Identification and characterization of non-coding RNAs and their associated proteins involved in cellular stress responses

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat)

submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by
Karol Jerzy Rogowski

Berlin, 2019

The experiments of this work were conducted from March 2014 to February 2018 in the group of Prof. Dr. Markus Landthaler at the MDC Berlin Buch.

1st reviewer: Prof. Dr. Markus Landthaler

Max Delbruck Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany

2nd reviewer: Prof. Dr. Markus Wahl

Freie Universität Berlin, Germany

Date of Defense: 21.11.2019

Acknowledgements

First of all, I would like to thank Prof. Dr. Markus Landthaler for the opportunity to work in his group and for the supervision during the length of my PhD project. I am thankful for all the advice, guidance and patience he had for me.

I would also like to thank Prof. Dr. Gunter Meister (University of Regensburg) for his efforts as a co-supervisor of my work within The European non-coding RNA training network and for the kind gift of purified, recombinant Csy4 protein, which was used in this work.

Very warm thanks to all the members of the Landthaler laboratory for a great working atmosphere and all the help I received during our work together. Special thank you goes to Emanuel Wyler, Nico Kastelic and Miha Milek for the help with experimental procedures and bioinformatical analysis. To Emanuel Wyler and Miha Milek also for good advice and constructive criticism of my work. Additional thank you to Emmanuel Wyler for generation of PAR-CLIP data used in this work, all his amazing support and for help with translation of the summary of this thesis into German. Special warm thanks go to Ulrike Zinnall for all the support given and her invaluable help in all the situations where communication in German was required.

I also want to thank Prof. Dr. Markus Wahl for agreeing to be a reviewer of this work and Freie Universität Berlin for accepting me as PhD student.

A special thank you goes to Dr. Szymon Swiezewski (Polish Academy of Sciences) for inspiration, fruitful discussions and support during the length of my PhD project.

Finally, I would like to thank all the collaborators who performed computational analysis of high-throughput sequencing data presented in this thesis: Lorenzo Calviello (Floor lab, UCSF), Petar Glazar (Rajewsky lab, MDC) and Vedran Franke (Akalin lab, MDC).

Summary

Genotoxic stress is an important threat to the wellbeing of cells in the human body. Damages to genetic material contribute to the aging process and can lead to mutations that cause cancer or other diseases. The DNA damage response has therefore been studied extensively. With recent advancements in the field, it became clear that the study of lncRNAs and RBPs is necessary to broaden our understanding of DDR. This thesis aimed at discovery and characterization of lncRNAs and their associated proteins involved in DDR to provide novel insight into the regulation of this process.

Firstly, we have attempted to develop a new methodology that would allow to study RNA-protein interactions with high efficiency and ease of use. We have designed two new experimental approaches to overcome limitations of previously established methods. Their usefulness, however, turned out to be limited and we were unable to provide a reliable improvement to existing methodologies.

Secondly, we provided an extensive atlas of the transcriptional landscape of MCF-7 cells during the DDR. We employed three different high-throughput RNA sequencing approaches to provide a comprehensive analysis of the transcriptome. Our efforts led to identification of differentially expressed transcripts regardless of their polyadenylation or stability, or towards their presence in existing genome annotations. We have provided a foundation for future functional studies of lncRNAs potentially involved in DDR and investigated four lncRNAs upregulated in this condition in more detail.

Thirdly, we have provided a new important insight into regulatory mechanism of DDX3 – a protein, which has been previously reported to associate with lncRNAs and is implicated in both DDR and tumorigenesis. We have described the nature of DDX3-lncRNAs interactions on a transcriptome-wide scale and furthered our understanding of DDX3 mediated regulation of translation. We have determined the global effects of DDX3 depletion on the abundance of mRNAs and their translation efficiency. In combination with the analysis of DDX3-mRNA binding specificity those results show that the protein is required for translation initiation on subset of mRNAs harboring structured 5’ UTRs.

Zusammenfassung

Genotoxischer Stress ist eine ständige Bedrohung der Integrität menschlicher Zellen. Schäden im genetischen Material tragen zum Alterungsprozess bei und zur Entstehung von Krebs und anderen Krankheiten. Die DNA-Schadensantwort (DSA) ist daher ein intensiv bearbeitetes Forschungsfeld. Neue Erkenntnisse haben gezeigt, dass auch RNA-bindende Proteine (RBPs) und nicht-codierende RNAs (ncRNAs) vermehrt erforscht werden müssen für das Verständnis der DSA. Diese Doktorarbeit hatte zum Ziel, ncRNAs und die damit assoziierten RBPs zu identifizieren und charakterisieren, um die DSA besser zu verstehen.

Im ersten Teil haben wir versucht, eine neue Methode zu etablieren, um RBP-ncRNA-Interaktionen effizient und einfach zu identifizieren. Mit zwei neuen Prozeduren versuchten wir, die bisherigen Methoden zu verbessern. Ihre Nützlichkeit war aber beschränkt, und wir konnten keine nachhaltige Verbesserung der existierenden Protokolle erreichen.

Im zweiten Teil wurde ein ausführliches Verzeichnis der Veränderung im Transkriptom der Zellen während der DSA erstellt. Drei verschiedene Hochdurchsatz-Sequenziermethoden wurden verwendet, um Transkripte unabhängig von deren Polyadenylierungs-Status und Stabilität zu detektieren, und unabhängig davon ob sie schon annotiert sind. Wir haben damit das Fundament für zukünftige Forschungsprojekte gelegt, und vier ncRNAs auch näher untersucht.

Im dritten Teil haben wir die Rolle des DDX3-Proteins in der Regulation der zellulären Genexpression untersucht. Bekannt ist, dass DDX3 mit ncRNAs interagiert, und sowohl in der DSA wie auch bei der Tumorentstehung relevant ist. Wir haben die DDX3-ncRNA-Interaktionen Transkriptom-weit beschrieben, und die Rolle von DDX3 in der Regulation der Translation charakterisiert. Zusammengenommen zeigen unsere Ergebnisse, dass DDX3 notwendig ist, um bei einem Teil der Boten-RNAs die Translation effizient zu initiieren.

Table of content

1. Introduction

1.1. Non-coding RNAs in human biology

The central dogma of molecular biology put the role of RNA as a carrier of genetic information from DNA to production of proteins. The last decades of study, however aided by the development of high throughput sequencing techniques showed that RNA molecules play much more diverse roles in cells biology than originally appreciated. Alongside messenger RNAs (mRNAs), which provide a template for protein translation, and transcripts centered around production of proteins such as ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) other species of transcripts that are not directly involved in a process of translation have been discovered and functionally characterized. Discovery of their functionality directly challenged the notion that the sole function of RNA is to produce proteins and led to revision of the central dogma of molecular biology. Those transcripts are collectively named non-coding RNAs (ncRNAs) and are arbitrarily divided into two groups based on their size. Small ncRNAs are shorter than 200nt in length and long non-coding RNAs (lncRNAs) are longer than 200 nt. In mammalian cells there are three main subgroups of small ncRNAs: micro RNAs (miRNAs), short interfering RNAs (siRNAs) and piwi-interacting RNAs (piRNAs). Those classes of transcripts differ between each other in regard of origin, processing and their mode of action. They however share a role in regulating expression of other cellular transcripts. The roles of lncRNAs on the other hand are more diverse and they have been shown to regulate a variety of cellular processes. Overall the discoveries in the field of ncRNA made in last decades redefined our way of thinking about RNA and their cellular functions, and thus broadened our understanding of molecular biology.

1.1.1. Properties of small non-coding RNAs

The best studied group of small ncRNAs are miRNAs. They are ~19-24nt in length, single stranded transcripts that regulate gene expression on post transcriptional level. MiRNAs are produced from long RNAs called primary miRNAs (pri-miRNAs), which are transcribed by RNA polymerase II (Pol II). Those transcripts contain one or more about 80 base pair long stem loops, which are released to give rise to precursor miRNAs (pre-miRNAs) (Han et al., 2006). Generation of pre-miRNAs is dependent on the Microprocessor complex, which recognizes specific motifs in pri-miRNA and cleaves miRNA hairpins from the

primary transcript. The microprocessor complex is made of two proteins: an RNase III enzyme Drosha and the RNA binding protein DiGeorge Critical Region 8 (DGCR8), which are present in the complex as a hetero-dimer (Nguyen et al., 2015; Denli et al., 2004; Lee et al., 2003; Gregory et al., 2004). The resulting pre-miRNA is next transported to the cytoplasm via the Exportin 5 pathway (Yi et al., 2003; Lund et al., 2004) and is further cleaved by another RNAse III enzyme called Dicer 1 (DCR1) (Hutvagner et al., 2001; Taylor et al., 2013; Ma et al., 2012). DCR1 activity leads to the release of 21-23nt long double stranded miRNAs (Zhang et al., 2002; Ma et al., 2004). One of the strands of the duplex is bound by an Argonaute (Ago) protein forming the RNA induced silencing complex (RISC) (Liu et al., 2004; Meister et al., 2004; Rivas et al., 2005). The RISC can target transcripts through the sequence complementarity of the loaded miRNA and lead to gene silencing by target mRNA degradation or translational inhibition (Rand et al., 2005; Diederichs and Haber 2007). Usually, target recognition by miRNAs is mediated by base pairing between the seed sequence of the miRNA (nucleotides 2-8 at miRNA 5'end) and sequences in the 3' untranslated region (UTR) of the target RNA. The choice in miRNA mediated gene silencing by inhibition of translation or by transcript degradation appears to depend on degree and nature of sequence complementarity between the target RNA and the miRNA (Pasquinelli, 2012; Fabian and Sonenberg, 2012; Czech and Hannon, 2010). Throughout the years, miRNAs have been shown to be important regulators of many pathways in mammalian cells including tissue differentiation (Chen et al., 2006; Krichevsky et al., 2006). Their misregulation have also been associated with diseases, including viral infections (Triboulet et al., 2007) and cancer (Michael et al., 2003; Takamizawa et al., 2004; He et al., 2005; Hayashita et al., 2005). The relevance of miRNAs as regulators of cells biology is reinforced by the fact that more than half of the human transcriptome has been predicted to be regulated by miRNAs, which means that they potentially regulate most major human gene pathways (Krol et al., 2010; Winter et al., 2009).

1.1.2. Properties of long non-coding RNAs

Another group of RNA regulators that are being increasingly appreciated as regulators of cellular physiology are lncRNAs. They are generally classified as RNA molecules longer than 200 nt that have no or little predicted coding potential. Those molecules are pervasively transcribed from eukaryotic genomes and rise from both intergenic regions and protein coding gene bodies (Okazaki, et al., 2002; Affymetrix ENCODE Transcriptiome Project, 2009; Katayama et al., 2005; Laurent et al., 2012). The wealth of cellular lncRNAs is vast and

consists of transcripts of diverse characteristics including spliced and unspliced, nuclear and cytoplasmic, polyadenylated and non-polyadenylated molecules. In addition, there is a class of transcripts, which are subjected to a process of back splicing which joins their 3' and 5' ends resulting in a circular transcript (circRNA), which are generally considered to be lncRNAs (Salzman et al., 2012).

Although many lncRNAs are processed similarly to mRNAs, i.e. being subject to capping, splicing and polyadenylation, all of them were initially considered to be of little physiological relevance. When lncRNAs were discovered, RNA was believed to be solely an intermediate between information coded in DNA and protein synthesis. For this reason, long RNA molecules not involved in production of proteins were dubbed “transcriptional noise”. It was hypothesized that RNA polymerase II randomly initiates transcription throughout the genome, giving rise to both functional mRNA molecules and non-functional lncRNAs. Several features however pointed to functionality of certain lncRNAs, including conservation of promoters (Carninci et al., 2005; Derrien et al., 2012), regulation by transcription factors (Cawley et al., 2004; Guttman et al., 2009) and chromatin signatures typical for active gene expression (Guttman et al., 2009; Derrien et al., 2012). It was also shown that the number of RNA molecules not coding for proteins in a genome correlates well with organism complexity, in contrast to the number of protein coding sequences (Taft and Mattick, 2003; Liu et al., 2013). This observation pointed to the possibility that lncRNAs may constitute a novel layer of regulation of cell physiology. The notion that lncRNAs represent nothing more than transcriptional noise was directly challenged by the discovery of functional lncRNAs, the earliest notable example of which is Xist (Brown et al., 1992).

1.1.3. Functions of long non-coding RNAs

Xist (X inactive specific transcript) is a long non-coding RNA that is essential in X chromosome inactivation and formation of the Barr body in placental mammals. Xist coats the entire inactivated X chromosome and works to recruit Polycomb repressive complexes 1 and 2 (PRC1 and PRC2) (Zhao et al., 2008). This leads to deposition of repressive epigenetic modifications like histone H3 hypo-acetylation and H3K27 methylation resulting in transcriptional inactivation of almost the entire chromosome (Yang et al., 2010). Other lncRNAs have been shown to regulate diverse cellular processes including chromatin modification (Pandey et al., 2008; Khalil et al., 2009; Tsai et al., 2010), formation of specialized subnuclear organelles (Sunwoo et al, 2009; Bond et al., 2009), fine tuning the

miRNA pathways by sponging miRNA molecules (Du et al., 2016; Wu et al., 2017; Shan et al., 2018), regulation of splicing (Gonzalez et al. 2015; Romero-Barrios et al, 2018) and enhancer activity (Kim et al., 2010; Ørom et al., 2010). Moreover, many studies showed links between lncRNAs and differentiation and development (Amaral and Mattick 2008; St Laurent et al., 2016) as well as different diseases, including schizophrenia (Barry et al., 2014), cancer (Mourtada-Maarabouniet al., 2009; Gutschner et al., 2013; Liu et al., 2014), cardiovascular diseases and diabetes (Broadbent et al., 2008; Burd et al., 2010; Pasmant et al., 2011).

Despite indirect indications for lncRNAs functionality as a class of transcripts and increasing number of studies ascribing functions to diverse members of this group of molecules many questions and doubts remain. It is still a matter of speculation to what extent lncRNAs are transcriptional noise of the cell and how many lncRNA genes give rise to functional transcripts, or their transcription itself has a regulatory role. Determining functions of plethora of lncRNAs is an important challenge of modern biology. However, due to properties of non-coding transcripts their study is often difficult. Many lncRNAs have relatively low expression levels when compared to mRNAs. They are often transcribed transiently during development and/or, in specific cell types, resulting in low levels of transcripts detected in whole tissue samples (Derrien et al., 2012; Deveson et al., 2017). Low levels of lncRNA expression are sometimes explained by the notion that it is the process of their transcription rather than utilization of resulting transcript that exerts the function of lncRNAs. This notion is at least in part supported by recent studies showing that many long intergenic non-codind RNAs (lincRNAs) are inefficiently co-transcriptionally spliced and terminated in comparison to mRNAs, thus being rapidly degraded by the RNA exosome (Schlackow et al., 2017). This observation may reflect a mechanism in which indeed the functions of some lincRNAs are carried out by the process of transcription itself and resulting transcripts are non-functional and thus quickly degraded. An alternative explanation is that observed transcription of lincRNAs is non-functional and they require a specific cellular context to be stabilized to exert their function. Moreover, conservation of lncRNA sequences is significantly lower than conservation of protein coding genes (Mouse Genome Sequencing Consortium et al., 2002). They also contain relatively few defined RNA motifs within their sequences. It is possible, that many non-coding transcripts exert their functions through formation of secondary structures, thereby requiring less strict sequence conservation then their protein coding counterparts. Another issue that may complicate the evaluation of the biological relevance of ncRNAs is the fact that they may carry out highly specialized

functions. It is possible for genes to escape phenotypic screens such as knock-downs of some of them affect phenotypes not commonly examined in such experiments or have effects that are difficult to observe. This phenomenon is potentially problematic for functional study of all classes of genes, but it has been speculated to be more common in case of lncRNAs than of protein coding transcripts (Mattick 2018).

In recent years, the need to reevaluate the lack of protein coding potential of many lncRNAs became apparent. The initial classification of RNA molecules as non-coding was based on their lack of open reading frames containing at least 100 codons, associated with lack of ability to code for typical functional protein products (The FANTOM consortium et al., 2002; Deveson et al., 2017; Kang et al., 2017). With the advent of ribosome profiling technology, which evaluates direct association of the translation machinery with translated transcripts, it became apparent, however, that many short open reading frames can be translated into small peptides. Those peptides have been showed to rise from regions of mRNAs annotated as untranslated as well as from lncRNAs (Ji et al., 2015; Bazin et al., 2017). It has been shown that 40% transcripts previously considered as non-coding associate with actively translating ribosomes. This observation suggests that many transcripts classified previously as non-coding may in fact also function as mRNAs. The lncRNA-ribosome associations observed in ribosome profiling experiments, however, have to be considered with care. Although the experimental setup allows to distinguish between the RNAs that are associating with ribosomes in a non-consequential manner and those that are actually engaged in translation, it does not provide information on stability or functionality of produced peptides. It has been reported that translated lncRNAs are prevalently cytoplasmic and the majority of their peptide products are considered to be unstable by-products of ribosome activity (Ji et al., 2015). The prevalence of lncRNA-ribosome interactions not resulting in production of stable, functional peptides in human cells is supported by low numbers of peptides originating from lncRNAs being identified with mass spectrometry approaches (Sun et al., 2014; Crappé et al., 2015). The examples of lncRNA encoded small peptides with biologically relevant functions are, however, numerous (Kondo et al., 2007; Anderson et al., 2015; Nelson et al., 2016; Matsumoto et al., 2017; Huang et al., 2017; D'Lima et al., 2017). It is therefore important to consider regulatory roles of lncRNAs as potential peptide coding messages in their functional study, especially when their localization is cytoplasmic.

1.1.4. Challenges in study of long non-coding RNAs

Due to their specific properties, studying lncRNAs requires not only careful consideration, but also use of novel experimental approaches. Investigating transcript-protein interactions is of particular interest, as lncRNAs are known to exert their functions through interactions with proteins. Investigation of RNA sequences bound to a specific protein of interest has been achieved with success thanks to methods like RNA immunoprecipitation (RIP) and different variants of cross-linking and immunoprecipitation (CLIP) (Zhao et al., 2010; Ule et al., 2003; Hafner et al., 2010). However, de-novo identification of proteins interacting with a transcript of interest is more difficult. This is particularly challenging in case of lncRNAs, due to their general low abundance and stability. Typical RNA pull-down experiments involve incubation of immobilized in-vitro transcribed molecules of interest with cellular lysates to identify proteins that are capable of binding their sequence (Lee et al., 2013; Treiber et al., 2017). However, the usefulness of those approaches in the study of RNAs proved to be limited. The protein-RNA interaction that is being captured takes place in-vitro and may not reflect binding maintained in living cells. Perhaps for this reason classic RNA pull-down experiments are associated with high background levels (Treiber et al., 2017). To better understand the regulatory nature of cellular transcripts, unbiased and highly sensitive methods allowing the determination of the protein interactome of a specific transcript are necessary. A number of approaches that isolate specific RNAs expressed in cells that bear short sequence tags embedded in the transcript were developed to meet this requirement. In those systems, in vivo interactions are being captured and usage of cross-linking is possible. Variants of those methods utilize S1 aptamere, MS2 tagging, StreptoTag and others (Bachler et al., 1999; Srisawat and Engelke 2001; Yoon et al., 2012; Leppek and Stoecklin 2013). Those approaches turned out to be functional but were put to little use in the lncRNA field, probably due to their relatively low stringency and unsatisfying efficiency. Another, more recently developed family of methods, employed affinity purification of transcripts of interest from cellular lysates with biotinylated oligonucleotides. Those methods proved much more useful in identifying RNA-interacting proteins than previously developed systems. Three similar approaches were created independently: capture hybridization analysis of RNA targets (CHART), RNA Antisense Purification (RAP) and Chromatin Isolation by RNA Purification (ChIRP) (Simon et al., 2011; Engreitz et al., 2013; Chu et al., 2011). Originally, those methods were developed to study lncRNA interactions with chromatin, but later were combined with mass spectrometry to identify RNA-bound proteins. They utilized biotinylated

RNA or DNA probes of various lengths antisense to the target transcript for purification. All three provided a remarkable improvement in terms of specificity and efficiency compared to preexisting methods. None of them require usage of genetic engineering of cells to be used and can target endogenous transcripts. A drawback shared by them however is their cost, as for each target transcript a unique array of biotinylated oligonucleotides needs to be purchased. Moreover, a limiting step in identifying proteins bound by the target RNA is current mass spectrometry technology. To be identified in this approach, the amounts of proteins purified have to exceed a certain threshold, which in case of targeting lowly expressed RNAs is difficult to achieve. However, the mass spectrometry technologies are being constantly perfected, and their sensitivity is being increased (Li et al., 2017; Iwamoto and Shimada 2018). It seems likely that in the following years improvements in those technologies will allow for reliable identification of much lower amounts of proteins than currently possible. This will make identification of proteins interacting with a single transcript more feasible. For the time being, however, the usefulness of methods identifying proteins interacting with a single RNA of interest is restricted to relatively highly expressed targets. The limited toolbox of methods allowing effective and cost-efficient investigation of lncRNA interactors calls for development of new approaches.

1.1.5. Properties of circular RNAs

Another group of transcripts that has unexpectedly emerged as important regulators of cells biology are circRNAs. They do not have a free 5' or 3' terminus, are generally classified as lncRNAs and are much more stable then linear transcripts, because of a lack of free termini that could be targeted by exonucleases. The existence of cellular RNA molecules with circular configuration has been reported for the first time around four decades ago (Hsu and Coca-Prados 1979; Halbreich et al., 1980). Later examples of circRNAs being produced from different genes also expressing mRNAs have been reported, but their existence have been mostly attributed to splicing defects and their potential functions have not been studied in detail (Cocquerelle et al., 1992; Capel et al., 1993; Cocquerelle et al., 1993). Those early studies, however still yielded some information about circRNAs. Certain sequence properties required for their production and their high stability have been determined. Notably, circRNAs may exhibit a half-life of more than 48 h, whereas typical half-life of mRNAs are rarely longer than 10 h (Jeck and Sharpless 2014). They have also been shown to arise from both exonic and intronic regions of their host genes (Halbreich et al., 1980; Capel et al., 1993; Pasman et al., 1996; Braun et al., 1996). In the past years the advancement of high throughput

sequencing technology enabled in-depth analysis of circular transcripts. More than 20 000 eukaryotic circRNAs have been identified (Glazar et al., 2014) and many among them have tissue specific expression patterns. CircRNAs were showed to arise from intronic and exonic sequences of known protein coding genes, including 5' and 3' UTRs as well as from intergenic regions of the genome (Memczak et al., 2013; Zhang et al., 2013). Recently models for biosynthesis of circRNAs have been presented. Most of those transcripts seem to be produced by the backsplicing event of joining a 5' splice donor to an upstream 3' splice acceptor (Jeck et al., 2013). The process does not show any U2 or U12 spliceosome preference (Guo et al., 2014). It has also been reported that exons that are circularized are flanked by large introns (Salzman et al., 2012; Westholm et al., 2014) and that repetitive sequences in the flanking intron promotes this process (Jeck et al., 2013). Alongside the repetitive sequences RNA binding proteins such as Quaking (QKI), RNA-binding motif protein 20 (RBM20), the dsRNA binding/editing protein ADAR (Ivanov et al. 2015) and Muscleblind (MBL) have been shown to promote the backsplicing event (Ashwal-Fluss et al., 2014; Conn et al., 2015; Khan et al., 2016). Still, the investigation of the regulation of circRNA biogenesis is likely only at its beginning.

1.1.6. Functions of circular RNAs

The molecular functions of circRNAs are still poorly studied. The pioneering work on their roles in regulation of cells biology came from two independent studies. In the first study a circRNA antisense to the cerebellar degeneration-related protein 1 transcript (CDR1as) was discovered to be an important regulator in neuronal tissues, which is highly conserved in different animal species (Memczak et al., 2013). Multiple binding sites for miR-7 were identified on this transcript and it was shown that the circRNA is densely bound by miRNA effector complex. The nature of miR-7 binding sites on CDR1as was shown to be favorable for interaction with the silencing complex, but not extensive enough to promote degradation of the transcript. Thus, CDR1as was shown to be able to fine-tune the activity of miR-7 by acting as a sponge, which can sequester this miRNA from binding to its mRNA targets. CDR1as regulatory potential in animals was later studied in more details. This circRNA has been determined to be an important regulator of brain functions in mouse. Its interaction with miR-7 was confirmed in human and mouse brain. Mouse with CDR1as knock-out displayed neurological defects, which were associated with changed expression of genes in the brain, which regulate neuronal activity (Piwecka et al., 2017). The second pioneering study on roles of circRNAs reported a function of circular transcript named ciRS-7 originating from the

mouse Sry gene. This RNA acts as a molecular sponge for miR-138 (Hansen et al., 2013). The circRNA binding sites for miR-138 are mismatched in a way that allows RISC complex binding but prevents RNA cleavage. This interaction sequesters the miRNA molecules from binding to their mRNA target, thus fine-tuning their activity. Other studies have shown that different circRNAs contain binding sites for multiple miRNAs, suggesting that miRNA sponging may be a common feature of various circular transcripts (Zheng et al., 2016; Yang et al., 2016; Hsiao et al., 2017). A group of intron-retaining circRNAs have been also shown to act as miRNA decoys, but an additional role has been ascribed to them. They localize to promoters of genes they originate from, associate with Pol II and enhance the efficiency of transcription of their parental genes in cis (Li et al., 2015). Other circRNAs have been shown to play roles in processes like suppression of exonuclease mediated maturation of rRNA (Holdtet al., 2016) or modulating protein-protein interactions (Du et al., 2016). CircRNAs are generally classified as ncRNAs, however it has been shown that at least some of them actually are messengers used for protein translation. The indication that endogenous circular transcripts are being translated in human cells came from indirect lines of evidence first (Abe et al., 2015). Recently, however direct evidence for circRNA translation has been provided. In one study (Pamudurti et al., 2017) a group of circRNAs was shown to associate with translating ribosomes in *Drosophila*, presence of a protein coded by a specific circRNA in flies was confirmed and in-vivo translation of circRNA reporters was observed. In another study a circular transcript named circ-ZNF609 specifically controlling myoblast proliferation have been directly shown to be translated into protein (Legnini et al., 2017). Those discoveries directly challenge categorization of circRNAs as lncRNAs and showed that the family of circular transcripts is more diverse than previously expected. Other examples of different functions of circRNAs have been published, however most of those transcripts have not been studied to date and novel pathways in regulation of which they are involved are likely to be discovered. CircRNAs have been implied in the regulation of diverse processes in the human body as well as pathologies, including neurodegenerative diseases and cancer (Greene et al., 2017; Li et al., 2017; Bolha et al., 2017). It is evident that further study of circRNAs as well as linear lncRNAs is necessary to broaden our understanding of human biology and progression of pathologies.

1.2. Cellular response to genotoxic stress

Cells in the human body have to constantly deal with factors that threaten their homeostasis and cause them stress. All cellular components are subject to damage, however alterations of DNA as a repository of cells generic material are far more dangerous than lesions in proteins or RNAs, which are relatively quickly turned-over. Depending on the source of damage, DNA may experience lesions such as modifications of bases, single strand breaks (SSBs), or double strand breaks (DSBs). DNA damage has the potential to interfere with essential cellular processes like transcription (Svejstrup, 2010) or replication (Branzei and Foiani, 2005) and may be lethal for the cell. Damages to genetic material contribute to aging process and can lead to mutations that cause cancer or other diseases (Hoeijmakers, 2009). Efficient coping with DNA damage is critical as according to estimations each day every cell in human body faces tens of thousands DNA-damaging events (Lindahl, 2000). To deal with such a common yet dangerous threat, cells in the human body had to develop efficient mechanisms of coping with this so-called genotoxic stress. In response to damage of their genetic information cells initiate highly coordinated cascades of events. Repair of different kinds of DNA lesions is carried out by distinct, specialized repair pathways. Cellular mechanisms of dealing with genotoxic stress are generally known as DNA damage response (DDR) pathways. They are comprised of three major steps. First, the damage is located and recognized by sensor proteins. Next, repair factors are recruited to the site of damage. In the third step, the actual repair is carried out by effector proteins. If the damage is persistent or irreparable, DDR pathways signal cells to activate apoptosis (Roos and Kaina, 2006). The relevance of those pathways for protecting the genome is evidenced by the fact that mutations in factors responsible for DNA damage repair are among the most frequently accumulated in various cancers (Parikh et al., 2018).

1.2.1. DNA repair mechanism

There are five major repair pathways that resolve different types of DNA damage. Mismatch repair (MMR) corrects DNA mismatches, which mostly arise from mistakes of the proofreading activity of the replication polymerases (Ravi et al., 2006). Base excision repair (BER) is mainly responsible for removing modifications of the bases that do not distort the DNA helix (Kim and Wilson 2012). Lesions repaired by BER include for example SSBs, deaminated, oxidated and alkylated bases, which are commonly introduced to DNA by by-products of cells metabolism. Nucleotide excision repair (NER) removes a wide variety of helix-distorting lesions including base cross-links (Orlando 2013). DNA damage sites targeted

by NER are commonly induced by external factors, such as the UV component of sunlight. Homologous recombination (HR) and non-homologous end joining (NHEJ) repair double strand breaks (Mao et al., 2009). NHEJ relies on direct ligation of broken DNA strands. It is generally considered error prone as it doesn't include strict control for alterations or deletions in broken DNA ends. However, most of DSBs are repaired by NHEJ and errors do not occur frequently (Beucher et al., 2009). HR on the other hand is considered error free as repair is carried out with use of sister chromatid sequence as a template for faithful repair (Jasin and Rothstein 2013). This ensures that no alterations in DNA or deletions are introduced at the damaged site.

1.2.2. Immediate response to double strand breaks

The most dangerous type of DNA damage is DSB. If not repaired, they can lead to cell death and, if not repaired correctly, they can cause deletions or chromosomal aberrations (Cannan and pederson 2016). It has been shown that even a single DSB may lead to cell death (Bennet et al., 1993; Rich et al, 2000). Mistakes in DSB repair can potentially lead to development of cancer or other diseases (Moynahan et al., 2010; Merlo et al., 2016). Thus, proper control of those DNA lesions is critical for cell viability and maintenance of genome integrity. Central to the cellular response to DNA double strand breaks is a protein kinase called Ataxia Telangiectasia Mutated (ATM). Inactive ATM is present in the cell as a dimer. Upon DNA damage it is recruited to damage sites by the Mre11-Rad50-Nbs1 (MRN) mediator complex, which acts as a DSB sensor (Lee and Paull, 2005). This recruitment activates the kinase, which initiates complex signal transduction pathways facilitating the damage repair. ATM phosphorylates hundreds of substrates including other kinases, which then also become activated (Bekker-Jensen and Mailand 2010). Thus, the immediate DSB response is marked by extensive posttranslational modifications of cellular proteins. Through this, ATM is crucial for DNA repair, both locally at the site of damage and globally on cellular level. This includes cell cycle regulation, changes in transcription and pre-RNA splicing and alterations in chromatin structure. One of the immediate targets of ATM at sites of damage is the variant histone H2AX, which is being phosphorylated on Ser139 (γH2AX) (Burma et al., 2001). The process leads to recruitment of additional DDR proteins and amplification of H2AX phosphorylation further away from damage site. In consequence, DNA repair foci are formed. Phosphorylation of ATM targets also leads to recruitment of E3 ubiquitin ligases. which ubiquitinate histone proteins including H2AX (Huen et al., 2007) and linker histone H1 (Thorslund et al., 2015). This allows further recruitment of DNA repair

factors and leads to strong chromatin relaxation, creating a permissive environment for repair. At this stage, the choice between HR and NHEJ is made. Either protein 53BP1 or BRCA1 may be recruited to damage site. Their actions are largely antagonistic, and they promote NHEJ and HR respectively by recruitment of specific downstream repair proteins (Moynahan et al., 1999; Bothmer et al., 2010; Bunting et al., 2010). Choice between the two pathways seems to relay on the cell-cycle phase and the extent of DNA damage. NHEJ is active throughout all cell-cycle phases and is a major DSB repair pathway even in presence of active HR (Kakarougkas and Jeggo, 2014). HR functions only in S/G2 phase, where it is utilized to repair more complex DNA damage events. It is the pathway of choice when damage occurs in heterochromatic regions and when complex DSBs with multiple damages in close proximity are induced to DNA (Shibata et al., 2011; Kakarougkas et al., 2013).

1.2.2.1. Transcriptional activity of double strand breaks

For a long time, genomic regions harboring DSBs were believed to be fully transcriptionally inactive. Indeed, proper expression of genes harboring damaged DNA sequences is not possible and could interfere with the repair process, thus it has to be inhibited (Kruhlak et al., 2007; Shanbhag et al., 2010; Pankotai et al., 2012). It has been shown recently however, that a specialized class of transcripts is produced from the sites of DSBs. Short RNAs, about 21nt long, named DSB-induced small RNAs (diRNAs) and originating from sites of DSBs have been detected and discovered to be critical for damage repair (Wei et al., 2012; Francia et al., 2012). They are being produced similarly to miRNAs, depending on DICER and DROSHA proteins, but have been shown not to function in posttranscriptional gene silencing. Instead, they target directly the damage site they originate from, by associating with long RNA products transcribed from the same genomic locations (Michelini et al., 2017). Those long RNAs seem to be at the same time substrates used for diRNAs productions and docking stations for diRNAs targeting to the damage site. γH2AX deposition does not seem to be dependent on those short transcripts, but they are necessary for recruitment of downstream repair factors, maintenance of repair foci and removal of DSBs. This mechanism seems to be evolutionally conserved as functional diRNAs have also been detected in plants (Wei et al., 2012).

1.2.3. Cellular environment of DNA damage repair

In addition to triggering direct actions to repair DSBs DDR activates pathways that induce more global changes to cellular environment. The most obvious example of this is checkpoint activation, which stops progressing through the cell cycle providing time to carry out the DNA repair before replication or mitosis. To achieve this, ATM directly phosphorylates Chk2, which leads to its activation and phosphorylation of downstream targets resulting in cell cycle arrest at the G2/M checkpoint (Ahn et al., 2000). Another important example of global changes to cellular environment triggered by ATM in response to DNA damage is the effect it has on the gene expression programs of the cell. The central player in the changing transcriptional landscape of the cell in response to DSBs is the transcription factor p53 (Yogosawa and Yoshida 2018). Normally, the TP53/p53 protein is targeted for a rapid degradation and thus kept at low levels. Its activation in DDR relies mainly on the protein stabilization by ATM-dependent mechanisms. This is achieved through actions of multiple pathways, including direct phosphorylation of TP53 by ATM (Canman et al., 1998; Banin et al., 1998; Moumen et al., 2013). TP53 directly promotes transcription of many DNA repair genes. Stabilization of the protein leads to increased expression of its targets and reinforcement of cell cycle arrest. In case of persistent DNA damage, p53 activity may lead to senescence and apoptosis. The significance of TP53 for safeguarding the genome is underscored by the fact that it has been reported to be one of the most frequently mutated genes in human cancers (Lakin and Jackson 1999).

An additional level of gene expression regulation present in response to DSBs relies on posttranscriptional regulation. DNA damage leads to changes in expression of many miRNAs, which add to complexity of regulation of expression of genes important for DDR (Maes et al., 2008; Templin et al., 2011; Wagner-Ecker et al., 2010; Simone et al., 2009; Jasson et el., 2008). DDR proteins regulate miRNA abundance both on the level of transcription and maturation. MiRNAs in turn are important regulators of many DDR proteins including ATM, H2AX and BRCA1 (Hu et al., 2009; Ng et al., 2010; Song et al., 2011; Wu et al., 2013; Yan et al., 2010; Wang et al., 2011). The importance of miRNA pathways in regulating DDR is evident as down regulation or overexpression of various miRNAs has been shown to lead to impairment of proper DDR (He et al., 2007; Pichiorri et al., 2010; Cannell et al., 2010; Breaun et al., 2008).

1.2.4. Long non-coding RNAs regulate DNA damage response

Also, lncRNAs have been shown to be important regulators of cellular response to DNA damage. Involvement of lncRNAs in safeguarding the genome has been speculated before their involvement was directly shown as expression of many among them is changed in cancer cells (Mourtada-Maarabouni et al., 2009; Gutschner et al., 2013; Liu et al., 2014). Moreover, it has been shown that various non-coding transcripts expression are regulated by DDR pathways, including p53 (Huarte et al., 2010; Hung et al., 2011). The first example of such transcripts playing a direct role in response to DSBs was a study showing a low copy number, RNA polymerase II-transcribed, polyadenylated, uncapped transcript originating from a genomic region upstream of the cell cycle regulator cyclin D1 (CCND1) gene promoter (Wang et al., 2008). This RNA binds TLS, an RNA-binding protein that is an inhibitor of the histone acetyl transferase CBP/p300. When DDR is triggered, this ncRNA is upregulated, recruits TLS to CCND1 promoters and activate it. This leads to TLS interaction with CBP/p300 and repression of CCND1 transcription. Since that seminal discovery, more lncRNAs have been shown to play a role in regulating gene expression profiles of DDR genes through regulating the chromatin states of their targets. The lncRNA JADE regulates activity of its neighbor gene JADE1, which globally regulates histone H4 acetylation in DDR, contributing to establishment of specific transcriptional response to genotoxic stress (Wan et al., 2013). JADE regulates its target by facilitating BRCA1 interaction with CBP/p300, which allows for efficient induction of JADE1 expression after DNA damage. Another lncRNA, APTR, associates with PRC2 to target the promoter of the p21 gene, which is a direct p53 target, responsible for cell cycle regulation (Negishi et al., 2014). This interaction results in deposition of H3K27me3, thus leading to transcriptional silencing. Epigenetic regulation of gene expression is a classic mode of action of known lncRNAs and high numbers of those molecules differentially expressed in response to genotoxic stress indicates that this mechanism may be more widespread in DDR than currently appreciated.

Involvement of lncRNAs in response to DSBs is not restricted to regulating chromatin states of DDR gene promoters. They have also been shown to play roles in processes previously thought to depend solely on activity of DDR proteins. An example of this is exemplified by a lncRNA called damage induced noncoding (DINO) (Schmitt et al., 2016). This transcript is upregulated in DDR directly by p53. When expressed, DINO in turn associates with p53 to promotes its stabilization, thereby promoting a positive feedback loop. Depletion of DINO does not affect p53 phosphorylation, yet it is necessary for this

transcription factor's ability to induce expression of its target genes. Another example of lncRNA regulating p53 is a product of the WD repeat containing antisense to p53 (WRAP53) gene (Mahmoudi et al., 2009). WRAP53 is an interesting example of a gene that encodes both a protein coding transcript and a lncRNA. WRAP53β codes for a protein that has a function in DSB repair. The second one, WRAP53γ is a lncRNA with a function in DDR that is mechanistically independent of the function of its protein coding sibling. This transcript directly interacts with the 5' UTR of the p53 mRNA via RNA-RNA interactions to promote p53 expression. It appears that the function of WRAP53γ is mainly to stabilize the p53 mRNA, as when it is knocked down it results in lower levels of both p53 protein and mRNA. LncRNAs have also been shown to regulate direct repair of DSBs. A non-coding transcript called DNA damage-sensitive RNA1 (DDSR1) has been reported to play an important function directly in repair through HR (Sharma et al., 2015). DDSR1 has been shown to interact with factors associated with HR including BRCA1 and heterogeneous nuclear ribonucleoprotein U-like 1 (hnRNPUL1). Loss of DDSR1 does not impair localization of BRCA1 and its interactors to site of DSB, instead leads to their increased accumulation. This accumulation is associated with reduced efficiency of HR. Thus, DDSR1 has been shown to be important for events downstream from binding of initial HR factors to DSB. It has been speculated that the role of DDSR1 in this regulation is to mediate BRCA1-hnRNPUL1 interactions, as depletion of hnRNPUL1 results in similar HR defects. The exact molecular mechanism of this regulation is not known, but it is a fascinating example of previously unanticipated involvements of lncRNAs directly in DSB repair.

More lncRNAs have been shown to play important roles in cellular responses to DSBs, however their number is so far limited. This does not mean that DDR is not regulated by lncRNAs to a much higher extent than currently appreciated. Research of functions of lncRNAs in DDR is still in its infancy, however their relevance in response to genotoxic stress is unquestionable. Increasing lines of evidence point to lncRNAs playing much broader roles in DDR then currently known. As mentioned above, high throughput RNA sequencing experiments show many lncRNAs with expression values changed upon induction of genotoxic stress. Many more among those transcripts are misregulated in cancer cells (Mourtada-Maarabouniet al., 2009; Gutschner et al., 2013; Liu et al., 2014), pointing to their potential roles as tumor suppressors or proto-oncogenes. As of now, few of those lncRNAs were functionally characterized, thus new ways through which non-coding transcripts play important roles in safeguarding the genome are likely to be discovered. Studies of roles of

lncRNAs in DDR and cancer are also driven by the fact that they are potentially suitable for targeting in treatments. Some lncRNAs have been recognized as useful cancer biomarkers (Ji et al., 2003; Fradet et al., 2004; Shappell 2008; Rasool et al., 2016). Many of them are also shown to be expressed specifically in cancer cells, which makes them good candidates for targeting in new cancer therapies, potentially having no side effects to the patient (Arun et al., 2018; Renganathan and Felley-Bosco 2017).

