APRIL 25, 2018
VOLUME 118
NUMBER 8

pubs.acs.org/CR

CHEMICAL REVIEWS

RNA: FROM SINGLE MOLECULES TO MEDICINE

ACS Publications
Most Trusted. Most Cited. Most Read.

www.acs.org

CHEMICAL REVIEWS

Review

Cite This: *Chem. Rev.* 2018, 118, 4365−4403

pubs.acs.org/CR

From "Cellular" RNA to "Smart" RNA: Multiple Roles of RNA in Genome Stability and Beyond

Flavia Michelini,[*,†] Ameya P. Jalihal,[‡] Sofia Francia,[†,§] Chance Meers,[∥] Zachary T. Neeb,[⊥] Francesca Rossiello,[†] Ubaldo Gioia,[†] Julio Aguado,[†] Corey Jones-Weinert,[†] Brian Luke,[#,▽] Giuseppe Biamonti,[§] Mariusz Nowacki,[⊥] Francesca Storici,[∥] Piero Carninci,[○] Nils G. Walter,[‡] and Fabrizio d'Adda di Fagagna[*,†,§]

[†]IFOM − The FIRC Institute of Molecular Oncology, Milan, 20139, Italy

[‡]Single Molecule Analysis Group and Center for RNA Biomedicine, Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1055, United States

[§]Istituto di Genetica Molecolare, CNR - Consiglio Nazionale delle Ricerche, Pavia, 27100, Italy

[∥]School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

[⊥]Institute of Cell Biology, University of Bern, Baltzerstrasse 4, 3012 Bern, Switzerland

[#]Institute of Developmental Biology and Neurobiology, Johannes Gutenberg University, 55099 Mainz, Germany

[▽]Institute of Molecular Biology (IMB), 55128 Mainz, Germany

[○]RIKEN Center for Life Science Technologies, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama City, Kanagawa 230-0045, Japan

ABSTRACT: Coding for proteins has been considered the main function of RNA since the "central dogma" of biology was proposed. The discovery of noncoding transcripts shed light on additional roles of RNA, ranging from the support of polypeptide synthesis, to the assembly of subnuclear structures, to gene expression modulation. Cellular RNA has therefore been recognized as a central player in often unanticipated biological processes, including genomic stability. This ever-expanding list of functions inspired us to think of RNA as a "smart" phone, which has replaced the older obsolete "cellular" phone. In this review, we summarize the last two decades of advances in research on the interface between RNA biology and genome stability. We start with an account of the emergence of noncoding RNA, and then we discuss the involvement of RNA in DNA damage signaling and repair, telomere maintenance, and genomic rearrangements. We continue with the depiction of single-molecule RNA detection techniques, and we conclude by illustrating the possibilities of RNA modulation in hopes of creating or improving new therapies. The widespread biological functions of RNA have made this molecule a reoccurring theme in basic and translational research, warranting it the transcendence from classically studied "cellular" RNA to "smart" RNA.

CONTENTS

Special Issue: RNA: From Single Molecules to Medicine

Received: August 15, 2017
Published: March 30, 2018

ACS Publications © 2018 American Chemical Society

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

1. INTRODUCTION

In 1973 the world's first mobile phone call was made, giving birth to the era of cellular phones. Gradually, features such as text messaging, cameras, games, and music were added to the devices, but until recently these additions were considered extras with respect to the original main purpose of voice calling. During the past few decades, the integration of novel technologies and unprecedented connectivity into mobile phones catalyzed the paradigm shift from "cellular" to "smart" phones. Smartphones completely transformed consumer perception of their mobile devices, gradually becoming a virtual toolbox with a solution for almost every need.

A similar shift in perception has occurred within the scientific community during the last 60 years, surpassing the original view of RNA in Crick's "central dogma" as solely the messenger of genetic information. Initial discoveries of noncoding RNAs (ncRNAs) having a biological function independent from protein coding included tRNA (tRNA), rRNA (rRNA), and spliceosomal RNA. Since then, the list of additional roles assigned to transcripts has grown exponentially. Although the vast majority of the genome is transcribed,[1] current estimates indicate that only about 1.5% of it codes for proteins. It is now becoming evident that this nucleic acid is an extremely versatile molecule implicated in many different cellular processes, from structural support, to epigenetic modulation of gene expression, to maintenance of genome integrity. Consequently, numerous links between defects in noncoding RNA and human diseases have been described.[2] Thus, RNA has broken free from its original confined role of subordinate messenger for DNA to emerge as an indispensable smart tool for a multitude of cellular needs.

In this review, we cover the very topical notion that RNA, both coding and noncoding, is involved in the maintenance of genomic stability as an example of "smart RNA". We begin with a historical perspective on the emergence of the noncoding functions of RNA and of the RNA interference machinery. We discuss novel aspects of a recently discovered class of RNAs involved in DNA damage signaling and DNA repair,[3−7] as well as RNAs that guide genomic rearrangements[8,9] and maintain telomere homeostasis.[10−12] We then highlight the importance of using cutting-edge, single-molecule resolution techniques to study the location and biology of low-abundance, highly specialized RNA molecules. Finally, we discuss the exciting potential of targeting such RNAs with antisense tools as a viable therapeutic option. Throughout the article, we guide the reader to additional reviews that describe certain aspects in greater detail that, due to spatial constraints, we only mention.

Unlike the case of smartphones, the multiple functions of RNA are innate. In this respect, cellular RNA has always been smart, we just had to realize it.

2. LIVING IN AN RNA WORLD

2.1. Gene Hunting during the Genome Revolution

The way we study biology has dramatically changed in the past 20 years due to the genome revolution. Genome sequencing efforts have been paralleled by efforts to identify encoded genes. Although the scientific community has long been aware of the many different functions of RNA (ranging from ribozymes to splicing), the search for novel genes was strongly biased toward the identification of protein-coding elements because of the assumption that they would be the main components of the newly sequenced genomes. Dogmas, such as "one gene, one enzyme"[13] or "one gene, one polypeptide", have influenced the community far beyond their original message, long reinforcing a narrow protein-centric view of genetic information.

The quest for genes ended with the surprising finding that the human genome, as well as the genome of other vertebrates, encodes for only slightly more than 20,000 protein-coding genes.[14,15] This small number was a big surprise due to previous estimates suggesting that the human genome would contain more than 100,000 genes,[16] commensurate with the expectation for organisms such as humans that have a complex developmental body plan and central nervous system. For instance, the genome of *Caenorhabditis elegans* contains 19,000 protein-coding genes,[17] while the fruit fly *Drosophila melanogaster* genome has 14,000.[18] Thus, the number of protein-coding genes is similar in invertebrate and mammalian genomes. Next, the scientific community postulated that alternative splicing and combinatorial transcriptional control by transcription factors may contribute to human body complexity. However, these two phenomena are not sufficient to explain differences in complexity as they are also found in invertebrates.[19] Nonetheless, mainstream genome analysis has continued to focus on protein-coding genes for years, mostly for practical reasons: they are relatively easy to identify given their high expression, long open reading frames, and defined start and termination points.[20,21]

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

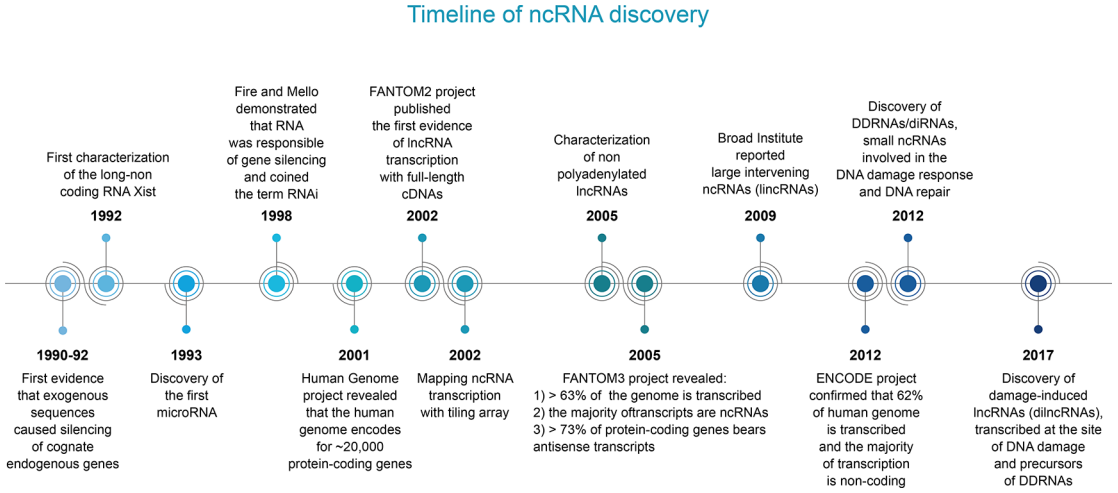

Figure 1. Timeline of the major breakthroughs in the "smart" RNA world, as discussed in this review.

2.2. Discovery of RNA Dark Matter

In the quest for protein-coding genes, efforts have been conducted to identify expressed polyadenylated RNAs.[21,22] In retrospect, the selection of cDNA that showed an open reading frame as evidence of protein-coding genes has caused a long delay in the discovery of human non-protein-coding transcripts longer than 200 nt, named long noncoding RNAs (lncRNAs)—this classification simply comes from experimental restrictions inherent to sequencing library preparation.

The concept that the genomes of organisms with high complexity are largely transcribed and that the main output is comprised of ncRNAs derives from analysis of mouse full-length cDNA collections and whole genome high-density tiling arrays, with a series of parallel studies identifying mounting evidence of widespread transcription. In 2002 (Figure 1), the FANTOM2 project found the first evidence of lncRNA transcription and also identified ~2500 antisense RNAs.[20] Subsequently, it became clear that nuclear, nonpolyadenylated lncRNAs are the major output of the genome.[23,24] In 2005 (Figure 1), the FANTOM3 project reached the conclusion that at least 63% of the genome is transcribed, most of which is comprised of ncRNAs.[1] Additionally, by cap analysis gene expression (CAGE) technology,[25,26] it has been shown that at least 73% of loci encoding for protein-coding genes are also transcribed from the antisense strand. This antisense transcription often results in products that regulate the activity of their respective sense mRNAs.[27] Furthermore, one of the many early high-throughput RNA sequencing techniques, which are collectively named next generation sequencing, unexpectedly found that even retrotransposon element (RE) expression is tightly regulated in mammalian cells and tissues, producing lncRNAs[28] which may in turn contribute to regulate RE expression. These and other studies also suggest that lncRNAs are somehow expressed at lower levels than protein-coding mRNAs, are often localized in the nucleus, and generally display time-, tissue-, or even cell-specific expression.

Despite initially being met with a large amount of healthy skepticism, the findings that pervasive transcription is one of the most abundant products of the genome have been supported by several prominent studies.[29] For instance, in the ENCODE project (Figure 1), next generation CAGE RNA-sequencing confirmed that 62% of the human genome is transcribed and that the main output is noncoding RNAs.[30,31] While experimental approaches clearly demonstrated that lncRNAs are indeed transcribed and display specific functions, the interpretation of their biological significance has proven difficult due to their low expression, stability, and conservation, relative to protein-coding RNAs.[32] Future studies will be required to disentangle the multiple functions of RNA, which therefore deserves the title "smart" as proposed in our review.

Interestingly, the early identification of functional small RNAs, or those less than 200 nt long, was readily accepted by the community. This acceptance was so easily given likely due to the previous discovery of the RNA interference (RNAi) pathway (see section 3).

The discovery of siRNA and miRNA spurred the identification of numerous other classes of small RNAs featuring defined lengths and functions. piRNAs, for example, are 28−29 nt long sncRNAs that associate with Miwi, Mili, and Piwi proteins and function to repress the transcription and mobilization of RE, thus contributing to preserve genome integrity in the germline.[89] While there is broad acceptance that piRNAs have a fundamental role in the repression of transposable elements, it is unclear why RE expression is not uniformly silenced, allowing for regulated RE activity during embryonic development,[90] in embryonic stem cells, and in induced pluripotent stem cells.[91]

Among the vast landscape of small RNAs, some of them do not show a specific length and their interacting partners are less characterized. For example, small RNAs that overlap with transcription starting sites (TSSs), known as promoter associated RNAs (PASRs), and transcription termination sites (TTSs), named termination associated RNAs (TASRs), have been identified but poorly characterized. Some of them have been implicated in regulation of transcription, but much remains to be learned about their biology.[92] One standing question regarding PASRs is to which extent they overlap with PROMPTs, a class of unstable RNA degraded by the exosome machinery.[93]

A large fraction of the small ncRNAs derives from processing of lncRNA precursors; thus, the regulation of lncRNA transcription and processing plays an important role in many aspects of small RNA biology. A good example is a novel class of small ncRNAs named DNA damage response RNAs, or DDRNAs, involved in the cellular response to DNA damage and in DNA repair (Figure 1).[61,94] DDRNAs are DROSHA- and DICER-dependent products of damage-induced lncRNAs, or dilncRNAs, transcribed by RNA polymerase II (RNAP II) at the site of DNA damage[95] (see section 4.2.2 for details).

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

Due to space constraints, not all known classes of short and long noncoding RNAs could have been discussed here, as the field has dramatically expanded in recent years.

2.3. Function Does Not Always Mean Protein-Coding

The concept that the protein-coding-centric view cannot explain vertebrate/mammalian complexity came from the pioneering insights of John Mattick (Garvan Institute of Medical Resarch, Australia). By analyzing the noncoding content of all available genomes, Mattick noticed that the fraction of noncoding DNA in a genome increases progressively from ~30% in prokaryotes to a staggering 98.5% in humans. This observation is in line with a role for many noncoding regions, that are in fact largely transcribed, in fine-tuning protein production during development and in participating in tissue homeostasis maintenance in higher eukaryotes and in particular in mammals.[19]

In parallel, genetic studies also established that a large fraction of the genetic information falls outside the boundaries of the exons of protein-coding genes. Genome wide association studies (GWAS) have identified a plethora of single nucleotide polymorphism (SNPs) associated with a large variety of human non-Mendelian diseases, the majority of which falls outside protein-coding regions, often mapping to novel promoter elements and enhancers.[96] Enhancer regions are sources of another class of lncRNAs called enhancers RNAs (eRNAs). Although their functions are not fully known, in some cases eRNAs are involved in transcription activation through chromatin looping.[97] Further, recent studies have identified lncRNAs putatively involved with the molecular cause of some human diseases.[98] Altogether, genetic evidence, together with sequence conservation at promoter or exons of lncRNAs, suggests that at least 19,000 human lncRNAs may be functional. A very recent study supports a cell-type specific role for lncRNAs in transcriptional activation.[99] In contrast with many small RNAs, such as miRNAs, lncRNAs lack a common, unified function, therefore requiring more intensive investigations to address their activity.

Many lncRNAs are restricted to the nucleus, where some have structural roles, as in the case of paraspeckles, subnuclear bodies constituted by ncRNAs and proteins.[100] Other nuclear lncRNAs are associated with chromatin where they contribute to regulation of the epigenome. A few examples include *HOTAIR*,[101] *XIST*,[102] and lncRNAs associated with imprinted loci. Novel technologies will be essential to map specific RNA-chromatin interactions and to dissect all the functions of chromatin-bound lncRNAs.

Antisense transcription is another essential source of ncRNA, impacting either positively or negatively on canonical gene expression.[103] When antisense RNAs are exported to the cytoplasm, they can also regulate RNA stability[27] as well as protein translation. For example, one peculiar class of antisense RNAs that act as translation regulators, SINEUPs, enhance the translation of the mRNAs they overlap through a SINE element embedded in the nonoverlapping part of the antisense.[104,105] Curiously, various lncRNAs may act as SINEUPs, independent of their origin, suggesting that RNA structure, rather than its primary sequence, is most important.[106]

It is therefore becoming apparent that ncRNAs play diverse and important functions in the cell. An additional layer of complexity is given by the unexpected engagement of ncRNAs, and components of their pathways, in other fundamental mechanisms of the cell, such as the response to DNA damage.

3. RNA INTERFERENCE PATHWAY

RNAi, the process by which RNAs inhibit gene expression by sequence-specifically base-pairing with other RNAs, was initially described in plants and fungi as a peculiar yet effective mechanism to preserve genome integrity and protect against viruses and transposons.[33,34] Later, RNAi was detected in a broad variety of other eukaryotic organisms[35−39] and acknowledged as a more general strategy through which cells finely tune gene expression at the post-transcriptional level. By now, RNAi has been used for over two decades as a tool to study and manipulate gene function.

The first evidence that a long double-stranded RNA (dsRNA) was responsible for triggering sequence-specific silencing of a target gene was provided in 1998 by Andrew Fire and Craig Mello, who coined the term RNA interference.[35] Shortly after, other groups proposed a different model for this phenomenon in which small RNAs, released by cleavage of long dsRNA precursors, were the actual effectors of the post-transcriptional gene silencing; hence, they named them short interfering RNAs (siRNAs).[40−43]

siRNAs are double-stranded RNA molecules, 20−25 base pairs in length, known to cause the degradation of the perfect complementary target RNA. siRNAs can be produced from RNA transcribed in the nucleus (endogenous siRNAs), or they can be virally derived or experimentally introduced as chemically synthesized dsRNA (exogenous siRNAs). Endogenous siRNAs have been described in plants and in *C. elegans*, and they can originate from overlapping sense and antisense transcripts[44] or from repeat-associated genomic regions.[45] To exert their function, siRNAs must be unwound and loaded into the RNA-induced silencing complex (RISC). The RISC complex contains Argonaute (AGO) proteins, which display the endonucleolytic activity responsible for cleavage of the target RNA. In mammals, there are four AGO proteins (1−4) that can participate to the RISC complex, but only AGO2 is catalytically active and functions for the direct degradation of the target mRNA.[46] Only one of the two strands of the siRNA duplex (the guide strand) is loaded into the RISC complex, whereas the other strand, known as the passenger strand, is released and degraded.[47] Endogenous siRNAs are thought to play an important role in defending genomes against transposable elements, as well as foreign nucleic acids, such as viruses.

Another class of endogenous small RNAs also capable of eliciting RNAi was discovered by Victor Ambros and colleagues in 1993, and they were later named microRNAs (miRNAs).[48−51] miRNAs have been revealed to play important roles in almost every cellular process investigated.[52] The biogenesis of most miRNAs requires the RNase III DICER and the Microprocessor complex, which is composed of the other RNase III DROSHA and the dsRNA binding protein DGCR8 (DiGeorge syndrome critical region 8).[53] In the canonical pathway for miRNA biogenesis, an RNA polymerase II (RNAP II) dependent, single stranded, and capped primary RNA (pri-miRNA) is first processed by the Microprocessor complex in the nucleus, transforming it into a ~70-nucleotide hairpin-structured precursor RNA (pre-miRNA), which is then exported to the cytoplasm. Interestingly, some pre-miRNAs are produced from very short introns, called mirtrons, as a result of splicing and debranching,[54] thereby bypassing the requirement of the Microprocessor complex. In the cytoplasm, cleavage by DICER, that works together with TRBP (transactivation-responsive RNA binding protein) and PACT (protein activator

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

of PKR),[55] results in a 20−23 nt miRNA duplex.[56] Differently from siRNAs, miRNAs can exert their function either by triggering the degradation of the cognate mRNAs or by preventing their translation. Efficient mRNA targeting requires base-pairing of nucleotides 2 to 8 at the 5′ end of the miRNA, the so-called "seed region", with the target mRNA. The degree of complementarity between the seed region and target mRNA determines if silencing is induced through translational repression, the potential results of imperfect complementarity, or through cleavage, the result of perfect complementarity. As for siRNAs, the guide strand is preferentially incorporated into the RISC complex, which came to be known as "miRISC" following the discovery of its association with miRNA. In some cases, the passenger strand (designated as miRNA*) can also enter the miRISC complex to guide gene silencing. The miRISC complex also contains members of the GW182 (glycine-tryptophan protein of 182 kDa) family, which coordinate translational inhibition and the consequent mRNA poly(A)-tail shortening.[57]

Components of RISC are thought to localize and function just in the cytoplasm. However, in human cells RNAi has been demonstrated to mediate repression of target RNAs in the nucleus as well.[58]

3.1. DICER, DROSHA, and DGCR8: Emerging New Roles in Transcription Regulation

A growing body of evidence has unveiled novel miRNA-independent functions for DICER and the Microprocessor complex, ranging from the maintenance of genome integrity to the modulation of alternative splicing.[59−64] Here, we focus on the unanticipated roles played by the Microprocessor complex and DICER in the regulation of transcription with important implications in controlling genome stability, sometimes independently from small RNA generation.

The function of eukaryotic RNAP II is not limited to faithful copy of the information encoded in the genome, but it takes part in crosstalk with a myriad of other factors involved in the excision, addition, and editing of ribonucleotides in the nascent transcript.[65] Among these factors, the Microprocessor complex and DICER turned out to be talkative interlocutors of the RNAP II machinery.

Microprocessor, initially found to be cotranscriptionally recruited to miRNA-encoding genomic loci,[66] has been lately shown to localize at many different non-miRNA genes, including at superenhancers.[67−70] In addition, the affinity of Microprocessor for hairpin structures in nascent RNA was demonstrated to be exploited by the cell to promote premature transcription termination of endogenous retroviral genes via stem-loop excision, a process independent from mature miRNA production.[70] This Microprocessor-mediated cut inevitably generates an additional 3′-end in the nascent transcript, consequently providing an early potential alternative transcription termination site and thus inhibiting retroviral gene expression. Interestingly, accurate transcription termination of many miRNA-containing lncRNAs relies on Microprocessor endonucleolytic activity rather than the canonical cleavage-and-polyadenylation pathway.[71] Moreover, human DROSHA has also been shown to enhance the expression of a subset of coding genes. Intriguingly, while transcriptional regulation exerted at these loci depends on DROSHA ability to interact with RNAP II, its catalytic activity is instead dispensable.[67]

It is now well established that many RNAi factors are not relegated to the cytoplasm, as initially proposed, but they are functionally active also in the nucleus of different eukaryotes, whereby they guide transcriptional gene silencing (TGS) through the deposition of repressive chromatin marks at silenced loci.[58,72] In recent years, a number of reports revealed that Dicer also plays direct roles in nuclear transcriptional regulation. For example, it has been shown that *S. pombe* mutants lacking Dicer (Dcr1) failed to remove stalled RNAP II at sites of collision between transcription and DNA replication, resulting in accumulation of recombinogenic DNA−RNA hybrids[73−77] and consequent genome instability. Interestingly, the catalytic-dead Dcr-1 mutant was still able to release RNAP II from these loci, suggesting that in this system the ability of Dicer to promote transcriptional termination is independent from the biogenesis of sncRNAs, similarly to the nuclease-independent function of Drosha in transcriptional regulation described above. Nevertheless, Dcr1-mediated sncRNAs could be detected at these loci in wild type strains, though their physiological role remains elusive.[78] The presence of Dicer in the nucleus of mammalian cells is instead a subject of debate. While murine DICER seems to be circumscribed to the cytoplasm,[79] several reports demonstrated its presence in the nuclei of human cells.[80−87] Hence, the functions played by DICER in human nuclei have started to be elucidated only recently. For example, human nuclear DICER was reported to localize to chromosomal regions with paused RNAP II, specifically in the proximity of transcription start sites and polyadenylation signals (PASs).[84−86] Occasionally, these sites were found to be associated with R-loops—three-stranded nucleic acid structures composed of a DNA−RNA hybrid and the displaced single-stranded DNA—that triggered the transcription of antisense RNAs.[85] The resulting dsRNA formation in turn could lead to the recruitment of DICER, together with other RNAi factors, and the consequent formation of heterochromatin at RNAP II-paused sites, ultimately enforcing TGS.[85,86] The presence of DICER at specific PASs, besides its role in transcriptional repression, suggests it may also control alternative transcription termination since DICER-dependent deposition of repressive chromatin marks surrounding such PASs may decrease RNAP II speed, ultimately imposing altered transcription termination at these sites and the production of alternative RNA variants.[84]

The Microprocessor complex and DICER are versatile factors, acting, in addition to their canonical roles, as nuclear transcriptional fine-tuners. Importantly, the ability of DROSHA, DGCR8, and DICER to slow down the transcription rate and mitigate DNA−RNA hybrid accumulation, which represents an intrinsic threat for genome integrity,[88] suggests a fail-safe mechanism in genome maintenance.

4. CROSSTALK BETWEEN ncRNA, RNAi, AND THE CELLULAR RESPONSE TO DNA DAMAGE

4.1. DNA Damage Response (DDR) Cascade and the Moonlighting Functions of DDR Factors as RNA Binding Proteins

The integrity of our genome is constantly threatened by endogenous and exogenous agents.[107] Cells have evolved a coordinated set of events to recognize the damage and promptly fix it, thus avoiding the replication and perpetuation of a compromised template. The DNA damage response (DDR) cascade is dependent on a broad variety of post-translational modifications such as phosphorylation, ubiquitination, sumoylation, poly(ADP-ribosylation), acetylation, and methylation.[107−110] These modifications are recognized by specific protein domains, thereby orchestrating the recruitment of DDR

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

factors to the DNA damage sites and ultimately the spreading of the signal throughout the cell. If the lesion cannot be repaired, persistent DDR activation may induce cell death by apoptosis or a permanent cell-cycle arrest called cellular senescence, both of which are known cellular intrinsic barriers to tumorigenesis.[111]

Until recently, the DDR signaling cascade was thought to consist entirely of proteins. The discovery of novel species of small ncRNAs directly implicated in upstream activation of the DDR in 2012[61,94] and of long ncRNAs induced at the site of DNA breaks in 2017[95] has radically changed this perspective (see section 4.2 for details). Moreover, other noncoding RNAs have been shown to serve as templates for DNA repair (see section 6 for details) or to guide genomic rearrangements (see section 7 for details). Interestingly, during the past decade, large-scale proteomic analyses and genome-wide screens have revealed that an unexpected proportion of RNA-binding proteins (RBPs) and proteins involved in transcription are involved in the DDR[112] and that, conversely, factors originally discovered as guardians of genomic integrity show an affinity for RNA.[113] Indeed the number of dual DNA- and RNA-binding proteins has surprisingly grown.[114]

Thus, the unexpected relationship between proteins involved in the DNA damage response and RNA may be important for the maintenance of genome stability.

4.1.1. DNA Damage Response to DNA Double-Strand Breaks. Among the different types of lesions that may threaten our genome, DSBs are the most dangerous since they may lead to loss of genetic materials and chromosomal rearrangements, predisposing cells to malignant transformation. The repair of a DSB relies on either homology-dependent or -independent mechanisms.[115,116] Homologous recombination (HR) is a homology-dependent and error-free mechanism that requires a homologous template, usually a sister chromatid, which allows accurate repair of postreplicative DSBs during S and G2 phases of the cell-cycle.[117] In contrast, classical nonhomologous end joining (C-NHEJ) is a homology-independent mechanism active throughout the entire cell-cycle; although highly efficient, its very simple mechanism of basic religation, without proof-reading, makes NHEJ amenable to errors and thus to introduce mutations.[118] NHEJ involves no or limited processing of DNA ends, while HR requires the formation of 3′ single-stranded overhangs. Thus, a critical step for the cellular choice between the two pathways is the DNA end resection.

DSBs are powerful activators of two large serine/threonine phosphatidylinositol 3-kinase-related kinases (PIKKs): ataxia telangiectasia mutated (ATM) and DNA-dependent protein kinase (DNA-PK)[119]). Exposure of single-stranded DNA is instead recognized by a third PIKK named ataxia telangiectasia and Rad3-related (ATR).[119] The recruitment of these apical kinases to the lesions leads to the local phosphorylation *in cis* of the histone variant H2AX at serine 139 (named γH2AX), a key step in the nucleation of the DDR. Following the first burst of H2AX phosphorylation, the ATM kinase phosphorylates many substrates, including mediator of DNA damage checkpoint 1 (MDC1) and p53 binding protein 1 (53BP1). These phosphorylation events fuel a positive feedback loop that facilitates the recruitment of additional ATM molecules to the DSB site[120,121] and the spreading of γH2AX up to megabases away from the lesion. This signal amplification results in the accumulation of numerous copies of various DDR factors at and flanking the DSB, forming cytologically visible foci.

The localization of DDR factors to DSBs has been described as a two-phase process in which the initial recruitment occurs by the direct recognition of the DNA lesion in a γH2AX-independent manner, followed by accumulation of DDR proteins at the damaged site in a γH2AX-dependent manner.[122,123] During the first phase, the DNA ends are promptly recognized by specialized factors: the KU70/KU80 heterodimer (or KU), the MRE11-RAD50-NBS1 (MRN) complex, or poly(ADP-ribose) polymerases (PARPs). A precise distinction of instances in which different DSB sensors are individually engaged, the timing of their recruitment, and whether they cooperate or compete with each other for the same DNA end are all issues that are only now becoming clearer. Likely, the nature of the break, the chromatin environment, the cell-cycle phase during the damaging event, and the cell type are all elements to take into account when attempting to answer these open questions.

4.1.1.1. KU. The KU70/KU80 heterodimer (KU) is a ring-shaped complex that encloses the DNA ends and recruits the DNA-PK catalytic subunit (DNA-PKcs), which phosphorylates itself, H2AX, and other targets, thereby initiating the classical nonhomologous end joining (C-NHEJ) repair pathway.[119] It has been reported that KU binds to the RNA component of telomerase both in yeast[124−126] and in human cells[127] (see section 8.1 for details about KU and telomere maintenance). KU has been found at promoter regions regulating gene expression[128,129] and also in a complex with elongating RNAP II.[130] Finally, it has been shown that treating pre-extracted cells with RNase A increases the detection by immunofluorescence of KU and other NHEJ factors at DSBs, indicating that a large fraction of these proteins are bound to RNA in the cell.[131]

4.1.1.2. MRN. The MRE11-RAD50-NBS1 (MRN) heterotrimeric complex plays important roles in detection and signaling of DSBs, as well as in initial processing of DNA ends prior to repair. The cohesin-like RAD50 protein interacts with MRE11 via its ATPase domains forming the globular head of the complex, and by dimerizing it ensures stable clamping and tethering of the complex to DNA ends.[132] NBS1 interacts with MRE11 and is instead responsible for the nuclear localization of the complex and for the local recruitment of ATM to DSBs where it gets activated. MRE11 possesses 3′-to-5′ exonuclease and 5′ overhang endonuclease activities, which, together with the auxiliary endonuclease C-terminal binding protein interacting protein (CtIP), are essential for the initial steps of DNA end resection. In mammalian cells, the MRN complex has not been involved in the C-NHEJ; however, together with CtIP, it regulates alternative NHEJ (alt-NHEJ), which utilizes short microhomologies to direct repair.[132]

Given the distinct pathways in which they act, MRN and KU were considered mutually exclusive at DNA ends. However, by high resolution microscopy a certain degree of colocalization between KU and MRN at individual DNA ends was observed.[131] Very recently, it has been shown that MRN can indeed access KU-blocked DNA ends by diffusion onto nucleosome-coated DNA.[133] Excitingly, two independent studies reported that KU functions as a protein block stimulating yeast MRN-CtIP endonuclease cleavage *in vitro*.[134,135]

So far, there is no evidence suggesting that MRN is capable of binding RNA. However, MRN does appear to be involved in RNAP II transcription following DNA damage. Indeed, RAD50 has been found to interact with RNAP II upon UV-irradiation.[136] Additionally, it has been demonstrated that all three components of the complex bind to RNAP II upon ionizing radiations and are important for damage-induced transcription at DSBs[95] (see section 4.2.2 for details).

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

4.1.1.3. PAR. PARylation is the process by which poly(ADP-ribose) polymerases (PARPs) covalently attach (poly)ADP-ribose (PAR) units to Glu, Lys, or Asp residues of acceptor proteins or PARP itself. The activity of the three major PARPs, PARP1, PARP2 and PARP3, is induced by DNA damage: PARP1 is activated by single-strand breaks, DNA cross-links, stalled replication forks, and DSBs; PARP2 recognizes gaps and flap structures; PARP3 selectively responds to DSBs.[137,138] PARPs are efficiently and transiently recruited to DSBs, for example during the first 5 min of laser microirradiation, where they have been proposed to orchestrate chromatin decondensation and the subsequent accessibility to the damage sites of a variety of factors, ranging from chromatin remodelers to transcription factors.[139] Indeed PARPs are known to promote chromatin decompaction at promoters and to facilitate the loading of RNAP II machinery at transcription start sites.[140]

PAR is a nucleic acid-like molecule, and it can be recognized by RNA-binding domains such as the RNA recognition motif (RRM) and the RGG motifs, regions rich in arginines (R) and glycines (G) present in several RNA-binding proteins. Thus, it is not surprising that PAR and RNA can compete for the same RBPs; an example of this is the case of the RNA-binding protein NONO which is recruited to DSBs in a PAR-dependent manner.[141] Therefore, changes in the local concentration of either of these molecules may dynamically alter the assembly of protein complexes, which, in turn, may affect cellular processes including the DDR.[142]

A good example of a factor involved in the DDR that can promiscuously bind to PAR chains and to RNA molecules is the heterochromatin protein HP1.[143,144] HP1 has been shown to rapidly localize to sites of DNA damage in a PAR-dependent manner, subsequently being displaced and then slowly recruited again.[145] Interestingly, HP1 requires an RNA component to bind to pericentric heterochromatin,[146] and its hinge domain can bind both DNA and RNA.[147]

Another example is FUS/TLS (fused in sarcoma/translocated in liposarcoma), a member of the FET family of RNA/DNA binding proteins,[148] which is a multifunctional factor with reported roles in splicing, transcription, mRNA export and translation, and the DDR. Initially identified as a fusion oncoprotein, FUS was later implicated in neurodegenerative diseases such as amyotrophic lateral sclerosis and frontotemporal lobar degeneration.[149] In response to DNA damage, FUS is rapidly and transiently recruited to DSBs, likely through the interaction between its RGG domain and PAR.[150−152] Interestingly, the same domain mediates the recruitment of FET proteins to paraspeckles by direct binding to the lncRNA NEAT1.[153] In the absence of FUS, the localization of some DDR factors, such as 53BP1, to the site of damage is reduced and the efficiency of both HR and NHEJ is compromised.[150,154] It has been shown that DSBs trigger local ncRNA transcription[95](see sections 4.2.2 for details). Given the ability of FUS to bind RNA, an exciting possibility is that FUS accumulation at DSBs could be modulated by RNA, in synergy or in competition with PAR chains.

The PAR-dependent localization to sites of DNA damage of several RNA binding factors occurs in two steps: beginning with a transient recruitment and ending with exclusion.[141,150,155] It is therefore tempting to speculate that such dynamic behavior can also be mediated by the damage-induced lncRNAs (see section 4.2.2 for details), possibly due to waves of transcription correlating with the bimodal dynamics of chromatin relaxation and compaction at the site of break.

4.1.1.4. DNA Damage Mediators: MDC1, 53BP1. MDC1, which directly binds γH2AX, has surprisingly been found in an RNA interactome capture screen, an unbiased approach to identify protein−polyA RNA direct interactions.[156] Interestingly, MOF (orthologue of *Drosophila* males absent on the first, or MYST1), a histone acetyltransferase shown to be important for MDC1 localization to DSBs, can also bind to ncRNA.[157] Moreover, it has been observed that irradiation-induced MDC1 foci are reduced upon treatment with RNase A or in the absence of DICER or DROSHA[61,158](see section 4.2.1 for details).

