolized *in vivo* cancer modeling,<sup>91</sup> and they are the targets of most of the current anticancer therapies, using synthetic small molecules<sup>92</sup> or even other proteins (e.g., antibodies<sup>93</sup>). Considering the profound and vast knowledge on protein medicinal chemistry that we have accumulated, the recent discovery that several ncRNAs can be translated into peptides was good news to many. Even more recently, additional open reading frames (ORFs) have been found to reside within the 5′ untranslated region (UTR) of hundreds of mRNAs. These “upstream ORFs (uORFs)” are short and contribute to regulating the translation of the longer coding sequences (CDSs) located downstream.<sup>94–98</sup>

With this mindboggling complexity in mind, in this review we aim to provide an overview of cases in which the categorization of “coding gene” vs “non-coding gene” is outdated, since the same gene produces coding and non-coding elements.<sup>99</sup> We also discuss the implications of this parsimonious usage of genetic units, in terms of independent or interdependent regulation of expression and concordant or discordant biological output in cancer. In addition, we point out the methodological advancements that are required to study bifunctional genes, with particular emphasis on the importance of *in vivo* modeling. Finally, we highlight the opportunities and challenges associated with the therapeutic implications of bifunctional genes in the development of anticancer therapies.

## BIFUNCTIONAL GENOMIC LOCI EXPRESS mRNAs AND ncRNAs

There are genomic loci that can be rightfully considered both coding and non-coding because they express both an mRNA and a ncRNA, in a mutually exclusive or coexisting fashion (Fig. 1). Such genomic loci are defined “hybrid” or “bifunctional”, and representative cancer-relevant examples are described below.

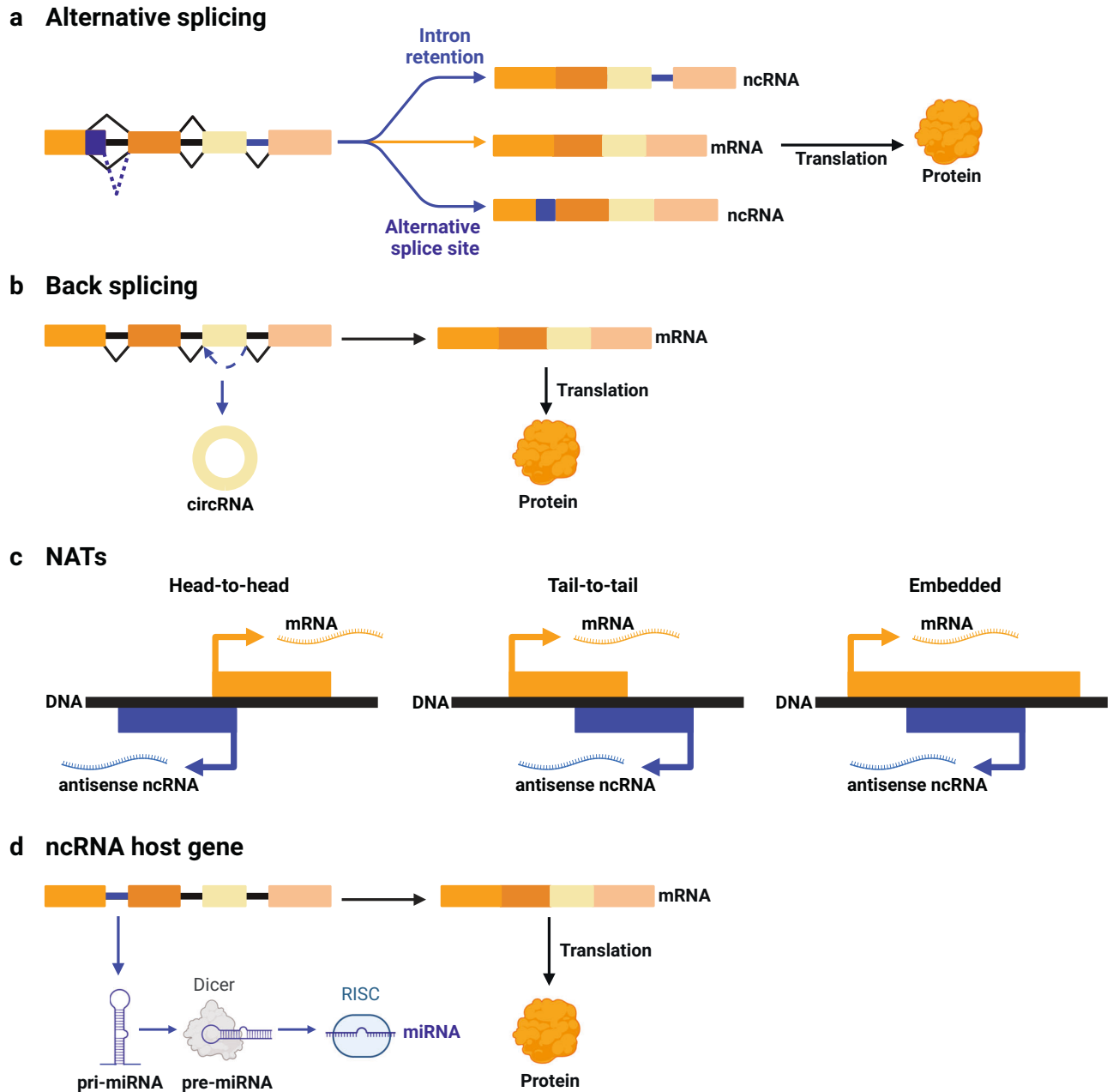
### Expression of an mRNA and a ncRNA depending on the exons spliced together

Alternative splicing (AS) is one of the mechanisms through which cells express more than one transcript from the same genomic locus, thereby enriching their transcriptome (and proteome). Alternative splicing comes in five main flavors (exon skipping, alternative 5′ splice site, alternative 3′ splice site, mutually exclusive exons, intron retention) and it can occur in conjunction with the use of alternative transcription start sites (promoters), or alternative polyadenylation sites.<sup>100–102</sup>

Several examples have been reported of genomic loci that undergo alternative splicing and end up expressing a coding transcript and a non-coding transcript with a role in cancer (Fig. 1a).<sup>103</sup> Two of such examples are described below.

The *Steroid Receptor RNA Activator 1 (SRA1)* gene was the first bifunctional gene to be characterized. It expresses Steroid Receptor co-Activator Protein (SRAP), an oncogenic protein that acts as positive transcriptional regulator of steroid receptors. In breast cancer SRAP is overexpressed and associated with worse prognosis,<sup>104</sup> while in prostate cancer it has been shown to potentiate the activity of the androgen receptor.<sup>105</sup> Due to intron 1 retention,<sup>106</sup> the *SRA1* gene expresses *lincRNA-SRA* as well. Oncogenic *lincRNA-SRA* is overexpressed in melanoma. In this context, it induces cell proliferation, migration, invasion, epithelial to mesenchymal transition (EMT), as well as metastasis in a xenograft model.<sup>107</sup> However, the molecular mechanism(s) of its action remain to be elucidated.

The *PPP1R10* gene expresses Protein Phosphatase-1 (PP-1) Nuclear Targeting Subunit (PNUTS), a ubiquitous nuclear



**Fig. 1 Bifunctional genomic loci express mRNAs and ncRNAs.** Genomic loci are defined bifunctional when they can be considered both coding (orange) and non-coding (blue), because they express both an mRNA and a ncRNA, through one of the following mechanisms. **a** The ncRNA is the product of alternative splicing, for example through the retention of introns or the choice of an alternative splice site. **b** In back-splicing, the joining of an upstream 3' splice acceptor with a 5' downstream splice donor leads to the production of a non-coding circRNA. **c** NATs are non-coding RNA molecules transcribed from the opposite DNA strand. According to the degree of overlap with sense coding mRNAs, NATs are defined as head-to-head (the overlap is in the 5' region, left), tail-to-tail (the overlap is in the 3' region, middle) or embedded (the overlap is complete, right). **d** Exons compose the mRNA that is translated into a protein, while introns compose a non-coding RNA such as a miRNA. The ncRNA is called "intragenic", while the coding gene in which it resides is called "host gene".

protein that binds to PP-1. In turn, PP-1 is a serine/threonine phosphatase mainly involved in chromosome decondensation at mitosis exit.<sup>108</sup> PNUTS is a positive modulator of PP-1 activity, thereby favoring the re-entry into interphase.<sup>109</sup> It is also involved in DNA damage repair.<sup>110</sup> In the context of breast cancer, *PPP1R10* gene expresses *lncRNA-PNUTS* as well, due to an alternative 3' splice site that is located in exon 12 and breaks the ORF encoding PNUTS protein. *lncRNA-PNUTS* expression is under the regulation of TGF- $\beta$ : through AKT2-dependent phosphorylation, TGF- $\beta$  causes the release of hnRNP E1 splicing repressor from a TGF- $\beta$  Activated Translational (BAT) element that is

positioned at the alternative splice site. In this way, alternative splicing can occur and *lncRNA-PNUTS* is expressed. In turn, *lncRNA-PNUTS* acts as an effector of TGF- $\beta$ -induced EMT, because it sponges ZEB1-targeting miR-205. *lncRNA-PNUTS* silencing is in fact associated with decreased tumor initiation and metastasis in a xenograft model.<sup>111</sup>

Among common mechanisms through which coding and non-coding transcripts are expressed from the same genomic locus, there is also back-splicing (Fig. 1b). There are several examples of genomic loci expressing a linear coding mRNA and a circular non-coding RNA with roles in cancer.<sup>84</sup>

Specifically, we have attributed an oncogenic role to circular Pokéon (circPOK) in the context of mesenchymal tumors. circPOK is an EcircRNA generated from the *Zbtb7a* gene through back-splicing of exon 2. It promotes tumorigenesis by binding to the complex formed by Interleukin enhancer binding Factor 2 and 3 (ILF2/3), hence sustaining the transcription/stability of multiple mRNAs encoding interleukins and angiogenic factors.<sup>112</sup> Interestingly, the linear transcript expressed from *Zbtb7a* gene encodes *Zbtb7a/Pokémon/Lrf*, a transcriptional repressor that acts as a tumor suppressor of mesenchymal tumorigenesis, by promoting the differentiation of mesenchymal stem cells.<sup>113</sup> The discordant functions of the linear coding transcript vs the circular non-coding transcript are confirmed by the opposite trend observed in their expression levels when mesenchymal tumors are compared to normal tissues: Pokémon levels are lower, while circPOK levels are higher. The decrease in Pokémon levels is due to enhanced post-transcriptional regulation by oncogenic miRNAs,<sup>112</sup> while the mechanism behind the increase in circPOK remains to be established.

### Expression of a sense mRNA and an antisense ncRNA

In cancer, there are quite a few examples of genomic loci that are transcribed from both DNA strands and end up expressing a sense coding mRNA and an antisense non-coding RNA, i.e., an NAT (Fig. 1c).<sup>65,66</sup> Two of such examples are described below.

As an example of *cis*-NAT, we highlight the oncogenic *ZEB1-AS1*. Zinc finger E-box-Binding homeobox 1 (*ZEB1*) is an oncogenic transcription factor overexpressed in many epithelial tumor types. *ZEB1* is a master regulator of EMT, promoting migration/invasion in vitro and metastasis in vivo. It also confers resistance to chemotherapy.<sup>114</sup> *ZEB-AS1* is transcribed in antisense orientation from the promoter region of *ZEB1* and in turn is involved in a positive feedback loop, sustaining *ZEB1* expression at two levels. It promotes *ZEB1* transcription epigenetically. This is because it can recruit Mixed Lineage Leukemia 1 (MLL1) histone methyltransferase on the *ZEB1* promoter, therefore ensuring that H3 histone gets methylated at K4 (H3K4me3) and that DNA is accessible for transcription.<sup>115</sup> In addition, *ZEB-AS1* sponges several anti-EMT miRNAs, including miR-200 family and miR-205, that in turn target *ZEB1*.<sup>116,117</sup>

As an example of *trans*-NAT, we mention the tumor-suppressive *CDR1as*. *Cerebellum Degeneration-Related antigen 1* (*CDR1*) is a low-expressed, poorly-characterized one-exon gene. *LINC00632* is a 5-exon lincRNA transcribed in antisense orientation compared to *CDR1*. *CDR1as* is generated by back-splicing of the 5th exon of *LINC00632*, which fully overlaps with *CDR1*.<sup>118</sup> *CDR1as* is highly expressed in the brain, where it belongs to a sophisticated network that tightly regulates miR-7 level/activity and is involved in correct embryonic development.<sup>119–124</sup> Conversely, *CDR1as* is expressed at low levels in all other tissues, except for melanocytes that share neural crest origin with the brain.<sup>125</sup> In cancer, *CDR1as* is reported as downregulated both in gliomas and in melanoma, and it has been implicated in the regulation of protein stability and function.<sup>126</sup> In gliomas, it stabilizes p53 by binding to it, hence preventing the binding of MDM2 and the consequent proteasome-dependent degradation.<sup>127</sup> In melanoma, *CDR1as* inactivates the pro-metastatic IGF2BP3 protein by sponging it.<sup>128</sup>

### Expression of an mRNA from exons and of a ncRNA from introns of the same pre-mRNA

Bifunctional genomic loci exist where exons compose the mRNA that is translated into a protein, while introns compose a ncRNA. Such ncRNAs are called “intragenic” and the coding genes in which they reside are referred to as “host genes”.<sup>129</sup>

This arrangement is very common with miRNAs (Fig. 1d). Hundreds of miRNA genes are in fact located within introns of protein-coding genes. They can be transcribed from their own promoters as independent transcriptional units, but more

commonly transcription starts from the promoter of the host gene and produces a bifunctional pre-mRNA: exons are spliced together into the mRNA, while the intronic pri-miRNA is further processed into mature miRNA(s). Consequently, miRNA(s) and host gene are co-expressed and co-involved in regulatory circuits.<sup>130–132</sup>

miR-106~25 cluster of miRNAs is composed of miR-106b, miR-93 and miR-25. It is located in the 13th intron of *Mini-Chromosome Maintenance protein 7* (*MCM7*) gene and is produced through pre-mRNA splicing. Both the miRNA cluster and the protein are aberrantly overexpressed in prostate cancer. MCM7 protein promotes the initiation of genome replication, while miR-106b~25 is responsible for the decrease of PTEN levels by binding to 3'UTR of *PTEN*. By generating a vector that drives the expression of *MCM7* exons as well as intron 13 (*PIG/MCM7i13*), we demonstrated that MCM7 protein and the miR-106~25 cluster can transform mouse embryonic fibroblasts in vitro. We then moved into the in vivo setting.

The *MCM7i13* construct was placed under the control of the prostate-specific rat Probasin promoter *ARR<sub>2</sub>PB* (*Pb/MCM7i13*). Nine transgenic lines were analyzed and clustered into low expressor (LE), medium expressor (ME) and high expressor (HE), according to the degree of overexpression of MCM7, miR-106b, miR-93 and miR-25. By analyzing a cohort of 1-year-old transgenic mice belonging to the various lines, we observed that the dorsolateral prostates (DLPs) displayed multifocal lesions whose severity correlated with transgene expression levels: while LE mice showed just hyperplasia, ME and even more HE mice showed typical histological features of prostatic intraepithelial neoplasia (PIN). In the DLPs of HE transgenic mice aberrant proliferation was confirmed by increased number of Ki67-positive cells, while the downregulation of *Pten* mRNA levels together with hyperactivation of Akt signaling was consistent with the activity of miR-106~25 as a *Pten*-targeting cluster. These data confirm that the MCM7 protein and the miR-106~25 cluster are 2 hits that can initiate prostate tumorigenesis in a tissue-specific mouse model. Interestingly, *Pb/MCM7* transgenic mice were generated as well. However, no signs of PIN were visible up to 1 year of age. This further confirms that MCM7 protein cannot initiate prostate tumorigenesis alone and that miR-106~25 non-coding cluster is indispensable.<sup>133</sup>

In conclusion, all these specific examples concur to highlight the complexity of gene expression and demonstrate that a single genomic locus can exert both coding functions, through the transcription of an mRNA and the synthesis of the corresponding protein, and non-coding functions, by expressing a ncRNA, with important biological outcomes.