1.2.5. RNA binding proteins in DNA damage response

Another group of regulators with roles in DDR that were previously underappreciated are RNA binding proteins (RBPs). The relevance of RBPs is evidenced for example by the above described roles of ncRNAs in the repair of DSBs. The functions that many lncRNAs exert in regulation of cellular physiology are carried out through their interaction with proteins. Processes like regulation of chromatin states are directly carried out by proteins, which are being directed to their targets by lncRNAs. Thus, the relevance of transcripts in regulation of DDR as described above has to be considered in concert with the relevance of their interacting RBPs. The importance of proteins with RNA binding properties in response to genotoxic stress is further supported by the fact that proteins with known function in DDR are highly enriched in experiments identifying the polyA-RNA-bound proteome (Baltz et al., 2012). Indeed, increasing number of canonical DNA repair proteins are reported to interact with RNA to exert their function. A good example of this are the above described cases of p53 and BRCA1 binding to lncRNA molecules. In addition to increasing evidence of canonical DDR proteins having RNA binding properties, canonical RBPs are being recognized as regulators of DNA repair. Many such proteins have been identified in screens for proteins post-translationally modified in DNA damage or for those required for DDR (Matsuoka et al., 2007; Paulsen et al., 2009; Adamson et al., 2012). Involvement of RBPs in response to genotoxic stress conceptually makes sense as cells experiencing it need to regulate their RNA metabolism to restrict normal transcription from sites of DNA damage and enhance expression of genes responsible for repair. Recently, a direct approach to identify RBPs differentially interacting with polyadenylated RNAs was employed. Proteins binding polyadenylated RNAs were identified by Milek and co-workers by utilization of mRNA interactome capture after UV crosslink of proteins and mass spectrometry (Milek et al., 2017). This study revealed 266 proteins with increased binding to RNA upon IR. It showed that alongside known RNA-binders, a subset of DDR proteins with no previously known RNA interactions associated with polyadenylated transcripts upon DDR. It also showed that over

45% of RBPs identified to increase RNA binding upon DDR were components of the nucleolar proteome including a RNA helicase called DEAD-Box Helicase 54 (DDX54). In the study, DDX54 was discovered to bind 3' splice sites of target mRNAs and to associate with spliceosomal proteins, thus regulating expression of DDR proteins including many targets of p53. The helicase knock-down cells exhibited lowered survival rates after exposure to IR. The study of DDX54 is an interesting example of involvement of RNA binding proteins in DDR and signifies the importance of regulation of RNA metabolism in response to genotoxic stress.

1.2.5.1. DDX3 is implicated in DNA damage response

Among RBPs with important roles in DDR, the DEAD-Box Helicase X-linked 3 (DDX3) seems particularly interesting. DDX3 is evolutionally conserved from yeast to human (Tarn and Chang 2009). It has a crucial role in organisms as knockdown of the helicase is embryonically lethal (Li et al., 2014). The helicase is involved in multiple processes regulating gene expression on the level of RNA. The major role of DDX3 seems to be in regulating translation. It has been shown that the protein interacts with translation initiation factors and facilitates translation of mRNAs (Lee et al., 2008; Lai et al., 2008). In addition, a role of the helicase in supporting the assembly of functional 80S ribosome has been reported (Geissler et al., 2012). Roles of DDX3 in promoting translation in human cells are in line with discoveries from yeast, where the DDX3 ortholog Ded1 was shown to be required for general translation. Inactivation of Ded1 leads to polysome collapse and global downregulation of translation (Chuang et al., 1997; de la Cruz 1997). In human cells DDX3 depletion was not shown to significantly affect global translation (Lai et al., 2008). However, *DDX3* is an essential gene and it could only be studied upon knock-down in human cells. Levels of DDX3 expression upon knock-down, although low, may be sufficient to maintain protein function. Thus, extent to which DDX3 regulates translation in human cells remains a matter of speculation. It has been proposed that the protein is specifically required for translation imitation of mRNAs containing structured 5' untranslated region (UTR) (Soto-Rifo et al., 2012). One study reported an opposite role of DDX3 in regulation of translation. It has been shown that the protein represses the cap-dependent translation by interacting with Eukaryotic Translation Initiation Factor 4E (eIF4E) and trapping it in an inactive complex (Shih et al., 2008). DDX3 is also an important regulator of formation of stress granules (SG). SG are cytoplasmic assemblies of untranslating mRNAs that aid cell survival upon environmental insults (Buchan and Parker 2009). DDX3 facilitates formation of SG and its interaction with eIF4E is necessary for this process (Shih et al., 2012). Yet another pathway in which the

helicase seems to be involved is post-transcriptional gene silencing. DDX3 has been shown to facilitate its efficiency and to co-localize in cytoplasm with Ago2 (Kasim et al., 2013). What makes DDX3 mediated RNA regulation even more complex is the fact that it has also been reported to associate with lncRNAs, albeit to a much lower extent in comparison to its binding to mRNAs (Oh et al., 2016). The relevance of DDX3-lncRNA interactions has not been studied, but it seems plausible that as a RNA helicase it may play a role in their biogenesis or turnover. The protein has also been directly shown to play a role in regulating response to genotoxic stress. Its functions in this process are not fully understood, however DDX3 is known to associate with p53, regulate its accumulation and control p53 dependent apoptotic pathways after induction of genotoxic stress (Sun et al., 2013). In addition, the previously mentioned study by Milek and co-workers showed that DDX3 binding to polyadenylated RNAs is increased in irradiated cells, suggesting that it may play a role in DDR.

Dysfunctions in DDX3 have been reported to be associated with diseases, most notably with cancers. DDX3 was shown to be a proto-oncogene, capable of promoting metastasis in breast cancer (Botlagunta et al., 2008). On the other hand, it has been shown to also function as a tumor suppressor, through its regulation of p21 gene expression (Chao et al., 2006). The seemingly contradicting functions of the helicase in tumorigenesis may be explained by the plethora of cellular processes in control of which DDX3 is involved. The discrepancies in reported functions of DDX3 in tumorigenesis may be at least in part explained in differences in cellular context of different cell lines and types of cancer studied. It is however clear that mutations in DDX3 are common in cancer. DDX3 appears to be especially relevant in medulloblastomas associated with aberrations in Wingless/Integrated (WNT) signaling pathway in children and in medulloblastomas associated with aberrations in sonic hedgehog (SHH) signaling pathway in adults, where it has been reported to be mutated in over 50% and over 60% of cases respectively (Northcott et al., 2012; Kool et al., 2014). Those mutations seem to occur exclusively in the helicase domains of the protein (Pugh et al., 2012; Epling et al., 2015; Floor et al., 2016), which points to a relevance of the RNA dependent functions of DDX3 in these types of cancer. The roles that DDX3 seems to play in metastasis led to development of therapies targeting it for cancer treatment (Bol et al., 2015; Heerma van Voss et al., 2015; Wilky et al., 2016). The exact mechanisms involving DDX3 in cancer remain not fully understood, however and further study of this multi-functional

helicase are warranted. In depth knowledge of DDX3 molecular functions may contribute to our understanding of cancer biology and help in designing new treatments.

2. Materials

2.1. Materials for cell culture

2.1.1 Cell lines

Flp-In T-REX HEK293 – a human kidney cell line immortalized with adenovirus 5 DNA containing a single stably integrated FRT site at a transcriptionally active genomic locus and a stably expressed Tet repressor (Invitrogen cat. no. R75007) was purchased from Thermo Fisher Scientific.

MCF-7 – a human breast adenocarcinoma cell line (ATCC cat. no. HTB-22) was purchased from ATCC

2.1.2. Chemicals and antibiotics for cell culture

Table 1. Chemicals used for cell culture.

Full name	Short name	Supplier	Product no.
Dulbecco's Modified Eagle Medium (DMEM) Highglucose	DMEM (highglucose)	Thermo Fisher Scientific	11965092
Opti-MEM I Reduced Serum Medium	Opti-MEM	Thermo Fisher Scientific	31985070
Fetal Bovine Serum	FBS	Thermo Fisher Scientific	26140079
L-Glutamine (200 mM), liquid	Glutamine	Thermo Fisher Scientific	25030081
Trypsin-EDTA (0.05%)	Trypsin	Thermo Fisher Scientific	25300054
Phosphate buffered saline, pH 7.4	PBS	Thermo Fisher Scientific	10010023
4-thio uridine	4sU	Chem Genes	RP-2304
Dimethyl sulfoxide	DMSO	Carl Roth GmbH	A994
Hygromycin	Hyg	Invivogen	ant-hg-5
Puromycin	Puro	Invivogen	ant-pr-1
Doxycicline	Dox	Sigma-Aldrich	D9891

2.2. Antibodies, siRNAs and plasmids

Table 2. Antibodies used in this work

Name	Origin	Supplier	Product no.
Primary antibodies			
DDX3	Rabbit	Bethyl Laboratories, Inc.	A300-474A
Vinculin	Mouse	Sigma-Aldrich	V9131
HA	Mouse	Covance	MMS-101P

Secondary antibodies			
Amersham ECL Rabbit IgG, HRP-linked	Donkey	Ge Healthcare Life Sciences	NA934
Amersham ECL Mouse IgG, HRP-linked	Sheep	Ge Healthcare Life Sciences	NA931

Table 3. siRNAs used in this work.

Name	Sequence (from 5' end)	supplier
siPOOL DDX3	GCAACAACTGTCCTCCACA; GCATACTATTACAGGGAAA; GAGGTGATGTATGAATACA; CTGCCAAACAAGCTAATAT; CAGTTAATAAGGTTTCAAA; CGCATGTACCAGCATCCAT; CTAGCCAAATGTGGGCATA; GGATCTCGTAGTGATTCAA; GGTTGGAAATATGTACATA; CCATAAATAATATAAGGAA; GGTCTGATAACTTGAAATA; GGCTCCAGAAGAGTAACAA; GTAACAAACTGAAATCTTT; GCACATTGCAATCCTCAAA; GGCATAATCAGTGACTTGT; GGCAAACTGTATTAAGTTA; CTCAAGTCACTGTAGCTTT; CTCCTTTGTTGTTGTCAAT; GGGAAGTCTAGCTTCTTCA; GACCTGAACTCTTCAGATA; GCTTACTATAGACTTCGTA; GGCAGATCATTAATTATGA; GAGAATTCATCTACTTAGA; GCAAGGATTCACTGACCTT; GGCCAAAGATGAGCATTGT; GAGCCTCAGATTCGTAGAA; CTTTAACGAGAGGAACATA; GTGTAAATAAAGTGCTCTT; CCATAACTTTCTGATGTTA; GCCAGAATGCGGTGATCAA	siTOOLs Biotech
siPOOL negative control	Proprietary	siTOOLs Biotech
siLINC00475	AGACAGAAGAUGUGGAAAAUU	Eurofins Genomics
siLINC01021	CAUAGAUGCAAUAAGGUUUUU	Eurofins Genomics
siTCERG1L-AS1	CAAUAUAUCCCAGUGUAAAUU	Eurofins Genomics
siUNC5B-AS1	AGGCCUUCCGCAAAGUGUUUU	Eurofins Genomics

Table 4. Plasmids used in this work.

Name	Origin	Description
pENTR4	Thermo Fisher Scientific	gateway entry vector
pFRT-TO-FLAG-HA-DEST	Spitzer et al., 2013	gateway destination vector
pFRT-TO-STREP-HA-	Glatter et al., 2009	gateway destination vector

DEST		
pFRT-TO-GFP	Spitzer et al., 2013	plasmid expressin GFP
pcDNA5-FRT-TO	Thermo Fisher Scientific	expression vector
pHR510-mAIDHA-RFP	System Biosciences	mAID CRISPR homology recombination vector background
pHR310-mAIDHA-RFP	System Biosciences	mAID CRISPR homology recombination vector background
px458	Ran et al., 2013	Cas9 and CRISPR guide RNA expressing vector background
pog44	Thermo Fisher Scientific	Flp-Recombinase Expression Vector
pEX-Csy4	Eurofins Genomics (gene synthesis)	vector containing human codon optimized Csy4 sequence
pEX-DDX3HA	Eurofins Genomics (gene synthesis)	vector containing DDX3 homology arms for CRISPR knock-in

2.3. List of oligonucleotides

Table 5. Primers used for molecular cloning.

Forward Primer	Sequence (from 5' end)	Reverse Primer	Sequence (from 5' end)	Target gene
HULC_BamHI.F	aggatccATGGGGGTG GAACTCATGATG	HULC_XhoI.R	tactcgagAAGAATGGACA TCATTTTATTTCATTTT AATTTAGTTTTGTTTA	HULC
Csy4HP.F	agatctgtagaaaGTTCAC TGCCGTATAGGCAG	Csy4HP.R	ggatcccccagcagCTGCCT ATACGGCAGTGAAC	-
DDX3XgRNA1i	caccGAAGCTACACA AGGTATAGTC	DDX3XgRNA1ii	aaacGACTATACCTTGT GTAGCTTC	-
DDX3XgRNA5i	caccGCCATATTAGCT TGTTTGGCA	DDX3XgRNA5ii	aaacTGCCAAACAAGC TAATATGGC	-
DDX3XgRNA11i	caccGTACCCCACCA GTCAACCCCC	DDX3XgRNA11ii	aaacGGGGGTTGACTG GTGGGGTAC	-
DDX3XgRNA14i	caccGATAACTCCCA GGGGGTTGAC	DDX3XgRNA14ii	aaacGTCAACCCCCTG GGAGTTATC	-
DDX3XgRNA24i	caccGAGGAAATTAT AACTCCCAGG	DDX3XgRNA24ii	aaacCCTGGGAGTTAT AATTTCCTC	-

Table 6. Primer used for amplification in qRT-PCR

Forward Primer	Sequence (from 5' end)	Reverse Primer	Sequence (from 5' end)	Target gene
HULC.Fq	AAACTCTGAAGTA AAGGCCGGAA	HULC.Fr	TTGCTTGATGCTT TGGTCTGT	HULC
LINC00475.Fq	GCAGCTTTCAGGA AGGAACACC	qLINC00475.Rq	AGCCATACAATC CCGTCCTGG	LINC00475
LINC01021.Fq	AAGGGGGAGCAT AAGCTAAGGAA	LINC01021.Rq	GTCACTCCTTTTT CCACTTCTTCCC	LINC01021

TCERG1L-AS1.Fq	CTCTGCTAAGTGGGACTCGGG	TCERG1L-AS1.Rq	GCCCAGTGTTCCTTCCTGAGA	TCERG1L-AS
UNC5B-AS1.Fq	CTCAAGAGGTTGGGACTGGGG	UNC5B-AS1.Rq	ATGCCAGCTTCCCCCAAAAAG	UNC5B-AS1
Csy4.Fq	ATCCCAATCTACGGGGCAAC	Csy4.Fr	GAGTCAGAACCACGGGACAA	Csy4
chr17TU.Fq	GTGACAGGTGGCCTCACTCA	chr17TU.Rq	TGCAGGGACGCAGATACGAC	C17hr:80 454 535-80 454 635
GAPDH.Fq	GTCTCCTCTGACTTCAACAGCG	GAPDH.Rq	ACCACCCTGTTGCTGTAGCCAA	GAPDH

Table 7. Primers used for 3’ RACE

Forward Primer	Sequence (from 5' end)	Reverse Primer	Sequence (from 5' end)	Target gene
LINC00475.Fr	ACACTCACCGGGGTTTCGAC	Universal Amplification Primer	propriatary (Thermo Fisher Scientific)	LINC00475
LINC01021.Fr	ACCAGGGTGAGTGAACAAAGGA	Universal Amplification Primer	propriatary (Thermo Fisher Scientific)	LINC01021
TCERG1L-AS1.Fr	TGTTCTGGACTGTCCCGTCG	Universal Amplification Primer	propriatary (Thermo Fisher Scientific)	TCERG1L-AS
UNC5B-AS1.Fr	CTTAGGGTCTGCGATCCGCC	Universal Amplification Primer	propriatary (Thermo Fisher Scientific)	UNC5B-AS1

2.4. Buffers and solutions

Table 8. Buffers and solutions used in this work

Buffer	Composition
2x SDS loading buffer	100 mM Tris-Cl (pH 6.8); 4% (w/v) SDS; 0.2% (w/v) bromophenol blue; 20% (v/v) glycerol; 200 mM DTT
TBS-T	20 mM Tris; 150 mM NaCl; 0.1% Tween 20; pH 7.6
Fixation Solution I	30%; 15% acetic acid
Fixation Solution II	25% ethanol; 4.1% (w/v) NaOAc; 0.3% (w/v) Na2S2O3-5H2O; 0.125% glutaraldehyde
Staining Solution	0.1 % $AgNO_3$; 0.011% formaldehyde
Developer Solution	2.5% Na2CO3; 0.011% formaldehyde
NP-40 lysis buffer	50 mM HEPES-KOH; 150 mM KCl; 2 mM EDTA; 1 mM NaF; 0.5% (v/v) NP-40; 0.5 mM DTT; complete EDTA-free protease inhibitor cocktail; pH 7.4
2x Proteinase K buffer	200 mM Tris-HCl; 300 mM NaCl; 25 mM EDTA; 2% (w/v) SDS; pH 7.5

Csy4 Elution Buffer	20 mM Tris (pH 7.5); 100 mM NaC; 5% (v/v) glycerol; 1 mM DTT; 500 mM imidazole
Strep Tactin Wash Buffer	50 mM HEPES-KOH; 500 mM NaCl; 2% (v/v) NP-40; 2 mM DTT; complete EDTA-free protease inhibitor cocktail; pH 7.4
Mammalian Polysome Buffer	TRIS-HCl (pH 7.4); 100 Mm; NaCl 750 mM; MgCl2 25 mM; DTT 1 mM; CHX 100 ug/m
Mammalian Polysome Lysis Buffer	TRIS-HCl (pH 7.4) 100 Mm; NaCl 750 mM; MgCl2 25 mM; DTT 1 mM; CHX 100 ug/ml; Triton X-100 1 % (v/v); Turbo DNase 25 U/ml
2x Formamide Loading Buffer	95% formamide; 20 mM EDTA (pH 8.0); 0.05% (w/v) bromophenol blue; 0.05% (w/v) xylene cyanol
1x TBE	220 mM Tris; 180 mM Boric acid; 5 mM EDTA pH 8.0
High Salt Wash Buffer	100 mM; Tris-HCl (pH 7.4); 10 mM EDTA; 1 M NaCl; 0.1% (v/v) Tween20

3. Methods

3.1. Mammalian cell culture

3.1.1. Media composition for cell culture

Q-medium – DMEM (high glucose), supplemented with 10% FBS and 2 mM L-glutamine.

3.1.2. Culturing of mammalian cells

Adherent HEK293 and MCF-7 cells were grown in Q-medium and subcultured two times a week. Confluent cells were first washed with PBS and incubated with 1x Trypsin-EDTA at 37ºC for a few minutes, until cells detached from the plate surface. Trypsinization was stopped by addition of serum containing media. A defined volume of cell suspension (typically 1/8 or 1/10 volume) was transferred to a new plate with fresh culture medium.

3.1.3. Freezing and thawing of mammalian cell lines

After reaching confluency on a 10 cm dish cells, were trypsinized and centrifuged for 2 min at 300 xg. The cell pellet was resuspended in 8 ml of Q-medium, containing 10% DMSO and cell suspensions were aliquoted in 2 ml cryovials at 1 ml/vial. Vials were placed in a Cryo-container and stored overnight at -80ºC. The next day the vials were placed in a liquid nitrogen tank for storage.

To reculture cells, cryovials containing frozen cells were thawed in a water bath at 37ºC. Cells were transferred to a sterile 15 ml tube and mixed with 8 ml of growth medium. Tubes were centrifuged for 2 min at 300 g. Supernatants were discarded from pelleted cells and pellets were resuspended in 10 ml of growth medium. Cell suspensions were transferred to fresh 10 cm plates and cells were cultured according to description in section 3.1.1.

3.1.4. Reverse siRNA transfection

SiRNAs were resuspended in RNAse-free water and aliquoted. On the day of transfection, 80 pM siRNAs was diluted in 150 µl Opti-MEM and mixed with 150 µl Opti-MEM, containing 9 µl Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher Scientific). After 5 min incubation at room temperature, siRNA-Lipofectamine mixtures were transferred to 6-well plates. Trypsinized MCF-7 cells in 1700 µl medium were added to each

well, resulting in final siRNA concentration of 40 nM. 24h after transfection, medium was exchanged, and cells were harvested 48h after transfection. Knock-down efficiency was confirmed with qRT-PCR.

Transfection with siPools was performed in the same way as for single siRNAs, with a final siRNA concentration used 3 nM instead of 40 nM.

3.1.5. Forward DNA transfection

Forward DNA transfection for transient expression

On the day of transfection 12.5 μg of plasmid DNA was diluted in 750 μl Opti-MEM and mixed with 750 μl Opti-MEM, containing 11 μl Lipofectamine 2000 DNA Transfection Reagent (Thermo Fisher Scientific). After 5 min incubation at room temperature DNA-Lipofectamine mixtures were transferred to cells grown on 10cm dishes (confluency around 60%), containing 8.5 ml of fresh medium. 24h after transfection medium was exchanged and doxycycline was added to final concentration of 1 μg/ml to induce expression from the plasmid. 48h after transfection, cells were washed with PBS and harvested.

For every transfection experiment, transfection with a plasmid expressing GFP protein was performed and used to estimate transfection efficiency by observation under fluorescent microscope. Only cells from transfections displaying more than around 70% successfully transfected cells for the positive control plasmid were used for experiments.

Forward DNA transfection for generation of stable cell lines

On the day of transfection 2.5 μg of plasmid DNA was diluted in 150 μl Opti-MEM and mixed with 150 μl Opti-MEM, containing 2 μl Lipofectamine 2000 DNA Transfection Reagent. After 5 min incubation at room temperature, DNA-Lipofectamine mixtures were transferred to cells grown on in 6-well format (confluency around 70%), containing 1.7 ml of fresh medium. 24h after transfection, medium was exchanged and 48h after transfection cells were trypsinized and seeded on 10cm dishes. They were further grown in media containing appropriate antibiotic for selection of transformed cells. Media was exchanged every 48h and after 14 days of selection, single cell colonies were picked and propagated for further analysis. Propagated positively selected cells were tested for expression of inserted coding sequences by Western analysis.

3.1.6. Inducing DNA damage by treating cells with ionizing radiation

MCF-7 cells were grown under standard culturing conditions. On the day of treatment cells on plates were transported in a closed Styrofoam box to the site of treatment. Cells were irradiated with 10 (for RNA sequencing experiments) or 6 (for colony formation assay) Gy of ionizing radiation using a cesium-137 γ-ray source. Cells were transported in a Styrofoam box to be placed back in the incubator and were cultured normally until they were used for analysis.

3.1.7. Colony formation assay

Colony formation assay was performed on MCF-7 cells after reverse transfection with appropriate siRNAs (section 3.1.3.) and mock transfected cells as negative control. Two batches of cells were prepared per transfection to test colony formation capabilities with and without treatment. Twenty-four hours after transfection cells grown in 6-well format were exposed to IR (section 3.1.5.), while control cells were left untreated. IR exposed and untreated cells were trypsinized and 1/32 volume of cell suspension was seeded in new 6-well plates. Cells were grown under standard conditions for two weeks. Afterwards media from the cells was removed and cells were washed with PBS. Cells were fixed for 2h in PBS containing 6.0% (v/v) glutaraldehyde and stained for 1h in PBS containing 0.5% w/v crystal violet. Stained colonies were counted using a stereomicroscope. Results of the clonogenic assay were analyzed with Microsoft Excel 2010. Average number of colonies from two biological replicates per each condition and each siRNA transfection, as well as standard deviation of the replicates were calculated. Statistical relevances of differences between numbers of colonies obtained with transfections with different siRNAs were calculated using student t-tests.

3.1.8. UV cross-linking of human cells

Growth media was removed from HEK293 cells. Cells were washed in ice cold PBS and placed uncovered on a tray filled with ice. Tray with ice and cell plates were placed in Stratagene 1800 device and irradiated with 0.2 J/cm2 total energy of 254-nm UV light. After cross-linking cells were scraped off, suspended in ice cold PBS and centrifuged for 2 min at 400 xg, 4°C. Supernatants were discarded and cells were used directly in downstream applications or flash frozen in liquid nitrogen and stored in -80°C for later use.

3.2. Biochemical methods

3.2.1. Western analysis

Protein samples were heated in SDS loading buffer for 5 min at 95 °C, and proteins were separated by SDS-PAGE using PageRuler Plus Prestained Protein Ladder or PageRuler Prestained Protein Ladder as size marker. Proteins from the polyacrylamide gel were transferred to nitrocellulose membranes (Thermo Fisher Scientific) using the TransBlot Semi-dry Western Blot system (Bio-Rad) with 20 V for 60 min. Membranes were blocked for 60 min at room temperature in blocking buffer (5% w/v powdered milk solution in TBS-T buffer) and subsequently incubated with primary antibody, diluted in blocking buffer, at 4°C overnight, with constant agitation. Membranes were washed 3x 10 min in TBS-T and incubated with HRP conjugated secondary antibodies diluted in blocking buffer for 60 min at room temperature. Membranes were washed 3x 10 min in TBS-T and protein signals were visualized by Amersham ECL Western Blotting Detection Reagent (GE healthcare) using ImageQuant LAS 4000 (Fujifilm). ImageQuant LAS 4000 software and ImageJ were used for image processing and analysis.

3.2.2. Silver staining of proteins in polyacrylamide gels

Protein samples were heated in SDS loading buffer for 5 min at 95°C and separated by SDS-PAGE using PageRuler Plus Prestained Protein Ladder or PageRuler Prestained Protein Ladder as protein size marker. After electrophoresis gels were incubated in 50 ml Fixation Solution I for 1h at room temperature and then, in 50 ml Fixation Solution II overnight at room temperature. The next day, gels were washed 3 times in 50 ml sterile water and incubated 30 min in 50 ml Staining Solution. After the incubation gels were rinsed with sterile water and incubated in 50 ml Developer Solution until protein bands became clearly visible on the gel, then the reaction was stopped by addition of 5 ml of 50 mM EDTA. Solution was discarded and gels were rinsed with sterile water. Visualized proteins were observed, and gel images were taken using ImageQuant LAS 4000 (Fujifilm).

3.2.3. RNA extraction

To isolate total RNA from cells, after removal of growth media, cells were lysed directly on plates with 1 ml of TRIzol reagent (Thermo Fisher Scientific) per 20 cm^2 of growth surface. Lysates were transferred to tubes and mixed with 0.2 ml chloroform per 1 ml

of TRIzol used. Samples were separated into phases by centrifugation for 15 min at 12 000 g, 4 ºC. Aqueous phase containing the RNA was moved to a new tube and mixed with 0.5 ml of isopropanol, per 1 ml of TRIzol used for cell lysis. Samples were incubated for 10 min and centrifuged for 15 min at 14 000 g, 4 ºC to pellet the RNA. Supernatant was discarded, RNA pellet was washed with 1 ml of 75% ethanol per 1 ml of TRIzol used for lysis. After removal of ethanol RNA pellets were air dried and resuspended in RNAse free water. RNA concentration was measured with Nanodrop and adjusted to around 200 ng/ml by addition of RNAse free water. 10X TURBO DNase buffer to 1X final concentration and 1 µl of TURBO DNAse (Thermo Fisher Scientific) per each 10 µg RNA were added to RNA samples and incubated at 37ºC for 30min. One sample volume of Phenol/Chloroform/ Isoamyl alcohol was mixed with samples and incubated for 5 min at room temperature. Samples were centrifuged to separate phases, aqueous phase was transferred to a new tube and mixed with 1/10 volume of 3 M NaCl solution and 1 volume of isopropanol. Samples were centrifuged for 15 min at 14 000 g, 4ºC to pellet RNA. Pellets were washed with 1 ml of 75% ethanol, ethanol was removed, pellets air dried and resuspended in RNAse free water. RNA concentration was measured with Nanodrop and its integrity checked with agarose gel electrophoresis.

To isolate RNA from liquid samples 3 volumes TRIzol LS reagent (Thermo Fisher Scientific) was mixed with one volume of samples and incubated for 2 min at room temperature. Samples were mixed with 0.2 ml chloroform per 0.75 ml of TRIzol LS used. Samples were separated into phases by centrifugation for 15 min at 12 000 g, 4 ºC. The aqueous phase containing the RNA was transferred to a new tube and mixed with equal amount of isopropanol. Samples were incubated for 10 min and centrifuged for 15 min at 14 000 g, 4 ºC to pellet the RNA. Supernatant was discarded, RNA pellet was washed with 1 ml of 75% ethanol per 0.75 ml of TRIzol LS used. After removal of ethanol RNA pellets were air dried, resuspended in RNAse free water and used in subsequent analysis.

3.2.4. First strand cDNA synthesis

Up to 5 µg of RNA was mixed with 0.5 µl oligo(dT) primer (100 µM), 1 µl 10 mM dNTP mix and RNase free water to final volume of 14 µl. Mixtures were heated to 65ºC for 5 min and incubated on ice for 1 min. Samples were briefly centrifuged and mixed with 4 µl 5X First-Strand Buffer, 1 µl 0.1 M DTT and 1 µl SuperScript III RT. Mixtures were incubated for 5 min at 25ºC, then for 60 min at 50ºC and for 15 min at 70ºC. To remove RNA complementary to obtained cDNA from the sample, 1 µl (2units) of *E. coli* RNase H was

added with subsequent incubation at 37ºC for 20 min. Prepared cDNA was diluted with RNase-free water and used directly for downstream applications or stored in -20ºC.

3.2.5. Real-Time quantitative PCR

Real-Time quantitative PCR was performed with a StepOne System (Applied Biosystems) real-time PCR cycler and SYBR Green PCR Master Mix (Thermo Fisher Scientific) in two or three technical replicates for each biological replicate.

Reaction components (for list of primers in section 2.3)

Component	Stock concentration	Final concentration
Sybr Green Master Mix	2x	1x
Forward primer	10 Mm	2 μM
Reverse primer	10 Mm	2 μM
Template	depending on specific sample	
MiliQ water	up to the final reaction volume	

Real-Time quantitative PCR reaction profile:

Preincubation	95ºC 10 min	
Amplification	95ºC 15 s 60ºC 1 min	x40 cycles
Melting Curve	95ºC 15 s 60ºC 1 min 95ºC 15 s	slow heating (0.3ºC/s)

Results from the qRT-PCR reaction were calculated in Microsoft Office Excel 2010 software.

Relative transcript level

First from technical replicates representing the same biological replicate (discarding at the same time results from technical replicates that were non-representative, when more than two technical replicates were performed) mean number of cycles needed for detection of product by the cycler was counted for *GAPDH* reference gene (x) and the target gene (y) in all investigated biological repeats. Next, obtained results were substituted into the formula 2^(x-y)=z. Obtained result for investigated samples were divided by result obtained for control

samples. Means from two biological replicates were calculated. Standard deviation of results obtained for biological replicates was calculated with Excel STDEV function.

Fold enrichment of pulled-down HULC transcript

First, from technical replicates representing the same biological replicate mean number of cycles needed for detection of product by the cycler was counted for the positive sample (x) and the negative control sample (y). Next, obtained results were substituted into the formula 2^-(x-y)=z. Means from two biological replicates were calculated. Standard deviation of results obtained for biological replicates was calculated with Excel STDEV function.

Percentage of HULC transcript eluted in pull-down experiment

First from technical replicates representing the same biological replicate mean number of cycles needed for detection of product by the cycler was counted for transcripts eluted with imidazole buffer (x) and for transcripts eluted with subsequent Proteinase K treatment (y). Next, obtained results were substituted into the formula: 2^x/(2^x+2^y)*100=z to calculate percentage of transcript eluted with imidazole and 2^y/(2^x+2^y)*100=z to calculate percentage of transcript subsequently eluted with Proteinase K. Means from two biological replicates were calculated. Standard deviation of results obtained for biological replicates was calculated with Excel STDEV function.

Percentage of input Csy4 transcript pulled-down from input sample

Initially, from technical replicates representing the same biological replicate mean number of cycles needed for detection of product by the cycler was counted for the input sample (x) and the pull-down sample (y). X was adjusted by subtracting an appropriate amount of cycles representing the difference between percent of total input sample used and the percent of total pull-down sample used (a). For example if 1% of input sample and 100% of pull-down sample were used 6.644 cycles were subtracted from x. Next, obtained results were substituted into the formula 100*2^(a-y)=z. Means from two biological replicates were calculated. Standard deviation of results obtained for biological replicates was calculated with Excel STDEV function.

3.2.6. Csy4 mediated RNA pull-down

3.2.6.1. RNA pull-down with in-vitro purified Csy4

Preparation of biotinylated Csy4 mutant protein

One mg of recombinant in-vitro purified Csy4 protein (a kind gift from the Meister laboratory, University of Regensburg) was mixed with biotinylation buffer to final volume of 455 µl. Then 45 µl 10 mg/ml EZ-Link maleimide-PEG2-biotin was added and samples were incubated overnight at room temperature. After biotinylation, protein was dialyzed using Slide-A-Lyzer Dialysis Cassette Kit (Thermo Fisher Scientific) to remove free maleimide-PEG2-biotin. First, cassettes were hydrated in PBS buffer for 2 min, then all of the Csy4 sample was applied to the cassette with a syringe and cassettes were incubated in 1000 ml of PBS overnight at 4ºC. Csy4 protein was removed from the dialysis cassette, aliquoted with final concentration of glycerol 50% and stored in -80ºC for further use.

Cellular lysis and preparation of input sample

Mock transfected cells and cells transfected with plasmids expressing HULC lncRNA tagged with Csy4 target stem loop on the 5' transcript end (section 3.1.4.) were UV cross-linked and lysed with 3 cell pellet volumes of NP-40 lysis buffer (containing complete EDTA-free protease inhibitor cocktail and 0.5 mM DTT) and incubated on ice for 10 min. Cellular debris was pelleted by centrifugation at 12 000 xg for 15 min at 4ºC. Supernatants were transferred to new RNase free tubes and 10% of the lysate from each sample was saved in a separate tube for later analysis.

Enrichment for HULC RNA from the lysate

3 pmol of biotinylated Csy4 protein was added per each 100 µl of cell extract. As a negative control equal volume of lysate was used without addition of Csy4. Mixtures were incubated for 1.5h on rotating wheel at 4ºC. Then, 0.7 µl of M-280 Streptavidin Dynabeads (Thermo Fisher Scientific) for every 100 µl of cellular extract was prepared by washing twice with 1 ml of NP-40 lysis buffer and mixed with the Csy4 incubated lysate and negative control sample. The mixtures were incubated for 1h on rotating wheel at 4ºC. Beads were concentrated on a magnet, cellular extracts were discarded and beads were washed three times in NP-40 lysis buffer containing proteinase inhibitors and 0.5 mM DTT. For each wash, beads were incubated with the buffer for 5 min at 4ºC on a rotating wheel. After the second wash, beads were transferred to fresh RNase-free tubes and the third wash was performed.

Subsequent steps of the protocol differ depending on which parameters of the pull-down approach were being assayed.

Initial test of successful HULC enrichment

Beads from Csy4 preincubated and negative control samples were mixed with 100 µl of 1x Proteinase K buffer and Proteinase K to final concentration of 2 mg/ml. Input samples, which were put aside in initial step of the protocol were mixed with 1 volume of 2x Proteinase K buffer and Proteinase K to final concentration of 2 mg/ml. Then, beads and input samples were incubated in 40ºC for 1h and mixed with three volumes of Trizol LS reagent and RNA extraction was performed (section 3.2.3.). Obtained RNA was subjected to reverse transcription (section 3.2.4.) and analyzed with RT-qPCR (section 3.2.5.) to compare enrichment of HULC RNA.

Test of imidazole driven elution from beads

After confirmation of successful target transcript enrichment described above, efficiency of imidazole driven elution was tested. Enzymatic activity of Csy4 mutant protein may be restored in presence of imidazole. This mode of elution has been shown to provide increased specificity when compared to traditional elution methods (Lee et al., 2013).

After enrichment for HULC RNA from the lysate, beads were first incubated with 50 µl of Csy4 Elution Buffer at 4ºC, overnight. Eluates were taken to fresh RNase-free tubes and mixed with 50 µl of 2x Proteinase K buffer and Proteinase K to final concentration of 1 mg/ml. To remaining beads 100 µl of 1x Proteinase K buffer and Proteinase K to final concentration of 1mg/ml were added. Both samples were incubated at 40ºC for 30 min. Then, mixed with three volumes of TRIzol LS reagent and RNA extraction was performed (section 3.2.3.). Obtained RNA was subjected to reverse transcription (section 3.2.4.) and analyzed with RT-qPCR (section 3.2.5.) to assay the fraction of the transcript eluted with imidazole and fraction of the transcript eluted with subsequent Proteinase K treatment.

Test of method specificity and efficiency

After confirmation of satisfactory efficiency of imidazole driven elution, subsequent experiments were performed with this mode of elution after enrichment for HULC RNA from the lysate. To test the specificity of the method, eluates and input samples were split in two, and one half was subjected to Proteinase K treatment and TRIzol LS extraction to analyze associated transcripts with qRT-PCR. Second half of samples was mixed with 25 mM MgCl to a final concentration of 2 mM and with 1 µl of Benzonase, then incubated at 37ºC for 30

min, boiled in SDS loading buffer at 95ºC for 5 min and used for analysis with Western blotting (section 3.2.1.) or Silver Staining (section 3.2.2.).

3.2.6.2. RNA pull-down with Csy4 expressed in human cells

Determination of successful expression of Csy4

HEK293 cells were transfected with plasmids coding for human codon optimized Csy4 protein N-end fused with FLAG/HA or STREP/HA epitopes, with its mRNA tagged with different number of repeats of Csy4 target stem loops in the 3' UTR (section 3.1.4.). Expression of Csy4 from the plasmids was assayed with Western analysis (section 3.2.1.). Satisfactory expression was only achieved with constructs expression STREP/HA version of the protein, which were used in subsequent experiments.

Cellular lysis and preparation of input sample

UV cross-linked HEK293 cells expressing human codon optimized Csy4 protein tagged with STREP/HA, with its mRNA tagged with different number of repeats of Csy4 target stem loops in the 3' end (section 3.1.4.) were subjected to cellular lysis and preparation of input sample in the same way as described section 3.2.6.1.

Enrichment for Csy4 mRNA from the lysate

One µl of MagStrep XT beads (Iba Life Sciences) for every 100 µl of cellular extract was prepared by washing twice with 1 ml of NP-40 lysis buffer and mixed with the lysate. The mixture was incubated for 1h on rotating wheel at 4ºC. Beads were concentrated with Dynal MPC-S magnetic particle concentrator, cellular extracts were discarded and beads were washed three times in Strep Tactin Wash Buffer. For each wash, beads were incubated with the buffer for 5 min at 4ºC on a rotating wheel. After the second wash beads were moved to fresh RNase-free tubes and the third wash was performed.

Elution of proteins and RNAs and test of method specificity and efficiency

To elute RNAs beads were mixed with 100 µl of 1x Proteinase K buffer and Proteinase K to final concentration of 2 mg/ml. Input samples, which were put aside in the initial step of the protocol were mixed with 1 volume of 2x Proteinase K buffer and Proteinase K to final concentration of 2 mg/ml. Then, beads and input samples were incubated in 40ºC for 1h and mixed with three volumes of Trizol LS reagent and RNA extraction was

performed (section 3.2.3.). Obtained RNA was subjected to reverse transcription (section 3.2.4.) and analyzed with qPCR (section 3.2.5.) to compare enrichment of Csy4 mRNA.

To elute proteins beads were mixed with 1x SDS loading buffer containing 2 mM MgCl2 and proteinase inhibitors and with 1 µl of Benzonase. Input samples, which were put aside in the initial step of the protocol were mixed with 1 volume of 2x SDS loading buffer containing 4 mM MgCl2 and proteinase inhibitors and with 1 µl of Benzonase. Then, beads and input samples were incubated at 37ºC for 30 min, boiled in SDS loading buffer at 95ºC for 5 min and used for analysis with Western blotting (section 3.2.1.).

3.2.7. RNA library preparation and sequencing

3.2.7.1. Preparation of polyadenylated RNA fraction

RNA samples were subjected to enrichment for polyadenylated transcripts with Oligo (dT)25 Dynabeads. 75µg of RNA was diluted to 100 µl volume with 10mM Tris-HCl pH 7.5 and 100 µl of Binding Buffer was added. Mixtures were incubated in 65ºC for 2 min and placed on ice. 1 mg of Oligo (dT)25 Dynabeads was placed in a RNase-free tube, washed once with Binding Buffer and placed on a magnet. Supernatants were discarded and RNA was added to the beads. Samples were mixed and incubated for 5 min at room temperature on a rotating wheel. Tubes were placed on a magnet for 2 min and supernatants were removed. 200 µl of Washing Buffer B was added to the beads and samples were mixed by pipetting. Tubes were placed on the magnet and supernatants removed. Wash with Buffer B was performed twice. To elute bead bound RNA 100 µl of 10 mM Tris-HCl was added and samples were incubated at 80ºC for 2 min. Then, tubes were placed on a magnet and supernatant containing eluted RNA was quickly moved to a new tube. The selection for polyadenylated RNAs was performed one more time on eluted transcripts in the same way, with a difference of elution step, which was performed in 20 µl Tris instead of 100 µl. Concentration of obtained RNA was measured with Nanodrop and its integrity checked with agarose gel electrophoresis.

