At the crossroads of RNA biology, genome integrity and cancer

Biswendu Biswas*, Rady Chaaban*, Shrena Chakraborty*, Alexandre Devaux*, Ana Luisa Dian*, Anna Minello*, Jenny Kaur Singh*, Stephan Vagner, Patricia Uguen, Sarah Lambert, Martin Dutertre, Aura Carreira

*equal contribution

CNRS UMR 3348 Genome integrity, RNA and Cancer, Institut Curie, University Paris-Saclay, 91401 Orsay

Short title: RNA, genome integrity and cancer

Summary: This article is the synthesis of the scientific presentations that took place during two international courses at Institute Curie, one on post-transcriptional gene regulation and the other on genome instability and human disease, that were joined together in their 2021 edition. This joined course brought together the knowledge on RNA metabolism and the maintenance of genome stability.

Résumé: Cet article est la synthèse des présentations faites lors de deux cours internationaux de l'Institut Curie, l'un portant sur la régulation post-transcriptionnelle des gènes et l'autre sur l'instabilité génomique et les maladies humaines, qui ont eu lieu de façon conjointe en 2021. Ce cours joint a connecté les connaissances sur le métabolisme des ARN et sur le maintien de la stabilité des génomes.

Keywords: RNA, post-transcriptional regulation, genome instability, DNA damage, DNA repair, R-loops

List of abbreviations

mRNAs	messenger RNAs
mNET-seq	mammalian Native Elongating Transcript sequencing
POINT-seq	Polymerase Intact Nascent Transcript sequencing
RNAPII	RNA Polymerase II
poly(A) tail	polyadenylate tail
CPA	Cleavage and Polyadenylation
CPF	Cleavage and Polyadenylation Factor
RBPs	RNA Binding Proteins
EJC	Exon Junction Complex
NMD	Nonsense-mediated decay
NIN	Ninein
ROS	Reactive oxygen species
OOPs	Orthogonal Organic Phase separation
SG	Stress Granules
RNP	ribonucleoprotein
tRNAs	transfer RNAs
rRNAs	ribosomal RNAs
DDR	DNA damage response
PARPi	poly-ADP-ribose polymerase 1 inhibitor
QIBC	quantitative imaging-based cytometry
MiDAS	Mitotic DNA synthesis
OK-seq	Okazaki fragments sequencing
DSB	DNA double-strand break
HR	Homologous recombination
NHEJ	Non-homologous end joining
ssDNA	Single-stranded DNA
TAD	Topologically associated domain
IPA	intronic polyadenylation
UV-C	ultraviolet C
dilncRNAs	Damage-induced long non-coding RNAs
DRIP	DNA-RNA immunoprecipitation

The mutual interactions between RNA metabolism and the DNA damage response have gained particular interest over the recent years and entail important clinical implications. This has led the organizers of two international courses at Institut Curie ("5th Course on Post-transcriptional gene regulation" and "3rd Course on Genome integrity and human diseases") to exceptionally join their efforts. On the one hand, the "post-transcriptional gene regulation" course focuses on the study of post-transcriptional regulations, from molecular mechanisms to genome-wide networks. On the other, the "genome integrity and human disease" course aims to introduce basic mechanisms contributing to the maintenance of genome stability from molecular mechanisms up to omics approaches. Here, we discuss some of the highlights of this 7-day course, sponsored by Institut Curie and Société Française du Cancer, that took place in real-time virtual format in April 12-20 2021.

Post-transcriptional gene regulation

Post-transcriptional regulation plays an important role in controlling gene expression by subjecting precursor of messenger RNAs (mRNAs) called pre-mRNAs to a host of maturation events (*i.e.* splicing and polyadenylation) before being exported to the cytoplasm. Pre-mRNA splicing mainly occurs co-transcriptionally. The lab of **Maria Carmo-Fonseca** (Instituto de Medicina Molecular João Lobo Antunes, PT) has contributed to determining splicing kinetics in Metazoans by using several technologies. Live-cell imaging of stable transgenes showed a transcription rate of about 4 kilobases per minute and a splicing reaction lasting only for a few seconds [1]. More recently, through genome-wide analyses of nascent transcripts in human cells with the mNET-seq (mammalian native elongating transcript sequencing) and POINT-seq (polymerase intact nascent transcript sequencing) techniques, her team showed that for most introns, splicing takes place co-transcriptionally, immediately after the 3' splice site is transcribed by RNA polymerase II (RNAPII) [2,3]. However, for some introns splicing is delayed. These data raise questions about the molecular mechanisms underlying such different splicing kinetics, and their consequences on gene expression, alternative splicing regulation and, potentially, DNA-RNA hybrid formation and genome stability (which are discussed below).