Recruitment of 53BP1 to damaged DNA requires the presence of both monoubiquitinated H2A on lysine 15 (H2AK15ub) and dimethylated histone H4 on lysine 20 (H4K20me2).[110] 53BP1 associates with methylated histones through its tandem Tudor domain, which is usually found in RNA-binding proteins. Indeed, it has been shown that 53BP1 can be immunoprecipitated together with RNA molecules from cell lysates, and RNase A treatment in permeabilized living cells dissociates 53BP1 from IR-induced foci, which can reassemble in an RNA-dependent manner.[61,158,159] Moreover, 53BP1, together with components of the C-NHEJ pathway, has been found in a complex with RNAP II in human cells[160] (see section 6.3 for details). More recently, 53BP1 has been demonstrated to immunoprecipitate selectively with dilncRNAs and DDRNAs generated at sites of DSBs, in a manner dependent on its tandem Tudor domain[95]-(see section 4.2.2 for details).

These examples, though not an exhaustive list, are those that best point to the emerging evidence of an intimate and complex relationship between the DDR factors and RNA, and likely more is yet to come.

4.2. Discovery of Damage-Induced Transcription at the Site of DNA Breaks

4.2.1. DNA Damage Response Small ncRNA (DDRNA).

Emerging evidence suggests that noncanonical transcription, in the form of damage-induced small noncoding RNAs (sncRNAs), occurs at DNA damage sites.[3−5,7,73,75] The existence of ncRNA species induced upon DNA damage and the involvement of RNA in DNA repair processes were originally reported in lower organisms such as yeast and fungi. For example, in *Neurospora crassa*, quelling and DNA damage-induced small RNAs (qiRNAs) are produced upon treatment with DNA damaging agents. qiRNA biogenesis involves a single-stranded precursor, called aberrant RNA (aRNA),[161] which is converted into double-stranded RNA by RNA-dependent RNA polymerase (RdRP) activity and processed into small RNA through the same mechanisms that generate RNAi. qiRNA generation shares the same genetic requirements of the HR pathway; indeed, it depends on on replication protein A (RPA)[161] and DNA replication.[162] Mutations of genes involved in qiRNA biogenesis sensitize *Neurospora* strains to DNA damage.[163]

qiRNA seem to be mainly induced from repetitive or foreign sequences. Indeed they are transcribed from rDNA, but also of multiple copies of transgenes, acting as transgene-specific endogenous siRNA counteracting the expansion of selfish genetic elements.[164] A similar phenomenon has also been described in rice where rDNA exposed to DNA damage locally generates high levels of RecQ DNA helicase- and RdRP-dependent double-stranded sncRNAs, which are required for cell viability after DNA damage exposure.[165]

Interestingly, in the yeast *Saccharomyces cerevisiae*, which lacks RNAi machinery,[166] pre-existing RNA can serve as the template for DNA synthesis during repair of a chromosomal DSB and

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

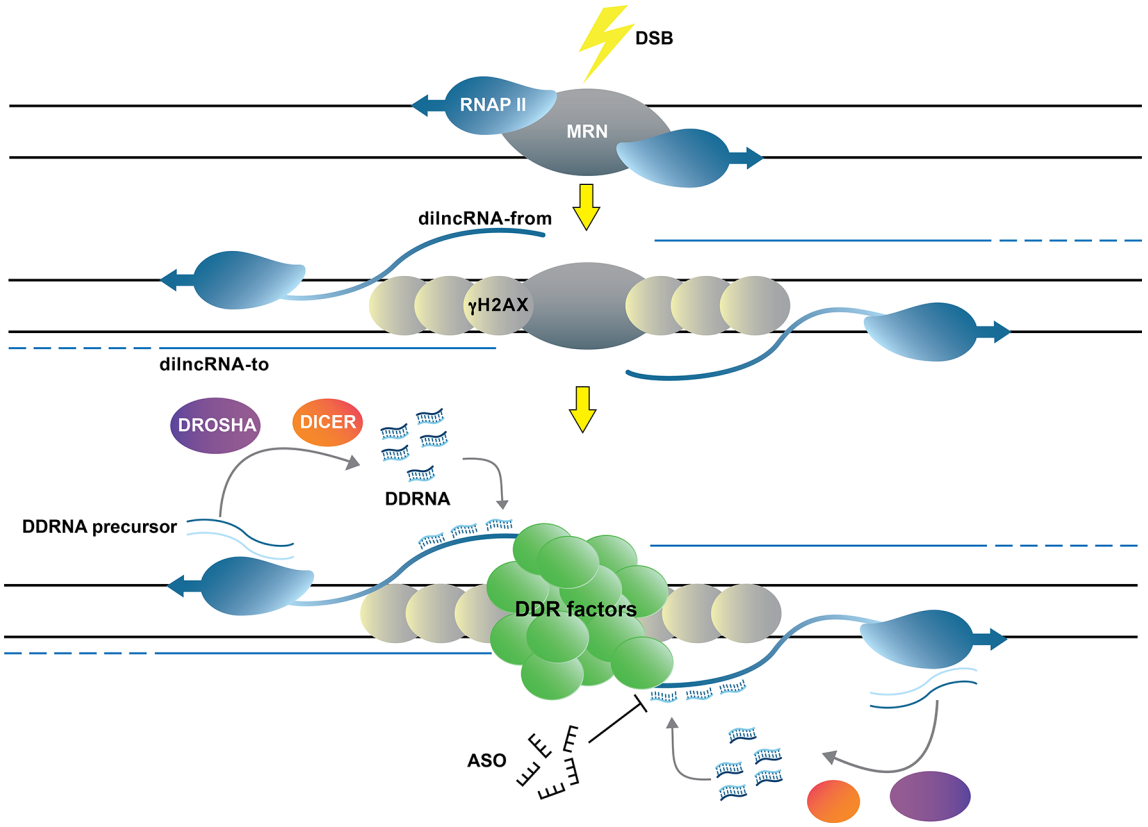

Figure 2. Proposed model for dilncRNA and DDRNA generation and activity at sites of DNA damage. Upon DSB induction, MRN recruits RNAP II to DNA ends triggering the bidirectional synthesis of dilncRNA-from (blue) and, less abundantly, of dilncRNA-to (light blue). DROSHA and DICER process the long double-stranded RNA, likely the outcome of paired or folded dilncRNAs, generating DDRNAs which, in turn, pair with nascent unprocessed single-stranded dilncRNAs. Together, DDRNAs and dilncRNAs bind to DDR factors and fuel DDR focus formation. Interfering with the interaction between dilncRNAs and DDRNAs through antisense oligonucleotides (ASOs) allows site-specific inhibition of DDR. Adapted from Michelini et al., *Nature Cell Biology* 2017.

thereby mediate recombination[167−169] (see section 6 for additional details). Another link that connects RNA to maintenance of genome integrity in *S. cerevisiae* lies in the RNA exonucleases Xrn1, Rrp6, and Trf4. These exonucleases were previously implicated in protecting genome stability from DNA−RNA hybrids and transcription-associated hyper-recombination[170] but have more recently been shown to control also the activation of Mec1/ATR during DSB-induced DDR.[171] While Xrn1 appears to be required for DNA end resection at the initial steps of HR, Rrp6 and Trf4 are dispensable for the resection process itself but are essential for replication protein A (RPA), a ubiquitous single-strand DNA binding protein, loading onto ssDNA.[171] RPA affinity for ssDNA is very high; thus, it is interesting that factors related to RNA processing are important for this interaction. However, differently from factors involved in qiRNA production in *Neurospora*, yeast Xrn1, Rrp6, and Trf4 are not required for completion of later steps of HR repair.[171]

In *D. melanogaster*, it has been shown that although exposed DNA ends of a plasmid are sufficient to induce the generation of sncRNAs,[172] they function to repress transcription of adjacent genes, rather than play a role in the DNA repair process.[173] However, it was recently proposed that splicing factors may stimulate sncRNA generation at a DSB generated by CRISPR-Cas9 downstream of an intron in cultured *Drosophila* cells.[174] Interestingly, the authors suggest that when RNAP II reaches the DNA end, the cotranscriptional spliceosome triggers a signal for the generation of an antisense transcript, potentially also stimulated by the formation of an R-loop, which then pairs with the sense transcript generating the dsRNA long precursor of the sncRNAs. They also hypothesize that a modification of the RNA polymerase complex may enable a strand switch and therefore allow for synthesis of a long RNA hairpin.[174]

Excitingly, different groups have established a direct link between DNA damage and the local generation of sncRNAs in mammalian cells. DDRNAs have the sequence of the damaged locus and are processed by the RNAi (see section 3.1 for details) machinery factors DROSHA and DICER[61] (Figure 2). The key difference between DDRNAs and canonical miRNAs is that DDRNAs can carry virtually any genomic sequence, as they are generated where DNA damage occurs. DDRNAs appear to be required for the full activation of DDR signaling[158] by mediating DDR foci assembly. In brief, DROSHA or DICER knockdown, but not the silencing of downstream RNAi effectors, impairs MDC1, the activated form of ATM and 53BP1 focal accumulation without affecting phosphorylation of H2AX.[158] In agreement with these findings, it has been shown that DICER gets phosphorylated upon DNA damage and translocates to the nucleus where it associates with DSB sites, being necessary for full recruitment of 53BP1 and MDC1.[87] A very recent work also confirmed that DROSHA and DICER, but not the silencing of downstream RNAi effectors, are necessary for 53BP1 focal accumulation and that DROSHA is involved in DNA repair by both HR and NHEJ.[175]

These results suggest that the focal concentration of diffusible DDR proteins can be regulated by sncRNAs. Indeed, the degradation of RNA by transient treatment with RNase A in a

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

permeabilized living cell dissociates 53BP1, MDC1, and pATM from DNA damage sites.[61,158,159] In a system in which a single DSB can be introduced in a traceable locus, RNase A treatment was sufficient to disassemble 53BP1 focus. Strikingly, upon incubation with RNA purified from cells damaged in parallel, but not from parental cells lacking the cleavable site, the 53BP1 focus reassembled. Similarly, incubation with total RNA extracted from cells in which DICER or DROSHA were silenced, or genetically inactivated, did not allow for DDR foci reformation following RNase A treatment.[61,158] Together, these data indicate that DDRNAs contain the sequence of the damaged site and that their generation depends on DROSHA and DICER. Indeed, NGS approaches confirmed the DSB-induced production of sequence-specific DDRNAs, displaying a size consistent with DICER and DROSHA products. When chemically synthesized and reintroduced into RNaseA-treated cells, DDRNAs allowed site-specific DDR focus formation, demonstrating that they can function *in trans* and in the absence of mRNAs.[61,158] Notably, DDRNAs are not required for the direct recognition of the DNA lesion, being instead stimulators of DDR foci assembly on γH2AX-decorated chromatin.[158] It makes sense, then, that the early DDR step of NBS1 association to sites of DNA damage is not sensitive to global RNA degradation,[158] similar to what has been described for the DNA damage sensor KU.[131] It thus seems that the modification of chromatin (γH2AX) and the local synthesis of DDRNA are the two events required to form the large structures known as DDR foci.

Other studies have confirmed the requirement of DICER- and DROSHA-dependent sncRNAs for the recruitment of DDR factors involved in DNA repair, such as RAD51 and BRCA1, together with histone modifier enzymes such as methyltransferase MMSET (WHSC1) and the acetyltransferase Tip60/KAT5.[176] Thus, sequence-specific sncRNAs may act as guiding molecules for the localization to and/or the activation of different utilities, such as for instance chromatin remodelers, at broken DNA ends.[75]

The existence of a class of similar 21 nt-long small RNAs, named DSB-induced RNAs (diRNAs), has been reported in *Arabidopsis thaliana* and in mammalian cells.[94] diRNAs are induced by DSBs in an ATR-dependent manner, are transcribed from the vicinity of the DSBs by plant RNA polymerase IV, and play a role in the RNA-directed DNA methylation (RdDM) pathway.[94] Differently from mammalian DDRNAs, diRNA biogenesis in plants requires not only DICER-like protein but also the activity of an RNA-dependent RNA polymerase, as well as AGO2.[94] In human cells, diRNAs generated from the sequence surrounding the DSB were shown to control recruitment of RAD51 to damaged sites via a direct interaction between the diRNA-AGO2 complex and RAD51, thus promoting HR-mediated DNA repair events.[177] In the proposed model, the diRNA-AGO2 complex anneals either to homologous broken DNA or to chromatin-bound transcripts originating from the target locus, suggesting that a homing mechanism via DNA:RNA or RNA:RNA paring may mediate the activity of diRNA *in trans* and influence DNA repair pathway choice. Another study in *A. thaliana* suggests that diRNAs do not act exclusively in HR-mediated repair but also play a role in NHEJ.[178] Consistent with a direct role played by local transcripts in the process of NHEJ, it has been shown that RNAP II and nascent mRNA associate with factors of classical NHEJ and that RNA can serve as template for error-free DNA repair in mammalian cells[160] (see section 6.3 for details).

Recently, the role of diRNAs and AGO2 in DNA repair has been challenged. By the use of CRISPR-Cas9 and TALEN technologies, it was shown that diRNAs are poorly induced upon DSB induction at endogenous genomic regions and that AGO2 inactivation does not affect HR in *A. thaliana* and in rice.[179] These controversial observations indicate that we are far from fully understanding the biogenesis and functions of DNA damage-associated sncRNAs.

Although they appear to be part of the same phenomenon, DDRNA and diRNA present essential differences both in the process of their biogenesis and in their function. A first difference is that sequencing of diRNAs reveals that they are generated starting from a few hundred bases away from the DNA break.[94] Given their above-mentioned involvement in DNA repair by HR, diRNAs might be produced starting where resection stops and dsDNA is left intact. On the other hand, sequencing showed that DDRNAs map very close to DNA ends.[61] This difference may also suggest that diRNAs are in fact generated after the initial steps of DDR signaling. Another peculiarity of diRNA biogenesis is the dependency on ATR, which primarily responds to the exposure of single-strand DNA. Because these ATR-activating events occur mainly during resection or replicative stress, a model where diRNAs might be generated after or concomitantly to resection is also supported.

The abundance of diRNAs in plants appears to be significantly higher than in mammalian cells, possibly due to the presence of RdRP activity.[94] Interestingly, high levels of pre-existing transcription of a transgene correlate with the generation of abundant diRNAs upon CRISPR/Cas9-induced DSBs.[179] However, these highly abundant diRNAs seem to be dispensable for HR. A possible reconciliation model proposed by the authors is that "primary diRNAs" are low abundant and play a role in DSB repair, while "secondary diRNAs" are more abundant, require active transcription, are not directly involved in DSB repair, but may trigger post transcriptional gene silencing. Indeed, secondary diRNAs may be amplified via a "ping-pong"-like mechanism, where primary diRNAs cleave their complementary long transcripts, which, in turn, are converted by RdRPs into double-stranded RNAs and processed by DICER-like proteins to generate a new pool of diRNAs.

A similar ping-pong mechanism, by which small RNAs suppress neighboring gene expression, has been proposed in *D. melanogaster*. According to these findings, endogenous small interfering RNAs (endo-siRNAs) are produced from a transfected linearized plasmid, mimicking DNA ends of a genomic DSB.[172,173] These data are in line with a role for break-derived sncRNAs in RNA quality control rather than DNA repair.

Whether these small RNAs originate from processing of pre-existing transcripts or from *de novo* transcription at sites of break in mammalian cells has been recently addressed and is discussed in depth in the next section.

4.2.2. Damage-Induced Long ncRNA (dilncRNA).

Prompted by the discovery of DDRNAs (see section 4.2.1 for details), the group of d'Adda di Fagagna (IFOM, Italy) in collaboration with the group of Nils Walter (University of Michigan, USA) has more recently probed the transcriptional landscape around a DSB in search of DDRNA precursors. By single-molecule fluorescent in situ hybridization (smFISH, see section 9.1 for details) and reverse transcription followed by quantitative PCR (RT-qPCR) techniques, a novel class of lncRNAs named damage-induced lncRNAs (dilncRNAs) transcribed by RNAP II upon damage from and toward the DNA ends was uncovered.[95,180] Induction of *de novo* transcription

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

from DSBs was demonstrated in various mammalian cellular systems. For these experiments, multiple endonucleases were used to generate DSBs at exogenous integrated constructs as well as endogenous genomic loci, in both transcribed and non-transcribed regions. dilncRNAs were generated regardless of transcription state, suggesting independence from pre-existing transcription or canonical promoter and enhancer elements. In a similar study by the same team, dilncRNA induction was also shown at dysfunctional telomeres[181] (see section 8.3 for details). In light of these observations, it is tempting to speculate that DSBs can themselves act as promoters; additional experiments, however, are needed to understand if the machinery needed for canonical RNAP II transcription, such as the preinitiation complex,[182] is required also for dilncRNA generation.

The apical DNA damage sensor MRN (see section 4.1.1 for details) has been shown to be required for RNAP II localization to the damaged site and for subsequent dilncRNA transcription.[95] Indeed, RNAP II immunoprecipitates with all three components of the MRN complex upon irradiation, although additional studies are needed to understand if the interaction is direct and through which domains and possibly modifications it is mediated. The role of MRN in the production of DDRNA/dilncRNA has been studied by knockdown experiments and by treatment with the small molecule mirin.[95] Inhibition of MRN activity by mirin reduces DDR focus reformation when DDRNAs were exogenously added to RNaseA-treated cells.[61] This could be because treatment with mirin inhibits dilncRNA synthesis,[95] thus reducing the localization of DDRNAs to the site of damage. RNAP II transcription is known to be stimulated by nicks or, more strongly, by a DSB with a 3′-overhang.[183,184] Since mirin inhibits both endo- and exonuclease activities of MRN,[185] it is tempting to speculate that either, or both, activities are required for RNAP II transcription from the DNA ends. Another possibility is that the reported ability of MRN to unwind DNA ends[186] is the step necessary to initiate RNAP II transcription from the DNA ends.

As discussed above (see section 4.1.1 for details), KU can bind RNA and RNAP II. Given the newly discovered role of MRN in damage-induced transcription, as well as the possibility of KU and MRN coexisting on the same DNA end[131,133] and of KU to stimulate MRN activity,[134,135] it will be interesting to determine the contribution, if any, of KU and its crosstalk with MRN in the production of ncRNAs at the site of DNA breaks.

In the proposed model (Figure 2), dilncRNAs divergent from and convergent to the DNA ends have the potential to pair and form a double-stranded RNA, which is processed by DROSHA, and then DICER, to generate DDRNAs. Accordingly, Michelini et al. demonstrated by qRT-PCR analyses that dilncRNAs accumulate in DROSHA-depleted damaged cells, while the products of DROSHA processing, called pre-DDRNAs, accumulate in the absence of DICER.[95] In the same samples, DDRNAs are induced upon DNA damage and decrease when DROSHA or DICER are silenced. The latter result is also consistent with the characterization of telomeric DDRNAs[181] (see section 8.3 for details).

The relevance of these ncRNA species in DDR signaling and in DNA repair comes from experiments preventing their transcription or their function. Indeed, a transient inhibition of RNAP II by small molecules, such as alpha-amanitin, prevents global DDR activation downstream of γH2AX and inhibits DNA repair.[95] Excitingly, antisense oligonucleotides (ASOs) (see section 10 for details) against dilncRNAs and DDRNAs (Figure 2) are able to reduce 53BP1 accumulation and DNA repair at individual genomic loci with an unprecedented degree of specificity.[95]

By intracellular single molecule high resolution localization and counting (iSHiRLoC, see section 9.4 for details), fluorescently labeled DDRNAs localize to the damaged site through base-pairing with unprocessed dilncRNAs emerging from the DSB, and this interaction is fundamental to fully activate the DDR.[95]

These events are not unprecedented. Indeed in *S. pombe*, small RNAs generated by DICER bind to a nascent transcript, which is also their precursor, and together maintain the epigenetic and genetic stability of the centromeric locus.[166] According to this so-called "nascent transcript" model, the unstable 2kb-long nascent transcript, synthesized by RNAP II preferably from one strand of the centromeric region,[187,188] is converted to dsRNA by RdRPs or by pairing to an antisense transcript. This double-stranded RNA is then processed either by an RNAi pathway-dependent mechanism involving DICER or by RNAi pathway-independent mechanisms, such as the RNA degradation pathway of the Trf4/Air2/Mtr4 (TRAMP), and the exosome complexes.[189] The resulting siRNA is loaded into the RITS (RNA-induced transcriptional silencing) complex, where the release of one of the two strands takes place generating mature Ago1-bound single-stranded siRNA.[45] The base pairings between the mature small RNAs component of the RITS complex and the nascent transcripts, but not the underlying DNA, are central for the recruitment of enzymes responsible for H3K9 methylation, a repressive histone mark, of the centromeric locus, and enforcement of transcriptional gene silencing.[190] Moreover, Ago1 slicing activity may contribute to the production of additional siRNAs and to the exhaustion of pericentromeric RNA in a self-sustaining loop. In this scenario, chromatin-associated nascent transcripts are not just the precursors of siRNAs, but they also act as local platforms for the coordinated assembly of chromatin remodelers guided by the siRNAs to the complementary target regions. A similar "nascent transcript" mechanism may also take place during the formation of a DDR focus. Michelini et al. showed that 53BP1 associates with *in situ* generated DDRNAs and dilncRNAs in a manner dependent on its Tudor domain.[95] The localization of DDRNAs to the damaged site through base-pairing with nascent dilncRNAs may represent one of the mechanisms by which 53BP1 is selectively recruited to DSBs, and potentially a common mechanism for the recruitment of other DDR proteins. However, it will be important to investigate whether this interaction is direct, as well as to assess the possible role of Argonaute proteins in the DDR version of the "nascent transcript" model.

Several reports have shown that DSBs within a transcriptional unit suppress canonical gene expression,[74,191] thus avoiding the transcription of a damaged template. This appears to be in contrast with the observed *de novo* transcription at DSBs. Once again, the literature on the *S. pombe* centromeric locus comes in handy. The apparent paradox of yeast cotranscriptional gene silencing, requiring a certain level of transcription to shut down transcription of specific genomic loci, has been recently solved. Indeed, the RNAi-mediated local concentration of chromatin remodelers and their residence time on the target sequence need to be above a certain threshold in order to switch off transcription and maintain the epigenetic marks.[190,192] It is therefore possible that damage-induced ncRNA transcription is a similarly tightly regulated mechanism shaping the chromatin surrounding a DSB to induce the suppression of pre-existing gene expression. In *S. pombe*, it has been demonstrated that *de novo* transcription is

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

induced at sites of DNA damage and that these newly synthesized RNA molecules anneal with their DNA templates resulting in transient DNA−RNA hybrids required for efficient DSB repair via HR.[193] Notably, two very recent studies, exploiting the same endonuclease-based model system, acknowledged the presence of DNA−RNA hybrids at DBSs also in mammalian cells.[175,194] In particular, it has been proposed that DROSHA is involved in the accumulation of DNA−RNA hybrids at DSBs[175] and that Senataxin, a well characterized R-loop helicase, is recruited at DSBs induced in transcribed genomic regions, where it removes such DNA−RNA hybrids, promotes RAD51 loading, and prevents translocations.[194] Future studies on this topic will have to take into account the existence of additional layers of complexity such as the kinetics of the events starting from DNA damage induction in a given cell type and cell-cycle phase. Indeed, the demand for novel approaches taking into account any heterogeneity in the cell population, such as single-cell resolution techniques (see section 9 for details), has become necessary among scientists that deal with quantitative analysis of siRNA-mediated epigenetic silencing.[190]

The discovery that each DDR focus relies not only on a common set of shared proteins but also on a set of RNA molecules generated *in situ*, that individually mark DDR events at distinct genomic loci, represents a leap forward in the understanding of the DDR pathways that may, in the future, be exploited for therapeutic purposes.

5. RNA SPLICING AND DNA DAMAGE

5.1. Splicing, a Brief Overview

Splicing is a complex mechanism by which noncoding intronic sequences are precisely removed from the primary gene transcript (pre-mRNA) to generate a mature mRNA molecule, and its regulation is key in all aspects of cell physiology and pathology. Intron removal is carried out by a large molecular machine, the spliceosome, which is assembled on the pre-mRNA in a stepwise manner and is composed of five small nuclear ribonucleoparticles, named snRNPs U1, U2, U4, U5, and U6, and a large number of proteins.[199] The spliceosome recognizes short sequence elements with a loose consensus at exon−intron boundaries (5′ and 3′ splice sites) as well as the branch point located near the 3′splice site.[200]

The weak and dynamic interactions between the spliceosome and the pre-mRNA can be modulated by RNA binding proteins (RBPs), which associate to splicing regulatory sequence elements. These elements are particularly relevant for the selection of splice sites that deviate from the consensus sequences (weak sites) and either stimulate (intronic and exonic enhancers) or repress (intronic and exonic silencers) their recognition, thus affecting the splicing outcome. The list of RBPs involved in this regulation is continuously expanding and includes, but is not limited to, the serine/arginine (SR) family of splicing factors and a group of proteins that bind to heterogeneous nuclear RNA (hnRNP proteins). The partial degeneration of splice site sequences and the possibility to modulate their recognition through protein complexes assembled on enhancers and silencers allow for numerous events of alternative splicing to occur for each transcript. By using various combinations of 5′ and 3′ splice sites, and the respective regulatory proteins that bind them, alternative splicing (AS) is capable of generating different mRNAs from a single pre-mRNA.[201] The vast majority (>90%) of human genes display AS events,[202] which are modulated not only during development in a cell-type dependent manner but also in response to a wide range of stimuli or stressing conditions, including DNA damage.[203−206]

Splicing decisions may be modulated by chromatin organization, in particular nucleosome positioning[207] and histone modifications,[208−210] and by the elongation rate of RNAP II.[201,211] The influence of this latter factor on splicing decisions

Box 1. RNA Modifications upon DNA Damage

The epitranscriptome is the collection of the chemical modifications of RNA. More than 100−150 different RNA modifications have been reported,[195,196] suggesting that a novel code awaits to be fully deciphered. This array of RNA base chemical alterations can, in principle, have an impact on several aspects of RNA biology by affecting RNA secondary structure, processing, stability, and interactions with other RNAs and RNA binding proteins. Apart from the known abundant modification of 5′-ends of mRNAs (e.g., 5′ 7-methylguanosine cap), the most commonly studied modifications are N^6-methyladenosine (m^6A), N^1-methyladenosine (m^1A), 5-methylcytosine (m^5C), inosine (I), and pseudouridine (Ψ), with the latter being also termed the fifth base of RNA because of its large quantity in the cell. However, so far the detection of modifications has been restricted to abundant RNAs such as tRNA and rRNAs. The generation of novel tools, such as antibodies specific against some modifications, allowed the extension of their study to less abundant RNAs, including noncoding RNAs. This, combined with novel chemical treatments that react specifically with RNA modifications and make them detectable through high-throughput sequencing methods, has led to a boom of "epitranscriptomics" studies. The potential to directly detect modified RNA bases through new sequencing technologies that avoid reverse transcription and cDNA amplification promises to further expand these opportunities.[197]

Recently, RNA modifications have been implicated in DNA damage repair.[198] Xiang and colleagues reported that an antibody against adenosine methylated at position 6 (m^6A) stains sites of laser-induced ultraviolet (UV) microirradiation. The signal is quickly but transiently (10 min) induced. This rapid kinetic is consistent with its dependency on PARP-1 activity, a response to single- and double-strand DNA breaks which is also very prompt and transient.

m^6A signals depend on methyltransferase-like 3 (METTL3), and its cofactor METTL14, and are removed by the fat mass and obesity-associated protein FTO, a demethylase, evidenced by inactivation of METTL3 and METTL14 preventing m^6A signals detection, while their intensity and persistence increase upon knock down of FTO. All three of these factors have been observed to accumulate at laser-induced DNA damage stripes.[198]

m^6A seems necessary for efficient repair of cyclobutane pyrimidine dimers (CPDs), a common UV-induced DNA lesion, since METTL3 inactivation delays CPD removal. Since the recruitment of DNA polymerase k (Pol k), a translesion DNA polymerase involved in UV-induced DNA damage, is impaired upon METTL3 and METTL14 inactivation, the authors suggested that m^6A is involved in its recruitment.[198] However, no m^6A reader or direct binding of Pol k to m^6A could be observed. Thus, the molecular mechanisms by which m^6A engages Pol k remain unknown.

Since m^6A is among the most abundant modifications, it is possible that additional, albeit less common, ones await to be discovered and characterized.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

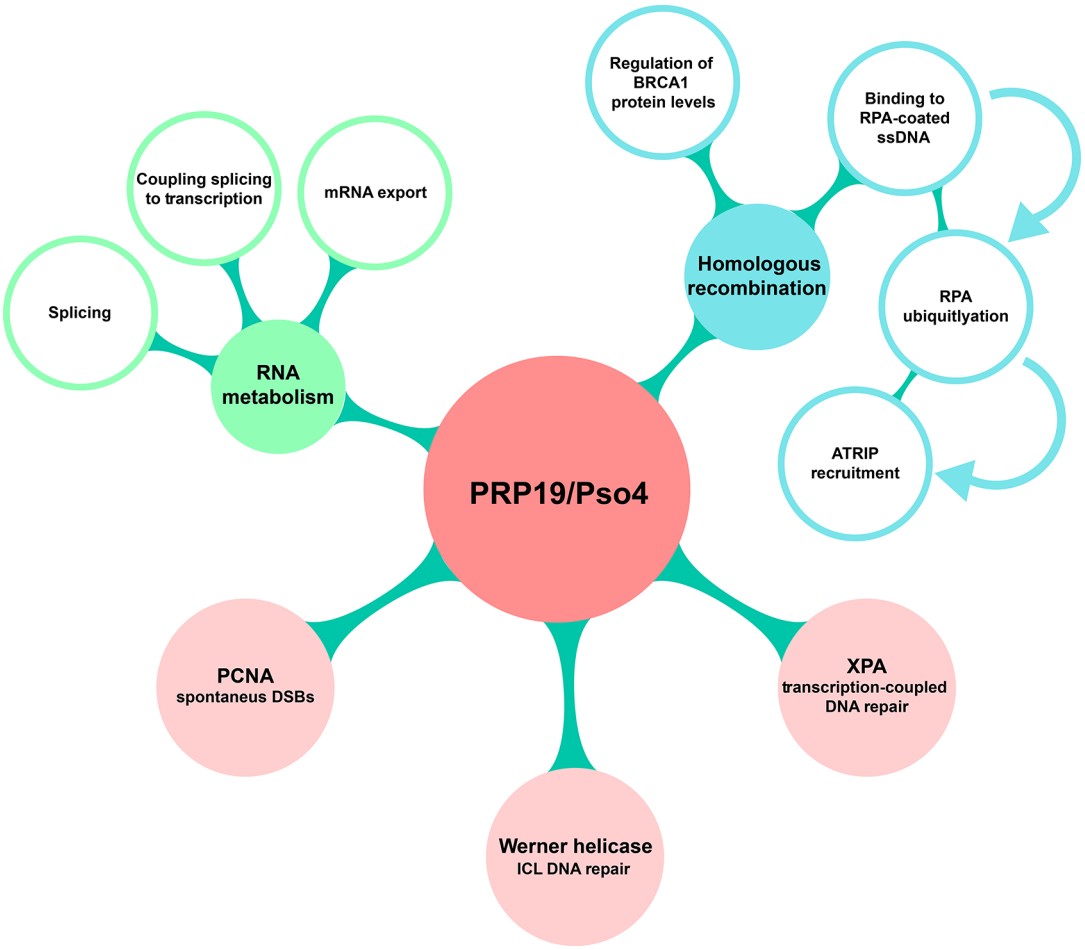

Figure 3. PRP19/Pso4 as an example of a splicing factor involved in the DNA damage response and repair. PRP19/Pso4 is a ubiquitin ligase involved in RNA splicing and mRNA export. It is also implicated in DNA repair: it colocalizes with the replication clamp PCNA, and its downregulation increases spontaneous DSBs; it interacts with Werner DNA helicase participating to the interstrand cross-link (ICL) repair; it associates to Xeroderma pigmentosum group A (XPA) protein playing a role in the transcription-coupled DNA repair; it participates to homologous recombination by regulating BRCA1 protein levels, binding to RPA-coated single-strand DNA, stimulating RPA ubiquitylation, and consequently ATRIP recruitment to stalled replication forks.

stems from the fact that the assembly of the spliceosome occurs cotranscriptionally;[212] thus, the elongation rate of RNAP II determines the time window available for a weak upstream splice site to interact with splicing factors before a competing stronger downstream splice site is transcribed.[213] A key player in coordinating transcription with splicing is the CTD (C-terminal domain) of RNAP II that acts as landing pad for numerous splicing factors.[214] The recruitment coupling model suggests that the phosphorylation status of the CTD, which is controlled by numerous factors, including DNA damage, determines the set of RBPs recruited to the transcriptional apparatus.[215]

By integrating different levels of regulatory events (chromatin organization, abundance of RBPs, post-translational modifications of RBPs and RNAP II), alternative splicing represents an ideal mechanism to finely tune gene expression in response to cell growth or stressing conditions, including DNA lesions.

5.2. Reciprocal Interaction between DNA Damage and Splicing

It is now becoming apparent that a reciprocal interaction exists between DNA damage generation and the regulation of alternative splicing. The impact of DNA damage on splicing profiles has been addressed in detail in several excellent reviews.[203−206,216] We will briefly discuss here only a few recent examples to illustrate how complex the interplay between DNA damage and regulation of AS can be.