## BIFUNCTIONAL mRNAs EXERT NON-CODING FUNCTIONS AS WELL

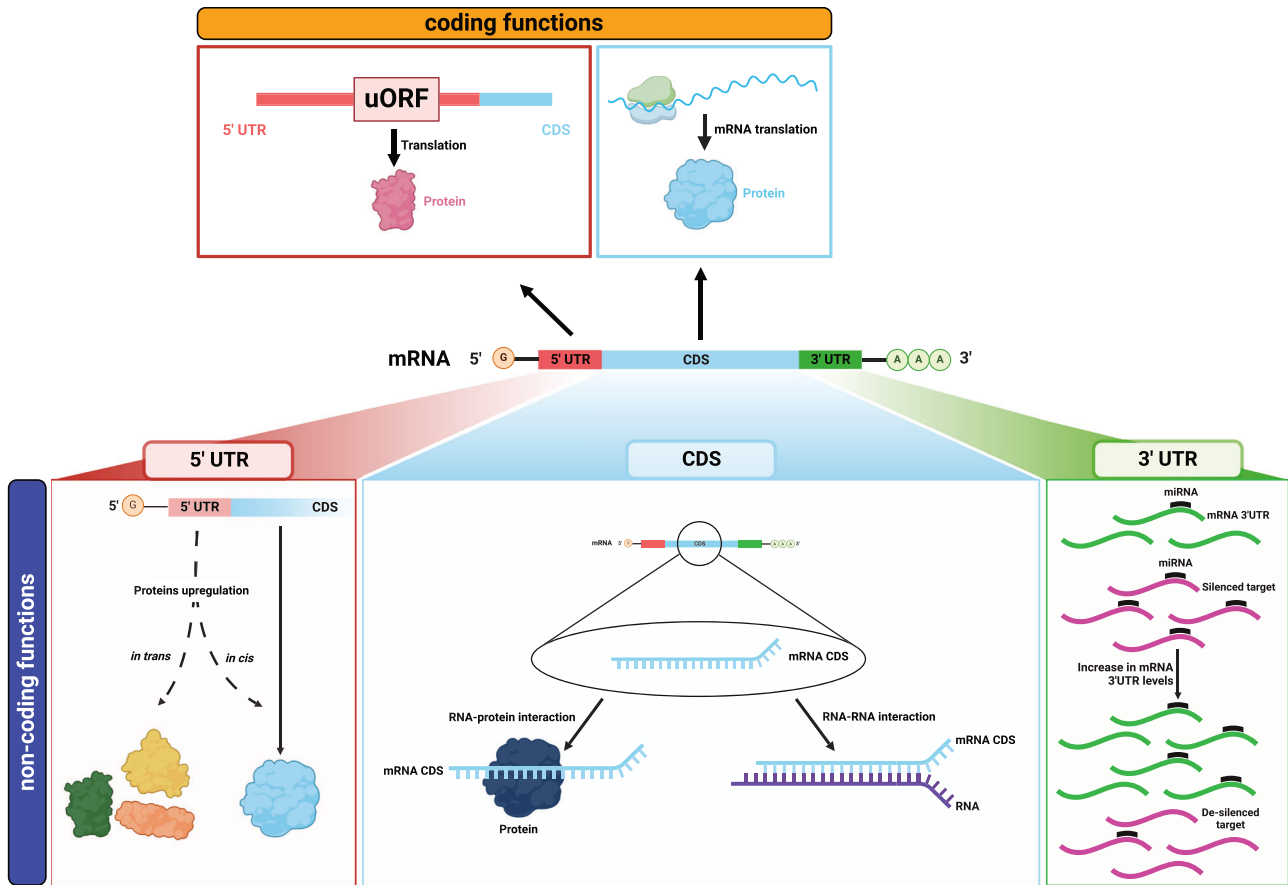
mRNAs are protein-coding RNAs, as their fate is to be exported from the nucleus to the cytoplasm, where they are translated into proteins. However, literature provides extensive evidence that mRNAs exert non-coding functions as well.

Here, we present examples of “bifunctional mRNAs” and of the role that they play in cancer because of the non-coding activities of their 5'UTR, CDS and 3'UTR (Fig. 2).

### The 5'UTR exerts non-coding functions

Two clear examples of 5'UTRs with non-coding functions are represented by the *c-Myc* and *Vascular Endothelial Growth Factor* (*VEGF*) genes. One (*c-MYC* P0 5'UTR) behaves as a tumor suppressor and the other (*VEGF* 5'UTR) as an oncogene.

The proto-oncogene *c-MYC* promotes cell proliferation and tumorigenesis. Although the most expressed *c-MYC* mRNAs are transcribed from the P1 and P2 transcription start sites, 5% of *c-MYC* transcripts are expressed from the upstream P0 transcription start site. P0 transcript is under the control of its own



**Fig. 2 Bifunctional mRNAs exert coding functions and non-coding functions.** In mRNA molecules (middle) three regions can be identified: the 5'UTR (red), the CDS (light blue) and the 3'UTR (green). (Top, orange) mRNAs are primarily protein-coding RNA molecules: they carry a primary ORF (the CDS), and they may also present a short uORF in the 5'UTR. (Bottom, blue) non-canonical non-coding functions have been attributed to mRNAs. The 5'UTR can exert non-coding functions *in cis* or *in trans*, by interacting with proteins. The CDS can be involved in non-coding RNA–protein or RNA–RNA interactions. The 3'UTR can exert non-coding functions within the concept of ceRNAs: due to MREs, 3'UTRs can sponge miRNAs, leading to the de-silencing of ceRNA partners that share MREs for the same miRNAs.

promoter and differs from the predominant P1 and P2 transcripts because of an extended 5'UTR (~640-nt extension). Blume et al.<sup>134</sup> discovered that the overexpression of P0 5'UTR strongly decreases the anchorage-independent growth of HeLa cells *in vitro*, and their ability to form tumors when xenografted in nude mice. Digging into the molecular mechanism underlying this phenomenon, they found that P0 5'UTR overexpression is associated with the upregulation of the c-MYC2 (p64) protein, which is translated from the P2 transcript. Such upregulation results in increased apoptotic cell death, likely a failsafe mechanism triggered in the presence of excessive oncogenic signaling.<sup>135</sup> Since no increase was observed in the level of the endogenous P2 transcript, the authors hypothesized that P0 5'UTR might work *in cis* by affecting the translation of the abovementioned c-MYC2 protein isoform, either directly through P0–P2 RNA interactions or indirectly through RBPs. It also remains to be elucidated which transcription factors control transcription from P0 vs P1/2 start site and how the choice is regulated in physiological and pathological conditions.

VEGF plays a critical role in tumorigenesis through promotion of neoangiogenesis.<sup>136</sup> However, the susceptibility of HCT116 colon cancer cells to chemotherapeutic agents (5-fluorouracil, etoposide and doxorubicin) is not fully rescued by recombinant VEGF protein, which raises the possibility that VEGF mRNA exerts coding-independent functions. Masuda et al.<sup>137</sup> unveiled that the overexpression of VEGF 5'UTR increases the anchorage-independent growth of HCT116 cells *in vitro*, and their ability to form tumors when xenografted in athymic nude mice. These

tumors were profiled by microarray analysis, showing the upregulation of anti-apoptotic genes and downregulation of pro-apoptotic genes. Specifically, the authors observed a decrease in IFN $\alpha$ /STAT1-dependent pro-apoptotic signaling pathway, which resulted in decreased sensitivity of xenografted tumors to 5-fluorouracil, a chemotherapeutic agent that is known to elicit an IFN $\alpha$ /STAT1-dependent pro-apoptotic response. The tumor-promoting function was mapped to a 270 nt-long region located between 475 nt and 745 nt of VEGF 5'UTR and persisted when cells were treated with the translation inhibitor cycloheximide, which proves that it is not related to protein synthesis. However, the exact molecular mechanism still needs to be elucidated. Considering the wide impact on transcriptome and the strong biological effects, the authors speculate that VEGF 5'UTR works *in trans* as a regulatory RNA that affects the expression/functions of a wide network of target proteins.<sup>137</sup>

### The CDS exerts non-coding functions

The CDS is the region of the mRNA whose fate is to be translated, as its nucleotide sequence instructs the amino acid sequence of the protein, and that is why mRNAs are transcripts that primarily exert a coding function. However, there are cancer-relevant examples by which the CDS exerts a non-coding function as well.

The p53 transcription factor is a tumor suppressor protein that prevents cancerous transformation, mainly acting as a keeper of genome integrity.<sup>138</sup> Under normal conditions p53 is kept inactive, while under stress conditions it gets rapidly activated. This is

possible due to the tight control exerted by the MDM2 protein on p53. MDM2 binds to p53 and, functioning as an E3 ligase, triggers the ubiquitination of p53 C-terminal domain, and consequent proteasome-mediated degradation in the cytoplasm. Under stress conditions, MDM2 quickly releases its hold, so that p53 level increases, gets localized to the nucleus, and unleashes its activity as a transcription factor. Consequently, cells stop cycling and either repair the damage or die by apoptosis. Once damage is repaired, multiple negative feedback loops place p53 back under MDM2 control, so that the permanent blockage of cell functions is avoided.<sup>139</sup>

Extensive literature has unveiled the mechanisms through which p53 level increases in response to stress. DNA damage induces multiple post-translational modifications of p53 protein (phosphorylation and acetylation) that weaken its interaction with MDM2. It also induces p14/ARF that inactivates MDM2 by sequestering it into the nucleolus. Finally, relevant to the main topic of this review, the DNA damage sensor ATM phosphorylates MDM2 at Ser395. This induces a conformational change in the protein and favors its interaction with a specific sequence within the CDS of *p53* mRNA. Because of this interaction, MDM2 switches from a negative to a positive regulator of p53: it cannot function as E3 ligase any longer, and it rather promotes *p53* mRNA translation, contributing to the increased levels of p53 protein.<sup>140–142</sup> Noteworthy, the BOX-I domain, which is the domain used by p53 protein to interact with MDM2 protein, is translated from the sequence used by *p53* mRNA to interact with MDM2 protein. Furthermore, the BOX-I domain is the most conserved region of p53 protein and has co-evolved with MDM2 protein. This attests that the coding and the non-coding functions of *p53* mRNA are equally important to ensure the fine regulation of p53 activation/deactivation.<sup>143</sup>

In normal somatic cells, progressive telomere shortening along with cell divisions results in replicative senescence. Conversely, cancer cells aberrantly express the Telomerase Reverse Transcriptase (TERT) enzyme that forms a ribonucleoprotein (RNP) with Telomerase RNA (TR) and uses it as a template to extend telomere length, counteracting telomere erosion. Since TR level exceeds TERT level, TERT-independent functions have been postulated for TR. Ivanyi-Nagy et al.<sup>144</sup> used a pull-down approach to map the RNA interactome of TR. Among the candidates further studied, *HIST1H1C*, the mRNA of H1.2 linker histone subtype, was found to exert a non-coding activity through its CDS. Within *HIST1H1C* CDS, nucleotides 334–348 form an RNA duplex with TR and were named Telomerase RNA InterActing Genetic Element (TRIAGE). The direct interaction of TRIAGE with TR does not affect TERT enzymatic activity, and yet it has a negative impact on telomere elongation. The authors hypothesize that, by base pairing with TR, TRIAGE sponges TERT RNP away from telomeres, impairing their elongation.

### The 3'UTR exerts non-coding functions

It is well established that 3'UTRs exert regulatory functions and have an impact on tumorigenesis. For example, we have recently reported in a zebrafish model of melanoma that the presence of the 3'UTR impacts on BRAFV600E-driven tumorigenesis. The strong melanoma driver effect of the CDS of reference BRAFV600E (BRAFV600E-ref) is in fact suppressed in the presence of the corresponding 3'UTR.<sup>145</sup>

3'UTRs are post-transcriptional regulators of gene expression, and they exert both *in cis* and *in trans* functions. *In cis* functions are mainly coding-related, as they involve the regulation of stability, localization, and translation of the mRNA itself.<sup>146,147</sup> By contrast, the mechanism through which 3'UTRs exert their non-coding function *in trans* falls within the concept of mRNAs as sponges for miRNAs, i.e., as ceRNAs. Hundreds of papers about this topic have been published in the last decade, most of which are in the context of cancer. Indeed, when 3'UTR-dependent functions

are at play, the overexpression of an oncogenic mRNA or the downregulation of a tumor-suppressive mRNA promotes cancer initiation and/or progression not only because of the encoded protein, but also because of the ceRNA activity of its 3'UTR. Three examples are described below.

In breast cancer the *CXCR4* mRNA promotes metastasis not only through its protein, the CXCR4 chemokine receptor, but also through the *CXCR4* 3'UTR that sponges the tumor-suppressive miR-146a, thus leading to the upregulation of TRAF6 and EGFR, two oncoproteins that activate the NF- $\kappa$ B pathway.<sup>148</sup>

The *PTEN* mRNA is renowned to suppress oncogenic PI3K/AKT signaling pathway through the PTEN protein, a phosphatase that dephosphorylates phosphatidylinositol-3,4,5-trisphosphate (PIP3) to phosphatidylinositol-4,5-bisphosphate (PIP2) and in so doing prevents downstream AKT activation.<sup>149</sup> However, the *PTEN* mRNA exerts a tumor-suppressive function also by sponging oncogenic miRNAs, and hence by sustaining the expression of tumor-suppressive ceRNA partners involved in other signaling pathways.<sup>70,150,151</sup> In turn, the ceRNA network of mRNAs that sustains the expression of the *PTEN* mRNA, and hence the inhibition of the PI3K/AKT signaling pathway, is even more extended<sup>34,150–155</sup> and fatally gets heavily affected by the widespread 3'UTR shortening that occurs in cancer.<sup>156</sup>

Quite interesting is also the case of the ZEB1 and ZEB2 transcription factors and master regulators of EMT (see “Expression of a sense mRNA and an antisense ncRNA” chapter above). In epithelial tumors, the ZEB1 protein is oncogenic because it promotes EMT not only directly (by repressing the transcription of miR-200 family), but also indirectly: it induced the transcription of *Integrin A1* mRNA, which in turn sponges the tumor-suppressive miR-181b away from *Adenylyl Cyclase 9* (*ADCY9*) mRNA. As a result, the levels of *ADCY9* protein increase, as well as those of cyclic AMP, which favors metastatic dissemination.<sup>157</sup> In melanoma *ZEB2* behaves as a tumor suppressor not only because ZEB2 protein activates MITF-dependent differentiation program,<sup>158</sup> but also because *ZEB2* 3'UTR has *PTEN* mRNA as ceRNA partner.<sup>151</sup>

For further examples of specific mRNAs with ceRNA activity, please refer to Table 1. Examples of extended ceRNA networks that revolve around mRNAs are reported in.<sup>153,159–161</sup>

### BIFUNCTIONAL ncRNAs EXERT CODING FUNCTIONS AS WELL

In the last years, polysome/ribosome profiling and mass spectroscopy analyses have provided evidence that, together with<sup>162</sup> or alternatively to<sup>163–165</sup> coding-independent activities, ncRNAs can carry ORFs that are translated into ncRNA-encoded peptides (ncPEPs).<sup>166–168</sup> The ~100 ncPEPs characterized so far are mostly short (< 100 aa), but longer examples exist.<sup>167,169–171</sup> Many of these ncPEPs have an impact in cancer, as described below (Fig. 3).

### Translation of pri-miRNAs

The presence of a short ORF has been demonstrated within the pri-miRNA of miR-200a and miR-200b, two miRNAs with tumor-suppressive activity that inhibit metastasis by blocking EMT (see “Expression of a sense mRNA and an antisense ncRNA” chapter and “The 3'UTR exerts non-coding functions” chapter above). When overexpressed in prostate cancer cells, these pri-miRNA-encoded peptides, called miPEP-200a and miPEP-200b, have been shown to inhibit migration and to impair EMT. However, the molecular mechanism of action and the potential synergy with the corresponding miRNAs remain to be established.<sup>172</sup>

### Translation of lincRNAs

Recent reports indicate that several lincRNAs contain ORFs that encode peptides and attest the importance of this feature in the development and progression of cancer.<sup>59</sup> Two examples are described below.