3.2.7.2. Depletion of polyadenylated RNA

RNA samples were subjected to depletion of polyadenylated transcripts with Oligo (dT)25 Dynabeads. 75µg of RNA was diluted to 100 µl volume with 10mM Tris-HCl pH 7.5 and 100 µl of Binding Buffer was added. Mixtures were incubated in 65ºC for 2 min and placed on ice. 1 mg of Oligo (dT)25 Dynabeads was placed in a RNase-free tube, washed

once with Binding Buffer and placed on a magnet. Supernatants were discarded and RNA was added to the beads. Samples were mixed and incubated for 5 min at room temperature on a rotating wheel. Tubes were placed on a magnet for 2 min and supernatants were moved to a new RNase-free tube. Then, a second round of depletion was performed as described above. Supernatants after second round of depletion were mixed with 1/10 volume of 3M NaCl and 1 volume of isopropanol. Samples were centrifuged for 30 min at 20 000 g, 4ºC to pellet RNA. Liquid was removed from the pellets, they were air dried and resuspended in 50 µl of RNase free water.

3.2.7.3. Depletion of ribosomal RNA

RNA samples were depleted of ribosomal RNAs with Ribo-Zero Gold Kit (Human/Mouse/Rat) from Illumina.

Per each sample, 225 µl of beads from the kit was placed in a RNase-free tube. Tubes were placed on a magnet and supernatants were removed from the beads. Beads were washed twice with 225 µl of RNase-free water and resuspended in 65 µl of Magnetic Bead Resuspension Solution by vortexing. Then, 1 µl of RiboGuard RNase Inhibitor was added to the beads.

Per each depletion reaction up to 5 µg of RNA was mixed with 4 µl Ribo-Zero rRNA Reaction Buffer, 10 µl Ribo-Zero Removal Solution and RNase-free water to final volume of 40 µl. Mixtures were incubated at 68ºC for 10 min. Then tubes were briefly centrifuged and incubated at room temperature for 5 min. RNA samples were next added to the resuspended beads and immediately mixed by pipetting. Then, samples were vortexed for 15 s and incubated at room temperature for 5 min. Tubes were open and placed on a magnet. 90 µl of supernatant was carefully removed and placed in a new RNase-free tube. Samples were mixed with 10 µl of 3M NaAc and 1.5 µl Glycogen. Then, 300 µl of 100% ethanol was mixed with the samples, which were next incubated at -80ºC for 2h. RNA was pelleted by centrifugation for 1h at 20 000 g, 4ºC. Supernatants were removed from RNA pellets, which were then air dried and resuspended in RNase-free water.

3.2.7.4. Preparation of stranded RNA libraries

Stranded RNA sequencing library were produced using the NEXTflex Directional RNA-Seq Kit V2 (Bioo Life Science).

To fragment RNA 100 ng of RNA per sample was mixed with 5 µl of NEXTflex RNA Fragmentation buffer and RNase-free water to a final sample volume of 19 µl. Samples were incubated at 95ºC for 10 min and immediately after placed on ice. To each sample, 1 µl of NEXTflex First Strand Synthesis Primer was added and incubated for 5 min at 65ºC. Samples were placed on ice and mixed with 1 µl of SuperScript III RT and 4 µl of NEXTflex Directional First Strand Synthesis Buffer Mix. To synthesize first strand of cDNA mixtures were incubated in a thermocycler first at 25ºC for 10 min, then 50ºC for 50 min and at 70ºC for 15 min. For second cDNA strand synthesis samples were mixed with 25 µl of NEXTflex Directional Second Strand Synthesis Mix and incubated at 16ºC for 60 min.

Next, nucleic acids were purified with AMPure XP Beads (Beckman Coulter). 90 µl of beads was mixed with each sample and mixtures were incubated at room temperature for 5 min. Beads were concentrated on a magnet and supernatants were discarded. Concentrated beads were washed twice by incubating with 200 µl of 80% ethanol. After ethanol was removed tubes were removed from the magnet and air dried for 2 min. Beads were resuspended in 17 µl Resuspension Buffer and incubated for 2 min in room temperature. Tubes were placed again on a magnet and beads were concentrated for 5 min. 16 µl of supernatants was moved to new tubes. Adenylation of nucleic acids was performed by mixing samples with 4.5 µl of NEXTflex Adenylation Mix and incubation in a thermocycler at 37ºC for 30 min and then 70ºC for 5 min. To ligate adapters samples were mixed with 27.5 µl NEXTflex Ligation Mix and 2 µl NEXTflex RNA-Seq Barcode Adapter (different barcode was used for each sample). Samples were incubated at 30ºC for 10 min and reaction products were purified with AMPure XP Beads. 50 µl of beads were mixed with each sample and incubated at room temperature for 5 min. Beads were concentrated on a magnet for 5 min and supernatants were removed. Concentrated beads were washed twice by incubating with 200 µl of 80% ethanol twice. After ethanol was removed tubes were removed from the magnet and air dried for 2 min. Beads were resuspended in 51 µl of Resuspension Buffer and incubated at room temperature for 2 min. Samples were placed again on a magnet for 5 min and supernatants were moved to new tubes. The whole bead purification procedure was repeated ending with elution of nucleic acids from the beads with 36 µl Resuspension Buffer and transferring 35 µl of supernatants to new tubes.

To PCR amplify obtained nucleic acids samples were mixed with 1 µl NEXTflex Uracil DNA Glycosylase, 12 µl NEXTflex PCR Master Mix and 2 µl NEXTflex Primer Mix and subjected to a reaction with a following profile:

37	30 min	
98	2 min	
98	30 s	
65	30 s	Repeat 15 cycles
72	60 s	
72	4 min	

PCR products were purified with AMPure XP Beads. Purification was performed as described above for total of 2 rounds of purification. First purification was performed with 40 µl of beads and elution with 51 µl of Resuspention Buffer, second with 40 µl of beads and elution with 32 µl of the buffer.

3.2.7.5. Preparation and sequencing of metabolically labeled RNA ftaction

Cells were grown in Q-medium containing 1 nM 4-sU for 5 min prior to RNA extraction with TRIzol reagent (section 3.2.3.). 70 µg of RNA per sample was mixed with 3.5 µg of *D. melanogaster* 4-sU labelled RNA, to provide internal control for quantification in analysis of the data generated in the experiment. RNA mixtures were biotinylated using MTSEA biotin-XX (Biotium). Reaction mixtures were containing 73.5 µg RNA, 10 mM HEPES (pH7.5), 1 mM EDTA, and 5 µg of the biotin in total volume of 250 µl. Samples were incubated for 30 min at room temperature in the dark.

After biotinylation samples were mixed with one sample volume of Phenol/Chloroform/ Isoamyl alcohol and incubated for 5 min at room temperature. Samples aqueous phases were recovered using Phase-Lock-Gel tubes (5PRIME) and RNA was precipitated using 1/10 sample volume 3M NaCl and 1 sample volume of isopropanol. RNA was pelleted by centrifugation at 20 000 xg, 4ºC. Pellets were washed with 1 ml of 75% ethanol, ethanol was removed, pellets air dried and resuspended in 50 µl RNAse free water. RNA was incubated at 65ºC for 10 min and fallowed by incubation on ice for 5 min. Biotinylated RNA was isolated from the samples using µMACS Streptavidin MicroBeads (Miltenyi). 200 µl of beads was added to each sample and incubated at room temperature for 15 min. Next, µColumns were placed on the µMACS magnetic separator and were equilibrated with nucleic acid wash buffer. RNA-beads mixtures were appliead to the columns and flow throughs were discarded. Then, the columns were washed twice with 500 µl of warm High Salt Wash Buffer. 4-sU labelled RNA was eluted from the columns in two rounds, each with 100 µl of freshly prepared 100 mM DTT. Eluted RNA was purified using MinElute Spin

columns (Qiagen) according to manufacturer's instructions. Obtained RNA samples were measured with Qubit 2.0 Fluorometer (Invitrogen).

Entirety of obtained samples was subjected to depletion of rRNAs with Ribo-Zero Gold Kit (section 3.2.6.3.), fallowing preparation of sequencing libraries with TruSeq Stranded Total RNA Library Prep Kit (Illumina), according to manufacturer's instructions.

3.2.7.6. High Throughput Next Generation Sequencing

RNA libraries were first quantified with Qubit 2.0 Fluorometer (Invitrogen) and libraries from the same experiment were pooled together in equimolar ratios. Pooled libraries were analyzed with Agilent High Sensitivity DNA chip and submitted with results of the analyzes to Genomics Core Facility at MDC, which performed the sequencing reactions.

3.2.8. Ribosome profiling

The protocol for Ribosome profiling was adapted from Ingolia et al., 2009.

Cell harvesting and lysis

Flp-In T-REx HEK293 cells grown on 10cm plates, transfected with control siPool and with DDX3 targeting siPool were washed with 5 ml of PBS containing 100 µg/ml cyclohexamide, flash frozen on liquid nitrogen, kept on dry ice, lysed in 150 µl of Mammalian Polysome Lysis buffer, transferred to wet ice and harvested to RNase-free tubes. Next, with syringes lysates were squeezed up and down through 26G needles 20 times and cleared by centrifugation at 20 000 xg for 5 min in 4ºC. Supernatants were transferred to new tubes, aliquoted and stored at -80ºC. Levels of DDX3 protein in the lysates were checked with Western analysis before continuing with the experiment.

mRNA library preparation for normalization of ribosome profiling data

To determine the abundance of mRNA in ribosome profiling, necessary for normalization of data and its proper analysis, an aliquot of a lysate per sample was subjected to RNA extraction with TRIzol LS reagent (section 3.2.3.). Extracted RNA was used for mRNA library preparation using TruSeq Stranded mRNA Library Prep Kit (Illumina) according to the manufacturer's protocol.

RNA digestion, monosome purification and RNA extraction

Thawed lysates were treated with RNase I at a final concentration of 2.5 U/µl for 45 min at room temperature with slow agitation. Further RNase activity was stopped by addition of 8 µl SUPERase In (20 U/µl) per 240 µl of lysate. Next, Illustra MicroSpin Columns S-400 HR were equilibrated with 3 ml of Mammalian Polysome Buffer, centrifuged 600 xg for 4 min at 4ºC, moved to RNase-free tubes and used to enrich for ribosome complexes. After the lysates entered the columns, they were centrifuged 600 xg for 2 min at 4ºC. To extract RNA from the flow through, samples were mixed with 3 volumes of TRIzol LS reagent, incubated for 10 min at room temperature, then mixed with 1/5 sample volume of chloroform, incubated again at room temperature for 3 min and centrifuged 15 000 xg for 15 min at 4ºC. RNA from aqueous phase of samples was purified and concentrated with Zymo-Spin IIC columns. Purified nucleic acids were eluted from the columns with 30 µl of MiliQ water and their concentration was measured with Nanodrop.

rRNA depletion and gel purification of ribosome footprints

Five µg of RNA per sample were used for rRNA depletion with Ribo-Zero Gold Kit (section 3.2.6.3.). Next, samples and two size markers 27 and 30 nt long were mixed with Formamide Loading Buffer, incubated at 95 ºC for 30 sec and separated in 17% Urea gels runned at 30W for 90 min using Owl Vertical Electrophoresis Systems (Thermo Fisher Scientific). The gel was stained in 120 ml of 1:10 000 SYBR Gold solution in 1x TBE buffer for 40 minutes with slow agitation and nucleic acids were observed with ImageQuant LAS-4000. Slices corresponding to ribosome footprints in each sample (between 27 and 30 nt), as well as 27 and 30 nt were excised from gel. Gel slices were eluted by incubation with 350 µl of 0.3M NaCl solution per slice overnight at 4ºC with shaking at 1 200 RPM. Maximum possible amount of liquid was transferred from tubes containing gel slices to new tubes and precipitated by addition of 1 µl Glycoblue and 3 volumes of 100% EtOH. Samples were centrifuged at 20 000 xg for 1 h in 4ºC to pellet nucleic acids. Supernatants were removed, pellets were air dried and resuspended in 19 µl of MiliQ water.

PNK treatment and purification

To resuspended nucleic acids 5 µl of 10x PNK Buffer, 0.5 µl 0.1 M DTT, 0.5 µl 100 mM ATP, 0.5 µl 10% Triton X-100, 2 µl PNK and 22.5 MiliQ water was added and samples were incubated 37ºC for 30 min. Next, 150 µl of MiliQ water, and 200 µl of phenol/chloroform/isoamylalcohol was added to the samples, which were then vortexed and

centrifuged at 13 000 xg for 2 min at 4ºC. Aqueous phase was moved from samples to new tubes, mixed with 20 µl of 3M NaCl and 1 µl of Glycoblue and precipitated with 3 volumes of 100% EtOH. Nucleic acids were pelleted by centrifugation at 20 000 xg for 1 h at 4ºC. Supernatants were removed, pellets were air dried and resuspended in 6 µl of MiliQ water.

3' adapter ligation and gel purification of adapter-ligated fragments

Resuspended 27 and 30 nt size markers were mixed together and split in two tubes. To all samples 2 µl 10x RNA ligase buffer, 10 µl 50% PEG-8000 and 1 µl 50 µM pre-adenylated 3' adapter were added. Samples were mixed by pipetting, denatured at 95ºC for 30 sec and kept on ice for 5 min. Next, to all samples, except for one of the tubes containing combined size markers (used as unligated control) 1 µl T4 RNA ligase 2, K227Q truncated was added. All samples were incubated for 16 h at 16ºC with slow agitation. Samples were mixed with 20 µl 2x Formamide Loading Buffer and denatured at 95ºC for 30 sec and RNA was separated in 15% Urea gel run at 30W for 90 min using Owl Vertical Electrophoresis Systems (Thermo Fisher Scientific). The gel was stained in 120 ml of 1:10 000 SYBR Gold solution in 1x TBE buffer for 40 minutes with slow agitation and nucleic acids were observed with ImageQuant LAS-4000. Slices corresponding to RNA fragments ligated to 3' adapters were excised from the gel. Gel slices were eluted by incubation with 350 µl of 0.3M NaCl solution per slice over night at 4ºC with shaking at 1 200 RPM. Maximum possible amount of liquid was transferred from tubes containing gel slices to new tubes and precipitated by addition of 1 µl Glycoblue and 3 volumes of 100% EtOH. Samples were centrifuged at 20 000 xg for 1 h in 4ºC to pellet nucleic acids. Supernatants were removed, pellets were air dried.

5' adapter ligation and gel purification of adapter-ligated fragments

Nucleic acid pellets were resuspended directly in 5 µl MiliQ water, 2 µl 10x RNA ligase buffer, 2 µl DMSO, 2 µl 10 nM ATP, 0.5 µl 100 µM 5' adapter. Next, 7.5 µl of 50% PEG-8000 was added and samples were mixed by pipetting. Samples were denatured at 95ºC for 30 sec and incubated on ice for 5 min. Samples containing size markers were split in two to serve as ligated and unligated size controls. 1 µl T4 RNA ligase 1 was added to all samples except for the unligated size control and all were incubated in thermos block for 2 h at 37ºC, shaking at 500 RPM. Samples were mixed with 1 volume of 2x Formamide Loading Buffer and denatured at 95ºC for 30 sec and RNA was separated in 15% Urea gel run at 30W for 90 min using Owl Vertical Electrophoresis Systems (Thermo Fisher Scientific). Gel was stained in 120 ml of 1:10 000 SYBR Gold solution in 1x TBE buffer for 40 minutes with slow agitation and nucleic acids were observed with ImageQuant LAS-4000. Slices corresponding

to RNA fragments ligated to 3' adapters were excised from gel. Gel slices were eluted by incubation with 350 µl of 0.3M NaCl solution per slice over night at 4ºC with shaking at 1.200 RPM. Maximum possible amount of liquid was moved from samples containing gel slices to new tubes and precipitated by addition of 1 µl Glycoblue and 3 volumes of 100% EtOH. Samples were centrifuged at 20 000 xg for 1 h in 4ºC to pellet nucleic acids. Supernatants were removed, pellets were air dried.

Reverse Transcription and PCR amplification

RNA pellets were used in a reverse transcription reaction (section 3.2.4.), with all ingredients of the reaction scaled down to final reaction volume of 15 µl. Resulting cDNA samples were diluted to a final volume of 80 µl, with RNase-free water. For PCR amplification 30 µl of resulting cDNA sample was mixed with 30 µl 5x Phusion Buffer, 3.75 µl 10 mM dNTPs, 0.75 µl 100 µM RPI1 primer, 0.75 µl 100 µM RPI reverse primer (with barcode), 1.5 µl Phusion DNA Polymerase and 83.25 µl RNase-free water. Mixtures were subjected to a PCR reaction with a following profile:

98	40 s	
98	10 s	
65	30 s	Repeat 15 cycles
72	15 s	
72	5 min	

PCR were precipitated with 450 µl ethanol at -20ºC and DNA was pelleted by centrifugation for 30 min at 13 000 g, 4ºC. DNA pellets were dried and resuspended in 25 µl of RNse-free water. 5 µl 6x DNA Loading Dye was mixed with each sample and nucleic acids were separated on 2.5% Low Melting agarose gel, with SYBR gold. DNA bands were visualized using a UV-transilluminator and PCR amplicons of the right size were excised and purified using Zymo Gel Purification Kit (Zymo Research).

3.2.9. Analysis of High-Throughput sequencing results

Sequencing results obtained from MCF-7 cells were converted into fastq files by N. Kastelic or M. Milek using cassava software at the host institute. Data was demultiplexed using Flexbar software (Dodt et al., 2012) and mapped to genome annotation with TopHat (Trapnell et al., 2012). Analysis of differential gene expression was performed using

SeqMonk (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/) and DESeq2 (Love et al., 2014) software.

Analysis of sequencing results for differential expression of circRNAs was performed by Petar Glazar from Rajewsky laboratory at the host institution.

Analysis of ribosome profiling sequencing results (ribosome footprints and mRNA abundance) was performed by Lorenzo Calviello from University of California, San Francisco (UCSF).

Analysis of sequencing results of 4-sU labeled fractions of RNA (section 3.2.7.5.) was performed by Vedran Franke from the laboratory of Altuna Akalin, using an improved version of the antisense transcript detection algorithm introduced in Wyler et al., 2017.

3.2.9. Molecular cloning

Sequence of Csy4 target stem loop (Lee et al., 2013) was ordered as two partially complementary oligonucleotides (section 2.3), which were hybridized and subjected to reaction with High Fidelity Phusion polymerase (Thermo Fisher Scientific) to obtain a double stranded product and was cloned to appropriate plasmids with standard restriction enzyme based cloning procedure.

HULC transcript was amplified from HEK293 derived cDNA with PCR using appropriate primers (section 2.3) and cloned to pcDNA5/FRT/TO plasmid background using restriction enzyme cloning. Resulting plasmid was used for further cloning with restriction enzymes to insert Csy4 target stem loop sequence to the 3' end HULC sequence.

Human codon optimized Csy4 mutant protein DNA sequence was cloned using restriction enzymes from pEX-Csy4 plasmid to pRNTR4 plasmid background. Resulting plasmid was used for further cloning with restriction enzymes to insert Csy4 target stem loop sequence to the 5' UTR of Csy4 sequence between the STOP codon and polyadenylation signal. Additional Csy4 target stem loop sequences were added to resulting plasmid with restriction enzyme cloning resulting in versions containing 1, 2 and 4 stem loop sequences. Obtained plasmids containing various number of Csy4 target stem loops were cloned to pFRT/TO/DEST/FLAG/HA and pFRT/TO/DEST/STREP/HA plasmid backgrounds with Gateway LR Clonase (Thermo Fisher Scientific).

DDX3 homology arms were cloned using restriction enzymes from pEX-DDX3HA to pHR110-mAIDHA-RFP and pHR510-mAIDHA-RFP-T2A plasmid backgrounds.

Sequences of guide RNAs targeting to genomic sequence in close proximity to DDX3 stop codon were ordered as pairs of partially complementary oligonucleotides (section 2.3), hybridized and cloned into px458 plasmid background.

3.2.10. Rapid Amplification of cDNA Ends

3’ RACE was performed with 3' RACE System for Rapid Amplification of cDNA Ends kit (Thermo Fisher Scientific), using RNA isolated with TRIzol Reagent (section. 3.2.3.) from MCF-7 cells 16h after exposure to 10 Gy of IR.

Synthesis of First Strand cDNA with 3’ adaptor sequence

5 µg of total RNA was mixed with RNase-free water to a final volume of 11 µl, 1 µl of 10 µM Adapter Primer solution was added, mixtures were incubated at 70ºC for 10 min and kept on ice for 1 min. Samples were centrifuged briefly and mixed with 2 µl 10X PCR buffer, 2 µl 25 mM MgCl2, 1 µl 10 mM dNTP mix and 2 µl 0.1 M DTT. Samples were briefly centrifuged again, mixed with 1 µl SuperScript II room temperature and incubated at 42ºC for 50 min. To terminate the reaction samples were heated to 70ºC for 15 min and cooled on ice. Samples were mixed with 1 µl of RNase H and incubated at 37ºC for 20 min.

Amplification of the target cDNA

To amplify the 3’ ends of transcript of interest 2 µl of cDNA was mixed with 5 µl 10X PCR buffer, 3 µl 25 mM MgCl2, 1 µl 10 mM dNTP mix, 1 µl of appropriate gene specific primer (section 2.3), 1 µl Universal Amplification Primer 0.5 µl Taq DNA polymerase and 36.5 µl RNase-free water. PCR reaction was performed with a standard program for Taq polymerase. with annealing temperature 60 ºC and extension time 2 min. PCR products were size separated with agarose gel electrophoresis, using ethidium bromide and GeneRuler 1kb DNA ladder. Observed amplicons were excised from the gel, purified with QIAquick Gel Extraction Kit (Qiagen) and cloned to pJET1.2 vector background, with CloneJET PCR Cloning Kit (Thermo Fisher Scientific). Plasmids obtained in the cloning procedure were used for bacterial transformation, bacteria were grown on agar plates with appropriate antibiotic selection and used in downstream analysis.

Identification of amplified 3’ cDNA ends.

Plasmids from three bacterial colonies per each cloned 3’ cDNA end were isolated with QIAprep Spin Miniprep Kit (Qiagen) and send for separate sanger sequencing reactions with pJET1.2 forward and reverse sequencing primers. Obtained sequences were compared with

known sequences of lncRNAs of interest available at https://lncipedia.org/, sequences of matching isoforms were determined and used in follow up experiments.

4. Results

4.1. Development of new RNA pull-down methodology

Quantitative determination of *in vivo* RNP composition is challenging. This is largely due to low sensitivity and specificity of existing methods to detect proteins interacting with a single RNA transcript of moderate to low expression level. We have addressed the need for a method which allows quantitative determination of *in-vivo* RNP composition combined with ease of use. Therefore, we based the development of new biochemical method on the Csy4 protein to specifically isolate a single RNA transcript from complex mixtures such as cell lysates. Csy4 is an endoribonuclease from *Pseudomonas aeruginosa* involved in CRISPR RNA processing. In bacteria it binds to a specific 16-nt long stem loop structure and cleaves the RNA at the base of the stem loop. The protein has been shown to have markedly higher affinity than that of molecules previously used for in-vivo RNA pull-down approaches (Stockley et la., 1995; Sternberg et al., 2012; Leppek and Stoecklin 2013). We have reasoned that this feature of Csy4 makes it a good candidate for design of RNA pull-down methodology that may outperform existing approaches. The utilization of Csy4 for RNA pull-down is possible thanks to development of a catalytically inactive version of this protein, which allows stable binding to the target stem loop without subsequent cleavage. The catalytic activity of the protein can be rescued in the presence of imidazole (Lee et al., 2013). This version of Csy4 has been successfully used for development of an in-vitro RNA pull down method (Fig. 1A). In order to attempt combining high efficiency with ease of use as well as to capture in-vivo RNA-protein interactions, we developed an experimental setup utilizing recombinant Csy4.

We have developed two different approaches of Csy4 mediated in-vivo RNA pull-down. In the first one the transcript of interest is tagged with Csy4 target stem loop and expressed in cells. Interactions between the transcript and its associated proteins take place in living cells and can be preserved by the formation of covalent protein-RNA cross-links facilitated by UV irradiation. Cells are lysed and target transcript is enriched for with in-vitro purified, tagged Csy4 conditional enzyme (Fig. 1B). In the second approach both the stem loop tagged transcript of interest and tagged Csy4 mutant protein are expressed in cells. RNA-protein interactions are preserved by cross-linking, cells are lysed and Csy4-hairpin-RNP complexes are isolated (Fig. 1C).

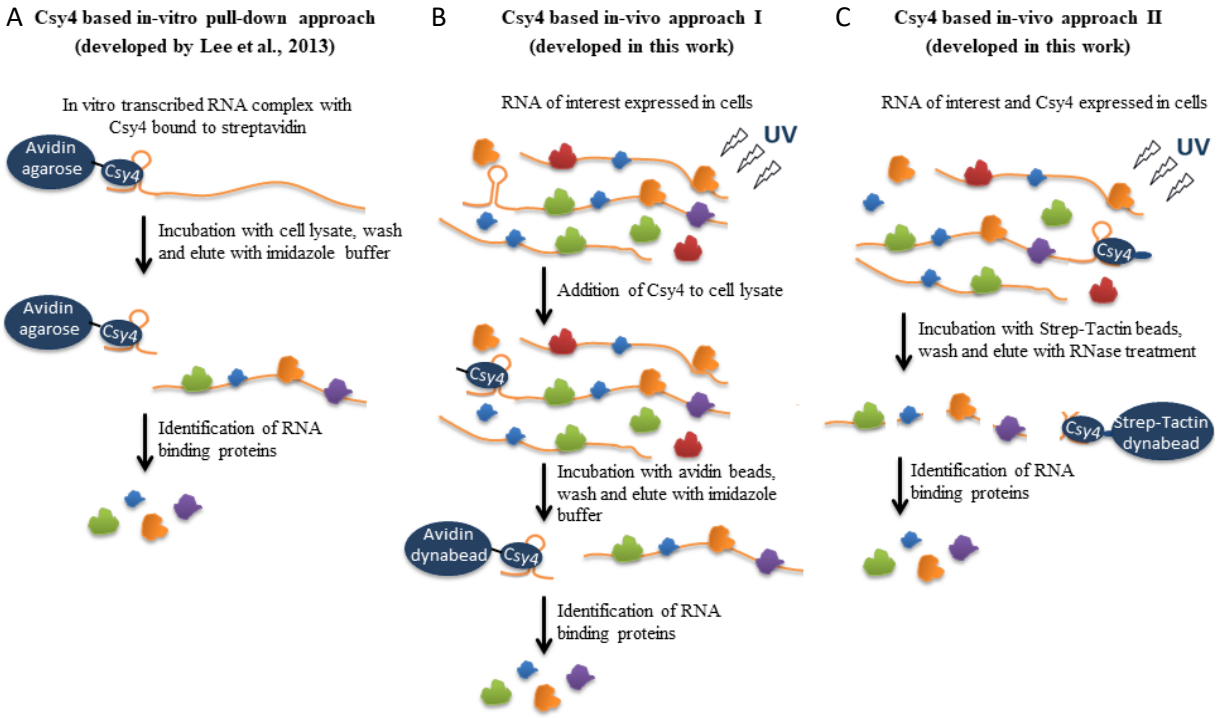


Figure 1. Graphic representation of Csy4 based pull-down strategies. A – Scheme of in-vitro Csy4-based RNA affinity purification. Avidin agarose, biotinylated Csy4, and Csy4 hairpin-tagged transcript are preincubated to make a ternary complex, which is then incubated with uncross-linked cell extract. After washing, the RNA–protein complexes are eluted by cleaving the Csy4 hairpin tag and releasing RNA–protein complexes upon the addition of imidazole. Developed by Lee et al., 2013. B – Scheme of experimental in-vivo Csy4-based RNA affinity purification developed in this work. Cells expressing Csy4 hairpin-tagged transcript are UV crosslinked. Biotinylated Csy4 protein and avidin dynabeads are first incubated with cell extract. After washing, the RNA–protein complexes are eluted by cleaving off the Csy4 hairpin tag and releasing RNA–protein complexes on the addition of imidazole. C – Scheme of second experimental *in-vivo* Csy4-based RNA affinity purification approach developed in this work. Cells expressing STREP-tagged Csy4 protein and hairpin-tagged transcript are UV crosslinked. Strep-Tactin dynabeads are incubated with cell extract. After washing RNA bound proteins are released by RNase treatment.

4.1.1. RNA pull-down with in-vitro purified Csy4 protein

In our first attempt we decided to express a transcript of interest fused with Csy4 target hairpin at its 5' end in HEK293 cells and use a recombinant catalytically inactive version of Csy4 protein to isolate the Csy4-hairpin-RNA from cellular lysates. Our goal was

to be able to isolate endogenous transcripts that could be tagged with the stem loop sequence through genetic engineering. For the proof of concept experiments, however, we have relied on overexpressing the target transcript in cells. This way we could address the usefulness of the method more quickly and easily. The transcript we have chosen for this purpose was lncRNA called Hepatocellular Carcinoma Up-Regulated Long Non-Coding RNA (HULC). HULC is around 500 nt long RNA with known protein interactors (Hammerle et al., 2013). HULC protein interactors are normally expressed in HEK293 cells, which should allow to validate efficiency of pull-down of target transcript associated proteins. It has also been shown that overexpression of this lncRNA in human cells does not negatively influence cell viability (Yan et al., 2017).

4.1.1.1. Enrichment for target transcript with *in-vitro* purified Csy4

First, we confirmed overexpression of tagged HULC transcript in transfected HEK293 cells with qRT-PCR (Fig. 2A). Next, we incubated biotinylated Csy4 mutant (provided by Meister laboratory, University of Regensburg) with cellular lysates from UV cross-linked cells expressing HULC transcript tagged with Csy4 target stem loop. To assay if Csy4 is capable of binding to its target RNA in cellular lysates we compared the recovery of target transcript after incubation with or without Csy4 followed by enrichment by avidin beads. Beads from positive and negative control samples were treated with proteinase K, followed by TRIzol extraction and qRT-PCR detection of the recovered target transcript. We successfully confirmed the recovery of HULC RNA and determined its 200 fold enrichment in the positive control when compared to the negative (Fig. 2B). Next, we sought to determine if the reactivation of Csy4 catalytic activity with imidazole to release the target transcript from the beads is possible. To this end we have performed two rounds of elution. First, Csy4-RNP complexes captured on beads were incubated with buffer containing imidazole. The second round of elution on the same beads was performed in the presence of proteinase K. Both eluates were analyzed with qRT-PCR. The analysis has shown that around 70% of input HULC was successfully eluted with imidazole treatment (Fig. 2C). The elution of HULC with imidazole was satisfactory. We decided to use it in further experiments as it has been shown to provide more specific release of target transcripts than traditional elution (Lee et al., 2013).

4.1.1.2. Specificity of RNA pull-down with in-vitro purified Csy4

Next, we sought to assay the specificity of the pull-down approach. We have incubated Csy4 with cellular lysates, including cells not expressing the target transcript as a negative control. RNP complexes were enriched for and released from beads by activating Csy4 catalytic activity with imidazole. We have split eluates in two and prepared one half for analysis of pulled-down RNA and one for analysis of pulled-down proteins. We assayed recovered levels of HULC target transcript relative to GAPDH mRNA with qRT-PCR. We chose GAPDH mRNA for this analysis, since it was used before to assay specificity of a similar RNA pull-down methodology (Yoon et al., 2012; Yoon and gorospe 2016). Although HULC was clearly enriched when compared to levels of GAPDH, the mRNA was still clearly detectable with qRT-PCR. Levels of HULC transcript relative to levels of GAPDH mRNA in pull-down samples were on average around 8 times higher than its levels in input (Fig. 2D). The observed presence of target unrelated transcript may signify a problem with specificity in pull-down approach.

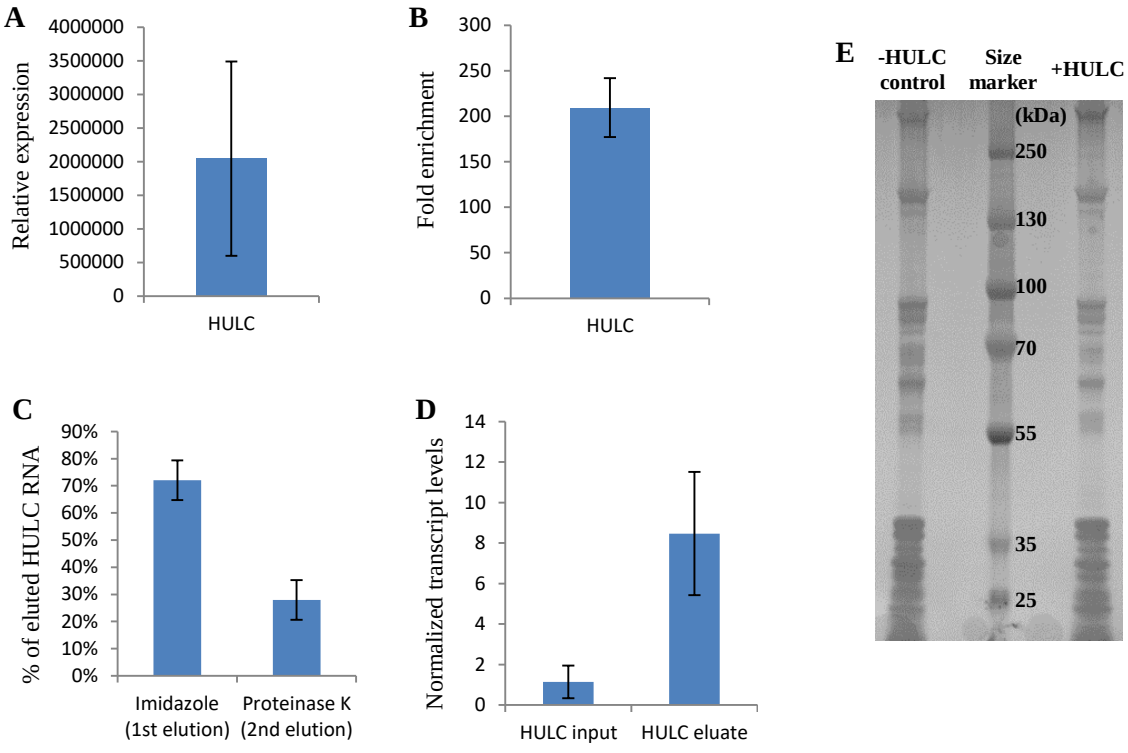


Figure 2. Efficiency and specificity of RNA pull-down with in-vitro purified Csy4 mutant protein. A – Levels of HULC in cells transfected with a plasmid expressing HULC tagged with Csy4 target stem loop, relative to HULC levels in cells transfected with plasmid expressing free GFP, normalized to GAPDH. Average of two biological replicates with error

bars representing standard deviation is shown. B – Fold enrichment of tagged target transcript HULC pulled-down with Csy4 from transfected HEK293 cells, calculated over negative control with streptavidin beads without Csy4 protein, measured with qRT-PCR. Average of two biological replicates with error bars representing standard deviation is shown. C – Fractions of HULC target transcript successfully released from the beads with imidazole treatment and subsequent proteinase K elution, measured with qRT-PCR (100% = sum of the HULC transcript released with imidazole and with proteinase K treatment). Averages of two biological replicates with error bars representing standard deviation are shown. D – HULC transcript levels in input and Csy4 pull-down sample, normalized to levels of GAPDH, signifying the increase of HULC to GAPDH ratios after pull-down, measured with qRT-PCR. Averages of two biological replicates with error bars representing standard deviation are shown. E – Silver Stained SDS-Page gel picture showing proteins recovered in Csy4 pull-down experiment. Proteins recovered in negative control where pull-down was performed on lysates from cells not expressing the tagged HULC target transcript and proteins recovered from cells expressing tagged HULC are shown. A representative example out of three biological replicates performed.

To assay the possibility of detecting HULC-interacting proteins, we decided to check the protein composition of the eluate. Eluates from lysates that contained or lacked overexpressed HULC were separated by SDS-PAGE and visualized with silver staining. The experiment showed presence of wide range of proteins with different molecular sizes in pull-down samples. This, pattern was similar to the negative control sample (Fig. 2E). The fact that imidazole driven elution from beads works through cleaving off Csy4-bound transcripts suggests that the high experimental background may be caused by unspecific Csy4 binding to non-target RNA. We have decided to redesign our pull-down approach to try to increase its specificity.

4.1.2. RNA pull-down with Csy4 protein expressed in human cells

We have attempted to pull-down Csy4 target transcript from cells expressing both the tagged transcript of interest and tagged Csy4 mutant protein to allow the binding to take place within the cells. We have developed constructs encoding human codon optimized Csy4 mutant protein fused on its N-terminal end to FLAG/HA- or STREP/HA-tags. Using HEK293 Flp-in cells we tried to develop cell lines stably expressing tagged Csy4 protein under

inducible promoter in which we could tag endogenous transcripts of interest with Csy4 target stem loops using genetic engineering or express tagged transcripts transiently from plasmids. Our attempts however, failed as no cells survived the antibiotic selection.

We next decided to express Csy4 in HEK 293 cells transiently from plasmids for proof of concept experiments. For this purpose, we developed constructs encoding human codon optimized Csy4 mutant protein with N-terminal FLAG/HA- or STREP/HA-tags as well as different number of Csy4 target stem loops in its 3' UTR (Fig. 3A). Although correct sequence of all used constructs was confirmed with sequencing for lack of mutations or sequence rearrangements, before they were used, the expression of FLAG/HA-tagged Csy4 was only detected on low levels according to Western analysis (Supplementary Fig. 1A). Expression of STREP/HA-tagged version of the protein, however, was successfully detected with Western blot with observed molecular weight similar to the predicted one (Supplementary Fig. 1B).

We used UV cross-linked, STREP/HA and target hairpin tagged Csy4 expressing HEK293 cells for the pull-down experiment. We have used three variants of the construct containing 1, 2 and 4 repeats of Csy4 target sequence in the 3' end of the transcript (Fig. 3A). After enrichment for Csy4 with the beads half of them was proteinase K treated to recover isolated transcripts. The second half was treated with RNase and boiled in SDS to recover pulled-down proteins for analysis. The fact that Csy4 protein contacts with its target RNA may be stabilized by cross-linking in this approach demands a different elution method than activation of Csy4 enzymatic activity (Fig. 1C). We have tested the successful isolation of Csy4 from the lysates by Western analysis to make sure that potential low efficiency of the pull-down is not due to inefficient Csy4 capture. Lysates before incubation with Strep-Tactin beads and protein elutes were subjected to Western analysis. We have observed relatively efficient capture of the protein by the beads (Fig. 3B). An estimation of Csy4 protein pull-down efficiency based on the Western blot image analysis with ImageJ software showed that we were able to pull-down around 15-20% of the protein from lysates (Fig. 3C). Next, we have used the transcript eluates to determine the target transcript pull-down efficiency and specificity with qRT-PCR. We have observed that in contrast to the previous approach after enrichment for the target transcript signal from target unrelated RNA GAPDH was detected on very low levels, whereas Csy4 mRNA signal was clearly enriched (Fig. 3D). The fraction of input Csy4 mRNA recovered however, was on levels of a fraction of percent, reaching

around 0.024%, 0.014% and 0.01% for transcripts tagged with 1, 2 and 4 repeats of Csy4 target stem loop respectively (Fig. 3E), which was far from satisfactory.