Besides splicing, another key step of pre-mRNA maturation is 3' end processing, which generally consists in RNA cleavage at the polyadenylation site and in the addition of a polyadenylate tail (poly(A) tail), that is not encoded in DNA. This cleavage and polyadenylation (CPA) process is coupled to transcription termination. In addition, the poly(A) tail promotes the nucleo-cytoplasmic export, translation and stability of the mRNA, and is the substrate of deadenylases that trigger mRNA degradation. The lab of **Lori Passmore** (MRC Laboratory of Molecular Biology, Cambridge, UK) determined the molecular structure of a key multiprotein component of the yeast CPA machinery, Cleavage and Polyadenylation Factor (CPF), by using *in vitro* reconstitution experiments, cryo-electron microscopy and X-ray crystallography [4,5]. They also determined the structural determinants of poly(A) tail recognition by two deadenylases. [6]. These studies provided insights into the molecular mechanisms of CPA and deadenylation processes in yeast, with implications for human cells since most protein components of these machineries are highly conserved in humans. Moreover, both polyadenylation (*e.g.*, FIP1L1, CPSF3/CPSF73, NUDT21/CFIm25) and deadenylation factors (*e.g.*, CCR4-NOT complex components) have been involved in cancer [7,8]. In this context, recent work from Lori Passmore and collaborators showed that CPSF3 is a druggable node in acute myeloid leukemia and Ewing sarcoma [9].

During co-transcriptional RNA processing, RNA-binding proteins (RBPs) are deposited on transcripts, and some of them follow mRNAs into the cytoplasm and impact their fate. This is the case of the exon junction complex (EJC), a protein complex that is deposited on transcripts during splicing and promotes their nucleo-cytoplasmic export and translation. The EJC is best known for its role in triggering nonsense-mediated decay (NMD) of mRNAs that have a

premature stop codon located upstream of the last exon junction, which is the case of many aberrant transcripts that are produced by mutated genes or upon splicing factor mutations in diseases like cancer. About twenty years ago, **Hervé Le Hir** (now at Institut de Biologie of ENS, Paris, FR) discovered the EJC and its role in various processes, including NMD, mRNA export, translation and specific localization within the cytoplasm [10]. More recently, his lab showed that the EJC is localized at basal bodies during ciliogenesis in mouse neural stem cells and mediates the localization of the NIN (ninein) mRNA to centrioles, that form basal bodies and where the NIN protein plays a key function. Furthermore, EJC down-regulation impairs ciliogenesis [11]. These findings may explain the involvement of the EJC in neural stem cell division and human neurodevelopmental disorders.

Once in the cytoplasm, mRNAs are subjected to translational regulation. Alterations in the mRNA translation machinery can impact diverse cellular aspects, such as cell proliferation and metabolism. **Davide Ruggero** (UCSF Helen Diller Family Comprehensive Cancer Center, US) presented the activity of a general translation initiation factor, eIF4E, that binds to the "cap" structure located at the 5' end of all mRNAs. His lab showed that eIF4E is essential for translating a subset of mRNAs implicated in the regulation of intracellular ROS levels, which is critical for tumour cell survival. Interestingly, recently they demonstrated that eIF4E heterozygous mice are resistant to diet-induced obesity, suggesting that diminished eIF4E levels can promote enhanced metabolic fitness [12].

Anne Willis (MRC Toxicology Unit, University of Cambridge, UK) provided evidence that mRNA translation of ribosomal proteins and nuclear-encoded mitochondrial factors is altered in malignant mesothelioma [13]. In an effort to identify RBPs involved in the cellular response to toxic injury and external stress, Anne Willis also presented the development of experimental approaches such as Orthogonal Organic Phase Separation (OOPs) that allows to discriminate RNA-bound proteins [14].

mRNA translation is also modified upon stress induction within stress-induced membrane-less organelles, so called stress granules (SG). **Jeffrey A Chao** (Friedrich Miescher Institute for Biomedical Research, Basel, CH) presented an elegant technique based on single-molecule imaging to analyse mRNA translation in living cells. In contrast to the current model that postulates mRNA translation is inhibited in SG, he showed that translation in SG is not a rare event, and that the SG environment does not directly inhibit translation [15].

The link between cytoplasmic membrane-less ribonucleoprotein (RNP) granules and mRNA fate was presented by **Dominique Weil** (Institute of Biology Paris-Seine, Paris, FR) who focuses on P-bodies. Her contribution in the development of a method to purify these granules enabled the characterization of RNA and protein content present in P-bodies using RNA-Seq and mass spectrometry, respectively. The comparison of the transcriptome of P-bodies before or after silencing mRNA decay or other repression factors revealed that GC content shapes mRNA storage and decay. Indeed, AU-rich and GC-rich mRNAs globally follow different decay pathways and the global GC content of mRNAs in P-bodies are closely linked to mRNA stability, translation rate, RBP and miRNA binding [16]. This study proposes an integrated view of the post-transcriptional control of mRNA translation and mRNA stability.

More than just an mRNA storage compartment, P-bodies have been discovered to be implicated in a variety of polarized and non-polarized cells to compartmentalize protein synthesis. **Florence Besse** (Institut de Biologie Valrose, FR) has evidence that the targeting of the transcripts to their destination is operated by RNP granules which contain RNA cargo and regulatory proteins [17]. Using high resolution live imaging techniques, she showed that the targeting of mRNAs in neuronal RNP granules is a dynamic and reversible mechanism [18]. She has also found that defects in this process are linked to neurodegenerative diseases.