**Table 1.** List of mRNAs that exert ceRNA activity in human cancer due to their 3'UTR.

3'UTR	miRNA	DE-SILENCED mRNA	TUMOR TYPE	SPONGING ACTIVITY OUTPUT	IN VIVO MODELS	Ref #
AEG-1	miR-30a	Snail and Vimentin	Non Small Cell Lung Cancer	Oncogenic	-	214
BCL11B	miR-17, miR-93, miR-20a, miR-20b, miR-106a, and miR-106b	MICA and MICB	Colorectal Cancer	Tumor-suppressive	xenograft	215
CCR2	miR-125b	STARD13	Breast Cancer	Tumor-suppressive	xenograft	216
CD44	miR-512-3p, miR-491, and miR-671	FN1	Breast Cancer	Oncogenic	xenograft	217
CD44	miR-328	Col1a1	Breast Cancer	Oncogenic	xenograft	217
CD44	miR-216a, miR-330, and miR-608	CDC42	Breast Cancer	Tumor-suppressive	xenograft	218
CD44	miR-34a (+ miR-373 and miR-520c)	ULBP2	Liver Cancer Stem Cells	Tumor-suppressive	-	219
CDH5, HOXD1, and HOXD10	miR-9, miR-10b, and miR-125b	STARD13	Breast Cancer	Tumor-suppressive	xenograft + patients	220
c-Myc	miR-34a	CD44	Urothelial Carcinoma	Oncogenic	xenograft + patients	221
CNOT6L	miRs-17, 19a, 19b, 20a, 20b, and 106b	PTEN	Colorectal Cancer and Prostate Cancer	Tumor-suppressive	ref. <sup>151</sup>	150
CXCR4	miR-146a	TRAF6 and EGFR	Breast Cancer	Oncogenic	-	148
CYP4Z1	miR-9	E-Cadherin	Breast Cancer	Tumor-suppressive	-	222
CYP4Z2P	miR-211, miR-125a-3p, miR-197, miR-1226, and miR-204	CYP4Z1	Breast Cancer	Oncogenic	chick embryo + rat	223,224
DNMT3B and TET3	miR-4465	many, including PTEN	Colorectal Cancer	Tumor-suppressive	-	155
E2F6	miR-193a	c-KIT	Ovarian Cancer	Oncogenic	xenograft + patients	225
FOXO1	miR-9	E-Cadherin	Breast Cancer	Tumor-suppressive	-	226
ITGA1	miR-181b	ADCY9	Lung Cancer (+713P, 307P, 344LN, 393LN, 393P, 412P, 344SQ, 344P, 531LN1, 531LN2, 531P1, and 531P2 murine cell lines)	Oncogenic	xenograft	157
LAMC1	miR-124	CD151	Hepatocellular Carcinoma	Oncogenic	xenograft + patients	227
MMP1	miR-188-5p	SOX4 and CDK4	Oral Squamous Cell Carcinoma	Oncogenic	patients	228
PCNX	miR-26, miR-182, miR-340, and miR-506	Skp2	Non Small Cell Lung Cancer	Oncogenic	patients	229
PDL1 and LDHA	miR-34a	PDL1 and LDHA	Triple Negative Breast Cancer (+4T1 murine cell line)	Oncogenic	xenograft + patients	230
PIK3C2A	miR-124	CD151	Hepatocellular Carcinoma	Oncogenic	xenograft + patients	231
PTCH1	miR-101-3p	SLC39A6	Non Small Cell Lung Cancer	Oncogenic	-	232
SNAI1	multiple migration/invasion-related miRNAs	Vimentin and E-cadherin	Ovarian Cancer	Oncogenic	-	233
SNAI2	-	MARCKS	Ovarian Cancer	Oncogenic	-	234
STARD13	miR-125b	TP53INP1	Breast Cancer	Tumor-suppressive	-	235



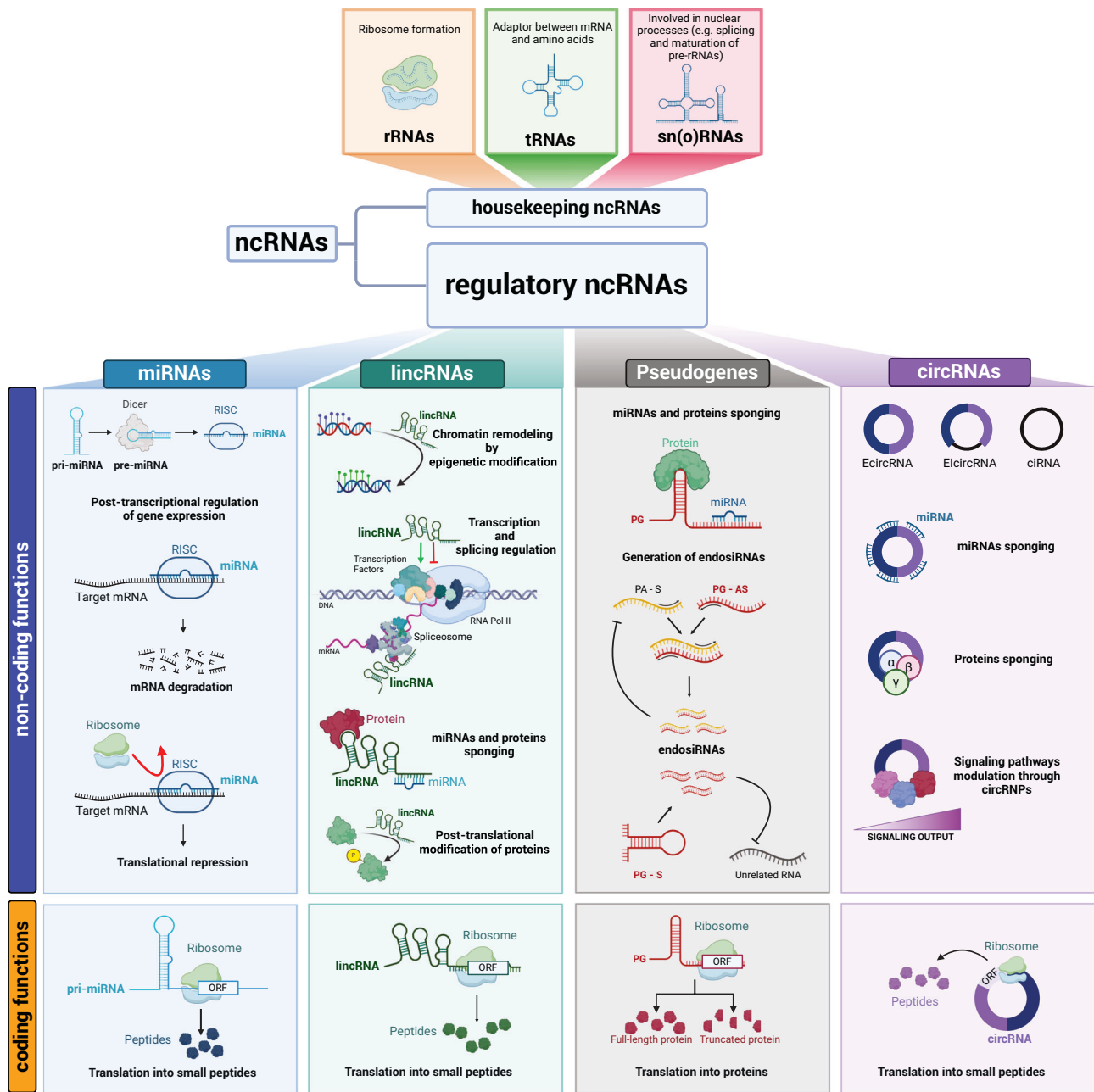
Table 1. continued

3'UTR	miRNA	DE-SILENCED mRNA	TUMOR TYPE	SPONGING ACTIVITY OUTPUT	IN VIVO MODELS	Ref #
STAR13	miR-125b	BMF	Breast Cancer	Tumor-suppressive	xenograft	236
STAR13	network	YAP/TAZ	Breast Cancer	Tumor-suppressive	xenograft	237
STAR13	miR-340, miR-448, miR-374, miR-203, let-7, miR-216b, miR-316, miR-23, and miR-153 (potential)	FAS	Hepatocellular Carcinoma	Tumor-suppressive	patients	238
TNRC6B	-	PTEN	Prostate Cancer	Tumor-suppressive	-	154
USP3	miR-224	SMAD4	Colorectal Cancer	Tumor-suppressive	xenograft + patients	239
<b>VAPA/PTEN</b>	<b>miRs-17, 19a, 20a, 20b, 26b, 106a, and 106b</b>	<b>PTEN/VAPA</b>	<b>Colorectal Cancer and Prostate Cancer</b>	<b>Tumor-suppressive</b>	-	150
Versican	miR-133a, miR-199a*, miR-144, and miR-431	CD34 and Fibronectin	Hepatocellular Carcinoma	Oncogenic	transgenic mice + patients	240
Versican	miR-199a-3p and miR-144	RB1	Breast Cancer (+4T1 murine cell line)	Tumor-suppressive	xenograft + transgenic mice	152
Versican	miR-144 and miR-136	PTEN	Breast Cancer (+4T1 murine cell line)	Tumor-suppressive	xenograft + transgenic mice	152
XIAP	miR-29a-5p	FSCN1	Breast Cancer and Colorectal Cancer	Oncogenic	xenograft	241
XIAP	let-7a-5p	HMGGA2	Hepatocellular Carcinoma	Oncogenic	-	242
<b>ZEB2/PTEN</b>	<b>at least four miRNAs (miR-181, miR-200b, miR-25, and miR-92a)</b>	<b>PTEN/ZEB2</b>	<b>Colorectal Cancer and Melanoma (+TB13602 murine cell line)</b>	<b>Tumor-suppressive</b>	<b>xenograft + transgenic mice</b>	151

Green: tumor-suppressive

Red: oncogenic

Bold: examples described in the text



**Fig. 3 Bifunctional ncRNAs exert non-coding functions and coding functions.** ncRNAs can be divided into two groups: housekeeping ncRNAs (upper panels) and regulatory ncRNAs (lower panels). Among housekeeping ncRNAs there are rRNAs, tRNAs, snRNAs and snoRNAs. Regulatory ncRNAs are further divided according to their length. Among short ncRNAs (< 200 nt) there are miRNAs (light blue). Long ncRNAs (lincRNAs,  $\geq 500$  nt) include lincRNAs (light green), pseudogenetic RNAs (PGs, light gray), NATs (see Fig. 1c), and circRNAs (light purple). The primary function of miRNAs is non-coding (blue) and relates to post-transcriptional regulation of gene expression: mRNA degradation or translational repression mediated by the RISC complex. Among non-coding functions of lincRNAs there are: chromatin remodeling by epigenetic modification, transcription and splicing regulation, sponging of miRNAs and proteins, post-translational modification of proteins. Non-coding functions of PGs can be parental gene (PA)-related or unrelated. They include sponging of miRNAs and proteins, and mRNA degradation through endosRNAs. S sense strand, AS antisense strand. circRNAs can be composed by only exons (EcircRNAs), both exons and introns (ElcircRNAs), or only introns (ciRNAs). They exert non-coding functions by acting as sponges for miRNAs and proteins, and by forming circRNPs that modulate signaling pathways. These groups of regulatory ncRNAs can all exert non-canonical coding functions as well, because they can be translated into ncPEPs, i.e., small peptides or proteins (orange).

*LINC00665* is a lincRNA aberrantly expressed in more than 15 cancer types. It is involved in several signaling pathways (Wnt/ $\beta$ -Catenin, TGF- $\beta$ , NF- $\kappa$ B, PI3K/AKT, and MAPK) and has attracted attention as a diagnostic and prognostic marker. In most of the cases, *LINC00665* is an oncogenic lincRNA that exerts non-coding activities by working as ceRNA for more than 20 miRNAs, as well as by binding to chromatin remodeling factors and transcriptional

regulators.<sup>163</sup> By contrast, in triple-negative breast cancer (TNBC), it acts as a tumor-suppressive lincRNA that expresses CIP2A-BP. This is a 52-aa peptide that in vitro inhibits migration and invasion of TNBC cells. Mechanistically, CIP2A-BP impairs the PI3K/AKT signaling pathway by binding to Cellular Inhibitor of PP2A (CIP2A). Consequently, CIP2A no longer binds to PP2A phosphatase, and in turn PP2A inactivates AKT by dephosphorylation at Thr308 and

**Table 2.** List of lincRNAs and PGs that are translated into peptides/proteins with a role in human cancer.

lincRNA	ENCODED PEPTIDE	TUMOR TYPE	PEPTIDE EXPRESSION LEVEL	PEPTIDE ACTIVITY OUTPUT	PEPTIDE BIOLOGICAL ACTIVITY	IN VIVO MODELS	Ref #
LINC00266-1	RBRP	Colorectal cancer	↑	Oncogenic	Binding to the m6A reader IGF2BP1. This results in increased c-Myc expression.	xenograft + patients	243
LINC00278	YY1BM	Esophageal squamous cell carcinoma	↓	Tumor-suppressive	Separation of YY1 from AR. This results in eEF2k downregulation.	xenograft + patients	244
<b>LINC00665</b>	<b>CIP2A-BP</b>	Triple negative breast cancer	↓	Tumor-suppressive	Binding to CIP2A. This results in increased dephosphorylation of AKT by PP2A, hence decreased signaling through the PI3K/AKT pathway.	xenograft + transgenic mice + patients	173
LINC00675	FORCP	Colorectal cancer	↓	Tumor-suppressive	Promotion of cancer cell apoptosis and suppression of tumorigenesis, possibly through the regulation of BRI3BP function.	xenograft	245
LINC00908	ASRPS	Triple negative breast cancer	↓	Tumor-suppressive	Impairment of STAT3 phosphorylation. This results in decreased VEGF expression.	xenograft + transgenic mice + patients	246
<b>LINC00998</b>	<b>SMIM30</b>	Hepatocellular carcinoma	↑	Oncogenic	Binding to YES1. This results in increased signaling through the MAPK pathway.	xenograft + patients	175
PG	ENCODED PEPTIDE	TUMOR TYPE	PEPTIDE EXPRESSION LEVEL	PEPTIDE ACTIVITY OUTPUT	PEPTIDE BIOLOGICAL ACTIVITY	IN VIVO MODELS	
BRAF pseudogene	BRAF pseudogene peptide	Thyroid cancer (+CHO hamster cell line and NIH3T3 murine cell line)	↑	Oncogenic	Promotion of signaling through the MAPK pathway. This leads to NIH3T3 cell transformation.	xenograft + patients	247
CRIPTO3	CRIPTO3 protein	Cancer (+F9 murine cells (cripto <sup>-/-</sup> , Nodal <sup>+</sup> ))	↑	Oncogenic	Activation of Nodal signaling pathway.	patients	248
MAPK6P4	P4-135aa	Glioblastoma	↑	Oncogenic	Phosphorylation and stabilization of KLF15.	xenograft + patients	249
<b>NANOGP8</b>	<b>NANOGP8 protein</b>	Prostate cancer	↑	Oncogenic	Cancer initiation in combination with c-MYC.	transgenic mice	192
NA88-A	NA88-A peptide	Melanoma	↑	Tumor-suppressive	Antigenic peptide that is recognized by CD8 <sup>+</sup> T cells.	–	250
OCT4-PG1	OCT4-PG1 protein	Chronic myeloid leukemia	↑	Oncogenic	Alteration of multidrug resistance phenotype by directly interacting with OCT4, SOX2, and NANOG and indirectly with ABC transporters.	–	251
STK24P1	P1-121aa	Glioblastoma	↑	Oncogenic	Phosphorylation and stabilization of ELF2. This results in increased VEGFR2 and VE-cadherin expression.	xenograft + patients	252

Bold: examples described in the text

Ser473 residues. Conversely, the EMT-promoting TGF- $\beta$  signaling pathway actively downregulates CIP2A-BP. Specifically, the TGF- $\beta$  effector SMAD4 upregulates *4E-BP1* transcription. In turn, 4E-BP1 binds to eIF4F and inhibits translation of many proteins, including CIP2A-BP.<sup>173</sup>

To study the role of the CIP2A-BP peptide in tumorigenesis in vivo, a knock-in (KI) mouse model was developed. Specifically, homologous recombination was used to knock in *CIP2A-BP* ORF,

preceded by a loxed STOP codon, into the *ROSA26* (*R26*) locus, which ensures robust transcription under the *R26* promoter. KI mice were then crossed with *CMV-Cre* mice so that homozygous *CIP2A-BP*<sup>+/+</sup> mice were obtained, which show constitutive and ubiquitous CIP2A-BP translation. Finally, *CIP2A-BP*<sup>+/+</sup> mice were further crossed with breast cancer-prone *MMTV-PyMT* mice, obtaining *MMTV-PyMT*;*CIP2A-BP*<sup>+/+</sup> mice. *MMTV-PyMT* is the most widely used genetically engineered model (GEM) of breast cancer

in the mouse. In this model, Murine Mammary Tumor Virus (MMTV) Long Terminal Repeat (LTR) promoter restricts the expression of Polyoma Middle T antigen (PyMT) to the mammary epithelium. This leads to the formation of multifocal primary tumors that have high tendency to metastasize to the lung.<sup>174</sup> In accordance with in vitro data, the combined expression of PyMT and CIP2A-BP in the mammary epithelium, which characterizes *MMTV-PyMT;CIP2A-BP<sup>+/+</sup>* mice, is associated with decreased AKT phosphorylation in primary tumors and decreased number of metastatic nodules in the lungs.<sup>173</sup>

Another interesting example is represented by *LINC00998*. In hepatocellular carcinoma (HCC), *LINC00998* is upregulated and is associated with worse prognosis. This lincRNA is under the transcriptional control by c-MYC and expresses a 59-aa peptide termed SMIM30, which shows oncogenic properties both in vitro and in xenograft mouse models. In fact, SMIM30 overexpression results in increased proliferation, migration and invasion. Mechanistically, SMIM30 localizes at the plasma membrane, where it binds to YES1, a member of SRC tyrosine kinase family. Binding to SMIM30 increases membrane anchoring of YES1 and facilitates downstream activation of oncogenic MAPK signaling pathway.<sup>175</sup>

Additional examples of lincRNA-encoded peptides are listed in Table 2 (upper).

