UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXVI CICLO DEL DOTTORATO DI RICERCA IN

BIOMEDICINA MOLECOLARE

Recruitment of H3.3 histone chaperon activity to R-loops by the Splicing-Factor Proline and Arginine Rich (SFPQ) stabilizes repetitive elements in cancer cells

Settore scientifico-disciplinare: BIO/11

Dottorando:
Alessandro Ferrando

Coordinatore:
Prof. Germana Meroni

Supervisore di tesi:
Prof. Stefan Schoeftner

Co-supervisore di tesi:
Prof. Licio Collavin

ANNO ACCADEMICO 2022/2023

Abstract

R-loops are triple-stranded nucleic acid structures containing an RNA moiety paired with its antisense DNA template strand and the loop of the displaced, single-stranded DNA (ssDNA). RNA:DNA hybrids form under normal physiological conditions and exert multiple biological functions. However, persistent R-loops have been shown to mediate replication stress, DNA damage and drive genomic instability, with a particular relevance at vertebrate telomeres.

We recently found that the RNA-binding protein SFPQ has a critical role in limiting R-loops formation at telomeres of human cancer cell. In line with this, loss-of-function SFPQ cells show increased levels of RNA:DNA hybrids, replication stress and DNA damage markers and subsequent genomic instability at telomeres level.

Here we show that SFPQ function is not limited to telomeres, but expands to other non-coding repeat regions in the human genome, as depicted by ChIP-seq analysis, paralleled by increased R-loop levels and genomic instability.

Moreover, SFPQ action is mediated by the its novel interactor DAXX, that allows R-loop suppression by inserting H3.3 histone variant.

Finally, RNA-seq analysis pointed out a SFPQ - R-loop dependent signature of innate immune response activation, unveiling a novel connection between non coding genome regulation and immune response, suggesting SFPQ as putative target for sarcoma treatment.

Table of contents

1 Introduction

1.1 R-loop

R-loops are triple-stranded nucleic acid structures that contain RNA:DNA hybrids, formed by a RNA molecule hybridized to its cognate DNA template, and the resulting displaced, single-stranded DNA (ssDNA) [1].

X-ray crystallography structural studies reveal that RNA:DNA hybrid adopt a particular conformation that is an intermediate between the A-form of nucleic acid, typical of double-stranded RNA (dsRNA) and the B-form which is peculiar of double-stranded DNA (dsDNA) [2].

R-loops were first discovered as transcription byproduct in bacteria[3], and subsequently detected in all relevant model organisms, as reviewed here [4]–[7].

Mapping experiments revealed that R-loops form all across the genome, both in coding and non-coding portions [8].

Two models have been proposed for R-loop formation. R-loops form prevalently co-transcriptionally, *in cis*, driven by negative supercoiling of DNA upstream of the progressing RNA polymerase (RNA-pol) that was shown to increase the likelihood of the nascent transcript to anneal to its complementary DNA strand [9]–[11].

Recent evidence support also the formation of R-loops *in trans*, a RAD51 dependent mechanism that supports invasion and base-pairing of RNAs to a complementary, yet spatially distant DNA sequence, forcing the looping of the non-homologous strands (**Figure 1**) [12], [13].

R-loop forming regions are typically G-rich [14], [15] presenting GC- and AT-skew, thus indicating an asymmetric nucleotide distribution in such regions [16]–[22]. The asymmetry in also reflected in the nucleotide strand abundance, with cytosines on the template strand and guanines on the non-template strand [16], [17], [23]. On top of this, G/C-rich R-loops can be further stabilized by folding of the displaced ssDNA into G-quadruplex (G4) structures. G4 structures form by intra or intermolecular stacking of two or more guanine tetrads, held together by Hoogsteen hydrogen bonds [24]–[27]. Due to the high energy level of guanin-guanin bonds, once formed, G4 structures can promote R-loop stability, thus impairing their resolution [9], [28].