More than 150 chemical modifications of RNAs have recently been identified [19], which includes N6-methyladenosine (m6A), the most widespread modification on mammalian mRNAs, pseudouridine (Ψ), ribose methylations (Nm), N1-methyladenosine (m1A), 5-

methylcytidine (m5C), N7-methylguanosine (m7G) and N4-acetylcytidine (ac4C) among many others. These modifications harbor the potential of regulating the properties of RNAs and have emerged as critical regulators of gene expression, highlighting the importance of understanding their nature and role in biology and disease [20]. Although N4-acetylcytidine (ac4C) is possibly one of the most highly conserved mechanisms of enzymatic modification of RNA, especially in tRNAs and rRNAs, the function of this cytidine acetylation, as well as its role in biology and disease, have yet to be elucidated [21]. **Shraga Schwartz** (Weizmann Institute of Science, Rehovot, IL) presented a novel chemical approach for quantitative mapping of ac4C at single-nucleotide resolution in order to study hyperthermophiles archaea. With this new insight, ac4C appeared as an essential modification for these microorganisms to resist extreme temperatures [21].

Recent advances in the DNA Damage Response

Many chemotherapeutic drugs used in combination with radiotherapy kill cancer cells by damaging DNA, and many of them, target DNA replication-based processes given the highly replicative nature of cancer cells. Moreover, the DNA damage response (DDR) factors and the pathways themselves are potential targets to improve anti-cancer therapies. Such strategy is beautifully illustrated by the clinical use of PARP inhibitors (PARPi) that target the DNA damage sensors poly-ADP-ribose polymerase 1 and 2 (PARP1/2). Dr **Matthias Altmeyer** (Department of Molecular Mechanisms of Disease, University of Zurich, CH) uses quantitative imaging-based cytometry (QIBC), a high-content microscopy approach to quantify the chromatin association of DDR factors and relevant parameters such as cytotoxicity according to cell cycle progression in single cells. Applying QIBC to investigate the cell response to PARPi, Dr Altmeyer revealed that these drugs have an impact outside S-phase-specific DNA damage response. By testing a panel of cell lines, it was possible to predict PARPi resistance or hypersensitivity and to delineate distinct cell responses to different PARPi in a quantitative manner [22]. One deleterious consequence of failures in S-phase progression is unfinished DNA replication resulting in under-replicated regions, also known as fragile sites, when cells enter mitosis. To overcome unfinished DNA replication, a process named MiDAS (Mitotic DNA synthesis) is active in mitosis to replicate under-replicated regions using a form of break-induced replication [23]. Briefly, under-replicated DNA is cleaved by various endonucleases, generating a break from which DNA synthesis is initiated in a conservative mode, as opposed to the canonical semi-conservative DNA synthesis. If MiDAS fails, 53BP1 nuclear bodies shield inherited genomic lesions from repair or degradation in G1. As a consequence of the inherited genomic lesions, the innate immune response can be activated through cGAS-STING pathway. Overall, the use of single cell experiments combined with technologies to map fragile sites and their behavior in a cell cycle specific manner provides a better description and understanding of the DDR.

Dr **Chunlong Chen** (Institut Curie, FR) has presented how genome-wide studies help to understand the dynamics of DNA replication in normal and challenged conditions. His lab focuses on understanding the spatio-temporal program of the human genome aiming to better understand how this program is deregulated in cancer cells or can be targeted to improve anti-cancer therapies. Deregulation of the DNA replication program threatens genome stability and is often observed in cancer cells. Chen's team has pioneered the development of deep sequencing of Okazaki fragments (OK-Seq), that mark the synthesis of the lagging strand, thus providing crucial information about replication fork directionality genome-wide [24]. His team has also developed the Repli-Seq approach that allows the timing and replication dynamics of any specific locus to be investigated [25]. Combining these approaches, Dr Chen has presented how gene transcription landscape impact on DNA replication dynamics with the emerging concept that transcription during S-phase is a source of replication stress leading to recurrent genome instability when transcription-replication conflicts are not dealt with properly [26]. For example, his results have confirmed the concept that large genes embedding long-transcription units

strongly delay replication completion resulting in fragile sites at which DNA synthesis is not completed before cells enter mitosis.

The most toxic DNA lesions induced by anticancer therapies are DNA double strand breaks (DSB) which are repaired either by the non-homologous end-joining (NHEJ) or homologous recombination (HR) pathway. This last pathway has been the focus of intense research since inactivation of HR (caused by BRCA1 or BRCA2 mutations for example) leads to predisposition to breast and ovarian cancers. Repair pathway choice between NHEJ and HR is under the control of several factors that prevent or favor the resection of a DSB in which single stranded DNA (ssDNA) is generated, a process essential for HR-dependent repair. The presentation of Dr **Dipanjan Chowdhury** (Division of Radiation and Genome Stability, Dana-Farber Cancer Institute, Boston, US) highlighted a novel factor involved in the resection of DSBs, identified in a loss-of-function CRISPR screen. This screen was focused on the identification of factors causing PARPi resistance or platinum-based therapy in BRCA mutated cell lines. This screen led to the discovery of DYNLL1 as a novel inhibitor of DSB end-resection [27]. Mechanistically, DYNLL1 interacts with the nuclease MRE11 (which mutations cause Ataxia Telangiectasia like disorder) to impair its activity. Moreover, decrease in DYNLL1 expression in carcinomas with low BRCA1 expression reduced genomic alterations. Together, this work highlights an important new factor in DSB-repair influencing responses to cancer therapies.

