



Review

The Functional Meaning of 5'UTR in Protein-Coding Genes

Natalia Ryczek ^{*,†} , Aneta Łyś [†] and Izabela Makałowska ^{*}

Institute of Human Biology and Evolution, Adam Mickiewicz University in Poznań, Uniwersytetu Ponańskiego 6, 61-614 Poznań, Poland

* Correspondence: nataliaryczek.nr@gmail.com (N.R.); izabela.makalowska@amu.edu.pl (I.M.)

† This authors contributed equally to this work.

Abstract: As it is well known, messenger RNA has many regulatory regions along its sequence length. One of them is the 5' untranslated region (5'UTR), which itself contains many regulatory elements such as upstream ORFs (uORFs), internal ribosome entry sites (IRESs), microRNA binding sites, and structural components involved in the regulation of mRNA stability, pre-mRNA splicing, and translation initiation. Activation of the alternative, more upstream transcription start site leads to an extension of 5'UTR. One of the consequences of 5'UTRs extension may be head-to-head gene overlap. This review describes elements in 5'UTR of protein-coding transcripts and the functional significance of protein-coding genes 5' overlap with implications for transcription, translation, and disease.

Keywords: 5'UTR; protein-coding genes; uORF; head-to-head genes overlap; miRNA

1. Introduction

The 5' untranslated region is the mRNA domain that contains plenty of elements such as upstream ORFs (uORFs), internal ribosome entry sites, microRNA binding sites, and structural components involved in the regulation of mRNA stability, pre-mRNA splicing, and translation initiation. Deregulation of cis-regulatory elements or secondary structures within the 5'UTRs may cause a change in gene expression [1]. This shows the functional importance of 5'UTR in control of gene expression. Furthermore, a growing number of evidence demonstrates that mutations within 5'UTRs are often linked with diseases, including cancer [2–4]. Oncogenes and tumor suppressors require precise regulation and often express transcripts containing various and atypically long 5' untranslated regions with new regulatory elements, such as uORFs or secondary structures [5,6]. Moreover, extension of 5' UTR due to the use of an alternative promoter may lead to an overlap with a gene located on the opposite DNA strand. Such overlap at 5' ends may be associated with additional regulatory functions [7,8]. This review summarizes the roles of 5'UTRs and the regulatory mechanisms in which these sequences are involved.

2. Upstream Open Reading Frames

The recent development of ribosome profiling techniques enabled the discovery of many translationally active uORFs in the human genome [9]. These upstream reading frames can both produce proteins called uPeptides and have a regulatory role related to translation control of main ORF (mORF). It has also been shown that the most common start codons for uORFs are canonical AUG and non-canonical CUG, not AUG, AAG, and AGG as previously thought [10]. The functional changes in the AUG and CUG codons were demonstrated in samples from cancer patients, which may indicate their relationship with cancer biology [10]. The use of non-canonical start codons results from interactions with various trans-acting initiation factors and structural elements within mRNA [11]. In addition, translation initiation from non-canonical start codons results from, among others, stress conditions and may play key roles in the different regulatory mechanisms [12].



Citation: Ryczek, N.; Łyś, A.; Makałowska, I. The Functional Meaning of 5'UTR in Protein-Coding Genes. *Int. J. Mol. Sci.* **2023**, *24*, 2976. <https://doi.org/10.3390/ijms24032976>

Academic Editors: Naoyuki Kataoka and Rui Henrique

Received: 28 December 2022

Revised: 20 January 2023

Accepted: 26 January 2023

Published: 3 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Interestingly, not all uORF are translated into a protein (Figure 1A) and some have a regulatory function. An example is gene *ATF4*, whose expression is unregulated in response to stress, which is regulated by uORFs present in 5'UTRs. The mechanisms involved in *ATF4* regulation are called leaky scanning. In this phenomenon, ribosome bypass uORF without uProtein formation [13–15]. This results in a positive regulation of mORF translation by enhanced reinitiation after termination of uORF translation. The same uORF can be translated into uPeptide under stress-free conditions, down-regulating mORF expression [16–18]. uORF leaky scanning was observed for the *ATF4* gene in head and neck squamous carcinomas. It has been proven that up-regulation of the *DDX3* gene, whose product plays a regulatory role in translation, increases the translation of the *ATF4* mRNA through leaky scanning of its uORF. The ATF4 transcriptional factor positively influences the progression of neoplastic cells, enhancing their mesenchymal properties [19]. The mechanism of leaky scanning resulting in the enhanced translation of mORFs was also observed in the *Arabidopsis thaliana*. Under the hypoxia condition, the increased ribosome presence in the mORFs regions was noticed, which results from the reduced uORF translation. Such mechanism may be a response to the challenging conditions [20].

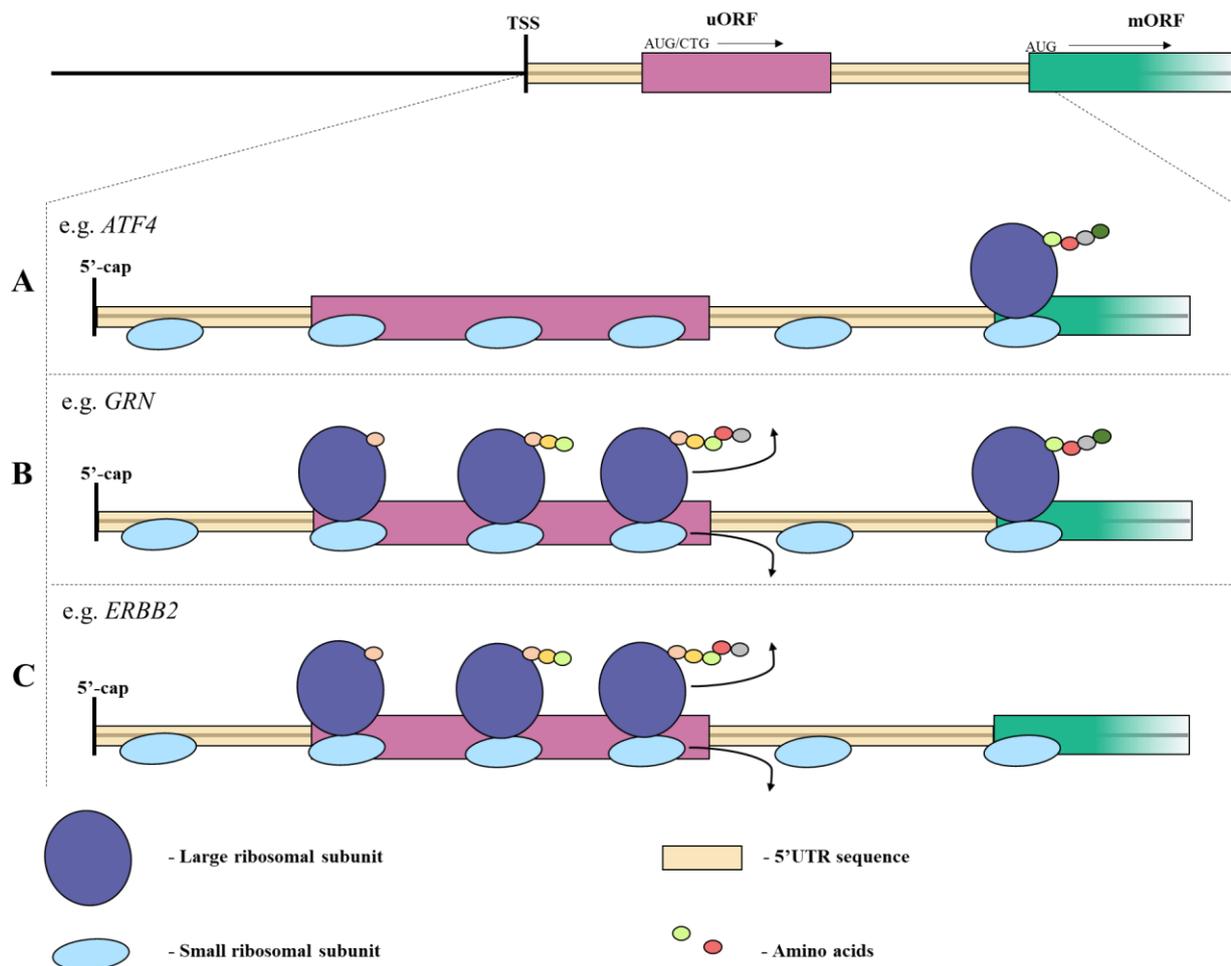


Figure 1. The most common uORF-mediated translation regulation pathways. (A) The leaky scanning phenomenon—uORF ribosome bypass without uProtein formation; (B) Production of uProteins by ribosomes, partial disassociation of the translational factors, and re-initiation of the mORF translation; (C) Production of uProteins by ribosomes, disassociation of the translational factors, and no translation of mORF.

Leaky scanning was also observed in non-stressful cell conditions. The mRNA of the *BACE1* gene consists of six uORFs, which inhibit or reduce the translation of the *BACE1*

protein. Deregulation of the inhibitory uORFs leads to increased *BACE1* expression, which was linked with the pathogenesis of Alzheimer's disease [21,22]. Another example is the uORF located in the *APOBEC3G* mRNA that serves as repressor for mORF translation, regulating *APOBEC3G* expressions through leaky scanning, and re-initiation mechanism. Interestingly, the same uORF is used by the HIV-1 Vif protein to repress the *APOBEC3G* translation and redirect mRNA into stress granules. This repression is beneficial for the virus because the product of this gene inhibits its replication [23].

Another mechanism that regulates the translation of mORFs by uORFs is related to disassociation of non-essential factors. When non-essential factors disassociate after uORF translation e.g., the large ribosomal subunit, but the eIF3 stays on mRNA, the remaining factors necessary to initiate mORF translation may be re-recruited (Figure 1B) [24–26]. However, the efficiency of this re-recruitment depends on the distance between uORF and mORF [26]. An example of such regulation is the *GRN* gene which was identified to express transcripts with short and long 5'UTRs. An uORF within the long 5'UTR of *GRN* mRNA inhibits translation of the progranulin protein. It has been demonstrated that transcripts of the *GRN* gene with shorter 5'UTR do not contain an uORF, hence these shorter transcripts are not translationally repressed. The presence of the long 5'UTR is associated with a reduction of *GRN* levels [27]. This regulation mechanism was also observed in squamous cell carcinoma (SCC), where the uORF translation initiation of the *SOX2* gene is an important mechanism of tumor initiation properties [28]. Repressed expression of progranulin from long 5'UTR mRNA increases the risk for frontotemporal lobar degeneration (FTLD), which belongs to a group of progressive brain diseases [29]. Alternatively spliced 5'UTRs of the oncosuppressor *DLG1* represent another example on how different 5' regions could impact gene expression and carcinogenesis. The longer *DLG1* 5'UTR contains an upstream short ORF, which inhibits translation of a downstream open reading frame [30]. It has been demonstrated that a higher expression level of transcripts with shorter 5'UTR results in an increased *DLG1* protein level during epithelial junctions' formation in colon cells. On the other hand, a higher amount of the larger mRNA isoform, causing a decrease in *DLG1* protein level, was observed during monocyte-to-macrophage differentiation [31]. Moreover, the results of RT-qPCR analysis have shown up-regulation of the large 5'UTR *DLG1* mRNA isoform in cells with malignant potential [30]. Therefore, it is plausible that the mechanism of the 5'UTR splicing could control *DLG1* protein abundance and potentially impact oncogenic processes by changing the *DLG1* levels [5].

uORFs as regulatory elements are able to control mORFs in a cis-acting manner, *inter alia*, via ribosomal subunits in eukaryotes [24,32]. The 43S preinitiation complex (PIC) PIC is made of many different proteins, including: eIF1, eIF1A, eIF2 ternary complex, eIF3, and 40S small ribosomal subunit [33–35]. The 43S PIC attaches to a 5'-cap structure of mRNA and slides the mRNA strand towards the 3' end in search of a translation initiation codon [25,36]. Repression of mORF occurs when the critical translation initiation factors, e.g., eIF3 or eIF2, ternary complex disconnect after translating the uORFs and do not bind again downstream to initiate the translation of mORF (Figure 1C). For example, in *Saccharomyces cerevisiae*, the *NDC80* gene expresses two mRNA isoforms: the canonical *NDC80* mRNA isoform (*NDC80^{ORF}*) and 5'-end-extended *NDC80* long undecoded transcript isoform (*NDC80^{luti}*). In the elongated region, *NDC80^{luti}* contains the regulatory uORFs that is blocking translation into the Ndc80 protein [37,38]. This mechanism was also observed in breast cancer, where the Her-2 receptor encoded by *ERBB2* gene is over-expressed compared to the non-cancer tissues. In non-cancer cells the Her-2 translation from mORF is repressed by the presence of uORF in 5'UTR. The over-expression of Her-2 in breast cancer cells is mediated by its 3'UTR interactions with uORF, blocking the formation of the uPeptide [39].