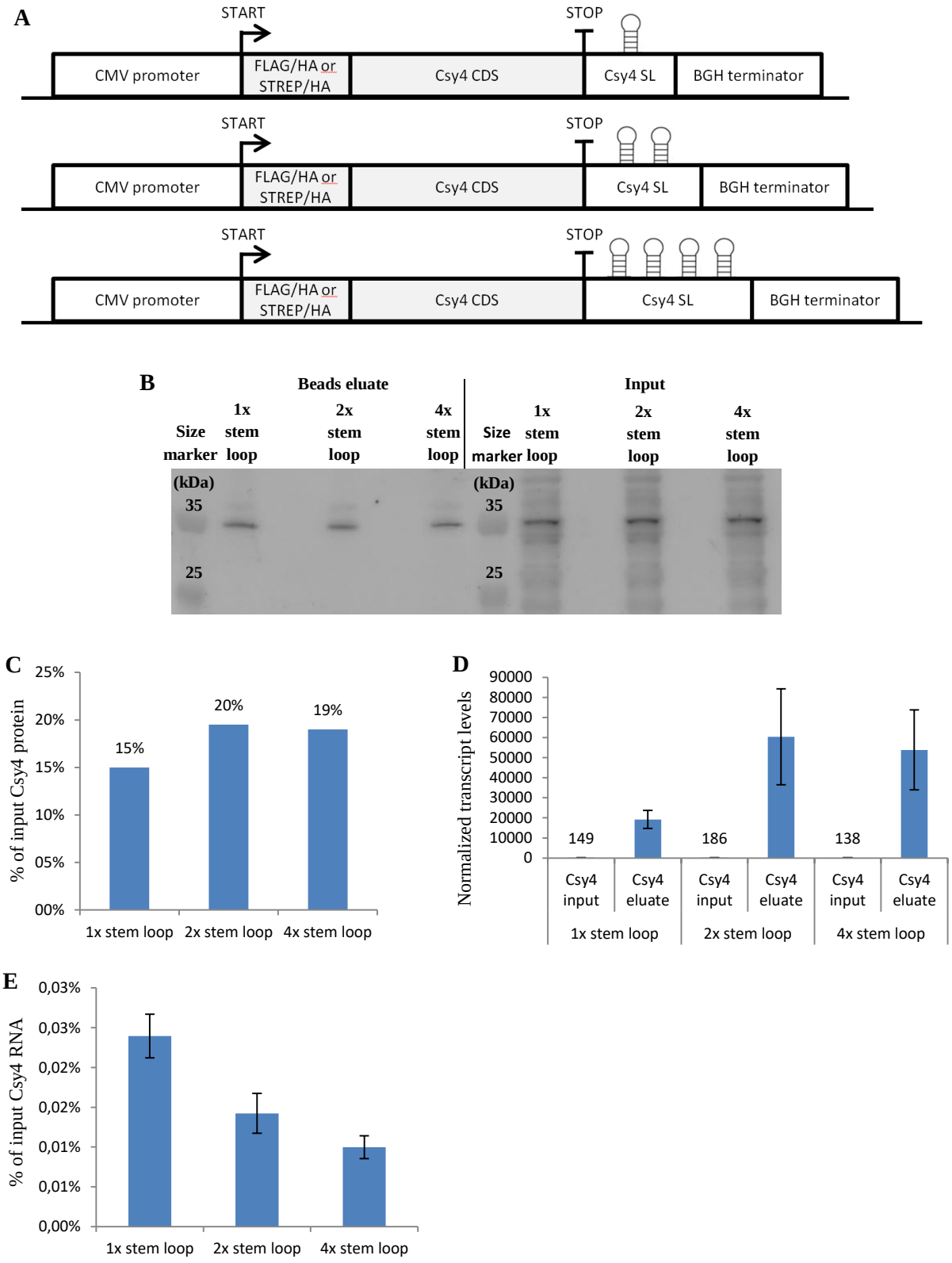


Figure 3. Efficiency of RNA pull-down with Csy4 protein expressed in human HEK293 cells.

A – A scheme representing Csy4 FLAG/HA or STREP/HA tagged expression cassettes containing different number of repeats of Csy4 target stem loops (SL) in 3' UTR in plasmids used for transient expression of the protein in HEK293 cells. B – A representative Western blot image with HA antibodies showing levels of Csy4 protein in input samples and in bead eluates from pull-down experiment performed on cells expressing Csy4 mRNA tagged with 1, 2 and 4 Csy4 target stem loops. 1% of total input volume and 10% of total eluate volume per sample was loaded on gel. C – Amounts of Csy4 protein successfully eluted from beads in pull-down performed on cells expressing Csy4 mRNA tagged with 1, 2 and 4 repeats of Csy4 target stem loop. The amounts of protein are presented as percentage of protein levels detected in input samples, quantified from image shown in figure A using ImageJ software quantification tool. D – Csy4 mRNA levels in input and Csy4 pull-down samples from cells expressing Csy4 mRNA tagged with different number of Csy4 target stem loops, normalized to levels of GAPDH, signifying the increase of Csy4 mRNA to GAPDH ratios after pull-down, measured with qRT-PCR. Results from one biological replicate are shown with error bars representing standard deviation between technical replicates. E – Amounts of Csy4 target transcript successfully pulled down from cells expressing Csy4 mRNA containing 1, 2 and 4 repeats of Csy4 target stem loop, measured with qRT-PCR, values represent amounts of pulled-down Csy4 mRNA as percentages of Csy4 mRNA in input samples. Averages of two biological replicates with error bars representing standard deviation are shown.

4.2. LncRNAs expressed in DNA damage response

In order to identify differentially expressed lncRNAs in DDR, we have performed strand-specific RNA sequencing experiments on cells exposed to genotoxic stress. Human breast adenocarcinoma MCF-7 cells were treated with 10 Gy of ionizing radiation (IR) to induce DSBs, and RNA from cells 4, 8, 16 and 24h after exposure was sequenced. To gain a comprehensive picture of expression of lncRNAs, we have performed separate sequencing of polyadenylated, non-polyadenylated and nascent fractions of transcripts.

4.2.1. Polyadenylated lncRNAs expressed in DNA damage response

Investigation of differentially expressed polyA+ RNAs by sequencing using the Gencode.v19 genome annotation revealed a total of 507 genes with altered expression values

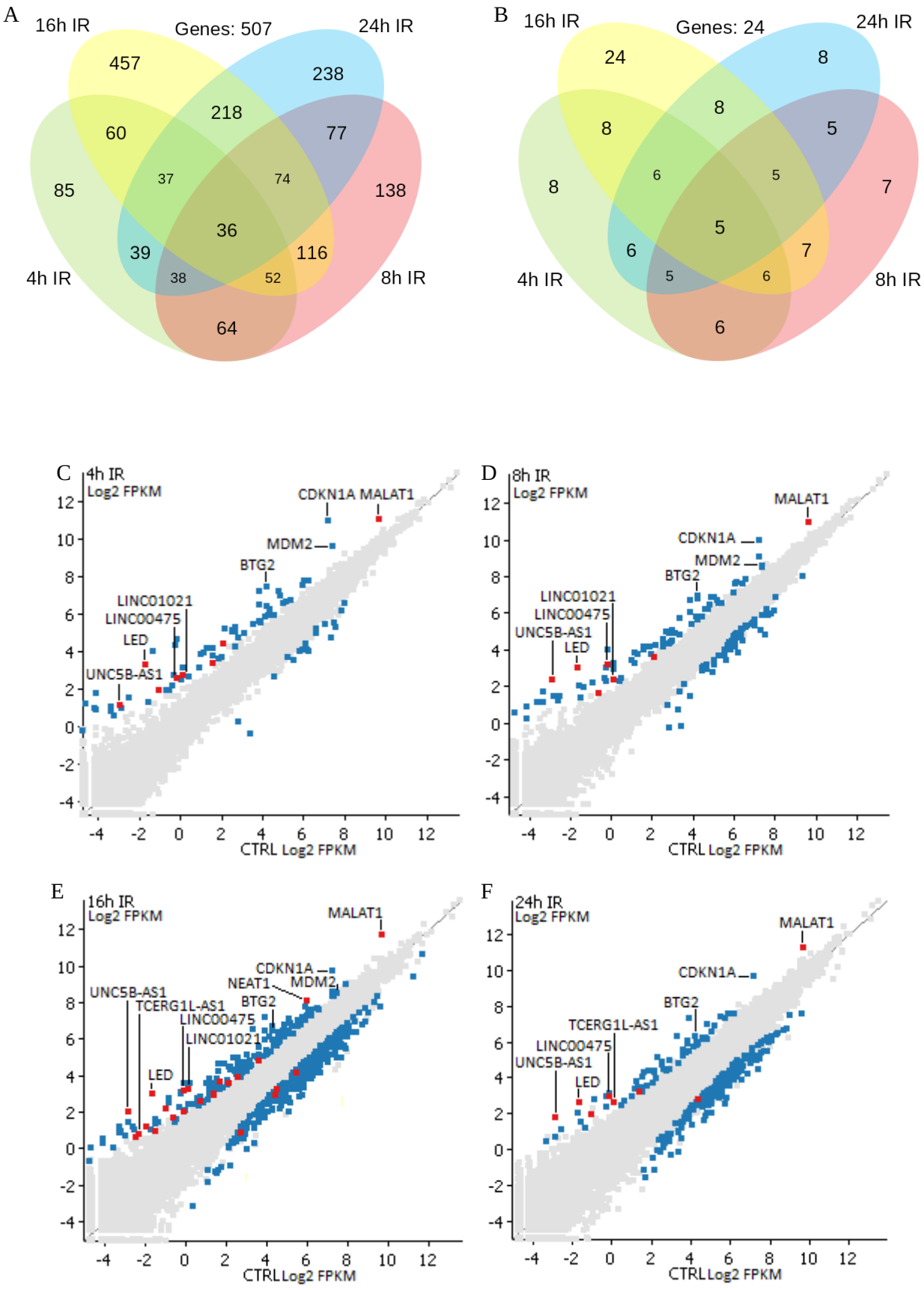


Figure 4. Summary of polyA+ RNA sequencing experiment results. A – Four ellipse Venn diagram showing numbers of genes identified as differentially expressed in cells in different

time points after IR treatment. B - Four ellipse Venn diagram showing numbers of lncRNA genes identified as differentially expressed in cells in different time points after IR treatment. From C to F – Scatter plots showing results of the sequencing experiment. Log2 transformed FPKM expression values of genes in control cells and cells after exposure to IR of two averaged biological replicates are shown. Transcriptomes were mapped to Gencode.v19 genome annotation and analyzed with DESeq2 and SeqMonk software. Genes with no significant expression change are marked in grey, differentially expressed genes are marked with blue for mRNAs and red for lncRNAs. C – cells 4h after IR treatment. D – cells 8h after IR treatment. E – cells 16h after IR treatment. F – cells 24h after IR treatment.

in cells 4, 8, 16 and 24h after induction of DNA damage (Fig. 4A). Among those transcripts were mRNAs known to be typically upregulated in presence of DSB, including CDKN1A, MDM2 and BTG2, as well as known lncRNAs (Fig. 4C – F). In total, 24 different transcripts with no protein coding potential were shown to be differentially expressed, in at least one time point of IR treatment (Fig. 4B). Among them all but three had no known molecular functions in human cells at the time when this work was performed. The three differentially expressed transcript that were previously functionally characterized were MALAT1, NEAT1 and LINC00086. To gain a more comprehensive picture of lncRNAs differentially expressed in DDR, we have repeated the sequencing analysis with a MCF-7 specific lncRNA annotation generated by GRO-seq (Sun et al., 2015). We were able to detect additional 5 lncRNAs differentially expressed in different time points of DDR that were not present in Gencode.v19 genome annotation (Supplementary Table 1). None of those transcripts had their molecular functions known at the time of this work being performed.

4.2.2. Non-polyadenylated lncRNAs expressed in DNA damage response

Before we analyzed the results of sequencing of transcripts depleted of polyadenylated and ribosomal RNAs, we sought to confirm that the depletion was successful. We have compared read coverage observed in depleted and in polyA+ samples. We have observed that relative amount of reads mapping to exons is strongly decreased in depleted samples, while relative mapping to genes is similar with polyA+ samples (Supplementary Fig. 2). This observation represents increased mapping of reads to unspliced pre-mRNAs, which lack a poly-A tail and are typically enriched in samples depleted of polyadenylated RNAs.

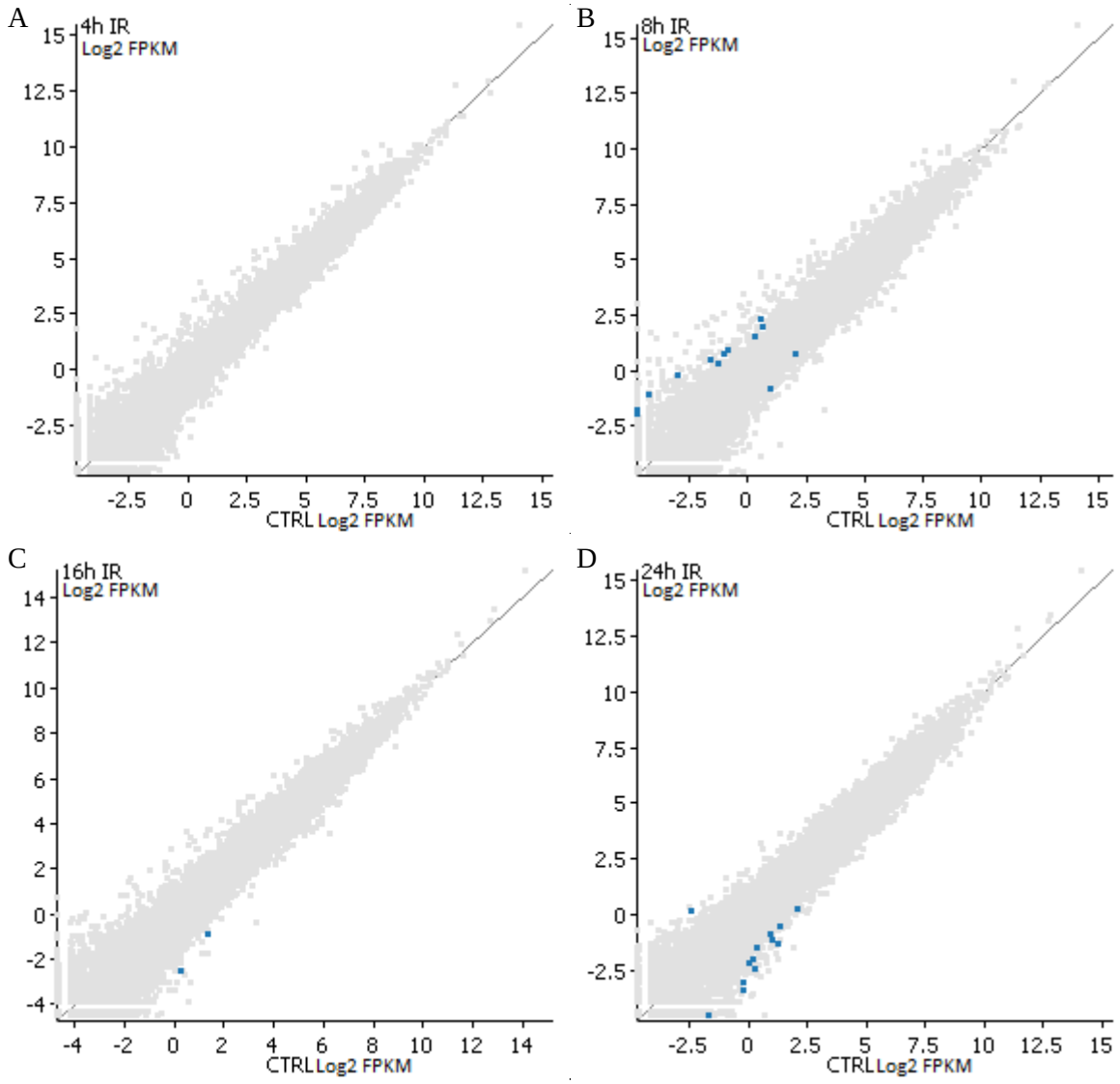


Figure 5. Scatter plots showing results of polyA-, ribo zero RNA sequencing experiments. Log2 transformed FPKM expression values of genes in control cells and cells after exposure to IR of two averaged biological replicates are shown. Transcriptomes were mapped to Gencode.v19 genome annotation and analyzed with Deseq2 and SeqMonk software. Differentially expressed genes that were not detected in polyA+ RNA sequencing are marked with blue. A – cells 4h after IR treatment. B – cells 8h after IR treatment. C – cells 16h after IR treatment. D – cells 24h after IR treatment.

Sequencing of RNAs depleted of polyadenylated and ribosomal RNAs showed upregulation of many of the same transcripts as sequencing of polyadenylated RNAs. This is in line with the fact that most RNAs are upregulated on transcriptional level in DDR and they can be captured before polyadenylation takes place. Analysis of expression of transcripts not detected in sequencing of polyadenylated RNAs revealed additional 24 differentially

expressed transcripts. 0, 13, 2 and 13 such transcripts were detected in 4h, 8h, 16h and 24h after treatment with IR, respectively (Fig. 5). All of those additionally identified transcripts were lncRNAs, with no known molecular functions at the time of this work being performed.

4.2.3. Detection of nascent RNAs in DNA damage response

To gain a more complete picture of gene expression dynamics in DDR we have performed additional RNA sequencing experiments utilizing metabolic 4-thiouridine (4sU) labeling. In this approach, we have sequenced the nascent transcriptome using incorporated 4-sU with 5 min labeling performed before RNA extraction. Experiments were performed on control cells and cells 4, 8, 16 and 24h after IR treatment. Data analysis of the sequencing results was performed by Vedran Franke from the laboratory of Altuna Akalin, using an improved version of the antisense transcript detection algorithm introduced in Wyler et al., 2017. Using this approach, we have identified various novel putative transcription units, which were not present in available gene annotations and were detected specifically in cells after IR treatment. Among them were transcriptional units, that are neither antisense to known genes nor appear to be associated with their promoters (Fig. 6). Those putative transcripts most likely represent novel lincRNAs.

Nascent RNA sequencing provides information on the amount of transcription, not on the levels of mature RNAs, thus it can capture transcripts with little influence of their turn-over rates. To verify the putative novel lncRNAs, we have checked if accumulation of reads mapping to their genomic regions can be also observed in polyA+ and polyA- sequencing experiments. Interestingly, we have found that one of investigated genomic region displayed strong increase in accumulation of mapped reads in nascent, polyA+ and polyA- sequencing experiments (Fig. 7A) and the second one displayed an increase only in nascent and polyA- experiments (Fig. 7B). This observation suggests that the first transcript is a polyadenylated RNA and that the second one is a non-polyadenylated RNA. We verified the expression of the putative polyadenylated transcript in DDR with qRT-PCR, using cDNA synthesized with oligo(dT) primer. The experiment confirmed induction of the putative lncRNA (Fig. 7C). The amplicon obtained in qRT-PCR was subjected to Sanger sequencing and its sequence matched with intended target was confirmed.

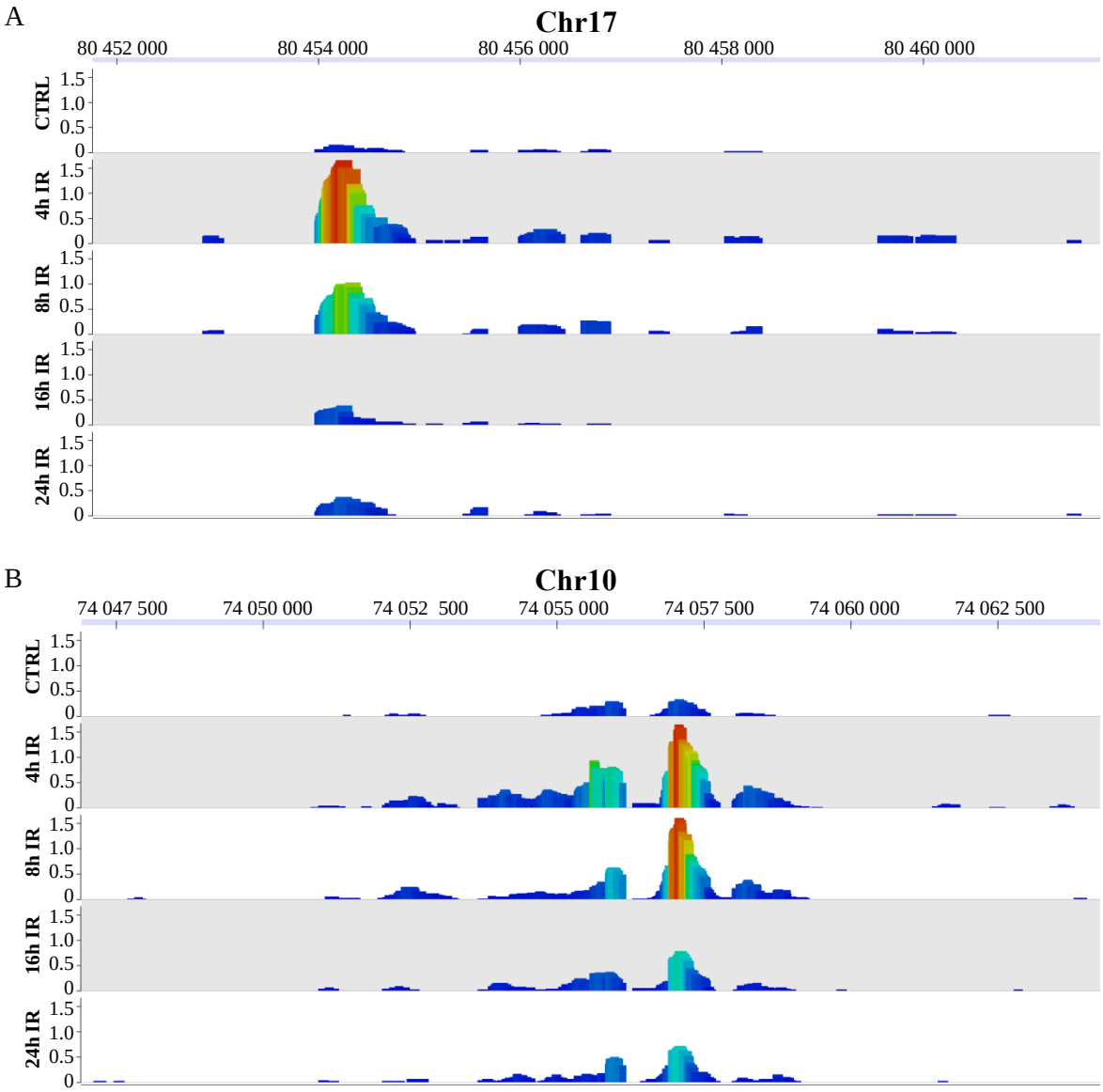


Figure 6. Putative novel lincRNAs detected with nascent RNA sequencing. Coverage profiles of reads (values corrected for number of mapped reads) from two averaged biological replicates mapping to two genomic regions, which were not annotated in available genome annotations are shown. A – region on chromosome 17 (only reads mapping to forward DNA strand are visualized). B – region on chromosome 10 (only reads mapping to reverse DNA strand are visualized).

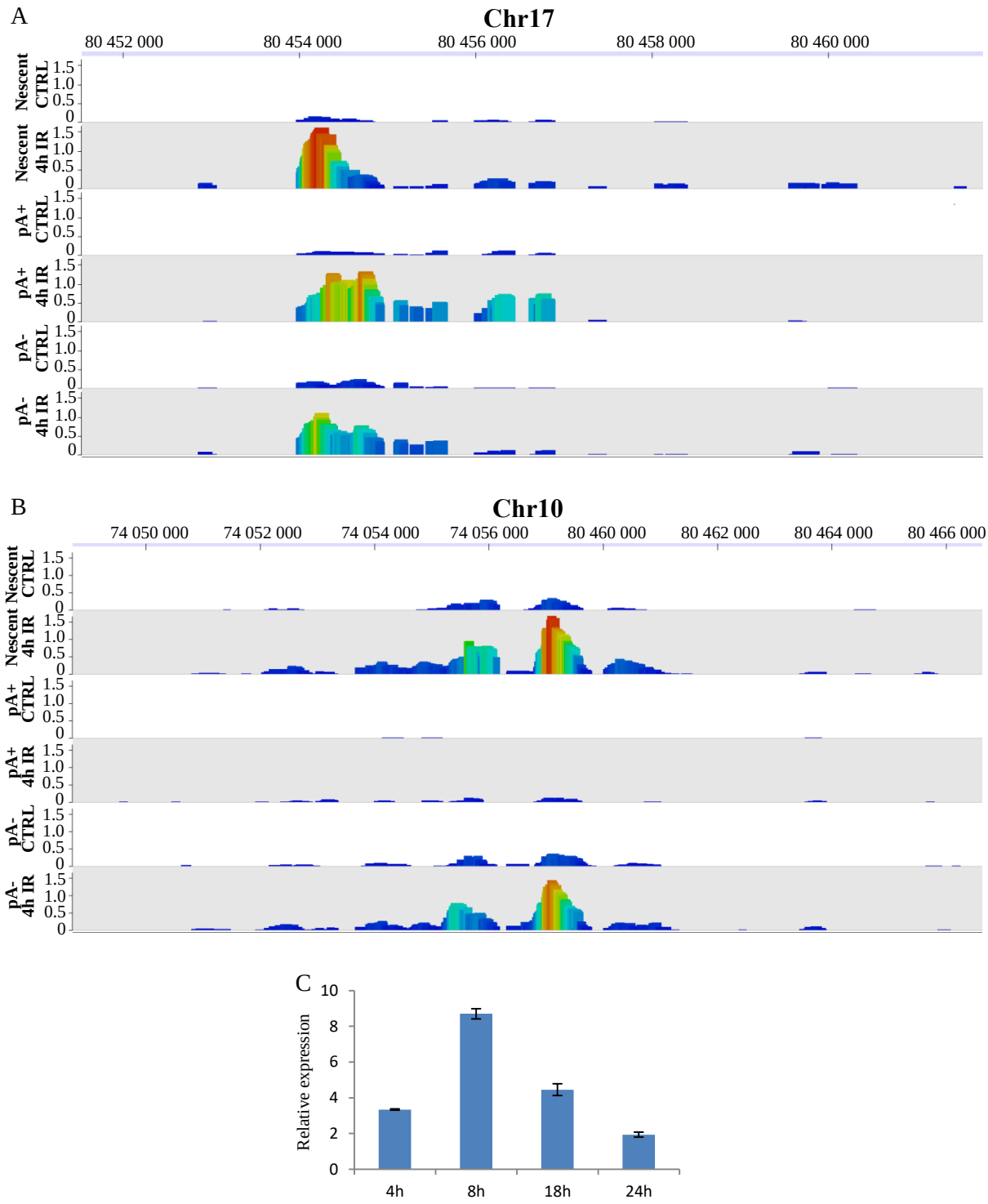


Figure 7. Validation of detection of putative novel lincRNAs. A and B – Coverage profiles of reads (values corrected for number of mapped reads) from two averaged biological replicates of nascent, polyA+ and poyA- RNA sequencing experiments mapping to two genomic regions, which were not annotated in available genome annotations are shown. A – region on chromosome 17 (only reads mapping to forward DNA strand are visualized). B – region on chromosome 10 (only reads mapping to reverse DNA strand are visualized). C – Expression

of putative novel lncRNA originating from chromosome 17 in MCF-7 cells 4, 8, 16 and 24h after exposure to IR, relative to expression in untreated cells, normalized to GAPDH. Averages of two biological replicates with error bars representing standard deviation are shown.

4.2.4. Circular RNAs expressed in DNA damage response

Sequencing results of samples depleted for rRNAs and polyA+ RNAs were also used to determine differential expression of circRNAs. This part of data analysis was performed as a part of collaboration by Peter Glazar from the Nicolaus Rajewsky laboratory. We were able to reproducibly detect 910, 2502, 3124, 1555, 2390 circRNAs in control cells and cells 4h, 8h, 16h and 24h after exposure to IR respectively. To select from detected circRNAs potentially relevant for DDR, we focused on those with log2FC >2 or <2. The analysis revealed hundreds of circRNAs with increased expression values upon DDR. This increase in expression prevalently comes from detection of circRNAs in IR, which were not detected in control cells (Fig. 8A). We did not observe any circRNAs with decreased expression values in response to DNA damage. This observation does not seem to be an artifact of sequencing depth, as we did not obtain more mapped reads in IR treated samples than in controls (Supplementary Fig. 3). The upregulated circRNAs originated almost exclusively from protein coding host genes. The total number of circRNAs detected to be upregulated after IR treatment in at least one time point was 1311. They originated from 939 different genes, with some genes being hosts to multiple circRNAs. Gene ontology analysis of those host genes revealed that protein products of 675 of them are subjects to phosphorylation, a posttranslational modification that is a hallmark of DDR. Functional analysis of genes hosting circRNAs upregulated in response to IR showed that the most prevalent categories among them are Ubl conjugation pathway, protein transport and DNA damage (Fig. 8B). Some functionally characterized circRNAs are known to regulate expression of their protein-coding siblings, thus it is tempting to speculate that detected circRNAs are involved in regulation of the same processes as their protein coding counterparts. To see if any upregulated circRNAs have a regulatory potential towards linear products of their host genes, we looked for mRNA-circRNA pairs that both displayed changed expression in DDR according to polyA+ and polyA- ribo zero sequencing results. All quantified circRNAs originated from host genes that also produced detectable levels of linear

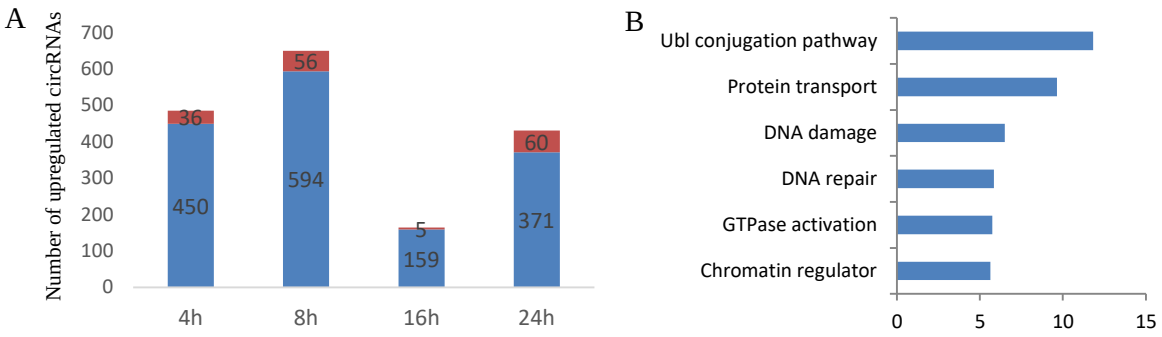


Figure 8. A – Number of circRNAs detected to be upregulated in sequencing of RNA depleted of polyadenylated and ribosomal transcripts in different time points after IR treatment in two biological replicates in MCF-7 cells. Blue – circRNAs detected after IR treatment, not expressed in untreated control cells. Red – upregulated circRNAs that were also expressed in control cells. B – GO enrichment analysis of genes being hosts to circRNAs upregulated in DDR. Top six "biological function" categories shown are ranked by Benjamini-Hochberg adjusted P-value (Benjamini and Hochberg 1995).

Table 9. Summary of information on mRNA-circRNA pairs originating from the same host gene that were differentially expressed in DDR. Average expression changes of two biological replicates are shown.

Host gene name	Expression change of linear and circular transcript detected in DDR time point	Expression change of linear transcript (log2FC)	circRNA ID	Expression change of circular transcript (log2FC)	Gene involved in
PIF1	4h	-3.90	chr15:65115976-65116553	detected only in DDR	DNA damage response
MCM6	8h	-1.40	chr2:136620176-136622733	detected only in DDR	DNA replication
DEPDC1B	8h	-1.37	chr5:59934576-59943360	3.78	Wnt signaling pathway
MCM10	8h	-1.27	chr10:13214375-13228280	detected only in DDR	DNA damage response
GLS2	8h	2.28	chr12:56871443-56872046	detected only in DDR	Glutamine metabolism
UNG	8h	-1.44	chr12:109539706-109541416	detected only in DDR	DNA damage response
DLGAP5	8h	-1.78	chr14:55647930-55650471	detected only in DDR	Cell cycle regulation
XG	8h	2.26	chrX:2700106-2700169	detected only in DDR	Plasma membrane
PIK3R3	16h	1.19	chr1:46521466-46543285	detected only in DDR	Regulation of protein-tyrosine kinases activity
KIF14	16h	-2.09	chr1:200550328-200561368	detected only in DDR	Cytokinesis
MCM6	16h	-1.90	chr2:136620176-136622733	detected only in DDR	DNA replication
KLHL24	16h	1.99	chr3:183361267-183369064	detected only in DDR	Protein ubiquitination
VPS13C	16h	1.34	chr15:62299506-62306191	detected only in DDR	Intracellular trafficking and secretion
POLE2	24h	-2.13	chr14:50131343-50141145	2.17	DNA replication
BRIP1	24h	-1.84	chr17:59853761-59857762	2.00	DNA damage response
WDHD1	24h	-1.67	chr14:55448257-55467710	detected only in DDR	DNA replication
CENPI	24h	-2.06	chrX:100356046-100364987	detected only in DDR	Centromere complex assembly

RNAs. However, we were able to find only a handful of cases in which expression of both mRNA and circRNA were changed at the same time point after exposure to IR (Table 9). Most of genes hosting those pairs had a DDR related function. It is possible that those circRNAs regulate their protein coding siblings *in cis* in response to genotoxic stress.

4.2.5. Validation of expression changes of selected lncRNAs upregulated in DNA damage response

For further study we have selected four lncRNAs displaying highest upregulation in response to IR according to our RNA sequencing experiments, that were not functionally characterized before (Table 10). Two lincRNAs, LINC00475 and LINC01021, and two antisense lncRNAs, TCERG1L-AS1 and UNC5B-AS1. TCERG1L-AS1 sequence is contained within the last intron of Transcription Elongation Regulator 1 Like (TCERG1L) protein coding gene. Expression of the mRNA from this gene was not detected in our sequencing experiments. UNC5B-AS1 sequence is contained within the first intron of Unc-5 Netrin Receptor B (UNC5B) protein coding gene, which is expressed in MCF-7 cells according to our sequencing experiments. UNC5B encodes a netrin receptor localized in plasma membrane. Netrin proteins are regulators of nervous system development, involved in the process of axon guidance (Sun et al., 2011). UNC5B is involved in netrin signaling, but it has an additional function in regulation of apoptosis and has been shown to promote TP53-dependent apoptosis (Tanikawa et al., 2003). We have validated expression changes of selected lncRNAs in DDR with RT-qPCR using primers targeting parts of sequence with high coverage in RNA sequencing experiment (Fig. 9). To assay the potential role of selected lncRNAs in DDR we decided to attempt to knock them down with siRNAs. To be able to design efficient siRNAs, we first attempted to determine specific sequences of transcripts of interest in MCF-7 cells. Transcripts of interest were discovered and annotated in different cell lines and given relatively low pressure towards sequence conservation of lncRNAs, their sequences may not be exactly the same in MCF-7 cell as they are in reference assemblies and annotations. Moreover, some of the selected RNAs have multiple isoforms, being subjected to alternative splicing and polyadenylation. For example, LINC00475 has over 20 annotated alternative polyadenylation sites. Therefore, knowing at least part of a precise sequence of target transcripts may be important for efficient silencing under our experimental conditions. To determine suitable sequences to target, we performed a 3’ RACE on selected transcripts. For this experiment, cDNA derived from MCF7 cells 16h after induction of DNA damage

was used. We were able to determine around 400nt long 3' sequence for each of the transcripts and clarify which isoforms were expressed (Table 10).

Table 10. Information on lncRNAs differentially expressed in DDR selected for further study. Average expression changes from two biological replicates are shown.

Transcript name in Gencode.v19	Other transcript names	Detected in timepoints after IR	Properties	Highest detected log2FC	Molecular functions	Number of annotated polyadenylation sites	Number of polyadenylation sites detected in MCF-7
RP11-790G19.2	UNC5B-AS1	4h, 8h, 16h	Antisense to UNC5B	5.25	not known	2	1
LINC00475	lnc-CENPP	4h, 8h, 16h, 24h	Intergenic	3.30	not known	over 20	1
RP11-462G8.3	TCERG1L-AS1	16h, 24h	Antisense to TCERG1L	3.14	not known	1	1
RP11-46C20.1	LINC01021	4h, 8h, 16h	Intergenic	3.12	not known	12	2

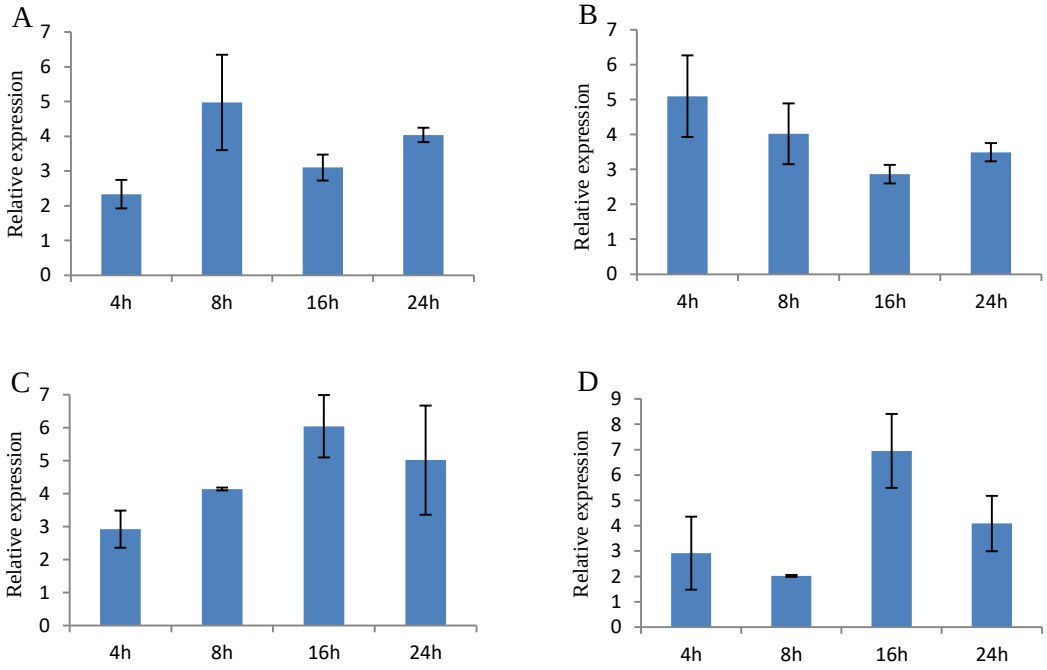


Figure 9. Expression of selected lncRNAs in MCF-7 cells 4, 8, 16 and 24h after exposure to IR, relative to expression in untreated cells, normalized to GAPDH. Averages of two biological replicates with error bars representing standard deviation are shown A – LINC00475. B – LINC01021. C – TCERG1L-AS1. D – UNC5B-AS1.

4.2.6. Assessment of functions of selected lncRNAs in DNA damage response

Two different siRNAs were designed to target each of selected lncRNAs within their 3' sequence determined with RACE. siRNA mediated knock-down efficiency was measured

on cells 16h after IR treatment using RT-qPCR. Among all tested siRNAs, at least one for each target transcript allowed successful knock-down of around 50%, except for siRNAs targeting LINC01021, for which we were able to lower its target expression by only around 13% (Fig. 10A). The knock-down efficiency of around 50% may not be sufficient to disrupt functions of the target transcript. We still decided to test the effects the lncRNAs may have on the cells' ability to cope with DDR. For this purpose, we performed colony formation assays upon RNAi treatment, an experiment commonly used to measure rates of cells survival after induction of genotoxic stress. If lncRNAs of interest have a regulatory role in DDR, depleting them should affect the cells' ability to cope with IR. We did not see any significant difference in the ability to form colonies between transfected and control cells after IR treatment (Fig. 10B). This observation may reflect a lack of relevant functions of the tested lncRNAs, or an insufficient knock-down achieved with siRNAs, that did not disrupt their function, or that the transcription of the lncRNA is important and not the transcript itself.

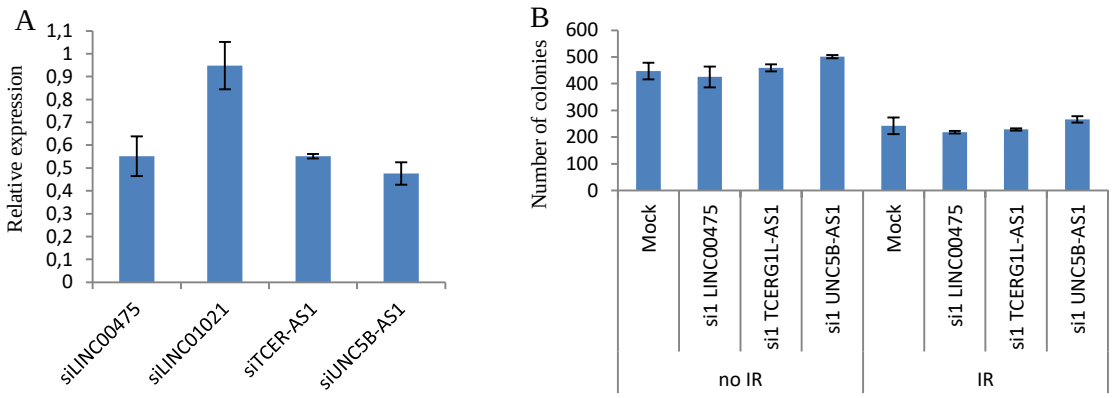


Figure 10. Knock-down efficiency of cells transfected with siRNAs targeting transcripts of interest and colony formation assay on transfected cells after transfection. A –expression values of lncRNAs of interest 48h after transfection with siRNAs and 16h after IR treatment relative to expression in mock treated cells 16h after IR treatment, normalized to expression of GAPDH. Averages of two biological replicates with error bars representing standard deviation are shown. B – average numbers of colonies formed by siRNA and mock transfected cells without and with IR treatment from two biological replicates, with error bars representing standard deviations are shown.

4.3. Regulation of human transcriptome by DDX3

A common feature of lncRNAs is that they are both regulated by proteins and often carry out their function in concert with RNA binding proteins (RBPs). Regulatory roles of RBPs in DDR are increasingly appreciated. The RNA helicase DDX3 has been reported to be one of the most highly enriched protein in the polyadenylated RNA-associated proteome in response to IR, which suggests that it may be another RNA binding protein involved in regulation of cellular responses to genotoxic stress (Milek et al., 2017). The protein is mostly studied in regard to its control of protein coding transcripts. It has been reported to control translation initiation on mRNAs with structured 5' UTRs (Lai et al., 2008; Oh et al., 2016; Soto-Rifo et al., 2012). DDX3X is an essential gene in eukaryotes (Sharma and Jankowsky 2014) and mutations are linked to diseases, including cancer (Epling et al., 2015; Jones et al., 2012; Kool et al., 2014; Oh et al., 2016; Pugh et al., 2012; Robinson et al., 2012; Valentin-Vega et al., 2016). The protein has also been reported to interact directly with lncRNAs, although to much lower extents than with mRNAs (Oh et al., 2016) and to be a part of larger lncRNA interacting protein complexes (Ribeiro et al., 2018). The functions that it plays through its association with lncRNAs are unknown and did not receive much attention. We have decided to study in more detail the nature of DDX3 interactions with RNAs and the significance they have for cells biology.