The substrate of the DDR is not the naked DNA but the chromatin, which plays a pivotal role in the signalling of DSB and their repair. The team of **Gaëlle Legube** (CBI, Toulouse, FR) has a particular interest in understanding how chromatin folding and modifications trigger the DDR. To do so, the team has developed unique tools to induce ~100 DSBs (DIvA) in cells and to analyse their signalling and repair according to the chromatin landscape [28]. The research of Legube's team previously reported that DSB within transcriptionally active regions are preferentially repaired by the HR pathway and that DSB mobility, clustering and nuclear positioning are key determinants of repair pathway choice [29–31]. More recently, her team has revealed how chromatin folding impacts early sensing of DSBs. One of the first events at DSB sites is the phosphorylation of the histone variant H2AX by the sensor kinase ATM (mutations of which cause Ataxia Telangiectasia disorder), a histone modification known as γH2AX. γH2AX is visible in the form of sub-nuclear foci since this modification can spread up to 50 kb around the site of the DSB. The mechanistic insight into the rapid spreading of γH2AX around DSB sites was missing. Gaelle Legube presented her most recent research explaining that topologically associated domains (TAD), that are self-interacting genomic regions, are pivotal to establish the early steps of the DDR. TAD boundaries assist in the establishment of γH2AX via one-sided cohesion-mediated loop extrusion on both sides of the DSB [32]. Taken together, she proposes that TADs are functional units of the DDR to establish γH2AX chromatin domains that promote DSB signalling and repair foci.

The links between DNA damage and RNA biology

It is now well established that DNA damage widely impacts gene expression at the level of transcription, but also at multiple post-transcriptional levels [33]. In particular, **Martin Dutertre** (Institut Curie, FR) showed that alternative splicing is widely regulated in cancer cell response and resistance to genotoxic anticancer agents, such as topoisomerase inhibitors [34–36]. In recent years, many genes have been found to contain alternative polyadenylation sites within annotated introns, and their use generates so-called intronic polyadenylation (IPA) isoforms with alternative last exons. Martin Dutertre showed that IPA isoforms are widely regulated by camptothecin and doxorubicin (topoisomerase I and II inhibitors, respectively) but with different genome-wide patterns: mainly, down-regulation events in the case of doxorubicin, and equal proportions of up- and down-regulation events in the case of camptothecin [35]. He also presented data showing the widespread regulation of IPA isoforms by cisplatin, another genotoxic anticancer agent. IPA isoform regulation is enriched in genes related to the DDR, cell

cycle and cell death, and he identified IPA isoforms that impact cell sensitivity to genotoxic agents. In addition, while the regulation of IPA isoforms has been mainly studied at the level of transcript synthesis, splicing and polyadenylation, he presented unpublished genome-wide analyses of their cytoplasmic regulation and of their translation status, by using 3'-seq (RNA-seq focused on the 3'-end of polyA+ RNA) on subcellular compartments and polysome fractions. These analyses reveal diverse fates and translational outcomes of IPA isoforms. Finally, he discussed the increasing evidence for reciprocal links between pre-mRNA 3' end processing (cleavage and polyadenylation) and genome stability [37].

The regulation of gene expression at multiple levels by DNA-damaging agents has been particularly characterized in the case of ultraviolet C (UV-C) irradiation. **Jesper Svejstrup** (University of Copenhagen, DK) showed, by using genome-wide GRO-seq analyses, that UV-C cell irradiation causes an inhibition of transcription elongation within 45 min. This is followed by an inhibition of transcription initiation within 2 to 4 h due to RNAPII degradation. Finally, transcription restarts between 12 and 24 h after irradiation thanks to RNAPII recovery, which is compromised in Cockayne syndrome B cells [38]. His lab also showed by RNA-seq analysis that elongation inhibition by UV-C irradiation is accompanied by an increase in the relative levels of IPA *versus* full-length mRNA isoforms in many long genes. In the *ASCC3* gene, IPA generates a transcript isoform with a non-coding function that antagonizes the function of the ASCC3 protein -encoded by the full-length mRNA- in transcription recovery following UV-C irradiation [39]. Jesper Svejstrup and his collaborators found that the ASCC3 protein is also involved in the management of translational stress due to ribosome collisions [40]. Recently, they discovered that collided ribosomes are coactivators of cGAS, a sensor of cytosolic DNA that activates interferon-stimulated genes and thereby the innate immune response [41]. This finding may be relevant to understand inflammation caused by viral infections.