There is also a possibility that specific uORFs will repress the mORF translation through a particular peptide sequence [40,41]. The uORFs presence in the 5'UTRs may also affect the stability of mRNA through the nonsense-mediated mRNA decay (NMD) process [42,43]. However, not all uORFs can lead to NMD. Most often it concerns uORFs

with canonical start codons. When translation can only be initiated from non-canonical start codons after canonical ones are eliminated, the resulting mRNAs are mostly not sensitive to NMD [44].

It has been shown that the genes in cancer cells have more uORFs than in healthy cells [45,46]. It has also been proven that mutations that arise within uORFs can lead to diseases in humans. An example is gonadal dysgenesis that is caused by the mutations in the uORF of the human *SRY* gene [47,48]. The presence of the uORF mutation in the human *IRF6* gene contributes to the emergence of the Van der Woude and popliteal pterygium syndromes [49] and in the human *GCH1* gene uORF can lead to symptoms of the Levodopa-responsive dystonia [50]. An interesting case of mutation in the 5'UTR of the human *HAMP* gene encoding hepcidin antimicrobial peptide was identified in the Portuguese Family. Mutation was associated with the appearance of a new translation start codon within the Kozak sequence. As a result, ribosomes translate the uORF protein and not the main hepcidin antimicrobial peptide, which leads to the hereditary hemochromatosis in young people [51]. A mutation identified in the 5'UTR of the *CDKN2A* tumor suppressor has a similar consequence and leads to translation from a novel AUG. It has been found that this mutation predisposes to melanoma [52]. Another example is an increased expression level of *TGF-β3* due to the G to A transition in the eleventh uORF. This mutation causes arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC) [53]. An expansion of the GGC repeats in the 5' UTR of the *N2C* gene leads to the translation of the uORF and emergence of the polyglycine rich uPeptide. This protein is accumulated in the cells intranuclear inclusions leading to the neuronal intranuclear inclusion disease [54]. In patients with pituitary adenoma and well-differentiated pancreatic cancer, a 4-bp deletion was identified within the 5'UTR sequence of the gene encoding the cyclin-dependent kinase inhibitor p27 (KIP1)—*CDKN1B*. The role of this atypical tumor suppressor is to regulate the cycle, proliferation, and differentiation of the cell. The above-described deletion is located within the uORF sequence and causes a stop codon shift. Consequently, a longer uPeptide is formed and the distance to the mORF is shortened, which leads to a decrease in the *CDKN1B* and p27 protein level [55]. It has also been shown that the loss of the start codon by uORF of the *CEBPB* gene leads to the formation of two short forms of proteins: C/EBP α and C/EBP β . Moreover, increased levels of the short form of C/EBP α have been associated with acute myeloid leukemia and C/EBP β with breast cancer [56].

Peptides resulting from the translation of uORF can also play a regulatory role in various processes. It has been shown that a uPeptide called MP31 arising from the uORF of the *PTEN* gene limits the lactate-pyruvate conversion process taking place in mitochondria [57]. Another example is a uORF of protein kinase C-eta (PKC- η) mRNA that encodes a peptide (uPEP2). This peptide exhibits the kinase inhibitory activity, which through typical protein kinase C pseudo-substrate motif auto-inhibits the catalytic kinase activity of all members of the protein kinase C family. In consequence, uPEP2 positively affects chemotherapy of breast cancer through inhibiting the cancer cell proliferation, survival, invasion, and metastasis [58]. There are also other examples of peptides generated from uORF located in 5' UTRs of genes, including: *UL4*, *TRPM7*, *DDIT3*, *PPP1R15A*, and *PTP4A-2* [24,59].

3. Head-to-Head Overlapping Genes Phenomenon

Genes that sequences partially or completely share the locus on the same or opposite DNA strand are defined as overlapping genes [7]. The first overlapping genes were already identified in 1969 in the coliphage genome [60] and soon were found to be common in viral and bacterial genomes [61]. However, until the beginning of this century, there was limited evidence of the existence of overlapping protein-coding genes in eucaryotic genomes. Nevertheless, multiple studies have revealed that this phenomenon is more common than previously thought. Overlapping genes have been detected in many organisms such as plants [62], yeast [63], fish [7], flies [64], mice [65], and humans [66,67]. Based on the recent analysis, Chen et al. claimed that 25.8% of the human protein-coding genes overlap [68]. Ho et al. identified 2541 overlapping gene pairs in the human genome. Among them,

473 pairs overlap in the head-to-head orientation (i.e., with their 5' ends) [69]. Results of bioinformatics research conducted on samples from 73 human tissues and cell lines revealed that from a total of 15,778 of protein-coding genes, 582 pairs overlap at their 5' ends [70].

The majority of 5' overlapping genes were identified based on bioinformatics analyzes. However, some were discovered experimentally. Two examples are human *DLG4* and *VLCAD* genes. The 5'-untranslated region of *DLG4* overlaps with the entire 5'UTR and 62 bp of the coding sequence of *VLCAD*. Despite overlapping at the 5' ends and sharing regulatory components, both genes are highly expressed at the same time in various tissues [71]. Overlapping genes could show tissue-specific expression patterns [72,73], co-expression [71], and co-regulation [74,75]. Some studies suggest that coordinated gene expression is common for genes located on opposite strands and overlapping in the 5'UTRs [76]. On the other hand, there are also works showing that the phenomena of gene overlap could lead to promoter competition and a negative expression correlation of involved genes [69,73]. However, recent studies demonstrated that this genomic arrangement may lead to a higher expression level of at least one gene from overlapping pair [70,77].

In the majority of cases, analysis of the human transcriptome have shown that an overlap between two genes located on opposite DNA strands is not a stable feature and depends on which alternative transcription start site (TSS) is activated. The utilization of more distal TSS causes an extension of the 5'UTR, resulting in genes' overlapping [70]. As an example, when simultaneously utilizing alternative TSSs, human genes *FBXL15* and *PSD* were identified to overlap at their 5' ends in brain tissue. The overlap between these genes occurs only when distal TSSs are used [78]. The overlap at the 5' ends is thought to be involved in various regulatory events, such as transcriptional interference [79] and RNA:RNA duplex formation [80]. However, despite the number of studies, the biological meaning of this genomic architecture is still debatable.

4. RNA Duplexes

Owing to sequence complementarity transcripts of overlapping genes may form RNA:RNA duplexes, which can affect transcription and translation of involved genes [80]. The formation of RNA:RNA duplexes may interfere with the alternative splicing process [81], cover miRNA-binding sites [82,83], or influence mRNA translation [84]. Kudla et al., have proved that the blockage of mRNA 5'UTR in *E. coli* results in poor translation efficiency [85]. However, the detection of RNA:RNA duplexes is technically difficult, mainly due to the instability of double-stranded RNA in eukaryotic cells. An additional complication in the identification of mRNA duplexes comes from the fact that this RNA:RNA hybrid can trigger a process leading to cutting double-stranded RNA into short duplexes—precursors of endo-siRNA [86,87]. It has also been demonstrated that the currently used methods have too low of a sensitivity and specificity to allow accurate identification of duplexes. So far, only a few RNA:RNA duplexes between two protein-coding transcripts have been experimentally validated. The *Wrap53* and *p53* gene pair is one of the best-characterized examples, for which RNA duplexes formation was confirmed in human cell lines [88].

A key tumor suppressor gene, *p53*, is a transcription factor whose function has been identified in numerous biological processes, including metabolism, senescence, or cell cycle arrest and apoptosis in response to cellular stress [88–90]. To avoid this response, *p53* is often inactivated in cancer cells, allowing survival and tumor progression [89]. The identification and characterization of natural antisense transcript *Wrap53* positioned opposite to the *p53* DNA strand revealed that both genes overlap at the 5' ends (Figure 2A) [88]. The *Wrap53* termed for WD40-encoding RNA antisense to *p53* encodes the WRAP53 protein (also denoted WDR79), crucial for cellular trafficking of small Cajal body-specific RNAs (scaRNAs) and recruitment of the telomerase enzyme to Cajal bodies [91–93]. *Wrap53* utilizes three alternative transcriptional start sites (1α , 1β , and 1γ) [92]. Usage of distal TSS (1α) results in the elongation of 5'UTR of the transcript described as WRAP53- 1α . WRAP53- 1α directly overlaps the first exon of the *p53* gene and regulates *p53* mRNA

at the transcriptional level [88]. Studies revealed that both genes interact at the 5' ends, increasing the stability of *p53* mRNA and enhancing the production of the p53 protein in response to cellular stress [88,94]. It has been found that *Wrap53* regulates *p53* via RNA:RNA hybridization (Figure 2B). Inhibition of *Wrap53/p53* duplex formation reduces *p53* abundance, confirming that this RNA:RNA interaction protects *p53* from degradation and sustains its expression in human cells [88]. Since the *p53* expression level is regulated by *Wrap53*, any alterations of *Wrap53* levels could contribute to tumorigenesis [92]. Over-expression of *Wrap53* has been associated with several types of tumors, including head and neck squamous cell carcinoma (HNSCC) [94,95], development of esophageal squamous cell carcinoma (ESCC) [96], or tumor progression in non-small cell lung cancer (NSCLC) [97].

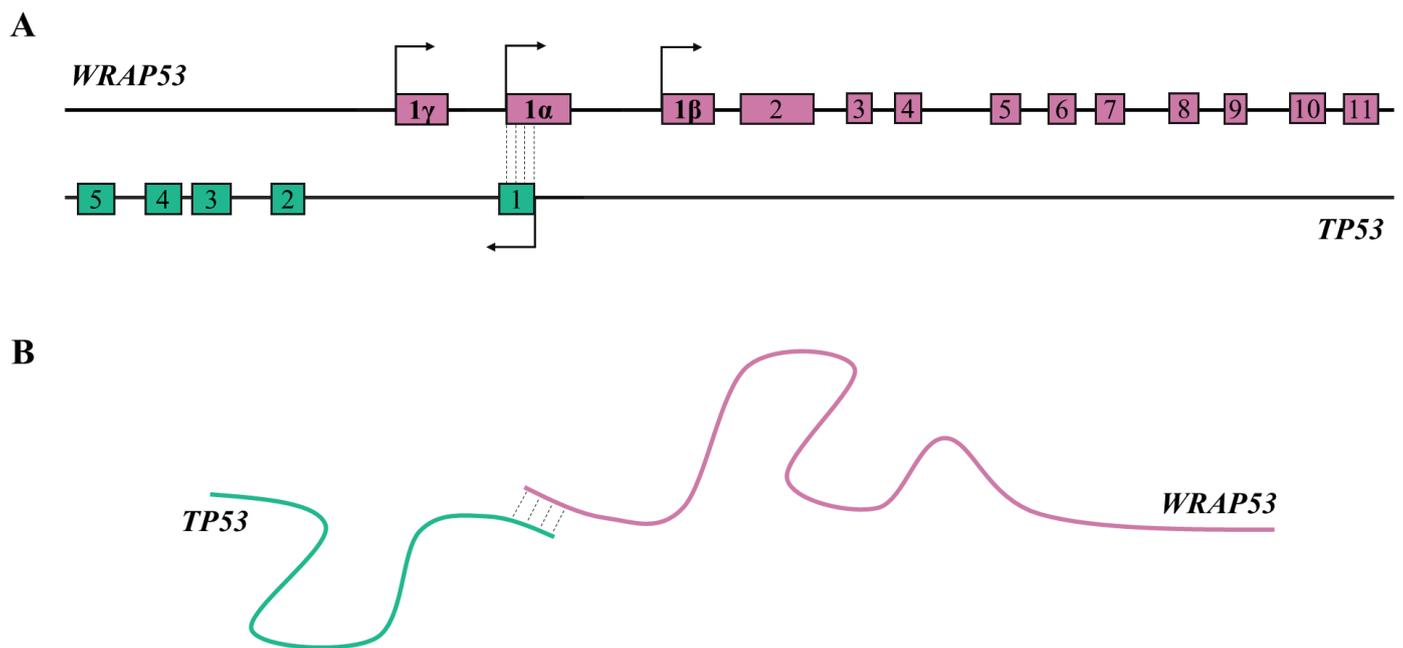


Figure 2. The *WRAP53* and *TP53* genes (A) and its RNAs duplex formation (B). The dotted lines demonstrate overlapping region between these genes (A) or created RNA duplex (B) and the arrows show the direction of transcription. Numbered regions indicate the coding sequences/exons of the *WRAP53* and *TP53* genes.