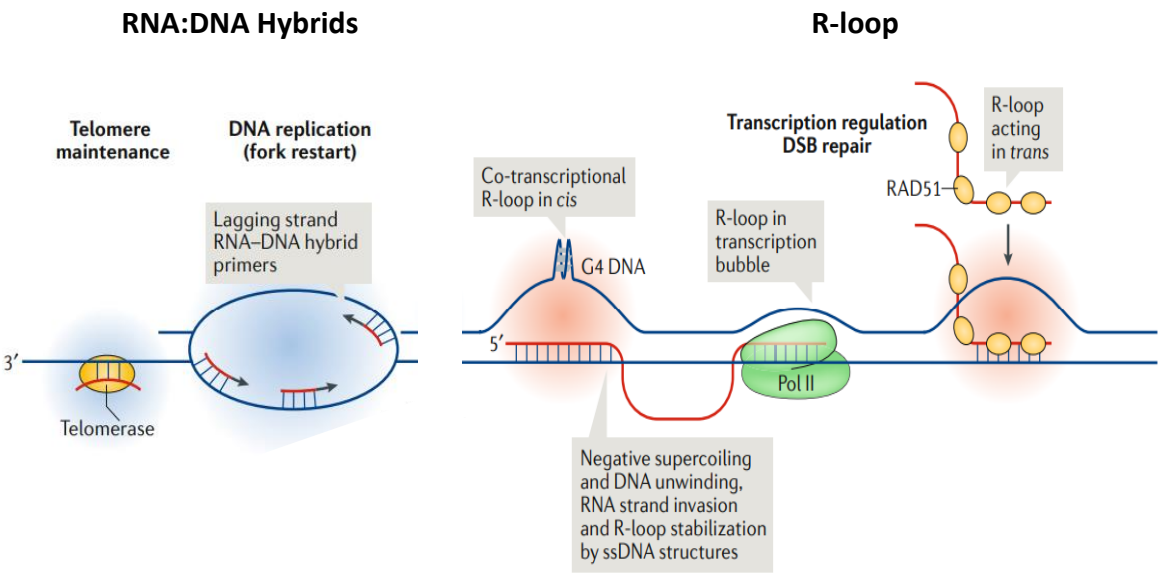


Figure 1. Models of RNA:DNA hybrid and R-loop formation.

RNA:DNA hybrids and R-loops are physiologically present at different locations in the genome, carrying out many distinct functions. Left, RNA:DNA hybrids are reported to mediate Telomeres elongation and DNA replication. Right, R-loops are shown to occur both in cis, by hybridizing downstream the transcribing RNA pol II, and in trans, by lncRNA in RAD51 recombinase-dependent manner. These interactions between RNA and DNA have important implications for DNA repair, replication, and gene regulation. Image adapted from Niehrs & Luke, 2020 [10].

1.2 Physiological roles of R-loop structures

R-loops represent prime sites for DNA damage mediated by replication-transcription conflicts between the replisome and RNAP complexes moving along DNA, increasing mutation rate [29], [30]. However, besides these deleterious effect, R-loops has been reported control central physiological processes [31]–[35] across different species [34], [36]–[38]. Most classic example range from R-loop driven immunoglobulin class switch recombination (CSR) promote antibody isotype diversity in activated B cells [39]–[41] to formation of RNA primers for the replication of the bacteriophage T4 DNA [42], the natural ColE1 plasmid [43], and mitochondrial DNA [44].

R-loops have been identified to play crucial role in controlling chromatin structure and gene transcription, either activating or repressing it [17], [33].