### Translation of PGs

It was estimated that dozens of pseudogenic proteins exist, and some can be longer than 100 aa.<sup>176</sup> They are mostly expressed from processed pseudogenes, therefore they do not share the same promoter with their parental genes and, although they might be highly homologous in sequence, they get expressed in distinct tissues, subcellular compartments, or pathophysiological conditions. Pseudogenic proteins specifically expressed in cancer might also carry mutations that further alter their functioning.<sup>68,177</sup>

NANOGP8 is an example of pseudogenic protein with an oncogenic role. Together with LIN28, OCT4 and SOX2, NANOG forms a core network of transcription factors that regulate self-renewal of stem cells. Cancer cells often re-acquire stem-like properties and one of the mechanisms is the aberrant expression of NANOGP8, the NANOG-like protein translated from *NANOGP8* processed pseudogene.<sup>178,179</sup> NANOGP8 is detectable in various cancer types, including prostate cancer and cancers of the gastrointestinal tract, and its levels are particularly high in the cancer stem cell subpopulation.<sup>180–185</sup> Accordingly, NANOGP8 displays oncogenic properties: it promotes clonogenicity, survival, proliferation, migration, anchorage-independent growth and resistance to anticancer drugs, both in vitro and in xenograft models.<sup>186–190</sup> Furthermore, a NANOGP8 signature is associated with worse prognosis.<sup>191</sup>

Importantly, a transgenic mouse model was developed to study the contribution of NANOGP8 to prostate tumorigenesis in vivo. The ORF of *NANOGP8* was placed under the control of the prostate-specific Probasin promoter mentioned above (*Pb/NANOGP8*). After monitoring a cohort of transgenic animals for up to 2 years, the authors did not observe histological evidence of hyperplasia or PIN. However, when they crossed *Pb/NANOGP8* transgenic mice with *Pb/c-MYC* transgenic mice, they observed an exacerbated phenotype, with thicker epithelial layers containing more atypical cells. These results indicate that NANOGP8 is not sufficient to initiate prostate cancer, although it cooperates with c-MYC.<sup>192</sup> Interestingly, similar results were obtained in additional transgenic models. When NANOGP8 was overexpressed in epithelial organs through the *Cytokeratin 14 (K14)* promoter, no spontaneous tumor development was detected even after a prolonged time of observation.<sup>193</sup> Analogously, the targeted overexpression of Nanog in the mammary gland was not sufficient to induce mammary tumors. However, in the presence of the concomitant overexpression of Wnt-1, Nanog

contributed to the decreased mouse survival due to highly enhanced metastatic burden.<sup>194</sup> In summary, NANOGP8 overexpression/inhibition causes a strong increase/decrease in prostate cancer cell line growth, when xenografted into immunodeficient mice,<sup>180–182,186,191,195</sup> but NANOGP8 overexpression is unable to initiate prostate tumorigenesis in a transgenic mouse.<sup>192,193</sup> These results show the importance of GEMs to assess the specific role played by the gene of interest in each phase of cancer development, from initiation to metastasis.

Additional examples of PG-encoded peptides are listed in Table 2 (lower).

### Translation of circRNAs

About 1% of circRNAs are translated.<sup>81</sup> This mainly applies to circRNAs composed only by exons (EcircRNAs), as they localize in the cytoplasm.<sup>196</sup> Translation remains polysome-dependent, but, in the absence of 5' cap and 3' polyA tail, it is cap-independent. Three main mechanisms have been identified thus far. One relies on the presence of an Internal Ribosome Entry Site (IRES), a sequence that can directly recruit ribosomes to initiate translation. Another cap-independent translation mechanism is mediated by N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) residues. It has been reported that at least 13% of circRNAs carry m<sup>6</sup>A modifications and that just one modification per circRNA is sufficient to promote the recruitment of the m<sup>6</sup>A reader YTHDF3, as well as eIF3A and eIF4s (A, B, G2) translation initiation factors. Finally, rolling translation occurs when the circRNA harbors the ATG, but not the stop codon. Translation is carried on in an infinite circle, until it is interrupted through a mechanism termed -1 Programmed Ribosomal Frameshifting (-1 PRF)-mediated Out-of-frame Stop Codon (OSC).<sup>197–199</sup>

circRNA translation is a field that has just started to be explored. Approaches for the detection of circRNA-encoded peptides<sup>200</sup> and tools for their study<sup>201</sup> have been developed. Yet, they require further refinement to exclude the unintended detection of peptides produced by the linear counterparts.<sup>200</sup> Nevertheless, solid experimental evidence is accumulating about circRNA-encoded peptides that play key roles in cancer.<sup>202–204</sup> Two examples are described below.

In glioblastoma multiforme (GBM), AKT3-174aa ncPEP is encoded by *hsa\_circ\_0017250/circAKT3*, a circRNA that comprises exons 3–7 of the *AKT3* gene. Together with AKT1 and 2, AKT3 is an oncogenic kinase with a well-established role in the PI3K/AKT signaling pathway. Since they activate key downstream effectors of this pathway, AKTs are involved in many aspects of tumor initiation and progression, up to metastasis and drug resistance. They are also actively investigated as therapeutic targets.<sup>205</sup> AKT3-174aa shows decreased expression levels in GBM tissues compared to adjacent normal brain tissues and displays onco-suppressive properties: when overexpressed, it decreases GBM cell proliferation in vitro and in xenograft models. It also increases sensitivity to radiation. Furthermore, higher AKT3-174aa levels are associated with better prognosis in patients. Mechanistically, AKT3-174aa competes with AKT3 for binding to phosphoinositide-dependent kinase-1 (PDK1). In so doing, it prevents AKT3 activation by phosphorylation on Thr308. Therefore, AKT3-174aa should be considered as a dominant-negative AKT3 isoform that restrains AKT3 activity through a negative feedback loop.<sup>164</sup>

In GBM and TNBC, FBXW7-185aa ncPEP is encoded by *hsa\_circ\_022705/circFBXW7*, a circRNA that comprises exons 3–4 of the *FBXW7* gene. FBXW7 is a tumor-suppressive E3 ubiquitin ligase that targets multiple oncogenic proteins, including c-MYC, for proteasome-dependent degradation.<sup>206</sup> FBXW7-185aa shows decreased expression levels in cancer tissues compared to adjacent normal tissues and has tumor-suppressive properties: when overexpressed, it decreases cancer cell proliferation in vitro and in xenograft models. Furthermore, higher FBXW7-185aa levels are associated with better prognosis in patients. Mechanistically, FBXW7-185aa has a strong affinity for Ubiquitin Specific Protease

**Table 3.** List of circRNAs that are translated into peptides with a role in human cancer.

circRNA	ENCODED PEPTIDE	TUMOR TYPE	PEPTIDE EXPRESSION LEVEL	PEPTIDE ACTIVITY OUTPUT	PEPTIDE BIOLOGICAL ACTIVITY	IN VIVO MODELS	Ref #
circAKT3	AKT3-174aa	Glioblastoma	↓	Tumor-suppressive	Competitive binding to PDK1. This results in decreased AKT3 phosphorylation, hence decreased signaling through the PI3K/AKT pathway.	xenograft + patients	164
circASK1	ASK1-272aa	Lung adenocarcinoma	↓	Tumor-suppressive	Competitive binding to AKT1. In this way ASK1 is released from phosphorylation-mediated inactivation.	patients	253
circAXIN1	AXIN1-295aa	Gastric cancer	↑	Oncogenic	Binding to APC. This prevents the interaction of APC with AXIN1, abolishing its inhibitory effect. As a result, signaling through the Wnt/ $\beta$ -Catenin pathway increases.	xenograft + patients	254
circ $\beta$ -Catenin	$\beta$ -Catenin-370aa	Non-small cell lung cancer	↑	Oncogenic	Competitive binding to GSK3 $\beta$ , which prevents the GSK3 $\beta$ -mediated degradation of $\beta$ -Catenin. This results in increased signaling through the Wnt/ $\beta$ -Catenin pathway.	patients	255
		Liver cancer	↑	Oncogenic		xenograft + patients	256
circCHEK	circCHEK1-246aa	Multiple myeloma	↑	Oncogenic	Promotion of cell proliferation through chromosomal instability; enhancement of macrophage-osteoclast differentiation.	xenograft	257
circDIDO1	DIDO1-529aa	Gastric cancer	↓	Tumor-suppressive	Promotion of the ubiquitin-mediated degradation of PRDX2.	xenograft + patients	258
circE-Cad	C-E-Cad	Glioblastoma	↑	Oncogenic	Promotion of EGFR signaling.	xenograft + patients	259
circEIF6	EIF6-224aa	Triple negative breast cancer	↑	Oncogenic	Promotion of the MYH9/Wnt/ $\beta$ -Catenin signaling pathway.	xenograft + patients	260
circFBXW7	FBXW7-185aa	Glioma	↓	Tumor-suppressive	Binding to USP28. This favors the proteasome-dependent degradation of c-Myc induced by FBXW7.	xenograft + patients	207
		Triple negative breast cancer	↓			xenograft + patients	162
circFGFR1	circFGFR1p	Cancer	↓	Tumor-suppressive	Negative regulation of FGFR1.	–	261
circFNDC3B	circFNDC3B-218aa	Colorectal cancer	↓	Tumor-suppressive	Inhibition of the Snail-FBP-EMT axis.	xenograft + patients	262
circGprc5a	circGprc5a-peptide	Bladder cancer	↑	Oncogenic	Binding to Gprc5A. This results in increased signaling through the GPCR pathway.	patients	263
circHEATR5B	HEATR5B-881aa	Glioblastoma	↓	Tumor-suppressive	Mediation of the inhibitory effect of circHEATR5B.	xenograft + patients	264
circ-HER2	HER2-103	Triple negative breast cancer	↑	Oncogenic	Binding to EGFR and HER3. This results in increased EGFR signaling.	xenograft + patients	265
circMAPK1	MAPK1-109aa	Gastric cancer	↓	Tumor-suppressive	Inhibition of MAPK1 phosphorylation. This results in decreased signaling through the MAPK pathway.	xenograft + patients	266

**Table 3.** continued

circRNA	ENCODED PEPTIDE	TUMOR TYPE	PEPTIDE EXPRESSION LEVEL	PEPTIDE ACTIVITY OUTPUT	PEPTIDE BIOLOGICAL ACTIVITY	IN VIVO MODELS	Ref #
circMAPK14	circMAPK14-175aa	Colorectal cancer	↓	Tumor-suppressive	Competitive binding to MKK6. This results in decreased MAPK14 phosphorylation and leads to proteasome-dependent degradation of FOXC1.	xenograft + patients	267
circMAP3K4	circMAP3K4-455aa	Hepatocellular carcinoma	↑	Oncogenic	Alteration of the nuclear distribution of AIF.	xenograft + patients	268
circPLCE1	circPLCE1-411	Colorectal carcinoma	↓	Tumor-suppressive	Dissociation of the HSP90α/RPS3 complex, followed by proteasome-dependent degradation of RSP3. As a result, signaling through the NF-κB pathway decreases.	xenograft + PDX + patients	269
circPPP1R12A	cPPP1R12A-73aa	Colon cancer	↑	Oncogenic	Promotion of Hippo-YAP signaling pathway.	xenograft + patients	270
circSHPRH	SHPRH-146aa	Glioma	↓	Tumor-suppressive	Promotion of proteasome-dependent degradation of PCNA.	xenograft + patients	271
circSMO	SMO-193aa	Glioblastoma	↑	Oncogenic	Promotion of HH signaling pathway (Shh/Gli1/FUS/SMO-193aa/SMO).	xenograft + patients	272
circUBE4B	circUBE4B-173aa	Esophageal squamous cell carcinoma	↑	Oncogenic	Promotion of MAPK1 phosphorylation. This results in increased signaling through the MAPK pathway.	xenograft + patients	273
circ0000437	CORO1C-47aa	Endometrial cancer	↓	Tumor-suppressive	Inhibition of VEGF expression.	xenograft + patients	274
<b>circPINTexon2</b>	<b>PINT87aa</b>	Glioblastoma	↓	Tumor-suppressive	Binding to PAF1. This results in the inhibition of the transcriptional elongation of multiple oncogenes.	xenograft + patients	209
ecircCUX1	p113	Neuroblastoma	↑	Oncogenic	Formation of the p113/ZRF1/BRD4 transcriptional regulatory complex.	xenograft + patients	275

Bold: examples described in the text

28 (USP28), a de-ubiquitinating enzyme that prevents c-MYC degradation. By binding to USP28, FBXW7-185aa impairs its binding to c-MYC, and hence enhances the proteasome-dependent degradation induced by FBXW7.<sup>162,207</sup> Interestingly, in TNBC hsa\_circ\_022705 not only favors FBXW7 activity through FBXW7-185aa ncPEP, but also sustains FBXW7 expression by sponging miR-197-3p.<sup>162</sup>