Recent studies have shown that two transcript isoforms of the *Wrap53* gene have distinct and opposite functions in cancer cell lines. Elongated at 5'UTR isoform *WRAP53-1α* regulates *p53* expression, while the shorter transcript *WRAP53-1β* has no regulatory effect on p53 [98]. However, *WRAP53-1β* plays an important function in the repair of DNA double-strand breaks [99]. It has been shown that knockdown of *WRAP53-1α* suppresses cell migration in the A549 cell line, whereas depletion of *WRAP53-1β* promoted cell migration in these cells. Moreover, *WRAP53-1α* deficiency promoted H1975 cell invasion, but knockdown of *WRAP53-1β* had no significant effect on H1975 cells. Altogether, these results suggest that compared to the shorter isoform, the 5'UTR elongated transcript of *Wrap53* has different effects on *p53* and NSCLC cells [98].

Several consequences of the 5'-ends interaction between *p53* and *Wrap53* have been shown over the years. Interestingly, bioinformatics research revealed that in breast cancer the part of 5'UTR of *Wrap53* mRNA that overlaps the *p53* transcript contains a binding site for has-miR-4732-5p. Therefore, has-miR-4732-5p miRNA by binding to 5'UTR of the *Wrap53*, may block *Wrap53/p53* duplex formation and disturb the *p53* level [83].

There are some other instances of the regulatory role of RNA:RNA duplexes formed by two mRNA sequences. In human cells, splicing of *TRα1* is regulated via mRNA-mRNA interactions with *Rev-erba* [100]. Moreover, in gastric cancer, human genes *WDR83* and *DHPS* regulate and increase their mutual stability via the formation of a RNA:RNA

duplex [101]. An example also comes from *Arabidopsis thaliana*, where formation of RNA duplexes with regulatory consequences was detected for *POR1-OCA2* [102] and *SRO5-P5CDH* gene pairs [103].

Approaches focusing on identifying RNA duplexes in living cells have been significantly improved over the past years. More and more evidence for duplexes formed by protein-coding genes transcripts is coming from high-throughput techniques. Sharma et al. described the LIGR-seq method, which allows the study of global-scale RNA:RNA interactions in vivo. However, the yield of RNA–RNA interaction detected in this study was very low and constituted only 1029 reads [104]. In more precise methods called PARIS [105,106], 232,031 reads of RNA–RNA interactions were obtained but it is unknown how many of them were formed by two mRNAs [107]. A high-throughput method named SPLASH resulted in 4026 reads, including 990 mRNA–mRNA interactions in human cell lines [108]. All these methods enable global RNA interactome analysis but still have to be refined and new approaches need to be developed.

5. Transcriptional Interference

The phenomena of overlapping genes could lead to the downregulation of both genes via transcriptional interference mechanisms. TI occurs when one transcriptional process inhibits another transcriptional process taking place on the opposite DNA strand at the same time [79]. Over the past few years, the regulation of gene expression through transcriptional interference has been confirmed in various organisms, from viruses to metazoans [109]. Mechanisms of TI have been observed in *Escherichia coli* [110] and *Saccharomyces cerevisiae* [111]. Transcriptional interference has been also identified to play an important role during embryonic development in *Drosophila melanogaster* [112].

Four main mechanisms of transcriptional interference have been defined: promoter competition in the initiation phase of transcription, polymerase collisions, ‘sitting duck’ interference, and occlusion in the elongation phase of transcription [79]. Promoter competition occurs when promoters of head-to-head overlapping genes compete for the RNA polymerase II (RNAP II) complex [79]. In eukaryotes, promoter competition also happens when two promoters of overlapping genes share the same transcriptional factors binding sites. When these factors bind to the enhancers of one promoter, the second promoter enhancers are prevented from interacting and activating [113]. The mechanism of polymerase collision occurs in the elongation phase when the RNAP II complex from one strand acts as a physical barrier and prevents the RNAP II complex on the other strand from progressing [114,115]. In the head-to-head overlapping region, two proceeding polymerase complexes collide, resulting in untimely transcription termination and a decrease in the expression level of both involved genes [79,80]. According to the results of a bioinformatics study, the longer the overlap region, the greater the probability of RNAP II complexes collision [114]. ‘Sitting duck’ interference is considered when the weaker promoter’s RNAP II complex slowly progresses from the open complex to the extension complex and might be pushed out by the elongation complex from the stronger, convergent promoter [79]. The elongation of RNA polymerase may also inhibit transcription from the downstream promoter in a process described as promoter occlusion [116]. An extending RNAP II launched from the tandem promoter temporarily occupies the downstream promoter and its activator binding sites. As a result, the time in which RNAP II can bind to the downstream promoter is limited [79]. Transcriptional interference could lead to a negative expression correlation of overlapping genes [69,73]. However, recent studies have demonstrated that the utilization of overlapping TSSs may lead to, on average, higher genes expression [70,77]. Studies performed by Rosikiewicz et al. [70] on TSS-seq data from 73 human tissues and cell lines also showed that genes overlapping at the 5′ ends do not have a negative expression correlation.

Transcriptional interference is not limited to overlapping genes. Tandem transcriptional interference applies to the situation when the process of transcription from an upstream promoter inhibits the transcription of a co-oriented downstream promoter [117]. Brar et al., have shown that in budding yeast over 190 genes produce 5′UTR elongated

mRNA isoforms. It has been confirmed that in some of them the expression of 5' extended mRNA inhibits the expression from the downstream promoter [118,119]. For example, in *Saccharomyces cerevisiae*, TI occurs during cell differentiation and regulates the expression of the *NDC80* gene [37]. The *NDC80* gene encodes a subunit of the Ndc80 outer kinetochore complex, which is required for the connection of spindle microtubules to kinetochores during yeast meiosis [38,120]. In meiotic cells, the *NDC80* gene expresses two mRNA isoforms. During the prophase stage of meiosis, the expression of the canonical *NDC80* mRNA isoform (*NDC80^{ORF}*) is repressed by the transcription of a 5'-end-extended *NDC80* long undecoded transcript isoform (*NDC80^{luti}*) [37,38]. The transcription of a 5' elongated *NDC80* mRNA isoform provides a repressive chromatin state and leads to transcriptional interference of a downstream promoter of *NDC80^{ORF}*. Results of these studies demonstrate that the 5'UTR elongated *NDC80* transcript isoform plays a regulatory role and its expression impacts the cell by causing transcriptional interference at the *NDC80^{ORF}* promoter during meiotic prophase [37]. Since transcripts with elongated 5'UTR are also expressed in higher eukaryotes transcriptional interference caused by 5'UTR, elongated isoforms might be widespread across species [37]. Interestingly, in human prostate cancer cells, the constitutively activated upstream promoter of the *ATF3* gene leads to a higher expression of *ATF3* in stress response, while expression from the downstream promoter is suppressed [121]. It is proposed that the upstream promoter of the human *ATF3* gene could inhibit the transcription from the downstream promoter due to delayed transcriptional interference [79,106].

6. 5'UTR Sequences as Potential miRNA and Protein Binding Sites

6.1. 5'UTRs and Interactions with miRNAs

The regulation of gene expression by the presence of miRNA binding sites in mRNAs is well known [122–124]. Most of the interactions between miRNAs and mRNAs are related to the 3'UTRs, which has been confirmed by many studies [125–128]. However, there is strong evidence of a major role for the 5'UTR sequence in miRNA drive regulation of expression [129]. It has also been shown that structures present at the 5'-end of transcripts may influence interactions with miRNAs and the higher the degree of secondary structures in 5'UTR fragment, the greater the chance of these interactions. The big importance of secondary structures close to the 5' cap site on the formation of interactions between the miRNA and the 5'UTRs has been also demonstrated [130].

There is evidence for the positive miRNA regulation of mRNAs encoding ribosomal proteins (RP). miR-10a by joining with the 5'UTR sequence of RP mRNAs enhances the synthesis of ribosomal proteins positively influencing their biogenesis. It has been shown that most RP mRNAs regulated in this way had a 5'TOP motif associated with the presence of C residues at the sequence beginning [131]. Another example comes from colorectal cancer (CRC), where miRNA—miR-532-5p is involved in the process of carcinogenesis. Increased expression of this miRNA has been associated with a decreased expression of the runt-related transcription factor 3 (*RUNX3*) gene in CRC. Studies conducted on human HT—29 CRC cells have shown that miR-532-5p binds specifically to the 5'UTR region of the *RUNX3* mRNA, which in turn increases the viability and proliferation of cancer cells, positively influencing their expansion processes [132].

There is also evidence for miRNA targeting 5'UTRs in viruses. An interesting mechanism was noticed during the replication of the hepatitis C virus (HCV). This virus has a positive (+) strand RNA. In the 5'UTR region of this viral sequence, there are two hepatic-specific miR-122 binding sites. The attachment of this miRNA to both binding sites within the 5'UTR of the viral genome enables up-regulation of replication [133]. Several binding sites for different miRNAs have also been described in the 5'UTR of SARS-CoV-2, suggesting that also here they may have a regulatory function [134,135].

6.2. 5'UTRs and Interactions with RNA Binding Proteins

The 5'UTR sequence has a critical role in the recruitment of ribosomes to mRNA as well as in many processes related to the mechanisms regulating translation. Translation control by 5'UTR may result from direct induction or by obliterating the ability to bind to RNA binding proteins (RBPs). An example of translation control through 5'UTR-protein interactions is the *STAT3* mRNA translation inhibition through poly(rC)-binding protein 1 (PCBP1). Two leucine residues of the PCBP1 protein bind to *STAT3* 5'UTR and, by reducing the amount of oncogenic protein STAT3, serve as a tumor suppressor [136]. Another study has shown that translation of the *ELAVL4* gene is associated with alternative 5'UTR sequences of the distinct mRNA isoforms in neuronal development. This translation is regulated by RNA binding protein—Celf1 in the development of glutamatergic neurons. Dysregulation of these mechanisms can lead to neurological disorders and tumor formation [137]. In *Caenorhabditis elegans*, the LIN41 protein represses translation of the *lin-29A* gene through interaction with the 5'UTR of *lin-29A* mRNA. It has been suggested that the attachment of LIN41 to the 5'UTR sequence of the *lin-29A* gene mediates mRNA degradation by recruiting other proteins, including ribonucleases [138]. There is evidence that some RBP may act differently on UTR sequences in a position-dependent manner by binding to the 5'UTRs or 3'UTRs. Examples of this mechanism are metazoan iron regulatory proteins (IRPs). When they interact with the 5'UTRs, they inhibit mRNA translation, while when they bind to the 3'UTRs sequence, they increase the stability of the mRNA, making it less prone to degradation [139]. A similar mechanism was also observed in *Drosophila melanogaster*, where in female flies there is a need for *MSL-2* mRNA translation inhibition. The Sex-lethal (SXL) protein binds both 5'- and 3'UTR. When SXL binds to the *MSL-2* 3'UTR sequence it inhibits the attachment of 43S ribosomal preinitiation complexes and when it binds to the 5'UTR sequence, it blocks the ribosomal complexes scanning for the translation initiation codon of mRNA [140].

The 5'UTRs contain internal ribosome entry sites (IRESs) that enable the binding of the 40S ribosomal subunits to the mRNA [141]. Therefore, mutations in IRES could impact translation and as a result various diseases [142,143]. In the nervous system, the C to T transition in *Connexin-32* mRNA leads to the loss of IRES-mediated translation of *Connexin-32*. This mutation is responsible for the development of X-linked Charcot–Marie–Tooth disease (CMTX) [144]. It has also been reported that mutation in the IRES of the *c-Myc* mRNA enhanced its translation and influenced multiple myeloma progressions [145,146].

The 5'UTR sequence has its role in the regulation of mouse insulin biosynthesis. Proteins PABP, HuD and Protein-Disulfide Isomerase are involved in the regulation of glucose level. In the low/basal glucose conditions the PABP and HuD form a translation inhibitory complex by binding to the 5'UTR of the *Ins2* gene mRNA. When the glucose level rises, this complex is affected by Protein-Disulfide Isomerase, causing insulin translation to be stimulated [147].

The regulatory role of 5'UTRs in translation may also involve formation of secondary structures. L-ferritin mRNA contains a stem-loop structure IRE within 5'UTR. Disruptions of the IRE inhibits its interaction with iron-regulatory proteins, resulting in abnormal L-ferritin production. Changes in L-ferritin levels have been reported in the hereditary hyperferritinaemia/cataract syndrome (HHCS) [148,149]. In colorectal cancer, the cellular inhibitor of PP2A—CIP2A protein regulates translation of the *MYC* gene mRNA and by increasing the level of MYC protein leads to carcinogenesis. A suggested explanation for this phenomenon is the presence of the IRES elements and G-quadruplexes structures in the 5'UTR sequence of *MYC* mRNA [150]. The G-quadruplex structure is another regulatory element that is frequently observed in 5'UTR sequences. They have been revealed by various laboratory techniques that include the following: antibody arrays and RGB-1 coupled with small molecules [151], SHALiPE structural analysis [152], crystallization using iSpinach aptamer fluorescent complex with the 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI) [153], and luminescent-based techniques [154]. The structure of G-quadruplex was identified in the 5'UTR sequence of the hepatocyte nuclear factor 4-alpha

(*HNF4 α*) gene. The inhibition of the HNF4 α translation through interaction between 5'UTR G-quadruplex and RNA binding proteins can lead to carcinogenesis in the human liver [155,156]. Another mechanism based on the 5'UTR secondary structure was observed in human skeletal muscles. The DHX36 helicase is responsible for unwinding the G-quadruplex structures. It binds to the 5'UTR secondary structure of the *GNAI2* gene mRNA. When DHX36 rewinds the G-quadruplex structure, the Gnai2 protein is up produced in muscle resident stem cells. This induces proliferation and thus increases the ability of these cells to regenerate [157].