A novel link between AS and the DDR has recently been discovered: detained introns (DIs), a new class of introns that exhibit delayed splicing.[217] DI-containing transcripts are usually retained in the cell nucleus and form a reservoir of ready-to-use molecules, that, for example, can be called upon under conditions of impaired transcription. Notably, a subset of DIs, waiting in the nucleus for a signal, is spliced in response to DNA damage. Following DNA damage, a coordinated expression of specific splicing variants occurs, among which it is worth mentioning MDM4 and BCLAF1 that, respectively, control p53 and BRCA1 functions. The impact that DNA damage has on the splicing profile of BCLAF1 transcripts has important biological consequences. Indeed, in response to doxorubicin the pool of nuclear BCLAF1 transcripts containing DIs is halved while the level of protein-coding mRNA is up-regulated.[217] Moreover, BCLAF1 protein is excluded, along with the splicing factor THRAP3 and RNAP II, from DNA damage sites in a process that depends on ATM activity.[218] As a consequence, BCLAF1 protein is available to form a complex with BRCA1 phosphorylated by ATM. This complex recruits a number of splicing proteins, including Prp8, U2AF65, U2AF35, and SF3B, to a set of genes

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

involved in DNA damage signaling and repair, thus connecting DDR signaling activation with cotranscriptional splicing and mRNA stability.[219]

Since DNA damage can control the splicing profile of genes involved in the DDR, it would be expected that splicing inhibition may play a role in the response to DNA damage. This hypothesis has been recently verified by showing that a short period of splicing inhibition prior to irradiation impairs IR-induced DNA damage foci formation.[220,221] Furthermore, two natural compounds that affect the assembly of the spliceosome, namely the macrolide pladienolide B, which targets the splicing factor 3B subunit 1 (SF3B1) of the U2 snRNP,[222] and the biflavonoid isoginkgetin, which prevents the recruitment of the U4/U6.U5 tri-snRNP,[223] reduce ubiquitylation of damaged chromatin which is required for the assembly of DNA repair complexes. In particular, splicing inhibition impairs the recruitment to damaged sites of WRAP53β, RNF168, 53BP1, BRCA1, and RAD51, without affecting γH2AX and MDC1 signals, which are known to be recruited in a ubiquitin independent manner. This effect is due to the reduced expression of the short-lived E3 ubiquitin ligase RNF8,[220] rather than a direct effect on DNA damage foci stability. Interestingly, the decreased expression of RNF8 partially explains the defective DNA repair observed after depletion of various splicing factors, thereby demonstrating the importance of splicing factors to genome stability.[224]

One of the best characterized examples of a splicing factor with a role in the DDR is the ubiquitin ligase PRP19, also known as Pso4 for Psoralen 4 gene, which is part of a large multiprotein complex comprising six additional subunits.[225] PRP19/Pso4 acts at several levels of RNA metabolism (Figure 3): it modifies PRP3, a component of U4 snRNP, with a nonproteolytic ubiquitin chain that enhances protein−protein interactions and stabilizes the U4/U6.U5 complex;[226] it interacts with RNAP II and recruits the TREX complex, which is involved in mRNA export, to transcribed genes;[227] it forms a complex with U2AF65, which participates in the CTD-dependent coupling of splicing to transcription.[228]

A large body of data implicates PRP19/Pso4 in the DDR (Figure 3). The PRP19/Pso4 was initially identified as an essential DNA repair factor in *S. cerevisiae*[229] and it is one of the numerous human RBPs implicated in DNA repair.[230] Accordingly, its down-regulation increases the sensitivity of human cells to spontaneous DSBs as well as to hydroxyurea or PARP inhibitor treatments. This may be related to the fact that PRP19/Pso4 colocalizes with the replication clamp PCNA both during unperturbed cell-cycle and in response to replication stress inducers such as hydroxyurea or camptothecin.[231,232] In addition, PRP19/Pso4 participates in the interstand cross-link DNA repair pathway by interacting with Werner DNA helicase[233] and in the transcription-coupled DNA repair pathway through association with Xeroderma pigmentosum group A (XPA) protein.[234] PRP19/Pso4 also plays a role in the homologous recombination (HR) pathway by regulating the protein levels of BRCA1 and the generation of single-stranded DNA at DSBs.[231] This latter function most likely involves the ability of PRP19/Pso4 to bind RPA-coated single-stranded DNA.[232,235] Binding to RPA is required for PRP19/Pso4 localization to sites of DNA damage and for the ensuing RPA ubiquitylation, which facilitates the recruitment of ATRIP and the recovery of stalled replication forks. This mechanism shows strong similarities to what was previously described for DSB repair where ubiquitylation is required for γH2AX to act as a platform for the assembly of DDR complexes.[236]

PRP19/Pso4 may be one of the better characterized splicing factors that also plays a role in the DDR, but it is surely not the only one. Indeed, several proteins directly or indirectly involved in splicing associate with sites of DNA damage in a PAR-dependent manner, as mentioned above (see section 4.1.1 for details). In spite of this common feature, these RBPs participate in different DNA repair pathways. Thus, for instance SFPQ and NONO are two multifunctional DNA- and RNA-binding proteins involved in the catalytic step of the splicing reaction, in nuclear retention of defective RNAs and in DNA repair, stimulating NHEJ and repressing HR.[141] Another RBP, RBMX/hnRNPG, implicated in tissue-specific regulation of gene transcription and alternative splicing, is a positive regulator of HR.[224] However, the involvement of RBMX/hnRNPG in HR does not depend on its recruitment to sites of DNA damage but instead on its ability to control BRCA2 expression. Remarkably, some RBPs influence different steps of the assembly of repair foci. An example is FUS (see section 4.1.1 for details), whose depletion impairs the formation of DNA repair foci after treatment with topoisomerase II poison etoposide.[154] Depletion of RBM14, another RBP, stabilizes γH2AX foci[237] by reducing the recruitment of the NHEJ factors XRCC4 and XLF to damaged chromatin.[238] The transient recruitment of the RBP hnRNPUL1 to DNA damage sites requires both the MRN complex[239] and PARP1[240] and is necessary for the full activation of the ATR signaling pathway. Moreover, hnRNPUL1 stimulates DSB resection and HR by promoting the association of the BLM helicase to DNA breaks.[239] Transient association of RBPs with damaged areas appears to be a common theme. Indeed RBPs, including THRAP3, BCLAF1,[218] hnRNPC, and hnRNPK,[224] exhibit a prolonged exclusion from irradiated areas. Redistribution of these proteins requires both active transcription and the activity of PIKKs.[155] Interestingly, inhibition of PIKK prevents displacement of RBPs from sites of damage and favors the formation of DNA−RNA hybrids, suggesting that the displacement is part of a general mechanism to prevent unwanted DNA−RNA hybrids.[155]

Overall, these examples reveal the existence of tight connections between splicing regulation, the assembly of DNA repair complexes, and the activation of checkpoint pathways. RBPs appear to have a central role in the coordination of all these events. However, the underlying molecular mechanisms are still a matter of investigation, and in particular it is unclear whether or not RNA molecules are involved in these dynamic processes.

6. RNA-TEMPLATED DNA REPAIR IN YEAST AND MAMMALS

RNA molecules synthesized during transcription are complementary to the DNA strand that served as their template. Early work demonstrated that RNA could play an indirect role in genome modification and DSB repair if converted into a DNA copy (cDNA) and stitched into damaged sites via NHEJ in yeast and mammalian cells.[241−244] Not only can these cDNA molecules be inserted in a nonhomologous manner at sites of DSBs, but cDNA can also function as a homologous donor template to accurately repair DSBs via homologous recombination (HR) in budding yeast.[169] However, can an RNA molecule serve directly as a template for repairing/modifying DNA without the need of being converted into cDNA?[245,246] Indeed, RNA-containing DNA oligonucleotides can serve as templates for gene editing on plasmid or chromosomal DNA in *Escherichia coli*.[247−249] Similarly, RNA-containing and RNA-only oligonucleotides can serve as RNA donor templates for DSB repair, a

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

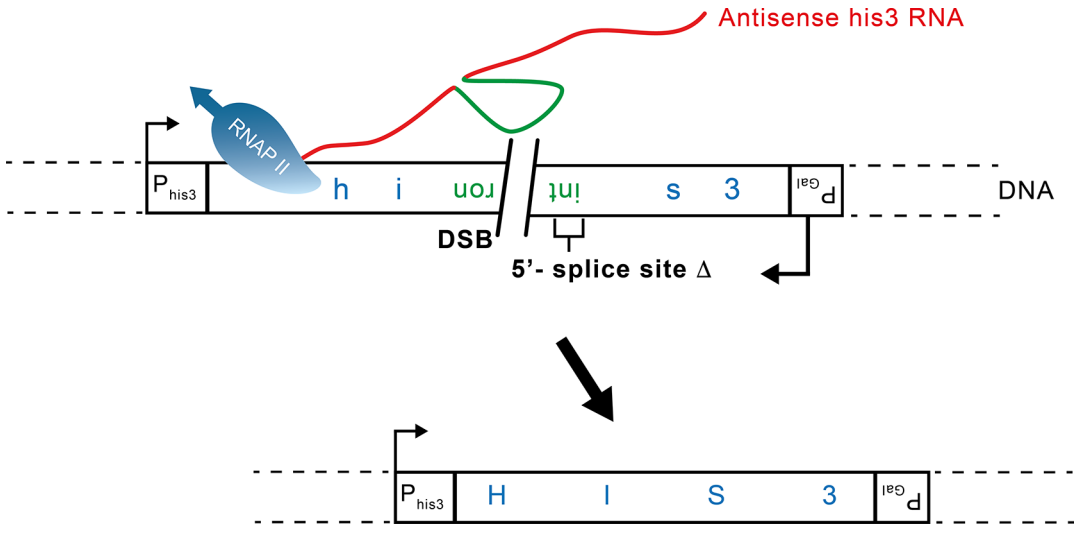

Figure 4. Diagram of the system to detect RNA-templated DSB repair. The *his3* gene is transcribed in the antisense orientation under an inducible promoter and contains an artificial intron that can only be spliced out from the antisense *his3* transcript. Splicing of the antisense RNA and DSB repair by the spliced RNA results in removal of the intron and restoration of a functional *HIS3* gene, which generates His^+ cells. Deletion of the 5′-splice site within the intron sequence is indicated.

phenomenon observed in yeast and human cells.[168,248] In addition, artificial long RNA templates injected in ciliate cells can guide genomic rearrangements[250] (see section 7 for details). RNA-templated DNA modifications have been proposed to explain the high-frequency non-Mendelian loss of heterozygosity in rice.[251] Moreover, *cis*- and *trans*-splicing mechanisms of chromosomal translocation suggest that chimeric RNAs generated by intergenic splicing may play a direct role to guide chromosomal rearrangements.[252−256] A proof of concept that RNA transcripts are recombinogenic and can directly alter the genetic information in chromosomal DNA derives from experiments performed in budding yeast.[169] Given these observations, the importance of RNA-templated repair becomes apparent.

6.1. Molecular Mechanisms

Keskin et al. demonstrated that in *S. cerevisiae* an endogenous transcript can serve as template for repair of a chromosomal DSB *in cis*.[169] The genetic assay was based on the antisense RNA-dependent repair of a nonfunctional histidine auxotrophic marker gene (*his3*). Briefly, an artificial intron (AI) is inserted in reverse orientation relative to *his3*, and antisense transcription is induced (Figure 4). While the AI cannot be spliced out of the sense *his3* transcript, it can be spliced out of antisense transcript. Following the generation of a DSB inside the AI, the pre-existing *his3* antisense transcripts is used as a template for HR, resulting in a functional *HIS3* gene lacking the intronic sequence (Figure 4). While accurate DSB repair of *his3* is seen in wild-type yeast cells by the formation of histidine prototrophic (His^+) colonies, it is dependent on the reverse transcriptase (RT) activity of yeast retrotransposons, indicating that repair in wild-type cells proceeds through a cDNA intermediate. However, the inability to detect direct RNA-templated DSB repair in wild-type yeast cells may be due to a limitation of the assay used. Indeed, direct RNA-templated DSB repair in wild-type yeast cells is blocked by the function of ribonucleases H (RNase H1 and H2) that cleave the RNA strand of DNA−RNA hybrids. Once the activity of RNase H enzymes is removed, DSB repair is detectable even in the absence of the reverse transcriptase.[169] These results demonstrate the existence of direct RNA-templated DSB repair.

Support for a direct RNA-templated DSB repair mechanism mediated by transcript RNA *in cis* is provided by the dependence on splicing of the antisense RNA. In fact, removal of the 5′- splice site (Figure 4) eliminates the formation of His^+ colonies. Furthermore, sequencing data and Southern blot analysis support the accurate repair by *cis*-acting RNA, rather than ectopic integration of cDNA transcript from other regions of the yeast genome.[169] Interestingly, even in the absence of the DSB, His^+ colonies are still detectable.[169] This finding suggests that the antisense RNA transcript can even modify DNA without induction of damage, possibly through spontaneous DSBs or nicks in the DNA. Overall, these results demonstrate that RNA can directly transfer genetic information to chromosomal DNA *in cis* with or without the induction of a DSB, revealing the existence of a mechanism in which genetic information can flow back from RNA to DNA, beyond the special case of reverse transcription postulated by the "central dogma" of molecular biology.[243]

6.2. How Does RNA-Templated DSB Repair Work?

Since RNA functions *in cis* as a donor template in DSB repair of *his3* in the assay described above, the mechanism of DSB repair by RNA is HR. Instead, the sensitivity to RNase H activity indicates that DNA−RNA hybrids must form to transfer information from RNA to DNA. Previous work showed that the RecA recombinase of *Escherichia coli* can promote formation of DNA−RNA hybrids.[257,258] Yeast RNA-templated DSB repair is strongly dependent on the recombinase Rad52, a fundamental protein in DNA repair by HR.[169,259] However, knockout of the *RAD52* gene, while reducing the frequency of DSB repair by RNA by a factor of 10, does not eliminate DSB repair by RNA, indicating that Rad52-independent RNA-templated DSB repair mechanisms do exist. These results in yeast are supported by *in vitro* experiments corroborating the ability of the Rad52 protein to catalyze the annealing of RNA to DNA.[169] Recently, it was shown that purified yeast or human Rad52 protein can catalyze an inverse strand-exchange reaction with DNA or with RNA *in vitro*, a property not observed using the RecA homologue Rad51 recombinase or yeast Rad59, which is important for strand annealing.[259] While RPA inhibits inverse strand exchange

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

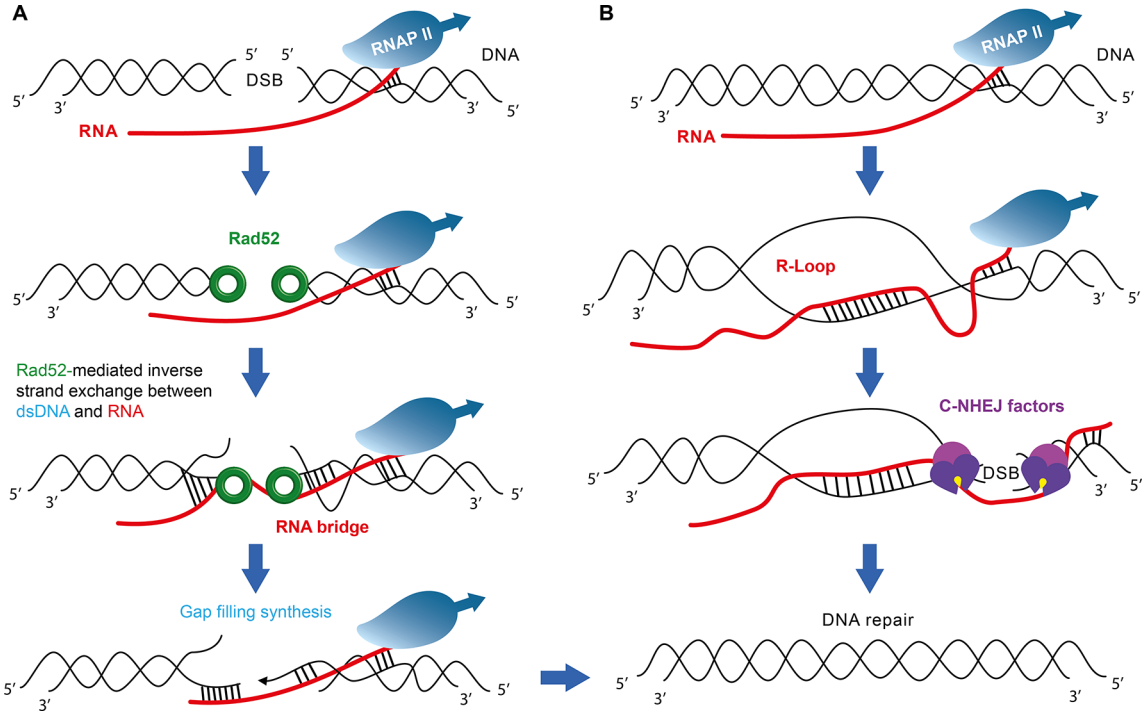

Figure 5. Models of DSB repair guided by RNA. (A) Model of RNA-templated DSB repair via Rad52-mediated inverse RNA strand exchange. The RNA transcript generated from a genomic region that experiences a DNA DSB can anneal with broken DNA ends with the aid of Rad52 in an inverse strand-exchange reaction forming an DNA−RNA hybrid that bridges the broken DNA ends and enables transfer of genetic information from RNA to DNA and accurate repair of the DSB. (B) Hypothetical model of RNA-mediated NHEJ repair of DSB. The RNA transcript forms an R-loop before the DSB occurs. Via a bridging mechanism, likely without the need of a gap-filling step, RNA guides the C-NHEJ protein complex to perform accurate DSB repair.

between two DNA molecules, it stimulates Rad52-mediated inverse strand exchange between DNA and RNA, possibly via protein−protein interaction with Rad52.[259] Rad52 also promotes inverse RNA strand exchange with short-tailed or even blunt-ended double-stranded DNA. These results parallel *in vivo* studies demonstrating that RNA-templated DSB repair is stimulated by the overexpression of either yeast or human Rad52 N-terminal domain (NTD).[259] Rad52 NTD retains the catalytic ability to promote inverse RNA strand exchange but lacks the Rad51 and RPA binding domains.[259] Furthermore, null mutations of the *RAD51* or *RAD59* genes increased the frequency of DSB repair by RNA in yeast.[169,259] This outcome is thought to occur by curbing the ability of DNA ends to recombine with sister chromatids, funneling repair to an RNA-templated pathway.[169] Moreover, impairment of DNA end processing by defects in *SAE2*, *EXO1*, or *MRE11* genes, which are important for DNA end resection following DSB, either increased or had no change in the frequency of DSB repair by RNA.[259] These data support a model in which Rad52 catalyzes inverse strand exchange between RNA and a nonresected, or little-resected end of DNA at the DSB. RNA then guides break repair by bridging the broken DNA ends and is used as a template for DNA synthesis to fill the gap, a mechanism that could be mediated by cellular DNA polymerases.[169,259] If resection is long, RNA-templated DNA repair may require reverse transcriptase for more extensive polymerization.

6.3. Involvement of NHEJ Mechanisms

Recently, it has been found that C-NHEJ may play a role in RNA-mediated DSB repair. Following DSBs introduction via bleomycin or ionizing radiation (IR), RNAP II immunoprecipitated with various C-NHEJ and recombination proteins, including LigIV, XRCC4, KU-70, Polμ, DNA-PK, Rad51, and Rad52.[160] Differently, alt-NHEJ proteins were absent or far less abundant in RNAP II complexes.[160] The authors reasoned that C-NHEJ proteins may have a role in DSB repair in actively transcribed genes and explored this further. However, in this study, little information is provided on the roles of recombination proteins, which have previously been documented to function at DSBs in active genes.[260] ChIP and quantitative PCR do indeed support the presence of C-NHEJ proteins (53BP1 and LigIV) at sites of DSBs in actively transcribed genes.[160] Importantly, C-NHEJ components were found associated with nascent RNA transcripts by RNA-ChIP and this association significantly decreased following treatment of permeabilized cells with RNase H prior to RNA-ChIP,[160] indicating the formation of DNA−RNA hybrids at DSB sites. This led the authors to suggest that C-NHEJ proteins may aid in an RNA-templated DNA repair mechanism. While RNA-donor oligonucleotides could repair a DSB in human cells *in trans*[248] and an actively transcribed DNA could increase the frequency of end joining ligation of a linearized plasmid in human cells either directly or via RNA sequences *in trans*,[160] it would be important to determine whether nascent pre-mRNA can template DSB repair *in cis* in mammalian cells. Following up on this possibility, in search of RNA-templated DNA polymerase activity, nuclear extracts of HEK-293 cells strikingly had the capability to copy an RNA template *in vitro*, independently of the major mammalian retrotransposon long interspersed elements (LINE1).[160] This result highlights the possibility that cellular DNA polymerases may have some RT activity, as shown for yeast replicative polymerases,[168] bacterial and archaeal polymerases,[261] and some mammalian polymerases.[262]

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

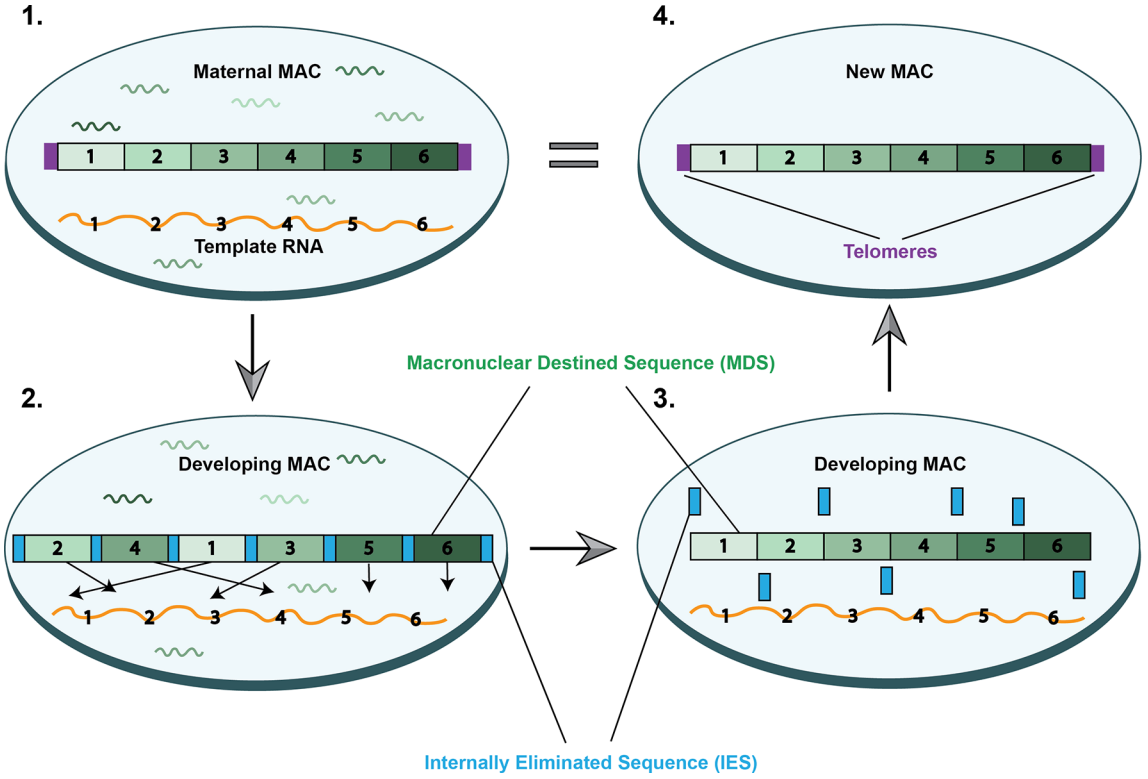

Figure 6. RNA-mediated genome unscrambling in *Oxytricha*. (1) Macronuclear nanochromosomes are transcribed bidirectionally in the maternal macronucleus, giving rise to template guide RNAs. Twenty-seven nt piRNAs are depicted as wavy green lines and play a protective role during DNA elimination in this process. (2) Template guide RNAs are transported to the developing macronucleus where they mediate the rearrangement of macronuclear destined sequences (MDSs). (3) Internally eliminated sequences (IESs) are removed from developing macronuclear nanochromosomes. (4) Development of the new macronucleus is complete after *de novo* telomere addition and several rounds of nanochromosome amplification.

6.4. Models of DSB Repair Mediated by RNA

Overall, these studies unveil an unexpected direct role of RNA in the DSB repair process: RNA may act as a template in repair of DSBs occurring in transcribed DNA.[160,169,243,259,263,264] An HR model based on the results of experiments in *S. cerevisiae* suggests that a DSB occurring in an actively transcribed gene can be repaired *in cis* by the transcribed RNA as a bridging template for DNA repair. This process is aided by the inverse strand-exchange activity of Rad52 on dsDNA ends that have limited end resection (Figure 5A).

In cases of extensive resection, RNA-templated DSB repair could proceed with the aid of an RT. In addition, the RNA transcript can mediate DNA modifications in the absence of Rad52.[169,264] Remarkably, RNA retains some ability to modify its DNA gene *in cis* even in the absence of an induced DSB. In this scenario, the RNA partially hybridized to DNA may form an R-loop structure with the intact dsDNA. The failure to remove R-loops from the DNA duplex leads to increase in DNA damage, recombination rates, mutation frequencies, and loss of heterozygosity.[76,265] It is generally thought that the majority of R-loop-induced genomic instability stems from encounters between the DNA replication machinery and the altered chromatin environment in the vicinity of an R-loop.[266,267] If a spontaneous or induced DSB occurs near the R-loop site, repair by C-NHEJ may occur, with the RNA facilitating end ligation by C-NHEJ proteins through end-bridging (Figure 5B). Thus, the RNA transcript could be a donor in DSB repair either to allow HR or to guide C-NHEJ, possibly depending on the cell-cycle phase, the types of DSB lesions, and the extent of DNA end resection.

7. GENOMIC REARRANGEMENTS AND RNA: LESSONS FROM CILIATES

Probably the most striking evidence for a physiological role of RNA in controlling genome stability is in ciliates. This is because in the ciliated protozoans sncRNAs have been shown to be involved in the epigenetic transmission of information between maternal nuclei and their derivatives, mediating large-scale genomic rearrangements and elimination or retention of specific DNA sequences.[8]

7.1. Oxytricha

All ciliates, including the stichotrich *Oxytricha trifallax*, are characterized by nuclear dimorphism. These large unicellular ciliated protists contain two separate sources of genetic information: a transcriptionally silent germline micronucleus that is exchanged during matings, and a transcriptionally active somatic macronucleus containing tens of thousands of amplified gene-sized DNA molecules called "nanochromosomes" that are transcribed during asexual growth of the cells.[268] These macronuclear nanochromosomes are the smallest known genomic DNA molecules in nature, with an average size of 3.2 kb, and are present at 100−100,000 copies per macronucleus.[269,270] The micronuclear genome closely resembles that of a canonical eukaryotic genome with many genes organized on long chromosomes. However, micronuclear genes are typically interrupted by many short nongenic DNA sequences called internally eliminated sequences (IESs). For approximately 3,500 of these genes (~20% of genes in the *Oxytricha* genome), the macronuclear destined sequences (MDSs) that are connected upon IES removal exist in a

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

nonlinear, scrambled order.[271] When a mating occurs under the desired environmental conditions, two ciliates fuse and the process of macronuclear development from a newly acquired diploid micronucleus begins. At this point ciliates undergo a polytene chromosome stage (repeated rounds of micronuclear DNA replication without nuclear division, leading to large, banded chromosomes), eliminate more than 90% of their noncoding micronuclear germline genome (transposable elements, repetitive satellite sequences and IESs), fragment their chromosomes, and then sort and reorder the many thousands of nonlinear macronuclear destined sequences (MDSs) that will form functional genes. Ligation of MDSs, *de novo* telomere addition, and amplification of macronuclear nanochromosomes to the appropriate high copy number completes the development of a new, functional macronucleus (For general reviews of the process of macronuclear development in ciliates see refs 272−275).

Previous work has illustrated the roles of RNAs in mediating IES recognition/removal and the unscrambling events that ultimately take place during *Oxytricha* macronuclear development. Although the junctions of MDSs and IESs contain short direct repeat sequences that are likely involved (so-called "pointers"), they seem to act more as a structural requirement for unscrambling and DNA splicing, rather than for recognition by the necessary protein machinery.[276] Instead, maternal guide RNA templates that are transcribed in the maternal macronucleus from the nanochromosomes have been hypothesized to mediate this massive genomic rearrangement process.[277,] Long noncoding sense and antisense RNA transcripts, corresponding to entire macronuclear DNA molecules, can be detected, peaking at 12−24 h postconjugation, and these are transported to the newly developing macronucleus to provide guide templates for the correct rearrangement, deletion, and sometimes inversion of the micronuclear DNA sequences[250] (Figure 6). Microinjection of synthetic double-stranded DNA or RNA versions of alternatively rearranged nanochromosomes into the macronucleus of mating cells leads to changes in the reordering of MDSs, not only in the injected cells, but in offspring as well, suggesting epigenetic inheritance through these RNA templates.[250]

Recently, it has also been reported that Oxytricha produce and store RNA copies of whole somatic nanochromosomes during macronuclear development, which are derived from the maternal macronucleus before degradation. More than 60% of Oxytricha nanochromosomes have a corresponding RNA-cached copy, whose levels fluctuate throughout development, suggesting that not all developing macronuclear chromosomes undergo DNA rearrangements simultaneously.[278] While extensive studies of the gene expression program during macronuclear development in Oxytricha have implicated hundreds of proteins playing roles during these developmental processes, much remains to be elucidated when it comes to biogenesis, processing, and function of sncRNAs.[270,279]

A novel class of macronuclear-derived 27 nt small RNAs, called 27macRNAs, that are highly upregulated after *Oxytricha* conjugation, peak at 24 h postmixing of complementary mating types.[280,281] These 27mers are derived from the parental macronucleus as opposed to the micronucleus, have a strong 5′ U bias, and do not possess a 2′-O-CH_3 group modification at their 3′ end, typical of certain classes of small RNAs in other ciliates.[280−282] These 27macRNAs have been shown to associate with a PIWI homologue called Otiwi1 and specify which segments of micronuclear DNA will remain protected from degradation throughout macronuclear development. It has been suggested that this may occur through methylation and hydroxymethylation of cytosine residues within the DNA sequences to be eliminated.[280,283] Indeed, microinjection experiments of 27 nt RNAs containing a 5′ U, corresponding to IES regions of the genome to be eliminated, lead to their retention after the completion of the macronuclear development program. However, the relationship between the PIWI-associated 27macRNAs and the long noncoding dsRNA "guide templates" implicated in MDS rearrangements remains unknown.

Ciliates have evolved two extraordinary genomes that demonstrate the complexity of epigenetic inheritance and DNA manipulation in eukaryotes. To date, although the general timing of events involved in macronuclear development has been fairly well characterized, the molecular mechanisms underlying many of these processing events remain poorly understood. Genome-wide studies and high throughput sequencing of mRNAs expressed throughout ciliate macronuclear development have allowed the identification of many factors likely playing roles in the numerous RNA-mediated processes occurring during this time. A disproportionate number of the genes identified as upregulated encode proteins that are involved in DNA and RNA metabolism processes, with the majority of these genes encoding evolutionarily conserved proteins in higher level eukaryotes. A recent study in *Oxytricha* shows that a striking number of differentially expressed macronuclear development genes in ciliates are preferentially expressed in animal germline cells, illustrating that ciliates possess a highly conserved and primordial set of factors involved in germline and stem cell maintenance.[279] Thus, ciliates offer a unique and convenient system to study the influence of noncoding RNAs on genome integrity and transgenerational inheritance.

7.2. *Paramecium* and *Tetrahymena*

Macronuclear development has been more extensively studied in the distantly related ciliates *Tetrahymena* and *Paramecium*, where it has also been shown that epigenetic information from the parental macronucleus guides the elimination and retention of DNA sequences in the developing macronucleus. During the sexual life cycle of these ciliates, the entire parental micronuclear genome is transcribed bidirectionally to produce long, double-stranded RNAs early on in macronuclear development.[284] In *Paramecium*, these double-stranded RNA precursors are cleaved by DICER-like enzymes DCL2 and DCL3, to produce a class of 25 nt small RNAs, called scan RNAs (scnRNAs),[285−288] which are transported to the parental macronucleus where those with homologous macronuclear sequence are degraded. The remaining scnRNAs corresponding to micronuclear-specific sequences survive this filtering step and are transported to the developing macronucleus where, in association with PIWI proteins Ptiwi1/9, they "scan" the genome and mark IESs for excision and elimination.[288,289] Notably, this is the opposite of *Oxytricha*, where PIWI-associated 27 nt piRNAs mark DNA sequences for retention. Although the mechanism of DNA excision and elimination requires further investigation, it has been shown to depend on a "domesticated" piggyBac transposase called PiggyMac.[290−293] When IESs are excised from the developing macronuclear chromosomes, they have been shown to circularize, or concatamerize before circularization, depending on their size, to act as templates for the transcription of a second class of small RNAs called iesRNAs.[293−295] Precursors of iesRNAs are processed by the DICER-like enzyme DCL5, to

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

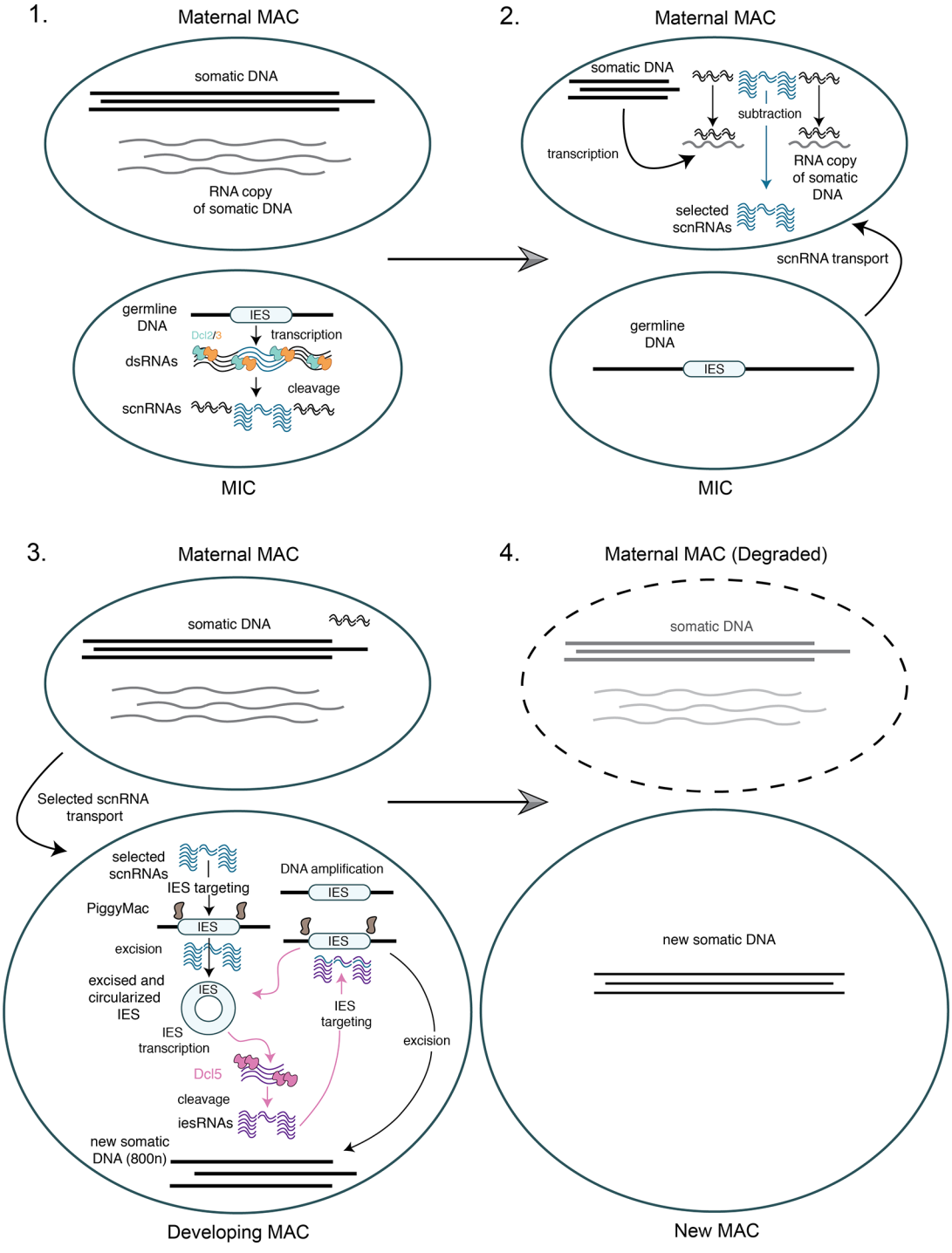

Figure 7. The scnRNA model for genomic rearrangements in *Paramecium*. (1) The micronculear genome is transcribed bidirectionally to produce long dsRNAs, which are processed by DICER-like enzymes DLC2/3 to produce 25 nt long scnRNAs. (2) scnRNAs are transported to the maternal macronucleus where they "scan" the macronculear genome via interaction with RNA transcripts of somatic DNA. scnRNAs pairing to homologous MDS sequences are filtered out and degraded, leaving only those corresponding to micronuclear-specific sequence or IESs. (3) Selected scnRNAs are transported to the developing macronucleus where they target the excision of IESs by the excisase PiggyMac. Excised IESs circularize and are transcribed to dsRNAs that are processed by the DICER-like enzyme DCL5 to produce iesRNAs. iesRNAs then ensure the precise and efficient excision of all remaining IESs from the developing macronuclear genome. (4) Development of the new macronucleus is complete, with the newly formed macronuclear genome matching that of the maternal macronucleus. The maternal macronucleus is degraded during the last steps of macronuclear development.

produce 22−31 nt small RNAs complementary to the sequence of excised IESs with a bias toward IES ends.[288] A second class of sRNAs are also produced later in macronuclear development in *Tetrahymena*, but it remains unclear whether or not these late-scnRNAs play the same role as iesRNAs in *Paramecium*.[296] While iesRNAs have more variation in length than scnRNAs and peak in expression much later during macronuclear development, both species possess a strong 5′ U bias. IesRNAs have been

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

implicated in genome quality control and help to ensure the complete and precise removal of all remaining IESs matching these sequences from the amplified chromosomes (approximately 800n) in *Paramecium*, leading to a new, functional macronucleus[288] (Figure 7). The general events of this "scnRNA model" also occur during *Tetrahymena* macronuclear development, although the specific details and associated factors may vary. It is worth noting that while *Tetrahymena* and *Paramecium* eliminate IESs during macronuclear development, these ciliates do not possess scrambled micronuclear genes such as the stichotrichs. In addition, their macronuclear chromosomes are much larger, coding for hundreds of genes instead of just one or two, typical of *Oxytricha* nanochromosomes.[268,270]

7.3. Epigenetic Inheritance in Ciliates, a Lesson from Plants

Although epigenetic inheritance has been well characterized and studied in ciliates, many questions still remain in the field. It is still poorly understood how the precursors to the different sRNA classes are initially transcribed and the processing machinery used to target particular genomic regions for either retention or elimination. Using the model organism *A. thaliana* for comparison, which displays another well studied system of epigenetic inheritance, some inferences can be drawn, although significant differences exist. In addition to the canonical RNAP II, *A. thaliana* possesses two additional nuclear multisubunit RNA polymerases, named RNA polymerase IV and RNA polymerase V, which play nonredundant roles in RNA-mediated gene-silencing.[297] RNA polymerase IV is used to transcribe the precursors to siRNAs, while RNA polymerase V is responsible for transcribing nascent transcripts necessary for AGO-associated siRNA targeting. Like plants, ciliates have undergone whole genome duplications that have led to paralogous transcription subunits and machinery, including paralogs of the first and second largest subunits of RNAP II (RPB1 and RPB2), reminiscent of plant Pol IV and Pol V. In *Oxytricha* it has been shown that the RNAP II machinery paralogs play a role in development independent from transcribing sRNA precursors, but it is unclear if these separate paralogous RNAP II subunits, also upregulated in *Paramecium* and *Tetrahymena*, are associated with general transcriptional machinery.[298] In addition, it is unknown how regions of the developing macronuclear genome are targeted for elimination or retention. In the field, this has been hypothesized to depend on production of nascent RNA transcripts at particular loci, but this has not been shown directly. Further studies are necessary to elucidate the underlying mechanisms responsible for genome rearrangements in ciliates, but with the power of next generation sequencing (NGS) of entire genomes and epigenomes, along with reverse genetic approaches, it will be possible to fill in the remaining gaps in our knowledge of these processes.