4.3.1. DDX3 interacts primarily with protein coding transcripts

We have performed a PAR-CLIP experiment on cells overexpressing FLAG-tagged DDX3X protein in HEK293 Flip-in cells (Fig. 11A). The experiment was performed by Emanuel Wyler in the host laboratory and data analysis was performed by Lorenzo Calviello from University of California, San Francisco (UCSF). Before this work, iCLIP experiments on DDX3X protein were published, reporting the RNA binding sites of the protein (Oh et al., 2016). The PAR-CLIP approach, however, allows utilization of T to C transitions in analysis of sequencing results, which are introduced specifically to uridines that were crosslinked to the bound protein. Those diagnostic transitions should allow increased specificity in analysis of DDX3X interacting sites on RNA. For our experiment, HEK293 Flp-In cells expressing inducible FLAG-tagged DDX3X protein were treated with 4-thiouridine (4sU) and RNA-protein complexes were cross-linked using UV-A at 365 nm (Hafner et al., 2010). DDX3X cross-linked RNA fragments we determined using high-throughput sequencing and originated primarily from coding transcripts (Fig. 11B). This is in line with previously published data,

which showed that a major regulatory role of DDX3 is related to mRNA and translation regulation (Oh et la., 2016; Shih et al., 2008). In order to better understand regulatory roles that DDX3 plays through RNA interactions, we decided to further focus our work on the protein binding to mRNAs.

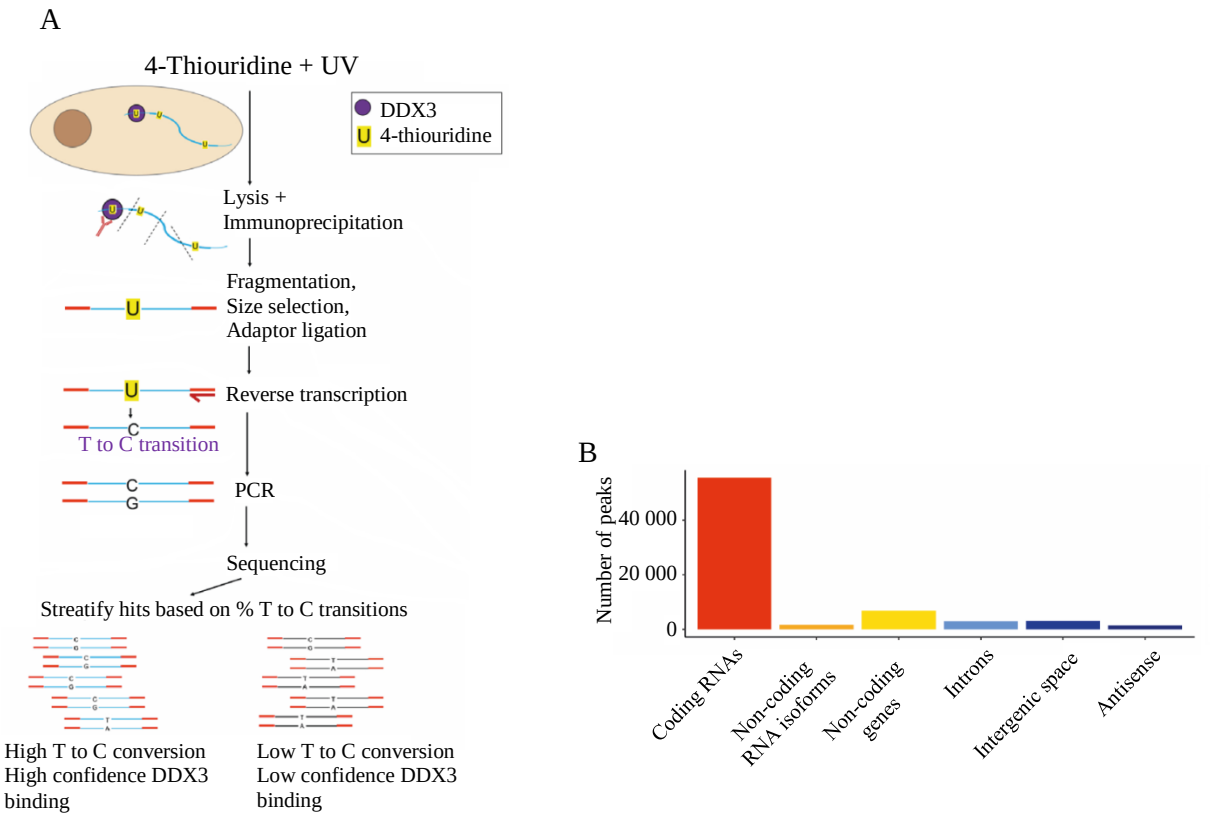


Figure 11. DDX3 binds prevalently to mRNAs (Figure generated by Lorenzo Calviello). A – A representation of the PAR-CLIP experiment workflow allowing for determination of general and high confidence protein binding sites in the transcriptome during data analysis. B – Number of DDX3 binding sites among different classes of non-rRNA transcripts, determined with two biological replicates of PAR-CLIP (analysis performed by Emanuel Wyler and Lorenzo Calviello).

4.3.2. A subset of human transcriptome is affected by DDX3 depletion

To be able to identify transcripts that are regulated by DDX3, we performed ribosome profiling and mRNA sequencing experiments. We have attempted to develop HEK293 cell lines expressing endogenous DDX3 protein tagged with an auxin-inducible degron (AID) on

the C-terminus using the CRISPR/Cas9 system. AID allows for a rapid depletion of a protein in eukaryotic cells upon addition of auxin, a plant hormone not present in eukaryotic cells and neutral to their physiology (Natsume et al., 2016). In presence of auxin, proteins bearing the AID tag become subject to ubiquitination and degradation, which is able to knock-out the protein within 24h (Natsume et al., 2016). We have reasoned that a complete depletion of DDX3 with an AID-based system would be beneficial to our analysis. Residual expression of a target in knock-down approaches may confer functions that are difficult to assay and may obscure the full picture of function of the protein of interest. We were successful in selecting for viable cells bearing an AID knock-in in the terminus of DDX3-coding sequence, however failed to induce the protein degradation with auxin (data not shown). We did not determine the precise reason for this failure at the time of performing this study and instead decided to use siRNAs for functional studies of DDX3.

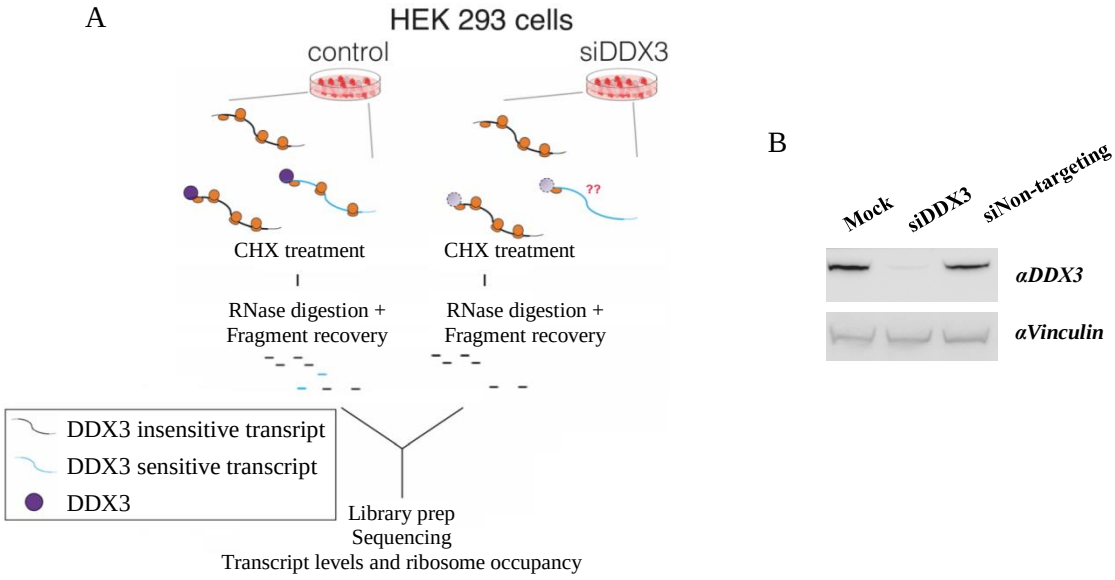


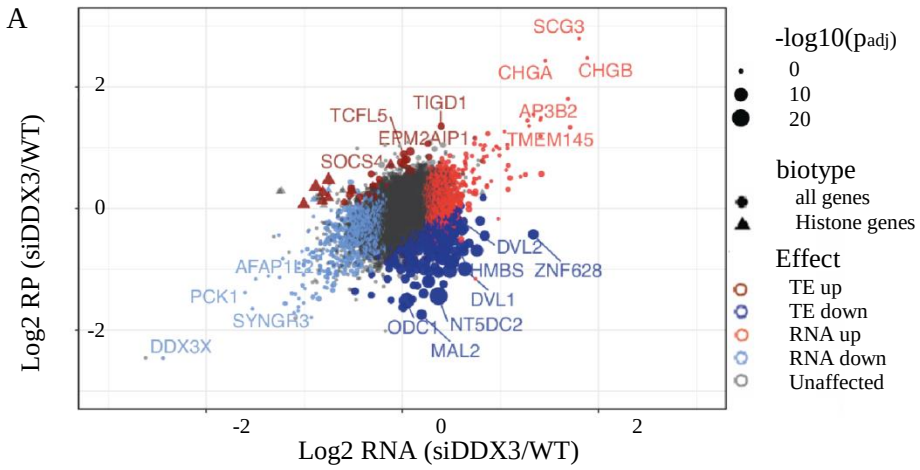
Figure 12. Ribosome profiling approach performed on cells depleted of DDX3 protein (Figure generated by Lorenzo Calviello). A – A workflow of ribosome profiling experiment. B – Image of Western blot analysis showing levels of DDX3X protein and Vinculin loading control in HEK293 control cells and cells 48h after transfection with either non-targeting control, or DDX3-targeting siPOOLs.

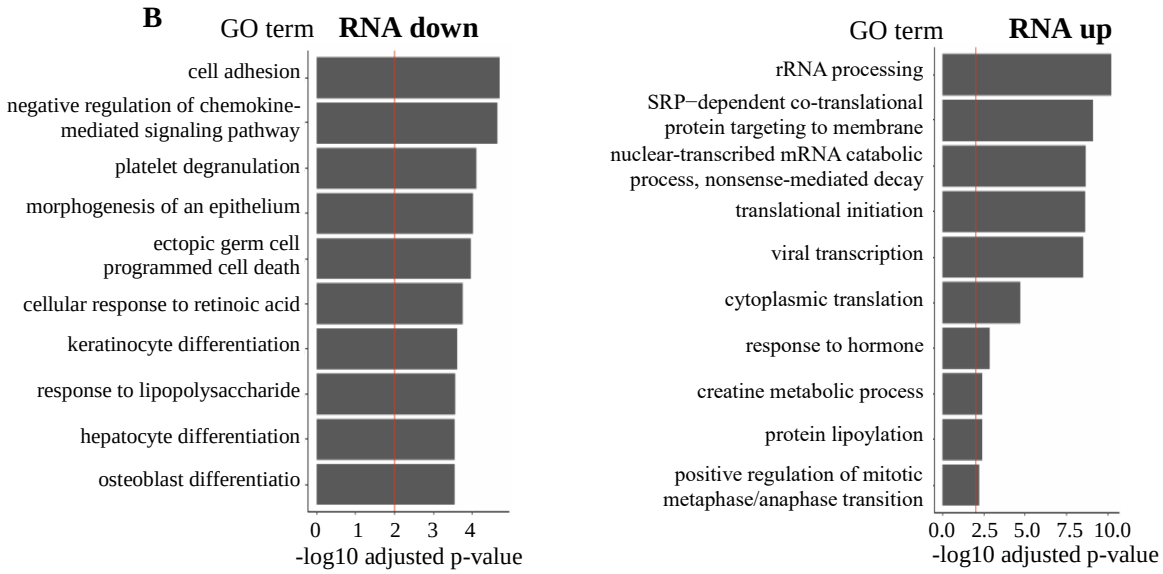
We have used HEK293 cells transfected with siRNAs targeting DDX3 mRNA or control siRNAs (Fig. 12). The computational analysis for this work was performed by

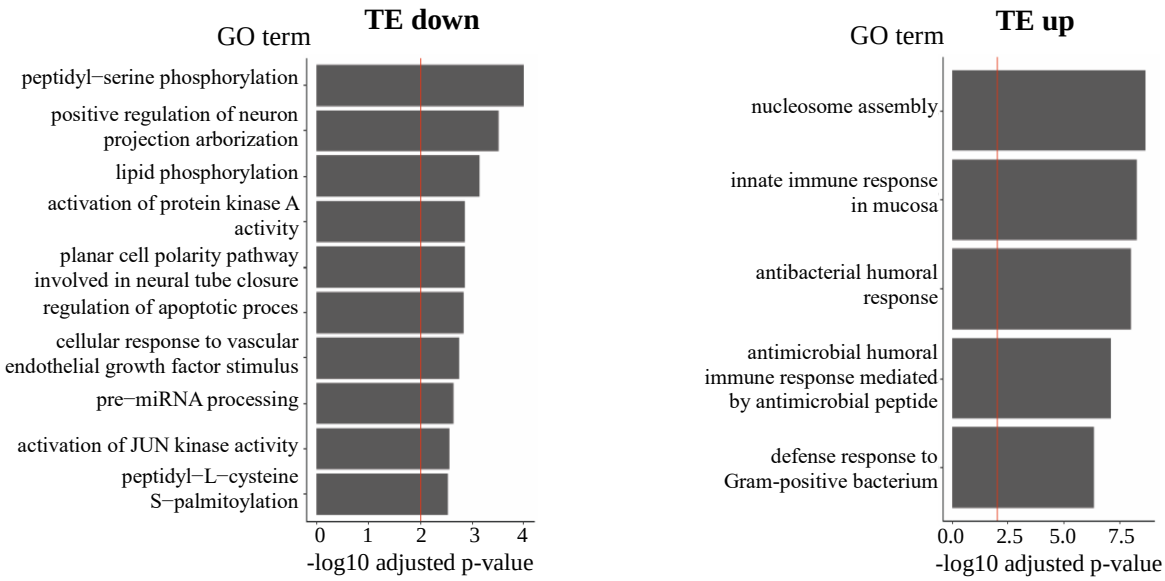
Lorenzo Calviello. We have ribosome profiling data to determine changes in translation and used the mRNA sequencing data to normalize for changes in transcript abundance in order to be able to calculate translation efficiency (TE) changes upon DDX3 depletion. We observed that most expressed transcripts do not display significant changes in their abundance nor in TE upon depletion of DDX3 (Table 11 and Fig. 12A). About 6% and 7% of expressed mRNAs were downregulated and upregulated, respectively. This observation is in line with previously published data, where changes of expression level of DDX3 did not strongly affect abundance of transcripts (Shih et al., 2008). We have also determined that only around 2% of all expressed mRNAs displayed changed TE (Table 11). Among those changes, decreased TE upon DDX3 knockdown was more prominent. Those results are in line with data published before, which reported that DDX3 does not affect translation globally, but appears to be necessary for controlling this process on a subgroup of mRNAs containing structured 5' UTRs (Lai et al., 2008). At the same time, they are in contrast with a report stating that changes of DDX3 levels have dramatic effects on protein synthesis (Shih et al., 2008). We have also performed a gene ontology (GO) analysis of transcripts affected by DDX3 depletion (Fig. 13B). An example of a gene affected by DDX3 depletion is ODC1 (Fig. 13C), which is known to be subject to translational control and is implicated in protecting cells from DNA damage and in cancer. (Gerner and Meyskens 2004; Hogarty et al., 2008).

Table 11. Summary of number of genes that change translation and mRNA expression levels after DDX3 knock-down. TE – translation efficiency (the ratio of ribosome protected footprints to RNA). Determined from two averaged biological replicates of ribosome profiling and mRNA sequencing.

	Not significant	**RNA level down**	**RNA level up**	**TE down**	**TE up**
Number of genes	10365	792	913	249	34
% of total	83.91	6.41	7.39	2.02	0.28


A
SCG3
CHGA
CHGB
TIGD1
TCFL5
AR3B2
ERM2AIP1
TMEM145
SOCS4
DVL2
AFAP1L2
HMBS
ZNF628
PCK1
DVL1
SYNGR3
ODC1
NT5DC2
DDX3X
MAL2
Log2 RP (siDDX3/WT)
Log2 RNA (siDDX3/WT)
-log10(padj)
0
10
20
biotype
all genes
Histone genes
Effect
TE up
TE down
RNA up
RNA down
Unaffected


B
GO term
RNA down
cell adhesion
negative regulation of chemokine-mediated signaling pathway
platelet degranulation
morphogenesis of an epithelium
ectopic germ cell programmed cell death
cellular response to retinoic acid
keratinocyte differentiation
response to lipopolysaccharide
hepatocyte differentiation
osteoblast differentiatio
-log10 adjusted p-value
GO term
RNA up
rRNA processing
SRP−dependent co-translational protein targeting to membrane
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay
translational initiation
viral transcription
cytoplasmic translation
response to hormone
creatine metabolic process
protein lipoylation
positive regulation of mitotic metaphase/anaphase transition
-log10 adjusted p-value


GO term
TE down
peptidyl−serine phosphorylation
positive regulation of neuron projection arborization
lipid phosphorylation
activation of protein kinase A activity
planar cell polarity pathway involved in neural tube closure
regulation of apoptotic proces
cellular response to vascular endothelial growth factor stimulus
pre−miRNA processing
activation of JUN kinase activity
peptidyl−L−cysteine S−palmitoylation
-log10 adjusted p-value
GO term
TE up
nucleosome assembly
innate immune response in mucosa
antibacterial humoral response
antimicrobial humoral immune response mediated by antimicrobial peptide
defense response to Gram-positive bacterium
-log10 adjusted p-value

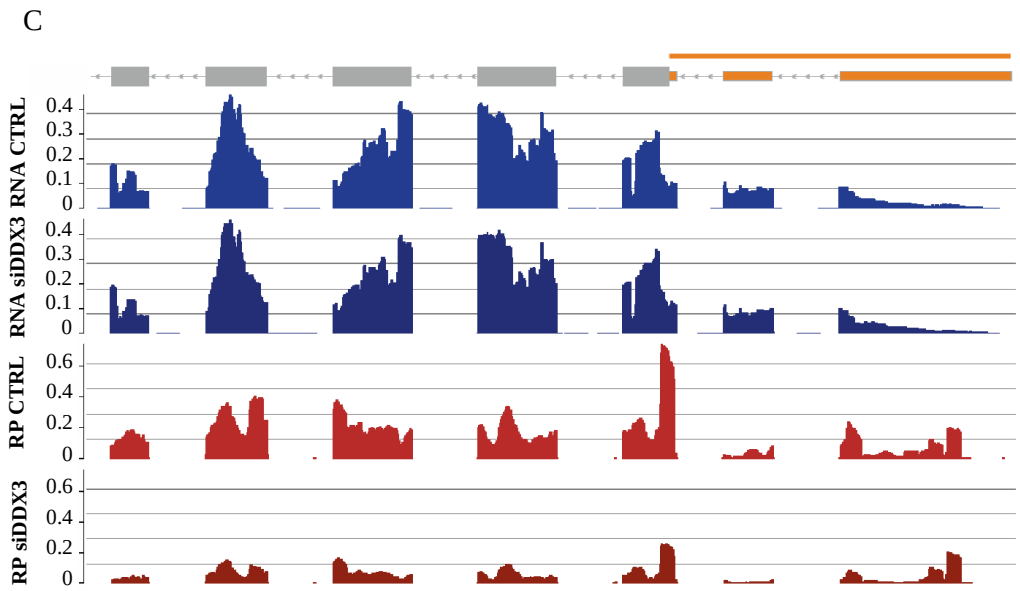


Figure 13. Depletion of DDX3 affects translation (Figure generated by Lorenzo Calviello). A – Representation of differential expression and changes in ribosome occupancy on mRNAs upon DDX3 knock-down determined with averaged data from two biological replicates of mRNA sequencing and ribosome profiling. RP – ribosome profiling. B – Gene ontology terms enriched in sets of genes transcripts of which change their level or TE upon DDX3 depletion. C – Effects of DDX3 knock-down on translation and expression of ODC1 mRNAs. Averaged results from two biological replicates are shown.

4.3.3. Depletion of DDX3 affects translation initiation

The yeast ortholog Ded1 has been reported to translationally regulate its target mRNAs through interactions with their 5' UTRs (Guenthler et al., 2018; Gupta et al., 2018). A similar regulation has been observed in case of individual mRNAs in human cells (Ku et al., 2018; Lai et al., 2008; Lai et al., 2010; Soto-Rifo et al., 2012). However, this regulation has not been investigated in human cells on a transcriptome-wide scale. We have focused our attention on 5' UTRs in analysis of generated data to investigate DDX3-mediated regulation of this structure in whole transcriptome. We have observed that depletion of DDX3 with siRNAs led to increased detection of ribosome footprints in 5' UTRs when compared to the coding sequences (Fig. 14A). This effect was most pronounced for the transcripts that had their TE reduced upon depletion of the protein (Fig. 14A and Fig. 14B). Accumulation of ribosomes in 5' UTRs upon depletion of DDX3 may reflect defects in initiating translation by assembled ribosomes, for which DDX3 may be required. This line of reasoning is in

agreement with previously reported functions of DDX3 and its association with translation initiation factors (Lee et al., 2008; Soto-Rifo et al., 2012). To test the hypothesis that DDX3 regulates translation initiation of mRNAs with structured 5' UTRs, we analyzed properties of 5' UTRs of transcripts affected and unaffected by DDX3 depletion. We have not seen any strong differences in 5' UTR lengths of transcripts differentially regulated by DDX3 (Fig. 14C). However, GC content of analyzed regions was clearly increased in transcripts that displayed lower TE upon DDX3 depletion, especially when compared with genes that displayed lower TE (Fig. 14C). GC content of a transcript is in general a good predictor for the presence of secondary structure (Galtier and Lobry 1997; Dohm et al., 2008), including RNA G-quadruplexes (rG4) (Murat et al., 2018). We have analyzed the predicted secondary structures in 5' UTRs of RNAs and indeed found that in case of transcripts with lower TE upon DDX3 depletion, the predicted folding free energy was lower than that of other transcripts (Fig. 14D). Those findings support the hypothesis that there is a general connection between DDX3 mediated translation regulation and structure of 5' UTRs. They are also in line with recently reported DDX3 in-vitro interaction with rG4 structure from a 5' UTR of a human transcript (Herdy et al., 2018).

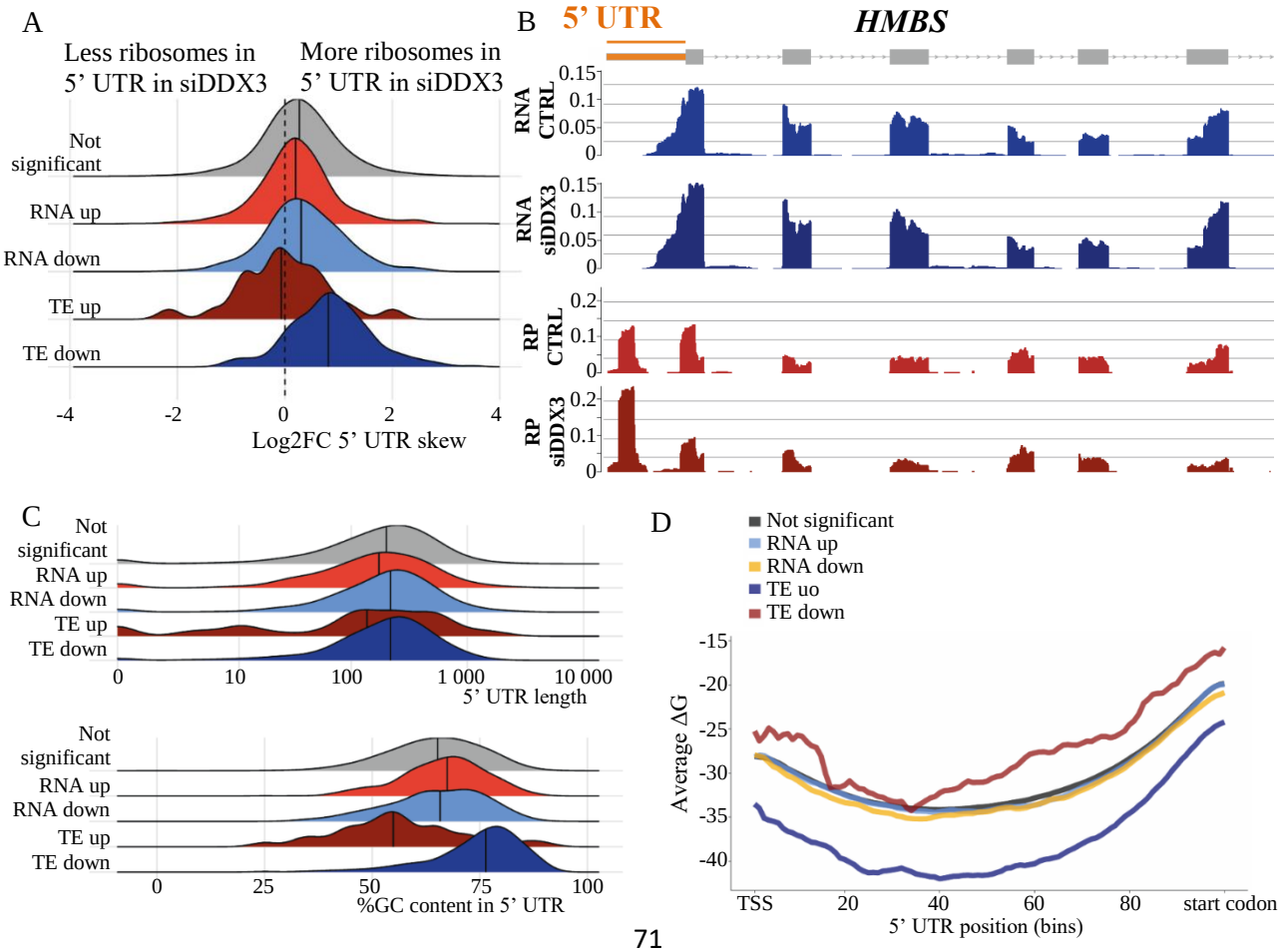


Figure 14. Transcripts sensitive to DDX3 depletion have more structured 5' UTRs (Figure generated by Lorenzo Calviello). A – Representation of the ratio of fold-change in ribosome occupancy in the 5' UTR to the coding sequence in transcripts affected and unaffected by DDX3 depletion, averaged from two biological replicates. B – An example of a *Hydroxymethylbilane Synthase* (*HMBS*) gene displaying increase in ratio of ribosome occupancy in the 5' UTR when compared to the coding sequence. Averaged result from two biological replicates are shown. C – Representation of number of mRNAs having different characteristics of their of 5' UTRs (5' UTR length – upper panel; 5' UTR GC content – lower panel) for groups of transcripts affected and unaffected by DDX3 depletion, determined from two averaged biological replicates. D – A representation of predicted Gibbs free energy values along 5' UTR sequences of transcripts affected and unaffected by DDX3 depletion.

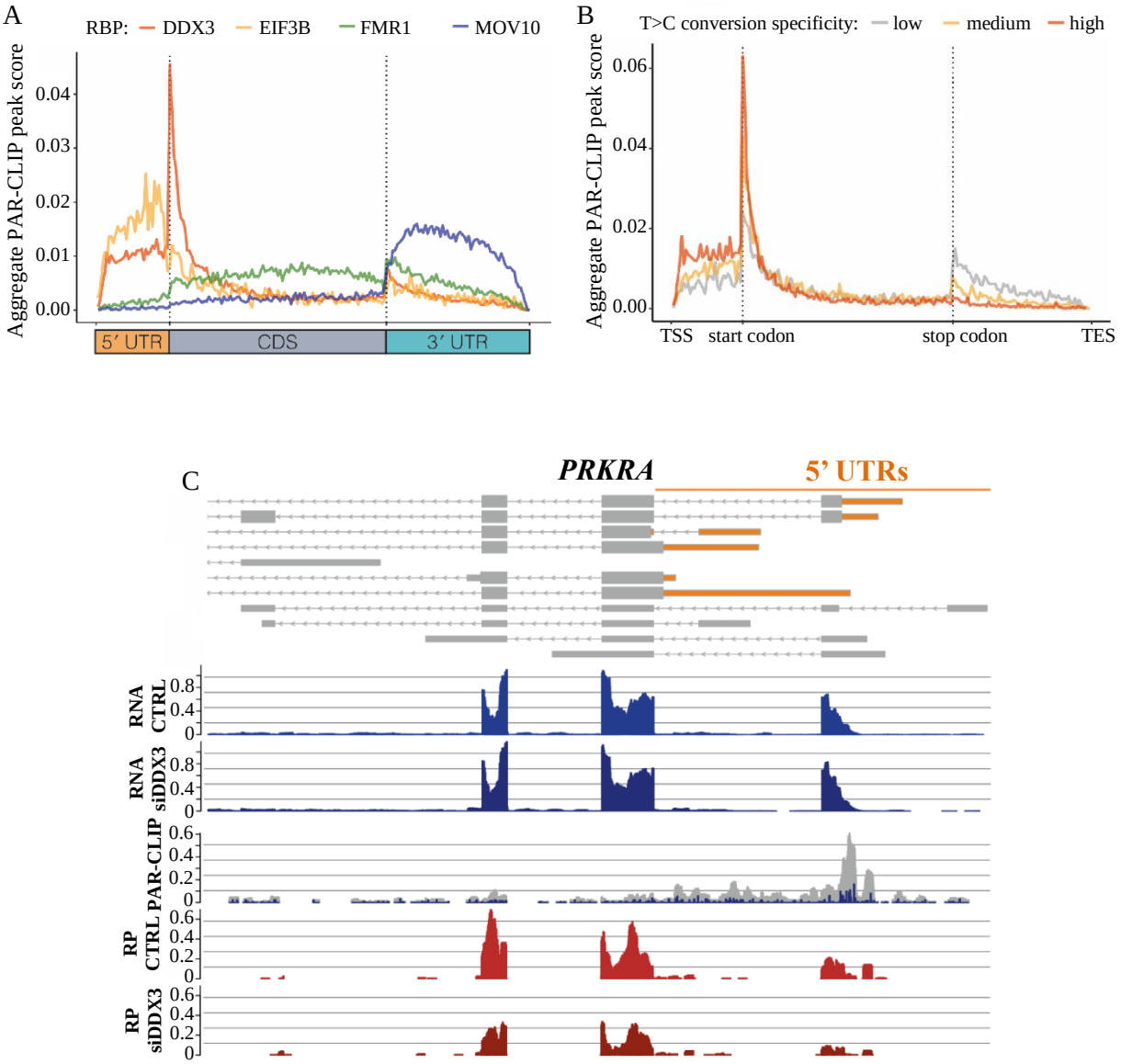
4.3.4. DDX3 specifically binds to 5' UTRs

To better understand DDX3 mediated regulation of mRNAs and their 5' UTRs, we analyzed the aforementioned PAR-CLIP data in combination with results obtained with ribosome profiling and mRNA sequencing. We have used sequenced RNA fragments to determine general DDX3 binding sites and the advantage of presence of diagnostic T to C transitions in those reads to determine high confidence binding sites (Fig. 15A and B). By analyzing sequenced non-rRNA RNA fragments that were identified as DDX3 bound in PAR-CLIP experiment, we have determined that the protein binds to most mRNAs expressed in HEK293 cells (Table 12). The majority of bound transcripts were not affected by DDX3 depletion on level of transcript abundance nor TE. To gain more insight into the context of DDX3-binding to RNAs we decided to assay the averaged binding of the protein to specific regions of the transcript by averaging detected reads across all expressed transcripts in a metagene analysis. We also compared DDX3 binding to binding pattern of Eukaryotic Translation Initiation Factor 3 Subunit B (eIF3B), Fragile X Mental Retardation 1 (FMR1) and Mov10 RISC Complex RNA Helicase (MOV10) proteins, which were determined by PAR-CLIP experiments before this work and are available in the POSTAR2 database (Zhu et al., 2019). EIF3B is a part of the translation initiation complex (Aitken et al., 2016), FMR1 associates with elongating ribosomes (Chen et al., 2014) and MOV10 binds to 3' UTRs (Sievers et al., 2012; Gregersen et al., 2014). We have observed that the strongest DDX3 binding happened in 5' UTRs, with the highest peak near the start codon (Fig. 15A). The protein binding was also observed in the coding sequence and in 3' UTRs, albeit at lower levels than the binding to the 5' UTR. This binding pattern correlates best with binding of

eIF3B (Fig. 15A), which is a protein having core functions in translation initiation. The protein associates with 40S ribosome and facilitates the recruitment of factors creating pre-initiation ribosome complex (Wagner et al., 2016). The metagene analysis of high confidence binding sites represented by T to C transitions in the PAR-CLIP approach determined an even clearer prevalence of DDX3 binding near the start codon (Fig. 15B). This binding pattern was observed also for genes that displayed reduced TE upon DDX3 knock-down, examples of which are PRKRA and ODC1 (Fig. 15C). In addition to high frequency of binding to mRNAs, we have determined DDX3 binding to 18S ribosomal RNA (Fig. 15D). The major interaction site was mapped to helix 16 of the 18S rRNA. This position of DDX3 may provide access to resolve mRNA secondary structures during inspection by the scanning ribosome. The observed cross-link site is opposite to the eIF4B binding site, a factor important for ribosome function (Walker et al., 2013; Sen et al., 2016). This is also in line with previously observed association of DDX3 with eIF3 (Lee et al., 2008). Taken together, those results are consistent with previously reported association of DDX3 with translation initiation machinery and its requirement for translation initiation of specific mRNAs (Soto-Rifo et al., 2012), which bears resemblances to translation regulation of the DDX3 ortholog Ded1 in yeast (Sen et al., 2015).

Table 12. Summary of effects of DDX3 depletion on transcriptome. Numbers of transcripts effected on the level of abundance and TE, as well as numbers of transcripts directly bound by DDX3 in differently effected groups of transcripts are shown Determined from two averaged biological replicates of ribosome profiling and mRNA sequencing.

	All detected	**RNA level up**	**RNA level down**	**TE up**	**TE down**
Detected genes	12054	617	792	34	249
DDX3-bound (% of detected genes)	8869 (73.6%)	492 (79.7%)	496 (62.6%)	28 (82.4%)	211 (84.7%)
DDX3-bound, high T to C transition frequency (% of DDX3-bound)	5932 (66.9%)	332 (67.5%)	336 (67.7%)	20 (71.4%)	168 (79.6%)


A
RBP: DDX3 EIF3B FMR1 MOV10
Aggregate PAR-CLIP peak score
0.04
0.03
0.02
0.01
0.00
5′ UTR
CDS
3′ UTR
B
T>C conversion specificity: low medium high
Aggregate PAR-CLIP peak score
0.06
0.04
0.02
0.00
TSS
start codon
stop codon
TES
C
PRKRA
5' UTRs
RNA CTRL
RNA siDDX3
PAR-CLIP
RP CTRL
RP siDDX3
0.8
0.4
0
0.6
0.2

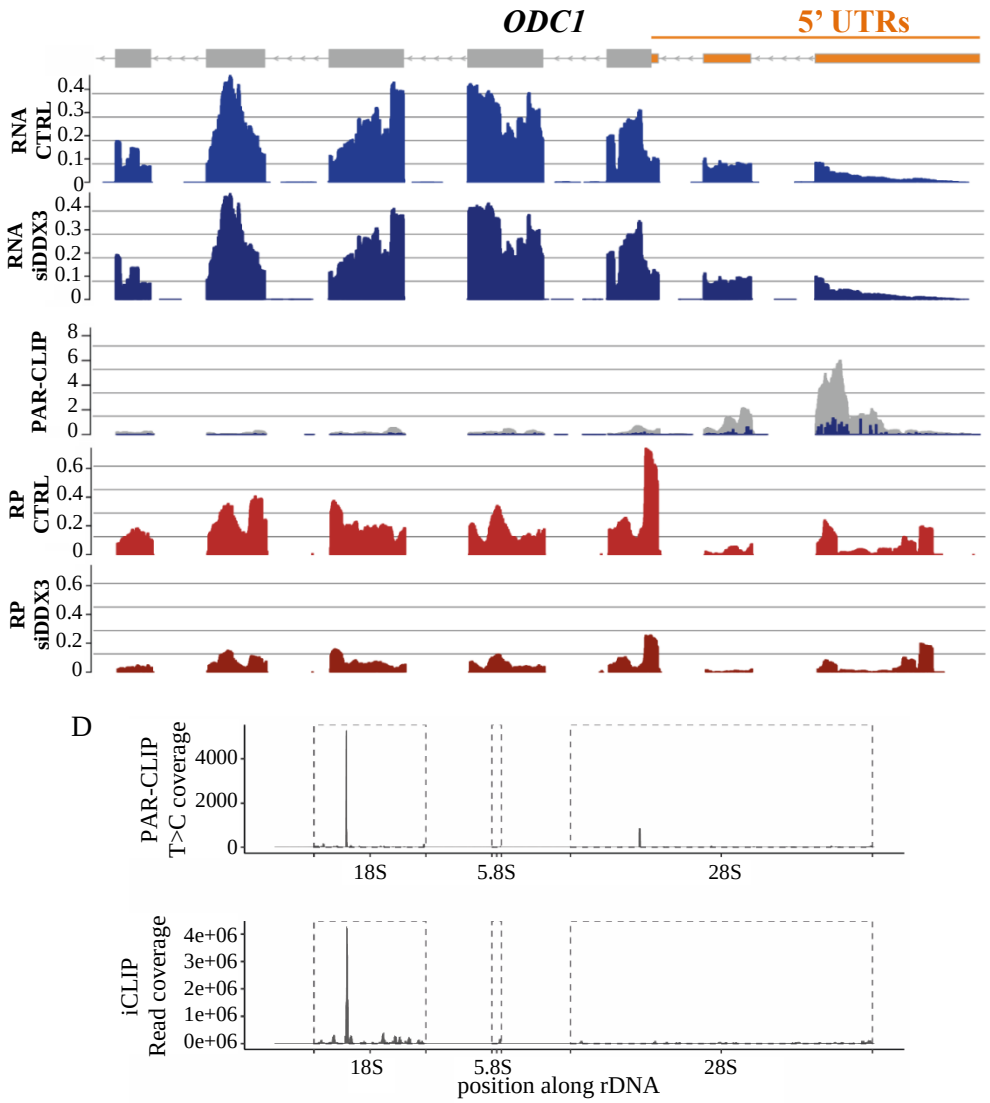


Figure 15. DDX3 binds to 5' UTRs of mRNAs, similarly to canonical translation initiation factor (Figure generated by Lorenzo Calviello). A – Metagene plot of DDX3, eIF3B (a canonical translation initiation factor), FMR1 (associates with elongating ribosomes) and MOV10 (a helicase binding to 3' UTRs) PAR-CLIP data showing protein occupancy along the transcript body. B – Metagene plot of DDX3 PAR-CLIP data utilizing peaks of T to C transitions from two averaged biological replicates, showing protein occupancy along the transcript body. C – Examples of two genes (PRKRA and ODC1) whose transcripts were identified to be bound by DDX3 and displayed reduced TE upon the protein knock-down. In Panel PAR-CLIP grey color represents mapping of reads from the experiment and dark blue high frequency of T to C transitions. Averaged result from two biological replicates are shown. D – Representation of sites of high confidence DDX3 binding to ribosomal RNA, assayed by peaks of T to C transitions from two averaged biological replicates of PAR-CLIP

(upper panel) and compared to binding determined with iCLIP by Oh et al., 2016 (bottom panel).

4.3.5. DDX3 binds preferentially to GC-rich mRNA regions

The DDX3 binding to 5' UTRs of transcripts that display decreased TE observed in this study allowed us to investigate a connection between 5' UTR structure and transcripts whose translation depends on DDX3. This connection has been observed before for certain examples of human mRNAs (Ku et al., 2018; Lai et al., 2008; Lai et al., 2010; Soto-Rifo et al., 2012) and for DDX3 ortholog from yeast Ded1 (Guenther et al., 2018; Gupta et al., 2018). We first checked if DDX3 displays preference for binding to structured 5' UTRs. We have observed that PAR-CLIP peaks with high T to C transitions occur in GC-rich regions with high specificity (Fig. 16A). Next, we checked how T to C conversions, which represent DDX3 binding, change across bodies of mRNAs that are sensitive and insensitive to the protein knock-down. We have found that enriched DDX3 binding to 5' UTRs is specific to transcripts that displayed decreased TE upon the protein knock-down (Fig. 16B). Those results suggest that there is a direct relationship between DDX3-depletion and the observed decrease in TE. They confirm the previous observations that DDX3 associates with protein complexes controlling translation initiation. The helicase however is required for translation initiation of only of subset of mRNAs. Our observations support the notion that those mRNAs require DDX3 activity to resolve RNA structures in their 5' UTRs to permit translation initiation by the ribosome.