7. 5'UTRs and Their Other Implications

As was described above, transcripts with alternative 5'-untranslated regions, by affecting mRNA stability and translational efficiency, can determine gene expression [158,159]. These mRNA variants may also differ in the length of 5' untranslated regions (UTRs) [160,161], display distinct expression patterns [162], and have various and even opposite biological functions [98,163]. By expression in a tissue-specific manner, these transcripts may regulate protein expression and control developmental and physiological processes [159].

The deregulation of translation, via the 5'UTR region, is associated with many diseases, including cancers [164–166]. In fact, long and complex 5'UTRs are more frequently expressed by oncogenes and tumor suppressor genes [167]. *BRCA1*, *Mdm2*, and *AXIN2* are just a few examples of cancer related genes that utilize alternative 5'UTRs [159,168,169].

The tumor suppressor gene *BRCA1* is involved in breast and ovarian cancer development, mainly through decreased levels of *BRCA1* mRNA. Two *BRCA1* transcripts that use alternative promoters show distinct expression patterns [159,168]. Transcript with longer 5'UTR is expressed only in breast cancer tissue, whereas *BRCA1* mRNA with shorter 5'UTR has been found in cancerous and noncancerous breast tissues [168]. Elongated at the 5' end mRNA of the *BRCA1* gene has a more complex secondary structure and upstream AUG codons, which have been shown to reduce translational efficiency compared to a shorter isoform [161,170]. Deregulated expression of the *BRCA1* increases the amount of 5' extended transcript and causes a decrease in the BRCA1 protein noticed in breast and ovarian cancers [168].

The *BCRP* gene that encodes breast cancer resistance protein also expresses mRNAs with different 5'UTRs. *BCRP* is frequently over-expressed in human cancers and leukemia, causing multidrug resistance in chemotherapy [164,171]. It has been found that *BCRP* gene utilizes tissue-specific promoters that produce at least three 5'UTR variants. The expression level of the *BCRP* mRNA isoform with the longest 5'UTR has been increased in the drug-resistant cells. In human breast cancer, cell lines selected with Adriamycin and verapamil (MCF-7/AdrVp) transcripts with extended 5'UTR constituted 47%; in ovarian carcinoma, cells selected with topotecan (Igrov1/T8) 71%; and with mitoxantrone (Igrov1/MX3) cell, 66% of the total *BCRP* mRNA transcripts, implying that this isoform is predominant in drug-resistant cells. Furthermore, mRNA containing longer 5'UTR is more efficiently translated compared to isoforms with shorter 5'UTR, thus can regulate BCRP protein levels [164]. The high expression level of *BCRP* is also associated with a poor prognosis for acute myeloid leukemia (AML) patients [172]. Results of analysis of the *BCRP* transcripts isoforms from pediatric AML samples revealed that a novel 5'UTR located 90 kb upstream of the exon 2 translation initiation site was expressed in 40% of the bone marrow samples and in all acute megakaryoblastic leukemia (AML FAB-M7) cases. Interestingly, expression from this upstream promoter wasn't detected in non-hematopoietic cell lines. These findings suggest that *BCRP* utilizes a tissue-specific promoter and high expression of transcripts isoforms generated from this promoter may have a negative impact on M7 AML subtype [173].

The expression level of the ER β protein is strongly influenced by the expression of alternative 5'UTRs. ER β is estrogen receptor β , which regulates cell growth and differentiation. Studies have demonstrated that two alternative 5'UTRs (longer UTRa and shorter UTRb) of *Er β* have tissue-specific distribution. Additionally, evidence implies that 5'UTRs may determine downstream splicing events, potentially influencing ER β function [167,174].

An activating transcription factor 3 (ATF3) determines cell fate due to the regulation of stress response and its expression was increased in the human prostate [175]. The *ATF3* alternative promoter described as P1 was identified ~43.5 kb upstream of the P2 promoter. Multiple transcriptional start sites and various 5'UTRs of the P1 transcripts were detected. Transcripts with the longest 5'UTR contain inhibitory elements, whereas the stimulatory elements are present in the shorter 5'UTRs. These features are especially important in response to serum or oxidoreductive stress. Furthermore, results suggest that the upstream P1 promoter is constitutively activated in human cancer cells [121].

8. Conclusions

The 5'UTR sequences via different mechanisms play an important role in the regulation of both transcription and translation. Some of these mechanisms enable the proper functioning of the cell, while others lead to pathological changes, which are summarized in Table 1. Many of these mechanisms have been well understood and explained, while others require further investigation and clarification. With the constant development of laboratory techniques, such as ribosome profiling, crystallization, and genome/transcriptome editing tools, we can expect a more thorough elucidation of all the mechanisms involving 5'UTRs in near future.

Table 1. List of genes that various regulatory elements of the 5'UTR are involved in diseases.

Regulatory Elements	Gene	Disease	References
De-regulation of uORFs	<i>BACE1</i>	Alzheimer's disease	[21,22]
	<i>ATF4</i>	Head and neck squamous carcinomas	[19]
	<i>Her-2</i>	Breast cancer	[39]
	<i>SOX2</i>	Squamous cell carcinoma (SCC)	[28]
	<i>CDKN2A</i>	Melanoma	[52]
	<i>SRY</i>	Gonadal dysgenesis	[47,48]
	<i>IRF6</i>	Van der Woude and Popliteal Pterygium Syndromes	[49]
	<i>GCH1</i>	Levodopa-responsive dystonia	[50]
	<i>HAMP</i>	Hereditary hemochromatosis	[51]
	<i>N2C</i>	Neuronal intranuclear inclusion disease (NIID)	[54]
	<i>CDKN1B</i>	Pituitary adenoma, Pancreatic cancer	[55]
	<i>CEBPB</i>	Acute myeloid, Breast cancer	[56]
	<i>EPHB1</i>	Breast cancer, Colon cancer	[46]
	<i>MAP2K6</i>	Colon adenocarcinoma	[46]
Length of 5'UTR	<i>TGF-β3</i>	Arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC)	[53]
	<i>DLG1</i>	High risk of human papillomavirus (HPV)	[5,30]
	<i>BRCA1</i>	Hereditary breast-ovarian cancer syndrome	[159,168]
	<i>BCRP</i>	Breast cancer Acute myeloid leukemia (AML)	[164] [173]
	<i>ERβ</i>	Breast and lung cancer	[167,174]
IRES	<i>GRN</i>	Progressive brain diseases	[27]
	<i>ATF3</i>	Prostate cancer	[121]
	<i>Connexin-32</i>	X-linked Charcot–Marie–Tooth disease (CMTX)	[144]
miRNA and protein binding sites	<i>c-Myc</i>	Multiple myeloma patients	[145,146]
	<i>RUNX3</i>	Colorectal cancer	[132]
	<i>STAT3</i>	Several types of tumors	[136]
Secondary and stem-loop structures	<i>ELAVL4</i>	Neurological disorders	[137]
	L-ferritin	Hereditary hyperferritinemia/cataract syndrome (HHCS)	[148,149]
	<i>MYC</i>	Colorectal cancer	[150]
	<i>HNF4α</i>	Liver cancer	[155,156]

Author Contributions: Conceptualization, N.R.; writing—original draft preparation, N.R. and A.L.; writing—review and editing, N.R., A.L. and I.M.; supervising, I.M.; funding acquisition, I.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Center of Poland, grant number 2017/25/B/NZ2/01519.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Schuster, S.L.; Hsieh, A.C. The Untranslated Regions of MRNAs in Cancer. *Trends Cancer* **2019**, *5*, 245–262. [[CrossRef](#)]
- Mularoni, L.; Sabarinathan, R.; Deu-Pons, J.; Gonzalez-Perez, A.; López-Bigas, N. OncodriveFML: A General Framework to Identify Coding and Non-Coding Regions with Cancer Driver Mutations. *Genome Biol.* **2016**, *17*, 128. [[CrossRef](#)] [[PubMed](#)]
- Cottrell, K.A.; Chaudhari, H.G.; Cohen, B.A.; Djuranovic, S. PTRE-Seq Reveals Mechanism and Interactions of RNA Binding Proteins and MiRNAs. *Nat. Commun.* **2018**, *9*, 301. [[CrossRef](#)] [[PubMed](#)]
- Kutchko, K.M.; Sanders, W.; Ziehr, B.; Phillips, G.; Solem, A.; Halvorsen, M.; Weeks, K.M.; Moorman, N.; Laederach, A. Multiple Conformations Are a Conserved and Regulatory Feature of the RB1 5' UTR. *RNA* **2015**, *21*, 1274–1285. [[CrossRef](#)] [[PubMed](#)]
- Marziali, F.; Dizanzo, M.P.; Cavatorta, A.L.; Gardiol, D. Differential Expression of DLG1 as a Common Trait in Different Human Diseases: An Encouraging Issue in Molecular Pathology. *Biol. Chem.* **2019**, *400*, 699–710. [[CrossRef](#)] [[PubMed](#)]
- Lammich, S.; Buell, D.; Zilow, S.; Ludwig, A.-K.; Nuscher, B.; Lichtenthaler, S.F.; Prinzen, C.; Fahrenholz, F.; Haass, C. Expression of the Anti-Amyloidogenic Secretase ADAM10 Is Suppressed by Its 5'-Untranslated Region. *J. Biol. Chem.* **2010**, *285*, 15753–15760. [[CrossRef](#)]
- Makalowska, I.; Lin, C.-F.; Makalowski, W. Overlapping Genes in Vertebrate Genomes. *Comput. Biol. Chem.* **2005**, *29*, 1–12. [[CrossRef](#)]
- Sanna, C.R.; Li, W.-H.; Zhang, L. Overlapping Genes in the Human and Mouse Genomes. *BMC Genom.* **2008**, *9*, 169. [[CrossRef](#)]
- Renz, P.F.; Valdivia-Francia, F.; Sendoel, A. Some like It Translated: Small ORFs in the 5'UTR. *Exp. Cell Res.* **2020**, *396*, 112229. [[CrossRef](#)]
- McGillivray, P.; Ault, R.; Pawashe, M.; Kitchen, R.; Balasubramanian, S.; Gerstein, M. A Comprehensive Catalog of Predicted Functional Upstream Open Reading Frames in Humans. *Nucleic Acids Res.* **2018**, *46*, 3326–3338. [[CrossRef](#)]
- Kearse, M.G.; Wilusz, J.E. Non-AUG Translation: A New Start for Protein Synthesis in Eukaryotes. *Genes Dev.* **2017**, *31*, 1717–1731. [[CrossRef](#)] [[PubMed](#)]
- Andreev, D.E.; Loughran, G.; Fedorova, A.D.; Mikhaylova, M.S.; Shatsky, I.N.; Baranov, P.V. Non-AUG Translation Initiation in Mammals. *Genome Biol.* **2022**, *23*, 111. [[CrossRef](#)] [[PubMed](#)]
- Carrara, M.; Sigurdardottir, A.; Bertolotti, A. Decoding the Selectivity of EIF2 α Holophosphatases and PPP1R15A Inhibitors. *Nat. Struct. Mol. Biol.* **2017**, *24*, 708–716. [[CrossRef](#)] [[PubMed](#)]
- Young, S.K.; Willy, J.A.; Wu, C.; Sachs, M.S.; Wek, R.C. Ribosome Reinitiation Directs Gene-Specific Translation and Regulates the Integrated Stress Response. *J. Biol. Chem.* **2015**, *290*, 28257–28271. [[CrossRef](#)]
- Vattem, K.M.; Wek, R.C. Reinitiation Involving Upstream ORFs Regulates ATF4 mRNA Translation in Mammalian Cells. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11269–11274. [[CrossRef](#)]
- Baird, T.D.; Palam, L.R.; Fusakio, M.E.; Willy, J.A.; Davis, C.M.; McClintick, J.N.; Anthony, T.G.; Wek, R.C. Selective mRNA Translation during EIF2 Phosphorylation Induces Expression of IBTK α . *MBoC* **2014**, *25*, 1686–1697. [[CrossRef](#)]
- Szamecz, B.; Rutkai, E.; Cuchalová, L.; Munzarová, V.; Herrmannová, A.; Nielsen, K.H.; Burela, L.; Hinnebusch, A.G.; Valášek, L. EIF3a Cooperates with Sequences 5' of UORF1 to Promote Resumption of Scanning by Post-Termination Ribosomes for Reinitiation on GCN4 mRNA. *Genes Dev.* **2008**, *22*, 2414–2425. [[CrossRef](#)]
- Young, S.K.; Wek, R.C. Upstream Open Reading Frames Differentially Regulate Gene-Specific Translation in the Integrated Stress Response. *J. Biol. Chem.* **2016**, *291*, 16927–16935. [[CrossRef](#)]
- Chen, H.-H.; Yu, H.-I.; Yang, M.-H.; Tarn, W.-Y. DDX3 Activates CBC-EIF3-Mediated Translation of UORF-Containing Oncogenic MRNAs to Promote Metastasis in HNSCC. *Cancer Res.* **2018**, *78*, 4512–4523. [[CrossRef](#)]
- Juntawong, P.; Girke, T.; Bazin, J.; Bailey-Serres, J. Translational Dynamics Revealed by Genome-Wide Profiling of Ribosome Footprints in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E203–E212. [[CrossRef](#)]
- Zhou, W.; Song, W. Leaky Scanning and Reinitiation Regulate BACE1 Gene Expression. *Mol. Cell. Biol.* **2006**, *26*, 3353–3364. [[CrossRef](#)] [[PubMed](#)]
- Mihailovich, M.; Thermann, R.; Grohovaz, F.; Hentze, M.W.; Zacchetti, D. Complex Translational Regulation of BACE1 Involves Upstream AUGs and Stimulatory Elements within the 5' Untranslated Region. *Nucleic Acids Res.* **2007**, *35*, 2975–2985. [[CrossRef](#)] [[PubMed](#)]