8. INVOLVEMENT OF RNA IN TELOMERE MAINTENANCE

8.1. Consequences of Dysfunctional Telomeres

Telomeres are the distal tips of linear chromosomes, composed of short, guanosine-rich hexameric tandem repeats. In humans the 5′-TTAGGG-3′/3′-CCCTAA-5′ sequence is repeated approximately 2000 times to generate a telomere length between 10 and 15 kb pairs. Telomeres pose a particular conundrum for the cell due to the fact that they resemble the DNA end structure typically present at DSBs.[299,300] Unlike DSBs, which must be repaired to ensure cell survival, telomeres actively inhibit DNA repair and DDR signaling. This is achieved by the recruitment of a specific set of proteins, collectively called "Shelterin", which directly inhibit DDR at telomeres, and by assuming a secondary structure, reinforced by Shelterin, called the t-loop, in which the tip of the telomere loops back on itself, thereby hiding the end.[10−12,301−303]

Despite the fact that telomeres are refractory to DNA repair activities, KU (see section 4.1.1 for details) is associated with normal telomeres. While C-NHEJ inhibition is achieved by the Shelterin component TRF2 that prevents the recruitment of Ligase IV from telomeres,[301,304] the presence of KU seems to be important as a second line of inhibition of HR and alt-NHEJ pathways, in a Shelterin-free environment.[305] Beyond DNA repair inhibition, localization of KU to telomeres could also be important for telomere length regulation through its binding to the RNA component of telomerase, TERC.[127] This is an additional example of functional RNAs interacting with DNA repair proteins.

Telomeric "repair"—for instance a fusion of a telomeric chromosome end to another telomere or to an interstitial telomere repeat-containing region—can lead to genomic rearrangements, with consequent potential changes in ploidy, and eventually may contribute to cell transformation.[306−309] Proliferating cells not expressing telomerase, or using other mechanisms of telomere elongation, eventually accumulates telomeres that are critically short, or "dysfunctional". Dysfunctional telomeres are akin to exposed DNA ends of DSBs and are promptly recognized by the DDR machinery.[111,306,310,311] Telomere shortening can occur gradually over multiple population doublings as a result of the so-called "end-replication problem", in which the cell is unable to replicate all the way to the end of the telomere lagging strand. Shortening can also be due to the abrupt loss of telomeric material via DSB formation, potentially as a consequence of DNA replication stress. Critically short telomeres, as well as DSBs within telomeric repeats, trigger a DNA damage response that cannot support efficient DNA repair,[301] provoking a protracted, likely permanent, DNA damage induced-checkpoint that arrests cell-cycle progression. The permanent cell-cycle arrest associated with the protracted DDR caused by unrepaired DNA ends[312] is referred to as cellular senescence.[111,313]

Some cells can respond to telomeric shortening by *de novo* telomere elongation through telomerase via catalytic extension of the telomere, or HR, using other sources of telomeric material as templates for extension. Telomerase is a reverse transcriptase (TERT) that carries its own RNA template (TERC), and it is preferentially recruited to the shortest telomeres, presumably due to the absence of telomerase inhibitory proteins, to maintain their length and hence avert senescence.[314−316] Telomerase is the primary means of telomere lengthening in stem cells, where it is expressed at low, but detectable, levels.[317] Telomerase is also responsible for the maintenance of telomeres in approximately 85−90% of human malignancies.[318] In most of the remaining tumor types, the HR-based Alternative Lengthening of Telomeres (ALT) mechanism is activated, which relies on a form of HR between a chromosomal telomere and other telomeric material for elongation. However, some reports have recently challenged the idea that an essential feature of cancer is the acquisition of a telomere maintenance mechanism. Some patient-derived melanoma and neuroblastoma cells do not express telomerase nor activate the ALT mechanism, and indeed, their telomeres shorten during serial passages in culture. The phenotype associated with these cancer cells has been referred to as ever-shorter telomeres.[319,320] In support of this notion,

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

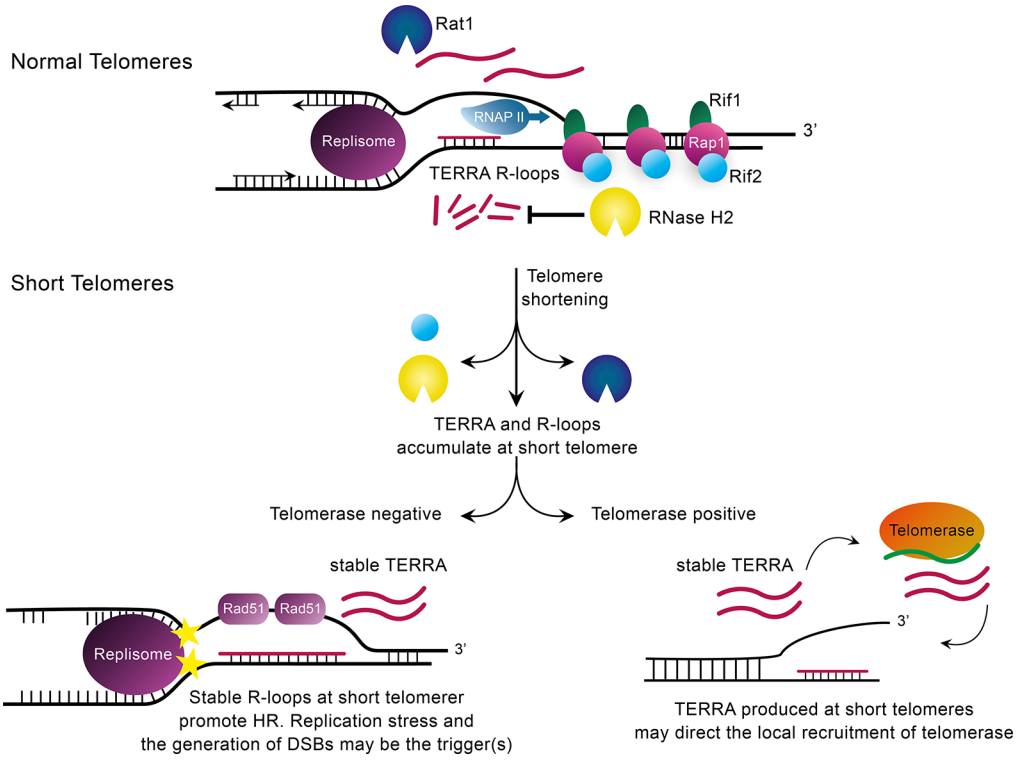

Figure 8. Roles of TERRA in telomere elongation. At normal length telomeres, TERRA is transcribed and forms R-loops. To allow unperturbed replication, TERRA and its R-loops are degraded by Rat1 and RNase H2, respectively. Due to the absence of Rat1 at shortened telomeres, TERRA becomes stabilized and accumulates (bottom right). In telomerase positive cells, TERRA from a short telomere can colocalize in the nucleoplasm with telomerase and then reassociate specifically with the shortened telomere. This type of regulation may facilitate telomerase in finding the correct telomere to elongate. Although R-loops also accumulate at short telomeres, it is not clear how they affect the regulation or telomerase. R-loops accumulate at shortened telomeres due to the mislocalization of Rif2 and hence RNase H2. In telomerase negative cells telomeric R-loops promote the recruitment of the HR machinery (Rad51) to critically short telomeres and contribute to DDR activation. This may be due to an encounter between R-loops and the DNA replication machinery and eventually the generation of a DSB. HR promotes length maintenance by using a homologous telomere as a template to elongate from. It is not understood if stable TERRA (not in an DNA−RNA hybrid) also contributes to HR at short telomeres. Adapted from Graf et al., *Cell* 2017.

bioinformatics analysis of a large cohort of human tumors (18,430 samples) has recently reported that approximately a fifth of the analyzed samples neither expressed telomerase nor harbored alterations in ATRX or DAXX genes, which are commonly mutated in ALT.[321] These new findings add an additional layer of complexity for cancer treatment because they suggest that, at least in some cases, prevention of telomere shortening is not required for oncogenesis nor for cancer progression, thus potentially blunting therapeutic approaches targeting telomere maintenance mechanisms.

Recent studies from multiple laboratories have shown that the regulation of ncRNA transcribed from telomeric regions plays an important role at damaged/shortened telomeres to promote activation of the DDR and, hence, their repair.[181,322−326] In addition to chromosome ends, telomeric ncRNA associates with multiple nontelomeric loci to stimulate transcription.[327] Therefore, the regulation of ncRNA at telomeres may be a critical determinant with regard to the rate at which a cell enters a state of replicative senescence.

8.2. Telomere Repeat-Containing RNA

Telomere Repeat-Containing RNA (TERRA) is one example of a lncRNA harboring telomeric repeat sequences.[12,328] TERRA transcription is initiated in the subtelomeric region and continues into the telomeric repeats, although it is unlikely to reach chromosome ends.[329−332] Recently, there have been important observations describing how TERRA is regulated at different telomeric states.[12] TERRA levels are tightly regulated with respect to cell-cycle. In G1 the levels are maintained low, but at the G1/S transition TERRA is transiently upregulated and then subsequently degraded as the cells progress through the S phase and into G2.[326,333,334] In the yeast *S. cerevisiae*, the degradation throughout the S phase is carried out by the 5′ to 3′ RNA exonuclease Rat1 so that TERRA is removed at approximately the time when telomeres are replicated by DNA polymerase and/or extended by telomerase. When telomeres become short, this precise regulation is altered due to the inability of Rat1 to associate with short telomeres[326] (Figure 8).

In *S. cerevisiae*, TERRA produced from a critically short telomere has the ability to associate with telomerase in the nucleoplasm, which is then recruited specifically to the shortened telomere where the TERRA molecule was produced, presumably to promote telomerase-mediated elongation.[324] Consistently, in *S. pombe* it was demonstrated that polyadenylated TERRA levels increase upon telomere shortening, and these TERRA molecules, in turn, associate with the telomerase enzyme.[323] Moreover, in *S. pombe* experimentally induced expression of TERRA from a single telomere results in telomerase-mediated elongation exclusively at the telomere overproducing TERRA. Similar experiments performed in human cells have also demonstrated that TERRA and telomerase interact in cell extracts.[327,335] Taken together it appears that TERRA may function as a S.O.S. signal of

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

sorts at short telomeres, to fetch and direct telomerase to the telomere in need of elongation[336] (Figure 8, bottom right).

Since TERRA degradation is impaired specifically at short telomeres in late S phase (when telomerase acts), this may increase the chances that a productive TERRA-telomerase interaction occurs to facilitate elongation of the short telomere. In future studies, it will be important to interrogate the relationship between TERRA cell-cycle regulation, telomere replication, and telomerase in more depth. Furthermore, it remains unresolved as to how TERRA may help to direct telomerase to the "right" telomere. Is it the subtelomeric sequence information that is important to form DNA−RNA hybrids, or are RNA−protein interactions the key? There are also inconsistencies that remain to be addressed, such as the observation that the overexpression of TERRA in *S. pombe* leads to telomere elongation[323] while in *S. cerevisiae* this leads to telomere shortening,[331,337] although the discrepancy may be due to different expression levels in the different model systems. It is also not understood why the impairment of the RNA exonuclease Rat1 or the NAD-dependent deacetylase Sir2, in budding yeast, does not lead to telomere lengthening, despite high TERRA levels in these mutants. Although these open questions remain, there is ample evidence implicating TERRA as an important intermediate to promote the telomerase-mediated repair of shortened telomeres.

Another important feature of TERRA is that it can form DNA−RNA hybrids, which likely lead to the formation of R-loops.[322,325,338,339] Similar to TERRA levels, telomeric R-loops are regulated in an identical cell-cycle dependent manner. RNase H2, which degrades the RNA moiety of an DNA−RNA hybrid, gets recruited to telomeres, approximately at the time of telomere replication, and promotes R-loop removal.[326] In a manner once again reminiscent of the regulation of TERRA levels, TERRA R-loops are no longer degraded in a timely manner when telomeres become critically short, due to the inability of RNase H2 to properly localize to shortened telomeres (Figure 8). This delay in R-loop resolution likely results in an encounter between the replication machinery and R-loops, an event that triggers HR.[326] The persistence of TERRA R-loops is an important feature of critically short telomeres in the absence of telomerase. If R-loops are removed, through RNase H1 overexpression, cells enter replicative senescence at an accelerated rate and the HR machinery fails to associate with short telomeres.[325,326] In contrast, when R-loops are allowed to accumulate at telomeres, through RNase H2 deletion, the rate of senescence is significantly reduced. Therefore, at normal length telomeres, TERRA and its R-loops are produced at each cell-cycle at the G1/S transition, only to be degraded, which likely facilitates replication passage and does not promote elongation by telomerase or HR. When a telomere is damaged/shortened, the G1/S up-regulation of TERRA occurs in a timely manner but the subsequent degradation is rendered defective, so that TERRA and R-loops persist into late S phase and promote telomerase and HR-mediated elongation, respectively (Figure 8). It will be interesting to determine the significance of G1/S-specific TERRA up-regulation, as it may occur as a precautionary measure in the case of telomere shortening, to regulate telomeric replication origins or to regulate gene expression elsewhere in the genome.[327] In human cells TERRA R-loops may perform similar functions, as DNA−RNA hybrids are enriched at telomeres in human Immunodeficiency, Centromeric instability and Facial anomalies (ICF) patient cells, which have extremely short telomeres and are largely responsible for DDR activation at chromosome ends.[322]

Telomeric R-loops are also important beyond their role during replicative senescence. Cancer cells that employ the HR-based ALT mechanism show increased TERRA levels and telomeric R-loops. The overexpression of RNase H1 impedes telomere maintenance in these cells, while its depletion causes rapid telomere loss.[339,340] It will be important to understand how RNase H1 contributes to HR in ALT cells as well as during replicative senescence. Similarly, the overexpression of RNase H1 in yeast cells prevents the generation of type II survivors, which are considered the yeast ALT equivalent.[341] The microbial pathogen *Trypanosoma brucei* uses telomeric TERRA R-loops to induce HR-dependent antigen switching to evade immune detection and increase pathogenesis.[342] Therefore, the use of TERRA R-loops at telomeres to stimulate HR is evolutionary conserved, although exploited for different means.

In human cells, telomere dysfunction induced by removal of TRF2 leads to increased TERRA levels at all transcribed telomeres.[332,343] Furthermore, the TRF2 homodimerization domain, which induces chromatin compaction[344] and prevents DDR activation,[345] represses TERRA transcription independently of p53 and does not rely on ATM-dependent DDR signaling.[332] The UUAGGG-repeat array of TERRA transcripts directly bind to SUV39H1 H3K9 histone methyltransferase, sustaining the accumulation of the heterochromatic mark H3K9me3 at dysfunctional telomeres.[332] Similarly, TERRA has been reported to accumulate H3K9me3 at telomeres,[346] indicating a functional role of TERRA in heterochromatin reorganization at telomeres. In contrast to the idea of TERRA transcription arising from several individual subtelomeres,[332] recent reports propose that TERRA transcription is restricted to one, or two at most, subtelomeres in mouse and human cells.[347,348] Further studies are needed to clarify these apparent inconsistencies.

Although much effort has been focused on understanding TERRA's function at telomeres, it has recently been shown that TERRA also has nontelomeric functions.[327] TERRA physically associates with thousands of nontelomeric loci where it frequently positively regulates transcription. TERRA binding sites overlap strikingly with those of ATRX—a chromatin-remodeling protein known to aid deposition of H3K9me3 at telomeres—and the two appear to compete at genomic sites. Indeed, TERRA and ATRX physically interact, and TERRA can displace ATRX from DNA templates, suggesting that it may remove ATRX from chromatin. In agreement, upon TERRA depletion, ATRX foci accumulate in the nucleus, both at telomeres and elsewhere. It is noteworthy that TERRA is upregulated in ALT cancers where ATRX is frequently mutated. In such a scenario, TERRA would be potentially unleashed to activate transcription via removal of ATRX, thereby inhibiting deposition of repressive H3K9me3 chromatin modifications; it will be interesting to determine how R-loops, ATRX, and TERRA-regulated transcription are coordinated. Moreover, gapmer oligonucleotides (see section 10 for details) targeting telomeric repeats have been used to fully deplete TERRA in mouse cells,[327] leading to an increase of telomere dysfunction as well as a greater occurrence of other multiple telomeric pathologies, such as loss or duplication of the telomeric repeats and fusions between sister chromatids.

These findings, together, indicate a functional role for TERRA transcripts in the maintenance of telomere integrity in both mouse and human cells.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

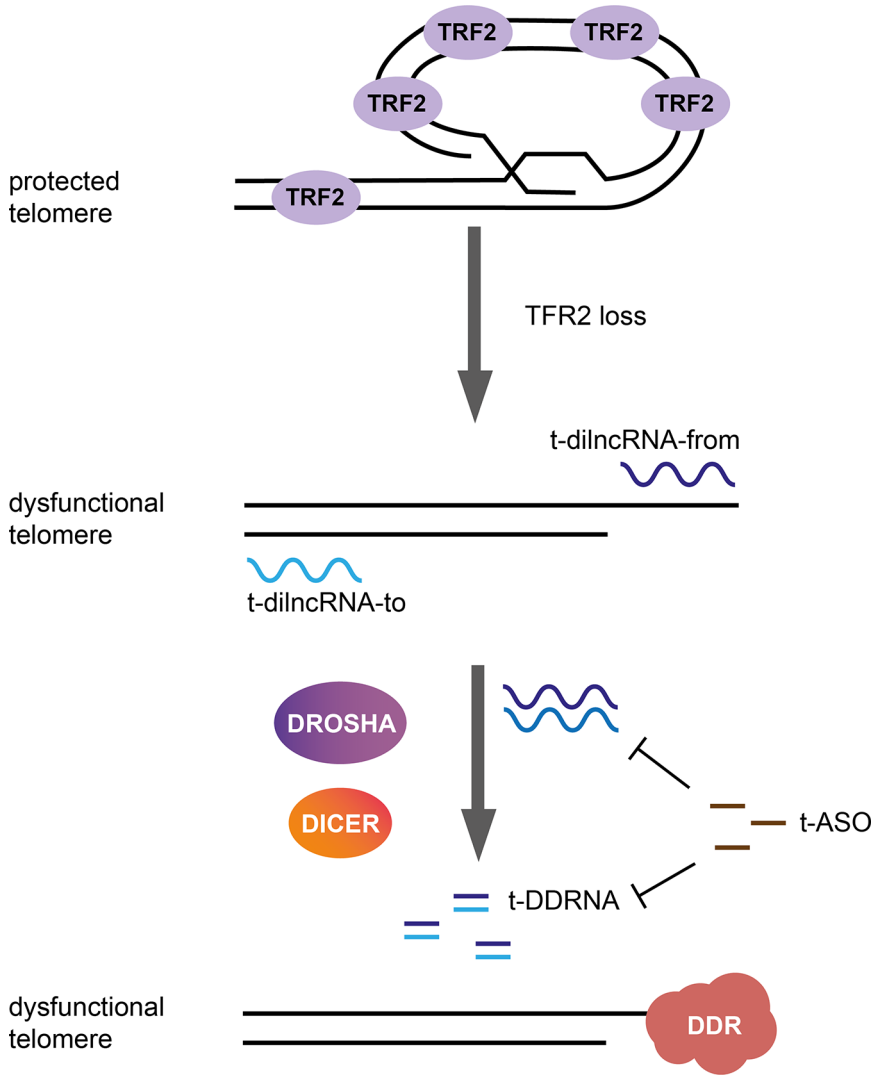

Figure 9. Proposed model for the generation of telomeric dilncRNAs and DDRNAs at dysfunctional telomeres. Upon TRF2 loss, C-rich t-dilncRNA-from (purple) and the G-rich t-dilncRNA-to (light blue) are induced. DROSHA and DICER process the long double-stranded RNA, likely the outcome of paired or folded t-dilncRNAs, generating t-DDRNAs. Antisense oligonucleotides containing telomeric sequences t-ASOs interfere with t-dilncRNA and t-DDRNA functions and allow inhibition of DDR at dysfunctional telomeres.

8.3. Other Noncoding Telomeric Transcripts

TERRA is not the only ncRNA transcript produced from telomeric DNA. In *S. pombe* a C-rich telomeric transcript (transcribed in a telomere to centromere direction) referred to as ARIA has been reported. Unlike TERRA, ARIA is composed purely of telomeric repeats and does not harbor subtelomeric sequences.[349,350] Similar to TERRA, ARIA levels increase when telomeres are rendered dysfunctional by deleting components of the telomere specific Shelterin-like complex in fission yeast. This behavior is reminiscent of the transcriptional induction of both dilncRNA-from and dilncRNA-to species upon DNA damage.[95]

An interesting observation comes from mouse embryonic stem cells, where a positive correlation has been found between the heterochromatin status of telomeres and the level of telomere specific small RNAs (tel-sRNAs), that seem not to be TERRA degradation products nor DICER products.[351] It has been hypothesized that tel-sRNAs belong to the piRNA family as the chemical features, the size, the resistance to β-elimination treatment, are conserved. tel-sRNA may be involved in the establishment of the heterochromatic state at telomeres. Consistently, they are regulated through epigenetic mechanisms, and their levels are positively correlated with the telomeric heterochromatin status.[351] It remains to be determined whether TERRA serves as the precursor molecule for the generation of tel-sRNA.

As has been demonstrated for DSBs[95] (see section 4.2.2 for details), mammalian dysfunctional telomeres, induced via TRF2 removal or functional impairment, also produce dilncRNAs[181] (Figure 9). Telomeric dilncRNAs (t-dilncRNAs) arise from the transcription of both G-rich and C-rich strands of deprotected telomeres in mammalian cells. If, as canonical dilncRNAs, t-dilncRNAs are transcribed from the DNA end, they are unlikely to contain subtelomeric elements; thus, they are different from TERRA. Also, t-dilncRNAs are processed by DROSHA and DICER to produce short RNAs, called telomeric DNA damage response RNAs (t-DDRNAs). These sncRNAs are essential to ensure full activation of DDR signaling at dysfunctional telomeres and repair in the form of NHEJ-mediated fusions. It will be interesting to determine whether t-DDRNAs also play a role in the repair of telomeres by HR or in the maintenance of ALT telomeres. Moreover, the interplay between TERRA and t-DDRNAs has not been investigated. Similar to what has been

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

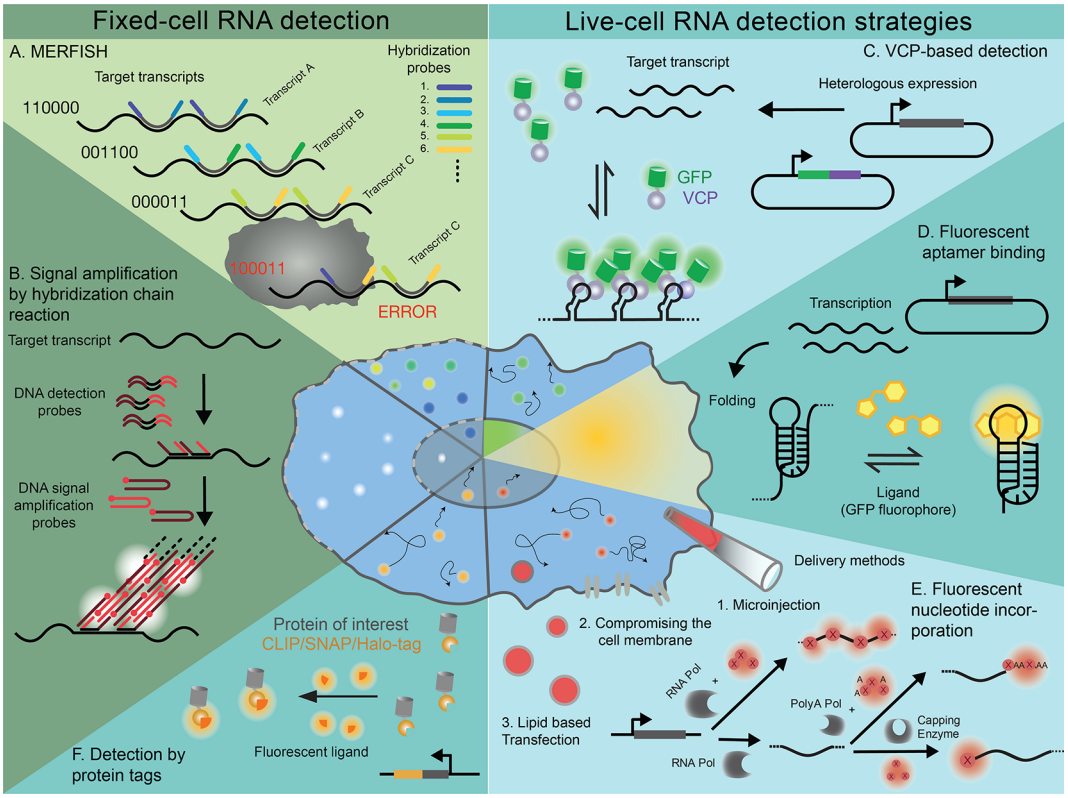

Figure 10. Illustrations of recent major developments in intracellular RNA detection methods.

demonstrated for endonuclease-driven DSBs[95] (see section 4.2.2 for details), the use of ASOs (see section 10 for more detail) inhibiting the functions of t-dilncRNAs and t-DDRNAs efficiently reduces DDR activation at dysfunctional telomeres both in cultured cells and in mouse models. This suggests that telomeric DDR is amenable to specific control by exogenous agents, that have the potential to be developed as therapeutic agents.

9. SINGLE MOLECULE FLUORESCENCE METHODS FOR RNA DETECTION

9.1. Importance of Intracellular RNA Detection in the DNA Damage Response Field

With numerous lines of evidence pointing to RNA's diverse noncoding functions, the imperative to study cellular RNAs in their native context has never been greater. Recently, single molecule fluorescence techniques have been successfully applied to understand and enumerate the intracellular functions and properties of various RNAs and RNA−protein complexes.[352−355] Furthermore, the involvement of both long and short noncoding RNAs in guiding the DDR machinery to sites of DNA damage was discovered in part using observations from single molecule fluorescence microscopy[95] (see section 4.2.2). Rapid progress in intracellular single molecule detection methods to study RNAs and their binding partners has been driven by their inherent advantages over ensemble-averaging techniques. Most salient among these advantages are their abilities to detect, with great sensitivity, concentrations as low as those of most endogenous pathway components (eliminating the need for artificial overexpression, which bears the risk of overwhelming said pathway) and to reveal critical heterogeneities in the maturation, functionality, and spatiotemporal distribution of macromolecules. For a full historical perspective of the field of single molecule fluorescence-based RNA detection, we refer the reader to Pitchiaya et al. 2014.[353] Here we present an overview of key advances in fluorescence-based intracellular RNA detection methods of relevance to the study of the DDR (Figure 10).

9.2. Detecting RNA in Fixed Cells

In situ hybridization (ISH) techniques localize nucleic acids in fixed tissue samples by detecting bound oligonucleotide probes with sequence complementarity to the target transcript. Since the first demonstration of the technique by Gall et al. in 1969,[356] many versions of their protocol, varying in probing strategy and detection method, have been proposed over the years, with the objective of increasing spatial resolution, detection sensitivity, and throughput of RNA transcripts.[357]

While fluorescent *in situ* hybridization (FISH) has been used for the detection of nucleic acids for more than 30 years, its application for single transcript detection and counting has become possible only in the past decade. The premise of single molecule FISH (smFISH) is that multiple fluorophore-labeled nucleic acid probes bind over the length of a target transcript, allowing individual transcripts to be detected as bright spots against a dark background. Automated imaging and counting of these spots then yields information about the subcellular localization of transcripts and cell-to-cell variability in transcript number. Unlike many sequencing technologies, these intracellular transcriptomics methods are not limited to poly adenylated RNAs and are well suited for the study of noncoding RNAs. The sensitivity and ability of these technologies to detect even single transcripts make them particularly suited to studying RNAs involved in the DDR. For example dilncRNAs were detected using smFISH[95] (see section 4.2.2).

9.2.1. Multiplexed Detection. A major challenge for intracellular RNA detection in the past has been the limited

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

number of species that can be probed simultaneously. In this context, multiplexing refers to the process of detecting multiple significantly distinct molecular species in a biological sample. As with protein detection, multiplexing RNA detection can be achieved by multicolor labeling so that each target species (or probe) is tagged with a fluorophore of different emission maximum.[358,359] However, multicolor detection today is limited by the number of colors that can be simultaneously imaged using conventional, single molecule sensitive light microscopy, to at most 7 distinct colors.

An alternative approach to multiplex RNA detection is sequential barcoding. With this strategy, the same diverse transcripts are repeatedly probed using different probe sequences. Individual transcripts are detected in multiple rounds as fluorescent spots, allowing high confidence detection of each molecule over multiple rounds of probing, without the need for multiple colors. To reprobe the same transcript using a different sequence, hybridized RNA probes can be degraded and washed away between successive rounds of hybridization (seqFISH),[360] or the fluorophores on the hybridized probes can be photobleached, allowing the same wavelengths to be used for subsequent rounds of imaging (MERFISH).[361] The sequence of hybridization rounds in which a single spot was detected then allows detection errors to be corrected and false positives to be eliminated, thereby increasing the sensitivity of detection (Figure 10A). Finally, suitable image registration and spot detection algorithms allow seqFISH/MERFISH to theoretically detect and locate hundreds of different transcript sequences at the single molecule level, making them powerful tools to study single cell transcriptomes with spatial information. A complementary technique, FISSEQ, employs rolling circle cDNA amplification to sequence transcripts *in situ*, providing added nucleotide-level resolution to transcripts in cells.[362] These tools are poised to transform the field of RNA quantification and sequencing and are well positioned to aid the discovery of novel, rare, noncoding RNA species, such as dilncRNAs.

9.2.2. Advances in Signal Amplification. The primary challenge for single molecule detection is the need for a sufficiently bright spot signal, to enable super-resolution (~10−20 nm) localization of single transcripts. Common smFISH strategies aim to solve this issue by using multiple (>10) labeled probes to decorate the length of the transcript, and/or the use of illumination strategies such as confocal or HILO illumination to decrease background fluorescence.

One method that has gained attention recently, termed hybridization chain reaction (HCR), achieves signal amplification via self-complementary fluorescent probes that allow single molecule resolution even without specialized illumination schemes. HCR uses fluorescent probes that can self-assemble into long chains or branched structures, where each link is a probe that is bound to two others (Figure 10B). Increased interest in HCR for single transcript detection has led to improved HCR protocols that have been applied to demonstrate RNA detection in whole zebrafish embryos.[363] These advances promise to make single molecule methods more accessible to the general research community by reducing the need for specialized microscopy equipment.

9.2.3. Detecting RNA in Living Cells. Cellular responses to changes in the environment often involve physiological adaptations that occur over time-scales of seconds to minutes. Here, live-cell analysis becomes necessary in order to capture rapid and dynamic physiological processes in real-time. The primary requirements for successful live-cell, real-time RNA visualization strategies are the ability to label transcripts for observation in living cells while retaining their biological functionality, the delivery of labeled RNAs into the cell, and the ability to image the labeled RNAs with high spatial and temporal resolution. The generally lower cellular abundance of RNA transcripts compared to proteins helps to discern closely spaced single molecules, making them easier to study than proteins using these methods.