The links between genome stability and RNA biology are not limited to post-transcriptional gene regulation by DNA damage. Indeed, in the last decade, non-coding RNAs have emerged as pivotal players in the maintenance of genome stability in response to DNA damage (see also next section). Dr **Fabrizio d'Adda di Fagagna** (IFOM, Milano, IT) has made major discoveries in this field by establishing that DSBs are actively transcribed by RNAPII although some details of this mechanism remain to be elucidated [42]. His team has proposed that RNAPII is recruited to DSB to generate ncRNAs named damage-induced long non-coding RNAs (dilncRNAs) that are important to foster DSB signaling and DDR foci by a liquid-liquid phase-separation process [43]. The inhibition of dilncRNAs using antisense oligonucleotides (called AOS) lead to site-specific inhibition of the DDR, affecting DSB repair by NHEJ and HR. These results bring about the concept that DSB-induced transcription is essential to fully activate the DDR. Moreover, the use of telomere specific AOS to inhibit the DDR induced by telomere shortening restored tissue homeostasis in animal models, providing a potential clinical option to treat diseases associated to accelerated aging [44,45]. Finally, Dr Fabrizio d'Adda di Fagagna presented his recent work on the mechanism of recruitment of RNAPII to DSBs: he showed that RNAPII recruitment requires the MRN complex which cleaves the dsDNA providing an entry point for RNAPII binding and transcription initiation [46]. **Monika Gullerova** (University of Oxford, UK) has also contributed to describe how ncRNAs are generated at DSBs. Her lab showed that in human and mouse cells, a fraction of the endoribonuclease Dicer –which is best known for its cytoplasmic role in microRNA processing– is present in the nucleus, phosphorylated upon DNA damage, recruited to DSBs, and processes damage-induced dsRNA [47,48]. Furthermore, she showed that Dicer depletion delays the DDR by impaired recruitment of repair factors MDC1 and 53BP1 [48]. She presented novel work pertaining to a novel noncanonical pathway of ncRNA processing by Dicer.

RNA-containing structures such DNA-RNA hybrids or ribonucleotide insertions have recently emerged as essential players in the maintenance of genome integrity. RNA in the form of R-loops is the subject of study of **Fred Chédin** lab (UC Davis, US). These 3-stranded nucleic acid structures are composed of a DNA-RNA hybrid and a displaced ssDNA strand which form co-

transcriptionally. R-loops have physiological roles such as transcription regulation but may also impose pathological consequences for the cell. The Chedin lab has pioneered several high-throughput technologies based on DRIP (DNA-RNA immunoprecipitation) such as DRIP-seq and DRIPc-seq to map R-loops at the single molecule level in the mammalian genome using the RNA-DNA hybrid specific antibody S9.6. Using these technologies, they have found that there are around 300 R-loops/per cell and that R-loops are on average ~300 bp in size. Chedin focused his talk on the physiological role of R-loops: these functions include helping to pause the RNA pol II as well as transcription termination. Interestingly, he introduced the concept that these structures can transiently absorb or relax negative super-coiling which in turn may impact promoter activation and/or replication origin firing. Super-helicity also drives R-loop formation; thus, the dynamic formation and resolution of R-loops may contribute to the regulation of gene expression genome-wide acting as an epigenetic mark [49,50].

R-loops can be a source of genome instability in particular in the context of replication, the focus of **Karlene Cimprich** team (Standford Univ., US). In their lab, they recently developed a technique based on DRIPseq called qDRIP which allows not only strand-specific mapping of DNA-RNA hybrids as DRIPc-seq but facilitates the comparison of the DNA-RNA hybrid content between different biological conditions using synthetic DNA-RNA hybrids as internal standards [51]. Although most laboratories have studied R-loops in the nuclear compartment, one of Cimprich's lab most surprising recent findings is the presence of R-loops also in the cytoplasm. The role of these R-loops, how they form and why is currently under investigation.

Contributing to the question of the origin of R-loops tackled by Chedin and Cimprich labs; **Benoit Palancade** (Institut Jacques Monod, FR) described a series of elegant experiments performed in the yeast *S. cerevisiae* model system to demonstrate that R-loops form preferentially in intron-less genes and reciprocally, intronic sequences protect from R-loop formation [52]. These findings have led his team to hypothesize that introns protect from genome instability. In his talk, Benoit Palancade also brought up one of the conundrums in the field that is whether DNA-RNA hybrids are an obstacle or a necessary intermediate for DSB repair: He discussed that nascent RNA at DSBs may serve as a template for repair by HR in a process he describes as transcription-associated recombination (TAR); whereas other labs have shown that the accumulation of DNA-RNA hybrids during transcription can have deleterious consequences for genetic integrity, as mentioned above. This conundrum was further discussed by **Aura Carreira** (Institut Curie, FR). Her lab recently focused on the RNA helicase DDX5, a novel partner of the breast cancer susceptibility protein BRCA2. Her team found that BRCA2 and DDX5 localize at DNA-RNA hybrids at induced DSBs of actively transcribed regions. Using a missense variant of BRCA2 detected in breast cancer patients that reduces the association between the two proteins, they could show that BRCA2 and DDX5 cooperate to resolve DNA-RNA hybrids at DSBs whereas in cells bearing the variant repair by HR is delayed [53]. Thus, in this scenario, DNA-RNA hybrids appear to be deleterious for repair by HR.

In conclusion, this course has underscored the multiple and reciprocal links between RNA biology and genome integrity (**Figure 1**). Indeed, DNA damage and replication stress widely impact gene expression and RNA metabolism at multiple post-transcriptional levels. Conversely, post-transcriptional regulation of DDR genes impacts genome integrity; moreover, RNA processing -which is extensively coupled to transcription and chromatin- is involved in both the generation and repair of DNA damage. The study of these links is a recently expanding field that enhances our understanding of genome biology (by integrating genome expression, replication and integrity) and sheds new light on the molecular mechanisms of cancer development and therapy.