1.2.1 Regulation of Gene Expression and Chromatin Structure

R-loops have been linked with the imposition of post-translational histone modifications during transcriptional activation or repression [16], [34], [45]. In the context of chromatin condensation, a study in yeast *S. pombe* showed that single-stranded DNA of R-loops located in centromeric regions can recruit the RNA-induced transcriptional silencing (RITS) complex to promote pericentric heterochromatin formation [46]. In addition to this, R-loop formation at centromeric and pericentromeric regions has been associated with the imposition of H3S10P, a repressive chromatin compaction mark, known to accumulate in mitosis and meiosis, impacting on chromosome segregation [45], [47], [48]. On the other hand, R-loops at CpG rich promoter regions, are associated with an euchromatic pattern with elevated levels of histone H3 lysine 4 trimethylation (H3K4me3), as well as deposition of H4K20me1 and H3K79me2, promoting initiation and/or elongation of transcription [16], [17]. R-loops epigenetic action has also been reported to promote gene activation and transcription. Studies showed that the presence of R-loops in promoter regions can protect genes form the action of the *de novo* DNA methyltransferase 3B1 (DNMT3B1), ensuring gene activation. This suggests a model where the presence of R-loops prevents the access of DNMT3B1 promoting local hypomethylation and subsequent activation of transcription [16]. The regulation of mRNA transcription activation through R-loops has been demonstrated to depend on non-coding RNAs (ncRNAs) [1]. In human colon adenocarcinoma cell lines, an antisense RNA was shown to generate a R-loop at the CpG island containing promoter of the human vimentin gene (VIM), enhancing chromatin opening and binding of the transcription factor p65, core component of the NF-κB complex [49]. A recent study conducted in mouse embryonic stem cells (ESCs) demonstrated that R-loops located at specific gene promoters can recruit the transcriptional activating acetyltransferase, Tip60-p400, while concomitantly impeding the binding of the Polycomb Repressive Complex 2 (PRC2), that is typically involved in ESC differentiation, thus strongly regulating gene regulation and cell fate. Accordingly, overexpression of RNase H1, a RNase able to specifically, degrade the RNA component of a R-loop without affecting the DNA strands, reduces the presence of TIP60 and p400 at their target genes, promoting the recruitment of PRC2. Remarkably, another study in mouse ESCs found that R-loops recruit PRC1 and PRC2 towards a specific subset of PRC-repressed target genes, promoting the imposition of repressive chromatin marks, such as histone H2A Lysine

119 ubiquitination (H2AK119ub) and histone H3 Lysine 27 tri-methylation (H3K27me3), respectively. Removing R-loops led to the activation of PRC target genes that contain R-loops [50]. Thus, R-loops can act as epigenetic regulators that modulate gene expression in a context dependent manner, supporting an important role of R-loops in the control of gene expression. However, R-loops not only regulate gene expression by such mechanisms, but can directly regulated gene transcription and promoter activity. R-loops have been abundantly mapped at promoters of RNA polII transcribed gene regions in human [33], [34], [51]–[53]. Concomitantly, it was found that R-loops activates transcription of antisense lncRNAs associated to many protein-coding genes and enhancers genome-wide. In line with this, an overexpression of RNase H1 showed a reduction of R-loop structures at promotor regions leading to a decrease of antisense lncRNAs transcripts [54], [55]. The opposite role of R-loop in modulating epigenetic (allowing both activating or repressing signatures) indicates that R-loops at promoters have double role in activating or inactivating the transcription [17], [33], [34]. R-loops can be an obstacle during transcription elongation, causing collision with the progression of RNA Pol [56], [57]. RNA Pol II pauses or partially blocks transcriptional elongation in the presence of RNA:DNA hybrids located in a unit of transcription [58]. This phenomenon occurs prevalently near promoters, causing gene silencing [59]. Thus suggests that impaired formation of R-loops during transcription leads to RNA Pol pausing or backtracking, affecting gene expressions.

1.2.2. Termination

Genome-wide analysis showed an enrichment of R-loops at G-rich RNA Pol II termination sites [17]. In a normal context, termination of transcription is linked with the cleavage of nascent mRNAs at polyadenylation poly(A) sites. Potential G-quadruplex-forming sequences downstream of the poly(A) site can lead to transcription termination via R-loop formation, an effect that is reverted by overexpression of RNase H [60], [61]. This suggests that R-loops are implicated in transcription termination in natural context. Stalling of RNA Pol II at termination site can promote the pairing of the nascent RNA with its DNA template forming R-loops. It was found that DNA-RNA helicases such as Senataxin and DHX9 are able to resolve R-loop structures formed in termination sites in order to properly terminate transcription [60], [62], [63], thus confirming a central role of R-loop in proper transcription termination.

1.3 R-loop and genomic instability

1.3.1. Genomic instability driven by R-loop

The presence of persistent, unscheduled R-loops represent a risk for genome instability [30], [64], [65]. R-loops have been reported as I) hotspots for mutations driven by reactive oxygen species, nucleases, and other agents that target the displaced single-stranded DNA loop, II) impeding replication fork progression, III) causes of transcription-replication conflicts, leading to replication fork stalling or collapse, known as replication stress or IV) interfering with DNA repair by physically blocking the access of repair enzymes to DNA lesions [64], [66].