Strategies for fluorescently labeling RNAs can be broadly classified into methods that label RNA secondary structures and those that label specific nucleotides within the sequence.

9.3. Strategies for RNA Secondary Structure-Based Labeling

The most widely adopted RNA labeling strategy to detect transcripts in living cells has been the use of RNA-binding viral coat proteins (VCPs).[364] This method, first demonstrated by Robert Singer's group, exploits the high specificity and affinity with which VCPs such as the MS2- or PP7-coat proteins bind with their cognate RNA stem-loop structures. These stem-loop sequences are inserted within untranslated regions (usually the 3′UTR) of the transcript of interest, typically in multiple copies (8 to 96). These modified transcripts are then expressed along with fluorescently tagged VCPs (usually expressed as fusions with GFP or mCherry). The fluorescent signal from multiply bound VCPs allows individual RNA molecules to be visualized as single transcripts even without super-resolution techniques (Figure 10C).

A number of groups have independently developed live-cell single molecule translation reporter systems by combining VCP-based mRNA labeling strategies with intracellular protein immunolabeling methods.[365−368] In these methods, intracellular, fluorescent, typically single-chain antibodies (scAb) bind to antigenic sites present on the protein of interest, thereby serving as fluorescent probes for the protein. Classical fluorescent protein tags are of limited utility to study fast processes such as translation elongation when genetically encoded along with the peptide under study, because the maturation of a protein tag itself occurs over longer time scales than those of translation (the fastest maturing GFP variants fold in ~10 min whereas translation occurs over seconds to minutes).[369,370] However, moving the tag from the protein under study to an antipeptide scAb probe in these immonolabeling methods allows rapid processes such as peptide elongation to now be studied in real-time at single molecule resolution, where the detection is only limited by diffusion of the probes and antibody−antigen affinity.[371,372] Because the binding of each fluorescent antibody is reversible, these detection methods are robust against loss of signal from photobleaching of individual fluorescent tags, making them suitable for time-lapse imaging of proteins, albeit with the caveat of substantial fluorescence background from unbound probe. It is conceivable that these live-cell immunolabeling methods can be used to detect sites of DNA damage or monitor recruitment of protein components with greater ease than classical fluorescent protein tags.

Another development in the field of protein detection has arisen from work on protein appendages such as Halo-, CLIP-, and SNAP-tags that can be labeled with specific suicide substrates.[373] These protein tags are genetically encoded, thereby retaining the specificity of fluorescent proteins, but are more versatile, as they can covalently couple to specific membrane-permeable fluorescent ligands (Figure 10F). The advantage of using these small chemical ligands for labeling in cell culture is that they can be easily added, and the unbound probe easily

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

washed away, making fluorescent labeling amenable to pulse-chase experiments. This technology has been applied to study various aspects of telomere biology[374] and protein translation,[368,375] and it is a powerful technology for studying single proteins.

A complementary RNA-detection method using fluorescently tagged, inactivated RNA guided-Cas9 enzyme for intracellular RNA tracking was reported recently,[376] further extending the fluorescent toolset using an extant ribonucleoparticle.

9.4. Strategies for Direct RNA Labeling

Some of the major disadvantages of endogenous labeling—the need for significant sequence modification and protein overexpression—can be overcome by site-specific incorporation of modifiable nucleotides directly into the transcript. Commercial availability of fluorophores with improved photostability and pH tolerance, and development of chemical RNA synthesis and covalent conjugation technologies, allow RNA transcripts to be conveniently labeled *in vitro* for *in vivo* visualization.[377−379]

NTPs conjugated with fluorescent dyes can be used for sequence-specific or nonspecific cotranscriptional labeling (Figure 10E). The ability to control the number of fluorophores incorporated into a transcript further offers the opportunity of stepwise photobleaching analysis, in which the number of fluorophores, and hence labeled molecules, present in individual fluorescent spots can be counted. The intensity traces from individual spots, representing photobleaching curves, can be analyzed to reveal >10 fluorophores per spot.[379]

The primary challenge of labeling by chemical modifications is the delivery of covalently modified RNAs into cells. Diverse delivery strategies in use today range from variable-dosage methods such as vesicle endocytosis (e.g., lipid-based transfection) or delivery via membrane permeabilization (e.g., electroporation, permeabilization by detergents and bacterial pore-forming toxins), to defined-dosage methods such as microinjection, which have been reviewed recently.[380] Microinjection-based, defined-dosage iSHiRLoC (intracellular single molecule high resolution localization and counting) has been successfully applied to understand the temporal evolution of miRNA maturation, target-binding, turnover, and subcellular localization in the cell.[352,354,381]

iSHiRLoC was applied recently to study the localization of the small DDRNA cleavage products of DROSHA and DICER. In addition to being used to observe localization of DDRNAs to sites of DNA damage, this tool has been used to investigate the functional role of these RNAs in DDR focus formation by virtue of its ability for controlled RNA delivery[95] (see section 4.2.2). Together, live-cell single-RNA and -protein visualization technologies present a formidable toolbox that allows novel molecular functions in DDR to be probed in situ, in real-time.

9.5. Other Strategies for RNA Detection

Another class of RNA detection methods relies on the ability of a small molecule ligand, such as a GFP-fluorophore mimic—difluoro-4-hydroxybenzylidene imidazolinone—to emit enhanced fluorescence upon binding to an RNA aptamer, such as the prototypical "Spinach" aptamer[355] (Figure 10D). Such aptamer sequences can be inserted into transcripts, and the enhanced fluorophore intensity of the ligand upon binding both reduces background and can be used to read out the concentration level of these RNAs. These aptamer-based methods have been applied to detect toxic RNA aggregates,[382] to detect RNA modification activity,[383] and as metabolite-sensors.[384] Orthogonal to the protein-based RNA detection strategies discussed above, these methods hold great promise for intracellular RNA detection. The discovery of brighter turn-on and higher-affinity dye-aptamer combinations[385,386] suggests that this technology may find broader applicability for single molecule detection in the future.

Multiple developments in the field of fluorescence-based intracellular RNA/ribonucleoprotein detection over the past decade are a testament to the growing awareness of the great importance of spatial and temporal information for understanding cellular RNA biology. The discoveries enabled by the technical advances discussed here continue to underscore the importance of RNA in cellular physiology, ultimately reaffirming the rise of smart RNAs, including in the DDR.

10. ANTISENSE OLIGONUCLEOTIDES AS LAB TOOLS AND THERAPEUTIC AGENTS

Given the emerging role of several distinct RNAs in multiple physiological processes often of clinical relevance, interfering with RNA functions can be exploited as therapeutic strategies. One of the most powerful sets of tools to achieve this are antisense oligonucleotides (ASOs) which bind to their RNA target directly through Watson−Crick base pairing.

Based on their mechanism of action, ASOs can be divided into two classes. Gapmers have a central DNA region of 8−12 nucleotides, flanked by 2−3 chemically modified nucleotides on each side. These modifications are designed to increase affinity and stability of binding (see below). When bound to their RNA target, the central part forms a DNA−RNA hybrid, generating a substrate for the activity of cellular RNase H enzymes that degrades the RNA strand.[387]

Blockers do not require a specific position of the modified nucleotides and are often referred as "mixmers". Typically modified nucleotides are present every 2−3 deoxynucleotides. Very short ASOs (8 nucleotide-long) containing only modified nucleotides have also been described and shown to be effective.[388] Blockers do not activate degradation by nucleases, but they instead impose a steric block, preventing the interaction between their targets and other molecules, such as other nucleic acids or proteins.

In particular if an impact *in vivo* is desired, ASOs require chemical modifications which improve their stability, boost their binding to the target, reduce off-target effects, and decrease toxicity.[389] The phosphorothioate (PS) backbone is widely used to improve nuclease stability and pharmacokinetics mainly through its increased hydrophobicity (thus increased cell membranes permeability) compared to the natural phosphate group and through its capacity to avidly bind serum proteins such as albumin, thus avoiding clearance by kidneys.[390] As alternatives to the PS backbone, two uncharged chemistries are also used: the phosphorodiamidate morpholino oligomer (PMO)[391] and the peptide nucleic acid (PNA)[392] backbone. These modifications increase the stability and the binding affinity to their target; however, differently from the PS modification, they are not suitable for the RNase H-mediated degradation of the target RNA.

To further increase the binding affinity and nuclease resistance, sugar modifications have been developed, which are typically inserted at the 2′ position.[387] The most commonly used are the 2′-*O*-methyl (2′-OMe), 2′-*O*-methoxyethyl (2′-MOE), and 2′-fluoro (2′-F) modifications of RNA. Other sugar modifications include conformationally constrained nucleotides, such as locked nucleic acid (LNA), constrained ethyl (cEt), and tricyclo-DNA (tcDNA). These nucleotide analogues have an

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

extra bridge arising from the 2′ position, reducing the torsional flexibility of the sugar backbone and the entropy of duplex formation, strongly increasing the affinity and the specificity to the target. However, the introduced chirality of the backbone associated with these chemistries remains an unresolved issue.

Compared to small molecule drugs, ASOs are easier and faster to design, since the only information needed for their development is the sequence of the target RNA. In this regard they are prototypically regarded as the ideal tool to inhibit the functions of "undruggable" targets, typically cell factors lacking an enzymatic activity. In addition, biodistribution, toxicity, and stability are defined mainly by chemical and structural architecture, such as chemical modifications of sugars, bases, and backbone. It is therefore possible to predict to some extent the behavior of different drugs against different targets on the bases of previous studies of similar molecules with different sequences. However, in some cases the sequence itself can influence ASO toxicity, mainly caused by off-target effects, which can also be hybridization-independent. These toxic effects include pro-inflammatory events, immunostimulation, and liver and kidney toxicity.[393] This sequence-specific toxicity is in most cases unpredictable; however, it can be reduced by altering the ASO length and/or the position of the chemically modified nucleotides.

Perhaps the most important challenge for ASO-based therapeutics is the delivery to its target.[394] Differently from most drugs, which are small (less than 500 Da) and hydrophobic, ASOs typically weigh a few kDa and contain many negative charges.[389]

To be effective *in vivo* ASOs must first escape circulation and reach the target tissue. Their biodistribution is different in various organs;[394] for example, in the liver, endothelium is characterized by fenestrations between the cells, allowing for a more efficient delivery. Some organs are instead completely inaccessible, like the brain. Indeed, ASOs cannot cross the blood-brain barrier. To overcome this issue, ASOs can be administered locally, in the case of brain target by intrathecal injection into the cerebrospinal fluid.[395] An extra benefit of this strategy is that the blood brain barrier prevents ASOs to enter the bloodstream and be cleared by kidneys.

A second obstacle to overcome is the subcellular localization of ASO.[396] Unconjugated ASOs are taken up into cells by endocytosis and pinocytosis, in the absence of a delivery agent. But in order to reach their target, ASOs must escape from the endosomes to get into the cytoplasm or the nucleus. This is a slow process, and the efficiency can vary in different cell types.

Many ASO-based therapies target mature mRNAs to induce gene silencing. This can be achieved through RNase H-mediated mRNA degradation, or translation inhibition. Other ASOs are instead complementary to pre-mRNAs, blocking the donor or the acceptor splice site, thereby preventing the binding of splice factors.[397] Both of these strategies can be applied to pathologies caused by overexpression, or a gain of function mutation, of a specific gene, or to correct an irregular splicing event, thus modulating the biosynthesis of different protein isoforms with a possible therapeutic function. Beyond affecting gene expression levels, ASOs targeting mRNA sequences can also prevent the formation of detrimental nuclear structures associated with mutated RNA. For example, in repeat expansion diseases, such as familial amyotrophic lateral sclerosis, Huntington disease, and spinocerebellar ataxias, triplets or hexanucleotides with a high GC content in the transcribed portion of a gene are expanded. The mutated transcript, through multivalent base-pairing forms RNA foci by phase separation that have been proposed to disrupt cellular homeostasis by sequestering various RNA binding proteins.[398−400] These aberrant nuclear structures have recently been shown to be effectively disrupted by complementary ASOs.[401]

In the last few decades our knowledge about ncRNAs has dramatically increased, and their role in many pathologies has been unveiled. Novel therapeutic approaches based on the use of ASOs could exploit the targeting of ncRNAs, which are virtually undruggable by small molecule inhibitors. Anti-miRs are ASOs complementary to the mature miRNA sequence; blocking a single miRNA could lead to transcription derepression of many different genes,[402] while targeting a common seed sequence allows a single ASO to block a family of miRNAs.[388] miRNAs have a role in many different diseases, including cancer, diabetes, infections, and cardiovascular diseases.[403] In the past years, promising anti-miRs have been designed and successfully tested *in vivo* in animal models, and some have reached the clinical trials stage. One example is Mirvirasen, an ASO targeting the liver miR-122, which shows a strong antiviral activity in chronic hepatitis C infection.[404]

ASO-based therapy may be the best inhibitory method when a target ncRNA acts in the nucleus because, differently from siRNA-driven knock down, ASOs do not require the RISC complex acting in the cytoplasm. This is the case of many lncRNAs. The Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) is a very abundant nuclear lncRNA, which is upregulated in numerous cancers. ASO-mediated MALAT1 downregulation can reduce tumor progression and metastasis formation in mouse mammary or lung carcinoma models.[405,406] Although these results are promising, special attention should be dedicated to the specificity issue and to control experiments for off target effects.[407]

A recent example of a novel species of ncRNA targeted by ASOs is the inhibition of dilncRNA and DDRNA functions (as described in sections 4.2.2 and 8.3). In a model of telomere deprotection, DDR signaling and repair at telomeres are suppressed by using ASOs complementary to t-DDRNA sequences.[181] Silencing DDR activation at the telomeres could have the potential to treat some pathologies associated with telomere dysfunctions.[408] Excitingly, ASOs complementary to dilncRNA and DDRNA transcripts generated at a specific damaged genomic locus inhibit local 53BP1 focus formation, without affecting the DDR activation in other genomic locations within the same cell[95] (as already described in section 4.2.2). These data reveal that ASOs can be a suitable tool to interfere with the DNA damage signaling and repair events in a sequence-specific manner, potentially inducing genome instability and cell death specifically only in cells bearing a particular damaged DNA sequence.

However, despite their simple design and almost 40 years of research and development efforts since the first example of an ASO-based approach was proposed,[409] only a few ASO drugs have been approved for use in clinics. The first one was Fomivirsen (Vitravene, Isis Pharmaceuticals/Novartis Ophthalmics), a 21-mer phosphorothioate oligodeoxynucleotide, which was approved by the FDA in 1998 to treat cytomegalovirus retinitis by intraocular administration. It targets the viral mRNA encoding for immediate-early (IE)-2 protein, inhibiting its translation.[410] Another ASO approved for use in clinics is Mipomersen (Kynamro, Kastle Therapeutics/Ionis Pharmaceuticals). Mipomersen is systemically delivered[411] and has been approved for the treatment of homozygous familial hyper-

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

cholesterolemia, which is characterized by high plasma concentrations of low-density lipoproteins. It is a 20-mer phosphorothioate 2′-methoxyethoxy (MOE) gapmer that induces RNase H1-mediated degradation of apoB mRNA. Eteplirsen, also known as Exondys 51 (Sarepta Therapeutics), is a 30-mer phosphorodiamidate morpholino oligomer for treatment of Duchenne muscular dystrophy.[412] It is designed to induce skipping of exon 51 of the dystrophin protein, generating a shorter mRNA that encodes for a partially active isoform.

Very recently Nusinersen (Spinraza, Ionis Pharmaceuticals) has been approved. This is a 18-mer phosphorothioate 2′-*O*-methoxyethoxy oligonucleotide with all cytidines methyl-modified at the 5′-position.[413] It is indicated for types 1, 2, and 3 spinal muscular atrophy (SMA) and acts by blocking a splice site in the SMN1 and SMN2 mRNA, causing the inclusion of exon 7.

The above-reported examples illustrate that RNA molecules are quite promising therapeutic targets and highlight how critical our continuous efforts in understanding RNA systems biology are.

11. CONCLUSIONS AND PERSPECTIVES

In the introduction to this review we provocatively proposed a comparison between smartphones, objects that have transformed our daily lives with their diverse functions, and RNA molecules, which have increased the complexity of cellular processes as we know them. The obvious caveat is that while cellular phones evolved into smartphones by acquiring novel uses and functions, the multiple activities ascribed to RNA have always been intrinsic to its nature, just awaiting to be discovered.

The very concept of functionality can be a challenge when referring to an RNA molecule: a fixed length or sequence and interacting protein partners may not be sufficient to separate what is "junk" from what is functional. Essential mechanisms can thus seem invisible even to the eye of skilled and insightful scientists. Indeed, more collaborative research together with the advancement of cutting-edge technologies will help us reveal additional functions of RNA that contribute to the complexity of human life.[414] Importantly, this knowledge will be essential to develop novel effective antisense-based therapeutic approaches.

Developments in the field of fluorescence-based detection of RNA and RNA binding proteins over the past few years are a testament to the growing interest in incorporating spatial and temporal information into the study of RNA and cell biology. These advances are bringing us closer to understanding the numerous protein-coding and noncoding functions played by RNAs in normal and disease physiology. In the future, combinations of the techniques outlined in this review are likely to further increase the amount of information that can be extracted from microscopic observations leading to a new era of fluorescent single molecule imaging in the life sciences.

In this review, we have discussed how RNA molecules contribute to protecting and repairing the genome, guide genomic rearrangements, regulate telomere homeostasis, and mediate epigenetic transcriptional silencing. The nature of transcription at damaged chromatin is becoming increasingly apparent with emerging evidence involving transcription, splicing, and RNA processing factors and with RBPs being recruited to the sites of DNA damage and being necessary for full DDR activation. How local *de novo* transcription[95] coexists with transcriptional repression of the surrounding chromatin remains unclear and requires further investigation.[74,77] The ncRNAs generated at DSBs may be responsible for the transcriptional inhibition of surrounding canonical genes by recruiting chromatin remodeler complexes, a model that is reminiscent of nascent transcripts at centromeric regions in yeast.[190]

To ensure efficient signaling and repair of DNA damage, DDR proteins must relocate to the right place at the right time, assembling at DSB sites in a coordinated manner. Although the DNA damage response is an extensively studied pathway, the precise mechanism by which a cell detects and shields DNA lesions is still under debate. An exciting hypothesis places RNA at the apical levels of the DDR cascade. Since RNA is capable of assembling and organizing a compartment in the cell by liquid phase separation,[415,416] it is conceivable that it could be key to create a colloidal structure that holds and protects the DNA break and dynamically regulates access of DNA damage signaling and repair factors.

As proposed by Thomas Kuhn,[417] in science, bursts of discoveries on a particular subject are often followed by periods of relative slow, steady progress when every key question seems to have been answered, until a totally unexpected twist occurs. When it comes to smart RNA, the feeling is that the best is yet to come: the burst we are currently experiencing is likely to become a monumental explosion.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: flavia.michelini@gmail.com.
*E-mail: fabrizio.dadda@ifom.eu.

ORCID

Nils G. Walter: 0000-0002-7301-1275
Fabrizio d'Adda di Fagagna: 0000-0002-9603-5966

Notes

The authors declare the following competing financial interest(s): F.M., S.F., F.R., and F.dA.d.F. are inventors on the following parent patent application: PCT/EP2013/059753, and related applications. F.R., J.A., C.J.W., and F.dA.d.F. are inventors on the following parent patent application: PCT/EP2016/068162, and related applications. The remaining authors declare no competing financial interests.

Biographies

Flavia Michelini was born and educated in Italy. She received her Bachelor's and Master's degrees from "Tor Vergata" University in Rome. During her Bachelor's laboratory training, she evaluated the effect of zinc ion on rotavirus infection. During her Master's laboratory training at the Italian Higher Institute of Health, she studied the cytogenetic aberrations of mitotically reactivated senescent cells. In 2010, she moved to Milan where she joined Fabrizio d'Adda di Fagagna's laboratory at IFOM (the FIRC Institute of Molecular Oncology) as a Ph.D. student, working on a novel class of small noncoding RNAs (DDRNAs), involved in the DNA damage response. In 2014, she obtained her Ph.D. in Molecular Medicine from the European School of Molecular Medicine (SEMM). As a postdoctoral fellow in the same laboratory, she uncovered that DNA double-strand breaks trigger the transcription of long noncoding RNAs (dilncRNAs), which are DDRNA precursors and are important for the site-specific signaling of an appropriate DNA damage response and DNA repair. She is currently back in Rome, looking for the next job opportunity to continue her scientific career.

Ameya Jalihal received a B. Tech degree in Biotechnology from SASTRA University, Thanjavur, India, in 2015. His experiences in the laboratories

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

of Dr. Christian Ray (University of Kansas, Lawrence) and Dr. Kavita Babu (IISER Mohali) have shaped his interest in biological dynamics. He is currently enrolled as a doctoral candidate at the University of Michigan, Ann Arbor, where he is a part of the RNA and Single Molecule Analysis group headed by Prof. Nils Walter. He uses single-molecule, live-cell fluorescence microscopy approaches to study mechanisms of eukaryotic translation regulation as part of his thesis research.

Sofia Francia received her degree in Biology from "Università Statale di Milano", Italy, in 2003. As a master student, she spent two years in the laboratory of Prof. Marco Foiani, where she shaped her interest in cancer biology and DNA damage. In late 2003, she joined the laboratory headed by Fabrizio d'Adda di Fagagna at the FIRC Institute of Molecular Oncology Foundation (IFOM) in Milan, investigating the cellular response to DNA damage in senescence and cancer. She obtained her Ph.D. in Molecular Medicine from the European School of Molecular Medicine, in Milan, in 2009, and after that she continued as a postdoctoral fellow at IFOM. During these years, she discovered that noncoding RNAs play an unexpected function in the cellular response to DNA damage, a study published in *Nature*. In 2012, she spent a couple of months with a short term EMBO fellowship in the laboratory of Prof. Jiri Bartek at the Institute of Molecular and Translational Medicine Faculty of Medicine and Dentistry, Palacky University Olomouc, Czech Republic. Since 2015, she has been a researcher of the Istituto di Genetica Molecolare-Consiglio Nazionale delle Ricerche (IGM-CNR) in Pavia, heading a small research team. She obtained a permanent position in 2017.

Zach Neeb was born and raised in California, USA. He obtained both his B.S. and Ph.D. degrees at the University of California Santa Cruz in the Program for Biomedical Sciences and Engineering, with an emphasis on Molecular, Cell and Developmental Biology. Zach studied sRNA biogenesis and function in the ciliate *Oxytricha trifallax* during his Ph.D. in Professor Alan M. Zahler's lab, a member of The Center for Molecular Biology of RNA at UC Santa Cruz. Zach is currently a Postdoctoral Fellow studying macronuclear development in the ciliate *Paramecium tetraurelia* in Professor Mariusz Nowacki's lab, a member of the Institute of Cell Biology at the University of Bern in Switzerland.

Francesca Rossiello is a Postdoctoral fellow at the FIRC Institute of Molecular Oncology (IFOM), Milan, in DNA damage response and cellular senescence lab led by Dr. Fabrizio d'Adda di Fagagna. She obtained a Master's degree in Medical Biotechnologies and Molecular Medicine at University of Bari in 2008 and a Ph.D. in Molecular Medicine at the European School of Molecular Medicine in Milan in 2014.

Ubaldo Gioia obtained his Ph.D. in Genetics and Molecular Biology from "Sapienza" University of Rome in 2008 for his work on the biogenesis and functions of microRNAs in neuronal differentiation, in the laboratory of Dr. Elisa Caffarelli at the Institute of Molecular Biology and Pathology at National Research Council in Rome. He then worked as a postdoctoral fellow in Prof. Irene Bozzoni's group at "Sapienza" University of Rome, on the role of long noncoding RNAs in neuronal pathologies. In 2013, he joined Fabrizio d'Adda di Fagagna's laboratory at IFOM (the FIRC Institute of Molecular Oncology) in Milan, to study the functions of noncoding RNAs in the DNA damage response and repair and in neurodegeneration.

Julio Aguado obtained his B.S. in Biochemistry and Molecular Biology from the University of Navarra, Pamplona (Spain), in 2012. He was, then, awarded a Marie Curie fellowship from the Seventh Framework Programme of the European Union to pursue his Ph.D., which is currently taking place at The FIRC Institute of Molecular Oncology, Milan (Italy), under the supervision of Dr. Fabrizio d'Adda di Fagagna. His research is focused on the role of telomeric RNA at dysfunctional telomeres and its impact on senescence and aging.

Corey Jones-Weinert received his Bachelor's of Science in Molecular, Cellular, and Developmental biology from the University of Washington in 2013. There, he began his research career studying the development of zebrafish. Soon after, he moved to the Fred Hutchinson Cancer Research Center where he worked to clarify the role of phospho-signaling during post replication repair and to identify novel DNA damage-induced phosphorylation events. In 2014, he moved to Milan, Italy, where he now works as a Ph.D. student in the group of Fabrizio d'Adda di Fagagna at the FIRC Institute of Molecular Oncology (IFOM). Corey's current research is focused on deciphering the role of DNA damage response RNAs (DDRNAs) and damage induced long noncoding RNAs (dilncRNAs) at the telomere.

Mariusz Nowacki received his Ph.D. from University of Paris VI and Ecole Normale Supérieure in 2005. He did a postdoc at Princeton University from 2005, and since 2010 he has worked as a professor of Genetics at University of Bern in Switzerland. His research focuses on the mechanisms of RNA-mediated epigenetic inheritance and RNA-guided genome editing in ciliates.

Francesca Storici received her Ph.D. in Molecular Genetics from the International School of Advanced Studies in Trieste, Italy (1998). She was a postdoctoral fellow at the National Institute of Environmental and Health Sciences (NIEHS, NIH) in North Carolina until 2007 and then a research assistant professor at the Gene Therapy Center of the University of North Carolina at Chapel Hill, NC. She joined the Georgia Institute of Technology as an assistant professor in August 2007 and became Distinguished Cancer Scientist of the Georgia Research Alliance. In 2013, she was promoted to Associate Professor with tenure. In 2016, she became Howard Hughes Medical Institute Faculty Scholar. Just recently, she was promoted to Full Professor. Her research is on DNA damage, repair, and gene editing.

Piero Carninci was born and educated in Italy. He moved to Japan in 1995 at RIKEN and became tenure researcher in 1997. He has developed technologies to analyze the transcribed part of the genome (transcriptome), such as the cap-trapper and the CAGE. These technologies have been broadly used in the RIKEN FANTOM projects and allowed identifying noncoding RNAs as the major output of the mammalian genome and providing comprehensive maps of the mammalian promoters. From April in 2013, he has been the Director of the Division of Genomic Technologies and the Deputy Director of RIKEN Center for Life Science Technologies. He has published more than 320 papers and book chapters, edited books, and is a member of the editorial boards of various scientific journals.

Nils G. Walter was born in 1966 in Frankfurt am Main, Germany. He received his "Vordiplom" (B.S.) and "Diploma" (Masters) from the Technical University of Darmstadt after performing research with Hans-Günther Gassen on the physiochemical characterization of a protein dehyrogenase enzyme. He earned his Dr. Ing. while studying molecular in vitro evolution of DNA and RNA using fluorescence techniques with Nobel laureate Manfred Eigen at the Max-Planck-Institute for Biophysical Chemistry, Göttingen. For his postdoctoral studies, he turned to RNA enzymes under the guidance of John M. Burke at the University of Vermont in Burlington, Vermont. He is currently the Francis S. Collins Collegiate Professor of Chemistry, Biophysics, and Biological Chemistry in the College of Literature, Science and the Arts of the University of Michigan in Ann Arbor, Michigan. His research interests focus on noncoding RNA through the lens of single molecule fluorescence techniques. He founded and currently directs the Single Molecule Analysis in Real-Time (SMART) Center, as well as cofounded and currently codirects the Center for RNA Biomedicine at Michigan.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

Fabrizio d'Adda di Fagagna is a cell and molecular biologist that studies the involvement of the DNA damage response (DDR) pathways in physiologically relevant processes such as aging and cancer. Fabrizio is very patient with his postdocs. Fabrizio obtained his Ph.D. in Molecular Genetics at the International School for Advanced Studies (SISSA) working at the International Centre for Genetic Engineering (ICGEB) in Trieste (Italy) under the supervision of Mauro Giacca and Arturo Falaschi on the transcriptional regulation of HIV-1. He then moved to Cambridge (UK) to work in the group of Steve Jackson at the now Gurdon Institute. Here, most notably, he demonstrated that replicative cellular senescence is the outcome of DDR activation caused by the direct recognition of critically short telomeres. Fabrizio set up his own research group at IFOM (FIRC Institute of Molecular Oncology) in Milan (Italy) in 2003. Here, he demonstrated that oncogene activation is an intrinsically genotoxic event that causes DDR activation and cellular senescence establishment. More recently, he proposed a unifying model for cellular senescence establishment based on persistent DNA damage at telomeres. Fabrizio's most recent exciting finding is the discovery of an unanticipated role of noncoding RNAs in the direct activation of the DDR. This discovery fuels most of his present investigative efforts also with the aim of exploring the potential translation of these findings into interventions useful in cancer and aging-related diseases. In 2014, he was awarded a permanent position "for exceptional merits" at the Italian National Research Council (CNR) in Pavia, where he runs a laboratory. Dr. Fabrizio d'Adda di Fagagna is an EMBO member and received several awards including the European Association for Cancer Research (EACR) Young Cancer Researcher Award and the EMBO Young Investigator Award. He is a recipient of an ERC advanced grant.

ACKNOWLEDGMENTS

We are thankful to Cinzia Villa for editing figures. F.d'A.d.F.'s lab was supported by the Associazione Italiana per la Ricerca sul Cancro, AIRC (application 12971), Human Frontier Science Program (contract RGP 0014/2012), Cariplo Foundation (grant 2010.0818 and 2014-0812), Marie Curie Initial Training Networks (FP7 PEOPLE 2012 ITN (CodAge)), Fondazione Telethon (GGP12059), Association for International Cancer Research (AICR-Worldwide Cancer Research Rif. N. 14-1331), Progetti di Ricerca di Interesse Nazionale (PRIN) 2010−2011, the Italian Ministry of Education Universities and Research EPIGEN Project, an European Research Council advanced grant (322726). S.F. was supported by Collegio Ghislieri and Fondazione Cariplo (Grant rif. 2014-1215). F.d'A.d.F. and S.F. were supported by AriSLA (project DDRNA and ALS). N.G.W. is supported by National Institutes of Health (NIH) R01 grants GM062357, GM118524, and GM122803, as well as a University of Michigan Comprehensive Cancer Center/Biointerfaces Institute Research Grant. A.J. is supported by the NIH Cellular and Molecular Biology Training Grant T32-GM007315. M.N. is supported by European Research Council Grants 260358 "EPIGENOME" and 681178 "G-EDIT", Swiss National Science Foundation Grants 31003A_146257 and 31003A_166407, and National Center of Competence in Research RNA and Disease. F.S. is supported by the National Institute of General Medical Sciences, NIGMS, of the NIH, grant number GM115927, the National Science Foundation fund with grant number MCB-1615335, and the Howard Hughes Medical Institute Faculty Scholar grant 55108574. C.M. is supported by the GAANN fellowship P200A150130-16. G.B. is supported by AIRC grant number 15273. B.L.'s lab is supported by CancerTelSys (01ZX1302) in the E:med program of the German Federal Ministry of Education and Research (BMBF) and the DFG Heisenberg Program (LU1709/2).