References

[1] Martin RM, Rino J, Carvalho C, Kirchhausen T, Carmo-Fonseca M. Live-Cell Visualization of Pre-mRNA Splicing with Single-Molecule Sensitivity. Cell Reports 2013;4:1144–55. https://doi.org/10.1016/j.celrep.2013.08.013.

[2] Nojima T, Gomes T, Grosso ARF, Kimura H, Dye MJ, Dhir S, et al. Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. Cell 2015;161:526–40. https://doi.org/10.1016/j.cell.2015.03.027.

[3] Sousa-Luís R, Dujardin G, Zukher I, Kimura H, Weldon C, Carmo-Fonseca M, et al. POINT technology illuminates the processing of polymerase-associated intact nascent transcripts. Molecular Cell 2021;81:1935-1950.e6. https://doi.org/10.1016/j.molcel.2021.02.034.

[4] Casañal A, Kumar A, Hill CH, Easter AD, Emsley P, Degliesposti G, et al. Architecture of eukaryotic mRNA 3′-end processing machinery. Science 2017;358:1056–9. https://doi.org/10.1126/science.aao6535.

[5] Hill CH, Boreikaitė V, Kumar A, Casañal A, Kubík P, Degliesposti G, et al. Activation of the Endonuclease that Defines mRNA 3′ Ends Requires Incorporation into an 8-Subunit Core Cleavage and Polyadenylation Factor Complex. Molecular Cell 2019;73:1217-1231.e11. https://doi.org/10.1016/j.molcel.2018.12.023.

[6] Tang TTL, Stowell JAW, Hill CH, Passmore LA. The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. Nat Struct Mol Biol 2019;26:433–42. https://doi.org/10.1038/s41594-019-0227-9.

[7] Nourse J, Spada S, Danckwardt S. Emerging Roles of RNA 3′-end Cleavage and Polyadenylation in Pathogenesis, Diagnosis and Therapy of Human Disorders. Biomolecules 2020;10:915. https://doi.org/10.3390/biom10060915.

[8] Chalabi Hagkarim N, Grand RJ. The Regulatory Properties of the Ccr4–Not Complex. Cells 2020;9:2379. https://doi.org/10.3390/cells9112379.

[9] Ross NT, Lohmann F, Carbonneau S, Fazal A, Weihofen WA, Gleim S, et al. CPSF3-dependent pre-mRNA processing as a druggable node in AML and Ewing's sarcoma. Nat Chem Biol 2020;16:50–9. https://doi.org/10.1038/s41589-019-0424-1.

[10] Hir HL, Saulière J, Wang Z. The exon junction complex as a node of post-transcriptional networks. Nat Rev Mol Cell Biol 2016;17:41–54. https://doi.org/10.1038/nrm.2015.7.

[11] Kwon OS, Mishra R, Safieddine A, Coleno E, Alasseur Q, Faucourt M, et al. Exon junction complex dependent mRNA localization is linked to centrosome organization during ciliogenesis. Nat Commun 2021;12:1351. https://doi.org/10.1038/s41467-021-21590-w.

[12] Conn CS, Yang H, Tom HJ, Ikeda K, Oses-Prieto JA, Vu H, et al. The major cap-binding protein eIF4E regulates lipid homeostasis and diet-induced obesity. Nat Metab 2021;3:244–57. https://doi.org/10.1038/s42255-021-00349-z.

[13] Grosso S, Marini A, Gyuraszova K, Voorde JV, Sfakianos A, Garland GD, et al. The pathogenesis of mesothelioma is driven by a dysregulated translatome. Nat Commun 2021;12:4920. https://doi.org/10.1038/s41467-021-25173-7.

[14] Villanueva E, Smith T, Queiroz RML, Monti M, Pizzinga M, Elzek M, et al. Efficient recovery of the RNA-bound proteome and protein-bound transcriptome using phase separation (OOPS). Nat Protoc 2020;15:2568–88. https://doi.org/10.1038/s41596-020-0344-2.

[15] Mateju D, Chao JA. Stress granules: regulators or by-products? FEBS J 2021:febs.15821. https://doi.org/10.1111/febs.15821.

[16] Courel M, Clément Y, Bossevain C, Foretek D, Vidal Cruchez O, Yi Z, et al. GC content shapes mRNA storage and decay in human cells. ELife 2019;8:e49708. https://doi.org/10.7554/eLife.49708.

[17] Pushpalatha KV, Besse F. Local Translation in Axons: When Membraneless RNP Granules Meet Membrane-Bound Organelles. Front Mol Biosci 2019;6:129. https://doi.org/10.3389/fmolb.2019.00129.

[18] Formicola N, Heim M, Dufourt J, Lancelot A-S, Nakamura A, Lagha M, et al. Correction: Tyramine induces dynamic RNP granule remodeling and translation activation in the Drosophila brain. ELife 2021;10:e70755. https://doi.org/10.7554/eLife.70755.

[19] Boccaletto P, Machnicka MA, Purta E, Piątkowski P, Bagiński B, Wirecki TK, et al. MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Research 2018;46:D303–7. https://doi.org/10.1093/nar/gkx1030.