1.3.1.1 Hypermutation in ssDNA

The R-loop ssDNA filament was shown to be more prone to mutations due to its increased sensitivity to the action of DNA-damaging agents, but also the action of the activation-induced deaminase (AID) [67]. This process is exemplified by R-loops formed at immunoglobulin S regions where AID was reported to be essential for immunoglobulin Class Switch Recombination and hypermutation in somatic cells [68], [69]. Mechanistically, AID converts dC into dU residues in the ssDNA filament that subsequently triggers the initiation of base excision repair pathway through activation of uracil DNA glycosylase enzyme [70]. This enzyme excises the uracil base to create an abasic site and apurinic/apyrimidinic endonucleases remove the abasic site to create a single strand break (SSB) that will be then repaired by the Poly ADP-ribose polymerase 1 (PARP1) [71]–[73]. Other types of ssDNA modifications, such as oxidation by oxidative stress (8-oxoguanine) or methylation from the methyl-donor S-Adenosylmethionine (SAM) can as well affect R-loop structures, leading to DSB and genomic instability if exacerbated [74].

1.3.1.2 Transcription-replication conflict

The co-temporal activity of two cellular machineries at the same genomic region may cause an encounter, threatening genomic stability. Collision of the replication fork and the transcription machinery causes transcription-replication collision (TRC), leading to replication stress[57]. If the progression of the two complexes occur simultaneously and co-directionally, this might lead to a “mild” collision, causing RNA pol displacement

without significant impact on the DNA [75]. By contract, if moving in convergent directions, the RNA pol and the replication complex can undergo head-on collision, which induce pausing and blockage of the replication fork, that can lead to fork collapse and DNA breaks [64],[76].

Exacerbated transcription or unproper resolution of transcription-associated R-loop can dramatically increase transcription-replication conflict (TRC) by physically impeding fork progression, causing replication fork stalling and damage, best for R-loop hotspots such as telomeres, rDNA regions, and CpG islands, leading to excessive DNA damage and genomic instability [1], [77]–[81].

Noticeable, TCR, along with DNA lesions or secondary DNA structures such as G-quadruplex, hairpins or R-loop *per se*, can lead to replication stress [7], [9], [64], [82]–[85]. Mutations in R-loop resolving factors, such as Topoisomerases or helicases, lead to the accumulation of R-loop mediated DNA breaks, chromosome fragility, also replication fork stalling and TRC [66], [86]–[88]. In support of this vision, overexpression of RNase H1 reduces DNA damage, ameliorating the stress condition [6], [89].

1.3.2. R-Loops and DNA damage signalling

1.3.2.1 ATR and ATM pathway

To protect genome integrity, cells have evolved DNA damage signalling pathways that activates DNA damage response (DDR), used also for resolving replication stress.

The major regulators of the DDR are the ataxia telangiectasia and Rad3-related (ATR) and ataxia telangiectasia mutated (ATM) serine/threonine protein kinases that orchestrate the response to classing double-stranded DNA breaks responses as well as replication stress damage [90], [91]. Studies *in vitro* and *in vivo* suggest that ATR and ATM activation have distinct outcomes in DNA damage responses. ATM is largely activated by DSBs, in contrast, ATR signals a variety of DNA breaks and replication stress conditions [92].

Upon DSBs (also caused by prolonged R-loop mediated replication stress), ATM recruited by the MRE11-RAD50-NBS1 (MRN) complex, phosphorylates histone variant gamma (γ) H2AX at serine 139 and the effector kinase Chk2, leading to cell cycle arrest and DNA repair [87], [93], [94].

Along with sensing and resolving DNA breaks and replication stress, ATR kinase physiologically senses and suppresses R-loop [31], [95]. The R-loop displaced ssDNA is bound by the ssDNA binding by RPA that, upon self-phosphorylation, leading to the recruitment of ATRIP (ATR-interacting protein) and ATR kinase. ATRIP-ATR recruitment leads to activation of ATR by *in trans* phosphorylation of Threonine 1989, allowing the recruitment of the heterotrimeric ring-shaped 9-1-1 complex (RAD9-RAD1-HUS1) and TopBP1[96]. Complex formed activates downstream signalling, such as phosphorylation of Chk1, triggering cell cycle arrest for DNA repair and recruitment of R-loop resolution machineries [90]. In line with this model, inhibition of ATR prevents the recruitment of the RNA helicase DDX19 and Senataxin to R-loops, impairing their resolution [97], [98].