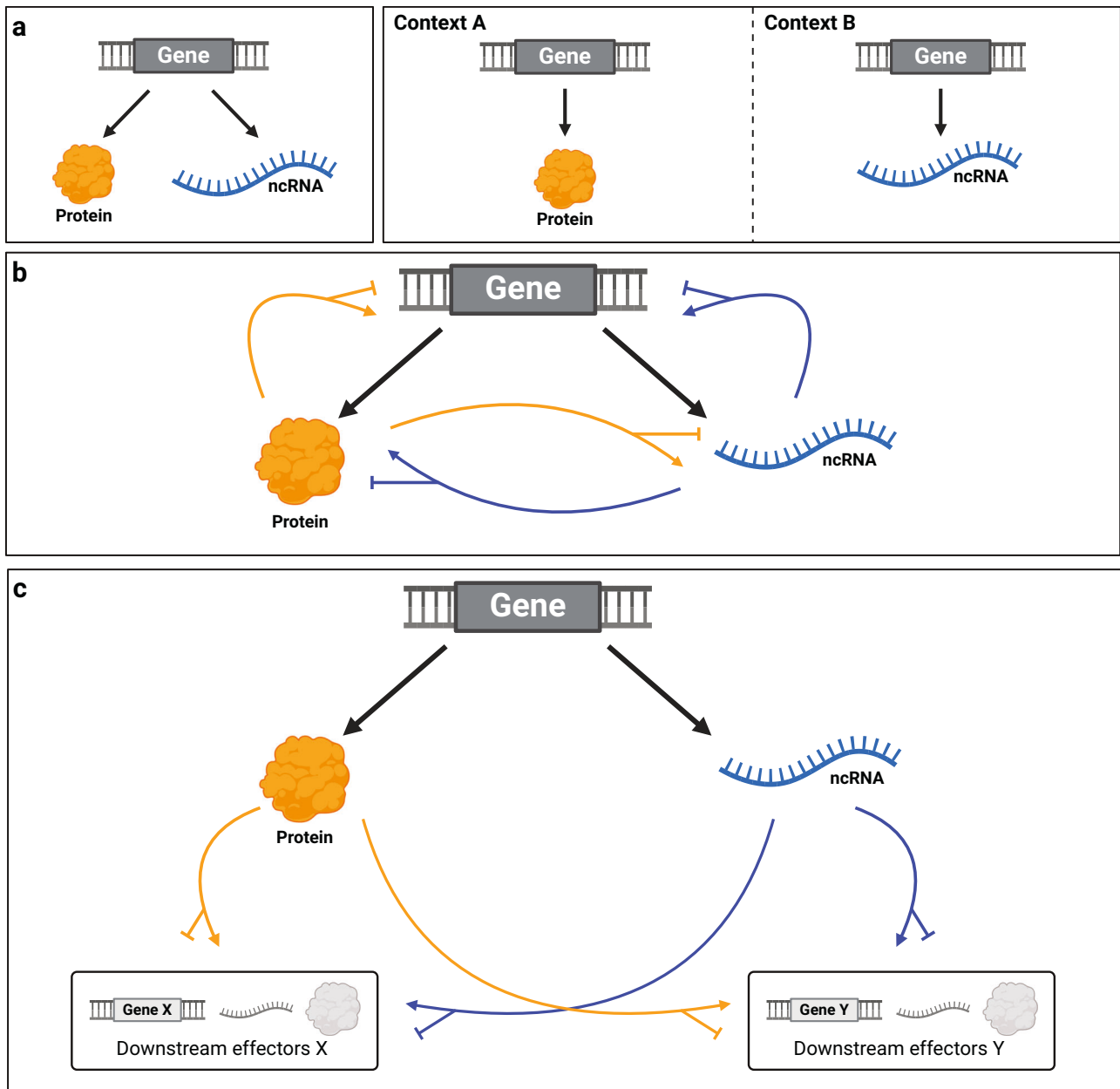
Additional examples of circRNA-encoded peptides are listed in Table 3.

### CONCLUDING REMARKS

In the last few years, it has become evident that the complexity of our genome vastly exceeds the simple organization into genetic units that code for proteins. Proteins are certainly the building blocks of cellular and organismal structures, but hundreds of thousands of non-coding RNAs are also at play, and in turn peptides of various lengths can be translated outside the canonical CDs. Therefore, the intricacy of the coding vs non-coding interplay is progressively unraveling with mindboggling scenarios. As a paradigmatic example of the blurry boundary between what is coding and what is non-coding, we highlight *long intergenic non-protein-coding RNA p53-induced transcript (LINC-PINT)*, a nuclear lincRNA under the transcriptional control

of p53. *LINC-PINT* exerts an oncosuppressive role, although the molecular mechanism varies from cancer type to cancer type. In colon cancer, it works as a long non-coding RNA: it binds to the Polycomb Repressive Complex 2 and in so doing it prevents the transcription of pro-proliferation and pro-survival genes.<sup>208</sup> Conversely, in glioblastoma, although it does not show coding capabilities per se, *LINC-PINT* undergoes back-splicing of exon 2. The EcircRNA that is generated, named circPINTexon2, contains an IRES and is translated into an 87-aa peptide. In turn, it is the PINT87aa ncPEP to be endowed with a tumor-suppressive role: in the nucleus, it directly interacts with RNA Polymerase II-Associated factor 1 (PAF1) and inhibits the elongation of the primary transcript of several oncogenes.<sup>209</sup>

In spite of such a variegated use of coding and non-coding elements, in Fig. 4 we attempt to categorize the modalities of expression and function of the coding and non-coding products derived from the same bifunctional gene: the protein and the non-coding RNA are co-expressed (e.g., in TNBC, hsa\_circ\_022705 is translated into FBXW7-185aa ncPEP and it also works as sponge for miR-197-3p<sup>162</sup>), or they are expressed in distinct tissues/physiopathologic conditions (e.g., *LINC00665* works as ceRNA for miRNAs in several cancer types, with the exception of TNBC where it is instead translated into CIP2A-BP peptide<sup>163</sup>) (Fig. 4a). The protein and the non-coding RNA act *in cis* on their own gene,



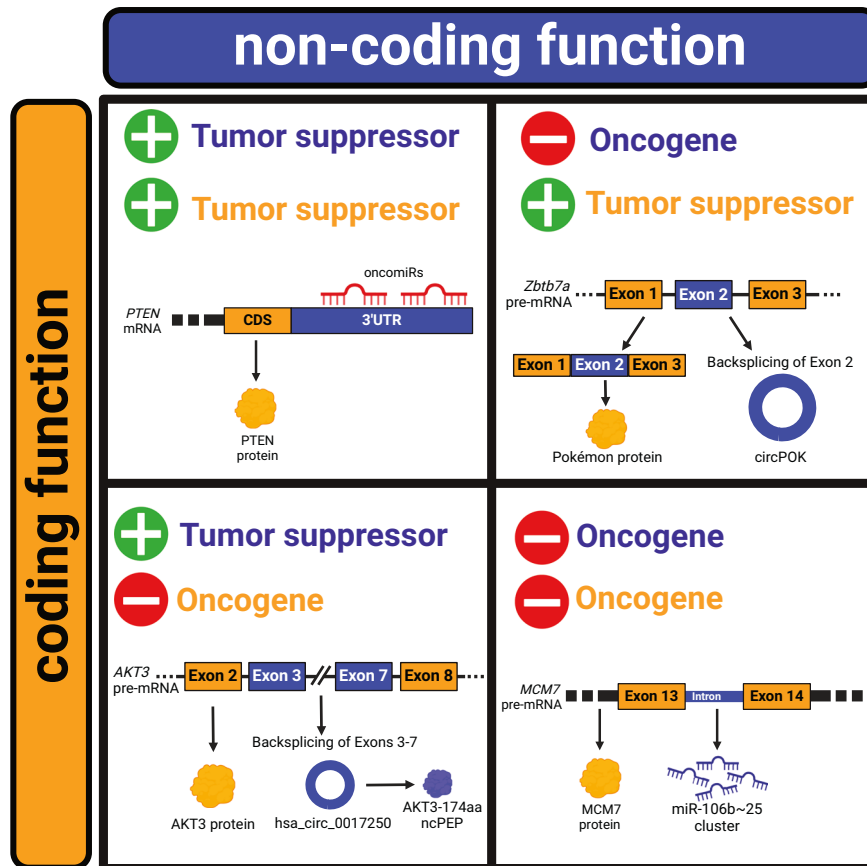
**Fig. 4 The complexity of expression and function of the coding and non-coding products derived from the same bifunctional gene.** The coding product is represented as a generic orange protein and the non-coding product as a generic blue ncRNA. **a** Possible scenarios for gene expression. The protein and the ncRNA are expressed together in the same context (left), or separately in two different contexts (right). **b** Possible *in cis* regulatory mechanisms. The protein and the ncRNA positively or negatively regulate the gene from which they originate, or each other. **c** Possible *in trans* regulatory mechanisms. The protein and the ncRNA positively or negatively regulate downstream effectors (other genes, RNAs or proteins) that can be distinct or the same for both.

regulating each other's expression and/or activity (e.g., *ZEB-AS1 cis-NAT* sustains ZEB1 protein expression<sup>115</sup>) (Fig. 4b). Alternatively, they act *in trans* on other genes or their products (e.g., MCM7 promotes DNA replication and miR-106b~25 cluster downregulates PTEN expression<sup>133</sup>) (Fig. 4c). Finally, their activities can be concordant (e.g., p53 protein and the CDS that encodes it<sup>140–142</sup>) or discordant (e.g., AKT3 protein vs AKT3-174aa ncPEP<sup>164</sup>).

In addition, we highlight that concordant or discordant molecular activities result in a concordant or discordant impact on tumorigenesis: the protein and the non-coding RNA can be both tumor suppressors or both oncogenes, but oncogenic and tumor-suppressive activities can also coexist (see Fig. 5 for specific examples).

How can we disentangle the specific role exerted by coding and non-coding products of the same genetic unit, when we face such a degree of molecular and biological complexity?

To ablate the genetic unit through classic homologous recombination-mediated knockout is a coarse approach that will most certainly lead to profoundly misleading results, due to the concomitant inactivation of two or more players. However, the knockout of the bifunctional gene can be coupled with the add-back of one of its functional products at the time. This strategy suffers from the limitation of triggering the expression of the added-back (non-)coding RNA at supra-physiological or under-physiological levels. Nevertheless, it helps define the function of each product per se, which should precede the study of functional cross-talks among multiple products.



**Fig. 5 Concordant or discordant impact of the coding and non-coding products derived from the same bifunctional gene in cancer.** Graphical representation of the possible combinations of cancer-related functions of coding (orange) and non-coding (blue) partners. Upper left panel: both the coding and the non-coding products have tumor-suppressive properties. *PTEN* mRNA encodes the tumor suppressor PTEN protein and, due to its 3'UTR, it can sponge oncogenic miRNAs. Upper right panel: the coding product is a tumor suppressor, while the non-coding product is an oncogene. *Zbtb7a* pre-mRNA undergoes canonical splicing and generates a mature mRNA that encodes the Pokémon protein with tumor suppressor properties; the same pre-mRNA also undergoes back-splicing of exon 2, leading to the formation of oncogenic circPOK. Bottom left panel: the coding product is an oncogene, while the non-coding product is a tumor suppressor. *AKT3* pre-mRNA is translated into AKT3 oncogenic protein, but exons 3–7 undergo back-splicing, leading to the production of a circRNA (*hsa\_circ\_0017250*) that exerts tumor-suppressive effects through its translation into the AKT3-174aa ncPEP. Bottom right panel: both the coding and the non-coding products have oncogenic properties. *MCM7* pre-mRNA is spliced to produce a mature mRNA that is translated into the oncogenic MCM7 protein; in addition, intron 13 hosts oncogenic miR-106b~25 cluster. From a therapeutic point of view, the optimal approach is to enhance/restore tumor-suppressive activities (green "plus" symbol) and/or, on the other hand, abolish/inhibit oncogenic activities (red "minus" symbol). This is easier in the case of bifunctional genes whose products have concordant outputs, while it can be ineffective or even deleterious in the case of a discordant output.

Alternatively, more precise ways to remodel the genome can be used, such as CRISPR-mediated editing. For example, CRISPR technology allows to surgically mutagenize the MREs present in the 3'UTR of a given mRNA, preventing its ability to work as ceRNA. Nevertheless, pitfalls hide even behind these apparently "cleaner" genetic interventions. If the coding and the non-coding RNAs regulate each other's expression, then the alteration of one will inevitably affect the other as well. Therefore, it will not be possible to establish whether the functional outcome is a direct or indirect consequence of the alteration introduced. Going back to the example mentioned above, the removal of MREs will certainly deprive the mRNA of its non-coding ceRNA activity *in trans*, but it will also have an effect *in cis*, as it will result in an increase in the stability/translation of the mRNA itself, and hence in the level of the corresponding protein, since it is no longer targeted by miRNAs.

Irrespective of the approach used to study a bifunctional gene *in vitro*, we emphasize the importance of corroborating the obtained results with appropriate *in vivo* models. So far the *in vivo* characterization of non-canonical functions of coding and non-coding RNAs has been almost exclusively performed in xenograft

models. Transgenic mouse models can be counted on one hand, while knockout mouse models are even fewer and all fall outside cancer research. For example, Masumoto et al.<sup>210</sup> set up a knockout GEM to study the role played by *LINC00961* and its encoded polypeptide SPAR (Small regulatory Polypeptide of Amino acid Response) in muscle regeneration. Xenografted animals are easy to handle, and they produce quite consistent results in a relatively short time. Nevertheless, we should rely on them just as a first readout of cell autonomous outcomes. To fully grasp the involvement of the most promising (non-)coding candidates on disease/cancer initiation, progression, and response to pharmacological treatment, it is then necessary to develop appropriate GEMs in immunocompetent hosts.<sup>91,211</sup>

Moving from basic to translational research, bifunctional genes offer concrete therapeutic opportunities. This rests on the flourishing interest in RNA-based drugs, which are under investigation in alternative to or in combination with "classical" chemical inhibitors of protein function and have gained full acceptance during COVID-19 emergency.<sup>212,213</sup> Bifunctional genes that produce only tumor suppressors or only oncogenes are easier



to deal with. By contrast, bifunctional genes that produce one tumor suppressor and one oncogene need to be approached with caution: we must avoid the unintended impairment of the tumor-suppressive activity or the boosting of the oncogenic activity caused by our intervention.

In conclusion, we show that the study of protein-coding and non-coding products, their functions and regulation is essential to understand many genetic units in our genome. We also propose that bifunctional genes should be classified as such, distinguishing them from only coding ones and only non-coding ones. However, we point out that much more functional validation is needed towards such classification, both in vitro and in vivo, using the appropriate technological tools for genetic manipulation currently available or to be developed in the foreseeable future. We cannot deny that some intricacies are so thick that may prove daunting or almost impossible to resolve. Nonetheless, we are confident that the widespread transcription and partly overlapping translation of our genome is paving the way for unprecedented opportunities of discovery, drug development and disease treatment.