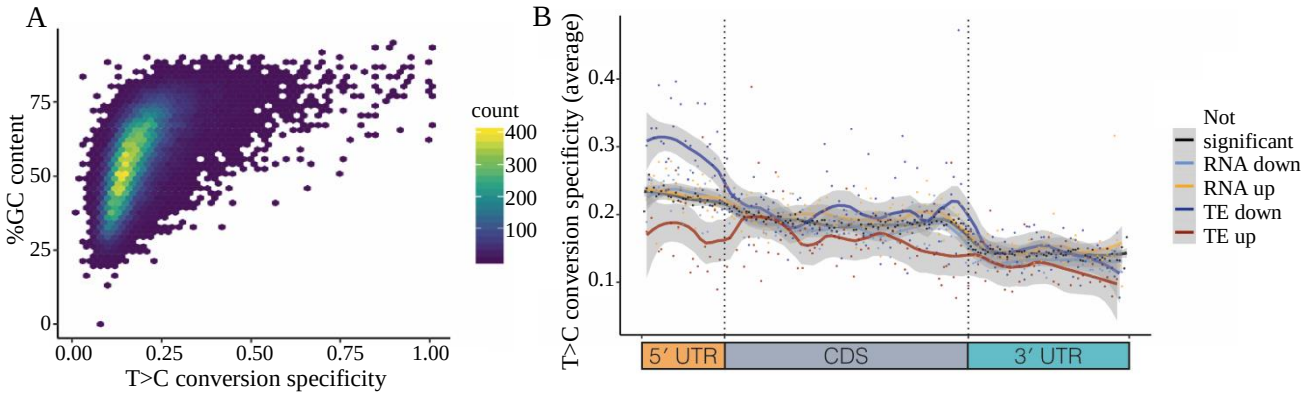


Figure 16. DDX3 binding regulates translation of GC-rich transcripts (Figure generated by Lorenzo Calviello). A – Representation of GC content of regions of DDX3 binding determined by peaks of T to C transitions from averaged two biological replicates of PAR-CLIP experiments. B – A metagene plot of DDX3 high confidence binding sites determined by peaks of T to C transitions from two

averaged biological replicates of PAR-CLIP experiments along transcript bodies of genes affected and unaffected by depletion of DDX3.

4.3.6. DDX3 specifically binds 5’ ends of lncRNAs

DDX3 prevalently associates with mRNAs, but its binding to lncRNAs has also been observed (Fig. 11B). This interaction has been reported before, but it has not been analyzed (Oh et al., 2016). We have decided to investigate the nature of DDX3 binding to lncRNAs using PAR-CLIP data. We have observed that protein binding along averaged lncRNA transcript body has a similar pattern as observed for mRNAs (compare Fig. 15B with Fig. 17A), with most prevalent binding spotted near the 5’ end of transcripts. This binding preference was observed for both spliced and unspliced lncRNAs (Fig. 17B).

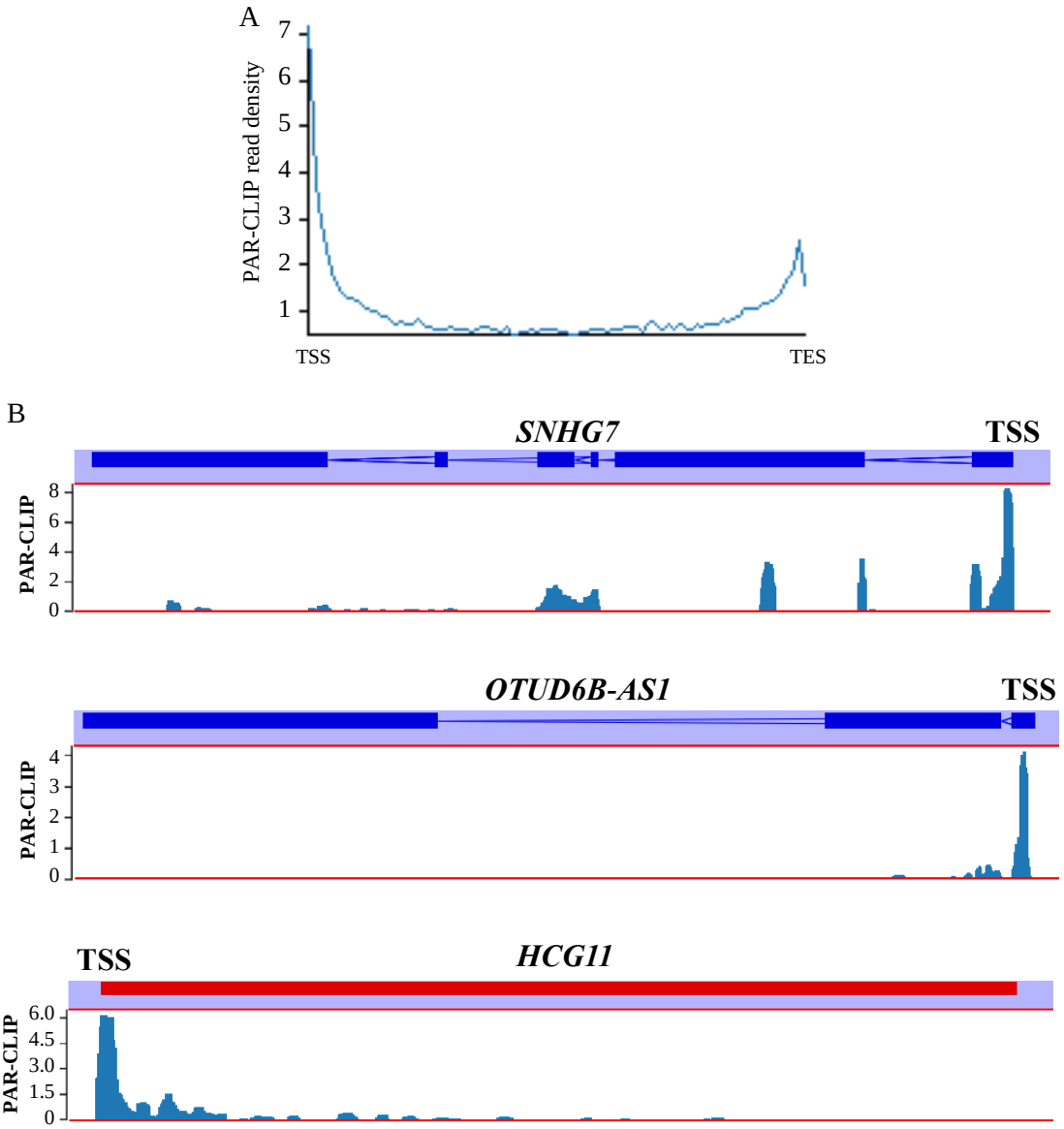


Figure 17. DDX3 binds preferentially near 5' ends of lncRNAs. A – Metagene plot of DDX3 PAR-CLIP data averaged from two biological replicates, showing protein occupancy along the length of averaged lncRNA transcript body from two averaged biological replicates. B – Representative examples of three lncRNA genes whose transcripts were identified to be bound by DDX3. Coverage profiles of mapped PAR-CLIP reads containing T to C transitions from two averaged biological replicates are shown. Color of gene annotation represents the strand from which transcripts originates (red – sense, blue – antisense).

5. Discussion

5.1. The need for novel RNA pull-down methodologies

Numerous approaches to enrich for the RNA of interest along with its protein interactors have been developed prior to this work. The traditional in-vitro approaches, where RNA of interest is in-vitro transcribed, immobilized and then incubated with a cellular lysate are associated with high levels of binding by molecules which under physiological conditions do not interact with the transcript of interest (Treiber et al., 2017). This problem seriously hampers usefulness of those methods in identifying biologically relevant RNA-protein interactions. This issue was addressed in the approach utilizing Csy4 molecule for the *in-vitro* RNA pull-down, where conditional version of the enzyme allowed for specific elution of the transcript bound by it (Lee et al., 2013). This has lowered levels of background signal obtained in the approach but did not eliminate the problem.

While work presented in this thesis was being performed RNA pull-down approaches relying on affinity purification with biotinylated oligonucleotides were being reported to be successful in identification of proteins associating with single transcript in combination with mass spectrometry (Chu et al., 2015; Warneford-Thomson et al., 2017). Due to their high efficiency they may be considered a golden standard for identification of proteins interacting with single RNA of interest. Their major disadvantage, however, is high cost of an array of biotinylated oligonucleotides that has to be purchased for each target RNA. This has been a limiting factor in popularization of the approach.

A wide array of affinity purification methods utilizing an RNA sequence which can be fused to a transcript expressed in cells and is bound by a specific molecule has been developed in the past (Bachler et al., 1999; Hartmuth et al., 2002; Srisawat and Engelke 2001; Youngman and Green 2005; Yoon et al., 2012). The advantage of those approaches in comparison to in-vitro pull-downs is that they allow capture of RNA-protein interactions that form in living cells. This allows capturing the RNA-protein interactions that better reflect in-vivo contacts of those molecules and limit experimental background. The optional use of cross-linking to induce covalent bonds between target transcript and associated with it proteins is another advantage that may further increase faithful capture of in-vivo interactions. Those methods were successful in purification and identification of molecules associated with target transcript in combination with highly sensitive methods like next-generation sequencing and Western analysis (Yoon et al., 2012; Yoon and Gorospe 2016). Their use for

de-novo identification of proteins associated with a transcript of interest has been limited and required overexpression of the targeted RNA (Li and Altman 2002; Dong et al., 2015). Those methods have not been often utilized owing to their technical challenges. In particular, low affinity or specificity of target RNA recognition have limited the widespread use of approaches based on aptamers, and bacteriophage coat protein binding sites. We have attempted to overcome those limitations by utilizing Csy4 protein for an in-vivo RNA pull-down. The protein has been shown to have exceptionally high affinity towards its target (equilibrium dissociation constant, Kd = 50 pM), and sequence specificity (Sternberg et al., 2012). This is a markedly higher affinity than that of molecules previously used for *in-vivo* RNA pull-down approaches based on bacteriophage MS2 coat protein and S1m streptavidin binding aptamer. MS2 interaction with its target stem loop has K_d = 20 nM (Stockley et la., 1995) and S1m interaction with streptavidin has K_d = 29 nM (Leppek and Stoecklin 2013). We have reasoned that utilization of a protein with such a high binding affinity has a chance to outperform approaches based on molecules with comparably low affinities towards their targets.

5.1.1. RNA pull-down with in-vitro purified Csy4 displays low target capture specificity

The utilization of Csy4 for pull-down approaches is possible thanks to development of a catalytically inactive version of this protein, which allows stable binding to the target stem loop without subsequent cleavage. The mutant version of Csy4 has its histidine residue 29 mutated to alanine (H29A) resulting in a conditional enzyme inactive under normal conditions. The catalytic activity can be rescued in the presence of imidazole (Lee et al., 2013). Our first approach utilizing recombinant, *in-vitro* purified Csy4 protein incubated with lysates from cells expressing HULC transcript tagged with Csy4 target hairpin succeeded in enriching for the transcript. It was, however associated with high experimental background levels reflected by apparent isolation of wide array of unrelated proteins (Fig. 2E) and of target unrelated transcript GAPDH (Fig. 2D). In this approach we have used conditional Csy4 mutant protein, which allowed imidazole driven elution of bound transcripts (Fig. 2C). This mode of elution has been shown before to specifically release transcripts directly bound by the protein and reduce background signal of unrelated molecules (Lee et al., 2013). Thus, observed enrichment for target unrelated transcript GAPDH and wide array of proteins in both negative and positive controls most likely reflects indiscriminate interactions of Csy4 protein with cellular RNAs. Csy4 has been shown to interact with its targets with high specificity under favorable *in-vitro* conditions (Sternberg et al., 2012). Moreover, Csy4 target sequence

is not normally present in the human transcriptome and Csy4 protein was not observed to target sequences other than the target stem loop when both molecules were expressed in human cells (Borchardt et al., 2015). Therefore, we have hypothesized that the environment of cellular lysate had a negative influence on activity of the protein, which resulted in its unspecific interactions with RNAs and observed levels of isolated proteins in our samples. Low specificity of this approach rendered it not useful for intended purpose and prompted us to test the second one.

5.1.2. RNA pull-down with Csy4 expressed in cells displays low target capture efficiency

In our second approach we expressed in cells the tagged Csy4 protein, mRNA of which was at the same time tagged with Csy4 target stem loops. In this system the interaction between the protein and its target was allowed to take place within cells and could be cross-linked in-vivo. Csy4 has been successfully expressed in human cells in a previous study retaining its enzymatic activity (Borchardt et al., 2015). Thus, capturing mutant Csy4-target transcript interactions in human cells seemed plausible. In this approach unlike in the previously tested one we did not detect high levels of target unrelated GAPDH transcript (Fig. 3D), even though the mode of elution used in this approach was not specific for Csy4 bound transcripts. This reinforced our view, that an issue of high experimental background levels observed in previously tested system reflected unspecific Csy4-RNA interactions under conditions of cellular lysate. However, the enrichment for the target mRNA was very modest in our second system (Fig. 3E). Our result is far from satisfactory when compared with affinity purification methods that we were hoping to outperform. Methods relying on use of MS2 coat protein and of S1m RNA binding aptamer achieve fraction of input target RNA recovered of ~1% and ~5% respectively (Leppek and Stoecklin 2013). Given the fact that Csy4 affinity towards its target is markedly higher than that of MS2 and S1m the results we achieved were unsatisfactory. Quantitative determination of in-vivo RNP composition requires isolation of significant amounts of RNAs from cellular lysates and it has been reliably achieved on endogenous transcripts only with methods utilizing biotinylated oligonucleotides, which were successful in recovering around 50% fraction of input target RNA (Simon et al., 2011). To be useful for intended purposes our system would have to at least perform better than S1m and MS2 based methods.

The low efficiency of enrichment for Csy4 mRNA could not be explained by the failure to enrich for the Csy4 protein from the lysate, as this step of the experiment appeared

to work relatively well in our hands (Fig. 3B and Fig. 3C). The disproportion between amounts of isolated Csy4 protein and isolated Csy4 tagged mRNA we have observed in our experiments suggests that Csy4 does not interact with its target efficiently in cells. We have assumed that after transfection levels of Csy4 protein in human cells are higher than those of its host mRNA as is true for most mRNAs in human cells (Silva and Vogel 2016; Liu et al., 2016), thus we did not expect low protein to transcripts ratios to be a limiting step in this approach. In a similar experimental pull-down design utilizing tagged MS2 coat protein both the protein and its target transcript are expressed under the same promoter as Csy4 in our approach and no problem with protein to transcript ratios was reported (Yoon et al., 2012). The low Csy4 protein to mRNA ratio in transfected cells is, however, a possible explanation of observed results. This could be caused by rapid Csy4 protein degradation in human cells. We did not, however observe any evidence for the protein degradation with Western analysis. Another possible reason for low Csy4 protein to mRNA ratio in transfected cells is inefficient protein translation. It seems possible that the presence of Csy4 target stem loop in 3' UTR of Csy4 mRNA has a negative effect on efficiency of translation. However, this assumption is contradicted by previous study showing that levels of a reporter protein encoded by a mRNA containing Csy4 target stem loop in its 3' UTR were not affected in presence of the Csy4 mutant protein in human cells (Borchardt et al., 2015).

Another possible explanation for low efficiency of the pull-down approach utilizing Csy4 protein and its mRNA tagged with Csy4 target stem loops is that folding of Csy4 target stem loop RNA sequences is inefficient. This could lead to lack of appropriate structures that can be bound by the protein. Inefficient folding has been shown to be a limiting step in case of pull-down with S1m RNA aptamer. That limitation could be circumvented by increasing the number of repeats of sequence forming S1m structure on the target transcript (Leppek and Stoecklin 2013). In our approach, however increasing the amount of Csy4 target stem loops on the Csy4 mRNA did not increase the pull-down efficiency (Fig. 3E). Thus, suggesting that inefficient RNA folding was not the limiting step in our approach. Another potential explanation for inefficient Csy4-RNA interaction in cells is that the protein may be actively removed from its target by RNA helicases. Testing this hypothesis, however, is difficult and we did not address it in our work.

Taken together we have determined that using recombinant in-vitro purified Csy4 mutant protein for the pull-down with cellular lysates is not suitable for efficient enrichment of Csy4 stem loop tagged transcripts. This approach did not show promise for further

optimization and for use for identification of proteins associated with a transcript of interest. Our second approach with Csy4 and target transcript both expressed in cells did not meet our expectations either. We were, however, unable to clearly determine the reason for its low efficiency. It is possible that further optimization of the method may result in a system useful in identification of RNA-protein interaction with single transcript resolution. Our results, however, are not encouraging.

5.2. Landscape of non-coding transcriptome in DNA damage response

We have determined the transcriptional landscape of MCF-7 cells in response to DNA damage induced by gamma-irradiation. Our goal was to identify novel lncRNAs involved in DDR. By separately sequencing polyadenylated, non-polyadenylated and nascent fractions of the transcriptome, we were able to analyze differential expression of diverse classes of transcripts independently of the properties of their 3' ends or turnover rates. Our approach allowed us to determine differential expression not only of linear transcripts, but also of circRNAs, for which differential expression in DDR has not been reported previously. By performing sequencing experiments on cells in various time points after induction of DDR, we were able to dynamically capture transcript up- or downregulation in response to genotoxic stress. Our work provides an extended insight into the transcriptional landscape during the DDR.

In the polyA+ sequencing experiment, alongside mRNAs typically upregulated in DDR, we have initially detected 24 differentially expressed lncRNAs (Fig. 4). Of those non-coding transcripts, 21 did not have a known molecular function at the time when this analysis was performed. A subsequent analysis with an MCF-7 specific genome annotation revealed 5 additional differentially expressed transcripts, which were not annotated in Gencode.v19 genome annotation. We sought to identify additional lncRNAs differentially expressed in DDR by sequencing non-polyadenylated fractions of transcripts. In this experiment, we have identified additional 24 differentially expressed lncRNAs (Fig. 5).

Standard RNA sequencing experiments use whole-cell RNA extractions, thus limiting gene expression analysis to steady-state RNA abundance levels. We have complemented our efforts in investigating transcriptomes of cells exposed to DNA damage by performing nascent RNA sequencing. With this method we were able to enrich for unstable transcripts. This data provides a valuable insight into expression of short-lived transcripts expressed in DDR, which are difficult to capture with traditional sequencing approaches. Using nascent

RNA sequencing we were able to detect novel transcription units, which were not present in available genome annotations. Two novel transcription units that we presented in this work are intergenic and longer than 200 nt (Fig. 6). Interestingly, both transcripts appear not to be rapidly turned-over in cells as we could also detect their signatures with sequencing of steady state RNAs (Fig. 7A and B). Moreover, we were able to successfully quantify one of those transcripts using qRT-PCR from cDNA library derived from polyadenylated RNAs. Our results show that nascent RNA sequencing approach may be used for identification of novel transcriptional units with success. The dataset we have generated will also be useful in future studies of novel and short lived RNAs with potential functions in DDR.

5.2.1. Evaluation of functions of LINC00475, LINC01021, TCERG1L-AS1 and UNC5B-AS1 in regulation of DNA damage response required further study

From the lncRNAs identified as differently expressed in DDR, we have selected those displaying highest expression changes and confirmed their upregulation with qRT-PCR (Fig. 9). We have attempted to assay roles selected transcripts may play in cell survival after induction of IR. To this end we have knocked-down their expression using siRNAs and performed colony formation assay after knock-down of three lncRNAs that we were able to deplete. We did not, however, detect any significant influence on cell survival after the knock-down (Fig. 10). This observation may reflect lack of regulatory roles played by selected lncRNAs in DDR. It may, however, also be caused by inefficient knock-down of those transcripts by siRNAs. We were able to knock-down three out of four selected lncRNAs and only to around 50% of their initial levels. This level of knock-down may be insufficient to disrupt function of targeted transcripts. Alternatively, it is possible that knocked-down lncRNA genes maintain their function through the act of transcription itself and resulting transcripts themselves are non-functional. Functionality of those lncRNAs should be assayed further with a different approach then siRNA knock-down, such as CRISPR-Cas9 mediated deletions. Potentially, a more efficient knock-down of lncRNAs may be achieved with RNase H activate antisense oligonucleotides (ASOs), which utilizes an endogenous RNAse H enzyme, which is prevalently localized in the nucleus. This method of knock-down has been shown to be more efficient then siRNAs for some lncRNA targets, especially when the transcript is localized to the nucleus (Lennox and Behlke 2016). This could be particularly important in the case of lncRNAs assayed in this study, as according to the NONCODE database (http://www.noncode.org/), only one of them (UNCB5B-AS1) prevalently localizes to the cytoplasm.

5.2.2. Circular RNAs are likely involved in regulation of DNA damage response

In addition to the aforementioned analysis of differential gene expression, we have performed an analysis of differential expression of circRNAs. The sequencing experiment of transcripts depleted of polyadenylated and ribosomal transcripts provides an enrichment of circRNAs, which is suitable for such an analysis (Linda and Salzman 2016). To our knowledge, this is the first report of differentially expressed circRNAs in DDR. We have detected high number of upregulated circRNAs in all time points of DDR. Interestingly, the majority of upregulated circRNAs were not detected in IR untreated control samples (Fig. 8A). We detected a wealth of circRNAs that seem to be DDR-specific in MCF-7 cells. Those transcripts prevalently originated from protein coding genes. We have performed a GO analysis of genes harboring circRNAs upregulated in DDR and observed that they were involved in plethora of regulatory functions, including DDR-related processes (Fig. 8B). The expression of the respective host genes was mainly unchanged. We were able to find 17 mRNA-circRNA pairs originating from the same gene, which both displayed differential expression in DDR (Table 1). Most of those genes had known functions related to DDR. Many functionally characterized circRNAs have a function in regulating expression of their mRNA siblings. Therefore, it is tempting to speculate that circRNAs from the above-mentioned mRNA-circRNA pairs have regulatory roles in DDR by influencing the levels of the respective mRNAs. Elaborating these relationships might be interesting for future studies.

In summary, we have provided a comprehensive picture of temporal transcriptome changes of human cells upon DDR, including the differential expression of circRNAs, which was not reported before this work. We did not succeed in identifying novel functional lncRNAs regulating DDR, but we have provided information potentially useful in future studies of those molecules.

5.3. DDX3 regulates translation of a subset of its target mRNAs

We have determined for the first-time transcriptome-wide effects of DDX3 depletion on translation efficiency in human cells. In combination with investigation of DDX3 binding sites and measuring mRNA abundance upon knock-down, this allowed us to provide a new important insight into DDX3 modes of action and to verify previously reported information on DDX3 function and. Although DDX3 is an essential gene and according to our data binds to most expressed mRNAs in human cells, we observed that its depletion affects translation efficiency of a relatively small subset of targets. This observation may reflect a bias in our

approach, which failed to detect changes in translation efficiency of additional DDX3 targets. This bias may come from the limitations of siRNA technology in completely depleting the protein of interest from cells, resulting in levels of protein, which still may be enough to maintain most of the protein functions in cells. The results we have produced argue that the main role of the protein is to promote translation of a subset of transcripts. The DDX3 gene is altered in many diseases. Our discoveries further the understanding of DDX3 functions, which may be useful for understanding human pathologies in which this protein is implicated.

5.3.1. DDX3 binding to coding and non-coding transcripts displays a similar pattern

Using PAR-CLIP, we have defined protein coding transcripts as major non-rRNA targets of DDX3 binding (Fig. 11B). Those findings are in line with previous work, which used iCLIP to determine DDX3 binding sites to human transcriptome (Oh et al., 2016) and with previously reported functions of DDX3 which are related to regulation of mRNA translation (Lai et al., 2010; Lee et al., 2008). Binding to non-coding transcripts was also observed, but its frequency was modest when compared to the binding to mRNAs. We did not investigate potentially regulatory functions of DDX3 association with lncRNAs. In light of our results discussed below we can however put forward a hypothesis explaining it. It is possible that DDX3 association with lncRNAs reflects binding related to scanning through the transcriptome for start codons by the ribosome, with which the protein associates. In such a scenario, binding to lncRNAs would not reflect regulatory roles related to non-coding transcripts and would be just a side effect of ribosome activity. It is also possible that DDX3 actually participates in translation of transcripts generally considered as non-coding. It has been reported that many transcripts considered as lncRNAs associate with ribosomes and give rise to peptides, which however may be largely non-functional (Ji et al., 2015; Bazin et al., 2017). However, ribosomes function in cytoplasm and most human lncRNAs are localized to the nucleus. This fact suggests that lncRNA-DDX3 interactions are most likely independent of ribosomes. Interestingly we have observed that DDX3 binding to lncRNAs along transcript bodies exhibits a similar pattern as binding to mRNA, with highest peaks occurring in the 5’ end of transcripts (Fig. 17). Regulatory relevance of this binding remains unknown. The most obvious possibility is that DDX3 utilizes its helicase activity to resolve RNA secondary structures in 5’ ends of lncRNAs to regulate their interactions with other proteins.

5.3.2. DDX3-mRNA binding pattern suggests a role in regulation of translation initiation

We have determined that DDX3 binds most of the expressed mRNAs and that this binding preferentially takes place in transcripts 5' UTRs around start codons. The observed binding pattern along the transcript body is very similar to that of eIF3B (Fig. 15A). EIF3B is part of translation initiation complex that has been reported to directly interact with DDX3 (Lee et al., 2008). Those results suggest that DDX3 regulatory function on protein coding transcripts is centered on their 5' UTRs and supports findings from other studies, which reported DDX3 function in regulation of translation initiation through interactions with proteins belonging to the translation initiation complex (Lee et al., 2008; Soto-Rifo et al., 2012). Association of DDX3 with the translation initiation machinery is further supported by the observation that the protein binds directly to the 40S ribosome (Fig. 15D). The binding site was detected on helix 16 of the 18S rRNA, which is similar to binding sites identified in previous studies (Valentin-Vega et al., 2016; Oh et al., 2016). The position of the binding is compatible with interactions between DDX3 and translation initiation factors (Lee et al., 2008). It is also in line with observations from experiments performed in yeast, where the DDX3 ortholog Ded1 cooperates with Translation Initiation Factor 3 (TIF3), an ortholog of human eIF4B, in translation initiation (Sen et al., 2016).

5.3.3. DDX3 depletion affects mRNA abundance most likely through an indirect mechanism

Our investigation of the effects of DDX3 depletion using siRNA pools revealed subsets of transcripts that are affected on the level of their abundance as well as translation efficiency. Changes in mRNA levels were more commonly observed than changes in translation efficiency (Table 1). One explanation of this is that it may be caused indirectly, through changes of expression of proteins known to influence mRNA abundance. Among the genes displaying decreased translation efficiency upon DDX3 depletion we have found some involved in the process of transcription, such as E2F1, E2F4, TCF3, HES6, and ELK1. Changes of their level could indeed influence abundance of mRNAs in cells. Another possible explanation is that DDX3 is involved directly in regulation of mRNA stability. This potential mode of action of the protein is, however, unexplored and remains a matter of speculation, as it has not been addressed in this work or up to our knowledge in any other studies.

5.3.4. DDX3 regulates translation of mRNAs with structured 5' UTRs

The number of transcripts, whose translation efficiency is affected by DDX3 depletion, was small when compared to the number of transcripts identified as bound by the protein (Table 2). The widespread DDX3 binding to mRNAs may reflect this protein's general association with ribosomes initiating translation. The protein may associate with ribosomes regardless of whether it is necessary for translation of a given transcript. This potential scenario would explain why the DDX3 depletion only significantly affects translation of specific mRNA targets. This assumption is in line with previous studies, which reported that DDX3 can associate with translation initiation factors directly, in absence of RNA (Lee et al., 2008). We have observed that most of the transcripts affected by DDX3 depletion on the level of translation were identified as bound by the protein in the PAR-CLIP experiment (Table 2). Our observation that DDX3 depletion results in specific shift of ribosomes on those transcripts from coding sequences to 5' UTRs (Fig. 14A) is in line with DDX3 role in translation initiation as accumulation of ribosomes in 5' UTRs most likely reflects impairment of this process.

DDX3 has been reported in previous studies to facilitate translation initiation of specific mRNAs with structured 5' UTRs (Soto-Rifo et al., 2012). Through our transcriptome-wide approaches, we have determined that this regulation is more widespread than previously shown. We have established a direct link between lowered translation efficiency of mRNAs upon DDX3 depletion and levels of GC in their 5' UTRs, which reflects a high potential to form secondary structures (Fig. 14C and Fig. 14D). Our observations led us to believe that DDX3 most likely controls translation initiation on mRNAs harboring complex 5' UTRs. To further validate this notion, we have again looked into DDX3 binding sites established with PAR-CLIP and determined that high confidence protein binding takes place in GC rich regions (Fig. 16A) and that it is specifically increased in 5' UTRs of mRNAs displaying lower translation efficiency upon DDX3 depletion (Fig. 16B). We believe that those findings strongly indicate that the major function of DDX3 in human cells is to facilitate translation initiation on transcripts harboring secondary RNA structures in their 5' UTRs, most likely through its helicase activity. This mode of action could play a part in regulation of DDR, in which DDX3 displays increased binding to polyadenylated RNAs (Milek et al., 2017). It is possible that increased binding of the protein to RNAs in DDR reflects its direct regulatory role in translation initiation of upregulated DDR genes. In line with this assumption we have

observed in our experiments that a number of genes involved in regulation of DDR was bound by DDX3 and displayed lowered TE upon DDX3 depletion.

5.3.5. DDX3 may have opposing roles in regulating translation initiation

In our analysis of translation efficiency upon DDX3 knockdown, we have observed a small subset of transcripts bound by DDX3, which displayed an increase in their translation efficiency upon the protein knock-down (Table 2). This effect is not in line with the major role of DDX3 our work has defined as promoting the translation initiation of transcripts with structured 5' UTRs. It is possible that identification of small number of transcripts displaying increase in translation efficiency is simply an artifact of our analysis and does not reflect direct role of DDX3 in negative regulation of translation. In line with this assumption is the fact that we have detected mRNAs coding for histone proteins in this group of transcripts (Fig. 13A). Histone mRNAs do not bear a typical 3' end and do not have polyA tails (Marzluff et al., 2017). We have determined mRNA abundance used for calculation of translation efficiency using polyadenylated fraction of transcripts. Thus, observed changes in translation efficiency of histone mRNAs may be a result of an experimental bias and require a more tailored approach for verification. The observed increase in translation efficiency may, however, also be a result of direct regulation mediated by the protein. In one previous study, DDX3 has been shown to negatively regulate cap-dependent translation of a specific mRNA both in vivo and in vitro in a luciferase reporter assay (Shih et al., 2008). In a later study that mechanism was also linked to DDX3 mediated formation of stress granules in which DDX3 is involved through its interactions with translation initiation factor independently of its ATPase and helicase activities (Shih et al., 2012). It seems possible that DDX3 has dual roles in binding to transcripts 5' UTRs, facilitating translation initiation on targets bearing secondary RNA structures and under specific circumstances inhibiting translation initiation on others. Presence of DDX3 on 5' UTRs of most expressed transcripts, which are not effected by DDX3 knock-down could reflect function of the protein in providing readiness of translation initiation machinery to quickly inhibit translation in face of potentially dangerous stresses. A secondary function of DDX3 as a conditional translation inhibitor could also explain our aforementioned observation that some mRNAs bound by this protein displayed lowered translation efficiency upon its depletion. This hypothesis provides an interesting potential follow-up to results presented in this work. Identification of changes of DDX3 binding to mRNAs upon stress induced formation of SGs and of associated changes in translation efficiency could provide important new insight into DDX3 functions.

5.3.6. Limitations of the approach used to study DDX3 functions

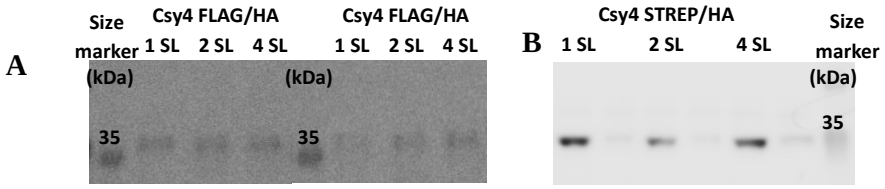
Functional study of DDX3 is made more difficult by the fact that it is an essential protein in eukaryotic cells. Traditional protein knock-outs cannot be studied due to loss of cells viability. Knock-down approaches, like the one used in this study, are associated with a certain level of the protein of interest still being expressed in cells (Fig. 12B). Those residual protein levels can obscure the full picture of regulatory roles played by the protein in knock-down experiments. Moreover, the time required to achieve efficient knock-down of a protein with siRNAs provides cells with a possibility to adapt to gradual depletion of that protein. It cannot be ruled out that such an adaptation took place in case of our siRNA knock-down of DDX3. In this scenario full picture of regulatory roles played by DDX3 could be obscured by activity of proteins having some degree of functional redundancy to DDX3. To overcome those issues, we have attempted to develop stable cell lines allowing for a transient and quick (within hours) DDX3 knock-out using the auxin induced degron system. Those efforts, however, were not successful and further attempts to create an experimental set up to study effects of complete DDX3 depletion are warranted.

Taken together we were able to gain important new insights into regulatory roles of DDX3. Our results establish that the major role of the protein is regulating translation initiation through binding to ribosomes and 5’ UTRs of target mRNAs harboring RNA secondary structures. This mode of action is similar to that of DDX3 yeast ortholog Ded1, with a marked difference that Ded1 regulates translation initiation of a much broader fraction of the transcriptome than DDX3 (Chuang et al. 1997; de la Cruz et al. 1997).

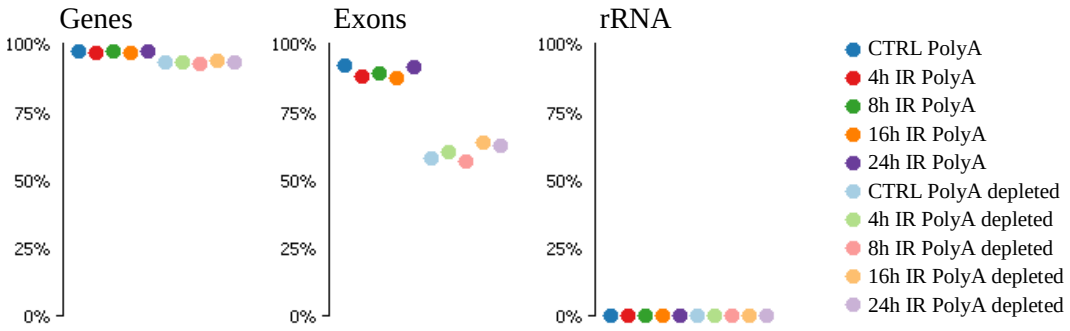
6. Supplementary tables and figures

Supplementary Table 1. List of lncRNAs differentially expressed in different time points after induction of IR, which were identified using MCF-7 specific lncRNA annotation (Sun et al., 2015) and were not annotated in Gencode.v19 genome annotation. Averaged FPKM values from two biological replicates are shown.

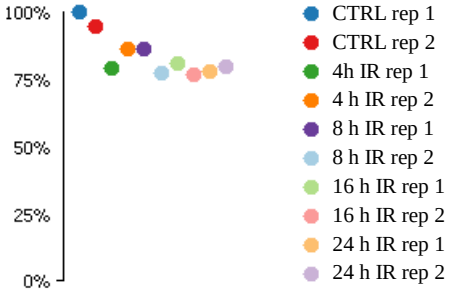
Gene name	Genomic coordinates	CTRL FPKM	4h IR FPKM	8h IR FPKM	16h IR FPKM	24h IR FPKM
LncRNA231:1	chr3:46686311-46702478	0.7	5.9	2.9	2.2	1.9
LncRNA1868:1	chr9:138999328-139000876	1.7	2.3	4.2	6.0	5.6
LncRNA1819:1	chr7:155000588-155010565	0.4	3.2	3.4	2.7	2.0
LncRNA1548:1	chr11:77930120-77936399	1.8	3.3	4.0	8.3	5.3
LncRNA641:1	chr18:23747816-23751320	3.5	2.6	2.3	1.2	0.9



Supplementary Figure 1. Expression of tagged Csy4 protein in HEK293 cells. A – Western blot image with HA antibodies showing levels of Csy4 protein after transfection with plasmid encoding FLAG/HA tagged Csy4 protein (predicted molecular weight of FLAG/HA tagged Csy4 protein is 26.67 kDa), with different amounts of Csy4 target stem loops (SL). B – Western blot image with HA antibodies showing levels of Csy4 protein after transfection with plasmid encoding STREP/HA tagged Csy4 protein (predicted molecular weight of STREP/HA tagged Csy4 protein is 29.28 kDa), with different amounts of Csy4 target stem loops



Supplementary Figure 2. Percentages of all sequenced reads mapping to genes. exons and rRNA in sequencing experiments performed for polyA enriched fraction of transcripts and fraction of transcripts after sequential depletion of polyadenylated and ribosomal RNAs. Shown are average results of two biological replicates per condition.



Supplementary Figure 3. Differences in relative number of mapped reads obtained in each biological replicate of sequencing experiment of transcripts depleted of polyadenylated and ribosomal RNAs (100% represents the library with the highest number of mapped reads).

7. Bibliography

Abe, N., Matsumoto, K., Nishihara, M., Nakano, Y., Shibata, A., Maruyama, H., Shuto, S., Matsuda, A., Yoshida, M., Ito, Y., Abe, H., 2015. Rolling Circle Translation of Circular RNA in Living Human Cells. Sci Rep 5, 16435. https://doi.org/10.1038/srep16435

Adamson, B., Smogorzewska, A., Sigoillot, F.D., King, R.W., Elledge, S.J., 2012. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat. Cell Biol. 14, 318–328. https://doi.org/10.1038/ncb2426

Affymetrix ENCODE Transcriptome Project, Cold Spring Harbor Laboratory ENCODE Transcriptome Project, 2009. Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. Nature 457, 1028–1032. https://doi.org/10.1038/nature07759

Ahn, J.Y., Schwarz, J.K., Piwnica-Worms, H., Canman, C.E., 2000. Threonine 68 phosphorylation by ataxia telangiectasia mutated is required for efficient activation of Chk2 in response to ionizing radiation. Cancer Res. 60, 5934–5936.

Aitken, C.E., Beznosková, P., Vlčkova, V., Chiu, W.-L., Zhou, F., Valášek, L.S., Hinnebusch, A.G., Lorsch, J.R., 2016. Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex. Elife 5. https://doi.org/10.7554/eLife.20934

Amaral, P.P., Mattick, J.S., 2008. Noncoding RNA in development. Mamm. Genome 19, 454–492. https://doi.org/10.1007/s00335-008-9136-7

Anderson, D.M., Anderson, K.M., Chang, C.-L., Makarewich, C.A., Nelson, B.R., McAnally, J.R., Kasaragod, P., Shelton, J.M., Liou, J., Bassel-Duby, R., Olson, E.N., 2015. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell 160, 595–606. https://doi.org/10.1016/j.cell.2015.01.009

Arun, G., Diermeier, S.D., Spector, D.L., 2018. Therapeutic Targeting of Long Non-Coding RNAs in Cancer. Trends in Molecular Medicine 24, 257–277. https://doi.org/10.1016/j.molmed.2018.01.001

Ashwal-Fluss, R., Meyer, M., Pamudurti, N.R., Ivanov, A., Bartok, O., Hanan, M., Evantal, N., Memczak, S., Rajewsky, N., Kadener, S., 2014. circRNA biogenesis competes with pre-mRNA splicing. Mol. Cell 56, 55–66. https://doi.org/10.1016/j.molcel.2014.08.019

Awwad, S.W., Abu-Zhayia, E.R., Guttmann-Raviv, N., Ayoub, N., 2017. NELF-E is recruited to DNA double-strand break sites to promote transcriptional repression and repair. EMBO Rep. 18, 745–764. https://doi.org/10.15252/embr.201643191

Bachler, M., Schroeder, R., von Ahsen, U., 1999. StreptoTag: a novel method for the isolation of RNA-binding proteins. RNA 5, 1509–1516.

Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., Landthaler, M., 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell 46, 674–690. https://doi.org/10.1016/j.molcel.2012.05.021

Banin, S., Moyal, L., Shieh, S.-Y., Taya, Y., Anderson, C.W., Chessa, L., Smorodinsky, N.I., Prives, C., Reiss, Y., Shiloh, Y., Ziv, Y., 1998. Enhanced Phosphorylation of p53 by ATM in Response to DNA Damage. Science 281, 1674–1677. https://doi.org/10.1126/science.281.5383.1674

Barnes, C., Kanhere, A., 2016. Identification of RNA-Protein Interactions Through In Vitro RNA Pull-Down Assays. Methods Mol. Biol. 1480, 99–113. https://doi.org/10.1007/978-1-4939-6380-5_9

Barry, G., Briggs, J.A., Vanichkina, D.P., Poth, E.M., Beveridge, N.J., Ratnu, V.S., Nayler, S.P., Nones, K., Hu, J., Bredy, T.W., Nakagawa, S., Rigo, F., Taft, R.J., Cairns, M.J., Blackshaw, S., Wolvetang, E.J., Mattick, J.S., 2014. The long non-coding RNA Gomafu is acutely regulated in response to neuronal activation and involved in schizophrenia-associated alternative splicing. Mol. Psychiatry 19, 486–494. https://doi.org/10.1038/mp.2013.45

Bazin, J., Baerenfaller, K., Gosai, S.J., Gregory, B.D., Crespi, M., Bailey-Serres, J., 2017. Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. PNAS 114, E10018–E10027. https://doi.org/10.1073/pnas.1708433114

Bekker-Jensen, S., Mailand, N., 2010. Assembly and function of DNA double-strand break repair foci in mammalian cells. DNA Repair (Amst.) 9, 1219–1228. https://doi.org/10.1016/j.dnarep.2010.09.010

Bennett, C.B., Lewis, A.L., Baldwin, K.K., Resnick, M.A., 1993. Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid. Proceedings of the National Academy of Sciences 90, 5613–5617. https://doi.org/10.1073/pnas.90.12.5613

Beucher, A., Birraux, J., Tchouandong, L., Barton, O., Shibata, A., Conrad, S., Goodarzi, A.A., Krempler, A., Jeggo, P.A., Löbrich, M., 2009. ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. The EMBO Journal 28, 3413–3427. https://doi.org/10.1038/emboj.2009.276

Bol, Guus M., Vesuna, F., Xie, M., Zeng, J., Aziz, K., Gandhi, N., Levine, A., Irving, A., Korz, D., Tantravedi, S., Voss, M.R.H. van, Gabrielson, K., Bordt, E.A., Polster, B.M., Cope, L., Groep, P. van der, Kondaskar, A., Rudek, M.A., Hosmane, R.S., Wall, E. van der, Diest, P.J. van, Tran, P.T., Raman, V., 2015. Targeting DDX3 with a small molecule inhibitor for lung cancer therapy. EMBO Molecular Medicine 7, 648–669. https://doi.org/10.15252/emmm.201404368

Bol, Guus Martinus, Xie, M., Raman, V., 2015. DDX3, a potential target for cancer treatment. Molecular Cancer 14, 188. https://doi.org/10.1186/s12943-015-0461-7

Bolha, L., Ravnik-Glavač, M., Glavač, D., 2017a. Circular RNAs: Biogenesis, Function, and a Role as Possible Cancer Biomarkers. Int J Genomics 2017. https://doi.org/10.1155/2017/6218353

Bolha, L., Ravnik-Glavač, M., Glavač, D., 2017b. Long Noncoding RNAs as Biomarkers in Cancer. Dis Markers 2017. https://doi.org/10.1155/2017/7243968

Bond, C.S., Fox, A.H., 2009. Paraspeckles: nuclear bodies built on long noncoding RNA. J. Cell Biol. 186, 637–644. https://doi.org/10.1083/jcb.200906113

Borchardt, E.K., Vandoros, L.A., Huang, M., Lackey, P.E., Marzluff, W.F., Asokan, A., 2015. Controlling mRNA stability and translation with the CRISPR endoribonuclease Csy4. RNA 21, 1921–1930. https://doi.org/10.1261/rna.051227.115

Bothmer, A., Robbiani, D.F., Feldhahn, N., Gazumyan, A., Nussenzweig, A., Nussenzweig, M.C., 2010. 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. Journal of Experimental Medicine 207, 855–865. https://doi.org/10.1084/jem.20100244

Botlagunta, M., Vesuna, F., Mironchik, Y., Raman, A., Lisok, A., Winnard, P., Mukadam, S., Van Diest, P., Chen, J.H., Farabaugh, P., Patel, A.H., Raman, V., 2008. Oncogenic role of DDX3 in breast cancer biogenesis. Oncogene 27, 3912–3922. https://doi.org/10.1038/onc.2008.33

Brannan, C.I., Dees, E.C., Ingram, R.S., Tilghman, S.M., 1990. The product of the H19 gene may function as an RNA. Mol. Cell. Biol. 10, 28–36.

Branzei, D., Foiani, M., 2005. The DNA damage response during DNA replication. Current Opinion in Cell Biology, Cell division, growth and death / Cell differentiation 17, 568–575. https://doi.org/10.1016/j.ceb.2005.09.003

Braun, C.J., Zhang, X., Savelyeva, I., Wolff, S., Moll, U.M., Schepeler, T., Ørntoft, T.F., Andersen, C.L., Dobbelstein, M., 2008. p53-Responsive MicroRNAs 192 and 215 Are Capable of Inducing Cell Cycle Arrest. Cancer Res 68, 10094–10104. https://doi.org/10.1158/0008-5472.CAN-08-1569

Braun, S., Domdey, H., Wiebauer, K., 1996. Inverse splicing of a discontinuous pre-mRNA intron generates a circular exon in a HeLa cell nuclear extract. Nucleic Acids Res. 24, 4152–4157.

Broadbent, H.M., Peden, J.F., Lorkowski, S., Goel, A., Ongen, H., Green, F., Clarke, R., Collins, R., Franzosi, M.G., Tognoni, G., Seedorf, U., Rust, S., Eriksson, P., Hamsten, A., Farrall, M., Watkins, H., PROCARDIS consortium, 2008. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. Hum. Mol. Genet. 17, 806–814. https://doi.org/10.1093/hmg/ddm352

Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafrenière, R.G., Xing, Y., Lawrence, J., Willard, H.F., 1992. The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell 71, 527–542. https://doi.org/10.1016/0092-8674(92)90520-M

Buchan, J.R., Parker, R., 2009. Eukaryotic Stress Granules: The Ins and Out of Translation. Mol Cell 36, 932. https://doi.org/10.1016/j.molcel.2009.11.020

Bunting, S.F., Callén, E., Wong, N., Chen, H.-T., Polato, F., Gunn, A., Bothmer, A., Feldhahn, N., Fernandez-Capetillo, O., Cao, L., Xu, X., Deng, C.-X., Finkel, T., Nussenzweig, M., Stark, J.M., Nussenzweig, A., 2010. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. Cell 141, 243–254. https://doi.org/10.1016/j.cell.2010.03.012

Burd, C.E., Jeck, W.R., Liu, Y., Sanoff, H.K., Wang, Z., Sharpless, N.E., 2010. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. PLoS Genet. 6, e1001233. https://doi.org/10.1371/journal.pgen.1001233

Burma, S., Chen, B.P., Murphy, M., Kurimasa, A., Chen, D.J., 2001. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. J. Biol. Chem. 276, 42462–42467. https://doi.org/10.1074/jbc.C100466200

Canman, C.E., Lim, D.-S., Cimprich, K.A., Taya, Y., Tamai, K., Sakaguchi, K., Appella, E., Kastan, M.B., Siliciano, J.D., 1998. Activation of the ATM Kinase by Ionizing Radiation and Phosphorylation of p53. Science 281, 1677–1679. https://doi.org/10.1126/science.281.5383.1677

Cannan, W.J., Pederson, D.S., 2016. Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin. J. Cell. Physiol. 231, 3–14. https://doi.org/10.1002/jcp.25048

Cannell, I.G., Kong, Y.W., Johnston, S.J., Chen, M.L., Collins, H.M., Dobbyn, H.C., Elia, A., Kress, T.R., Dickens, M., Clemens, M.J., Heery, D.M., Gaestel, M., Eilers, M., Willis, A.E., Bushell, M., 2010. p38 MAPK/MK2-mediated induction of miR-34c following DNA damage prevents Myc-dependent DNA replication. PNAS 107, 5375–5380. https://doi.org/10.1073/pnas.0910015107

Capel, B., Swain, A., Nicolis, S., Hacker, A., Walter, M., Koopman, P., Goodfellow, P., Lovell-Badge, R., 1993. Circular transcripts of the testis-determining gene Sry in adult mouse testis. Cell 73, 1019–1030.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impiombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary,

V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminiecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P.T., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavesi, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schönbach, C., Sekiguchi, K., Semple, C. a. M., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiura, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessy, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J., Hayashizaki, Y., FANTOM Consortium, RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group), 2005. The transcriptional landscape of the mammalian genome. Science 309, 1559–1563. https://doi.org/10.1126/science.1112014

Cawley, S., Bekiranov, S., Ng, H.H., Kapranov, P., Sekinger, E.A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A.J., Wheeler, R., Wong, B., Drenkow, J., Yamanaka, M., Patel, S., Brubaker, S., Tammana, H., Helt, G., Struhl, K., Gingeras, T.R., 2004. Unbiased Mapping of Transcription Factor Binding Sites along Human Chromosomes 21 and 22 Points to Widespread Regulation of Noncoding RNAs. Cell 116, 499–509. https://doi.org/10.1016/S0092-8674(04)00127-8

Chao, C.-H., Chen, C.-M., Cheng, P.-L., Shih, J.-W., Tsou, A.-P., Lee, Y.-H.W., 2006. DDX3, a DEAD box RNA helicase with tumor growth-suppressive property and transcriptional regulation activity of the p21waf1/cip1 promoter, is a candidate tumor suppressor. Cancer Res. 66, 6579–6588. https://doi.org/10.1158/0008-5472.CAN-05-2415

Chen, E., Sharma, M.R., Shi, X., Agrawal, R.K., Joseph, S., 2014. Fragile X mental retardation protein regulates translation by binding directly to the ribosome. Mol. Cell 54, 407–417. https://doi.org/10.1016/j.molcel.2014.03.023

Chen, J.-F., Mandel, E.M., Thomson, J.M., Wu, Q., Callis, T.E., Hammond, S.M., Conlon, F.L., Wang, D.-Z., 2006. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat. Genet. 38, 228–233. https://doi.org/10.1038/ng1725

Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., Chang, H.Y., 2011. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol. Cell 44, 667–678. https://doi.org/10.1016/j.molcel.2011.08.027

Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., Chang, H.Y., 2015. Systematic discovery of Xist RNA binding proteins. Cell 161, 404–416. https://doi.org/10.1016/j.cell.2015.03.025

Chuang, R.Y., Weaver, P.L., Liu, Z., Chang, T.H., 1997. Requirement of the DEAD-Box protein ded1p for messenger RNA translation. Science 275, 1468–1471.

Cocquerelle, C., Daubersies, P., Majérus, M.A., Kerckaert, J.P., Bailleul, B., 1992. Splicing with inverted order of exons occurs proximal to large introns. EMBO J. 11, 1095–1098.

Cocquerelle, C., Mascrez, B., Hétuin, D., Bailleul, B., 1993. Mis-splicing yields circular RNA molecules. FASEB J. 7, 155–160.

Conn, S.J., Pillman, K.A., Toubia, J., Conn, V.M., Salmanidis, M., Phillips, C.A., Roslan, S., Schreiber, A.W., Gregory, P.A., Goodall, G.J., 2015. The RNA binding protein quaking regulates formation of circRNAs. Cell 160, 1125–1134. https://doi.org/10.1016/j.cell.2015.02.014

Crappé, J., Ndah, E., Koch, A., Steyaert, S., Gawron, D., De Keulenaer, S., De Meester, E., De Meyer, T., Van Criekinge, W., Van Damme, P., Menschaert, G., 2015. PROTEOFORMER: deep proteome coverage through ribosome profiling and MS integration. Nucleic Acids Res. 43, e29. https://doi.org/10.1093/nar/gku1283

Czech, B., Hannon, G.J., 2011. Small RNA sorting: matchmaking for Argonautes. Nature Reviews Genetics 12, 19–31. https://doi.org/10.1038/nrg2916

de la Cruz, J., Iost, I., Kressler, D., Linder, P., 1997. The p20 and Ded1 proteins have antagonistic roles in eIF4E-dependent translation in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A. 94, 5201–5206.

Denli, A.M., Tops, B.B.J., Plasterk, R.H.A., Ketting, R.F., Hannon, G.J., 2004. Processing of primary microRNAs by the Microprocessor complex. Nature 432, 231–235. https://doi.org/10.1038/nature03049

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., Lagarde, J., Veeravalli, L., Ruan, X., Ruan, Y., Lassmann, T., Carninci, P., Brown, J.B., Lipovich, L., Gonzalez, J.M., Thomas, M., Davis, C.A., Shiekhattar, R., Gingeras, T.R., Hubbard, T.J., Notredame, C., Harrow, J., Guigó, R., 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene

structure, evolution, and expression. Genome Res. 22, 1775–1789. https://doi.org/10.1101/gr.132159.111

Deveson, I.W., Hardwick, S.A., Mercer, T.R., Mattick, J.S., 2017a. The Dimensions, Dynamics, and Relevance of the Mammalian Noncoding Transcriptome. Trends Genet. 33, 464–478. https://doi.org/10.1016/j.tig.2017.04.004

Deveson, I.W., Hardwick, S.A., Mercer, T.R., Mattick, J.S., 2017b. The Dimensions, Dynamics, and Relevance of the Mammalian Noncoding Transcriptome. Trends in Genetics 33, 464–478. https://doi.org/10.1016/j.tig.2017.04.004

Diederichs, S., Haber, D.A., 2007. Dual Role for Argonautes in MicroRNA Processing and Posttranscriptional Regulation of MicroRNA Expression. Cell 131, 1097–1108. https://doi.org/10.1016/j.cell.2007.10.032

D'Lima, N.G., Ma, J., Winkler, L., Chu, Q., Loh, K.H., Corpuz, E.O., Budnik, B.A., Lykke-Andersen, J., Saghatelian, A., Slavoff, S.A., 2017. A human microprotein that interacts with the mRNA decapping complex. Nat. Chem. Biol. 13, 174–180. https://doi.org/10.1038/nchembio.2249

Dodt, M., Roehr, J.T., Ahmed, R., Dieterich, C., 2012. FLEXBAR—Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. Biology (Basel) 1, 895–905. https://doi.org/10.3390/biology1030895

Dohm, J.C., Lottaz, C., Borodina, T., Himmelbauer, H., 2008. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res 36, e105. https://doi.org/10.1093/nar/gkn425

Dong, Y., Yang, J., Ye, W., Wang, Y., Ye, C., Weng, D., Gao, H., Zhang, F., Xu, Z., Lei, Y., 2015. Isolation of Endogenously Assembled RNA-Protein Complexes Using Affinity Purification Based on Streptavidin Aptamer S1. Int J Mol Sci 16, 22456–22472. https://doi.org/10.3390/ijms160922456

Du, W.W., Yang, W., Liu, E., Yang, Z., Dhaliwal, P., Yang, B.B., 2016. Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. Nucleic Acids Res. 44, 2846–2858. https://doi.org/10.1093/nar/gkw027

Du, Z., Sun, T., Hacisuleyman, E., Fei, T., Wang, X., Brown, M., Rinn, J.L., Lee, M.G.-S., Chen, Y., Kantoff, P.W., Liu, X.S., 2016. Integrative analyses reveal a long noncoding RNA-mediated sponge regulatory network in prostate cancer. Nature Communications 7, 10982. https://doi.org/10.1038/ncomms10982

Engreitz, J.M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E.S., Plath, K., Guttman, M., 2013. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X-chromosome. Science 341, 1237973. https://doi.org/10.1126/science.1237973

Epling, L.B., Grace, C.R., Lowe, B.R., Partridge, J.F., Enemark, E.J., 2015. Cancer-associated mutants of RNA helicase DDX3X are defective in RNA-stimulated ATP hydrolysis. J. Mol. Biol. 427, 1779–1796. https://doi.org/10.1016/j.jmb.2015.02.015

Fabian, M.R., Sonenberg, N., 2012. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. Nature Structural & Molecular Biology 19, 586–593. https://doi.org/10.1038/nsmb.2296

Floor, S.N., Condon, K.J., Sharma, D., Jankowsky, E., Doudna, J.A., 2016. Autoinhibitory Interdomain Interactions and Subfamily-specific Extensions Redefine the Catalytic Core of the Human DEAD-box Protein DDX3. J. Biol. Chem. 291, 2412–2421. https://doi.org/10.1074/jbc.M115.700625

Fradet, Y., Saad, F., Aprikian, A., Dessureault, J., Elhilali, M., Trudel, C., Mâsse, B., Piché, L., Chypre, C., 2004. uPM3, a new molecular urine test for the detection of prostate cancer. Urology 64, 311–315; discussion 315-316. https://doi.org/10.1016/j.urology.2004.03.052

Francia, S., Michelini, F., Saxena, A., Tang, D., de Hoon, M., Anelli, V., Mione, M., Carninci, P., d'Adda di Fagagna, F., 2012. Site-specific DICER and DROSHA RNA products control the DNA damage response. Nature 488, 231–235. https://doi.org/10.1038/nature11179

Geissler, R., Golbik, R.P., Behrens, S.-E., 2012. The DEAD-box helicase DDX3 supports the assembly of functional 80S ribosomes. Nucleic Acids Res. 40, 4998–5011. https://doi.org/10.1093/nar/gks070

Gerner, E.W., Meyskens Jr, F.L., 2004. Polyamines and cancer: old molecules, new understanding. Nature Reviews Cancer 4, 781–792. https://doi.org/10.1038/nrc1454

Glatter, T., Wepf, A., Aebersold, R., Gstaiger, M., 2009. An integrated workflow for charting the human interaction proteome: insights into the PP2A system. Mol Syst Biol 5, 237. https://doi.org/10.1038/msb.2008.75

Glažar, P., Papavasileiou, P., Rajewsky, N., 2014. circBase: a database for circular RNAs. RNA 20, 1666–1670. https://doi.org/10.1261/rna.043687.113

Gonzalez, I., Munita, R., Agirre, E., Dittmer, T.A., Gysling, K., Misteli, T., Luco, R.F., 2015. A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature. Nat. Struct. Mol. Biol. 22, 370–376. https://doi.org/10.1038/nsmb.3005

Greenberg, R.A., Sobhian, B., Pathania, S., Cantor, S.B., Nakatani, Y., Livingston, D.M., 2006. Multifactorial contributions to an acute DNA damage response by BRCA1/BARD1-containing complexes. Genes Dev 20, 34–46. https://doi.org/10.1101/gad.1381306

Greene, J., Baird, A.-M., Brady, L., Lim, M., Gray, S.G., McDermott, R., Finn, S.P., 2017. Circular RNAs: Biogenesis, Function and Role in Human Diseases. Front Mol Biosci 4. https://doi.org/10.3389/fmolb.2017.00038

Gregersen, L.H., Schueler, M., Munschauer, M., Mastrobuoni, G., Chen, W., Kempa, S., Dieterich, C., Landthaler, M., 2014. MOV10 Is a 5' to 3' RNA helicase contributing to UPF1 mRNA target degradation by translocation along 3' UTRs. Mol. Cell 54, 573–585. https://doi.org/10.1016/j.molcel.2014.03.017

Gregory, R.I., Yan, K.-P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., Shiekhattar, R., 2004. The Microprocessor complex mediates the genesis of microRNAs. Nature 432, 235–240. https://doi.org/10.1038/nature03120

Guenther, U.-P., Weinberg, D.E., Zubradt, M.M., Tedeschi, F.A., Stawicki, B.N., Zagore, L.L., Brar, G.A., Licatalosi, D.D., Bartel, D.P., Weissman, J.S., Jankowsky, E., 2018. The helicase Ded1p controls use of near-cognate translation initiation codons in 5′ UTRs. Nature 559, 130. https://doi.org/10.1038/s41586-018-0258-0

Guo, J.U., Agarwal, V., Guo, H., Bartel, D.P., 2014. Expanded identification and characterization of mammalian circular RNAs. Genome Biol. 15, 409. https://doi.org/10.1186/s13059-014-0409-z

Gupta, N., Lorsch, J.R., Hinnebusch, A.G., 2018. Yeast Ded1 promotes 48S translation pre-initiation complex assembly in an mRNA-specific and eIF4F-dependent manner. Elife 7. https://doi.org/10.7554/eLife.38892

Gutschner, T., Hämmerle, M., Eissmann, M., Hsu, J., Kim, Y., Hung, G., Revenko, A., Arun, G., Stentrup, M., Gross, M., Zörnig, M., MacLeod, A.R., Spector, D.L., Diederichs, S., 2013. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. Cancer Res. 73, 1180–1189. https://doi.org/10.1158/0008-5472.CAN-12-2850

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., Cabili, M.N., Jaenisch, R., Mikkelsen, T.S., Jacks, T., Hacohen, N., Bernstein, B.E., Kellis, M., Regev, A., Rinn, J.L., Lander, E.S., 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458, 223–227. https://doi.org/10.1038/nature07672

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., Ulrich, A., Wardle, G.S., Dewell, S., Zavolan, M., Tuschl, T., 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141, 129–141. https://doi.org/10.1016/j.cell.2010.03.009

Halbreich, A., Pajot, P., Foucher, M., Grandchamp, C., Slonimski, P., 1980. A pathway of cytochrome b mRNA processing in yeast mitochondria: specific splicing steps and an intron-derived circular DNA. Cell 19, 321–329.

Hämmerle, M., Gutschner, T., Uckelmann, H., Ozgur, S., Fiskin, E., Gross, M., Skawran, B., Geffers, R., Longerich, T., Breuhahn, K., Schirmacher, P., Stoecklin, G., Diederichs, S., 2013. Posttranscriptional destabilization of the liver-specific long noncoding RNA HULC

by the IGF2 mRNA-binding protein 1 (IGF2BP1). Hepatology 58, 1703–1712. https://doi.org/10.1002/hep.26537

Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., Sohn, S.Y., Cho, Y., Zhang, B.-T., Kim, V.N., 2006. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell 125, 887–901. https://doi.org/10.1016/j.cell.2006.03.043

Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., Kjems, J., 2013. Natural RNA circles function as efficient microRNA sponges. Nature 495, 384–388. https://doi.org/10.1038/nature11993

Hartmuth, K., Urlaub, H., Vornlocher, H.-P., Will, C.L., Gentzel, M., Wilm, M., Lührmann, R., 2002. Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. Proc. Natl. Acad. Sci. U.S.A. 99, 16719–16724. https://doi.org/10.1073/pnas.262483899

Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y., Takahashi, T., 2005. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res. 65, 9628–9632. https://doi.org/10.1158/0008-5472.CAN-05-2352

He, L., He, X., Lim, L.P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D., Jackson, A.L., Linsley, P.S., Chen, C., Lowe, S.W., Cleary, M.A., Hannon, G.J., 2007. A microRNA component of the p53 tumour suppressor network. Nature 447, 1130–1134. https://doi.org/10.1038/nature05939

He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J., Hammond, S.M., 2005. A microRNA polycistron as a potential human oncogene. Nature 435, 828–833. https://doi.org/10.1038/nature03552

He, Y., Zhang, D., Yang, Y., Wang, X., Zhao, X., Zhang, P., Zhu, H., Xu, N., Liang, S., 2018. A double-edged function of DDX3, as an oncogene or tumor suppressor, in cancer progression (Review). Oncol. Rep. 39, 883–892. https://doi.org/10.3892/or.2018.6203

Heerma van Voss, M.R., Vesuna, F., Trumpi, K., Brilliant, J., Berlinicke, C., de Leng, W., Kranenburg, O., Offerhaus, G.J., Bürger, H., van der Wall, E., van Diest, P.J., Raman, V., 2015. Identification of the DEAD box RNA helicase DDX3 as a therapeutic target in colorectal cancer. Oncotarget 6, 28312–28326. https://doi.org/10.18632/oncotarget.4873

Herdy, B., Mayer, C., Varshney, D., Marsico, G., Murat, P., Taylor, C., D'Santos, C., Tannahill, D., Balasubramanian, S., 2018. Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts. Nucleic Acids Res 46, 11592–11604. https://doi.org/10.1093/nar/gky861

Hoeijmakers, J.H.J., 2009. DNA damage, aging, and cancer. N. Engl. J. Med. 361, 1475–1485. https://doi.org/10.1056/NEJMra0804615

Hogarty, M.D., Norris, M.D., Davis, K., Liu, X., Evageliou, N.F., Hayes, C.S., Pawel, B., Guo, R., Zhao, H., Sekyere, E., Keating, J., Thomas, W., Cheng, N.C., Murray, J., Smith, J., Sutton, R., Venn, N., London, W.B., Buxton, A., Gilmour, S.K., Marshall, G.M., Haber, M., 2008. ODC1 is a critical determinant of MYCN oncogenesis and a therapeutic target in neuroblastoma. Cancer Res. 68, 9735–9745. https://doi.org/10.1158/0008-5472.CAN-07-6866

Holdt, L.M., Stahringer, A., Sass, K., Pichler, G., Kulak, N.A., Wilfert, W., Kohlmaier, A., Herbst, A., Northoff, B.H., Nicolaou, A., Gäbel, G., Beutner, F., Scholz, M., Thiery, J., Musunuru, K., Krohn, K., Mann, M., Teupser, D., 2016. Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. Nat Commun 7, 12429. https://doi.org/10.1038/ncomms12429

Hsiao, K.-Y., Lin, Y.-C., Gupta, S.K., Chang, N., Yen, L., Sun, H.S., Tsai, S.-J., 2017. Noncoding Effects of Circular RNA CCDC66 Promote Colon Cancer Growth and Metastasis. Cancer Res. 77, 2339–2350. https://doi.org/10.1158/0008-5472.CAN-16-1883

Hsu, M.T., Coca-Prados, M., 1979. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. Nature 280, 339–340.

Hu, H., Du, L., Nagabayashi, G., Seeger, R.C., Gatti, R.A., 2010. ATM is down-regulated by N-Myc–regulated microRNA-421. PNAS 107, 1506–1511. https://doi.org/10.1073/pnas.0907763107

Huang, J.-Z., Chen, M., Chen, D., Gao, X.-C., Zhu, S., Huang, H., Hu, M., Zhu, H., Yan, G.-R., 2017. A Peptide Encoded by a Putative lncRNA HOXB-AS3 Suppresses Colon Cancer Growth. Mol. Cell 68, 171-184.e6. https://doi.org/10.1016/j.molcel.2017.09.015

Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., Attardi, L.D., Regev, A., Lander, E.S., Jacks, T., Rinn, J.L., 2010. A Large Intergenic Noncoding RNA Induced by p53 Mediates Global Gene Repression in the p53 Response. Cell 142, 409–419. https://doi.org/10.1016/j.cell.2010.06.040

Huen, M.S.Y., Grant, R., Manke, I., Minn, K., Yu, X., Yaffe, M.B., Chen, J., 2007. RNF8 Transduces the DNA-Damage Signal via Histone Ubiquitylation and Checkpoint Protein Assembly. Cell 131, 901–914. https://doi.org/10.1016/j.cell.2007.09.041

Hung, T., Wang, Yulei, Lin, M.F., Koegel, A.K., Kotake, Y., Grant, G.D., Horlings, H.M., Shah, N., Umbricht, C., Wang, P., Wang, Yu, Kong, B., Langerød, A., Børresen-Dale, A.-L., Kim, S.K., van de Vijver, M., Sukumar, S., Whitfield, M.L., Kellis, M., Xiong, Y., Wong, D.J., Chang, H.Y., 2011. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nature Genetics 43, 621–629. https://doi.org/10.1038/ng.848

Hutvágner, G., McLachlan, J., Pasquinelli, A.E., Bálint, E., Tuschl, T., Zamore, P.D., 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293, 834–838. https://doi.org/10.1126/science.1062961

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., Weissman, J.S., 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324, 218–223. https://doi.org/10.1126/science.1168978

Iwamoto, N., Shimada, T., 2018. Recent advances in mass spectrometry-based approaches for proteomics and biologics: Great contribution for developing therapeutic antibodies. Pharmacol. Ther. 185, 147–154. https://doi.org/10.1016/j.pharmthera.2017.12.007

Iyer, R.R., Pluciennik, A., Burdett, V., Modrich, P.L., 2006. DNA Mismatch Repair: Functions and Mechanisms. Chem. Rev. 106, 302–323. https://doi.org/10.1021/cr0404794

Jasin, M., Rothstein, R., 2013. Repair of Strand Breaks by Homologous Recombination. Cold Spring Harb Perspect Biol 5. https://doi.org/10.1101/cshperspect.a012740

Jeck, W.R., Sharpless, N.E., 2014. Detecting and characterizing circular RNAs. Nat. Biotechnol. 32, 453–461. https://doi.org/10.1038/nbt.2890

Jeck, W.R., Sorrentino, J.A., Wang, K., Slevin, M.K., Burd, C.E., Liu, J., Marzluff, W.F., Sharpless, N.E., 2013. Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA 19, 141–157. https://doi.org/10.1261/rna.035667.112

Ji, P., Diederichs, S., Wang, W., Böing, S., Metzger, R., Schneider, P.M., Tidow, N., Brandt, B., Buerger, H., Bulk, E., Thomas, M., Berdel, W.E., Serve, H., Müller-Tidow, C., 2003. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene 22, 8031–8041. https://doi.org/10.1038/sj.onc.1206928

Ji, Z., Song, R., Regev, A., Struhl, K., 2015. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. eLife 4. https://doi.org/10.7554/eLife.08890

Jones, D.T.W., Jäger, N., Kool, M., Zichner, T., Hutter, B., Sultan, M., Cho, Y.-J., Pugh, T.J., Hovestadt, V., Stütz, A.M., Rausch, T., Warnatz, H.-J., Ryzhova, M., Bender, S., Sturm, D., Pleier, S., Cin, H., Pfaff, E., Sieber, L., Wittmann, A., Remke, M., Witt, H., Hutter, S., Tzaridis, T., Weischenfeldt, J., Raeder, B., Avci, M., Amstislavskiy, V., Zapatka, M., Weber, U.D., Wang, Q., Lasitschka, B., Bartholomae, C.C., Schmidt, M., von Kalle, C., Ast, V., Lawerenz, C., Eils, J., Kabbe, R., Benes, V., van Sluis, P., Koster, J., Volckmann, R., Shih, D., Betts, M.J., Russell, R.B., Coco, S., Tonini, G.P., Schüller, U., Hans, V., Graf, N., Kim, Y.-J., Monoranu, C., Roggendorf, W., Unterberg, A., Herold-Mende, C., Milde, T., Kulozik, A.E., von Deimling, A., Witt, O., Maass, E., Rössler, J., Ebinger, M., Schuhmann, M.U., Frühwald, M.C., Hasselblatt, M., Jabado, N., Rutkowski, S., von Bueren, A.O., Williamson, D., Clifford, S.C., McCabe, M.G., Collins, V.P., Wolf, S.,

Wiemann, S., Lehrach, H., Brors, B., Scheurlen, W., Felsberg, J., Reifenberger, G., Northcott, P.A., Taylor, M.D., Meyerson, M., Pomeroy, S.L., Yaspo, M.-L., Korbel, J.O., Korshunov, A., Eils, R., Pfister, S.M., Lichter, P., 2012. Dissecting the genomic complexity underlying medulloblastoma. Nature 488, 100–105. https://doi.org/10.1038/nature11284

Josson, S., Sung, S.-Y., Lao, K., Chung, L.W.K., Johnstone, P.A.S., 2008. Radiation modulation of microRNA in prostate cancer cell lines. Prostate 68, 1599–1606. https://doi.org/10.1002/pros.20827

Kakarougkas, A., Ismail, A., Klement, K., Goodarzi, A.A., Conrad, S., Freire, R., Shibata, A., Lobrich, M., Jeggo, P.A., 2013. Opposing roles for 53BP1 during homologous recombination. Nucleic Acids Res. 41, 9719–9731. https://doi.org/10.1093/nar/gkt729

Kakarougkas, A., Jeggo, P.A., 2014. DNA DSB repair pathway choice: an orchestrated handover mechanism. Br J Radiol 87. https://doi.org/10.1259/bjr.20130685

Kang, Y.-J., Yang, D.-C., Kong, L., Hou, M., Meng, Y.-Q., Wei, L., Gao, G., 2017. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. Nucleic Acids Res. 45, W12–W16. https://doi.org/10.1093/nar/gkx428

Kasim, V., Wu, S., Taira, K., Miyagishi, M., 2013. Determination of the role of DDX3 a factor involved in mammalian RNAi pathway using an shRNA-expression library. PLoS ONE 8, e59445. https://doi.org/10.1371/journal.pone.0059445

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engström, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B., Wahlestedt, C., RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Project Core Group), FANTOM Consortium, 2005. Antisense transcription in the mammalian transcriptome. Science 309, 1564–1566. https://doi.org/10.1126/science.1112009

Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., Regev, A., Lander, E.S., Rinn, J.L., 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc. Natl. Acad. Sci. U.S.A. 106, 11667–11672. https://doi.org/10.1073/pnas.0904715106

Khan, M.A.F., Reckman, Y.J., Aufiero, S., van den Hoogenhof, M.M.G., van der Made, I., Beqqali, A., Koolbergen, D.R., Rasmussen, T.B., van der Velden, J., Creemers, E.E., Pinto, Y.M., 2016. RBM20 Regulates Circular RNA Production From the Titin Gene. Circ. Res. 119, 996–1003. https://doi.org/10.1161/CIRCRESAHA.116.309568

Kim, T.-K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., Markenscoff-Papadimitriou, E., Kuhl, D.,

Bito, H., Worley, P.F., Kreiman, G., Greenberg, M.E., 2010. Widespread transcription at neuronal activity-regulated enhancers. Nature 465, 182–187. https://doi.org/10.1038/nature09033

Kim, Y.-J., Wilson, D.M., 2012. Overview of Base Excision Repair Biochemistry. Curr Mol Pharmacol 5, 3–13.

Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T., Wutz, A., 2004. A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. PLoS Biol. 2, E171. https://doi.org/10.1371/journal.pbio.0020171

Kondo, T., Hashimoto, Y., Kato, K., Inagaki, S., Hayashi, S., Kageyama, Y., 2007. Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. Nat. Cell Biol. 9, 660–665. https://doi.org/10.1038/ncb1595

Kool, M., Jones, D.T.W., Jäger, N., Northcott, P.A., Pugh, T.J., Hovestadt, V., Piro, R.M., Esparza, L.A., Markant, S.L., Remke, M., Milde, T., Bourdeaut, F., Ryzhova, M., Sturm, D., Pfaff, E., Stark, S., Hutter, S., Seker-Cin, H., Johann, P., Bender, S., Schmidt, C., Rausch, T., Shih, D., Reimand, J., Sieber, L., Wittmann, A., Linke, L., Witt, H., Weber, U.D., Zapatka, M., König, R., Beroukhim, R., Bergthold, G., van Sluis, P., Volckmann, R., Koster, J., Versteeg, R., Schmidt, S., Wolf, S., Lawerenz, C., Bartholomae, C.C., von Kalle, C., Unterberg, A., Herold-Mende, C., Hofer, S., Kulozik, A.E., von Deimling, A., Scheurlen, W., Felsberg, J., Reifenberger, G., Hasselblatt, M., Crawford, J.R., Grant, G.A., Jabado, N., Perry, A., Cowdrey, C., Croul, S., Zadeh, G., Korbel, J.O., Doz, F., Delattre, O., Bader, G.D., McCabe, M.G., Collins, V.P., Kieran, M.W., Cho, Y.-J., Pomeroy, S.L., Witt, O., Brors, B., Taylor, M.D., Schüller, U., Korshunov, A., Eils, R., Wechsler-Reya, R.J., Lichter, P., Pfister, S.M., ICGC PedBrain Tumor Project, 2014. Genome sequencing of SHH medulloblastoma predicts genotype-related response to smoothened inhibition. Cancer Cell 25, 393–405. https://doi.org/10.1016/j.ccr.2014.02.004

Krichevsky, A.M., Sonntag, K.-C., Isacson, O., Kosik, K.S., 2006. Specific microRNAs modulate embryonic stem cell-derived neurogenesis. Stem Cells 24, 857–864. https://doi.org/10.1634/stemcells.2005-0441

Krol, J., Loedige, I., Filipowicz, W., 2010. The widespread regulation of microRNA biogenesis, function and decay. Nat. Rev. Genet. 11, 597–610. https://doi.org/10.1038/nrg2843

Kruhlak, M., Crouch, E.E., Orlov, M., Montaño, C., Gorski, S.A., Nussenzweig, A., Misteli, T., Phair, R.D., Casellas, R., 2007. The ATM repair pathway inhibits RNA polymerase I transcription in response to chromosome breaks. Nature 447, 730–734. https://doi.org/10.1038/nature05842

Ku, Y.-C., Lai, M.-H., Lo, C.-C., Cheng, Y.-C., Qiu, J.-T., Tarn, W.-Y., Lai, M.-C., 2019. DDX3 Participates in Translational Control of Inflammation Induced by Infections and Injuries. Mol. Cell. Biol. 39. https://doi.org/10.1128/MCB.00285-18

Lai, M.-C., Chang, W.-C., Shieh, S.-Y., Tarn, W.-Y., 2010. DDX3 Regulates Cell Growth through Translational Control of Cyclin E1. Molecular and Cellular Biology 30, 5444–5453. https://doi.org/10.1128/MCB.00560-10

Lai, M.-C., Lee, Y.-H.W., Tarn, W.-Y., 2008. The DEAD-box RNA helicase DDX3 associates with export messenger ribonucleoproteins as well as tip-associated protein and participates in translational control. Mol. Biol. Cell 19, 3847–3858. https://doi.org/10.1091/mbc.e07-12-1264

Lai Wing Sun, K., Correia, J.P., Kennedy, T.E., 2011. Netrins: versatile extracellular cues with diverse functions. Development 138, 2153–2169. https://doi.org/10.1242/dev.044529

Lakin, N.D., Jackson, S.P., 1999. Regulation of p53 in response to DNA damage. Oncogene 18, 7644–7655. https://doi.org/10.1038/sj.onc.1203015

Larrea, A.A., Lujan, S.A., Kunkel, T.A., 2010. SnapShot: DNA mismatch repair. Cell 141, 730.e1. https://doi.org/10.1016/j.cell.2010.05.002

Latham, J.P., Searle, P.F., Mautner, V., James, N.D., 2000. Prostate-specific antigen promoter/enhancer driven gene therapy for prostate cancer: construction and testing of a tissue-specific adenovirus vector. Cancer Res. 60, 334–341.

Lee, C.-S., Dias, A.P., Jedrychowski, M., Patel, A.H., Hsu, J.L., Reed, R., 2008. Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. Nucleic Acids Res 36, 4708–4718. https://doi.org/10.1093/nar/gkn454

Lee, H.Y., Haurwitz, R.E., Apffel, A., Zhou, K., Smart, B., Wenger, C.D., Laderman, S., Bruhn, L., Doudna, J.A., 2013. RNA-protein analysis using a conditional CRISPR nuclease. Proc. Natl. Acad. Sci. U.S.A. 110, 5416–5421. https://doi.org/10.1073/pnas.1302807110

Lee, J.-H., Paull, T.T., 2005. ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. Science 308, 551–554. https://doi.org/10.1126/science.1108297

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., Kim, V.N., 2003. The nuclear RNase III Drosha initiates microRNA processing. Nature 425, 415–419. https://doi.org/10.1038/nature01957

Legnini, I., Di Timoteo, G., Rossi, F., Morlando, M., Briganti, F., Sthandier, O., Fatica, A., Santini, T., Andronache, A., Wade, M., Laneve, P., Rajewsky, N., Bozzoni, I., 2017. Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. Mol. Cell 66, 22-37.e9. https://doi.org/10.1016/j.molcel.2017.02.017

Lennox, K.A., Behlke, M.A., 2016. Mini-review: Current strategies to knockdown long non-coding RNAs. Journal of Rare Diseases & Treatment.

Lewejohann, L., Skryabin, B.V., Sachser, N., Prehn, C., Heiduschka, P., Thanos, S., Jordan, U., Dell'Omo, G., Vyssotski, A.L., Pleskacheva, M.G., Lipp, H.-P., Tiedge, H., Brosius, J., Prior, H., 2004. Role of a neuronal small non-messenger RNA: behavioural alterations in BC1 RNA-deleted mice. Behav. Brain Res. 154, 273–289. https://doi.org/10.1016/j.bbr.2004.02.015

Li, M., Ding, W., Sun, T., Tariq, M.A., Xu, T., Li, P., Wang, J., 2018. Biogenesis of circular RNAs and their roles in cardiovascular development and pathology. The FEBS Journal 285, 220–232. https://doi.org/10.1111/febs.14191

Li, Q., Zhang, P., Zhang, C., Wang, Y., Wan, R., Yang, Y., Guo, X., Huo, R., Lin, M., Zhou, Z., Sha, J., 2014. DDX3X regulates cell survival and cell cycle during mouse early embryonic development. J Biomed Res 28, 282–291. https://doi.org/10.7555/JBR.27.20130047

Li, X., Wang, W., Chen, J., 2017. Recent progress in mass spectrometry proteomics for biomedical research. Sci China Life Sci 60, 1093–1113. https://doi.org/10.1007/s11427-017-9175-2

Li, Y., Altman, S., 2002. Partial reconstitution of human RNase P in HeLa cells between its RNA subunit with an affinity tag and the intact protein components. Nucleic Acids Res 30, 3706–3711.

Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., Zhong, G., Yu, B., Hu, W., Dai, L., Zhu, P., Chang, Z., Wu, Q., Zhao, Y., Jia, Y., Xu, P., Liu, H., Shan, G., 2015. Exon-intron circular RNAs regulate transcription in the nucleus. Nat. Struct. Mol. Biol. 22, 256–264. https://doi.org/10.1038/nsmb.2959

Lindahl, T., Barnes, D.E., 2000. Repair of endogenous DNA damage. Cold Spring Harb. Symp. Quant. Biol. 65, 127–133.