23. Libre, C.; Seissler, T.; Guerrero, S.; Batische, J.; Verriez, C.; Stupfler, B.; Gilmer, O.; Weber, M.M.; Cimarelli, A.; Etienne, L.; et al. A Conserved UORF Impacts APOBEC3G Translation and Is Essential for Translational Inhibition by the HIV-1 Vif Protein. *Biomedicines* **2022**, *10*, 13. [[CrossRef](#)] [[PubMed](#)]
24. Takahashi, H.; Miyaki, S.; Onouchi, H.; Motomura, T.; Idesako, N.; Takahashi, A.; Murase, M.; Fukuyoshi, S.; Endo, T.; Satou, K.; et al. Exhaustive Identification of Conserved Upstream Open Reading Frames with Potential Translational Regulatory Functions from Animal Genomes. *Sci. Rep.* **2020**, *10*, 16289. [[CrossRef](#)]
25. Kozak, M. New Ways of Initiating Translation in Eukaryotes? *Mol. Cell. Biol.* **2001**, *21*, 1899–1907. [[CrossRef](#)]
26. Rajkowitsch, L.; Vilela, C.; Berthelot, K.; Ramirez, C.V.; McCarthy, J.E.G. Reinitiation and Recycling Are Distinct Processes Occurring Downstream of Translation Termination in Yeast. *J. Mol. Biol.* **2004**, *335*, 71–85. [[CrossRef](#)]
27. Capell, A.; Fellerer, K.; Haass, C. Progranulin Transcripts with Short and Long 5' Untranslated Regions (UTRs) Are Differentially Expressed via Posttranscriptional and Translational Repression. *J. Biol. Chem.* **2014**, *289*, 25879–25889. [[CrossRef](#)]
28. Sendoel, A.; Dunn, J.G.; Rodriguez, E.H.; Naik, S.; Gomez, N.C.; Hurwitz, B.; Levorse, J.; Dill, B.D.; Schramek, D.; Molina, H.; et al. Translation from Unconventional 5' Start Sites Drives Tumour Initiation. *Nature* **2017**, *541*, 494–499. [[CrossRef](#)]
29. Riedl, L.; Mackenzie, I.R.; Förstl, H.; Kurz, A.; Diehl-Schmid, J. Frontotemporal Lobar Degeneration: Current Perspectives. *Neuropsychiatr. Dis. Treat.* **2014**, *10*, 297–310. [[CrossRef](#)]
30. Cavatorta, A.L.; Facciuto, F.; Valdano, M.B.; Marziali, F.; Giri, A.A.; Banks, L.; Gardiol, D. Regulation of Translational Efficiency by Different Splice Variants of the Disc Large 1 Oncosuppressor 5'-UTR. *FEBS J.* **2011**, *278*, 2596–2608. [[CrossRef](#)]
31. Marziali, F.; Cavatorta, A.L.; Valdano, M.B.; Facciuto, F.; Gardiol, D. Transcriptional and Translational Mechanisms Contribute to Regulate the Expression of Discs Large 1 Protein during Different Biological Processes. *Biol. Chem.* **2015**, *396*, 893–902. [[CrossRef](#)] [[PubMed](#)]
32. Zhang, H.; Wang, Y.; Lu, J. Function and Evolution of Upstream ORFs in Eukaryotes. *Trends Biochem. Sci.* **2019**, *44*, 782–794. [[CrossRef](#)] [[PubMed](#)]
33. Thakur, A.; Hinnebusch, A.G. EIF1 Loop 2 Interactions with Met-TRNAⁱ Control the Accuracy of Start Codon Selection by the Scanning Preinitiation Complex. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E4159–E4168. [[CrossRef](#)] [[PubMed](#)]
34. Jackson, R.J.; Hellen, C.U.T.; Pestova, T.V. The Mechanism of Eukaryotic Translation Initiation and Principles of Its Regulation. *Nat. Rev. Mol. Cell. Biol.* **2010**, *11*, 113–127. [[CrossRef](#)]
35. Thakur, A.; Gaikwad, S.; Vijamarri, A.K.; Hinnebusch, A.G. EIF2 α Interactions with MRNA Control Accurate Start Codon Selection by the Translation Preinitiation Complex. *Nucleic Acids Res.* **2020**, *48*, 10280–10296. [[CrossRef](#)]
36. Ramakrishnan, V. Ribosome Structure and the Mechanism of Translation. *Cell* **2002**, *108*, 557–572. [[CrossRef](#)]
37. Chia, M.; Tresenrider, A.; Chen, J.; Spedale, G.; Jorgensen, V.; Ünal, E.; van Werven, F.J. Transcription of a 5' Extended MRNA Isoform Directs Dynamic Chromatin Changes and Interference of a Downstream Promoter. *eLife* **2022**, *6*, e27420. [[CrossRef](#)]
38. Meyer, R.E.; Chuong, H.H.; Hild, M.; Hansen, C.L.; Kinter, M.; Dawson, D.S. Ipl1/Aurora-B Is Necessary for Kinetochores Restructuring in Meiosis I in *Saccharomyces Cerevisiae*. *MBoC* **2015**, *26*, 2986–3000. [[CrossRef](#)]
39. Mehta, A.; Trotta, C.R.; Peltz, S.W. Derepression of the Her-2 UORF Is Mediated by a Novel Post-Transcriptional Control Mechanism in Cancer Cells. *Genes Dev.* **2006**, *20*, 939–953. [[CrossRef](#)]
40. Ito, K.; Chiba, S. Arrest Peptides: Cis-Acting Modulators of Translation. *Annu. Rev. Biochem.* **2013**, *82*, 171–202. [[CrossRef](#)]
41. Ramu, H.; Vázquez-Laslop, N.; Klepacki, D.; Dai, Q.; Piccirilli, J.; Micura, R.; Mankin, A.S. Nascent Peptide in the Ribosome Exit Tunnel Affects Functional Properties of the A-Site of the Peptidyl Transferase Center. *Mol. Cell* **2011**, *41*, 321–330. [[CrossRef](#)] [[PubMed](#)]
42. Ramani, A.K.; Nelson, A.C.; Kapranov, P.; Bell, I.; Gingeras, T.R.; Fraser, A.G. High Resolution Transcriptome Maps for Wild-Type and Nonsense-Mediated Decay-Defective *Caenorhabditis Elegans*. *Genome Biol.* **2009**, *10*, R101. [[CrossRef](#)] [[PubMed](#)]
43. Yi, Z.; Sanjeev, M.; Singh, G. The Branched Nature of the Nonsense-Mediated MRNA Decay Pathway. *Trends Genet.* **2021**, *37*, 143–159. [[CrossRef](#)] [[PubMed](#)]
44. Jaquet, V.; Wallerich, S.; Voegeli, S.; Túrós, D.; Vilorio, E.C.; Becskei, A. Determinants of the Temperature Adaptation of MRNA Degradation. *Nucleic Acids Res.* **2022**, *50*, 1092–1110. [[CrossRef](#)]
45. Ye, Y.; Liang, Y.; Yu, Q.; Hu, L.; Li, H.; Zhang, Z.; Xu, X. Analysis of Human Upstream Open Reading Frames and Impact on Gene Expression. *Hum. Genet.* **2015**, *134*, 605–612. [[CrossRef](#)]
46. Schulz, J.; Mah, N.; Neuenschwander, M.; Kischka, T.; Ratei, R.; Schlag, P.M.; Castaños-Vélez, E.; Fichtner, I.; Tunn, P.-U.; Denkert, C.; et al. Loss-of-Function UORF Mutations in Human Malignancies. *Sci. Rep.* **2018**, *8*, 2395. [[CrossRef](#)]
47. Calvo, S.E.; Pagliarini, D.J.; Mootha, V.K. Upstream Open Reading Frames Cause Widespread Reduction of Protein Expression and Are Polymorphic among Humans. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7507–7512. [[CrossRef](#)]
48. Poulat, F.; Desclozeaux, M.; Tuffery, S.; Jay, P.; Boizet, B.; Berta, P. Mutation in the 5' Noncoding Region of the SRY Gene in an XY Sex-Reversed Patient. *Hum. Mutat.* **1998**, *11*, S192–S194. [[CrossRef](#)]
49. Kondo, S.; Schutte, B.C.; Richardson, R.J.; Bjork, B.C.; Knight, A.S.; Watanabe, Y.; Howard, E.; de Lima, R.L.L.F.; Daack-Hirsch, S.; Sander, A.; et al. Mutations in IRF6 Cause Van Der Woude and Popliteal Pterygium Syndromes. *Nat. Genet.* **2002**, *32*, 285–289. [[CrossRef](#)]
50. Tassin, J.; Dürr, A.; Bonnet, A.M.; Gil, R.; Vidailhet, M.; Lücking, C.B.; Goas, J.Y.; Durif, F.; Abada, M.; Echenne, B.; et al. Levodopa-Responsive Dystonia. GTP Cyclohydrolase I or Parkin Mutations? *Brain* **2000**, *123 Pt 6*, 1112–1121. [[CrossRef](#)]