REFERENCES

(1) Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M. C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; et al. The Transcriptional Landscape of the Mammalian Genome. *Science* **2005**, *309*, 1559−1563.
(2) Esteller, M. Non-Coding RNAs in Human Disease. *Nat. Rev. Genet.* **2011**, *12*, 861−874.
(3) Ba, Z.; Qi, Y. Small RNAs: Emerging Key Players in DNA Double-Strand Break Repair. *Sci. China: Life Sci.* **2013**, *56*, 933−936.
(4) Sharma, V.; Misteli, T. Non-Coding RNAs in DNA Damage and Repair. *FEBS Lett.* **2013**, *587*, 1832−1839.
(5) d'Adda di Fagagna, F. A Direct Role for Small Non-Coding RNAs in DNA Damage Response. *Trends Cell Biol.* **2014**, *24*, 171−178.
(6) Chowdhury, D.; Choi, Y. E.; Brault, M. E. Charity Begins at Home: Non-Coding RNA Functions in DNA Repair. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 181−189.
(7) Wickramasinghe, V. O.; Venkitaraman, A. R. RNA Processing and Genome Stability: Cause and Consequence. *Mol. Cell* **2016**, *61*, 496−505.
(8) Nowacki, M.; Shetty, K.; Landweber, L. F. RNA-Mediated Epigenetic Programming of Genome Rearrangements. *Annu. Rev. Genomics Hum. Genet.* **2011**, *12*, 367−389.
(9) Yerlici, V. T.; Landweber, L. F. Programmed Genome Rearrangements in the Ciliate Oxytricha. *Microbiol. Spectrum* **2014**, *2*, 10.1128/microbiolspec.MDNA3-0025-2014
(10) Azzalin, C. M.; Lingner, J. Telomere Functions Grounding on Terra Firma. *Trends Cell Biol.* **2015**, *25*, 29−36.
(11) Cusanelli, E.; Chartrand, P. Telomeric Repeat-Containing RNA Terra: A Noncoding RNA Connecting Telomere Biology to Genome Integrity. *Front. Genet.* **2015**, *6*, 143.
(12) Rippe, K.; Luke, B. Terra and the State of the Telomere. *Nat. Struct. Mol. Biol.* **2015**, *22*, 853−858.
(13) Beadle, G. W.; Tatum, E. L. Genetic Control of Biochemical Reactions in Neurospora. *Proc. Natl. Acad. Sci. U. S. A.* **1941**, *27*, 499−506.
(14) Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; et al. Initial Sequencing and Analysis of the Human Genome. *Nature* **2001**, *409*, 860−921.
(15) 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.; et al. Initial Sequencing and Comparative Analysis of the Mouse Genome. *Nature* **2002**, *420*, 520−562.
(16) Liang, F.; Holt, I.; Pertea, G.; Karamycheva, S.; Salzberg, S. L.; Quackenbush, J. Gene Index Analysis of the Human Genome Estimates Approximately 120,000 Genes. *Nat. Genet.* **2000**, *25*, 239−240.
(17) C. elegans Sequencing Consortium. Genome Sequence of the Nematode C. Elegans: A Platform for Investigating Biology. *Science* **1998**, *282*, 2012−2018.
(18) Adams, M. D.; Celniker, S. E.; Holt, R. A.; Evans, C. A.; Gocayne, J. D.; Amanatides, P. G.; Scherer, S. E.; Li, P. W.; Hoskins, R. A.; Galle, R. F.; et al. The Genome Sequence of Drosophila Melanogaster. *Science* **2000**, *287*, 2185−2195.
(19) Mattick, J. S. RNA Regulation: A New Genetics? *Nat. Rev. Genet.* **2004**, *5*, 316−323.
(20) Okazaki, Y.; Furuno, M.; Kasukawa, T.; Adachi, J.; Bono, H.; Kondo, S.; Nikaido, I.; Osato, N.; Saito, R.; Suzuki, H.; et al. Analysis of the Mouse Transcriptome Based on Functional Annotation of 60,770 Full-Length Cdnas. *Nature* **2002**, *420*, 563−573.
(21) Imanishi, T.; Itoh, T.; Suzuki, Y.; O'Donovan, C.; Fukuchi, S.; Koyanagi, K. O.; Barrero, R. A.; Tamura, T.; Yamaguchi-Kabata, Y.; Tanino, M.; et al. Integrative Annotation of 21,037 Human Genes Validated by Full-Length Cdna Clones. *PLoS Biol.* **2004**, *2*, e162.
(22) Gerhard, D. S.; Wagner, L.; Feingold, E. A.; Shenmen, C. M.; Grouse, L. H.; Schuler, G.; Klein, S. L.; Old, S.; Rasooly, R.; Good, P.;

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

et al. The Status, Quality, and Expansion of the Nih Full-Length Cdna Project: The Mammalian Gene Collection (Mgc). *Genome Res.* **2004**, *14*, 2121−2127.
(23) Kapranov, P.; Cawley, S. E.; Drenkow, J.; Bekiranov, S.; Strausberg, R. L.; Fodor, S. P.; Gingeras, T. R. Large-Scale Transcriptional Activity in Chromosomes 21 and 22. *Science* **2002**, *296*, 916−919.
(24) Cheng, J.; Kapranov, P.; Drenkow, J.; Dike, S.; Brubaker, S.; Patel, S.; Long, J.; Stern, D.; Tammana, H.; Helt, G.; et al. Transcriptional Maps of 10 Human Chromosomes at 5-Nucleotide Resolution. *Science* **2005**, *308*, 1149−1154.
(25) Shiraki, T.; Kondo, S.; Katayama, S.; Waki, K.; Kasukawa, T.; Kawaji, H.; Kodzius, R.; Watahiki, A.; Nakamura, M.; Arakawa, T.; et al. Cap Analysis Gene Expression for High-Throughput Analysis of Transcriptional Starting Point and Identification of Promoter Usage. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 15776−15781.
(26) Kodzius, R.; Kojima, M.; Nishiyori, H.; Nakamura, M.; Fukuda, S.; Tagami, M.; Sasaki, D.; Imamura, K.; Kai, C.; Harbers, M.; et al. Cage: Cap Analysis of Gene Expression. *Nat. Methods* **2006**, *3*, 211−222.
(27) Katayama, S.; Tomaru, Y.; Kasukawa, T.; Waki, K.; Nakanishi, M.; Nakamura, M.; Nishida, H.; Yap, C. C.; Suzuki, M.; Kawai, J.; et al. Antisense Transcription in the Mammalian Transcriptome. *Science* **2005**, *309*, 1564−1566.
(28) Faulkner, G. J.; Kimura, Y.; Daub, C. O.; Wani, S.; Plessy, C.; Irvine, K. M.; Schroder, K.; Cloonan, N.; Steptoe, A. L.; Lassmann, T.; et al. The Regulated Retrotransposon Transcriptome of Mammalian Cells. *Nat. Genet.* **2009**, *41*, 563−571.
(29) Guttman, M.; Amit, I.; Garber, M.; French, C.; Lin, M. F.; Feldser, D.; Huarte, M.; Zuk, O.; Carey, B. W.; Cassady, J. P.; et al. Chromatin Signature Reveals over a Thousand Highly Conserved Large Non-Coding RNAs in Mammals. *Nature* **2009**, *458*, 223−227.
(30) Djebali, S.; Davis, C. A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; et al. Landscape of Transcription in Human Cells. *Nature* **2012**, *489*, 101−108.
(31) The ENCODE Project Consortium. An Integrated Encyclopedia of DNA Elements in the Human Genome. *Nature* **2012**, *489*, 57−74.
(32) Clark, M. B.; Amaral, P. P.; Schlesinger, F. J.; Dinger, M. E.; Taft, R. J.; Rinn, J. L.; Ponting, C. P.; Stadler, P. F.; Morris, K. V.; Morillon, A.; et al. The Reality of Pervasive Transcription. *PLoS Biol.* **2011**, *9*, e1000625 discussion e1001102.
(33) Napoli, C.; Lemieux, C.; Jorgensen, R. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in Trans. *Plant Cell* **1990**, *2*, 279−289.
(34) Romano, N.; Macino, G. Quelling: Transient Inactivation of Gene Expression in Neurospora Crassa by Transformation with Homologous Sequences. *Mol. Microbiol.* **1992**, *6*, 3343−3353.
(35) Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. Potent and Specific Genetic Interference by Double-Stranded RNA in Caenorhabditis Elegans. *Nature* **1998**, *391*, 806−811.
(36) Kennerdell, J. R.; Carthew, R. W. Use of Dsrna-Mediated Genetic Interference to Demonstrate That Frizzled and Frizzled 2 Act in the Wingless Pathway. *Cell* **1998**, *95*, 1017−1026.
(37) Ngo, H.; Tschudi, C.; Gull, K.; Ullu, E. Double-Stranded RNA Induces Mrna Degradation in Trypanosoma Brucei. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 14687−14692.
(38) Wianny, F.; Zernicka-Goetz, M. Specific Interference with Gene Function by Double-Stranded RNA in Early Mouse Development. *Nat. Cell Biol.* **2000**, *2*, 70−75.
(39) Yang, S.; Tutton, S.; Pierce, E.; Yoon, K. Specific Double-Stranded RNA Interference in Undifferentiated Mouse Embryonic Stem Cells. *Mol. Cell. Biol.* **2001**, *21*, 7807−7816.
(40) Hamilton, A. J.; Baulcombe, D. C. A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants. *Science* **1999**, *286*, 950−952.
(41) Zamore, P. D.; Tuschl, T.; Sharp, P. A.; Bartel, D. P. RNAi: Double-Stranded RNA Directs the Atp-Dependent Cleavage of Mrna at 21 to 23 Nucleotide Intervals. *Cell* **2000**, *101*, 25−33.
(42) Elbashir, S. M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T. Duplexes of 21-Nucleotide RNAs Mediate RNA Interference in Cultured Mammalian Cells. *Nature* **2001**, *411*, 494−498.
(43) Elbashir, S. M.; Lendeckel, W.; Tuschl, T. RNA Interference Is Mediated by 21- and 22-Nucleotide RNAs. *Genes Dev.* **2001**, *15*, 188−200.
(44) Borsani, O.; Zhu, J.; Verslues, P. E.; Sunkar, R.; Zhu, J. K. Endogenous Sirnas Derived from a Pair of Natural Cis-Antisense Transcripts Regulate Salt Tolerance in Arabidopsis. *Cell* **2005**, *123*, 1279−1291.
(45) Buker, S. M.; Iida, T.; Buhler, M.; Villen, J.; Gygi, S. P.; Nakayama, J.; Moazed, D. Two Different Argonaute Complexes Are Required for Sirna Generation and Heterochromatin Assembly in Fission Yeast. *Nat. Struct. Mol. Biol.* **2007**, *14*, 200−207.
(46) Meister, G. Argonaute Proteins: Functional Insights and Emerging Roles. *Nat. Rev. Genet.* **2013**, *14*, 447−459.
(47) Dueck, A.; Meister, G. Assembly and Function of Small RNA - Argonaute Protein Complexes. *Biol. Chem.* **2014**, *395*, 611−629.
(48) Lagos-Quintana, M.; Rauhut, R.; Lendeckel, W.; Tuschl, T. Identification of Novel Genes Coding for Small Expressed RNAs. *Science* **2001**, *294*, 853−858.
(49) Lau, N. C.; Lim, L. P.; Weinstein, E. G.; Bartel, D. P. An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis Elegans. *Science* **2001**, *294*, 858−862.
(50) Lee, R. C.; Ambros, V. An Extensive Class of Small RNAs in Caenorhabditis Elegans. *Science* **2001**, *294*, 862−864.
(51) Lee, R. C.; Feinbaum, R. L.; Ambros, V. The C. Elegans Heterochronic Gene Lin-4 Encodes Small RNAs with Antisense Complementarity to Lin-14. *Cell* **1993**, *75*, 843−854.
(52) Sayed, D.; Abdellatif, M. Micrornas in Development and Disease. *Physiol. Rev.* **2011**, *91*, 827−887.
(53) Ha, M.; Kim, V. N. Regulation of Microrna Biogenesis. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 509−524.
(54) Berezikov, E.; Chung, W. J.; Willis, J.; Cuppen, E.; Lai, E. C. Mammalian Mirtron Genes. *Mol. Cell* **2007**, *28*, 328−336.
(55) Heyam, A.; Lagos, D.; Plevin, M. Dissecting the Roles of Trbp and Pact in Double-Stranded RNA Recognition and Processing of Noncoding RNAs. *Wiley Interdiscip Rev. RNA* **2015**, *6*, 271−289.
(56) Krol, J.; Loedige, I.; Filipowicz, W. The Widespread Regulation of Microrna Biogenesis, Function and Decay. *Nat. Rev. Genet.* **2010**, *11*, 597−610.
(57) Pfaff, J.; Meister, G. Argonaute and Gw182 Proteins: An Effective Alliance in Gene Silencing. *Biochem. Soc. Trans.* **2013**, *41*, 855−860.
(58) Castel, S. E.; Martienssen, R. A. RNA Interference in the Nucleus: Roles for Small RNAs in Transcription, Epigenetics and Beyond. *Nat. Rev. Genet.* **2013**, *14*, 100−112.
(59) Calses, P. C.; Dhillon, K. K.; Tucker, N.; Chi, Y.; Huang, J. W.; Kawasumi, M.; Nghiem, P.; Wang, Y.; Clurman, B. E.; Jacquemont, C.; et al. Dgcr8Mediates Repair of Uv-Induced DNA Damage Independently of RNA Processing. *Cell Rep.* **2017**, *19*, 162−174.
(60) Chitale, S.; Richly, H. Dicer and Zrf1 Contribute to Chromatin Decondensation During Nucleotide Excision Repair. *Nucleic Acids Res.* **2017**, *45*, 5901.
(61) Francia, S.; Michelini, F.; Saxena, A.; Tang, D.; de Hoon, M.; Anelli, V.; Mione, M.; Carninci, P.; d'Adda di Fagagna, F. Site-Specific Dicer and Drosha RNA Products Control the DNA-Damage Response. *Nature* **2012**, *488*, 231−235.
(62) Han, J.; Pedersen, J. S.; Kwon, S. C.; Belair, C. D.; Kim, Y. K.; Yeom, K. H.; Yang, W. Y.; Haussler, D.; Blelloch, R.; Kim, V. N. Posttranscriptional Crossregulation between Drosha and Dgcr8. *Cell* **2009**, *136*, 75−84.
(63) Johanson, T. M.; Keown, A. A.; Cmero, M.; Yeo, J. H.; Kumar, A.; Lew, A. M.; Zhan, Y.; Chong, M. M. Drosha Controls Dendritic Cell Development by Cleaving Messenger RNAs Encoding Inhibitors of Myelopoiesis. *Nat. Immunol.* **2015**, *16*, 1134−1141.
(64) Knuckles, P.; Vogt, M. A.; Lugert, S.; Milo, M.; Chong, M. M.; Hautbergue, G. M.; Wilson, S. A.; Littman, D. R.; Taylor, V. Drosha Regulates Neurogenesis by Controlling Neurogenin 2 Expression Independent of Micrornas. *Nat. Neurosci.* **2012**, *15*, 962−969.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

(65) Bentley, D. L. Coupling Mrna Processing with Transcription in Time and Space. *Nat. Rev. Genet.* **2014**, *15*, 163−175.
(66) Morlando, M.; Ballarino, M.; Gromak, N.; Pagano, F.; Bozzoni, I.; Proudfoot, N. J. Primary Microrna Transcripts Are Processed Co-Transcriptionally. *Nat. Struct. Mol. Biol.* **2008**, *15*, 902−909.
(67) Gromak, N.; Dienstbier, M.; Macias, S.; Plass, M.; Eyras, E.; Caceres, J. F.; Proudfoot, N. J. Drosha Regulates Gene Expression Independently of RNA Cleavage Function. *Cell Rep.* **2013**, *5*, 1499−1510.
(68) Macias, S.; Plass, M.; Stajuda, A.; Michlewski, G.; Eyras, E.; Caceres, J. F. Dgcr8 Hits-Clip Reveals Novel Functions for the Microprocessor. *Nat. Struct. Mol. Biol.* **2012**, *19*, 760−766.
(69) Suzuki, H. I.; Young, R. A.; Sharp, P. A. Super-Enhancer-Mediated RNA Processing Revealed by Integrative Microrna Network Analysis. *Cell* **2017**, *168*, 1000−1014 e1015.
(70) Wagschal, A.; Rousset, E.; Basavarajaiah, P.; Contreras, X.; Harwig, A.; Laurent-Chabalier, S.; Nakamura, M.; Chen, X.; Zhang, K.; Meziane, O.; et al. Microprocessor, Setx, Xrn2, and Rrp6 Co-Operate to Induce Premature Termination of Transcription by RNApii. *Cell* **2012**, *150*, 1147−1157.
(71) Dhir, A.; Dhir, S.; Proudfoot, N. J.; Jopling, C. L. Microprocessor Mediates Transcriptional Termination of Long Noncoding RNA Transcripts Hosting Microrna. *Nat. Struct. Mol. Biol.* **2015**, *22*, 319−327.
(72) Burger, K.; Gullerova, M. Swiss Army Knives: Non-Canonical Functions of Nuclear Drosha and Dicer. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 417−430.
(73) D'Alessandro, G.; d'Adda di Fagagna, F. Transcription and DNA Damage: Holding Hands or Crossing Swords? *J. Mol. Biol.* **2016**, *429*, 3215.
(74) Capozzo, I.; Iannelli, F.; Francia, S.; d'Adda di Fagagna, F. Express or Repress? The Transcriptional Dilemma of Damaged Chromatin. *FEBS J.* **2017**, *284*, 2133.
(75) Francia, S. Non-Coding RNA: Sequence-Specific Guide for Chromatin Modification and DNA Damage Signaling. *Front. Genet.* **2015**, *6*, 320.
(76) Santos-Pereira, J. M.; Aguilera, A. R Loops: New Modulators of Genome Dynamics and Function. *Nat. Rev. Genet.* **2015**, *16*, 583−597.
(77) Marnef, A.; Cohen, S.; Legube, G. Transcription-Coupled DNA Double-Strand Break Repair: Active Genes Need Special Care. *J. Mol. Biol.* **2017**, *429*, 1277−1288.
(78) Castel, S. E.; Ren, J.; Bhattacharjee, S.; Chang, A. Y.; Sanchez, M.; Valbuena, A.; Antequera, F.; Martienssen, R. A. Dicer Promotes Transcription Termination at Sites of Replication Stress to Maintain Genome Stability. *Cell* **2014**, *159*, 572−583.
(79) Much, C.; Auchynnikava, T.; Pavlinic, D.; Buness, A.; Rappsilber, J.; Benes, V.; Allshire, R.; O'Carroll, D. Endogenous Mouse Dicer Is an Exclusively Cytoplasmic Protein. *PLoS Genet.* **2016**, *12*, e1006095.
(80) Cheng, W.; Qi, Y.; Tian, L.; Wang, B.; Huang, W.; Chen, Y. Dicer Promotes Tumorigenesis by Translocating to Nucleus to Promote Sfrp1 Promoter Methylation in Cholangiocarcinoma Cells. *Cell Death Dis.* **2017**, *8*, e2628.
(81) Doyle, M.; Badertscher, L.; Jaskiewicz, L.; Guttinger, S.; Jurado, S.; Hugenschmidt, T.; Kutay, U.; Filipowicz, W. The Double-Stranded RNA Binding Domain of Human Dicer Functions as a Nuclear Localization Signal. *RNA* **2013**, *19*, 1238−1252.
(82) Gagnon, K. T.; Li, L.; Chu, Y.; Janowski, B. A.; Corey, D. R. RNAi Factors Are Present and Active in Human Cell Nuclei. *Cell Rep.* **2014**, *6*, 211−221.
(83) Matsui, M.; Li, L.; Janowski, B. A.; Corey, D. R. Reduced Expression of Argonaute 1, Argonaute 2, and Trbp Changes Levels and Intracellular Distribution of RNAi Factors. *Sci. Rep.* **2015**, *5*, 12855.
(84) Neve, J.; Burger, K.; Li, W.; Hoque, M.; Patel, R.; Tian, B.; Gullerova, M.; Furger, A. Subcellular RNA Profiling Links Splicing and Nuclear Dicer1 to Alternative Cleavage and Polyadenylation. *Genome Res.* **2016**, *26*, 24−35.
(85) Skourti-Stathaki, K.; Kamieniarz-Gdula, K.; Proudfoot, N. J. R-Loops Induce Repressive Chromatin Marks over Mammalian Gene Terminators. *Nature* **2014**, *516*, 436−439.
(86) White, E.; Schlackow, M.; Kamieniarz-Gdula, K.; Proudfoot, N. J.; Gullerova, M. Human Nuclear Dicer Restricts the Deleterious Accumulation of Endogenous Double-Stranded RNA. *Nat. Struct. Mol. Biol.* **2014**, *21*, 552−559.
(87) Burger, K.; Schlackow, M.; Potts, M.; Hester, S.; Mohammed, S.; Gullerova, M. Nuclear Phosphorylated Dicer Processes Double-Stranded RNA in Response to DNA Damage. *J. Cell Biol.* **2017**, *216*, 2373.
(88) Tubbs, A.; Nussenzweig, A. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell* **2017**, *168*, 644−656.
(89) Thomson, T.; Lin, H. The Biogenesis and Function of Piwi Proteins and Pirnas: Progress and Prospect. *Annu. Rev. Cell Dev. Biol.* **2009**, *25*, 355−376.
(90) Peaston, A. E.; Evsikov, A. V.; Graber, J. H.; de Vries, W. N.; Holbrook, A. E.; Solter, D.; Knowles, B. B. Retrotransposons Regulate Host Genes in Mouse Oocytes and Preimplantation Embryos. *Dev. Cell* **2004**, *7*, 597−606.
(91) Fort, A.; Hashimoto, K.; Yamada, D.; Salimullah, M.; Keya, C. A.; Saxena, A.; Bonetti, A.; Voineagu, I.; Bertin, N.; Kratz, A.; et al. Deep Transcriptome Profiling of Mammalian Stem Cells Supports a Regulatory Role for Retrotransposons in Pluripotency Maintenance. *Nat. Genet.* **2014**, *46*, 558−566.
(92) Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project. Post-Transcriptional Processing Generates a Diversity of 5′-Modified Long and Short RNAs. *Nature* **2009**, *457*, 1028−1032.
(93) Chen, Y.; Pai, A. A.; Herudek, J.; Lubas, M.; Meola, N.; Jarvelin, A. I.; Andersson, R.; Pelechano, V.; Steinmetz, L. M.; Jensen, T. H.; et al. Principles for RNA Metabolism and Alternative Transcription Initiation within Closely Spaced Promoters. *Nat. Genet.* **2016**, *48*, 984−994.
(94) Wei, W.; Ba, Z.; Gao, M.; Wu, Y.; Ma, Y.; Amiard, S.; White, C. I.; Rendtlew Danielsen, J. M.; Yang, Y. G.; Qi, Y. A Role for Small RNAs in DNA Double-Strand Break Repair. *Cell* **2012**, *149*, 101−112.
(95) Michelini, F.; Pitchiaya, S.; Vitelli, V.; Sharma, S.; Gioia, U.; Pessina, F.; Cabrini, M.; Wang, Y.; Capozzo, I.; Iannelli, 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*, 1400−1411.
(96) Andersson, R.; Gebhard, C.; Miguel-Escalada, I.; Hoof, I.; Bornholdt, J.; Boyd, M.; Chen, Y.; Zhao, X.; Schmidl, C.; Suzuki, T.; et al. An Atlas of Active Enhancers across Human Cell Types and Tissues. *Nature* **2014**, *507*, 455−461.
(97) Li, W.; Notani, D.; Rosenfeld, M. G. Enhancers as Non-Coding RNA Transcription Units: Recent Insights and Future Perspectives. *Nat. Rev. Genet.* **2016**, *17*, 207−223.
(98) Hon, C. C.; Ramilowski, J. A.; Harshbarger, J.; Bertin, N.; Rackham, O. J.; Gough, J.; Denisenko, E.; Schmeier, S.; Poulsen, T. M.; Severin, J.; et al. An Atlas of Human Long Non-Coding RNAs with Accurate 5′ Ends. *Nature* **2017**, *543*, 199−204.
(99) Werner, M. S.; Sullivan, M. A.; Shah, R. N.; Nadadur, R. D.; Grzybowski, A. T.; Galat, V.; Moskowitz, I. P.; Ruthenburg, A. J. Chromatin-Enriched lncRNAs Can Act as Cell-Type Specific Activators of Proximal Gene Transcription. *Nat. Struct. Mol. Biol.* **2017**, *24*, 596−603.
(100) Sasaki, Y. T.; Ideue, T.; Sano, M.; Mituyama, T.; Hirose, T. Menepsilon/Beta Noncoding RNAs Are Essential for Structural Integrity of Nuclear Paraspeckles. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 2525−2530.
(101) Rinn, J. L.; Kertesz, M.; Wang, J. K.; Squazzo, S. L.; Xu, X.; Brugmann, S. A.; Goodnough, L. H.; Helms, J. A.; Farnham, P. J.; Segal, E.; et al. Functional Demarcation of Active and Silent Chromatin Domains in Human Hox Loci by Noncoding RNAs. *Cell* **2007**, *129*, 1311−1323.
(102) Cerase, A.; Pintacuda, G.; Tattermusch, A.; Avner, P. Xist Localization and Function: New Insights from Multiple Levels. *Genome Biol.* **2015**, *16*, 166.
(103) Carninci, P. Tagging Mammalian Transcription Complexity. *Trends Genet.* **2006**, *22*, 501−510.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

(104) Carrieri, C.; Cimatti, L.; Biagioli, M.; Beugnet, A.; Zucchelli, S.; Fedele, S.; Pesce, E.; Ferrer, I.; Collavin, L.; Santoro, C.; et al. Long Non-Coding Antisense RNA Controls Uchl1 Translation through an Embedded Sineb2 Repeat. *Nature* **2012**, *491*, 454−457.
(105) Zucchelli, S.; Cotella, D.; Takahashi, H.; Carrieri, C.; Cimatti, L.; Fasolo, F.; Jones, M. H.; Sblattero, D.; Sanges, R.; Santoro, C.; et al. Sineups: A New Class of Natural and Synthetic Antisense Long Non-Coding RNAs That Activate Translation. *RNA Biol.* **2015**, *12*, 771−779.
(106) Schein, A.; Zucchelli, S.; Kauppinen, S.; Gustincich, S.; Carninci, P. Identification of Antisense Long Noncoding RNAs That Function as Sineups in Human Cells. *Sci. Rep.* **2016**, *6*, 33605.
(107) Ciccia, A.; Elledge, S. J. The DNA Damage Response: Making It Safe to Play with Knives. *Mol. Cell* **2010**, *40*, 179−204.
(108) Jackson, S. P.; Durocher, D. Regulation of DNA Damage Responses by Ubiquitin and Sumo. *Mol. Cell* **2013**, *49*, 795−807.
(109) Polo, S. E.; Jackson, S. P. Dynamics of DNA Damage Response Proteins at DNA Breaks: A Focus on Protein Modifications. *Genes Dev.* **2011**, *25*, 409−433.
(110) Uckelmann, M.; Sixma, T. K. Histone Ubiquitination in the DNA Damage Response. *DNA Repair* **2017**, *56*, 92−101.
(111) Campisi, J.; d'Adda di Fagagna, F. Cellular Senescence: When Bad Things Happen to Good Cells. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 729−740.
(112) Nishida, K.; Kuwano, Y.; Nishikawa, T.; Masuda, K.; Rokutan, K. RNA Binding Proteins and Genome Integrity. *Int. J. Mol. Sci.* **2017**, *18*, 134110.3390/ijms18071341
(113) Dutertre, M.; Vagner, S. DNA-Damage Response RNA-Binding Proteins (Ddrbps): Perspectives from a New Class of Proteins and Their RNA Targets. *J. Mol. Biol.* **2017**, *429*, 3139.
(114) Hudson, W. H.; Ortlund, E. A. The Structure, Function and Evolution of Proteins That Bind DNA and RNA. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 749−760.
(115) Chapman, J. R.; Taylor, M. R.; Boulton, S. J. Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Mol. Cell* **2012**, *47*, 497−510.
(116) Helleday, T.; Lo, J.; van Gent, D. C.; Engelward, B. P. DNA Double-Strand Break Repair: From Mechanistic Understanding to Cancer Treatment. *DNA Repair* **2007**, *6*, 923−935.
(117) Karpenshif, Y.; Bernstein, K. A. From Yeast to Mammals: Recent Advances in Genetic Control of Homologous Recombination. *DNA Repair* **2012**, *11*, 781−788.
(118) Lieber, M. R. The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu. Rev. Biochem.* **2010**, *79*, 181−211.
(119) Blackford, A. N.; Jackson, S. P. Atm, Atr, and DNA-Pk: The Trinity at the Heart of the DNA Damage Response. *Mol. Cell* **2017**, *66*, 801−817.
(120) Bekker-Jensen, S.; Lukas, C.; Melander, F.; Bartek, J.; Lukas, J. Dynamic Assembly and Sustained Retention of 53bp1 at the Sites of DNA Damage Are Controlled by Mdc1/Nfbd1. *J. Cell Biol.* **2005**, *170*, 201−211.
(121) Stucki, M.; Clapperton, J. A.; Mohammad, D.; Yaffe, M. B.; Smerdon, S. J.; Jackson, S. P. Mdc1 Directly Binds Phosphorylated Histone H2ax to Regulate Cellular Responses to DNA Double-Strand Breaks. *Cell* **2005**, *123*, 1213−1226.
(122) Celeste, A.; Fernandez-Capetillo, O.; Kruhlak, M. J.; Pilch, D. R.; Staudt, D. W.; Lee, A.; Bonner, R. F.; Bonner, W. M.; Nussenzweig, A. Histone H2ax Phosphorylation Is Dispensable for the Initial Recognition of DNA Breaks. *Nat. Cell Biol.* **2003**, *5*, 675−679.
(123) Yuan, J.; Adamski, R.; Chen, J. Focus on Histone Variant H2ax: To Be or Not to Be. *FEBS Lett.* **2010**, *584*, 3717−3724.
(124) Stellwagen, A. E.; Haimberger, Z. W.; Veatch, J. R.; Gottschling, D. E. Ku Interacts with Telomerase RNA to Promote Telomere Addition at Native and Broken Chromosome Ends. *Genes Dev.* **2003**, *17*, 2384−2395.
(125) Peterson, S. E.; Stellwagen, A. E.; Diede, S. J.; Singer, M. S.; Haimberger, Z. W.; Johnson, C. O.; Tzoneva, M.; Gottschling, D. E. The Function of a Stem-Loop in Telomerase RNA Is Linked to the DNA Repair Protein Ku. *Nat. Genet.* **2001**, *27*, 64−67.
(126) Fisher, T. S.; Taggart, A. K.; Zakian, V. A. Cell Cycle-Dependent Regulation of Yeast Telomerase by Ku. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1198−1205.
(127) Ting, N. S.; Yu, Y.; Pohorelic, B.; Lees-Miller, S. P.; Beattie, T. L. Human Ku70/80 Interacts Directly with Htr, the RNA Component of Human Telomerase. *Nucleic Acids Res.* **2005**, *33*, 2090−2098.
(128) Shi, L.; Qiu, D.; Zhao, G.; Corthesy, B.; Lees-Miller, S.; Reeves, W. H.; Kao, P. N. Dynamic Binding of Ku80, Ku70 and Nf90 to the Il-2 Promoter in Vivo in Activated T-Cells. *Nucleic Acids Res.* **2007**, *35*, 2302−2310.
(129) De Zio, D.; Bordi, M.; Tino, E.; Lanzuolo, C.; Ferraro, E.; Mora, E.; Ciccosanti, F.; Fimia, G. M.; Orlando, V.; Cecconi, F. The DNA Repair Complex Ku70/86 Modulates Apaf1 Expression Upon DNA Damage. *Cell Death Differ.* **2011**, *18*, 516−527.
(130) Mo, X.; Dynan, W. S. Subnuclear Localization of Ku Protein: Functional Association with RNA Polymerase Ii Elongation Sites. *Mol. Cell. Biol.* **2002**, *22*, 8088−8099.
(131) Britton, S.; Coates, J.; Jackson, S. P. A New Method for High-Resolution Imaging of Ku Foci to Decipher Mechanisms of DNA Double-Strand Break Repair. *J. Cell Biol.* **2013**, *202*, 579−595.
(132) Paull, T. T.; Deshpande, R. A. The Mre11/Rad50/Nbs1 Complex: Recent Insights into Catalytic Activities and Atp-Driven Conformational Changes. *Exp. Cell Res.* **2014**, *329*, 139−147.
(133) Myler, L. R.; Gallardo, I. F.; Soniat, M. M.; Deshpande, R. A.; Gonzalez, X. B.; Kim, Y.; Paull, T. T.; Finkelstein, I. J. Single-Molecule Imaging Reveals How Mre11-Rad50-Nbs1 Initiates DNA Break Repair. *Mol. Cell* **2017**, *67*, 891−898 e894.
(134) Wang, W.; Daley, J. M.; Kwon, Y.; Krasner, D. S.; Sung, P. Plasticity of the Mre11-Rad50-Xrs2-Sae2 Nuclease Ensemble in the Processing of DNA-Bound Obstacles. *Genes Dev.* **2017**, *31*, 2331−2336.
(135) Reginato, G.; Cannavo, E.; Cejka, P. Physiological Protein Blocks Direct the Mre11-Rad50-Xrs2 and Sae2 Nuclease Complex to Initiate DNA End Resection. *Genes Dev.* **2017**, *31*, 2325−2330.
(136) Boeing, S.; Williamson, L.; Encheva, V.; Gori, I.; Saunders, R. E.; Instrell, R.; Aygun, O.; Rodriguez-Martinez, M.; Weems, J. C.; Kelly, G. P.; et al. Multiomic Analysis of the Uv-Induced DNA Damage Response. *Cell Rep.* **2016**, *15*, 1597.
(137) Beck, C.; Robert, I.; Reina-San-Martin, B.; Schreiber, V.; Dantzer, F. Poly(Adp-Ribose) Polymerases in Double-Strand Break Repair: Focus on Parp1, Parp2 and Parp3. *Exp. Cell Res.* **2014**, *329*, 18−25.
(138) Ray Chaudhuri, A.; Nussenzweig, A. The Multifaceted Roles of Parp1 in DNA Repair and Chromatin Remodelling. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 610.
(139) Izhar, L.; Adamson, B.; Ciccia, A.; Lewis, J.; Pontano-Vaites, L.; Leng, Y.; Liang, A. C.; Westbrook, T. F.; Harper, J. W.; Elledge, S. J. A Systematic Analysis of Factors Localized to Damaged Chromatin Reveals Parp-Dependent Recruitment of Transcription Factors. *Cell Rep.* **2015**, *11*, 1486−1500.
(140) Krishnakumar, R.; Kraus, W. L. Parp-1 Regulates Chromatin Structure and Transcription through a Kdm5b-Dependent Pathway. *Mol. Cell* **2010**, *39*, 736−749.
(141) Krietsch, J.; Caron, M. C.; Gagne, J. P.; Ethier, C.; Vignard, J.; Vincent, M.; Rouleau, M.; Hendzel, M. J.; Poirier, G. G.; Masson, J. Y. Parp Activation Regulates the RNA-Binding Protein Nono in the DNA Damage Response to DNA Double-Strand Breaks. *Nucleic Acids Res.* **2012**, *40*, 10287−10301.
(142) Teloni, F.; Altmeyer, M. Readers of Poly(Adp-Ribose): Designed to Be Fit for Purpose. *Nucleic Acids Res.* **2016**, *44*, 993−1006.
(143) Bartova, E.; Malyskova, B.; Komurkova, D.; Legartova, S.; Suchankova, J.; Krejci, J.; Kozubek, S. Function of Heterochromatin Protein 1 During DNA Repair. *Protoplasma* **2017**, *254*, 1233−1240.
(144) Lee, Y. H.; Kuo, C. Y.; Stark, J. M.; Shih, H. M.; Ann, D. K. Hp1 Promotes Tumor Suppressor Brca1 Functions During the DNA Damage Response. *Nucleic Acids Res.* **2013**, *41*, 5784−5798.
(145) Sulli, G.; Di Micco, R.; d'Adda di Fagagna, F. Crosstalk between Chromatin State and DNA Damage Response in Cellular Senescence and Cancer. *Nat. Rev. Cancer* **2012**, *12*, 709−720.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