## REFERENCES

- CRICK, F. Central dogma of molecular biology. *Nature* **227**, 561–563 (1970).
- Strader, L. C. et al. The complexity of transferring genetic information. *Mol. Cell* **83**, 320–323 (2023).
- Takahashi, H. & Carninci, P. Widespread genome transcription: new possibilities for RNA therapies. *Biochem. Biophys. Res. Commun.* **452**, 294–301 (2014).
- Beermann, J., Piccoli, M.-T., Viereck, J. & Thum, T. Non-coding RNAs in development and disease: background, mechanisms, and therapeutic approaches. *Physiol. Rev.* **96**, 1297–1325 (2016).
- Mattick, J. S. et al. Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* **24**, 430–447 (2023).
- Oo, J. A., Brandes, R. P. & Leisegang, M. S. Long non-coding RNAs: novel regulators of cellular physiology and function. *Pflug. Arch.* **474**, 191–204 (2022).
- Mattick, J. S. A Kuhnian revolution in molecular biology: most genes in complex organisms express regulatory RNAs. *BioEssays* **45**, e2300080 (2023).
- Sparber, P., Filatova, A., Khantemirova, M. & Skoblov, M. The role of long non-coding RNAs in the pathogenesis of hereditary diseases. *BMC Med. Genomics* **12**, 42 (2019).
- Chen, Y., Li, Z., Chen, X. & Zhang, S. Long non-coding RNAs: from disease code to drug role. *Acta Pharm. Sin. B* **11**, 340–354 (2021).
- Shang, R., Lee, S., Senavirathne, G. & Lai, E. C. microRNAs in action: biogenesis, function and regulation. *Nat. Rev. Genet.* **24**, 816–833 (2023).
- Morey, C. & Avner, P. Employment opportunities for non-coding RNAs. *FEBS Lett.* **567**, 27–34 (2004).
- Zhang, P., Wu, W., Chen, Q. & Chen, M. Non-coding RNAs and their integrated networks. *J. Integr. Bioinform.* **16**, 20190027 (2019).
- Zhang, Z., Zhang, J., Diao, L. & Han, L. Small non-coding RNAs in human cancer: function, clinical utility, and characterization. *Oncogene* **40**, 1570–1577 (2021).
- Xiong, Q., Zhang, Y., Li, J. & Zhu, Q. Small non-coding RNAs in human cancer. *Genes* **13**, 2072 (2022).
- García-Vilchez, R. et al. METTL1 promotes tumorigenesis through tRNA-derived fragment biogenesis in prostate cancer. *Mol. Cancer* **22**, 119 (2023).
- Lin, Y., Zheng, J. & Lin, D. PIWI-interacting RNAs in human cancer. *Semin. Cancer Biol.* **75**, 15–28 (2021).
- Ameres, S. L. & Zamore, P. D. Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14**, 475–488 (2013).
- Michaels, Y. S. et al. Precise tuning of gene expression levels in mammalian cells. *Nat. Commun.* **10**, 818 (2019).
- Cavallari, I. et al. The miR-200 family of microRNAs: fine tuners of epithelial-mesenchymal transition and circulating cancer biomarkers. *Cancers* **13**, 5874 (2021).
- Ali Syeda, Z., Langden, S. S., Munkhzul, C., Lee, M. & Song, S. J. Regulatory mechanism of microRNA expression in cancer. *Int. J. Mol. Sci.* **21**, 1723 (2020).
- Peng, Y. & Croce, C. M. The role of microRNAs in human cancer. *Signal Transduct. Target. Ther.* **1**, 15004 (2016).
- Zhou, K., Liu, M. & Cao, Y. New insight into microRNA functions in cancer: oncogene-microRNA-tumor suppressor gene network. *Front. Mol. Biosci.* **4**, 46 (2017).
- Si, W., Shen, J., Zheng, H. & Fan, W. The role and mechanisms of action of microRNAs in cancer drug resistance. *Clin. Epigenetics* **11**, 25 (2019).
- Svoronos, A. A., Engelman, D. M. & Slack, F. J. OncomiR or tumor suppressor? the duplicity of microRNAs in cancer. *Cancer Res.* **76**, 3666–3670 (2016).
- Cui, M. et al. Circulating microRNAs in cancer: potential and challenge. *Front. Genet.* **10**, 626 (2019).
- Smolarz, B., Durczyński, A., Romanowicz, H., Szyłło, K. & Hogendorf, P. miRNAs in cancer (review of literature). *Int. J. Mol. Sci.* **23**, 2805 (2022).
- El-Daly, S. M., Gouhar, S. A. & Abd Elmageed, Z. Y. Circulating microRNAs as reliable tumor biomarkers: opportunities and challenges facing clinical application. *J. Pharm. Exp. Ther.* **384**, 35–51 (2023).
- Kim, T. & Croce, C. M. MicroRNA: trends in clinical trials of cancer diagnosis and therapy strategies. *Exp. Mol. Med.* **55**, 1314–1321 (2023).
- Salmena, L., Poliseno, L., Tay, Y., Kats, L. & Pandolfi, P. P. A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? *Cell* **146**, 353–358 (2011).
- Ala, U. et al. Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc. Natl. Acad. Sci. USA* **110**, 7154–7159 (2013).
- Bosia, C., Pagnani, A. & Zecchina, R. Modelling competing endogenous RNA networks. *PLoS One* **8**, e66609 (2013).
- Figliuzzi, M., Marinari, E. & De Martino, A. MicroRNAs as a selective channel of communication between competing RNAs: a steady-state theory. *Biophys. J.* **104**, 1203–1213 (2013).
- Hausser, J. & Zavolan, M. Identification and consequences of miRNA-target interactions—beyond repression of gene expression. *Nat. Rev. Genet.* **15**, 599–612 (2014).
- Poliseno, L. & Pandolfi, P. P. PTEN ceRNA networks in human cancer. *Methods* **77–78**, 41–50 (2015).
- Bosia, C. et al. RNAs competing for microRNAs mutually influence their fluctuations in a highly non-linear microRNA-dependent manner in single cells. *Genome Biol.* **18**, 37 (2017).
- Chiu, H. S. et al. The number of titrated microRNA species dictates ceRNA regulation. *Nucleic Acids Res.* **46**, 4354–4369 (2018).
- Dhawan, A. Mathematical modeling of ceRNA-based interactions. *Methods Mol. Biol.* **2324**, 105–114 (2021).
- Yang, N., Liu, K., Yang, M. & Gao, X. ceRNAs in cancer: mechanism and functions in a comprehensive regulatory network. *J. Oncol.* **2021**, 4279039 (2021).
- Nitzan, M., Steiman-Shimony, A., Altviva, Y., Biham, O. & Margalit, H. Interactions between distant ceRNAs in regulatory networks. *Biophys. J.* **106**, 2254–2266 (2014).
- Paci, P., Colombo, T. & Farina, L. Computational analysis identifies a sponge interaction network between long non-coding RNAs and messenger RNAs in human breast cancer. *BMC Syst. Biol.* **8**, 83 (2014).
- Conte, F. et al. Role of the long non-coding RNA PVT1 in the dysregulation of the ceRNA-ceRNA network in human breast cancer. *PLoS One* **12**, e0171661 (2017).
- Zhang, Y. et al. Inferences of individual drug responses across diverse cancer types using a novel competing endogenous RNA network. *Mol. Oncol.* **12**, 1429–1446 (2018).
- Qian, D. et al. Integrated analysis of ceRNA network reveals prognostic and metastasis associated biomarkers in breast cancer. *Front. Oncol.* **11**, 670138 (2021).
- Chen, J. et al. Comprehensive analysis of ceRNA networks reveals prognostic lncRNAs related to immune infiltration in colorectal cancer. *BMC Cancer* **21**, 255 (2021).
- Xu, J., Xu, J., Liu, X. & Jiang, J. The role of lncRNA-mediated ceRNA regulatory networks in pancreatic cancer. *Cell Death Discov.* **8**, 287 (2022).
- Chan, J. & Tay, Y. Noncoding RNA:RNA regulatory networks in cancer. *Int. J. Mol. Sci.* **19**, 1310 (2018).
- Anastasiadou, E., Jacob, L. S. & Slack, F. J. Non-coding RNA networks in cancer. *Nat. Rev. Cancer* **18**, 5–18 (2018).
- Li, Z. et al. LncBook 2.0: integrating human long non-coding RNAs with multi-omics annotations. *Nucleic Acids Res.* **51**, D186–D191 (2023).
- Nojima, T. & Proudfoot, N. J. Mechanisms of lncRNA biogenesis as revealed by nascent transcriptomics. *Nat. Rev. Mol. Cell Biol.* **23**, 389–406 (2022).
- Hou, J., Zhang, G., Wang, X., Wang, Y. & Wang, K. Functions and mechanisms of lncRNA MALAT1 in cancer chemotherapy resistance. *Biomark. Res.* **11**, 23 (2023).
- Ma, L., Bajic, V. B. & Zhang, Z. On the classification of long non-coding RNAs. *RNA Biol.* **10**, 924–933 (2013).
- Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **43**, 904–914 (2011).
- Li, X., Wu, Z., Fu, X. & Han, W. lncRNAs: insights into their function and mechanics in underlying disorders. *Mutat. Res. Rev. Mutat. Res.* **762**, 1–21 (2014).
- Zhang, X. et al. Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. *Int. J. Mol. Sci.* **20**, 5573 (2019).
- Statello, L., Guo, C. J., Chen, L. L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **22**, 96–118 (2021).

56. Qian, Y., Shi, L. & Luo, Z. Long non-coding RNAs in cancer: implications for diagnosis, prognosis, and therapy. *Front. Med.* **7**, 612393 (2020).
57. Chen, B. et al. Targeting non-coding RNAs to overcome cancer therapy resistance. *Signal Transduct. Target. Ther.* **7**, 121 (2022).
58. Barczak, W. et al. Long non-coding RNA-derived peptides are immunogenic and drive a potent anti-tumour response. *Nat. Commun.* **14**, 1078 (2023).
59. Ransohoff, J. D., Wei, Y. & Khavari, P. A. The functions and unique features of long intergenic non-coding RNA. *Nat. Rev. Mol. Cell Biol.* **19**, 143–157 (2018).
60. Brockdorff, N. et al. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71**, 515–526 (1992).
61. Rinn, J. L. et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311–1323 (2007).
62. Hussien, B. M. et al. Role of lncRNA BANC1 in human cancers: an updated review. *Front. Cell Dev. Biol.* **9**, 689992 (2021).
63. Wen, X., Wu, Y., Lou, Y., Xia, Y. & Yu, X. The roles of linc-ROR in the regulation of cancer stem cells. *Transl. Oncol.* **28**, 101602 (2023).
64. Ghasemian, M., Babaahmadi-Rezaei, H., Khedri, A. & Selvaraj, C. The oncogenic role of SAMMSON lncRNA in tumorigenesis: A comprehensive review with especial focus on melanoma. *J. Cell Mol. Med.* **27**, 3966–3973 (2023).
65. Zhao, S., Zhang, X., Chen, S. & Zhang, S. Natural antisense transcripts in the biological hallmarks of cancer: powerful regulators hidden in the dark. *J. Exp. Clin. Cancer Res.* **39**, 187 (2020).
66. Santos, F., Capela, A. M., Mateus, F. & Nóbrega-Pereira, S. & Bernardes de Jesus, B. Non-coding antisense transcripts: fine regulation of gene expression in cancer. *Comput. Struct. Biotechnol. J.* **20**, 5652–5660 (2022).
67. Sisu, C. GENCODE pseudogenes. *Methods Mol. Biol.* **2324**, 67–82 (2021).
68. Poliseno, L. Pseudogenes: newly discovered players in human cancer. *Sci. Signal.* **5**, re5 (2012).
69. Nakamura-García, A. K. & Espinal-Enríquez, J. Pseudogenes in cancer: state of the art. *Cancers* **15**, 4024 (2023).
70. Poliseno, L. et al. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038 (2010).
71. Poliseno, L., Marranci, A. & Pandolfi, P. P. Pseudogenes in human cancer. *Front. Med.* **2**, 68 (2015).
72. Cheetham, S. W., Faulkner, G. J. & Dinger, M. E. Overcoming challenges and dogmas to understand the functions of pseudogenes. *Nat. Rev. Genet.* **21**, 191–201 (2020).
73. Chen, X. et al. Re-recognition of pseudogenes: from molecular to clinical applications. *Theranostics* **10**, 1479–1499 (2020).
74. Wang, C. & Lin, H. Roles of piRNAs in transposon and pseudogene regulation of germline mRNAs and lncRNAs. *Genome Biol.* **22**, 27 (2021).
75. Cooke, S. L. et al. Processed pseudogenes acquired somatically during cancer development. *Nat. Commun.* **5**, 3644 (2014).
76. Stasiak, M. et al. The world of pseudogenes: new diagnostic and therapeutic targets in cancers or still mystery molecules? *Life* **11**, 1354 (2021).
77. Vitiello, M. & Poliseno, L. CRISPR/Cas technologies applied to pseudogenes. *Methods Mol. Biol.* **2324**, 265–284 (2021).
78. Sun, M. et al. Systematic functional interrogation of human pseudogenes using CRISPRi. *Genome Biol.* **22**, 240 (2021).
79. Sisu, C. Pseudogenes as biomarkers and therapeutic targets in human cancers. *Methods Mol. Biol.* **2324**, 319–337 (2021).
80. Jeck, W. R. et al. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **19**, 141–157 (2013).
81. Cheng, D., Wang, J., Dong, Z. & Li, X. Cancer-related circular RNA: diverse biological functions. *Cancer Cell Int.* **21**, 11 (2021).
82. Liu, C. X. & Chen, L. L. Circular RNAs: characterization, cellular roles, and applications. *Cell* **185**, 2016–2034 (2022).
83. Feng, X. Y. et al. New insight into circRNAs: characterization, strategies, and biomedical applications. *Exp. Hematol. Oncol.* **12**, 91 (2023).
84. Tang, X. et al. Review on circular RNAs and new insights into their roles in cancer. *Comput. Struct. Biotechnol. J.* **19**, 910–928 (2021).
85. Ma, S., Kong, S., Wang, F. & Ju, S. CircRNAs: biogenesis, functions, and role in drug-resistant tumours. *Mol. Cancer* **19**, 119 (2020).
86. Xu, T. et al. CircRNAs in anticancer drug resistance: recent advances and future potential. *Mol. Cancer* **19**, 127 (2020).
87. Kristensen, L. S., Jakobsen, T., Hager, H. & Kjems, J. The emerging roles of circRNAs in cancer and oncology. *Nat. Rev. Clin. Oncol.* **19**, 188–206 (2022).
88. Qi, D., Ke, R., Huang, J. H. & Wu, E. Forging the future of circRNA therapeutics: unleashing synthetic potential and conquering challenges. *Mol. Ther. Nucleic Acids* **33**, 42–43 (2023).
89. Huang, D. et al. Tumour circular RNAs elicit anti-tumour immunity by encoding cryptic peptides. *Nature* **625**, 593–602 (2024).
90. Villanueva, M. T. Circular RNA vaccines expose cryptic peptides. *Nat. Rev. Drug Discov.* **23**, 103–103 (2024).
91. Patton, E. E. et al. Melanoma models for the next generation of therapies. *Cancer Cell* **39**, 610–631 (2021).
92. Zhong, L. et al. Small molecules in targeted cancer therapy: Advances, challenges, and future perspectives. *Signal Transduct. Target. Ther.* **6**, 201 (2021).
93. Jin, S. et al. Emerging new therapeutic antibody derivatives for cancer treatment. *Signal Transduct. Target. Ther.* **7**, 39 (2022).
94. Chen, J. et al. Pervasive functional translation of noncanonical human open reading frames. *Science* **367**, 1140–1146 (2020).
95. Wei, L. H. & Guo, J. U. Coding functions of “noncoding” RNAs. *Science* **367**, 1074–1075 (2020).
96. Lee, D. S. M. et al. Disrupting upstream translation in mRNAs is associated with human disease. *Nat. Commun.* **12**, 1515 (2021).
97. Hofman, D. A. et al. Translation of non-canonical open reading frames as a cancer cell survival mechanism in childhood medulloblastoma. *Mol. Cell* **84**, 261–276.e18 (2024).
98. Delaidelli, A., Oliveira de Santis, J. & Sorensen, P. H. Actions speak louder than ORFs: A non-canonical microprotein promotes medulloblastoma oncogenesis. *Mol. Cell* **84**, 188–190 (2024).
99. Hubé, F. & Francastel, C. Coding and non-coding RNAs, the frontier has never been so blurred. *Front. Genet.* **9**, 140 (2018).
100. Matlin, A. J., Clark, F. & Smith, C. W. J. Understanding alternative splicing: towards a cellular code. *Nat. Rev. Mol. Cell Biol.* **6**, 386–398 (2005).
101. Sammeth, M., Foissac, S. & Guigó, R. A general definition and nomenclature for alternative splicing events. *PLoS Comput. Biol.* **4**, e1000147 (2008).
102. Liu, Q., Fang, L. & Wu, C. Alternative splicing and isoforms: from mechanisms to diseases. *Genes* **13**, 401 (2022).
103. Dhamija, S. & Menon, M. B. Non-coding transcript variants of protein-coding genes—what are they good for? *RNA Biol.* **15**, 1025–1031 (2018).
104. Yan, Y. et al. Steroid receptor RNA activator protein (SRAP): a potential new prognostic marker for estrogen receptor-positive/node-negative/younger breast cancer patients. *Breast Cancer Res.* **11**, R67 (2009).
105. KAWASHIMA, H. et al. A novel steroid receptor co-activator protein (SRAP) as an alternative form of steroid receptor RNA-activator gene: expression in prostate cancer cells and enhancement of androgen receptor activity. *Biochem. J.* **369**, 163–171 (2003).
106. Hube, F. et al. Alternative splicing of the first intron of the steroid receptor RNA activator (SRA) participates in the generation of coding and noncoding RNA isoforms in breast cancer cell lines. *DNA Cell Biol.* **25**, 418–428 (2006).
107. Hong, C. H., Ho, J. C. & Lee, C. H. Steroid receptor RNA activator, a long non-coding RNA, activates p38, facilitates epithelial-mesenchymal transformation, and mediates experimental melanoma metastasis. *J. Invest. Dermatol.* **140**, 1355–1363.e1 (2020).
108. Allen, P. B., Kwon, Y. G., Nairn, A. C. & Greengard, P. Isolation and characterization of PNUTS, a putative protein phosphatase 1 nuclear targeting subunit. *J. Biol. Chem.* **273**, 4089–4095 (1998).
109. Landsverk, H. B., Kirkhus, M., Bollen, M., Küntziger, T. & Collas, P. PNUTS enhances in vitro chromosome decondensation in a PP1-dependent manner. *Biochem. J.* **390**, 709–717 (2005).
110. Landsverk, H. B. et al. The protein phosphatase 1 regulator PNUTS is a new component of the DNA damage response. *EMBO Rep.* **11**, 868–875 (2010).
111. Grelet, S. et al. A regulated PNUTS mRNA to lncRNA splice switch mediates EMT and tumour progression. *Nat. Cell Biol.* **19**, 1105–1115 (2017).
112. Guarnerio, J. et al. Intragenic antagonistic roles of protein and circRNA in tumorigenesis. *Cell Res.* **29**, 628–640 (2019).
113. Guarnerio, J. et al. A genetic platform to model sarcomagenesis from primary adult mesenchymal stem cells. *Cancer Discov.* **5**, 396–409 (2015).
114. Wu, H. T. et al. Oncogenic functions of the EMT-related transcription factor ZEB1 in breast cancer. *J. Transl. Med.* **18**, 51 (2020).
115. Su, W. et al. Long noncoding RNA ZEB1-AS1 epigenetically regulates the expressions of ZEB1 and downstream molecules in prostate cancer. *Mol. Cancer* **16**, 142 (2017).
116. Burk, U. et al. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep.* **9**, 582–589 (2008).
117. Ghafouri-Fard, S. et al. A review on the role of ZEB1-AS1 in human disorders. *Pathol. Res. Pr.* **245**, 154486 (2023).
118. Guo, Z., Cao, Q., Zhao, Z. & Song, C. Biogenesis, features, functions, and disease relationships of a specific circular RNA: CDR1as. *Aging Dis.* **11**, 1009 (2020).
119. Hansen, T. B. et al. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* **30**, 4414–4422 (2011).
120. Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H. & Bartel, D. P. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**, 1537–1550 (2011).
121. Memczak, S. et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**, 333–338 (2013).