51. Matthes, T.; Aguilar-Martinez, P.; Pizzi-Bosman, L.; Darbellay, R.; Rubbia-Brandt, L.; Giostra, E.; Michel, M.; Ganz, T.; Beris, P. Severe Hemochromatosis in a Portuguese Family Associated with a New Mutation in the 5'-UTR of the HAMP Gene. *Blood* **2004**, *104*, 2181–2183. [[CrossRef](#)]
52. Liu, L.; Dilworth, D.; Gao, L.; Monzon, J.; Summers, A.; Lassam, N.; Hogg, D. Mutation of the CDKN2A 5' UTR Creates an Aberrant Initiation Codon and Predisposes to Melanoma. *Nat. Genet.* **1999**, *21*, 128–132. [[CrossRef](#)] [[PubMed](#)]
53. Beffagna, G.; Occhi, G.; Nava, A.; Vitiello, L.; Ditadi, A.; Basso, C.; Bauce, B.; Carraro, G.; Thiene, G.; Towbin, J.A.; et al. Regulatory Mutations in Transforming Growth Factor-Beta3 Gene Cause Arrhythmic Right Ventricular Cardiomyopathy Type 1. *Cardiovasc. Res.* **2005**, *65*, 366–373. [[CrossRef](#)] [[PubMed](#)]
54. Boivin, M.; Deng, J.; Pfister, V.; Grandgirard, E.; Oulad-Abdelghani, M.; Morlet, B.; Ruffenach, F.; Negroni, L.; Koebel, P.; Jacob, H.; et al. Translation of GGC Repeat Expansions into a Toxic Polyglycine Protein in NIID Defines a Novel Class of Human Genetic Disorders: The PolyG Diseases. *Neuron* **2021**, *109*, 1825–1835.e5. [[CrossRef](#)] [[PubMed](#)]
55. Occhi, G.; Regazzo, D.; Trivellin, G.; Boaretto, F.; Ciato, D.; Bobisse, S.; Ferasin, S.; Cetani, F.; Pardi, E.; Korbonits, M.; et al. A Novel Mutation in the Upstream Open Reading Frame of the CDKN1B Gene Causes a MEN4 Phenotype. *PLoS Genet.* **2013**, *9*, e1003350. [[CrossRef](#)]
56. Wethmar, K.; Bégay, V.; Smink, J.J.; Zaragoza, K.; Wiesenthal, V.; Dörken, B.; Calkhoven, C.F.; Leutz, A. C/EBPbetaDeltaORF Mice—A Genetic Model for UORF-Mediated Translational Control in Mammals. *Genes Dev.* **2010**, *24*, 15–20. [[CrossRef](#)] [[PubMed](#)]
57. Huang, N.; Li, F.; Zhang, M.; Zhou, H.; Chen, Z.; Ma, X.; Yang, L.; Wu, X.; Zhong, J.; Xiao, F.; et al. An Upstream Open Reading Frame in Phosphatase and Tensin Homolog Encodes a Circuit Breaker of Lactate Metabolism. *Cell Metab.* **2021**, *33*, 128–144. [[CrossRef](#)] [[PubMed](#)]
58. Jayaram, D.R.; Frost, S.; Argov, C.; Liju, V.B.; Anto, N.P.; Muraleedharan, A.; Ben-Ari, A.; Sinay, R.; Smoly, I.; Novoplansky, O.; et al. Unraveling the Hidden Role of a UORF-Encoded Peptide as a Kinase Inhibitor of PKCs. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2018899118. [[CrossRef](#)]
59. Dever, T.E.; Ivanov, I.P.; Sachs, M.S. Conserved Upstream Open Reading Frame Nascent Peptides That Control Translation. *Annu. Rev. Genet.* **2020**, *54*, 237–264. [[CrossRef](#)]
60. Bovre, K.; Szybalski, W. Patterns of Convergent and Overlapping Transcription within the B2 Region of Coliphage Lambda. *Virology* **1969**, *38*, 614–626. [[CrossRef](#)]
61. Johnson, Z.I.; Chisholm, S.W. Properties of Overlapping Genes Are Conserved across Microbial Genomes. *Genome Res.* **2004**, *14*, 2268–2272. [[CrossRef](#)] [[PubMed](#)]
62. Osato, N.; Yamada, H.; Satoh, K.; Ooka, H.; Yamamoto, M.; Suzuki, K.; Kawai, J.; Carninci, P.; Ohtomo, Y.; Murakami, K.; et al. Antisense Transcripts with Rice Full-Length cDNAs. *Genome Biol.* **2003**, *5*, R5. [[CrossRef](#)] [[PubMed](#)]
63. David, L.; Huber, W.; Granovskaia, M.; Toedling, J.; Palm, C.J.; Bofkin, L.; Jones, T.; Davis, R.W.; Steinmetz, L.M. A High-Resolution Map of Transcription in the Yeast Genome. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5320–5325. [[CrossRef](#)] [[PubMed](#)]
64. Misener, S.R.; Walker, V.K. Extraordinarily High Density of Unrelated Genes Showing Overlapping and Intraintronic Transcription Units. Sequence Data from This Article Have Been Deposited with the GenBank Data Library under Accession Nos. AF098020 and AF170829. *Biochim. Biophys. Acta (BBA) Gene Struct. Expr.* **2000**, *1492*, 269–270. [[CrossRef](#)]
65. Bâtshake, B.; Sundelin, J. The Mouse Genes for the EP1 Prostanoid Receptor and the PKN Protein Kinase Overlap. *Biochem. Biophys. Res. Commun.* **1996**, *227*, 70–76. [[CrossRef](#)] [[PubMed](#)]
66. Kennerson, M.L.; Nassif, N.T.; Dawkins, J.L.; Dekroon, R.M.; Yang, J.G.; Nicholson, G.A. The Charcot-Marie-Tooth Binary Repeat Contains a Gene Transcribed from the Opposite Strand of a Partially Duplicated Region of The COX10 Gene. *Genomics* **1997**, *46*, 61–69. [[CrossRef](#)]
67. Veeramachaneni, V.; Makalowski, W.; Galdzicki, M.; Sood, R.; Makalowska, I. Mammalian Overlapping Genes: The Comparative Perspective. *Genome Res.* **2004**, *14*, 280–286. [[CrossRef](#)]
68. Chen, C.-H.; Pan, C.-Y.; Lin, W. Overlapping Protein-Coding Genes in Human Genome and Their Coincidental Expression in Tissues. *Sci. Rep.* **2019**, *9*, 13377. [[CrossRef](#)]
69. Ho, M.-R.; Tsai, K.-W.; Lin, W. A Unified Framework of Overlapping Genes: Towards the Origination and Endogenous Regulation. *Genomics* **2012**, *100*, 231–239. [[CrossRef](#)]
70. Rosikiewicz, W.; Sikora, J.; Skrzypczak, T.; Kubiak, M.R.; Makałowska, I. Promoter Switching in Response to Changing Environment and Elevated Expression of Protein-Coding Genes Overlapping at Their 5' Ends. *Sci. Rep.* **2021**, *11*, 8984. [[CrossRef](#)]
71. Zhou, C.; Blumberg, B. Overlapping Gene Structure of Human VLCAD and DLG4. *Gene* **2003**, *305*, 161–166. [[CrossRef](#)] [[PubMed](#)]
72. Lu, T.; Zhu, C.; Lu, G.; Guo, Y.; Zhou, Y.; Zhang, Z.; Zhao, Y.; Li, W.; Lu, Y.; Tang, W.; et al. Strand-Specific RNA-Seq Reveals Widespread Occurrence of Novel Cis-Natural Antisense Transcripts in Rice. *BMC Genom.* **2012**, *13*, 721. [[CrossRef](#)] [[PubMed](#)]
73. Ling, M.H.; Ban, Y.; Wen, H.; Wang, S.M.; Ge, S.X. Conserved Expression of Natural Antisense Transcripts in Mammals. *BMC Genom.* **2013**, *14*, 243. [[CrossRef](#)] [[PubMed](#)]
74. Engström, P.G.; Suzuki, H.; Ninomiya, N.; Akalin, A.; Sessa, L.; Lavorgna, G.; Brozzi, A.; Luzi, L.; Tan, S.L.; Yang, L.; et al. Complex Loci in Human and Mouse Genomes. *PLoS Genet.* **2006**, *2*, e47. [[CrossRef](#)] [[PubMed](#)]
75. Soldà, G.; Suyama, M.; Pelucchi, P.; Boi, S.; Guffanti, A.; Rizzi, E.; Bork, P.; Tenchini, M.L.; Ciccarelli, F.D. Non-Random Retention of Protein-Coding Overlapping Genes in Metazoa. *BMC Genom.* **2008**, *9*, 174. [[CrossRef](#)]

76. Trinklein, N.D.; Aldred, S.F.; Hartman, S.J.; Schroeder, D.I.; Otilar, R.P.; Myers, R.M. An Abundance of Bidirectional Promoters in the Human Genome. *Genome Res.* **2004**, *14*, 62–66. [[CrossRef](#)]
77. Conley, A.B.; King Jordan, I. Epigenetic Regulation of Human Cis-Natural Antisense Transcripts. *Nucleic Acids Res.* **2012**, *40*, 1438–1445. [[CrossRef](#)]
78. Rosikiewicz, W.; Suzuki, Y.; Makołowska, I. OverGeneDB: A Database of 5' End Protein Coding Overlapping Genes in Human and Mouse Genomes. *Nucleic Acids Res.* **2018**, *46*, D186–D193. [[CrossRef](#)]
79. Shearwin, K.E.; Callen, B.P.; Egan, J.B. Transcriptional Interference—A Crash Course. *Trends Genet.* **2005**, *21*, 339–345. [[CrossRef](#)]
80. Faghihi, M.A.; Wahlestedt, C. Regulatory Roles of Natural Antisense Transcripts. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 637–643. [[CrossRef](#)]
81. Beltran, M.; Puig, I.; Peña, C.; García, J.M.; Álvarez, A.B.; Peña, R.; Bonilla, F.; de Herreros, A.G. A Natural Antisense Transcript Regulates Zeb2/Sip1 Gene Expression during Snail1-Induced Epithelial–Mesenchymal Transition. *Genes Dev.* **2008**, *22*, 756–769. [[CrossRef](#)] [[PubMed](#)]
82. Faghihi, M.A.; Modarresi, F.; Khalil, A.M.; Wood, D.E.; Sahagan, B.G.; Morgan, T.E.; Finch, C.E.; St. Laurent, G., III; Kenny, P.J.; Wahlestedt, C. Expression of a Noncoding RNA Is Elevated in Alzheimer's Disease and Drives Rapid Feed-Forward Regulation of β -Secretase. *Nat. Med.* **2008**, *14*, 723–730. [[CrossRef](#)] [[PubMed](#)]
83. Pouladi, N.; Kouhsari, S.M.; Feizi, M.H.; Gavani, R.R.; Azarfam, P. Overlapping Region of P53/WRAP53 Transcripts: Mutational Analysis and Sequence Similarity with MicroRNA-4732-5p. *Asian Pac. J. Cancer Prev.* **2013**, *14*, 3503–3507. [[CrossRef](#)] [[PubMed](#)]
84. Ebralidze, A.K.; Guibal, F.C.; Steidl, U.; Zhang, P.; Lee, S.; Bartholdy, B.; Jorda, M.A.; Petkova, V.; Rosenbauer, F.; Huang, G.; et al. PU.1 Expression Is Modulated by the Balance of Functional Sense and Antisense RNAs Regulated by a Shared Cis-Regulatory Element. *Genes Dev.* **2008**, *22*, 2085–2092. [[CrossRef](#)] [[PubMed](#)]
85. Kudla, G.; Murray, A.W.; Tollervey, D.; Plotkin, J.B. Coding-Sequence Determinants of Gene Expression in *Escherichia Coli*. *Science* **2009**, *324*, 255–258. [[CrossRef](#)]
86. Werner, A.; Cockell, S.; Falconer, J.; Carlile, M.; Alnumeir, S.; Robinson, J. Contribution of Natural Antisense Transcription to an Endogenous siRNA Signature in Human Cells. *BMC Genom.* **2014**, *15*, 19. [[CrossRef](#)]
87. Sinturel, F.; Bréchemier-Baey, D.; Kiledjian, M.; Condon, C.; Bénard, L. Activation of 5'-3' exoribonuclease Xrn1 by cofactor Dcs1 is essential for mitochondrial function in yeast. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 8264–8269. [[CrossRef](#)]
88. Mahmoudi, S.; Henriksson, S.; Corcoran, M.; Méndez-Vidal, C.; Wiman, K.G.; Farnebo, M. WRAP53, a Natural P53 Antisense Transcript Required for P53 Induction upon DNA Damage. *Mol. Cell* **2009**, *33*, 462–471. [[CrossRef](#)]
89. Vilborg, A.; Wilhelm, M.T.; Wiman, K.G. Regulation of Tumor Suppressor P53 at the RNA Level. *J. Mol. Med.* **2010**, *88*, 645–652. [[CrossRef](#)]
90. Farnebo, M.; Bykov, V.J.N.; Wiman, K.G. The P53 Tumor Suppressor: A Master Regulator of Diverse Cellular Processes and Therapeutic Target in Cancer. *Biochem. Biophys. Res. Commun.* **2010**, *396*, 85–89. [[CrossRef](#)]
91. Tycowski, K.T.; Shu, M.-D.; Kukoyi, A.; Steitz, J.A. A Conserved WD40 Protein Binds the Cajal Body Localization Signal of ScaRNP Particles. *Mol. Cell* **2009**, *34*, 47. [[CrossRef](#)]
92. Farnebo, M. WRAP53, a Novel Regulator of P53. *Cell Cycle* **2009**, *8*, 2343–2346. [[CrossRef](#)] [[PubMed](#)]
93. Venteicher, A.S.; Abreu, E.B.; Meng, Z.; McCann, K.E.; Terns, R.M.; Veenstra, T.D.; Terns, M.P.; Artandi, S.E. A Human Telomerase Holoenzyme Protein Required for Cajal Body Localization and Telomere Synthesis. *Science* **2009**, *323*, 644–648. [[CrossRef](#)] [[PubMed](#)]
94. Mahmoudi, S.; Henriksson, S.; Farnebo, L.; Roberg, K.; Farnebo, M. WRAP53 Promotes Cancer Cell Survival and Is a Potential Target for Cancer Therapy. *Cell Death Dis.* **2011**, *2*, e114. [[CrossRef](#)]
95. Sun, C.; Luo, X.; Gou, Y.; Hu, L.; Wang, K.; Li, C.; Xiang, Z.; Zhang, P.; Kong, X.; Zhang, C.; et al. TCAB1: A Potential Target for Diagnosis and Therapy of Head and Neck Carcinomas. *Mol. Cancer* **2014**, *13*, 180. [[CrossRef](#)]
96. Rao, X.; Huang, D.; Sui, X.; Liu, G.; Song, X.; Xie, J.; Huang, D. Overexpression of WRAP53 Is Associated with Development and Progression of Esophageal Squamous Cell Carcinoma. *PLoS ONE* **2014**, *9*, e91670. [[CrossRef](#)] [[PubMed](#)]
97. Sun, Y.; Yang, C.; Chen, J.; Song, X.; Li, Z.; Duan, M.; Li, J.; Hu, X.; Wu, K.; Yan, G.; et al. Overexpression of WDR79 in Non-small Cell Lung Cancer Is Linked to Tumour Progression. *J. Cell. Mol. Med.* **2016**, *20*, 698–709. [[CrossRef](#)] [[PubMed](#)]
98. Zhu, Y.; Sun, W.; Jiang, X.; Bai, R.; Luo, Y.; Gao, Y.; Li, S.; Huang, Z.; Gong, Y.; Xie, C. *Differential Effects of WRAP53 Transcript Variants on the Biological Behaviours of Human Non-Small Cell Lung Cancer Cells*; Research Square: Durham, NC, USA, 2021.
99. Pouladi, N.; Abdolahi, S.; Farajzadeh, D.; Hosseinpour Feizi, M.A. Haplotype and Linkage Disequilibrium of TP53-WRAP53 Locus in Iranian-Azeri Women with Breast Cancer. *PLoS ONE* **2019**, *14*, e0220727. [[CrossRef](#)]
100. Salato, V.K.; Rediske, N.W.; Zhang, C.; Hastings, M.L.; Munroe, S. An Exonic Splicing Enhancer within a Bidirectional Coding Sequence Regulates Alternative Splicing of an Antisense mRNA. *RNA Biol.* **2010**, *7*, 179–190. [[CrossRef](#)]
101. Su, W.-Y.; Li, J.-T.; Cui, Y.; Hong, J.; Du, W.; Wang, Y.-C.; Lin, Y.-W.; Xiong, H.; Wang, J.-L.; Kong, X.; et al. Bidirectional Regulation between WDR83 and Its Natural Antisense Transcript DHPS in Gastric Cancer. *Cell Res.* **2012**, *22*, 1374–1389. [[CrossRef](#)]
102. Sinturel, F.; Navickas, A.; Wery, M.; Describes, M.; Morillon, A.; Torchet, C.; Benard, L. Cytoplasmic Control of Sense-Antisense mRNA Pairs. *Cell Rep.* **2015**, *12*, 1853–1864. [[CrossRef](#)] [[PubMed](#)]
103. Borsani, O.; Zhu, J.; Verslues, P.E.; Sunkar, R.; Zhu, J.-K. Endogenous siRNAs Derived from a Pair of Natural Cis-Antisense Transcripts Regulate Salt Tolerance in Arabidopsis. *Cell* **2005**, *123*, 1279–1291. [[CrossRef](#)] [[PubMed](#)]