(146) Maison, C.; Bailly, D.; Peters, A. H.; Quivy, J. P.; Roche, D.; Taddei, A.; Lachner, M.; Jenuwein, T.; Almouzni, G. Higher-Order Structure in Pericentric Heterochromatin Involves a Distinct Pattern of Histone Modification and an RNA Component. *Nat. Genet.* **2002**, *30*, 329−334.
(147) Canzio, D.; Larson, A.; Narlikar, G. J. Mechanisms of Functional Promiscuity by Hp1 Proteins. *Trends Cell Biol.* **2014**, *24*, 377−386.
(148) Deng, Q.; Holler, C. J.; Taylor, G.; Hudson, K. F.; Watkins, W.; Gearing, M.; Ito, D.; Murray, M. E.; Dickson, D. W.; Seyfried, N. T.; et al. Fus Is Phosphorylated by DNA-Pk and Accumulates in the Cytoplasm after DNA Damage. *J. Neurosci.* **2014**, *34*, 7802−7813.
(149) Sama, R. R.; Ward, C. L.; Bosco, D. A. Functions of Fus/Tls from DNA Repair to Stress Response: Implications for Als. *ASN Neuro* **2014**, *6*, 175909141454447 10.1177/1759091414544472
(150) Mastrocola, A. S.; Kim, S. H.; Trinh, A. T.; Rodenkirch, L. A.; Tibbetts, R. S. The RNA-Binding Protein Fused in Sarcoma (Fus) Functions Downstream of Poly(Adp-Ribose) Polymerase (Parp) in Response to DNA Damage. *J. Biol. Chem.* **2013**, *288*, 24731−24741.
(151) Rulten, S. L.; Rotheray, A.; Green, R. L.; Grundy, G. J.; Moore, D. A.; Gomez-Herreros, F.; Hafezparast, M.; Caldecott, K. W. Parp-1 Dependent Recruitment of the Amyotrophic Lateral Sclerosis-Associated Protein Fus/Tls to Sites of Oxidative DNA Damage. *Nucleic Acids Res.* **2014**, *42*, 307−314.
(152) Altmeyer, M.; Neelsen, K. J.; Teloni, F.; Pozdnyakova, I.; Pellegrino, S.; Grofte, M.; Rask, M. B.; Streicher, W.; Jungmichel, S.; Nielsen, M. L.; et al. Liquid Demixing of Intrinsically Disordered Proteins Is Seeded by Poly(Adp-Ribose). *Nat. Commun.* **2015**, *6*, 8088.
(153) Hirose, T.; Virnicchi, G.; Tanigawa, A.; Naganuma, T.; Li, R.; Kimura, H.; Yokoi, T.; Nakagawa, S.; Bénard, M.; Fox, A. H.; et al. Neat1 Long Noncoding RNA Regulates Transcription Via Protein Sequestration within Subnuclear Bodies. *Mol.* **2014**, *25*, 169−183.
(154) Wang, W. Y.; Pan, L.; Su, S. C.; Quinn, E. J.; Sasaki, M.; Jimenez, J. C.; Mackenzie, I. R.; Huang, E. J.; Tsai, L. H. Interaction of Fus and Hdac1 Regulates DNA Damage Response and Repair in Neurons. *Nat. Neurosci.* **2013**, *16*, 1383−1391.
(155) Britton, S.; Dernoncourt, E.; Delteil, C.; Froment, C.; Schiltz, O.; Salles, B.; Frit, P.; Calsou, P. DNA Damage Triggers Saf-a and RNA Biogenesis Factors Exclusion from Chromatin Coupled to R-Loops Removal. *Nucleic Acids Res.* **2014**, *42*, 9047−9062.
(156) Conrad, T.; Albrecht, A. S.; de Melo Costa, V. R.; Sauer, S.; Meierhofer, D.; Orom, U. A. Serial Interactome Capture of the Human Cell Nucleus. *Nat. Commun.* **2016**, *7*, 11212.
(157) Akhtar, A.; Zink, D.; Becker, P. B. Chromodomains Are Protein-RNA Interaction Modules. *Nature* **2000**, *407*, 405−409.
(158) Francia, S.; Cabrini, M.; Matti, V.; Oldani, A.; d'Adda di Fagagna, F. Dicer, Drosha and DNA Damage Response RNAs Are Necessary for the Secondary Recruitment of DNA Damage Response Factors. *J. Cell Sci.* **2016**, *129*, 1468−1476.
(159) Pryde, F.; Khalili, S.; Robertson, K.; Selfridge, J.; Ritchie, A. M.; Melton, D. W.; Jullien, D.; Adachi, Y. 53bp1 Exchanges Slowly at the Sites of DNA Damage and Appears to Require RNA for Its Association with Chromatin. *J. Cell Sci.* **2005**, *118*, 2043−2055.
(160) Chakraborty, A.; Tapryal, N.; Venkova, T.; Horikoshi, N.; Pandita, R. K.; Sarker, A. H.; Sarkar, P. S.; Pandita, T. K.; Hazra, T. K. Classical Non-Homologous End-Joining Pathway Utilizes Nascent RNA for Error-Free Double-Strand Break Repair of Transcribed Genes. *Nat. Commun.* **2016**, *7*, 13049.
(161) Lee, H. C.; Aalto, A. P.; Yang, Q.; Chang, S. S.; Huang, G.; Fisher, D.; Cha, J.; Poranen, M. M.; Bamford, D. H.; Liu, Y. The DNA/RNA-Dependent RNA Polymerase Qde-1 Generates Aberrant RNA and Dsrna for RNAi in a Process Requiring Replication Protein a and a DNA Helicase. *PLoS Biol.* **2010**, *8*, e100049610.1371/journal.pbio.1000496
(162) Zhang, Z.; Chang, S. S.; Zhang, Z.; Xue, Z.; Zhang, H.; Li, S.; Liu, Y. Homologous Recombination as a Mechanism to Recognize Repetitive DNA Sequences in an RNAi Pathway. *Genes Dev.* **2013**, *27*, 145−150.
(163) Lee, H. C.; Chang, S. S.; Choudhary, S.; Aalto, A. P.; Maiti, M.; Bamford, D. H.; Liu, Y. Qirna Is a New Type of Small Interfering RNA Induced by DNA Damage. *Nature* **2009**, *459*, 274−277.
(164) Catalanotto, C.; Pallotta, M.; ReFalo, P.; Sachs, M. S.; Vayssie, L.; Macino, G.; Cogoni, C. Redundancy of the Two Dicer Genes in Transgene-Induced Posttranscriptional Gene Silencing in Neurospora Crassa. *Mol. Cell. Biol.* **2004**, *24*, 2536−2545.
(165) Chen, H.; Kobayashi, K.; Miyao, A.; Hirochika, H.; Yamaoka, N.; Nishiguchi, M. Both Osrecq1 and Osrdr1 Are Required for the Production of Small RNA in Response to DNA-Damage in Rice. *PLoS One* **2013**, *8*, e55252.
(166) Holoch, D.; Moazed, D. RNA-Mediated Epigenetic Regulation of Gene Expression. *Nat. Rev. Genet.* **2015**, *16*, 71−84.
(167) Derr, L. K.; Strathern, J. N. A Role for Reverse Transcripts in Gene Conversion. *Nature* **1993**, *361*, 170−173.
(168) Storici, F.; Bebenek, K.; Kunkel, T. A.; Gordenin, D. A.; Resnick, M. A. RNA-Templated DNA Repair. *Nature* **2007**, *447*, 338−341.
(169) Keskin, H.; Shen, Y.; Huang, F.; Patel, M.; Yang, T.; Ashley, K.; Mazin, A. V.; Storici, F. Transcript-RNA-Templated DNA Recombination and Repair. *Nature* **2014**, *515*, 436−439.
(170) Wahba, L.; Gore, S. K.; Koshland, D. The Homologous Recombination Machinery Modulates the Formation of RNA-DNA Hybrids and Associated Chromosome Instability. *eLife* **2013**, *2*, e00505.
(171) Manfrini, N.; Trovesi, C.; Wery, M.; Martina, M.; Cesena, D.; Descrimes, M.; Morillon, A.; d'Adda di Fagagna, F.; Longhese, M. P. RNA-Processing Proteins Regulate Mec1/Atr Activation by Promoting Generation of Rpa-Coated Ssdna. *EMBO Rep.* **2015**, *16*, 221−231.
(172) Michalik, K. M.; Bottcher, R.; Forstemann, K. A Small RNA Response at DNA Ends in Drosophila. *Nucleic Acids Res.* **2012**, *40*, 9596−9603.
(173) Schmidts, I.; Bottcher, R.; Mirkovic-Hosle, M.; Forstemann, K. Homology Directed Repair Is Unaffected by the Absence of Sirnas in Drosophila Melanogaster. *Nucleic Acids Res.* **2016**, *44*, 8261−8271.
(174) Merk, K.; Breinig, M.; Bottcher, R.; Krebs, S.; Blum, H.; Boutros, M.; Forstemann, K. Splicing Stimulates Sirna Formation at Drosophila DNA Double-Strand Breaks. *PLoS Genet.* **2017**, *13*, e1006861.
(175) Lu, W. T.; Hawley, B. R.; Skalka, G. L.; Baldock, R. A.; Smith, E. M.; Bader, A. S.; Malewicz, M.; Watts, F. Z.; Wilczynska, A.; Bushell, M. Drosha Drives the Formation of DNA: RNA Hybrids around DNA Break Sites to Facilitate DNA Repair. *Nat. Commun.* **2018**, *9*, 532.
(176) Wang, Q.; Goldstein, M. Small RNAs Recruit Chromatin-Modifying Enzymes Mmset and Tip60 to Reconfigure Damaged DNA Upon Double-Strand Break and Facilitate Repair. *Cancer Res.* **2016**, *76*, 1904−1915.
(177) Gao, M.; Wei, W.; Li, M. M.; Wu, Y. S.; Ba, Z.; Jin, K. X.; Li, M. M.; Liao, Y. Q.; Adhikari, S.; Chong, Z.; et al. Ago2 Facilitates Rad51 Recruitment and DNA Double-Strand Break Repair by Homologous Recombination. *Cell Res.* **2014**, *24*, 532−541.
(178) Qi, Y.; Zhang, Y.; Baller, J. A.; Voytas, D. F. Histone H2ax and the Small RNA Pathway Modulate Both Non-Homologous End-Joining and Homologous Recombination in Plants. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **2016**, *783*, 9−14.
(179) Miki, D.; Zhu, P.; Zhang, W.; Mao, Y.; Feng, Z.; Huang, H.; Zhang, H.; Li, Y.; Liu, R.; Qi, Y.; et al. Efficient Generation of Dirnas Requires Components in the Posttranscriptional Gene Silencing Pathway. *Sci. Rep.* **2017**, *7*, 301.
(180) Storici, F.; Tichon, A. E. RNA Takes over Control of DNA Break Repair. *Nat. Cell Biol.* **2017**, *19*, 1382−1384.
(181) Rossiello, F.; Aguado, J.; Sepe, S.; Iannelli, F.; Nguyen, Q.; Pitchiaya, S.; Carninci, P.; d'Adda di Fagagna, F. DNA Damage Response Inhibition at Dysfunctional Telomeres by Modulation of Telomeric DNA Damage Response RNAs. *Nat. Commun.* **2017**, *8*, 13980.
(182) Sainsbury, S.; Bernecky, C.; Cramer, P. Structural Basis of Transcription Initiation by RNA Polymerase Ii. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 129−143.
(183) Kadesch, T. R.; Chamberlin, M. J. Studies of in Vitro Transcription by Calf Thymus RNA Polymerase Ii Using a Novel Duplex DNA Template. *J. Biol. Chem.* **1982**, *257*, 5286−5295.
(184) Dynan, W. S.; Burgess, R. R. In Vitro Transcription by Wheat Germ RNA Polymerase II. Initiation of RNA Synthesis on Relaxed, Closed Circular Template. *J. Biol. Chem.* **1981**, *256*, 5866−5873.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

(185) Shibata, A.; Moiani, D.; Arvai, A. S.; Perry, J.; Harding, S. M.; Genois, M. M.; Maity, R.; van Rossum-Fikkert, S.; Kertokalio, A.; Romoli, F.; et al. DNA Double-Strand Break Repair Pathway Choice Is Directed by Distinct Mre11 Nuclease Activities. *Mol. Cell* **2014**, *53*, 7−18.
(186) Liu, Y.; Sung, S.; Kim, Y.; Li, F.; Gwon, G.; Jo, A.; Kim, A. K.; Kim, T.; Song, O. K.; Lee, S. E.; et al. Atp-Dependent DNA Binding, Unwinding, and Resection by the Mre11/Rad50 Complex. *EMBO J.* **2016**, *35*, 743−758.
(187) Volpe, T. A.; Kidner, C.; Hall, I. M.; Teng, G.; Grewal, S. I.; Martienssen, R. A. Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. *Science* **2002**, *297*, 1833−1837.
(188) Djupedal, I.; Portoso, M.; Spahr, H.; Bonilla, C.; Gustafsson, C. M.; Allshire, R. C.; Ekwall, K. RNA Pol Ii Subunit Rpb7 Promotes Centromeric Transcription and RNAi-Directed Chromatin Silencing. *Genes Dev.* **2005**, *19*, 2301−2306.
(189) Buhler, M.; Haas, W.; Gygi, S. P.; Moazed, D. RNAi-Dependent and -Independent RNA Turnover Mechanisms Contribute to Heterochromatic Gene Silencing. *Cell* **2007**, *129*, 707−721.
(190) Shimada, Y.; Mohn, F.; Buhler, M. The RNA-Induced Transcriptional Silencing Complex Targets Chromatin Exclusively Via Interacting with Nascent Transcripts. *Genes Dev.* **2016**, *30*, 2571−2580.
(191) Iannelli, F.; Galbiati, A.; Capozzo, I.; Nguyen, Q.; Magnuson, B.; Michelini, F.; D'Alessandro, G.; Cabrini, M.; Roncador, M.; Francia, S.; et al. A Damaged Genome's Transcriptional Landscape through Multilayered Expression Profiling around in Situ-Mapped DNA Double-Strand Breaks. *Nat. Commun.* **2017**, *8*, 15656.
(192) Jih, G.; Iglesias, N.; Currie, M. A.; Bhanu, N. V.; Paulo, J. A.; Gygi, S. P.; Garcia, B. A.; Moazed, D. Unique Roles for Histone H3k9me States in RNAi and Heritable Silencing of Transcription. *Nature* **2017**, *547*, 463−467.
(193) Ohle, C.; Tesorero, R.; Schermann, G.; Dobrev, N.; Sinning, I.; Fischer, T. Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair. *Cell* **2016**, *167*, 1001−1013 e1007.
(194) Cohen, S.; Puget, N.; Lin, Y. L.; Clouaire, T.; Aguirrebengoa, M.; Rocher, V.; Pasero, P.; Canitrot, Y.; Legube, G. Senataxin Resolves RNA:DNA Hybrids Forming at DNA Double-Strand Breaks to Prevent Translocations. *Nat. Commun.* **2018**, *9*, 533.
(195) Lewis, C. J.; Pan, T.; Kalsotra, A. RNA Modifications and Structures Cooperate to Guide RNA-Protein Interactions. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 202−210.
(196) Helm, M.; Motorin, Y. Detecting RNA Modifications in the Epitranscriptome: Predict and Validate. *Nat. Rev. Genet.* **2017**, *18*, 275−291.
(197) Garalde, D. R.; Snell, E. A.; Jachimowicz, D.; Sipos, B.; Lloyd, J. H.; Bruce, M.; Pantic, N.; Admassu, T.; James, P.; Warland, A.; et al. Highly Parallel Direct RNA Sequencing on an Array of Nanopores. *Nat. Methods* **2018**, *15*, 201−206.
(198) Xiang, Y.; Laurent, B.; Hsu, C. H.; Nachtergaele, S.; Lu, Z.; Sheng, W.; Xu, C.; Chen, H.; Ouyang, J.; Wang, S.; et al. RNA M6a Methylation Regulates the Ultraviolet-Induced DNA Damage Response. *Nature* **2017**, *543*, 573−576.
(199) Wahl, M. C.; Will, C. L.; Lührmann, R. The Spliceosome: Design Principles of a Dynamic Rnp Machine. *Cell* **2009**, *136*, 701−718.
(200) Will, C. L.; Lührmann, R. Spliceosome Structure and Function. *Cold Spring Harbor Perspect. Biol.* **2011**, *3*, a003707.
(201) Kornblihtt, A. R.; Schor, I. E.; Allo, M.; Dujardin, G.; Petrillo, E.; Muñoz, M. J. Alternative Splicing: A Pivotal Step between Eukaryotic Transcription and Translation. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 153−165.
(202) Pan, Q.; Shai, O.; Lee, L. J.; Frey, B. J.; Blencowe, B. J. Deep Surveying of Alternative Splicing Complexity in the Human Transcriptome by High-Throughput Sequencing. *Nat. Genet.* **2008**, *40*, 1413−1415.
(203) Montecucco, A.; Biamonti, G. Pre-Mrna Processing Factors Meet the DNA Damage Response. *Front. Genet.* **2013**, *4*, 102.
(204) Giono, L. E.; Nieto Moreno, N.; Cambindo Botto, A. E.; Dujardin, G.; Muñoz, M. J.; Kornblihtt, A. R. The RNA Response to DNA Damage. *J. Mol. Biol.* **2016**, *428*, 2636−2651.
(205) Shkreta, L.; Chabot, B. The RNA Splicing Response to DNA Damage. *Biomolecules* **2015**, *5*, 2935−2977.
(206) Naro, C.; Bielli, P.; Pagliarini, V.; Sette, C. The Interplay between DNA Damage Response and RNA Processing: The Unexpected Role of Splicing Factors as Gatekeepers of Genome Stability. *Front. Genet.* **2015**, *6*, 142.
(207) Schwartz, S.; Ast, G. Chromatin Density and Splicing Destiny: On the Cross-Talk between Chromatin Structure and Splicing. *EMBO J.* **2010**, *29*, 1629−1636.
(208) Luco, R. F.; Allo, M.; Schor, I. E.; Kornblihtt, A. R.; Misteli, T. Epigenetics in Alternative Pre-Mrna Splicing. *Cell* **2011**, *144*, 16−26.
(209) Loomis, R. J.; Naoe, Y.; Parker, J. B.; Savic, V.; Bozovsky, M. R.; Macfarlan, T.; Manley, J. L.; Chakravarti, D. Chromatin Binding of Srp20 and Asf/Sf2 and Dissociation from Mitotic Chromosomes Is Modulated by Histone H3 Serine 10 Phosphorylation. *Mol. Cell* **2009**, *33*, 450−461.
(210) Sims, R. J., 3rd; Millhouse, S.; Chen, C. F.; Lewis, B. A.; Erdjument-Bromage, H.; Tempst, P.; Manley, J. L.; Reinberg, D. Recognition of Trimethylated Histone H3 Lysine 4 Facilitates the Recruitment of Transcription Postinitiation Factors and Pre-Mrna Splicing. *Mol. Cell* **2007**, *28*, 665−676.
(211) Kornblihtt, A. R. Coupling Transcription and Alternative Splicing. *Adv. Exp. Med. Biol.* **2007**, *623*, 175−189.
(212) Listerman, I.; Sapra, A. K.; Neugebauer, K. M. Cotranscriptional Coupling of Splicing Factor Recruitment and Precursor Messenger RNA Splicing in Mammalian Cells. *Nat. Struct. Mol. Biol.* **2006**, *13*, 815−822.
(213) de la Mata, M.; Lafaille, C.; Kornblihtt, A. R. First Come, First Served Revisited: Factors Affecting the Same Alternative Splicing Event Have Different Effects on the Relative Rates of Intron Removal. *RNA* **2010**, *16*, 904−912.
(214) Hsin, J. P.; Manley, J. L. The RNA Polymerase Ii Ctd Coordinates Transcription and RNA Processing. *Genes Dev.* **2012**, *26*, 2119−2137.
(215) David, C. J.; Manley, J. L. The RNA Polymerase C-Terminal Domain: A New Role in Spliceosome Assembly. *Transcription* **2011**, *2*, 221−225.
(216) Dutertre, M.; Lambert, S.; Carreira, A.; Amor-Gueret, M.; Vagner, S. DNA Damage: RNA-Binding Proteins Protect from near and Far. *Trends Biochem. Sci.* **2014**, *39*, 141−149.
(217) Boutz, P. L.; Bhutkar, A.; Sharp, P. A. Detained Introns Are a Novel, Widespread Class of Post-Transcriptionally Spliced Introns. *Genes Dev.* **2015**, *29*, 63−80.
(218) Beli, P.; Lukashchuk, N.; Wagner, S. A.; Weinert, B. T.; Olsen, J. V.; Baskcomb, L.; Mann, M.; Jackson, S. P.; Choudhary, C. Proteomic Investigations Reveal a Role for RNA Processing Factor Thrap3 in the DNA Damage Response. *Mol. Cell* **2012**, *46*, 212−225.
(219) Savage, K. I.; Gorski, J. J.; Barros, E. M.; Irwin, G. W.; Manti, L.; Powell, A. J.; Pellagatti, A.; Lukashchuk, N.; McCance, D. J.; McCluggage, W. G.; et al. Identification of a Brca1-Mrna Splicing Complex Required for Efficient DNA Repair and Maintenance of Genomic Stability. *Mol. Cell* **2014**, *54*, 445−459.
(220) Pederiva, C.; Böhm, S.; Julner, A.; Farnebo, M. Splicing Controls the Ubiquitin Response During DNA Double-Strand Break Repair. *Cell Death Differ.* **2016**, *23*, 1648−1657.
(221) Tanikawa, M.; Sanjiv, K.; Helleday, T.; Herr, P.; Mortusewicz, O. The Spliceosome U2 Snrnp Factors Promote Genome Stability through Distinct Mechanisms; Transcription of Repair Factors and R-Loop Processing. *Oncogenesis* **2016**, *5*, e280.
(222) Yokoi, A.; Kotake, Y.; Takahashi, K.; Kadowaki, T.; Matsumoto, Y.; Minoshima, Y.; Sugi, N. H.; Sagane, K.; Hamaguchi, M.; Iwata, M.; et al. Biological Validation That Sf3b Is a Target of the Antitumor Macrolide Pladienolide. *FEBS J.* **2011**, *278*, 4870−4880.
(223) O'Brien, K.; Matlin, A. J.; Lowell, A. M.; Moore, M. J. The Biflavonoid Isoginkgetin Is a General Inhibitor of Pre-Mrna Splicing. *J. Biol. Chem.* **2008**, *283*, 33147−33154.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

(224) Adamson, B.; Smogorzewska, A.; Sigoillot, F. D.; King, R. W.; Elledge, S. J. A Genome-Wide Homologous Recombination Screen Identifies the RNA-Binding Protein Rbmx as a Component of the DNA-Damage Response. *Nat. Cell Biol.* **2012**, *14*, 318−328.
(225) Grote, M.; Wolf, E.; Will, C. L.; Lemm, I.; Agafonov, D. E.; Schomburg, A.; Fischle, W.; Urlaub, H.; Lührmann, R. Molecular Architecture of the Human Prp19/Cdc5l Complex. *Mol. Cell. Biol.* **2010**, *30*, 2105−2119.
(226) Song, E. J.; Werner, S. L.; Neubauer, J.; Stegmeier, F.; Aspden, J.; Rio, D.; Harper, J. W.; Elledge, S. J.; Kirschner, M. W.; Rape, M. The Prp19 Complex and the Usp4sart3 Deubiquitinating Enzyme Control Reversible Ubiquitination at the Spliceosome. *Genes Dev.* **2010**, *24*, 1434−1447.
(227) Chanarat, S.; Seizl, M.; Strässer, K. The Prp19 Complex Is a Novel Transcription Elongation Factor Required for Trex Occupancy at Transcribed Genes. *Genes Dev.* **2011**, *25*, 1147−1158.
(228) David, C. J.; Boyne, A. R.; Millhouse, S. R.; Manley, J. L. The RNA Polymerase Ii C-Terminal Domain Promotes Splicing Activation through Recruitment of a U2af65-Prp19 Complex. *Genes Dev.* **2011**, *25*, 972−983.
(229) Grey, M.; Düsterhöft, A.; Henriques, J. A.; Brendel, M. Allelism of Pso4 and Prp19 Links Pre-Mrna Processing with Recombination and Error-Prone DNA Repair in Saccharomyces Cerevisiae. *Nucleic Acids Res.* **1996**, *24*, 4009−4014.
(230) 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.; et al. A Genome-Wide Sirna Screen Reveals Diverse Cellular Processes and Pathways That Mediate Genome Stability. *Mol. Cell* **2009**, *35*, 228−239.
(231) Abbas, M.; Shanmugam, I.; Bsaili, M.; Hromas, R.; Shaheen, M. The Role of the Human Psoralen 4 (Hpso4) Protein Complex in Replication Stress and Homologous Recombination. *J. Biol. Chem.* **2014**, *289*, 14009−14019.
(232) Marechal, A.; Li, J. M.; Ji, X. Y.; Wu, C. S.; Yazinski, S. A.; Nguyen, H. D.; Liu, S.; Jimenez, A. E.; Jin, J.; Zou, L. Prp19 Transforms into a Sensor of Rpa-Ssdna after DNA Damage and Drives Atr Activation Via a Ubiquitin-Mediated Circuitry. *Mol. Cell* **2014**, *53*, 235−246.
(233) Zhang, N.; Kaur, R.; Lu, X.; Shen, X.; Li, L.; Legerski, R. J. The Pso4Mrna Splicing and DNA Repair Complex Interacts with Wrn for Processing of DNA Interstrand Cross-Links. *J. Biol. Chem.* **2005**, *280*, 40559−40567.
(234) Kuraoka, I.; Ito, S.; Wada, T.; Hayashida, M.; Lee, L.; Saijo, M.; Nakatsu, Y.; Matsumoto, M.; Matsunaga, T.; Handa, H.; et al. Isolation of Xab2 Complex Involved in Pre-Mrna Splicing, Transcription, and Transcription-Coupled Repair. *J. Biol. Chem.* **2008**, *283*, 940−950.
(235) Wan, L.; Huang, J. The Pso4 Protein Complex Associates with Replication Protein a (Rpa) and Modulates the Activation of Ataxia Telangiectasia-Mutated and Rad3-Related (Atr). *J. Biol. Chem.* **2014**, *289*, 6619−6626.
(236) Smeenk, G.; Mailand, N. Writers, Readers, and Erasers of Histone Ubiquitylation in DNA Double-Strand Break Repair. *Front. Genet.* **2016**, *7*, 122.
(237) Yuan, M.; Eberhart, C. G.; Kai, M. RNA Binding Protein Rbm14 Promotes Radio-Resistance in Glioblastoma by Regulating DNA Repair and Cell Differentiation. *Oncotarget* **2014**, *5*, 2820−2826.
(238) Simon, N. E.; Yuan, M.; Kai, M. RNA-Binding Protein Rbm14 Regulates Dissociation and Association of Non-Homologous End Joining Proteins. *Cell Cycle* **2017**, *16*, 1−6.
(239) Polo, S. E.; Blackford, A. N.; Chapman, J. R.; Baskcomb, L.; Gravel, S.; Rusch, A.; Thomas, A.; Blundred, R.; Smith, P.; Kzhyshkowska, J.; et al. Regulation of DNA-End Resection by Hnrnpu-Like Proteins Promotes DNA Double-Strand Break Signaling and Repair. *Mol. Cell* **2012**, *45*, 505−516.
(240) Hong, Z.; Jiang, J.; Ma, J.; Dai, S.; Xu, T.; Li, H.; Yasui, A. The Role of Hnrpul1 Involved in DNA Damage Response Is Related to Parp1. *PLoS One* **2013**, *8*, e60208.
(241) Teng, S. C.; Kim, B.; Gabriel, A. Retrotransposon Reverse-Transcriptase-Mediated Repair of Chromosomal Breaks. *Nature* **1996**, *383*, 641−644.
(242) Ding, W.; Lin, L.; Chen, B.; Dai, J. L1 Elements, Processed Pseudogenes and Retrogenes in Mammalian Genomes. *IUBMB Life* **2006**, *58*, 677−685.
(243) Meers, C.; Keskin, H.; Storici, F. DNA Repair by RNA: Templated, or Not Templated, That Is the Question. *DNA Repair* **2016**, *44*, 17−21.
(244) Moore, J. K.; Haber, J. E. Capture of Retrotransposon DNA at the Sites of Chromosomal Double-Strand Breaks. *Nature* **1996**, *383*, 644−646.
(245) Trott, D. A.; Porter, A. C. Hypothesis: Transcript-Templated Repair of DNA Double-Strand Breaks. *BioEssays* **2006**, *28*, 78−83.
(246) Storici, F. RNA-Mediated DNA Modifications and RNA-Templated DNA Repair. *Curr. Opin. Mol. Ther.* **2008**, *10*, 224−230.
(247) Thaler, D. S.; Tombline, G.; Zahn, K. Short-Patch Reverse Transcription in Escherichia Coli. *Genetics* **1995**, *140*, 909−915.
(248) Shen, Y.; Nandi, P.; Taylor, M. B.; Stuckey, S.; Bhadsavle, H. P.; Weiss, B.; Storici, F. RNA-Driven Genetic Changes in Bacteria and in Human Cells. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **2011**, *717*, 91−98.
(249) Shen, Y.; Storici, F. Generation of RNA/DNA Hybrids in Genomic DNA by Transformation Using RNA-Containing Oligonucleotides. *J. Visualized Exp.* **2010**, DOI: 10.3791/2152.
(250) Nowacki, M.; Vijayan, V.; Zhou, Y.; Schotanus, K.; Doak, T. G.; Landweber, L. F. RNA-Mediated Epigenetic Programming of a Genome-Rearrangement Pathway. *Nature* **2008**, *451*, 153−158.
(251) Xu, P. Z.; Yuan, S.; Li, Y.; Zhang, H. Y.; Wang, X. D.; Lin, H. H.; Wu, X. J. Genome-Wide High-Frequency Non-Mendelian Loss of Heterozygosity in Rice. *Genome* **2007**, *50*, 297−302.
(252) Jividen, K.; Li, H. Chimeric RNAs Generated by Intergenic Splicing in Normal and Cancer Cells. *Genes, Chromosomes Cancer* **2014**, *53*, 963−971.
(253) Li, H.; Wang, J.; Mor, G.; Sklar, J. A Neoplastic Gene Fusion Mimics Trans-Splicing of RNAs in Normal Human Cells. *Science* **2008**, *321*, 1357−1361.
(254) Li, H.; Wang, J.; Ma, X.; Sklar, J. Gene Fusions and RNA Trans-Splicing in Normal and Neoplastic Human Cells. *Cell Cycle* **2009**, *8*, 218−222.
(255) Kowarz, E.; Dingermann, T.; Marschalek, R. Do Non-Genomically Encoded Fusion Transcripts Cause Recurrent Chromosomal Translocations? *Cancers* **2012**, *4*, 1036−1049.
(256) Kowarz, E.; Merkens, J.; Karas, M.; Dingermann, T.; Marschalek, R. Premature Transcript Termination, Trans-Splicing and DNA Repair: A Vicious Path to Cancer. *Am. J. Blood Res.* **2011**, *1*, 1−12.
(257) Kasahara, M.; Clikeman, J. A.; Bates, D. B.; Kogoma, T. Reca Protein-Dependent R-Loop Formation in Vitro. *Genes Dev.* **2000**, *14*, 360−365.
(258) Zaitsev, E. N.; Kowalczykowski, S. C. A Novel Pairing Process Promoted by Escherichia Coli Reca Protein: Inverse DNA and RNA Strand Exchange. *Genes Dev.* **2000**, *14*, 740−749.
(259) Mazina, O. M.; Keskin, H.; Hanamshet, K.; Storici, F.; Mazin, A. V. Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair. *Mol. Cell* **2017**, *67*, 19.
(260) Aymard, F.; Bugler, B.; Schmidt, C. K.; Guillou, E.; Caron, P.; Briois, S.; Iacovoni, J. S.; Daburon, V.; Miller, K. M.; Jackson, S. P.; et al. Transcriptionally Active Chromatin Recruits Homologous Recombination at DNA Double-Strand Breaks. *Nat. Struct. Mol. Biol.* **2014**, *21*, 366−374.
(261) Shi, C.; Shen, X.; Niu, S.; Ma, C. Innate Reverse Transcriptase Activity of DNA Polymerase for Isothermal RNA Direct Detection. *J. Am. Chem. Soc.* **2015**, *137*, 13804−13806.
(262) Murakami, E.; Feng, J. Y.; Lee, H.; Hanes, J.; Johnson, K. A.; Anderson, K. S. Characterization of Novel Reverse Transcriptase and Other RNA-Associated Catalytic Activities by Human DNA Polymerase Gamma: Importance in Mitochondrial DNA Replication. *J. Biol. Chem.* **2003**, *278*, 36403−36409.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