122. Hansen, T. B. et al. Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384–388 (2013).
123. Piwecka, M. et al. Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. *Science* **357**, eaam8526 (2017).
124. Kleaveland, B., Shi, C. Y., Stefano, J. & Bartel, D. P. A network of noncoding regulatory RNAs acts in the mammalian brain. *Cell* **174**, 350–362.e17 (2018).
125. Kristensen, L. S. et al. Spatial expression analyses of the putative oncogene ciRS-7 in cancer reshape the microRNA sponge theory. *Nat. Commun.* **11**, 4551 (2020).
126. Mecozzi, N., Vera, O. & Karreth, F. A. Squaring the circle: circRNAs in melanoma. *Oncogene* **40**, 5559–5566 (2021).
127. Lou, J. et al. Circular RNA CDR1as disrupts the p53/MDM2 complex to inhibit gliomagenesis. *Mol. Cancer* **19**, 138 (2020).
128. Hanniford, D. et al. Epigenetic silencing of CDR1as drives IGF2BP3-mediated melanoma invasion and metastasis. *Cancer Cell* **37**, 55–70.e15 (2020).
129. Boivin, V., Deschamps-Francoeur, G. & Scott, M. S. Protein coding genes as hosts for noncoding RNA expression. *Semin. Cell Dev. Biol.* **75**, 3–12 (2018).
130. Liu, B., Shyr, Y., Cai, J. & Liu, Q. Interplay between miRNAs and host genes and their role in cancer. *Brief. Funct. Genomics* **18**, 255–266 (2019).
131. Zeidler, M., Hüttenhofer, A., Kress, M. & Kummer, K. K. Intragenic microRNAs autoregulate their host genes in both direct and indirect ways — a cross-species analysis. *Cells* **9**, 232 (2020).
132. Wong, A. C. H. & Rasko, J. E. J. Splice and dice: intronic microRNAs, splicing and cancer. *Biomedicines* **9**, 1268 (2021).
133. Poliseno, L. et al. Identification of the *miR-106b~25* microRNA cluster as a proto-oncogenic *PTEN*-targeting intron that cooperates with its host gene *MCM7* in transformation. *Sci. Signal.* **3**, ra29 (2010).
134. Blume, S. et al. Inhibition of tumorigenicity by the 5'-untranslated RNA of the human c-myc P0 transcript. *Exp. Cell Res.* **288**, 131–142 (2003).
135. McMahon, S. B. MYC and the control of apoptosis. *Cold Spring Harb. Perspect. Med.* **4**, a014407 (2014).
136. Apte, R. S., Chen, D. S. & Ferrara, N. VEGF in signaling and disease: beyond discovery and development. *Cell* **176**, 1248–1264 (2019).
137. Masuda, K. et al. A novel tumor-promoting function residing in the 5' non-coding region of vascular endothelial growth factor mRNA. *PLoS Med.* **5**, e94 (2008).
138. Hernández Borrero, L. J. & El-Deiry, W. S. Tumor suppressor p53: Biology, signaling pathways, and therapeutic targeting. *Biochim. Biophys. Acta Rev. Cancer* **1876**, 188556 (2021).
139. Kung, C. P. & Weber, J. D. It's getting complicated — a fresh look at p53-MDM2-ARF triangle in tumorigenesis and cancer therapy. *Front. Cell Dev. Biol.* **10**, 818744 (2022).
140. Candeias, M. M. et al. p53 mRNA controls p53 activity by managing Mdm2 functions. *Nat. Cell Biol.* **10**, 1098–1105 (2008).
141. Naski, N. et al. The p53 mRNA-Mdm2 interaction. *Cell Cycle* **8**, 31–34 (2009).
142. Gajjar, M. et al. The p53 mRNA-Mdm2 interaction controls Mdm2 nuclear trafficking and is required for p53 activation following DNA damage. *Cancer Cell* **21**, 25–35 (2012).
143. Karakostis, K. & Fähræus, R. Shaping the regulation of the p53 mRNA tumour suppressor: the co-evolution of genetic signatures. *BMC Cancer* **19**, 915 (2019).
144. Ivanyi-Nagy, R. et al. The RNA interactome of human telomerase RNA reveals a coding-independent role for a histone mRNA in telomere homeostasis. *Elife* **7**, e40037 (2018).
145. De Paolo, R. et al. Differential impact of BRAFV600E isoforms on tumorigenesis in a zebrafish model of melanoma. *Cell Biosci.* **13**, 121 (2023).
146. Mayr, C. What are 3' UTRs doing? *Cold Spring Harb. Perspect. Biol.* **11**, a034728 (2019).
147. Chan, J. J., Tabatabaieian, H. & Tay, Y. 3'UTR heterogeneity and cancer progression. *Trends Cell Biol.* **33**, 568–582 (2023).
148. Zheng, T. et al. CXCR4 3'UTR functions as a ceRNA in promoting metastasis, proliferation and survival of MCF-7 cells by regulating miR-146a activity. *Eur. J. Cell Biol.* **94**, 458–469 (2015).
149. Lee, Y. R., Chen, M. & Pandolfi, P. P. The functions and regulation of the PTEN tumour suppressor: New modes and prospects. *Nat. Rev. Mol. Cell Biol.* **19**, 547–562 (2018).
150. Tay, Y. et al. Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell* **147**, 344–357 (2011).
151. Karreth, F. A. et al. In vivo identification of tumor-suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. *Cell* **147**, 382–395 (2011).
152. Lee, D. Y. et al. Expression of versican 3'-untranslated region modulates endogenous microRNA functions. *PLoS One* **5**, e13599 (2010).
153. Sumazin, P. et al. An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. *Cell* **147**, 370–381 (2011).
154. Zarringhalam, K. et al. Identification of competing endogenous RNAs of the tumor suppressor gene PTEN: A probabilistic approach. *Sci. Rep.* **7**, 7755 (2017).
155. Roquid, K. A. R., Alcantara, K. M. M. & Garcia, R. L. Identification and validation of mRNA 3'untranslated regions of DNMT3B and TET3 as novel competing endogenous RNAs of the tumor suppressor PTEN. *Int. J. Oncol.* **56**, 544–558 (2020).
156. Park, H. J. et al. 3' UTR shortening represses tumor-suppressor genes in trans by disrupting ceRNA crosstalk. *Nat. Genet.* **50**, 783–789 (2018).
157. Tan, X. et al. The epithelial-to-mesenchymal transition activator ZEB1 initiates a prometastatic competing endogenous RNA network. *J. Clin. Invest.* **128**, 1267–1282 (2018).
158. Caramel, J. et al. A switch in the expression of embryonic EMT-inducers drives the development of malignant melanoma. *Cancer Cell* **24**, 466–480 (2013).
159. Xu, J. et al. The mRNA related ceRNA–ceRNA landscape and significance across 20 major cancer types. *Nucleic Acids Res.* **43**, 8169–8182 (2015).
160. Chiu, H. S. et al. High-throughput validation of ceRNA regulatory networks. *BMC Genomics* **18**, 418 (2017).
161. Xu, X. et al. Gain of chromosome 1q perturbs a competitive endogenous RNA network to promote melanoma metastasis. *Cancer Res.* **82**, 3016–3031 (2022).
162. Ye, F. et al. 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**, 88–98 (2019).
163. Zhong, C., Xie, Z., Shen, J., Jia, Y. & Duan, S. LINC00665: an emerging biomarker for cancer diagnostics and therapeutics. *Cells* **11**, 1540 (2022).
164. Xia, X. et al. A novel tumor suppressor protein encoded by circular AKT3 RNA inhibits glioblastoma tumorigenicity by competing with active phosphoinositide-dependent kinase-1. *Mol. Cancer* **18**, 131 (2019).
165. Huang, X. et al. Circular RNA AKT3 upregulates PIK3R1 to enhance cisplatin resistance in gastric cancer via miR-198 suppression. *Mol. Cancer* **18**, 71 (2019).
166. Zhou, B. et al. Translation of noncoding RNAs and cancer. *Cancer Lett.* **497**, 89–99 (2021).
167. Xing, J., Liu, H., Jiang, W. & Wang, L. LncRNA-encoded peptide: functions and predicting methods. *Front. Oncol.* **10**, 622294 (2021).
168. Zhang, Y., Wang, X., Hu, C. & Yi, H. Shiny transcriptional junk: lncRNA-derived peptides in cancers and immune responses. *Life Sci.* **316**, 121434 (2023).
169. Liu, H. et al. ncEP: a manually curated database for experimentally validated ncRNA-encoded proteins or peptides. *J. Mol. Biol.* **432**, 3364–3368 (2020).
170. Dragomir, M. P. et al. Funcpep: a database of functional peptides encoded by non-coding rnas. *Noncoding RNA* **6**, 41 (2020).
171. Li, J. & Liu, C. Coding or noncoding, the converging concepts of RNAs. *Front. Genet.* **10**, 496 (2019).
172. Fang, J., Morsalin, S., Rao, V.N., Reddy, E. & Shyam, P. Decoding of non-coding DNA and non-coding RNA: pri-micro RNA-encoded novel peptides regulate migration of cancer cells. *J. Pharm. Sci. Pharm.* **3**, 23–27 (2017).
173. Guo, B. et al. Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression. *EMBO J.* **39**, e102190 (2020).
174. Attalla, S., Taifour, T., Bui, T. & Muller, W. Insights from transgenic mouse models of PyMT-induced breast cancer: Recapitulating human breast cancer progression in vivo. *Oncogene* **40**, 475–491 (2021).
175. Pang, Y. et al. Peptide SMIM30 promotes HCC development by inducing SRC/YES1 membrane anchoring and MAPK pathway activation. *J. Hepatol.* **73**, 1155–1169 (2020).
176. Troskie, R. L. et al. Long-read cDNA sequencing identifies functional pseudogenes in the human transcriptome. *Genome Biol.* **22**, 146 (2021).
177. Dubois, M. L. et al. UBB pseudogene 4 encodes functional ubiquitin variants. *Nat. Commun.* **11**, 1306 (2020).
178. Palla, A. R. et al. Reprogramming activity of NANOGP8, a NANOG family member widely expressed in cancer. *Oncogene* **33**, 2513–2519 (2014).
179. Singovski, G. et al. In vivo epigenetic reprogramming of primary human colon cancer cells enhances metastases. *J. Mol. Cell Biol.* **8**, 157–173 (2016).
180. Jeter, C. R. et al. Functional evidence that the self-renewal gene NANOG regulates human tumor development. *Stem Cells* **27**, 993–1005 (2009).
181. Jeter, C. R. et al. NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene* **30**, 3833–3845 (2011).
182. Zhang, K., Fowler, M., Glass, J. & Yin, H. Activated 5'flanking region of NANOGP8 in a self-renewal environment is associated with increased sphere formation and tumor growth of prostate cancer cells. *Prostate* **74**, 381–394 (2014).
183. Uchino, K. et al. Human Nanog pseudogene8 promotes the proliferation of gastrointestinal cancer cells. *Exp. Cell Res.* **318**, 1799–1807 (2012).
184. Zhang, J. et al. NANOG modulates stemness in human colorectal cancer. *Oncogene* **32**, 4397–4405 (2013).
185. Ma, X., Wang, B., Wang, X., Luo, Y. & Fan, W. NANOGP8 is the key regulator of stemness, EMT, Wnt pathway, chemoresistance, and other malignant phenotypes in gastric cancer cells. *PLoS One* **13**, e0192436 (2018).
186. Kawamura, N. et al. CRISPR/Cas9-mediated gene knockout of NANOG and NANOGP8 decreases the malignant potential of prostate cancer cells. *Oncotarget* **6**, 22361–22374 (2015).