The same pathway is active upon replication fork stalling or DNA breaks [92]. Mechanistically, stalled replication fork at R-loops sites (as well as transcription-replication head-on collision) results in fork reversal, that in turn recruits the MUS81 structure-specific endonuclease that specifically cleaves stalled or collapsed replication forks to solve such structures. The resulting DNA cleavage thus promotes ATR activation, G2/M checkpoint arrest and subsequent fork repair via homologous recombination (HR) pathway [87], [99]–[102]. Remarkably, R-loops and ATR were shown physiologically at centromeres of metaphase centromeres, mediating Aurora kinase B-mediated phosphorylation of histone H3 Serine 10 (H3S10P) to promote chromatin condensation and ensure faithful chromosome. [45], [103].

R-loop driven DNA damage that cannot be repaired in a correct manner ultimately leads to the accumulation of mutations and genome instability [104]. Accordingly R-loops have been linked to functional alteration in oncogenes (such as H-RAS Val-12 mutations) or tumour suppressor genes (e.g. BRCA1/2 mutations) [105]–[107].

1.3.2.2. Repair mechanisms linked to RNA:DNA hybrids

The final outcome of ATR/ATM signaling process is the activation of the DNA damage repair (DDR), that can be achieved through different mechanisms. The two major pathways involved are non-homologous end joining (NHEJ) and homologous recombination (HR) repair Activation of NHEJ results in a quick and error-prone rejoint of DNA ends without need for sequence similarity. By contrast, HR requires a sequence homologous to the one of the DNA break to be activate, but results in an error free repair [92] [108]. To achieve homologous sequence recognition, RAD51 protein recognises

DNA filaments and promotes strand exchange at DSB sites, forming nucleoprotein filaments during HR. Interestingly, the helicase Senataxin promotes RAD51 foci accumulation at DSB sites, regulating γ-H2AX signalling to promote DNA repair resolving R-loop. This allows to reduce abnormal re-joining of DNA breaks suggesting that a correct R-loop removal is necessary to minimize translocations, preserve genome integrity and cells survival following of DSB creation in active genes [109].

Alternatively, the Break induced recombination (BIR) pathway can take place to repair damaged DNA. Upon DSB, a RAD52 dependent DNA strand invasion takes place forming a D-loop, followed by initiation of DNA leading strand synthesis by DNA pol containing POL3D subunit. This will result in migrating replicating bubble and asynchronous synthesis of the lagging strand, resulting in DNA repair. Of note, RAD51 dependency in this process is still controversial [110]–[112].

DSB can also arise from R-loop removal through the transcription-coupled nucleotide excision repair (TC-NER) machinery [113]. It was reported that the Xeroderma pigmentosum group F and G (XPF and XPG) endonucleases are able to excise DNA non-template strand in the R-loops that block transcription machinery, forming a ssDNA lesions that can be transformed into DSBs after replication, that will be then repaired in a RAD52 dependent manner [113]–[115].

Curiously, DNA:RNA hybrids based DNA repair mechanism has been recently identified [101], [116], [117]. Upon DNA damage, DSB sites are transcribed by RNP II producing dilncRNAs (damage-induced long non-coding RNAs). RNAs produced at DNA damage sites were reported to engage in RNA:DNA hybrids formation, promoting DNA damage signalling [118]–[120]. DilincRNAs can also form dsRNA structures and then be processed by the Dicer pathway, generating diRNAs (DSB-induced small RNAs) [121]. diRNAs form a complex with Argonaute 2 and, by recruiting RAD51, promotes its localization to DSBs sites, possibly through diRNA hybridization with DNA [122].

1.3.2.3 When repairing is not worthy: the ALT phenotype

Immortality of cancer cells require the preservation telomeres at chromosome ends to avoid replicative senescence. In the absence of specialized mechanisms for telomere maintenance, linear chromosomes gradually shorten with each round of DNA replication, causing the so called end replication problem [123]. This leads to a progressive telomere shortening, ultimately leading to cellular senescence or apoptosis [124]. 85% of tumors maintain telomere length by upregulating telomerase, an enzyme that adds $(TTAGGG)_n$

repeats to telomeric ends, thus enabling indefinite cell proliferation and tumour progression [125]. However, approximately 15% of cancers, including osteosarcoma, glioma and PanNet cancer, achieve immortality in the absence of telomerase expression through an alternative mechanism known as the Alternative Lengthening of Telomeres (ALT) [126], [127].