(263) Wei, L.; Nakajima, S.; Bohm, S.; Bernstein, K. A.; Shen, Z.; Tsang, M.; Levine, A. S.; Lan, L. DNA Damage During the G0/G1 Phase Triggers RNA-Templated, Cockayne Syndrome B-Dependent Homologous Recombination. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E3495−3504.
(264) Keskin, H.; Meers, C.; Storici, F. Transcript RNA Supports Precise Repair of Its Own DNA Gene. *RNA Biol.* **2016**, *13*, 157−165.
(265) Aguilera, A.; Garcia-Muse, T. R Loops: From Transcription Byproducts to Threats to Genome Stability. *Mol. Cell* **2012**, *46*, 115−124.
(266) Castellano-Pozo, M.; Santos-Pereira, J. M.; Rondon, A. G.; Barroso, S.; Andujar, E.; Perez-Alegre, M.; Garcia-Muse, T.; Aguilera, A. R Loops Are Linked to Histone H3 S10 Phosphorylation and Chromatin Condensation. *Mol. Cell* **2013**, *52*, 583−590.
(267) Garcia-Pichardo, D.; Canas, J. C.; Garcia-Rubio, M. L.; Gomez-Gonzalez, B.; Rondon, A. G.; Aguilera, A. Histone Mutants Separate R Loop Formation from Genome Instability Induction. *Mol. Cell* **2017**, *66*, 597−609 e595.
(268) Prescott, D. M. The DNA of Ciliated Protozoa. *Microbiol Rev.* **1994**, *58*, 233−267.
(269) Aeschlimann, S. H.; Jonsson, F.; Postberg, J.; Stover, N. A.; Petera, R. L.; Lipps, H. J.; Nowacki, M.; Swart, E. C. The Draft Assembly of the Radically Organized Stylonychia Lemnae Macronuclear Genome. *Genome Biol. Evol.* **2014**, *6*, 1707−1723.
(270) Swart, E. C.; Bracht, J. R.; Magrini, V.; Minx, P.; Chen, X.; Zhou, Y.; Khurana, J. S.; Goldman, A. D.; Nowacki, M.; Schotanus, K.; et al. The Oxytricha Trifallax Macronuclear Genome: A Complex Eukaryotic Genome with 16,000 Tiny Chromosomes. *PLoS Biol.* **2013**, *11*, e1001473.
(271) Chen, X.; Bracht, J. R.; Goldman, A. D.; Dolzhenko, E.; Clay, D. M.; Swart, E. C.; Perlman, D. H.; Doak, T. G.; Stuart, A.; Amemiya, C. T.; et al. The Architecture of a Scrambled Genome Reveals Massive Levels of Genomic Rearrangement During Development. *Cell* **2014**, *158*, 1187−1198.
(272) Chalker, D. L.; Yao, M. C. DNA Elimination in Ciliates: Transposon Domestication and Genome Surveillance. *Annu. Rev. Genet.* **2011**, *45*, 227−246.
(273) Jahn, C. L.; Klobutcher, L. A. Genome Remodeling in Ciliated Protozoa. *Annu. Rev. Microbiol.* **2002**, *56*, 489−520.
(274) Nowacki, M.; Landweber, L. F. Epigenetic Inheritance in Ciliates. *Curr. Opin. Microbiol.* **2009**, *12*, 638−643.
(275) Prescott, D. M. Genome Gymnastics: Unique Modes of DNA Evolution and Processing in Ciliates. *Nat. Rev. Genet.* **2000**, *1*, 191−198.
(276) Prescott, D. M.; DuBois, M. L. Internal Eliminated Segments (Iess) of Oxytrichidae. *J. Eukaryotic Microbiol.* **1996**, *43*, 432−441.
(277) Prescott, D. M.; Ehrenfeucht, A.; Rozenberg, G. Template-Guided Recombination for Ies Elimination and Unscrambling of Genes in Stichotrichous Ciliates. *J. Theor. Biol.* **2003**, *222*, 323−330.
(278) Lindblad, K. A.; Bracht, J. R.; Williams, A. E.; Landweber, L. Thousands of RNA-Cached Copies of Whole Chromosomes Are Present in the Ciliate Oxytricha During Development. *RNA* **2017**, *23*, 1200.
(279) Neeb, Z. T.; Hogan, D. J.; Katzman, S.; Zahler, A. M. Preferential Expression of Scores of Functionally and Evolutionarily Diverse DNA and RNA-Binding Proteins During Oxytricha Trifallax Macronuclear Development. *PLoS One* **2017**, *12*, e0170870.
(280) Fang, W.; Wang, X.; Bracht, J. R.; Nowacki, M.; Landweber, L. F. Piwi-Interacting RNAs Protect DNA against Loss During Oxytricha Genome Rearrangement. *Cell* **2012**, *151*, 1243−1255.
(281) Zahler, A. M.; Neeb, Z. T.; Lin, A.; Katzman, S. Mating of the Stichotrichous Ciliate Oxytricha Trifallax Induces Production of a Class of 27 Nt Small RNAs Derived from the Parental Macronucleus. *PLoS One* **2012**, *7*, e42371.
(282) Kurth, H. M.; Mochizuki, K. 2'-O-Methylation Stabilizes Piwi-Associated Small RNAs and Ensures DNA Elimination in Tetrahymena. *RNA* **2009**, *15*, 675−685.
(283) Bracht, J. R.; Perlman, D. H.; Landweber, L. F. Cytosine Methylation and Hydroxymethylation Mark DNA for Elimination in Oxytricha Trifallax. *Genome Biol.* **2012**, *13*, R99.
(284) Chalker, D. L.; Yao, M. C. Nongenic, Bidirectional Transcription Precedes and May Promote Developmental DNA Deletion in Tetrahymena Thermophila. *Genes Dev.* **2001**, *15*, 1287−1298.
(285) Malone, C. D.; Anderson, A. M.; Motl, J. A.; Rexer, C. H.; Chalker, D. L. Germ Line Transcripts Are Processed by a Dicer-Like Protein That Is Essential for Developmentally Programmed Genome Rearrangements of Tetrahymena Thermophila. *Mol. Cell. Biol.* **2005**, *25*, 9151−9164.
(286) Mochizuki, K.; Gorovsky, M. A. Conjugation-Specific Small RNAs in Tetrahymena Have Predicted Properties of Scan (Scn) RNAs Involved in Genome Rearrangement. *Genes Dev.* **2004**, *18*, 2068−2073.
(287) Mochizuki, K.; Gorovsky, M. A. A Dicer-Like Protein in Tetrahymena Has Distinct Functions in Genome Rearrangement, Chromosome Segregation, and Meiotic Prophase. *Genes Dev.* **2005**, *19*, 77−89.
(288) Sandoval, P. Y.; Swart, E. C.; Arambasic, M.; Nowacki, M. Functional Diversification of Dicer-Like Proteins and Small RNAs Required for Genome Sculpting. *Dev. Cell* **2014**, *28*, 174−188.
(289) Duharcourt, S.; Butler, A.; Meyer, E. Epigenetic Self-Regulation of Developmental Excision of an Internal Eliminated Sequence in Paramecium Tetraurelia. *Genes Dev.* **1995**, *9*, 2065−2077.
(290) Baudry, C.; Malinsky, S.; Restituito, M.; Kapusta, A.; Rosa, S.; Meyer, E.; Betermier, M. Piggymac, a Domesticated Piggybac Transposase Involved in Programmed Genome Rearrangements in the Ciliate Paramecium Tetraurelia. *Genes Dev.* **2009**, *23*, 2478−2483.
(291) Cheng, C. Y.; Vogt, A.; Mochizuki, K.; Yao, M. C. A Domesticated Piggybac Transposase Plays Key Roles in Heterochromatin Dynamics and DNA Cleavage During Programmed DNA Deletion in Tetrahymena Thermophila. *Mol. Biol. Cell* **2010**, *21*, 1753−1762.
(292) Dubois, E.; Bischerour, J.; Marmignon, A.; Mathy, N.; Regnier, V.; Betermier, M. Transposon Invasion of the Paramecium Germline Genome Countered by a Domesticated Piggybac Transposase and the Nhej Pathway. *Int. J. Evol. Biol.* **2012**, *2012*, 436196.
(293) Kapusta, A.; Matsuda, A.; Marmignon, A.; Ku, M.; Silve, A.; Meyer, E.; Forney, J. D.; Malinsky, S.; Betermier, M. Highly Precise and Developmentally Programmed Genome Assembly in Paramecium Requires Ligase Iv-Dependent End Joining. *PLoS Genet.* **2011**, *7*, e1002049.
(294) Allen, S. E.; Hug, I.; Pabian, S.; Rzeszutek, I.; Hoehener, C.; Nowacki, M. Circular Concatemers of Ultra-Short DNA Segments Produce Regulatory RNAs. *Cell* **2017**, *168*, 990−999 e997.
(295) Betermier, M.; Duharcourt, S.; Seitz, H.; Meyer, E. Timing of Developmentally Programmed Excision and Circularization of Paramecium Internal Eliminated Sequences. *Mol. Cell. Biol.* **2000**, *20*, 1553−1561.
(296) Noto, T.; Kataoka, K.; Suhren, J. H.; Hayashi, A.; Woolcock, K. J.; Gorovsky, M. A.; Mochizuki, K. Small-RNA-Mediated Genome-Wide Trans-Recognition Network in Tetrahymena DNA Elimination. *Mol. Cell* **2015**, *59*, 229−242.
(297) Haag, J. R.; Pikaard, C. S. Multisubunit RNA Polymerases Iv and V: Purveyors of Non-Coding RNA for Plant Gene Silencing. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 483−492.
(298) Khurana, J. S.; Wang, X.; Chen, X.; Perlman, D. H.; Landweber, L. F. Transcription-Independent Functions of an RNA Polymerase Ii Subunit, Rpb2, During Genome Rearrangement in the Ciliate, Oxytricha Trifallax. *Genetics* **2014**, *197*, 839−849.
(299) d'Adda di Fagagna, F.; Teo, S.-H.; Jackson, S. P. Functional Links between Telomeres and Proteins of the DNA-Damage Response. *Genes Dev.* **2004**, *18*, 1781−1799.
(300) de Lange, T. How Telomeres Solve the End-Protection Problem. *Science* **2009**, *326*, 948−952.
(301) Fumagalli, M.; Rossiello, F.; Clerici, M.; Barozzi, S.; Cittaro, D.; Kaplunov, J. M.; Bucci, G.; Dobreva, M.; Matti, V.; Beausejour, C. M.; et al. Telomeric DNA Damage Is Irreparable and Causes Persistent DNA-Damage-Response Activation. *Nat. Cell Biol.* **2012**, *14*, 355−365.
(302) Bae, N. S.; Baumann, P. A Rap1/Trf2 Complex Inhibits Nonhomologous End-Joining at Human Telomeric DNA Ends. *Mol. Cell* **2007**, *26*, 323−334.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

(303) Doksani, Y.; Wu, J. Y.; de Lange, T.; Zhuang, X. Super-Resolution Fluorescence Imaging of Telomeres Reveals Trf2-Dependent T-Loop Formation. *Cell* **2013**, *155*, 345−356.
(304) Smogorzewska, A.; Karlseder, J.; Holtgreve-Grez, H.; Jauch, A.; de Lange, T. DNA Ligase Iv-Dependent Nhej of Deprotected Mammalian Telomeres in G1 and G2. *Curr. Biol.* **2002**, *12*, 1635.
(305) Sfeir, A.; de Lange, T. Removal of Shelterin Reveals the Telomere End-Protection Problem. *Science* **2012**, *336*, 593−597.
(306) Arnoult, N.; Karlseder, J. Complex Interactions between the DNA-Damage Response and Mammalian Telomeres. *Nat. Struct. Mol. Biol.* **2015**, *22*, 859−866.
(307) Davoli, T.; de Lange, T. Telomere-Driven Tetraploidization Occurs in Human Cells Undergoing Crisis and Promotes Transformation of Mouse Cells. *Cancer Cell* **2012**, *21*, 765−776.
(308) Davoli, T.; Denchi, E. L.; de Lange, T. Persistent Telomere Damage Induces Bypass of Mitosis and Tetraploidy. *Cell* **2010**, *141*, 81−93.
(309) Maciejowski, J.; Li, Y.; Bosco, N.; Campbell, P. J.; de Lange, T. Chromothripsis and Kataegis Induced by Telomere Crisis. *Cell* **2015**, *163*, 1641−1654.
(310) d'Adda di Fagagna, F.; Reaper, P. M.; Clay-Farrace, L.; Fiegler, H.; Carr, P.; Von Zglinicki, T.; Saretzki, G.; Carter, N. P.; Jackson, S. P. A DNA Damage Checkpoint Response in Telomere-Initiated Senescence. *Nature* **2003**, *426*, 194−198.
(311) Herbig, U.; Jobling, W. A.; Chen, B. P.; Chen, D. J.; Sedivy, J. M. Telomere Shortening Triggers Senescence of Human Cells through a Pathway Involving Atm, P53, and P21(Cip1), but Not P16(Ink4a). *Mol. Cell* **2004**, *14*, 501−513.
(312) Galbiati, A.; Beausejour, C.; d'Adda di Fagagna, F. A Novel Single-Cell Method Provides Direct Evidence of Persistent DNA Damage in Senescent Cells and Aged Mammalian Tissues. *Aging Cell* **2017**, *16*, 422−427.
(313) Rossiello, F.; Herbig, U.; Longhese, M. P.; Fumagalli, M.; d'Adda di Fagagna, F. Irreparable Telomeric DNA Damage and Persistent Ddr Signalling as a Shared Causative Mechanism of Cellular Senescence and Ageing. *Curr. Opin. Genet. Dev.* **2014**, *26*, 89−95.
(314) Teixeira, M. T.; Arneric, M.; Sperisen, P.; Lingner, J. Telomere Length Homeostasis Is Achieved Via a Switch between Telomerase-Extendible and -Nonextendible States. *Cell* **2004**, *117*, 323−335.
(315) Hemann, M. T.; Strong, M. A.; Hao, L. Y.; Greider, C. W. The Shortest Telomere, Not Average Telomere Length, Is Critical for Cell Viability and Chromosome Stability. *Cell* **2001**, *107*, 67−77.
(316) Liu, Y.; Kha, H.; Ungrin, M.; Robinson, M. O.; Harrington, L. Preferential Maintenance of Critically Short Telomeres in Mammalian Cells Heterozygous for Mtert. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 3597−3602.
(317) Lansdorp, P. M. Telomeres, Stem Cells, and Hematology. *Blood* **2008**, *111*, 1759−1766.
(318) Shay, J. W.; Wright, W. E. Telomeres and Telomerase in Normal and Cancer Stem Cells. *FEBS Lett.* **2010**, *584*, 3819−3825.
(319) Dagg, R. A.; Pickett, H. A.; Neumann, A. A.; Napier, C. E.; Henson, J. D.; Teber, E. T.; Arthur, J. W.; Reynolds, C. P.; Murray, J.; Haber, M.; et al. Extensive Proliferation of Human Cancer Cells with Ever-Shorter Telomeres. *Cell Rep.* **2017**, *19*, 2544−2556.
(320) Viceconte, N.; Dheur, M. S.; Majerova, E.; Pierreux, C. E.; Baurain, J. F.; van Baren, N.; Decottignies, A. Highly Aggressive Metastatic Melanoma Cells Unable to Maintain Telomere Length. *Cell Rep.* **2017**, *19*, 2529−2543.
(321) Barthel, F. P.; Wei, W.; Tang, M.; Martinez-Ledesma, E.; Hu, X.; Amin, S. B.; Akdemir, K. C.; Seth, S.; Song, X.; Wang, Q.; et al. Systematic Analysis of Telomere Length and Somatic Alterations in 31 Cancer Types. *Nat. Genet.* **2017**, *49*, 349−357.
(322) Sagie, S.; Toubiana, S.; Hartono, S. R.; Katzir, H.; Tzur-Gilat, A.; Havazelet, S.; Francastel, C.; Velasco, G.; Chedin, F.; Selig, S. Telomeres in Icf Syndrome Cells Are Vulnerable to DNA Damage Due to Elevated DNA: RNA Hybrids. *Nat. Commun.* **2017**, *8*, 14015.
(323) Moravec, M.; Wischnewski, H.; Bah, A.; Hu, Y.; Liu, N.; Lafranchi, L.; King, M. C.; Azzalin, C. M. Terra Promotes Telomerase-Mediated Telomere Elongation in Schizosaccharomyces Pombe. *EMBO Rep.* **2016**, *17*, 999−1012.
(324) Cusanelli, E.; Romero, C. A.; Chartrand, P. Telomeric Noncoding RNA Terra Is Induced by Telomere Shortening to Nucleate Telomerase Molecules at Short Telomeres. *Mol. Cell* **2013**, *51*, 780−791.
(325) Balk, B.; Maicher, A.; Dees, M.; Klermund, J.; Luke-Glaser, S.; Bender, K.; Luke, B. Telomeric RNA-DNA Hybrids Affect Telomere-Length Dynamics and Senescence. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1199−1205.
(326) Graf, M.; Bonetti, D.; Lockhart, A.; Serhal, K.; Kellner, V.; Maicher, A.; Jolivet, P.; Teixeira, M. T.; Luke, B. Telomere Length Determines Terra and R-Loop Regulation through the Cell Cycle. *Cell* **2017**, *170*, 72−85 e14.
(327) Chu, H. P.; Cifuentes-Rojas, C.; Kesner, B.; Aeby, E.; Lee, H. G.; Wei, C.; Oh, H. J.; Boukhali, M.; Haas, W.; Lee, J. T. Terra RNA Antagonizes Atrx and Protects Telomeres. *Cell* **2017**, *170*, 86−101 e116.
(328) Cusanelli, E.; Chartrand, P. Telomeric Noncoding RNA: Telomeric Repeat-Containing RNA in Telomere Biology. *Wiley Interdiscip Rev. RNA* **2014**, *5*, 407−419.
(329) Azzalin, C. M.; Reichenbach, P.; Khoriauli, L.; Giulotto, E.; Lingner, J. Telomeric Repeat Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. *Science* **2007**, *318*, 798−801.
(330) Nergadze, S. G.; Farnung, B. O.; Wischnewski, H.; Khoriauli, L.; Vitelli, V.; Chawla, R.; Giulotto, E.; Azzalin, C. M. Cpg-Island Promoters Drive Transcription of Human Telomeres. *RNA* **2009**, *15*, 2186−2194.
(331) Pfeiffer, V.; Lingner, J. Terra Promotes Telomere Shortening through Exonuclease 1-Mediated Resection of Chromosome Ends. *PLoS Genet.* **2012**, *8*, e1002747.
(332) Porro, A.; Feuerhahn, S.; Delafontaine, J.; Riethman, H.; Rougemont, J.; Lingner, J. Functional Characterization of the Terra Transcriptome at Damaged Telomeres. *Nat. Commun.* **2014**, *5*, 5379.
(333) Porro, A.; Feuerhahn, S.; Reichenbach, P.; Lingner, J. Molecular Dissection of Telomeric Repeat-Containing RNA Biogenesis Unveils the Presence of Distinct and Multiple Regulatory Pathways. *Mol. Cell. Biol.* **2010**, *30*, 4808−4817.
(334) Flynn, R. L.; Cox, K. E.; Jeitany, M.; Wakimoto, H.; Bryll, A. R.; Ganem, N. J.; Bersani, F.; Pineda, J. R.; Suva, M. L.; Benes, C. H.; et al. Alternative Lengthening of Telomeres Renders Cancer Cells Hypersensitive to Atr Inhibitors. *Science* **2015**, *347*, 273−277.
(335) Redon, S.; Reichenbach, P.; Lingner, J. Protein RNA and Protein Protein Interactions Mediate Association of Human Est1a/Smg6 with Telomerase. *Nucleic Acids Res.* **2007**, *35*, 7011−7022.
(336) Smekalova, E.; Baumann, P. Terra -a Calling Card for Telomerase. *Mol. Cell* **2013**, *51*, 703−704.
(337) Maicher, A.; Kastner, L.; Dees, M.; Luke, B. Deregulated Telomere Transcription Causes Replication-Dependent Telomere Shortening and Promotes Cellular Senescence. *Nucleic Acids Res.* **2012**, *40*, 6649−6659.
(338) Pfeiffer, V.; Crittin, J.; Grolimund, L.; Lingner, J. The Tho Complex Component Thp2 Counteracts Telomeric R-Loops and Telomere Shortening. *EMBO J.* **2013**, *32*, 2861−2871.
(339) Arora, R.; Lee, Y.; Wischnewski, H.; Brun, C. M.; Schwarz, T.; Azzalin, C. M. RNAseh1 Regulates Terra-Telomeric DNA Hybrids and Telomere Maintenance in Alt Tumour Cells. *Nat. Commun.* **2014**, *5*, 5220.
(340) Arora, R.; Azzalin, C. M. Telomere Elongation Chooses Terra Alternatives. *RNA Biol.* **2015**, *12*, 938.
(341) Yu, T. Y.; Kao, Y. W.; Lin, J. J. Telomeric Transcripts Stimulate Telomere Recombination to Suppress Senescence in Cells Lacking Telomerase. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 3377.
(342) Nanavaty, V.; Sandhu, R.; Jehi, S. E.; Pandya, U. M.; Li, B. Trypanosoma Brucei Rap1Maintains Telomere and Subtelomere Integrity by Suppressing Terra and Telomeric RNA:DNA Hybrids. *Nucleic Acids Res.* **2017**, *45*, 5785−5796.
(343) Feretzaki, M.; Lingner, J. A Practical Qpcr Approach to Detect Terra, the Elusive Telomeric Repeat-Containing RNA. *Methods* **2017**, *114*, 39−45.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

(344) Baker, A. M.; Fu, Q.; Hayward, W.; Victoria, S.; Pedroso, I. M.; Lindsay, S. M.; Fletcher, T. M. The Telomere Binding Protein Trf2 Induces Chromatin Compaction. *PLoS One* **2011**, *6*, e19124.
(345) Denchi, E. L.; de Lange, T. Protection of Telomeres through Independent Control of Atm and Atr by Trf2 and Pot1. *Nature* **2007**, *448*, 1068−1071.
(346) Arnoult, N.; Van Beneden, A.; Decottignies, A. Telomere Length Regulates Terra Levels through Increased Trimethylation of Telomeric H3k9 and Hp1alpha. *Nat. Struct. Mol. Biol.* **2012**, *19*, 948−956.
(347) Montero, J. J.; Lopez de Silanes, I.; Grana, O.; Blasco, M. A. Telomeric RNAs Are Essential to Maintain Telomeres. *Nat. Commun.* **2016**, *7*, 12534.
(348) Lopez de Silanes, I.; Grana, O.; De Bonis, M. L.; Dominguez, O.; Pisano, D. G.; Blasco, M. A. Identification of Terra Locus Unveils a Telomere Protection Role through Association to Nearly All Chromosomes. *Nat. Commun.* **2014**, *5*, 4723.
(349) Bah, A.; Wischnewski, H.; Shchepachev, V.; Azzalin, C. M. The Telomeric Transcriptome of Schizosaccharomyces Pombe. *Nucleic Acids Res.* **2012**, *40*, 2995.
(350) Greenwood, J.; Cooper, J. P. Non-Coding Telomeric and Subtelomeric Transcripts Are Differentially Regulated by Telomeric and Heterochromatin Assembly Factors in Fission Yeast. *Nucleic Acids Res.* **2012**, *40*, 2956.
(351) Cao, F.; Li, X.; Hiew, S.; Brady, H.; Liu, Y.; Dou, Y. Dicer Independent Small RNAs Associate with Telomeric Heterochromatin. *RNA* **2009**, *15*, 1274−1281.
(352) Pitchiaya, S.; Krishnan, V.; Custer, T. C.; Walter, N. G. Dissecting Non-Coding RNA Mechanisms in Cellulo by Single-Molecule High-Resolution Localization and Counting. *Methods* **2013**, *63*, 188−199.
(353) Pitchiaya, S.; Heinicke, L. A.; Custer, T. C.; Walter, N. G. Single Molecule Fluorescence Approaches Shed Light on Intracellular RNAs. *Chem. Rev.* **2014**, *114*, 3224−3265.
(354) Pitchiaya, S.; Heinicke, L. A.; Park, J. I.; Cameron, E. L.; Walter, N. G. Resolving Subcellular Mirna Trafficking and Turnover at Single-Molecule Resolution. *Cell Rep.* **2017**, *19*, 630−642.
(355) Paige, J. S.; Nguyen-Duc, T.; Song, W.; Jaffrey, S. R. Fluorescence Imaging of Cellular Metabolites with RNA. *Science* **2012**, *335*, 1194.
(356) Gall, J. G.; Pardue, M. L. Formation and Detection of RNA-DNA Hybrid Molecules in Cytological Preparations. *Proc. Natl. Acad. Sci. U. S. A.* **1969**, *63*, 378−383.
(357) Crosetto, N.; Bienko, M.; van Oudenaarden, A. Spatially Resolved Transcriptomics and Beyond. *Nat. Rev. Genet.* **2015**, *16*, 57−66.
(358) Toseland, C. P. Fluorescent Labeling and Modification of Proteins. *J. Chem. Biol.* **2013**, *6*, 85−95.
(359) Goncalves, M. S. Fluorescent Labeling of Biomolecules with Organic Probes. *Chem. Rev.* **2009**, *109*, 190−212.
(360) Lubeck, E.; Coskun, A.; Zhiyentayev, T.; Ahmad, M.; Cai, L. Single-Cell in Situ RNA Profiling by Sequential Hybridization. *Nat. Methods* **2014**, *11*, 360.
(361) Moffitt, J.; Zhuang, X. RNA Imaging with Multiplexed Error-Robust Fluorescence in Situ Hybridization (Merfish). *Methods Enzymol.* **2016**, *572*, 1.
(362) Lee, J. H.; Daugharthy, E. R.; Scheiman, J.; Kalhor, R.; Yang, J. L.; Ferrante, T. C.; Terry, R.; Jeanty, S. S.; Li, C.; Amamoto, R.; et al. Highly Multiplexed Subcellular RNA Sequencing in Situ. *Science* **2014**, *343*, 1360−1363.
(363) Choi, H. M.; Beck, V. A.; Pierce, N. A. Next-Generation in Situ Hybridization Chain Reaction: Higher Gain, Lower Cost, Greater Durability. *ACS Nano* **2014**, *8*, 4284−4294.
(364) Hocine, S.; Raymond, P.; Zenklusen, D.; Chao, J. A.; Singer, R. H. Single-Molecule Analysis of Gene Expression Using Two-Color RNA Labeling in Live Yeast. *Nat. Methods* **2013**, *10*, 119−121.
(365) Yan, X.; Hoek, T. A.; Vale, R. D.; Tanenbaum, M. E. Dynamics of Translation of Single Mrna Molecules in Vivo. *Cell* **2016**, *165*, 976−989.
(366) Pichon, X.; Bastide, A.; Safieddine, A.; Chouaib, R.; Samacoits, A.; Basyuk, E.; Peter, M.; Mueller, F.; Bertrand, E. Visualization of Single Endogenous Polysomes Reveals the Dynamics of Translation in Live Human Cells. *J. Cell Biol.* **2016**, *214*, 769−781.
(367) Morisaki, T.; Lyon, K.; DeLuca, K. F.; DeLuca, J. G.; English, B. P.; Zhang, Z.; Lavis, L. D.; Grimm, J. B.; Viswanathan, S.; Looger, L. L.; et al. Real-Time Quantification of Single RNA Translation Dynamics in Living Cells. *Science* **2016**, *352*, 1425−1429.
(368) Wu, B.; Eliscovich, C.; Yoon, Y. J.; Singer, R. H. Translation Dynamics of Single Mrnas in Live Cells and Neurons. *Science* **2016**, *352*, 1430−1435.
(369) Iizuka, R.; Yamagishi-Shirasaki, M.; Funatsu, T. Kinetic Study of De Novo Chromophore Maturation of Fluorescent Proteins. *Anal. Biochem.* **2011**, *414*, 17310.1016/j.ab.2011.03.036
(370) Thorn, K. Genetically Encoded Fluorescent Tags. *Mol. Biol. Cell* **2017**, *28*, 84810.1091/mbc.E16-07-0504.
(371) Tanenbaum, M. E.; Gilbert, L. A.; Qi, L. S.; Weissman, J. S.; Vale, R. D. A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell* **2014**, *159*, 635−646.
(372) Viswanathan, S. e. a. High-Performance Probes for Light and Electron Microscopy. *Nat. Methods* **2015**, *12*, 568−76.
(373) Grimm, J. B.; English, B. P.; Chen, J.; Slaughter, J. P.; Zhang, Z.; Revyakin, A.; Patel, R.; Macklin, J. J.; Normanno, D.; Singer, R. H.; et al. A General Method to Improve Fluorophores for Live-Cell and Single-Molecule Microscopy. *Nat. Methods* **2015**, *12*, 244−250.
(374) Schmidt, J. C.; Zaug, A. J.; Cech, T. R. Live Cell Imaging Reveals the Dynamics of Telomerase Recruitment to Telomeres. *Cell* **2016**, *166*, 1188−1197.
(375) Halstead, J. M.; Lionnet, T.; Wilbertz, J. H.; Wippich, F.; Ephrussi, A.; Singer, R. H.; Chao, J. A. Translation. An RNA Biosensor for Imaging the First Round of Translation from Single Cells to Living Animals. *Science* **2015**, *347*, 1367−1671.
(376) Nelles, D.; Fang, M.; O'Conell, M.; Xu, J.; Markmiller, S.; Doudna, J.; Yeo, G. Programmable RNA Tracking in Live Cells with Crispr/Cas9. *Cell* **2016**, *165*, 48810.1016/j.cell.2016.02.054
(377) Paredes, E.; Evans, M.; Das, S. R. RNA Labeling, Conjugation and Ligation. *Methods* **2011**, *54*, 251−259.
(378) Custer, T. C.; Walter, N. G. In Vitro Labeling Strategies for in Cellulo Fluorescence Microscopy of Single Ribonucleoprotein Machines. *Protein Sci.* **2017**, *26*, 1363−1379.
(379) Tsekouras, K.; Custer, T. C.; Jashnsaz, H.; Walter, N. G.; Presse, S. A Novel Method to Accurately Locate and Count Large Numbers of Steps by Photobleaching. *Mol. Biol. Cell* **2016**, *27*, 3601−3615.
(380) Stewart, M. P.; Sharei, A.; Ding, X.; Sahay, G.; Langer, R.; Jensen, K. F. In Vitro and Ex Vivo Strategies for Intracellular Delivery. *Nature* **2016**, *538*, 183−192.
(381) Pitchiaya, S.; Androsavich, J. R.; Walter, N. G. Intracellular Single Molecule Microscopy Reveals Two Kinetically Distinct Pathways for Microrna Assembly. *EMBO Rep.* **2012**, *13*, 709−715.
(382) Strack, R. L.; Jaffrey, S. R. Live-Cell Imaging of Mammalian RNAs with Spinach2. *Methods Enzymol.* **2015**, *550*, 129−146.
(383) Svensen, N.; Jaffrey, S. R Fluorescent RNA Aptamers as a Tool to Study RNA-Modifying Enzymes. *Cell Chem. Biol.* **2016**, *23*, 415−425.
(384) You, M.; Litke, J. L.; Jaffrey, S. R. Imaging Metabolite Dynamics in Living Cells Using a Spinach-Based Riboswitch. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E2756−2765.
(385) Dolgosheina, E. V.; Jeng, S. C. Y.; Panchapakesan, S. S. S.; Cojocaru, R.; Chen, P. S. K.; Wilson, P. D.; Hawkins, N.; Wiggins, P. A.; Unrau, P. J. RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking. *ACS Chem. Biol.* **2014**, *9*, 2412−2420.
(386) Dolgosheina, E. V.; Unrau, P. J. Fluorophore-Binding RNA Aptamers and Their Applications. *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 843−851.
(387) Khvorova, A.; Watts, J. K. The Chemical Evolution of Oligonucleotide Therapies of Clinical Utility. *Nat. Biotechnol.* **2017**, *35*, 238−248.
(388) Obad, S.; dos Santos, C. O.; Petri, A.; Heidenblad, M.; Broom, O.; Ruse, C.; Fu, C.; Lindow, M.; Stenvang, J.; Straarup, E. M.; et al. Silencing of Microrna Families by Seed-Targeting Tiny LNAs. *Nat. Genet.* **2011**, *43*, 371−378.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403

(389) Bennett, C. F.; Swayze, E. E. RNA Targeting Therapeutics: Molecular Mechanisms of Antisense Oligonucleotides as a Therapeutic Platform. *Annu. Rev. Pharmacol. Toxicol.* **2010**, *50*, 259−293.
(390) Eckstein, F. Phosphorothioates, Essential Components of Therapeutic Oligonucleotides. *Nucleic Acid Ther.* **2014**, *24*, 374−387.
(391) Summerton, J.; Weller, D. Morpholino Antisense Oligomers: Design, Preparation, and Properties. *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 187−195.
(392) Good, L.; Nielsen, P. E. Progress in Developing Pna as a Gene-Targeted Drug. *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 431−437.
(393) Frazier, K. S. Antisense Oligonucleotide Therapies: The Promise and the Challenges from a Toxicologic Pathologist's Perspective. *Toxicol. Pathol.* **2015**, *43*, 78−89.
(394) Juliano, R. L. The Delivery of Therapeutic Oligonucleotides. *Nucleic Acids Res.* **2016**, *44*, 6518−6548.
(395) Smith, R. A.; Miller, T. M.; Yamanaka, K.; Monia, B. P.; Condon, T. P.; Hung, G.; Lobsiger, C. S.; Ward, C. M.; McAlonis-Downes, M.; Wei, H.; et al. Antisense Oligonucleotide Therapy for Neurodegenerative Disease. *J. Clin. Invest.* **2006**, *116*, 2290−2296.
(396) Crooke, S. T.; Wang, S.; Vickers, T. A.; Shen, W.; Liang, X. H. Cellular Uptake and Trafficking of Antisense Oligonucleotides. *Nat. Biotechnol.* **2017**, *35*, 230−237.
(397) Rigo, F.; Seth, P. P.; Bennett, C. F. Antisense Oligonucleotide-Based Therapies for Diseases Caused by Pre-Mrna Processing Defects. *Adv. Exp. Med. Biol.* **2014**, *825*, 303−352.
(398) Miller, J. W.; Urbinati, C. R.; Teng-Umnuay, P.; Stenberg, M. G.; Byrne, B. J.; Thornton, C. A.; Swanson, M. S. Recruitment of Human Muscleblind Proteins to (Cug)(N) Expansions Associated with Myotonic Dystrophy. *EMBO J.* **2000**, *19*, 4439−4448.
(399) Conlon, E. G.; Lu, L.; Sharma, A.; Yamazaki, T.; Tang, T.; Shneider, N. A.; Manley, J. L. The C9orf72 Ggggcc Expansion Forms RNA G-Quadruplex Inclusions and Sequesters Hnrnp H to Disrupt Splicing in Als Brains. *eLife* **2016**, *5*,10.7554/eLife.17820.
(400) Timchenko, N. A.; Cai, Z. J.; Welm, A. L.; Reddy, S.; Ashizawa, T.; Timchenko, L. T. RNA Cug Repeats Sequester Cugbp1 and Alter Protein Levels and Activity of Cugbp1. *J. Biol. Chem.* **2001**, *276*, 7820−7826.
(401) Jain, A.; Vale, R. D. RNA Phase Transitions in Repeat Expansion Disorders. *Nature* **2017**, *546*, 243−247.
(402) Li, Z.; Rana, T. M. Therapeutic Targeting of Micrornas: Current Status and Future Challenges. *Nat. Rev. Drug Discovery* **2014**, *13*, 622−638.
(403) Rupaimoole, R.; Slack, F. J. Microrna Therapeutics: Towards a New Era for the Management of Cancer and Other Diseases. *Nat. Rev. Drug Discovery* **2017**, *16*, 203−222.
(404) Ottosen, S.; Parsley, T. B.; Yang, L.; Zeh, K.; van Doorn, L. J.; van der Veer, E.; Raney, A. K.; Hodges, M. R.; Patick, A. K. In Vitro Antiviral Activity and Preclinical and Clinical Resistance Profile of Miravirsen, a Novel Anti-Hepatitis C Virus Therapeutic Targeting the Human Factor Mir-122. *Antimicrob. Agents Chemother.* **2015**, *59*, 599−608.
(405) Gutschner, T.; Hammerle, M.; Eissmann, M.; Hsu, J.; Kim, Y.; Hung, G.; Revenko, A.; Arun, G.; Stentrup, M.; Gross, M.; et al. The Noncoding RNA Malat1 Is a Critical Regulator of the Metastasis Phenotype of Lung Cancer Cells. *Cancer Res.* **2013**, *73*, 1180−1189.
(406) Arun, G.; Diermeier, S.; Akerman, M.; Chang, K. C.; Wilkinson, J. E.; Hearn, S.; Kim, Y.; MacLeod, A. R.; Krainer, A. R.; Norton, L.; et al. Differentiation of Mammary Tumors and Reduction in Metastasis Upon Malat1 lncRNA Loss. *Genes Dev.* **2016**, *30*, 34−51.
(407) Watts, J. K.; Corey, D. R. Gene Silencing by Sirnas and Antisense Oligonucleotides in the Laboratory and the Clinic. *J. Pathol.* **2012**, *226*, 365−379.
(408) Armanios, M.; Blackburn, E. H. The Telomere Syndromes. *Nat. Rev. Genet.* **2012**, *13*, 693−704.
(409) Stephenson, M. L.; Zamecnik, P. C. Inhibition of Rous Sarcoma Viral RNA Translation by a Specific Oligodeoxyribonucleotide. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75*, 285−288.
(410) The Vitravene Study Group. A Randomized Controlled Clinical Trial of Intravitreous Fomivirsen for Treatment of Newly Diagnosed Peripheral Cytomegalovirus Retinitis in Patients with Aids. *Am. J. Ophthalmol.* **2002**, *133*, 467−474.
(411) Crooke, R. M.; Graham, M. J.; Lemonidis, K. M.; Whipple, C. P.; Koo, S.; Perera, R. J. An Apolipoprotein B Antisense Oligonucleotide Lowers Ldl Cholesterol in Hyperlipidemic Mice without Causing Hepatic Steatosis. *J. Lipid Res.* **2005**, *46*, 872−884.
(412) Stein, C. A. Eteplirsen Approved for Duchenne Muscular Dystrophy: The Fda Faces a Difficult Choice. *Mol. Ther.* **2016**, *24*, 1884−1885.
(413) Hua, Y.; Sahashi, K.; Hung, G.; Rigo, F.; Passini, M. A.; Bennett, C. F.; Krainer, A. R. Antisense Correction of Smn2 Splicing in the Cns Rescues Necrosis in a Type Iii Sma Mouse Model. *Genes Dev.* **2010**, *24*, 1634−1644.
(414) de Hoon, M.; Shin, J. W.; Carninci, P. Paradigm Shifts in Genomics through the Fantom Projects. *Mamm. Genome* **2015**, *26*, 391−402.
(415) Aguzzi, A.; Altmeyer, M. Phase Separation: Linking Cellular Compartmentalization to Disease. *Trends Cell Biol.* **2016**, *26*, 547−558.
(416) Courchaine, E. M.; Lu, A.; Neugebauer, K. M. Droplet Organelles? *EMBO J.* **2016**, *35*, 1603−1612.
(417) Kuhn, T. S. *The Structure of Scientific Revolutions*; University of Chicago Press, 1970.

DOI: 10.1021/acs.chemrev.7b00487
Chem. Rev. 2018, 118, 4365−4403