104. Sharma, E.; Sterne-Weiler, T.; O'Hanlon, D.; Blencowe, B.J. Global Mapping of Human RNA-RNA Interactions. *Mol. Cell* **2016**, *62*, 618–626. [[CrossRef](#)] [[PubMed](#)]
105. Lu, Z.; Gong, J.; Zhang, Q.C. PARIS: Psoralen Analysis of RNA Interactions and Structures with High Throughput and Resolution. *Methods Mol. Biol.* **2018**, *1649*, 59–84. [[CrossRef](#)] [[PubMed](#)]
106. Zhang, M.; Li, K.; Bai, J.; Velema, W.A.; Yu, C.; van Damme, R.; Lee, W.H.; Corpuz, M.L.; Chen, J.-F.; Lu, Z. Optimized Photochemistry Enables Efficient Analysis of Dynamic RNA Structuromes and Interactomes in Genetic and Infectious Diseases. *Nat. Commun.* **2021**, *12*, 2344. [[CrossRef](#)] [[PubMed](#)]
107. Lu, Z.; Zhang, Q.C.; Lee, B.; Flynn, R.A.; Smith, M.A.; Robinson, J.T.; Davidovich, C.; Gooding, A.R.; Goodrich, K.J.; Mattick, J.S.; et al. RNA Duplex Map in Living Cells Reveals Higher Order Transcriptome Structure. *Cell* **2016**, *165*, 1267–1279. [[CrossRef](#)]
108. Aw, J.G.A.; Shen, Y.; Wilm, A.; Sun, M.; Lim, X.N.; Boon, K.-L.; Tapsin, S.; Chan, Y.-S.; Tan, C.-P.; Sim, A.Y.L.; et al. In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation. *Mol. Cell* **2016**, *62*, 603–617. [[CrossRef](#)]
109. Palmer, A.C.; Egan, J.B.; Shearwin, K.E. Transcriptional Interference by RNA Polymerase Pausing and Dislodgement of Transcription Factors. *Transcription* **2011**, *2*, 9–14. [[CrossRef](#)]
110. Crampton, N.; Bonass, W.A.; Kirkham, J.; Rivetti, C.; Thomson, N.H. Collision Events between RNA Polymerases in Convergent Transcription Studied by Atomic Force Microscopy. *Nucleic Acids Res.* **2006**, *34*, 5416–5425. [[CrossRef](#)]
111. Prescott, E.M.; Proudfoot, N.J. Transcriptional Collision between Convergent Genes in Budding Yeast. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8796–8801. [[CrossRef](#)]
112. Petruk, S.; Sedkov, Y.; Riley, K.M.; Hodgson, J.; Schweisguth, F.; Hirose, S.; Jaynes, J.B.; Brock, H.W.; Mazo, A. Transcription of Bxd Noncoding RNAs Promoted by Trithorax Represses Ubx in Cis by Transcriptional Interference. *Cell* **2006**, *127*, 1209–1221. [[CrossRef](#)] [[PubMed](#)]
113. Conte, C.; Dastugue, B.; Vaury, C. Promoter Competition as a Mechanism of Transcriptional Interference Mediated by Retrotransposons. *EMBO J.* **2002**, *21*, 3908–3916. [[CrossRef](#)] [[PubMed](#)]
114. Osato, N.; Suzuki, Y.; Ikeo, K.; Gojobori, T. Transcriptional Interferences in Cis Natural Antisense Transcripts of Humans and Mice. *Genetics* **2007**, *176*, 1299–1306. [[CrossRef](#)] [[PubMed](#)]
115. Wight, M.; Werner, A. The Functions of Natural Antisense Transcripts. *Essays Biochem.* **2013**, *54*, 91–101. [[CrossRef](#)]
116. Palmer, A.C.; Ahlgren-Berg, A.; Egan, J.B.; Dodd, I.B.; Shearwin, K.E. Potent Transcriptional Interference by Pausing of RNA Polymerases over a Downstream Promoter. *Mol. Cell* **2009**, *34*, 545–555. [[CrossRef](#)]
117. Shuman, S. Transcriptional Interference at Tandem lncRNA and Protein-Coding Genes: An Emerging Theme in Regulation of Cellular Nutrient Homeostasis. *Nucleic Acids Res.* **2020**, *48*, 8243–8254. [[CrossRef](#)]
118. Brar, G.A.; Yassour, M.; Friedman, N.; Regev, A.; Ingolia, N.T.; Weissman, J.S. High-Resolution View of the Yeast Meiotic Program Revealed by Ribosome Profiling. *Science* **2012**, *335*, 552–557. [[CrossRef](#)]
119. Liu, Y.; Stuparevic, I.; Xie, B.; Becker, E.; Law, M.J.; Primig, M. The Conserved Histone Deacetylase Rpd3 and the DNA Binding Regulator Ume6 Repress BOI1's Meiotic Transcript Isoform during Vegetative Growth in *Saccharomyces Cerevisiae*. *Mol. Microbiol.* **2015**, *96*, 861–874. [[CrossRef](#)]
120. Tien, J.F.; Fong, K.K.; Umbreit, N.T.; Payen, C.; Zelter, A.; Asbury, C.L.; Dunham, M.J.; Davis, T.N. Coupling Unbiased Mutagenesis to High-Throughput DNA Sequencing Uncovers Functional Domains in the Ndc80 Kinetochore Protein of *Saccharomyces Cerevisiae*. *Genetics* **2013**, *195*, 159–170. [[CrossRef](#)]
121. Miyazaki, K.; Inoue, S.; Yamada, K.; Watanabe, M.; Liu, Q.; Watanabe, T.; Adachi, M.T.; Tanaka, Y.; Kitajima, S. Differential Usage of Alternate Promoters of the Human Stress Response Gene ATF3 in Stress Response and Cancer Cells. *Nucleic Acids Res.* **2009**, *37*, 1438–1451. [[CrossRef](#)]
122. Saliminejad, K.; Khorshid, H.R.K.; Soleymani Fard, S.; Ghaffari, S.H. An Overview of MicroRNAs: Biology, Functions, Therapeutics, and Analysis Methods. *J. Cell. Physiol.* **2019**, *234*, 5451–5465. [[CrossRef](#)] [[PubMed](#)]
123. Hill, M.; Tran, N. MiRNA Interplay: Mechanisms and Consequences in Cancer. *Dis. Model Mech.* **2021**, *14*, dmm047662. [[CrossRef](#)] [[PubMed](#)]
124. Michlewski, G.; Cáceres, J.F. Post-Transcriptional Control of MiRNA Biogenesis. *RNA* **2019**, *25*, 1–16. [[CrossRef](#)]
125. Libri, V.; Helwak, A.; Miesen, P.; Santhakumar, D.; Borger, J.G.; Kudla, G.; Grey, F.; Tollervy, D.; Buck, A.H. Murine Cytomegalovirus Encodes a MiR-27 Inhibitor Disguised as a Target. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 279–284. [[CrossRef](#)] [[PubMed](#)]
126. Ha, M.; Kim, V.N. Regulation of MicroRNA Biogenesis. *Nat. Rev. Mol. Cell. Biol.* **2014**, *15*, 509–524. [[CrossRef](#)]
127. Sundaram, G.M.; Common, J.E.A.; Gopal, F.E.; Srikanta, S.; Lakshman, K.; Lunny, D.P.; Lim, T.C.; Tanavde, V.; Lane, E.B.; Sampath, P. "See-Saw" Expression of MicroRNA-198 and FSTL1 from a Single Transcript in Wound Healing. *Nature* **2013**, *495*, 103–106. [[CrossRef](#)] [[PubMed](#)]
128. Navarro, E.; Mallén, A.; Hueso, M. Dynamic Variations of 3'UTR Length Reprogram the mRNA Regulatory Landscape. *Biomedicines* **2021**, *9*, 1560. [[CrossRef](#)]
129. Da Sacco, L.; Masotti, A. Recent Insights and Novel Bioinformatics Tools to Understand the Role of MicroRNAs Binding to 5' Untranslated Region. *Int. J. Mol. Sci.* **2012**, *14*, 480–495. [[CrossRef](#)]