ALT involves a distinct type of homology-dependent recombination (HDR) [126] that resembles break-induced replication (BIR), first observed in yeast [128]. ALT entails the invasion of one telomere filament (the recipient) into another (the donor) to achieve elongation. The donor or template DNA strand can be comprised of the sister-telomere, another telomere, or extrachromosomal telomeric DNA, and acts as the source to be replicated [129].

In the context of ALT, R-loops to have a central role in this process. Transcripts originated from RNA pol II driven transcription of the telomeric C-rich strand gives rise to the telomeric repeat containing RNA TERRA that is prone to engage in R-loop formation . Resulting replication stress and double-strand break (DSB) formation, turn stimulates homology-directed repair (HDR) between telomeres, thus directly causing the ALT phenotype[83], [84], [133].

A RAD52-dependent mechanisms facilitates strand invasion , leading to telomere extension through DNA synthesis mediated by PCNA-RFC-Pol δ [134]. Another pathway involves RAD51 in this process[135], [136].

1.4 R-loop regulators

Given the importance of RNA:DNA hybrids in maintaining genome stability and ensuring gene regulation, a notable number of proteins were found to control R-loop formation and resolution (**Figure 2**) [1], [8], [30], [35], [77], [137].

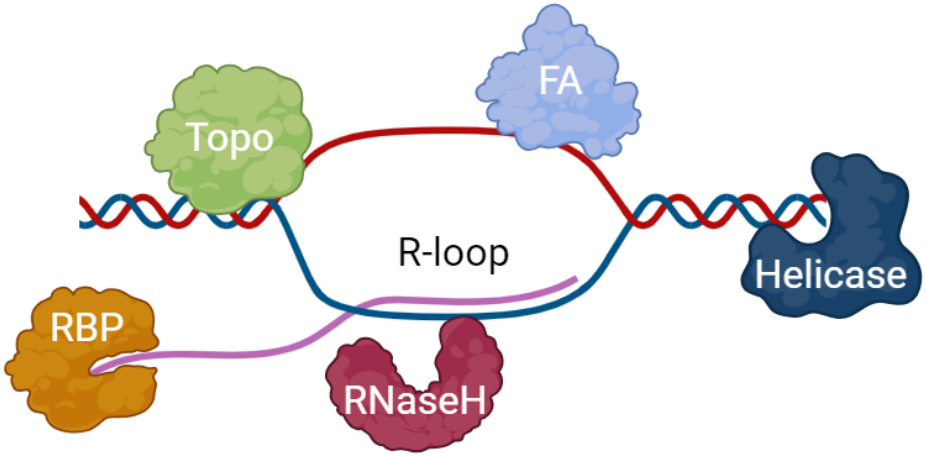


Figure 2. R-loop resolving factors

R-loops can be counteracted through distinct mechanisms. RNase H enzymes that degrade the RNA component of the hybrid structure, by contrast helicases unwind the R-loops during transcription termination. Alternatively, Topoisomerases resolves negative supercoiling to prevent R-loop formation. RNA-binding proteins (RBP) also play a role in preventing R-loop formation, as well as proteins of the Fanconi Anemia (FA) pathways. Image generated using BioRender.

1.4.1 RNase H protein family

RNaseH family proteins are ribonucleases composed of a N terminal RNA:DNA hybrid binding domain (named HBD) and a C terminal RNaseH resolves R-loops by degrading the RNA component of this structure [138]. While RNaseH1 is a monomeric protein, RNaseH2 in a heterotrimeric protein containing the RNaseH2A catalytic subunit, and two additional subunits (RNaseH2B and RNaseH2C), required for biological functions [139].Although RNase H proteins are highly evolutionally conserved and show high grade homology from yeast to human, different roles and mechanisms of action have been identified for the two proteins. RNases H1 requires a substrate with at least four ribonucleotides to be processed and has been found to be implicated in removing RNA primers in genomic as well as mitochondrial DNA [138], [140]. By contrast, RNaseH2 removes the RNA moiety of Okazaki fragments during DNA lagging strand replication and elimination of single ribonucleotide mis-incorporated during DNA synthesis [141].

R-Loops were shown to recruit RNases H1 through the action of the Replication protein A (RPA) complex. RPA works as a sensor of ssDNA, present at DNA damage but also in DNA:RNA hybrids. After binds to the ssDNA filament, the RPA32 subunit gets phosphorylated at serine 33 (RPA32ser33P), activating both DNA damage response and the local accumulation of RNaseH1 at RNA:DNA hybrids *in vitro* [92], [142]. Mutations in RPA32 binding domain of RNase H1 compromise the efficient removal of RNA:DNA hybrids, leading to an accumulation of R-loops and genomic instability. Thus, the interaction of RPA32 and RNase H1 is necessary to prevent R-loop formation, thus suppressing genomic instability [143].