[20] Schwartz S. Cracking the epitranscriptome. RNA 2016;22:169–74. https://doi.org/10.1261/rna.054502.115.

[21] Sas-Chen A, Thomas JM, Matzov D, Taoka M, Nance KD, Nir R, et al. Dynamic RNA acetylation revealed by quantitative cross-evolutionary mapping. Nature 2020;583:638–43. https://doi.org/10.1038/s41586-020-2418-2.

[22] Michelena J, Lezaja A, Teloni F, Schmid T, Imhof R, Altmeyer M. Analysis of PARP inhibitor toxicity by multidimensional fluorescence microscopy reveals mechanisms of sensitivity and resistance. Nat Commun 2018;9:2678. https://doi.org/10.1038/s41467-018-05031-9.

[23] Lezaja A, Altmeyer M. Dealing with DNA lesions: When one cell cycle is not enough. Current Opinion in Cell Biology 2021;70:27–36. https://doi.org/10.1016/j.ceb.2020.11.001.

[24] Petryk N, Kahli M, d'Aubenton-Carafa Y, Jaszczyszyn Y, Shen Y, Silvain M, et al. Replication landscape of the human genome. Nat Commun 2016;7:10208. https://doi.org/10.1038/ncomms10208.

[25] Brison O, El-Hilali S, Azar D, Koundrioukoff S, Schmidt M, Nähse V, et al. Transcription-mediated organization of the replication initiation program across large genes sets common fragile sites genome-wide. Nat Commun 2019;10:5693. https://doi.org/10.1038/s41467-019-13674-5.

[26] Promonet A, Padioleau I, Liu Y, Sanz L, Biernacka A, Schmitz A-L, et al. Topoisomerase 1 prevents replication stress at R-loop-enriched transcription termination sites. Nat Commun 2020;11:3940. https://doi.org/10.1038/s41467-020-17858-2.

[27] He YJ, Meghani K, Caron M-C, Yang C, Ronato DA, Bian J, et al. DYNLL1 binds to MRE11 to limit DNA end resection in BRCA1-deficient cells. Nature 2018;563:522–6. https://doi.org/10.1038/s41586-018-0670-5.

[28] Clouaire T, Rocher V, Lashgari A, Arnould C, Aguirrebengoa M, Biernacka A, et al. Comprehensive Mapping of Histone Modifications at DNA Double-Strand Breaks Deciphers Repair Pathway Chromatin Signatures. Molecular Cell 2018;72:250-262.e6. https://doi.org/10.1016/j.molcel.2018.08.020.

[29] Aymard F, Aguirrebengoa M, Guillou E, Javierre BM, Bugler B, Arnould C, et al. Genome-wide mapping of long-range contacts unveils clustering of DNA double-strand breaks at damaged active genes. Nat Struct Mol Biol 2017;24:353–61. https://doi.org/10.1038/nsmb.3387.

[30] Marnef A, Cohen S, Legube G. Transcription-Coupled DNA Double-Strand Break Repair: Active Genes Need Special Care. Journal of Molecular Biology 2017;429:1277–88. https://doi.org/10.1016/j.jmb.2017.03.024.

[31] Marnef A, Finoux A-L, Arnould C, Guillou E, Daburon V, Rocher V, et al. A cohesin/HUSH- and LINC-dependent pathway controls ribosomal DNA double-strand break repair. Genes Dev 2019;33:1175–90. https://doi.org/10.1101/gad.324012.119.

[32] Arnould C, Rocher V, Finoux A-L, Clouaire T, Li K, Zhou F, et al. Loop extrusion as a mechanism for formation of DNA damage repair foci. Nature 2021;590:660–5. https://doi.org/10.1038/s41586-021-03193-z.

[33] Dutertre M, Lambert S, Carreira A, Amor-Guéret M, Vagner S. DNA damage: RNA-binding proteins protect from near and far. Trends in Biochemical Sciences 2014;39:141–9. https://doi.org/10.1016/j.tibs.2014.01.003.

[34] Dutertre M, Sanchez G, De Cian M-C, Barbier J, Dardenne E, Gratadou L, et al. Cotranscriptional exon skipping in the genotoxic stress response. Nat Struct Mol Biol 2010;17:1358–66. https://doi.org/10.1038/nsmb.1912.

[35] Dutertre M, Chakrama FZ, Combe E, Desmet F-O, Mortada H, Polay Espinoza M, et al. A recently evolved class of alternative 3′-terminal exons involved in cell cycle regulation by topoisomerase inhibitors. Nat Commun 2014;5:3395. https://doi.org/10.1038/ncomms4395.

[36] Tanaka I, Chakraborty A, Saulnier O, Benoit-Pilven C, Vacher S, Labiod D, et al. ZRANB2 and SYF2-mediated splicing programs converging on ECT2 are involved in breast cancer cell resistance to doxorubicin. Nucleic Acids Research 2020;48:2676–93. https://doi.org/10.1093/nar/gkz1213.

[37] Dutertre M, Sfaxi R, Vagner S. Reciprocal Links between Pre-messenger RNA 3′-End Processing and Genome Stability. Trends in Biochemical Sciences 2021;46:579–94. https://doi.org/10.1016/j.tibs.2021.01.009.