Liu, G., Mattick, J.S., Taft, R.J., 2013. A meta-analysis of the genomic and transcriptomic composition of complex life. Cell Cycle 12, 2061–2072. https://doi.org/10.4161/cc.25134

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.-J., Hammond, S.M., Joshua-Tor, L., Hannon, G.J., 2004. Argonaute2 is the catalytic engine of mammalian RNAi. Science 305, 1437–1441. https://doi.org/10.1126/science.1102513

Liu, P.Y., Erriquez, D., Marshall, G.M., Tee, A.E., Polly, P., Wong, M., Liu, B., Bell, J.L., Zhang, X.D., Milazzo, G., Cheung, B.B., Fox, A., Swarbrick, A., Hüttelmaier, S., Kavallaris, M., Perini, G., Mattick, J.S., Dinger, M.E., Liu, T., 2014. Effects of a novel long noncoding RNA, lncUSMycN, on N-Myc expression and neuroblastoma progression. J. Natl. Cancer Inst. 106. https://doi.org/10.1093/jnci/dju113

Liu, Y., Beyer, A., Aebersold, R., 2016. On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell 165, 535–550. https://doi.org/10.1016/j.cell.2016.03.014

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., Kutay, U., 2004. Nuclear export of microRNA precursors. Science 303, 95–98. https://doi.org/10.1126/science.1090599

Ma, E., Zhou, K., Kidwell, M.A., Doudna, J.A., 2012. Coordinated activities of human dicer domains in regulatory RNA processing. J. Mol. Biol. 422, 466–476. https://doi.org/10.1016/j.jmb.2012.06.009

Ma, J.-B., Ye, K., Patel, D.J., 2004. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature 429, 318–322. https://doi.org/10.1038/nature02519

Maes, O.C., An, J., Sarojini, H., Wu, H., Wang, E., 2008. Changes in MicroRNA expression patterns in human fibroblasts after low-LET radiation. J. Cell. Biochem. 105, 824–834. https://doi.org/10.1002/jcb.21878

Mahmoudi, S., Henriksson, S., Corcoran, M., Méndez-Vidal, C., Wiman, K.G., Farnebo, M., 2009. Wrap53, a natural p53 antisense transcript required for p53 induction upon DNA damage. Mol. Cell 33, 462–471. https://doi.org/10.1016/j.molcel.2009.01.028

Mao, Z., Bozzella, M., Seluanov, A., Gorbunova, V., 2008. Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair (Amst) 7, 1765–1771. https://doi.org/10.1016/j.dnarep.2008.06.018

Marzluff, W.F., Koreski, K.P., 2017. Birth and Death of Histone mRNAs. Trends in Genetics 33, 745–759. https://doi.org/10.1016/j.tig.2017.07.014

Matsumoto, A., Pasut, A., Matsumoto, M., Yamashita, R., Fung, J., Monteleone, E., Saghatelian, A., Nakayama, K.I., Clohessy, J.G., Pandolfi, P.P., 2017. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. Nature 541, 228–232. https://doi.org/10.1038/nature21034

Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S.P., Elledge, S.J., 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 316, 1160–1166. https://doi.org/10.1126/science.1140321

Mattick, J.S., 2018. The State of Long Non-Coding RNA Biology. Noncoding RNA 4. https://doi.org/10.3390/ncrna4030017

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., Tuschl, T., 2004. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol. Cell 15, 185–197. https://doi.org/10.1016/j.molcel.2004.07.007

Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., Loewer, A., Ziebold, U., Landthaler, M., Kocks, C., le Noble, F., Rajewsky, N., 2013. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature 495, 333–338. https://doi.org/10.1038/nature11928

Michael, M.Z., O' Connor, S.M., van Holst Pellekaan, N.G., Young, G.P., James, R.J., 2003. Reduced accumulation of specific microRNAs in colorectal neoplasia. Mol. Cancer Res. 1, 882–891.

Michalik, K.M., Böttcher, R., Förstemann, K., 2012. A small RNA response at DNA ends in Drosophila. Nucleic Acids Res. 40, 9596–9603. https://doi.org/10.1093/nar/gks711

Michelini, F., Pitchiaya, S., Vitelli, V., Sharma, S., Gioia, U., Pessina, F., Cabrini, M., Wang, Y., Capozzo, I., Iannelli, F., Matti, V., Francia, S., Shivashankar, G.V., Walter, N.G., d'Adda di Fagagna, F., 2017. Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. Nature Cell Biology 19, 1400–1411. https://doi.org/10.1038/ncb3643

Milek, M., Imami, K., Mukherjee, N., Bortoli, F.D., Zinnall, U., Hazapis, O., Trahan, C., Oeffinger, M., Heyd, F., Ohler, U., Selbach, M., Landthaler, M., 2017. DDX54 regulates transcriptome dynamics during DNA damage response. Genome Res. 27, 1344–1359. https://doi.org/10.1101/gr.218438.116

Mohammad, F., Mondal, T., Guseva, N., Pandey, G.K., Kanduri, C., 2010. Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. Development 137, 2493–2499. https://doi.org/10.1242/dev.048181

Moreno-Gonzalez, I., Soto, C., 2011. Misfolded Protein Aggregates: Mechanisms, Structures and Potential for Disease Transmission. Semin Cell Dev Biol 22, 482–487. https://doi.org/10.1016/j.semcdb.2011.04.002

Moumen, A., Magill, C., Dry, K., Jackson, S.P., 2013. ATM-dependent phosphorylation of heterogeneous nuclear ribonucleoprotein K promotes p53 transcriptional activation in response to DNA damage. Cell Cycle 12, 698–704. https://doi.org/10.4161/cc.23592

Mourtada-Maarabouni, M., Pickard, M.R., Hedge, V.L., Farzaneh, F., Williams, G.T., 2009. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene 28, 195–208. https://doi.org/10.1038/onc.2008.373

Mouse Genome Sequencing Consortium, Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S.E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M.R., Brown, D.G., Brown, S.D., Bult, C., Burton, J., Butler, J., Campbell, R.D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A.T., Church, D.M., Clamp, M., Clee, C., Collins, F.S., Cook, L.L., Copley, R.R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K.D., Deri, J., Dermitzakis, E.T., Dewey, C., Dickens,

N.J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D.M., Eddy, S.R., Elnitski, L., Emes, R.D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G.A., Flicek, P., Foley, K., Frankel, W.N., Fulton, L.A., Fulton, R.S., Furey, T.S., Gage, D., Gibbs, R.A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T.A., Green, E.D., Gregory, S., Guigó, R., Guyer, M., Hardison, R.C., Haussler, D., Hayashizaki, Y., Hillier, L.W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D.B., Johnson, L.S., Jones, M., Jones, T.A., Joy, A., Kamal, M., Karlsson, E.K., Karolchik, D., Kasprzyk, A., Kawai, J., Keibler, E., Kells, C., Kent, W.J., Kirby, A., Kolbe, D.L., Korf, I., Kucherlapati, R.S., Kulbokas, E.J., Kulp, D., Landers, T., Leger, J.P., Leonard, S., Letunic, I., Levine, R., Li, J., Li, M., Lloyd, C., Lucas, S., Ma, B., Maglott, D.R., Mardis, E.R., Matthews, L., Mauceli, E., Mayer, J.H., McCarthy, M., McCombie, W.R., McLaren, S., McLay, K., McPherson, J.D., Meldrim, J., Meredith, B., Mesirov, J.P., Miller, W., Miner, T.L., Mongin, E., Montgomery, K.T., Morgan, M., Mott, R., Mullikin, J.C., Muzny, D.M., Nash, W.E., Nelson, J.O., Nhan, M.N., Nicol, R., Ning, Z., Nusbaum, C., O'Connor, M.J., Okazaki, Y., Oliver, K., Overton-Larty, E., Pachter, L., Parra, G., Pepin, K.H., Peterson, J., Pevzner, P., Plumb, R., Pohl, C.S., Poliakov, A., Ponce, T.C., Ponting, C.P., Potter, S., Quail, M., Reymond, A., Roe, B.A., Roskin, K.M., Rubin, E.M., Rust, A.G., Santos, R., Sapojnikov, V., Schultz, B., Schultz, J., Schwartz, M.S., Schwartz, S., Scott, C., Seaman, S., Searle, S., Sharpe, T., Sheridan, A., Shownkeen, R., Sims, S., Singer, J.B., Slater, G., Smit, A., Smith, D.R., Spencer, B., Stabenau, A., Stange-Thomann, N., Sugnet, C., Suyama, M., Tesler, G., Thompson, J., Torrents, D., Trevaskis, E., Tromp, J., Ucla, C., Ureta-Vidal, A., Vinson, J.P., Von Niederhausern, A.C., Wade, C.M., Wall, M., Weber, R.J., Weiss, R.B., Wendl, M.C., West, A.P., Wetterstrand, K., Wheeler, R., Whelan, S., Wierzbowski, J., Willey, D., Williams, S., Wilson, R.K., Winter, E., Worley, K.C., Wyman, D., Yang, S., Yang, S.-P., Zdobnov, E.M., Zody, M.C., Lander, E.S., 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562. https://doi.org/10.1038/nature01262

Moynahan, M.E., Chiu, J.W., Koller, B.H., Jasin, M., 1999. Brca1 controls homology-directed DNA repair. Mol. Cell 4, 511–518.

Moynahan, M.E., Jasin, M., 2010. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. Nat. Rev. Mol. Cell Biol. 11, 196–207. https://doi.org/10.1038/nrm2851

Murat, P., Marsico, G., Herdy, B., Ghanbarian, A., Portella, G., Balasubramanian, S., 2018. RNA G-quadruplexes at upstream open reading frames cause DHX36- and DHX9-dependent translation of human mRNAs. Genome Biol 19. https://doi.org/10.1186/s13059-018-1602-2

Natsume, T., Kiyomitsu, T., Saga, Y., Kanemaki, M.T., 2016. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. Cell Reports 15, 210–218. https://doi.org/10.1016/j.celrep.2016.03.001

Negishi, M., Wongpalee, S.P., Sarkar, S., Park, J., Lee, K.Y., Shibata, Y., Reon, B.J., Abounader, R., Suzuki, Y., Sugano, S., Dutta, A., 2014. A New lncRNA, APTR,

Associates with and Represses the CDKN1A/p21 Promoter by Recruiting Polycomb Proteins. PLoS One 9. https://doi.org/10.1371/journal.pone.0095216

Nelson, B.R., Makarewich, C.A., Anderson, D.M., Winders, B.R., Troupes, C.D., Wu, F., Reese, A.L., McAnally, J.R., Chen, X., Kavalali, E.T., Cannon, S.C., Houser, S.R., Bassel-Duby, R., Olson, E.N., 2016. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science 351, 271–275. https://doi.org/10.1126/science.aad4076

Ng, W.L., Yan, D., Zhang, X., Mo, Y.-Y., Wang, Y., 2010. Over-expression of miR-100 is responsible for the low-expression of ATM in the human glioma cell line: M059J. DNA Repair 9, 1170–1175. https://doi.org/10.1016/j.dnarep.2010.08.007

Nguyen, T.A., Jo, M.H., Choi, Y.-G., Park, J., Kwon, S.C., Hohng, S., Kim, V.N., Woo, J.-S., 2015. Functional Anatomy of the Human Microprocessor. Cell 161, 1374–1387. https://doi.org/10.1016/j.cell.2015.05.010

Nissim, L., Perli, S.D., Fridkin, A., Perez-Pinera, P., Lu, T.K., 2014. Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. Mol Cell 54, 698–710. https://doi.org/10.1016/j.molcel.2014.04.022

Northcott, P.A., Jones, D.T.W., Kool, M., Robinson, G.W., Gilbertson, R.J., Cho, Y.-J., Pomeroy, S.L., Korshunov, A., Lichter, P., Taylor, M.D., Pfister, S.M., 2012. Medulloblastomics: the end of the beginning. Nat. Rev. Cancer 12, 818–834. https://doi.org/10.1038/nrc3410

Oh, S., Flynn, R.A., Floor, S.N., Purzner, J., Martin, L., Do, B.T., Schubert, S., Vaka, D., Morrissy, S., Li, Y., Kool, M., Hovestadt, V., Jones, D.T.W., Northcott, P.A., Risch, T., Warnatz, H.-J., Yaspo, M.-L., Adams, C.M., Leib, R.D., Breese, M., Marra, M.A., Malkin, D., Lichter, P., Doudna, J.A., Pfister, S.M., Taylor, M.D., Chang, H.Y., Cho, Y.-J., 2016. Medulloblastoma-associated DDX3 variant selectively alters the translational response to stress. Oncotarget 7, 28169–28182. https://doi.org/10.18632/oncotarget.8612

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schönbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. a. M., Setou, M., Shimada, K., Sultana, R., Takenaka, Y.,

Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E., Hayashizaki, Y., FANTOM Consortium, RIKEN Genome Exploration Research Group Phase I & II Team, 2002. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 420, 563–573. https://doi.org/10.1038/nature01266

Ørom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., Guigo, R., Shiekhattar, R., 2010. Long noncoding RNAs with enhancer-like function in human cells. Cell 143, 46–58. https://doi.org/10.1016/j.cell.2010.09.001

Pamudurti, N.R., Bartok, O., Jens, M., Ashwal-Fluss, R., Stottmeister, C., Ruhe, L., Hanan, M., Wyler, E., Perez-Hernandez, D., Ramberger, E., Shenzis, S., Samson, M., Dittmar, G., Landthaler, M., Chekulaeva, M., Rajewsky, N., Kadener, S., 2017. Translation of CircRNAs. Mol Cell 66, 9-21.e7. https://doi.org/10.1016/j.molcel.2017.02.021

Pandey, R.R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-Dinardo, D., Kanduri, C., 2008. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol. Cell 32, 232–246. https://doi.org/10.1016/j.molcel.2008.08.022

Pankotai, T., Bonhomme, C., Chen, D., Soutoglou, E., 2012. DNAPKcs-dependent arrest of RNA polymerase II transcription in the presence of DNA breaks. Nat. Struct. Mol. Biol. 19, 276–282. https://doi.org/10.1038/nsmb.2224

Parikh, A.R., He, Y., Hong, T.S., Corcoran, R.B., Clark, J.W., Ryan, D.P., Catenacci, D.V.T., Chao, J., Fakih, M., Ross, J.S., Stephens, P.J., Miller, V.A., Ali, S.M., Schrock, A.B., 2018. Analysis of DNA damage response (DDR) genes and tumor mutational burden (TMB) across 17,486 carcinomas of the tubular GI tract: Implications for therapy. JCO 36, 43–43. https://doi.org/10.1200/JCO.2018.36.4_suppl.43

Pasman, Z., Been, M.D., Garcia-Blanco, M.A., 1996. Exon circularization in mammalian nuclear extracts. RNA 2, 603–610.

Pasmant, E., Sabbagh, A., Vidaud, M., Bièche, I., 2011. ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. FASEB J. 25, 444–448. https://doi.org/10.1096/fj.10-172452

Pasquinelli, A.E., 2012. MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. Nat. Rev. Genet. 13, 271–282. https://doi.org/10.1038/nrg3162

Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.-C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E., Meyer, T., Cimprich, K.A., 2009. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. Mol. Cell 35, 228–239. https://doi.org/10.1016/j.molcel.2009.06.021

Pichiorri, F., Suh, S.-S., Rocci, A., De Luca, L., Taccioli, C., Santhanam, R., Wenchao, Z., Benson, D.M., Hofmainster, C., Alder, H., Garofalo, M., Di Leva, G., Volinia, S., Lin, H.-J., Perrotti, D., Kuehl, M., Aqeilan, R.I., Palumbo, A., Croce, C.M., 2010. Down-regulation of p53-inducible microRNAs 192, 194 and 215 impairs the p53/MDM2 auto-regulatory loop in multiple myeloma development. Cancer Cell 18, 367–381. https://doi.org/10.1016/j.ccr.2010.09.005

Piwecka, M., Glažar, P., Hernandez-Miranda, L.R., Memczak, S., Wolf, S.A., Rybak-Wolf, A., Filipchyk, A., Klironomos, F., Cerda Jara, C.A., Fenske, P., Trimbuch, T., Zywitza, V., Plass, M., Schreyer, L., Ayoub, S., Kocks, C., Kühn, R., Rosenmund, C., Birchmeier, C., Rajewsky, N., 2017. Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. Science 357. https://doi.org/10.1126/science.aam8526

Pothof, J., Verkaik, N.S., van IJcken, W., Wiemer, E.A.C., Ta, V.T.B., van der Horst, G.T.J., Jaspers, N.G.J., van Gent, D.C., Hoeijmakers, J.H.J., Persengiev, S.P., 2009. MicroRNA-mediated gene silencing modulates the UV-induced DNA-damage response. EMBO J. 28, 2090–2099. https://doi.org/10.1038/emboj.2009.156

Pugh, T.J., Weeraratne, S.D., Archer, T.C., Pomeranz Krummel, D.A., Auclair, D., Bochicchio, J., Carneiro, M.O., Carter, S.L., Cibulskis, K., Erlich, R.L., Greulich, H., Lawrence, M.S., Lennon, N.J., McKenna, A., Meldrim, J., Ramos, A.H., Ross, M.G., Russ, C., Shefler, E., Sivachenko, A., Sogoloff, B., Stojanov, P., Tamayo, P., Mesirov, J.P., Amani, V., Teider, N., Sengupta, S., Francois, J.P., Northcott, P.A., Taylor, M.D., Yu, F., Crabtree, G.R., Kautzman, A.G., Gabriel, S.B., Getz, G., Jäger, N., Jones, D.T.W., Lichter, P., Pfister, S.M., Roberts, T.M., Meyerson, M., Pomeroy, S.L., Cho, Y.-J., 2012. Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature 488, 106–110. https://doi.org/10.1038/nature11329

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281–2308. https://doi.org/10.1038/nprot.2013.143

Rand, T.A., Petersen, S., Du, F., Wang, X., 2005. Argonaute2 Cleaves the Anti-Guide Strand of siRNA during RISC Activation. Cell 123, 621–629. https://doi.org/10.1016/j.cell.2005.10.020

Rasool, M., Malik, A., Zahid, S., Basit Ashraf, M.A., Qazi, M.H., Asif, M., Zaheer, A., Arshad, M., Raza, A., Jamal, M.S., 2016. Non-coding RNAs in cancer diagnosis and therapy. Noncoding RNA Res 1, 69–76. https://doi.org/10.1016/j.ncrna.2016.11.001

Renganathan, A., Felley-Bosco, E., 2017. Long Noncoding RNAs in Cancer and Therapeutic Potential. Adv. Exp. Med. Biol. 1008, 199–222. https://doi.org/10.1007/978-981-10-5203-3_7

Ribeiro, D.M., Zanzoni, A., Cipriano, A., Delli Ponti, R., Spinelli, L., Ballarino, M., Bozzoni, I., Tartaglia, G.G., Brun, C., 2018. Protein complex scaffolding predicted as a prevalent function of long non-coding RNAs. Nucleic Acids Res. 46, 917–928. https://doi.org/10.1093/nar/gkx1169

Rich, T., Allen, R.L., Wyllie, A.H., 2000. Defying death after DNA damage. Nature 407, 777–783. https://doi.org/10.1038/35037717

Rivas, F.V., Tolia, N.H., Song, J.-J., Aragon, J.P., Liu, J., Hannon, G.J., Joshua-Tor, L., 2005. Purified Argonaute2 and an siRNA form recombinant human RISC. Nat. Struct. Mol. Biol. 12, 340–349. https://doi.org/10.1038/nsmb918

Robinson, G., Parker, M., Kranenburg, T.A., Lu, C., Chen, X., Ding, L., Phoenix, T.N., Hedlund, E., Wei, L., Zhu, X., Chalhoub, N., Baker, S.J., Huether, R., Kriwacki, R., Curley, N., Thiruvenkatam, R., Wang, J., Wu, G., Rusch, M., Hong, X., Becksfort, J., Gupta, P., Ma, J., Easton, J., Vadodaria, B., Onar-Thomas, A., Lin, T., Li, S., Pounds, S., Paugh, S., Zhao, D., Kawauchi, D., Roussel, M.F., Finkelstein, D., Ellison, D.W., Lau, C.C., Bouffet, E., Hassall, T., Gururangan, S., Cohn, R., Fulton, R.S., Fulton, L.L., Dooling, D.J., Ochoa, K., Gajjar, A., Mardis, E.R., Wilson, R.K., Downing, J.R., Zhang, J., Gilbertson, R.J., 2012. Novel mutations target distinct subgroups of medulloblastoma. Nature 488, 43–48. https://doi.org/10.1038/nature11213

Romero-Barrios, N., Legascue, M.F., Benhamed, M., Ariel, F., Crespi, M., 2018. Splicing regulation by long noncoding RNAs. Nucleic Acids Res. 46, 2169–2184. https://doi.org/10.1093/nar/gky095

Roos, W.P., Kaina, B., 2006. DNA damage-induced cell death by apoptosis. Trends Mol Med 12, 440–450. https://doi.org/10.1016/j.molmed.2006.07.007

Salzman, J., Gawad, C., Wang, P.L., Lacayo, N., Brown, P.O., 2012. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. PLoS ONE 7, e30733. https://doi.org/10.1371/journal.pone.0030733

Schärer, O.D., 2013. Nucleotide Excision Repair in Eukaryotes. Cold Spring Harb Perspect Biol 5. https://doi.org/10.1101/cshperspect.a012609

Schlackow, M., Nojima, T., Gomes, T., Dhir, A., Carmo-Fonseca, M., Proudfoot, N.J., 2017. Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. Mol Cell 65, 25–38. https://doi.org/10.1016/j.molcel.2016.11.029

Schoeftner, S., Sengupta, A.K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., Wutz, A., 2006. Recruitment of PRC1 function at the initiation of X inactivation

independent of PRC2 and silencing. EMBO J. 25, 3110–3122. https://doi.org/10.1038/sj.emboj.7601187

Sen, N.D., Zhou, F., Harris, M.S., Ingolia, N.T., Hinnebusch, A.G., 2016. eIF4B stimulates translation of long mRNAs with structured 5' UTRs and low closed-loop potential but weak dependence on eIF4G. Proc. Natl. Acad. Sci. U.S.A. 113, 10464–10472. https://doi.org/10.1073/pnas.1612398113

Sen, N.D., Zhou, F., Ingolia, N.T., Hinnebusch, A.G., 2015. Genome-wide analysis of translational efficiency reveals distinct but overlapping functions of yeast DEAD-box RNA helicases Ded1 and eIF4A. Genome Res. 25, 1196–1205. https://doi.org/10.1101/gr.191601.115

Shan, Y., Ma, J., Pan, Y., Hu, J., Liu, B., Jia, L., 2018. LncRNA SNHG7 sponges miR-216b to promote proliferation and liver metastasis of colorectal cancer through upregulating GALNT1. Cell Death & Disease 9, 722. https://doi.org/10.1038/s41419-018-0759-7

Shanbhag, N.M., Rafalska-Metcalf, I.U., Balane-Bolivar, C., Janicki, S.M., Greenberg, R.A., 2010. ATM-dependent chromatin changes silence transcription in cis to DNA double-strand breaks. Cell 141, 970–981. https://doi.org/10.1016/j.cell.2010.04.038

Shappell, S.B., 2008. Clinical utility of prostate carcinoma molecular diagnostic tests. Rev Urol 10, 44–69.

Sharma, D., Jankowsky, E., 2014. The Ded1/DDX3 subfamily of DEAD-box RNA helicases. Crit. Rev. Biochem. Mol. Biol. 49, 343–360. https://doi.org/10.3109/10409238.2014.931339

Sharma, V., Khurana, S., Kubben, N., Abdelmohsen, K., Oberdoerffer, P., Gorospe, M., Misteli, T., 2015. A BRCA1-interacting lncRNA regulates homologous recombination. EMBO Rep 16, 1520–1534. https://doi.org/10.15252/embr.201540437

Shibata, A., Conrad, S., Birraux, J., Geuting, V., Barton, O., Ismail, A., Kakarougkas, A., Meek, K., Taucher-Scholz, G., Löbrich, M., Jeggo, P.A., 2011. Factors determining DNA double-strand break repair pathway choice in G2 phase. EMBO J. 30, 1079–1092. https://doi.org/10.1038/emboj.2011.27

Shieh, S.-Y., Ikeda, M., Taya, Y., Prives, C., 1997. DNA Damage-Induced Phosphorylation of p53 Alleviates Inhibition by MDM2. Cell 91, 325–334. https://doi.org/10.1016/S0092-8674(00)80416-X

Shih, J.-W., Tsai, T.-Y., Chao, C.-H., Wu Lee, Y.-H., 2008. Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. Oncogene 27, 700–714. https://doi.org/10.1038/sj.onc.1210687

Shih, J.-W., Wang, W.-T., Tsai, T.-Y., Kuo, C.-Y., Li, H.-K., Wu Lee, Y.-H., 2012. Critical roles of RNA helicase DDX3 and its interactions with eIF4E/PABP1 in stress granule

assembly and stress response. Biochem. J. 441, 119–129. https://doi.org/10.1042/BJ20110739

Sievers, C., Schlumpf, T., Sawarkar, R., Comoglio, F., Paro, R., 2012. Mixture models and wavelet transforms reveal high confidence RNA-protein interaction sites in MOV10 PAR-CLIP data. Nucleic Acids Res. 40, e160. https://doi.org/10.1093/nar/gks697

Silva, G.M., Vogel, C., 2016. Quantifying gene expression: the importance of being subtle. Mol Syst Biol 12. https://doi.org/10.15252/msb.20167325

Simon, M.D., Wang, C.I., Kharchenko, P.V., West, J.A., Chapman, B.A., Alekseyenko, A.A., Borowsky, M.L., Kuroda, M.I., Kingston, R.E., 2011. The genomic binding sites of a noncoding RNA. Proc Natl Acad Sci U S A 108, 20497–20502. https://doi.org/10.1073/pnas.1113536108

Simone, N.L., Soule, B.P., Ly, D., Saleh, A.D., Savage, J.E., Degraff, W., Cook, J., Harris, C.C., Gius, D., Mitchell, J.B., 2009. Ionizing radiation-induced oxidative stress alters miRNA expression. PLoS ONE 4, e6377. https://doi.org/10.1371/journal.pone.0006377

Skryabin, B.V., Sukonina, V., Jordan, U., Lewejohann, L., Sachser, N., Muslimov, I., Tiedge, H., Brosius, J., 2003. Neuronal untranslated BC1 RNA: targeted gene elimination in mice. Mol. Cell. Biol. 23, 6435–6441.

Song, L., Lin, C., Wu, Z., Gong, H., Zeng, Y., Wu, J., Li, M., Li, J., 2011. miR-18a impairs DNA damage response through downregulation of ataxia telangiectasia mutated (ATM) kinase. PLoS ONE 6, e25454. https://doi.org/10.1371/journal.pone.0025454

Soto-Rifo, R., Rubilar, P.S., Limousin, T., de Breyne, S., Décimo, D., Ohlmann, T., 2012. DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. EMBO J. 31, 3745–3756. https://doi.org/10.1038/emboj.2012.220

Spitzer, J., Landthaler, M., Tuschl, T., 2013. Rapid creation of stable mammalian cell lines for regulated expression of proteins using the Gateway® recombination cloning technology and Flp-In T-REx® lines. Meth. Enzymol. 529, 99–124. https://doi.org/10.1016/B978-0-12-418687-3.00008-2

Srisawat, C., Engelke, D.R., 2001. Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. RNA 7, 632–641.

St Laurent, G., Shtokalo, D., Tackett, M.R., Yang, Z., Eremina, T., Wahlestedt, C., Urcuqui-Inchima, S., Seilheimer, B., McCaffrey, T.A., Kapranov, P., 2012. Intronic RNAs constitute the major fraction of the non-coding RNA in mammalian cells. BMC Genomics 13, 504. https://doi.org/10.1186/1471-2164-13-504

St Laurent, G., Vyatkin, Y., Antonets, D., Ri, M., Qi, Y., Saik, O., Shtokalo, D., de Hoon, M.J.L., Kawaji, H., Itoh, M., Lassmann, T., Arner, E., Forrest, A.R.R., FANTOM consortium, Nicolas, E., McCaffrey, T.A., Carninci, P., Hayashizaki, Y., Wahlestedt, C., Kapranov, P., 2016. Functional annotation of the vlinc class of non-coding RNAs using

systems biology approach. Nucleic Acids Res. 44, 3233–3252. https://doi.org/10.1093/nar/gkw162

Sternberg, S.H., Haurwitz, R.E., Doudna, J.A., 2012. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. RNA 18, 661–672. https://doi.org/10.1261/rna.030882.111

Stockley, P.G., Stonehouse, N.J., Murray, J.B., Goodman, S.T., Talbot, S.J., Adams, C.J., Liljas, L., Valegård, K., 1995. Probing sequence-specific RNA recognition by the bacteriophage MS2 coat protein. Nucleic Acids Res 23, 2512–2518.

Sun, H., Chen, C., Shi, M., Wang, D., Liu, M., Li, D., Yang, P., Li, Y., Xie, L., 2014. Integration of mass spectrometry and RNA-Seq data to confirm human ab initio predicted genes and lncRNAs. Proteomics 14, 2760–2768. https://doi.org/10.1002/pmic.201400174

Sun, M., Zhou, T., Jonasch, E., Jope, R.S., 2013. DDX3 regulates DNA damage-induced apoptosis and p53 stabilization. Biochim. Biophys. Acta 1833, 1489–1497. https://doi.org/10.1016/j.bbamcr.2013.02.026

Sunwoo, H., Dinger, M.E., Wilusz, J.E., Amaral, P.P., Mattick, J.S., Spector, D.L., 2009. MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome Res. 19, 347–359. https://doi.org/10.1101/gr.087775.108

Svejstrup, J.Q., 2010. The interface between transcription and mechanisms maintaining genome integrity. Trends Biochem. Sci. 35, 333–338. https://doi.org/10.1016/j.tibs.2010.02.001

Szabo, L., Salzman, J., 2016. Detecting circular RNAs: bioinformatic and experimental challenges. Nat Rev Genet 17, 679–692. https://doi.org/10.1038/nrg.2016.114

Taft, R.J., Mattick, J.S., 2003. Increasing biological complexity is positively correlated with the relative genome-wide expansion of non-protein-coding DNA sequences. Genome Biology 5, P1. https://doi.org/10.1186/gb-2003-5-1-p1

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., Mitsudomi, T., Takahashi, T., 2004. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res. 64, 3753–3756. https://doi.org/10.1158/0008-5472.CAN-04-0637

Tarn, W.-Y., Chang, T.-H., 2009. The current understanding of Ded1p/DDX3 homologs from yeast to human. RNA Biol 6, 17–20.

Taylor, D.W., Ma, E., Shigematsu, H., Cianfrocco, M.A., Noland, C.L., Nagayama, K., Nogales, E., Doudna, J.A., Wang, H.-W., 2013. Substrate-specific structural rearrangements of human Dicer. Nat. Struct. Mol. Biol. 20, 662–670. https://doi.org/10.1038/nsmb.2564

Team*, T.F.C. and the R.G.E.R.G.P.I.& I., 2002. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 420, 563–573. https://doi.org/10.1038/nature01266

Templin, T., Paul, S., Amundson, S.A., Young, E.F., Barker, C.A., Wolden, S.L., Smilenov, L.B., 2011. Radiation-Induced Micro-RNA Expression Changes in Peripheral Blood Cells of Radiotherapy Patients. International Journal of Radiation Oncology*Biology*Physics 80, 549–557. https://doi.org/10.1016/j.ijrobp.2010.12.061

Thorslund, T., Ripplinger, A., Hoffmann, S., Wild, T., Uckelmann, M., Villumsen, B., Narita, T., Sixma, T.K., Choudhary, C., Bekker-Jensen, S., Mailand, N., 2015. Histone H1 couples initiation and amplification of ubiquitin signalling after DNA damage. Nature 527, 389–393. https://doi.org/10.1038/nature15401

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols 7, 562–578. https://doi.org/10.1038/nprot.2012.016

Treiber, T., Treiber, N., Plessmann, U., Harlander, S., Daiß, J.-L., Eichner, N., Lehmann, G., Schall, K., Urlaub, H., Meister, G., 2017. A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. Mol. Cell 66, 270-284.e13. https://doi.org/10.1016/j.molcel.2017.03.014

Triboulet, R., Mari, B., Lin, Y.-L., Chable-Bessia, C., Bennasser, Y., Lebrigand, K., Cardinaud, B., Maurin, T., Barbry, P., Baillat, V., Reynes, J., Corbeau, P., Jeang, K.-T., Benkirane, M., 2007. Suppression of microRNA-silencing pathway by HIV-1 during virus replication. Science 315, 1579–1582. https://doi.org/10.1126/science.1136319

Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., Chang, H.Y., 2010. Long noncoding RNA as modular scaffold of histone modification complexes. Science 329, 689–693. https://doi.org/10.1126/science.1192002

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., Darnell, R.B., 2003. CLIP identifies Nova-regulated RNA networks in the brain. Science 302, 1212–1215. https://doi.org/10.1126/science.1090095

Wagner, S., Herrmannová, A., Šikrová, D., Valášek, L.S., 2016. Human eIF3b and eIF3a serve as the nucleation core for the assembly of eIF3 into two interconnected modules: the yeast-like core and the octamer. Nucleic Acids Res 44, 10772–10788. https://doi.org/10.1093/nar/gkw972

Wagner-Ecker, M., Schwager, C., Wirkner, U., Abdollahi, A., Huber, P.E., 2010. MicroRNA expression after ionizing radiation in human endothelial cells. Radiation Oncology 5, 25. https://doi.org/10.1186/1748-717X-5-25

Walker, S.E., Zhou, F., Mitchell, S.F., Larson, V.S., Valasek, L., Hinnebusch, A.G., Lorsch, J.R., 2013. Yeast eIF4B binds to the head of the 40S ribosomal subunit and promotes mRNA recruitment through its N-terminal and internal repeat domains. RNA 19, 191–207. https://doi.org/10.1261/rna.035881.112

Wan, G., Hu, X., Liu, Y., Han, C., Sood, A.K., Calin, G.A., Zhang, X., Lu, X., 2013. A novel non-coding RNA lncRNA-JADE connects DNA damage signalling to histone H4 acetylation. EMBO J 32, 2833–2847. https://doi.org/10.1038/emboj.2013.221

Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K., Kurokawa, R., 2008. Induced ncRNAs Allosterically Modify RNA Binding Proteins in cis to Inhibit Transcription. Nature 454, 126–130. https://doi.org/10.1038/nature06992

Wang, X., Guo, L., 2008. Effect of tetramethyl ammonium hydroxide on rheological properties of aqueous zirconia suspensions with polyacrylate. Colloids and Surfaces A: Physicochemical and Engineering Aspects 324, 126–130. https://doi.org/10.1016/j.colsurfa.2008.04.007

Wang, Y., Yu, Y., Tsuyada, A., Ren, X., Wu, X., Stubblefield, K., Rankin-Gee, E.K., Wang, S.E., 2011. Transforming growth factor-β regulates the sphere-initiating stem cell-like feature in breast cancer through miRNA-181 and ATM. Oncogene 30, 1470–1480. https://doi.org/10.1038/onc.2010.531

Warneford-Thomson, R., He, C., Sidoli, S., Garcia, B.A., Bonasio, R., 2017. Sample Preparation for Mass Spectrometry-based Identification of RNA-binding Regions. J Vis Exp. https://doi.org/10.3791/56004

Wei, W., Ba, Z., Gao, M., Wu, Y., Ma, Y., Amiard, S., White, C.I., Rendtlew Danielsen, J.M., Yang, Y.-G., Qi, Y., 2012. A Role for Small RNAs in DNA Double-Strand Break Repair. Cell 149, 101–112. https://doi.org/10.1016/j.cell.2012.03.002

Westholm, J.O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., Celniker, S.E., Graveley, B.R., Lai, E.C., 2014. Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. Cell Rep 9, 1966–1980. https://doi.org/10.1016/j.celrep.2014.10.062

Wilky, B.A., Kim, C., McCarty, G., Montgomery, E.A., Kammers, K., DeVine, L.R., Cole, R.N., Raman, V., Loeb, D.M., 2016. RNA helicase DDX3: a novel therapeutic target in Ewing sarcoma. Oncogene 35, 2574–2583. https://doi.org/10.1038/onc.2015.336

Winter, J., Jung, S., Keller, S., Gregory, R.I., Diederichs, S., 2009. Many roads to maturity: microRNA biogenesis pathways and their regulation. Nature Cell Biology 11, 228–234. https://doi.org/10.1038/ncb0309-228

Wu, C.-W., Dong, Y.-J., Liang, Q.-Y., He, X.-Q., Ng, S.S.M., Chan, F.K.L., Sung, J.J.Y., Yu, J., 2013. MicroRNA-18a attenuates DNA damage repair through suppressing the

expression of ataxia telangiectasia mutated in colorectal cancer. PLoS ONE 8, e57036. https://doi.org/10.1371/journal.pone.0057036

Wu, X.-S., Wang, F., Li, H.-F., Hu, Y.-P., Jiang, L., Zhang, F., Li, M.-L., Wang, X.-A., Jin, Y.-P., Zhang, Y.-J., Lu, W., Wu, W.-G., Shu, Y.-J., Weng, H., Cao, Y., Bao, R.-F., Liang, H.-B., Wang, Z., Zhang, Y.-C., Gong, W., Zheng, L., Sun, S.-H., Liu, Y.-B., 2017. LncRNA-PAGBC acts as a microRNA sponge and promotes gallbladder tumorigenesis. EMBO reports e201744147. https://doi.org/10.15252/embr.201744147

Wyler, E., Menegatti, J., Franke, V., Kocks, C., Boltengagen, A., Hennig, T., Theil, K., Rutkowski, A., Ferrai, C., Baer, L., Kermas, L., Friedel, C., Rajewsky, N., Akalin, A., Dölken, L., Grässer, F., Landthaler, M., 2017. Widespread activation of antisense transcription of the host genome during herpes simplex virus 1 infection. Genome Biol. 18, 209. https://doi.org/10.1186/s13059-017-1329-5

Yan, D., Ng, W.L., Zhang, X., Wang, P., Zhang, Z., Mo, Y.-Y., Mao, H., Hao, C., Olson, J.J., Curran, W.J., Wang, Y., 2010. Targeting DNA-PKcs and ATM with miR-101 sensitizes tumors to radiation. PLoS ONE 5, e11397. https://doi.org/10.1371/journal.pone.0011397

Yan, H., Tian, R., Zhang, M., Wu, J., Ding, M., He, J., 2016. High expression of long noncoding RNA HULC is a poor predictor of prognosis and regulates cell proliferation in glioma. Onco Targets Ther 10, 113–120. https://doi.org/10.2147/OTT.S124614

Yang, F., Babak, T., Shendure, J., Disteche, C.M., 2010. Global survey of escape from X inactivation by RNA-sequencing in mouse. Genome Res. 20, 614–622. https://doi.org/10.1101/gr.103200.109

Yang, W., Du, W.W., Li, X., Yee, A.J., Yang, B.B., 2016. Foxo3 activity promoted by non-coding effects of circular RNA and Foxo3 pseudogene in the inhibition of tumor growth and angiogenesis. Oncogene 35, 3919–3931. https://doi.org/10.1038/onc.2015.460

Yi, R., Qin, Y., Macara, I.G., Cullen, B.R., 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. 17, 3011–3016. https://doi.org/10.1101/gad.1158803

Yogosawa, S., Yoshida, K., 2018. Tumor suppressive role for kinases phosphorylating p53 in DNA damage-induced apoptosis. Cancer Sci. 109, 3376–3382. https://doi.org/10.1111/cas.13792

Yoon, J.-H., Gorospe, M., 2016. Identification of mRNA-interacting factors by MS2-TRAP (MS2-tagged RNA affinity purification). Methods Mol Biol 1421, 15–22. https://doi.org/10.1007/978-1-4939-3591-8_2

Yoon, J.-H., Srikantan, S., Gorospe, M., 2012. MS2-TRAP (MS2-tagged RNA affinity purification): tagging RNA to identify associated miRNAs. Methods 58, 81–87. https://doi.org/10.1016/j.ymeth.2012.07.004

Youngman, E.M., Green, R., 2005. Affinity purification of in vivo-assembled ribosomes for in vitro biochemical analysis. Methods 36, 305–312. https://doi.org/10.1016/j.ymeth.2005.04.007

Zhang, H., Kolb, F.A., Brondani, V., Billy, E., Filipowicz, W., 2002. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. EMBO J. 21, 5875–5885.

Zhang, Y., Zhang, X.-O., Chen, T., Xiang, J.-F., Yin, Q.-F., Xing, Y.-H., Zhu, S., Yang, L., Chen, L.-L., 2013. Circular intronic long noncoding RNAs. Mol. Cell 51, 792–806. https://doi.org/10.1016/j.molcel.2013.08.017

Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., Lee, J.T., 2010. Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol. Cell 40, 939–953. https://doi.org/10.1016/j.molcel.2010.12.011

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.-J., Lee, J.T., 2008. Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome. Science 322, 750–756. https://doi.org/10.1126/science.1163045

Zheng, Q., Bao, C., Guo, W., Li, S., Chen, J., Chen, B., Luo, Y., Lyu, D., Li, Y., Shi, G., Liang, L., Gu, J., He, X., Huang, S., 2016. Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. Nat Commun 7, 11215. https://doi.org/10.1038/ncomms11215