130. Gu, W.; Xu, Y.; Xie, X.; Wang, T.; Ko, J.-H.; Zhou, T. The Role of RNA Structure at 5' Untranslated Region in MicroRNA-Mediated Gene Regulation. *RNA* **2014**, *20*, 1369–1375. [[CrossRef](#)]
131. Ørom, U.A.; Nielsen, F.C.; Lund, A.H. MicroRNA-10a Binds the 5'UTR of Ribosomal Protein MRNAs and Enhances Their Translation. *Mol. Cell* **2008**, *30*, 460–471. [[CrossRef](#)]
132. Zhang, J.; Zhou, W.; Liu, Y.; Liu, T.; Li, C.; Wang, L. Oncogenic Role of MicroRNA-532-5p in Human Colorectal Cancer via Targeting of the 5'UTR of RUNX3. *Oncol. Lett.* **2018**, *15*, 7215–7220. [[CrossRef](#)] [[PubMed](#)]
133. Kunden, R.D.; Khan, J.Q.; Ghezelbash, S.; Wilson, J.A. The Role of the Liver-Specific MicroRNA, MiRNA-122 in the HCV Replication Cycle. *Int. J. Mol. Sci.* **2020**, *21*, 5677. [[CrossRef](#)] [[PubMed](#)]
134. Baldassarre, A.; Paolini, A.; Bruno, S.P.; Felli, C.; Tozzi, A.E.; Masotti, A. Potential Use of Noncoding RNAs and Innovative Therapeutic Strategies to Target the 5'UTR of SARS-CoV-2. *Epigenomics* **2020**, *12*, 1349–1361. [[CrossRef](#)]
135. Ying, H.; Ebrahimi, M.; Keivan, M.; Khoshnam, S.E.; Salahi, S.; Farzaneh, M. miRNAs; a Novel Strategy for the Treatment of COVID-19. *Cell Biol. Int.* **2021**, *45*, 2045–2053. [[CrossRef](#)]
136. Li, Z.; Wang, X.; Jia, R. Poly(RC) Binding Protein 1 Represses the Translation of STAT3 through 5' UTR. *Curr. Gene Ther.* **2022**, *22*, 397–405. [[CrossRef](#)]
137. Popovitchenko, T.; Park, Y.; Page, N.F.; Luo, X.; Krsnik, Z.; Liu, Y.; Salamon, I.; Stephenson, J.D.; Kraushar, M.L.; Volk, N.L.; et al. Translational Derepression of Elavl4 Isoforms at Their Alternative 5' UTRs Determines Neuronal Development. *Nat. Commun.* **2020**, *11*, 1674. [[CrossRef](#)] [[PubMed](#)]
138. Aeschmann, F.; Kumari, P.; Bartake, H.; Gaidatzis, D.; Xu, L.; Ciosk, R.; Großhans, H. LIN41 Post-Transcriptionally Silences MRNAs by Two Distinct and Position-Dependent Mechanisms. *Mol. Cell* **2017**, *65*, 476–489.e4. [[CrossRef](#)]
139. Kühn, L.C. Iron Regulatory Proteins and Their Role in Controlling Iron Metabolism. *Metallomics* **2015**, *7*, 232–243. [[CrossRef](#)]
140. Beckmann, K.; Grskovic, M.; Gebauer, F.; Hentze, M.W. A Dual Inhibitory Mechanism Restricts Msl-2 mRNA Translation for Dosage Compensation in Drosophila. *Cell* **2005**, *122*, 529–540. [[CrossRef](#)]
141. Merrick, W.C. Cap-Dependent and Cap-Independent Translation in Eukaryotic Systems. *Gene* **2004**, *332*, 1–11. [[CrossRef](#)]
142. Godet, A.-C.; David, F.; Hantelys, F.; Tatin, F.; Lacazette, E.; Garmy-Susini, B.; Prats, A.-C. IRES Trans-Acting Factors, Key Actors of the Stress Response. *Int. J. Mol. Sci.* **2019**, *20*, 924. [[CrossRef](#)]
143. Marques, R.; Lacerda, R.; Romão, L. Internal Ribosome Entry Site (IRES)-Mediated Translation and Its Potential for Novel mRNA-Based Therapy Development. *Biomedicines* **2022**, *10*, 1865. [[CrossRef](#)] [[PubMed](#)]
144. Hudder, A.; Werner, R. Analysis of a Charcot-Marie-Tooth Disease Mutation Reveals an Essential Internal Ribosome Entry Site Element in the Connexin-32 Gene. *J. Biol. Chem.* **2000**, *275*, 34586–34591. [[CrossRef](#)] [[PubMed](#)]
145. Chappell, S.A.; LeQuesne, J.P.; Paulin, F.E.; deSchoolmeester, M.L.; Stoneley, M.; Soutar, R.L.; Ralston, S.H.; Helfrich, M.H.; Willis, A.E. A Mutation in the C-Myc-IRES Leads to Enhanced Internal Ribosome Entry in Multiple Myeloma: A Novel Mechanism of Oncogene de-Regulation. *Oncogene* **2000**, *19*, 4437–4440. [[CrossRef](#)] [[PubMed](#)]
146. Evans, J.R.; Mitchell, S.A.; Spriggs, K.A.; Ostrowski, J.; Bomsztyk, K.; Ostarek, D.; Willis, A.E. Members of the Poly (RC) Binding Protein Family Stimulate the Activity of the c-Myc Internal Ribosome Entry Segment In Vitro and In Vivo. *Oncogene* **2003**, *22*, 8012–8020. [[CrossRef](#)] [[PubMed](#)]
147. Pandey, P.R.; Sarwade, R.D.; Khalique, A.; Seshadri, V. Interaction of HuDA and PABP at 5'UTR of Mouse Insulin2 Regulates Insulin Biosynthesis. *PLoS ONE* **2018**, *13*, e0194482. [[CrossRef](#)]
148. Cazzola, M. Role of Ferritin and Ferroportin Genes in Unexplained Hyperferritinaemia. *Best Pract. Res. Clin. Haematol.* **2005**, *18*, 251–263. [[CrossRef](#)]
149. Vanita, V.; Hejtmancik, J.F.; Hennies, H.C.; Guleria, K.; Nürnberg, P.; Singh, D.; Sperling, K.; Singh, J.R. Sutural Cataract Associated with a Mutation in the Ferritin Light Chain Gene (FTL) in a Family of Indian Origin. *Mol. Vis.* **2006**, *12*, 93–99.
150. Denk, S.; Schmidt, S.; Schurr, Y.; Schwarz, G.; Schote, F.; Diefenbacher, M.; Armendariz, C.; Dejure, F.; Eilers, M.; Wiegering, A. CIP2A Regulates MYC Translation (via Its 5'UTR) in Colorectal Cancer. *Int. J. Color. Dis.* **2021**, *36*, 911–918. [[CrossRef](#)]
151. Katsuda, Y.; Sato, S.; Inoue, M.; Tsugawa, H.; Kamura, T.; Kida, T.; Matsumoto, R.; Asamitsu, S.; Shioda, N.; Shiroto, S.; et al. Small Molecule-Based Detection of Non-Canonical RNA G-Quadruplex Structures That Modulate Protein Translation. *Nucleic Acids Res.* **2022**, *50*, 8143–8153. [[CrossRef](#)]
152. Kwok, C.K.; Sahakyan, A.B.; Balasubramanian, S. Structural Analysis Using SHALiPE to Reveal RNA G-Quadruplex Formation in Human Precursor MicroRNA. *Angew. Chem.* **2016**, *128*, 9104–9107. [[CrossRef](#)]
153. Fernandez-Millan, P.; Autour, A.; Ennifar, E.; Westhof, E.; Ryckelynck, M. Crystal Structure and Fluorescence Properties of the Spinach Aptamer in Complex with DFHBI. *RNA* **2017**, *23*, 1788–1795. [[CrossRef](#)] [[PubMed](#)]
154. Chen, F.; Li, G.; Liu, H.; Leung, C.-H.; Ma, D.-L. G-Quadruplex-Based Detection of Glyphosate in Complex Biological Systems by a Time-Resolved Luminescent Assay. *Sens. Actuators B Chem.* **2020**, *320*, 128393. [[CrossRef](#)]
155. Guo, S.; Lu, H. Conjunction of Potential G-Quadruplex and Adjacent Cis-Elements in the 5' UTR of Hepatocyte Nuclear Factor 4-Alpha Strongly Inhibit Protein Expression. *Sci. Rep.* **2017**, *7*, 17444. [[CrossRef](#)] [[PubMed](#)]
156. Bugaut, A.; Balasubramanian, S. 5'-UTR RNA G-Quadruplexes: Translation Regulation and Targeting. *Nucleic Acids Res.* **2012**, *40*, 4727–4741. [[CrossRef](#)]
157. Chen, X.; Yuan, J.; Xue, G.; Campanario, S.; Wang, D.; Wang, W.; Mou, X.; Liew, S.W.; Umar, M.I.; Isern, J.; et al. Translational Control by DHX36 Binding to 5'UTR G-Quadruplex Is Essential for Muscle Stem-Cell Regenerative Functions. *Nat. Commun.* **2021**, *12*, 5043. [[CrossRef](#)]

158. Landry, J.-R.; Mager, D.L.; Wilhelm, B.T. Complex Controls: The Role of Alternative Promoters in Mammalian Genomes. *Trends Genet.* **2003**, *19*, 640–648. [[CrossRef](#)]
159. Hughes, T.A. Regulation of Gene Expression by Alternative Untranslated Regions. *Trends Genet.* **2006**, *22*, 119–122. [[CrossRef](#)]
160. Thorsen, K.; Schepeler, T.; Øster, B.; Rasmussen, M.H.; Vang, S.; Wang, K.; Hansen, K.Q.; Lamy, P.; Pedersen, J.S.; Eller, A.; et al. Tumor-Specific Usage of Alternative Transcription Start Sites in Colorectal Cancer Identified by Genome-Wide Exon Array Analysis. *BMC Genom.* **2011**, *12*, 505. [[CrossRef](#)]
161. Araujo, P.R.; Yoon, K.; Ko, D.; Smith, A.D.; Qiao, M.; Suresh, U.; Burns, S.C.; Penalva, L.O.F. Before It Gets Started: Regulating Translation at the 5' UTR. *Comp. Funct. Genom.* **2012**, *2012*, 475731. [[CrossRef](#)]
162. Chen, J.; Tresenrider, A.; Chia, M.; McSwiggen, D.T.; Spedale, G.; Jorgensen, V.; Liao, H.; van Werven, F.J.; Ünal, E. Kinetochores Inactivation by Expression of a Repressive mRNA. *eLife* **2017**, *6*, e27417. [[CrossRef](#)] [[PubMed](#)]
163. Wang, E.T.; Sandberg, R.; Luo, S.; Khrebtkova, I.; Zhang, L.; Mayr, C.; Kingsmore, S.F.; Schroth, G.P.; Burge, C.B. Alternative Isoform Regulation in Human Tissue Transcriptomes. *Nature* **2008**, *456*, 470–476. [[CrossRef](#)]
164. Nakanishi, T.; Bailey-Dell, K.J.; Hassel, B.A.; Shiozawa, K.; Sullivan, D.M.; Turner, J.; Ross, D.D. Novel 5' Untranslated Region Variants of BCRP mRNA Are Differentially Expressed in Drug-Selected Cancer Cells and in Normal Human Tissues: Implications for Drug Resistance, Tissue-Specific Expression, and Alternative Promoter Usage. *Cancer Res.* **2006**, *66*, 5007–5011. [[CrossRef](#)] [[PubMed](#)]
165. Marcel, V.; Hainaut, P. P53 Isoforms—A Conspiracy to Kidnap P53 Tumor Suppressor Activity? *Cell. Mol. Life Sci.* **2008**, *66*, 391. [[CrossRef](#)] [[PubMed](#)]
166. Davuluri, R.V.; Suzuki, Y.; Sugano, S.; Plass, C.; Huang, T.H.-M. The Functional Consequences of Alternative Promoter Use in Mammalian Genomes. *Trends Genet.* **2008**, *24*, 167–177. [[CrossRef](#)]
167. Smith, L. Post-Transcriptional Regulation of Gene Expression by Alternative 5'-Untranslated Regions in Carcinogenesis. *Biochem. Soc. Trans.* **2008**, *36*, 708–711. [[CrossRef](#)]
168. Sobczak, K.; Krzyzosiak, W.J. Structural Determinants of BRCA1 Translational Regulation. *J. Biol. Chem.* **2002**, *277*, 17349–17358. [[CrossRef](#)]
169. Hughes, T.A.; Brady, H.J.M. Regulation of Axin2 Expression at the Levels of Transcription, Translation and Protein Stability in Lung and Colon Cancer. *Cancer Lett.* **2006**, *233*, 338–347. [[CrossRef](#)]
170. Wang, J.; Lu, C.; Min, D.; Wang, Z.; Ma, X. A Mutation in the 5' Untranslated Region of the BRCA1 Gene in Sporadic Breast Cancer Causes Downregulation of Translation Efficiency. *J. Int. Med. Res.* **2007**, *35*, 564–573. [[CrossRef](#)]
171. Chen, L.; Manautou, J.E.; Rasmussen, T.P.; Zhong, X. Development of Precision Medicine Approaches Based on Inter-Individual Variability of BCRP/ABCG2. *Acta Pharm. Sin. B* **2019**, *9*, 659–674. [[CrossRef](#)]
172. Damiani, D.; Tiribelli, M.; Michelutti, A.; Geromin, A.; Cavallin, M.; Fabbro, D.; Pianta, A.; Malagola, M.; Damante, G.; Russo, D.; et al. Fludarabine-Based Induction Therapy Does Not Overcome the Negative Effect of ABCG2 (BCRP) over-Expression in Adult Acute Myeloid Leukemia Patients. *Leuk. Res.* **2010**, *34*, 942–945. [[CrossRef](#)] [[PubMed](#)]
173. Campbell, P.K.; Zong, Y.; Yang, S.; Zhou, S.; Rubnitz, J.E.; Sorrentino, B.P. Identification of a Novel, Tissue-Specific ABCG2 Promoter Expressed in Pediatric Acute Megakaryoblastic Leukemia. *Leuk. Res.* **2011**, *35*, 1321–1329. [[CrossRef](#)]
174. Smith, L.; Brannan, R.A.; Hanby, A.M.; Shaaban, A.M.; Verghese, E.T.; Peter, M.B.; Pollock, S.; Satheesha, S.; Szykiewicz, M.; Speirs, V.; et al. Differential Regulation of Oestrogen Receptor β Isoforms by 5' Untranslated Regions in Cancer. *J. Cell. Mol. Med.* **2010**, *14*, 2172–2184. [[CrossRef](#)] [[PubMed](#)]
175. Pelzer, A.E.; Bektic, J.; Haag, P.; Berger, A.P.; Pycha, A.; Schäfer, G.; Rogatsch, H.; Horninger, W.; Bartsch, G.; Klocker, H. The Expression of Transcription Factor Activating Transcription Factor 3 in the Human Prostate and Its Regulation by Androgen in Prostate Cancer. *J. Urol.* **2006**, *175*, 1517–1522. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.