[38] Tufegdžić Vidaković A, Mitter R, Kelly GP, Neumann M, Harreman M, Rodríguez-Martínez M, et al. Regulation of the RNAPII Pool Is Integral to the DNA Damage Response. Cell 2020;180:1245-1261.e21. https://doi.org/10.1016/j.cell.2020.02.009.

[39] Williamson L, Saponaro M, Boeing S, East P, Mitter R, Kantidakis T, et al. UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene. Cell 2017;168:843-855.e13. https://doi.org/10.1016/j.cell.2017.01.019.

[40] Juszkiewicz S, Speldewinde SH, Wan L, Svejstrup JQ, Hegde RS. The ASC-1 Complex Disassembles Collided Ribosomes. Molecular Cell 2020;79:603-614.e8. https://doi.org/10.1016/j.molcel.2020.06.006.

[41] Wan L, Juszkiewicz S, Blears D, Bajpe PK, Han Z, Faull P, et al. Translation stress and collided ribosomes are co-activators of cGAS. Molecular Cell 2021;81:2808-2822.e10. https://doi.org/10.1016/j.molcel.2021.05.018.

[42] Michelini F, Pitchiaya S, Vitelli V, Sharma S, Gioia U, Pessina F, et al. Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. Nat Cell Biol 2017;19:1400–11. https://doi.org/10.1038/ncb3643.

[43] Pessina F, Giavazzi F, Yin Y, Gioia U, Vitelli V, Galbiati A, et al. Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors. Nat Cell Biol 2019;21:1286–99. https://doi.org/10.1038/s41556-019-0392-4.

[44] Aguado J, Sola-Carvajal A, Cancila V, Revêchon G, Ong PF, Jones-Weinert CW, et al. Inhibition of DNA damage response at telomeres improves the detrimental phenotypes of Hutchinson–Gilford Progeria Syndrome. Nat Commun 2019;10:4990. https://doi.org/10.1038/s41467-019-13018-3.

[45] Di Micco R, Krizhanovsky V, Baker D, d'Adda di Fagagna F. Cellular senescence in ageing: from mechanisms to therapeutic opportunities. Nat Rev Mol Cell Biol 2021;22:75–95. https://doi.org/10.1038/s41580-020-00314-w.

[46] Sharma S, Anand R, Zhang X, Francia S, Michelini F, Galbiati A, et al. MRE11-RAD50-NBS1 Complex Is Sufficient to Promote Transcription by RNA Polymerase II at Double-Strand Breaks by Melting DNA Ends. Cell Reports 2021;34:108565. https://doi.org/10.1016/j.celrep.2020.108565.

[47] Burger K, Gullerova M. Nuclear re-localization of Dicer in primary mouse embryonic fibroblast nuclei following DNA damage. PLoS Genet 2018;14:e1007151. https://doi.org/10.1371/journal.pgen.1007151.

[48] Burger K, Schlackow M, Potts M, Hester S, Mohammed S, Gullerova M. Nuclear phosphorylated Dicer processes double-stranded RNA in response to DNA damage. Journal of Cell Biology 2017;216:2373–89. https://doi.org/10.1083/jcb.201612131.

[49] Chedin F, Benham CJ. Emerging roles for R-loop structures in the management of topological stress. Journal of Biological Chemistry 2020;295:4684–95. https://doi.org/10.1074/jbc.REV119.006364.

[50] Ginno PA, Lott PL, Christensen HC, Korf I, Chédin F. R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. Molecular Cell 2012;45:814–25. https://doi.org/10.1016/j.molcel.2012.01.017.

[51] Crossley MP, Bocek MJ, Hamperl S, Swigut T, Cimprich KA. qDRIP: a method to quantitatively assess RNA–DNA hybrid formation genome-wide. Nucleic Acids Research 2020;48:e84–e84. https://doi.org/10.1093/nar/gkaa500.

[52] Bonnet A, Grosso AR, Elkaoutari A, Coleno E, Presle A, Sridhara SC, et al. Introns Protect Eukaryotic Genomes from Transcription-Associated Genetic Instability. Molecular Cell 2017;67:608-621.e6. https://doi.org/10.1016/j.molcel.2017.07.002.

[53] Sessa G, Gómez-González B, Silva S, Pérez-Calero C, Beaurepere R, Barroso S, et al. BRCA2 promotes DNA-RNA hybrid resolution by DDX5 helicase at DNA breaks to facilitate their repair‡. EMBO J 2021;40. https://doi.org/10.15252/embj.2020106018.

Figure legend

Figure 1: **Reciprocal links between RNA biology and genome integrity**. Some RNA molecules are produced at sites of double-strand DNA breaks and promote the recruitment of repair factors. R-loops, which are made of DNA-RNA hybrids and a displaced DNA strand, are involved in transcription-replication conflicts, leading to replication stress and DNA damage. Defects in pre-messenger RNA splicing and cleavage/polyadenylation, which are generally coupled to transcription, can favor DNA damage. DNA damage signaling widely impacts post-transcriptional steps of gene expression, including RNA splicing, cleavage/polyadenylation, modification, export, localization, translation and degradation, all of which contribute to the regulation of genes involved in the DNA damage response (DDR). DNA is depicted in blue and RNA in orange. DSB, double-strand DNA break. (Created with BioRender.com).