187. Ishiguro, T. et al. Differential expression of nanog1 and nanogp8 in colon cancer cells. *Biochem. Biophys. Res. Commun.* **418**, 199–204 (2012).
188. Mattoo, A. R., Zhang, J., Espinoza, L. A. & Jessup, J. M. Inhibition of *NANOG/NANOGP8* downregulates MCL-1 in colorectal cancer cells and enhances the therapeutic efficacy of BH3 mimetics. *Clin. Cancer Res.* **20**, 5446–5455 (2014).
189. Jiang, Z., Liu, Y. & Wang, C. Oncogenic NanogP8 expression regulates cell proliferation and migration through the Akt/mTOR signaling pathway in human gastric cancer — SGC-7901 cell line. *Onco. Targets Ther.* **9**, 4859–4866 (2016).
190. Li, L. et al. *NANOGP8* expression regulates gastric cancer cell progression by transactivating DBC1 in gastric cancer MKN-45 cells. *Oncol. Lett.* **17**, 555–563 (2019).
191. Jeter, C. R. et al. *NANOG* reprograms prostate cancer cells to castration resistance via dynamically repressing and engaging the AR/FOXA1 signaling axis. *Cell Discov.* **2**, 16041 (2016).
192. Liu, B. et al. Transgenic overexpression of NanogP8 in the mouse prostate is insufficient to initiate tumorigenesis but weakly promotes tumor development in the Hi-Myc mouse model. *Oncotarget* **8**, 52746–52760 (2017).
193. Badeaux, M. A. et al. In vivo functional studies of tumor-specific retrogene NanogP8 in transgenic animals. *Cell Cycle* **12**, 2395–2408 (2013).
194. Lu, X., Mazur, S. J., Lin, T., Appella, E. & Xu, Y. The pluripotency factor nanog promotes breast cancer tumorigenesis and metastasis. *Oncogene* **33**, 2655–2664 (2014).
195. Sui, Y. et al. Roles of *NANOGP8* in cancer metastasis and cancer stem cell invasion during development of castration-resistant prostate cancer. *Ann. Transl. Med.* **9**, 45–45 (2021).
196. Lei, M., Zheng, G., Ning, Q., Zheng, J. & Dong, D. Translation and functional roles of circular RNAs in human cancer. *Mol. Cancer* **19**, 30 (2020).
197. Yang, Y. et al. Extensive translation of circular RNAs driven by N6-methyladenosine. *Cell Res.* **27**, 626–641 (2017).
198. Shi, Y., Jia, X. & Xu, J. The new function of circRNA: translation. *Clin. Transl. Oncol.* **22**, 2162–2169 (2020).
199. Wang, Y. et al. Expanding uncapped translation and emerging function of circular RNA in carcinomas and noncarcinomas. *Mol. Cancer* **21**, 13 (2022).
200. Hansen, T. B. Signal and noise in circRNA translation. *Methods* **196**, 68–73 (2021).
201. Mecozzi, N. et al. Genetic tools for the stable overexpression of circular RNAs. *RNA Biol.* **19**, 353–363 (2022).
202. Kong, S., Tao, M., Shen, X. & Ju, S. Translatable circRNAs and lncRNAs: Driving mechanisms and functions of their translation products. *Cancer Lett.* **483**, 59–65 (2020).
203. Sinha, T., Panigrahi, C., Das, D. & Chandra Panda, A. Circular RNA translation, a path to hidden proteome. *Wiley Interdiscip. Rev. RNA* **13**, e1685 (2022).
204. Zhang, L., Gao, H., Li, X., Yu, F. & Li, P. The important regulatory roles of circRNA-encoded proteins or peptides in cancer pathogenesis (review). *Int. J. Oncol.* **64**, 19 (2023).
205. He, Y. et al. Targeting PI3K/Akt signal transduction for cancer therapy. *Signal Transduct. Target. Ther.* **6**, 425 (2021).
206. Lan, H. & Sun, Y. FBXW7 E3 ubiquitin ligase: Degrading, not degrading, or being degraded. *Protein Cell* **10**, 861–863 (2019).
207. Yang, Y. et al. Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis. *J. Natl. Cancer Inst.* **110**, 304–315 (2018).
208. Marín-Béjar, O. et al. Pint lincRNA connects the p53 pathway with epigenetic silencing by the polycomb repressive complex 2. *Genome Biol.* **14**, R104 (2013).
209. Zhang, M. et al. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nat. Commun.* **9**, 4475 (2018).
210. Matsumoto, A. et al. *MTORC1* and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature* **541**, 228–232 (2017).
211. Honkala, A., Malhotra, S. V., Kummars, S. & Junttila, M. R. Harnessing the predictive power of preclinical models for oncology drug development. *Nat. Rev. Drug Discov.* **21**, 99–114 (2022).
212. Winkle, M., El-Daly, S. M., Fabbri, M. & Calin, G. A. Noncoding RNA therapeutics—challenges and potential solutions. *Nat. Rev. Drug Discov.* **20**, 629–651 (2021).
213. Zhu, Y., Zhu, L., Wang, X. & Jin, H. RNA-based therapeutics: an overview and prospectus. *Cell Death Dis.* **13**, 644 (2022).
214. Liu, K. et al. AEG-1 3′-untranslated region functions as a ceRNA in inducing epithelial-mesenchymal transition of human non-small cell lung cancer by regulating miR-30a activity. *Eur. J. Cell Biol.* **94**, 22–31 (2015).
215. Qian, M. et al. BCL11B regulates MICA/B-mediated immune response by acting as a competitive endogenous RNA. *Oncogene* **39**, 1514–1526 (2020).
216. Hu, J. et al. The CCR2 3′UTR functions as a competing endogenous RNA to inhibit breast cancer metastasis. *J. Cell Sci.* **130**, 3399–3413 (2017).
217. Rutnam, Z. J. & Yang, B. B. The non-coding 3′ UTR of CD44 induces metastasis by regulating extracellular matrix functions. *J. Cell Sci.* **125**, 2075–2085 (2012).
218. Jeyapalan, Z. et al. Expression of CD44 3′-untranslated region regulates endogenous microRNA functions in tumorigenesis and angiogenesis. *Nucleic Acids Res.* **39**, 3026–3041 (2011).
219. Weng, J. et al. CD44 3′-untranslated region functions as a competing endogenous RNA to enhance NK sensitivity of liver cancer stem cell by regulating ULBP2 expression. *Int. J. Biol. Sci.* **15**, 1664–1675 (2019).
220. Li, X. et al. STARD13-correlated ceRNA network inhibits EMT and metastasis of breast cancer. *Oncotarget* **7**, 23197–23211 (2016).
221. Chen, P. C. et al. c-Myc acts as a competing endogenous RNA to sponge miR-34a, in the upregulation of CD44, in urothelial carcinoma. *Cancers* **11**, 1457 (2019).
222. Wang, B. et al. CYP4Z1 3′UTR represses migration of human breast cancer cells. *Biochem. Biophys. Res. Commun.* **478**, 900–907 (2016).
223. Zheng, L., Li, X., Gu, Y., Lv, X. & Xi, T. Correction to: The 3′UTR of the pseudogene CYP4Z2P promotes tumor angiogenesis in breast cancer by acting as a ceRNA for CYP4Z1. *Breast Cancer Res. Treat.* **179**, 521–522 (2020).
224. Zheng, L., Li, X., Gu, Y., Lv, X. & Xi, T. The 3′UTR of the pseudogene CYP4Z2P promotes tumor angiogenesis in breast cancer by acting as a ceRNA for CYP4Z1. *Breast Cancer Res. Treat.* **150**, 105–118 (2015).
225. Cheng, F. H. C. et al. E2F6 functions as a competing endogenous RNA, and transcriptional repressor, to promote ovarian cancer stemness. *Cancer Sci.* **110**, 1085–1095 (2019).
226. Yang, J. et al. FOXO1 3′UTR functions as a ceRNA in repressing the metastases of breast cancer cells via regulating miRNA activity. *FEBS Lett.* **588**, 3218–3224 (2014).
227. Yang, Z. P., Ma, H. S., Wang, S. S., Wang, L. & Liu, T. LAMC1 mRNA promotes malignancy of hepatocellular carcinoma cells by competing for MicroRNA-124 binding with CD151. *IUBMB Life* **69**, 595–605 (2017).
228. Wang, C., Mao, C., Lai, Y., Cai, Z. & Chen, W. MMP1 3′UTR facilitates the proliferation and migration of human oral squamous cell carcinoma by sponging miR-188-5p to up-regulate SOX4 and CDK4. *Mol. Cell Biochem.* **476**, 785–796 (2021).
229. Li, J. et al. Pecanex functions as a competitive endogenous RNA of S-phase kinase associated protein 2 in lung cancer. *Cancer Lett.* **406**, 36–46 (2017).
230. Huang, X. et al. PDL1 And LDHA act as ceRNAs in triple negative breast cancer by regulating miR-34a. *J. Exp. Clin. Cancer Res.* **36**, 129 (2017).
231. Liu, T. et al. PIK3C2A mRNA functions as a miR-124 sponge to facilitate CD151 expression and enhance malignancy of hepatocellular carcinoma cells. *Oncotarget* **7**, 43376–43389 (2016).
232. Wan, X. et al. Co-expression analysis revealed PTCH1-3′UTR promoted cell migration and invasion by activating miR-101-3p/SLC39A6 axis in non-small cell lung cancer: implicating the novel function of PTCH1. *Oncotarget* **9**, 4798–4813 (2018).
233. Li, J., Yu, H., Xi, M., Ma, D. & Lu, X. The SNAI1 3′UTR functions as a sponge for multiple migration-/invasion-related microRNAs. *Tumour Biol.* **36**, 1067–1072 (2015).
234. Li, J., Wang, J., Yue, H. & Lu, X. SNAI2 3′untranslated region promotes the invasion of ovarian cancer cells by inducing MARCKS expression. *J. Cancer* **10**, 2480–2487 (2019).
235. Zheng, L. et al. StarD13 3′-untranslated region functions as a ceRNA for TP53INP1 in prohibiting migration and invasion of breast cancer cells by regulating miR-125b activity. *Eur. J. Cell Biol.* **97**, 23–31 (2018).
236. Guo, X. et al. Displacement of Bax by BMF mediates STARD13 3′UTR-induced breast cancer cells apoptosis in a miRNA-dependent manner. *Mol. Pharm.* **15**, 63–71 (2018).
237. Zheng, L. et al. STARD13-correlated ceRNA network-directed inhibition on YAP/TAZ activity suppresses stemness of breast cancer via co-regulating Hippo and Rho-GTPase/F-actin signaling. *J. Hematol. Oncol.* **11**, 72 (2018).
238. Zhang, H., Wang, F. & Hu, Y. STARD13 promotes hepatocellular carcinoma apoptosis by acting as a ceRNA for Fas. *Biotechnol. Lett.* **39**, 207–217 (2017).
239. Wang, Z. et al. Downregulated USP3 mRNA functions as a competitive endogenous RNA of SMAD4 by sponging miR-224 and promotes metastasis in colorectal cancer. *Sci. Rep.* **7**, 4281 (2017).
240. Fang, L. et al. Versican 3′-untranslated region (3′-UTR) functions as a ceRNA in inducing the development of hepatocellular carcinoma by regulating miRNA activity. *FASEB J.* **27**, 907–919 (2013).
241. Wu, Q. et al. XIAP 3′-untranslated region as a ceRNA promotes FSCN1 function in inducing the progression of breast cancer by binding endogenous miR-29a-5p. *Oncotarget* **8**, 16784–16800 (2017).
242. Wu, W. Y., Tao, S. Q., Wang, X. N., Lobie, P. E. & Wu, Z. S. XIAP 3′-untranslated region serves as a competitor for HMG2 by arresting endogenous let-7a-5p in human hepatocellular carcinoma. *Tumour Biol.* **39**, 1010428317719578 (2017).
243. Zhu, S. et al. An oncopeptide regulates m6A recognition by the m6A reader IGF2BP1 and tumorigenesis. *Nat. Commun.* **11**, 1685 (2020).

244. Wu, S. et al. A novel micropeptide encoded by Y-linked LINC00278 links cigarette smoking and AR signaling in male esophageal squamous cell carcinoma. *Cancer Res.* **80**, 2790–2803 (2020).
245. Li, X. L. et al. A small protein encoded by a putative lncRNA regulates apoptosis and tumorigenicity in human colorectal cancer cells. *Elife* **9**, e53734 (2020).
246. Wang, Y. et al. LncRNA-encoded polypeptide ASRPS inhibits triple-negative breast cancer angiogenesis. *J. Exp. Med.* **217**, jem.20190950 (2020).
247. Zou, M. et al. Oncogenic activation of MAP kinase by BRAF pseudogene in thyroid tumors. *Neoplasia* **11**, 57–65 (2009).
248. Sun, C. et al. CRIPTO3, a presumed pseudogene, is expressed in cancer. *Biochem. Biophys. Res. Commun.* **377**, 215–220 (2008).
249. Zhang, M. et al. Pseudogene MAPK6P4-encoded functional peptide promotes glioblastoma vasculogenic mimicry development. *Commun. Biol.* **6**, 1059 (2023).
250. Moreau-Aubry, A. et al. A processed pseudogene codes for a new antigen recognized by a CD8(+) T cell clone on melanoma. *J. Exp. Med.* **191**, 1617–1624 (2000).
251. Lettnin, A. P. et al. Silencing the OCT4-PG1 pseudogene reduces OCT-4 protein levels and changes characteristics of the multidrug resistance phenotype in chronic myeloid leukemia. *Mol. Biol. Rep.* **46**, 1873–1884 (2019).
252. Li, T. et al. A novel peptide P1-121aa encoded by STK24P1 regulates vasculogenic mimicry via ELF2 phosphorylation in glioblastoma. *Exp. Neurol.* **367**, 114477 (2023).
253. Wang, T. et al. A novel protein encoded by circASK1 ameliorates gefitinib resistance in lung adenocarcinoma by competitively activating ASK1-dependent apoptosis. *Cancer Lett.* **520**, 321–331 (2021).
254. Peng, Y. et al. A novel protein AXIN1-295aa encoded by circAXIN1 activates the Wnt/ $\beta$ -catenin signaling pathway to promote gastric cancer progression. *Mol. Cancer* **20**, 158 (2021).
255. Zhao, W., Zhang, Y. & Zhu, Y. Circular RNA circ $\beta$ -catenin aggravates the malignant phenotype of non-small-cell lung cancer via encoding a peptide. *J. Clin. Lab. Anal.* **35**, e23900 (2021).
256. Liang, W. C. et al. Translation of the circular RNA circ $\beta$ -catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome Biol.* **20**, 84 (2019).
257. Gu, C. et al. CHEK1 and circCHEK1\_246aa evoke chromosomal instability and induce bone lesion formation in multiple myeloma. *Mol. Cancer* **20**, 84 (2021).
258. Zhang, Y. et al. CircDIDO1 inhibits gastric cancer progression by encoding a novel DIDO1-529aa protein and regulating PRDX2 protein stability. *Mol. Cancer* **20**, 101 (2021).
259. Gao, X. et al. Circular RNA-encoded oncogenic E-cadherin variant promotes glioblastoma tumorigenicity through activation of EGFR-STAT3 signalling. *Nat. Cell Biol.* **23**, 278–291 (2021).
260. Li, Y. et al. circ-EIF6 encodes EIF6-224aa to promote TNBC progression via stabilizing MYH9 and activating the Wnt/ $\beta$ -catenin pathway. *Mol. Ther.* **30**, 415–430 (2022).
261. Chen, C.-K. et al. Structured elements drive extensive circular RNA translation. *Mol. Cell* **81**, 4300–4318.e13 (2021).
262. Pan, Z. et al. A novel protein encoded by circFNDC3B inhibits tumor progression and EMT through regulating Snail in colon cancer. *Mol. Cancer* **19**, 71 (2020).
263. Gu, C. et al. circGprc5a promoted bladder oncogenesis and metastasis through Gprc5a-targeting peptide. *Mol. Ther. Nucleic Acids* **13**, 633–641 (2018).
264. Song, J. et al. A novel protein encoded by ZCRB1-induced circHEATR5B suppresses aerobic glycolysis of GBM through phosphorylation of JMJD5. *J. Exp. Clin. Cancer Res.* **41**, 171 (2022).
265. Li, J. et al. Circular HER2 RNA positive triple negative breast cancer is sensitive to Pertuzumab. *Mol. Cancer* **19**, 142 (2020).
266. Jiang, T. et al. A novel protein encoded by circMAPK1 inhibits progression of gastric cancer by suppressing activation of MAPK signaling. *Mol. Cancer* **20**, 66 (2021).
267. Wang, L. et al. A novel tumour suppressor protein encoded by circMAPK14 inhibits progression and metastasis of colorectal cancer by competitively binding to MKK6. *Clin. Transl. Med.* **11**, e613 (2021).
268. Duan, J. L. et al. A novel peptide encoded by N6-methyladenosine modified circMAP3K4 prevents apoptosis in hepatocellular carcinoma. *Mol. Cancer* **21**, 93 (2022).
269. Liang, Z. X. et al. A novel NF- $\kappa$ B regulator encoded by circPLCE1 inhibits colorectal carcinoma progression by promoting RPS3 ubiquitin-dependent degradation. *Mol. Cancer* **20**, 103 (2021).
270. Zheng, X. et al. A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling. *Mol. Cancer* **18**, 47 (2019).
271. Zhang, M. et al. A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis. *Oncogene* **37**, 1805–1814 (2018).
272. Wu, X. et al. A novel protein encoded by circular SMO RNA is essential for Hedgehog signaling activation and glioblastoma tumorigenicity. *Genome Biol.* **22**, 33 (2021).
273. Lyu, Y. et al. A novel protein encoded by circUBE4B promotes progression of esophageal squamous cell carcinoma by augmenting MAPK/ERK signaling. *Cell Death Dis.* **14**, 346 (2023).
274. Li, F. et al. 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**, 101182 (2021).
275. Yang, F. et al. p113 isoform encoded by CUX1 circular RNA drives tumor progression via facilitating ZRF1/BRD4 transactivation. *Mol. Cancer* **20**, 123 (2021).

## ACKNOWLEDGEMENTS

We thank Maurizio S. Podda and all Polisen lab members for helpful discussions. We acknowledge many investigators in the field whose primary data could not be cited. This work was supported by ISPRO-Istituto per lo Studio, la Prevenzione e la Rete Oncologica (institutional funding to L.P.). It was also partially supported by AIRC-Associazione Italiana Ricerca sul Cancro (IG #25694 to L.P.). P.P.P. has been supported by the Renown Foundation and the PTEN Research Foundation. Cartoons created with BioRender.com.

## AUTHOR CONTRIBUTIONS

L.P. and P.P.P. conceptualized the review. M.L. gathered the information and created the figures. All authors contributed to writing and editing the text.


## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Laura Polisen or Pier Paolo Pandolfi.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024