1.4.2 Helicases

In addition to RNase H protein family, proteins with unwinding activity play a crucial role in controlling R-loop levels [144].

This is the case of RNA helicases, that resolve RNA:DNA hybrids by physically displacing the RNA component from the R-loop, thus allowing DNA filaments re-annealing [60]. RNA helicase Senataxin (SETX), mutated in amyotrophic lateral sclerosis 4 (ALS4) patients, and its *S. cerevisiae* counterpart Sen1 are implied in R-loop dependent genome stability and R-loop resolution. *S. cerevisiae* Sen1 helicase-inactive mutants (sen1-1) causes accumulation of R-loop and transcription-associated recombination [145]; this phenotype was also recapitulated in Setx−/− mice [146]. The DXH9 RNA helicase, as well as the historical Bloom's syndrome helicase (BLM) [147] and Werner syndrome (WRN) helicases [148], have been found to interact with R-loop, suppressing them and preventing R-loop-associated DNA damage.

Curiously, another member of the DEAD-box RNA helicase, DDX1, has been found to be able to resolve G-quadruplex structures in RNA molecule, promoting (instead of resolving) hybridization of RNA to DNA strand within R-loop structures to help Class Switch Recombination [149], pointing out the complex and finely tuned regulation of RNA:DNA hybrids. Only few examples were reported, but an increasing number of helicases are being involved in R-loop management and extensively reviewed [150], [151].

1.4.3 Topoisomerases

Topoisomerases, that control DNA topology during replication, transcription and recombination, have been demonstrated to prevent R-loop formation [152]–[154].

The evolutionary conserved Topoisomerase I (TOP1) is involved R-loop management by resolving transcription-associated negative supercoiling, thus preventing pairing of the nascent RNA with the DNA template (i.e. R-loop formation) [153], [155]–[157]. In line with this, loss of TOP1 in yeast and human causes negative DNA supercoiling accumulation ahead of the processing RNA Pol II, facilitating RNA:DNA formation [153], [158]–[160].

Along with TOP1, TOP2 was shown to be involved in R-loop management [160], being able to resolves the negative torsional stress at transcription site and playing important roles in DNA replication, recombination and chromatids segregation by generating DSBs [161], [162]. As many drugs are available for its inhibition and subsequent enzyme-mediated DNA damage, this protein in particular results of great interest for cancer chemotherapy [163]–[165].

As one may expect, DNA Topoisomerase III beta (TOP 3B) too is also able to resolve negative supercoiling that forms during Pol II elongation reducing R-loop formation [166].

1.4.4 Fanconi Anemia

The Fanconi Anemia (FA) protein family are involved in interstrand crosslink resolution, replication fork stability, and DDR activation [167]–[169]. Besides these canonical functions, a part of FA family proteins were seen involved in R-loop management and maintenance of genomic stability [170]–[172]. The central tumorsuppressor BRCA1 (FANCS) has been found to bind to R-loops at transcription termination sites and recruits Senataxin, resolving these atypical nucleic acid structures [173]. Loss of BRCA1 leads to increased R-loops formation and subsequent DNA damage, reflected by γ-H2AX accumulation and ssDNA breaks at these sites [173]. BRCA2 (FANCD1) too has been found regulating R-loops, by binding the branched structure of R-loops, facilitating the recruitment of RNase H1 and SETX, thus protecting the cell from replicative stress and genomic instability [174]. Its depletion causes cause R-loop accumulation and DDR [175]. FANCI, FANCD2, [176] and FANCA [175] (signalling members of the pathway) had been shown to be involved in such a process, as their depletion causes a dramatic increase in R-loop levels and subsequent DNA breaks in various cell models cells [177],

[178]. In particular, the monoubiquitinated FANCI/D2 complex binds to R-loops, representing an important step in R-loops response [176].

FANCM by contrast has been directly found involved in RNA:DNA hybrid unwinding through its ATPase/translocase activity, thus controlling ALT activity of telomerase negative cells [179] FANCR (RAD51) has been recently identified as a R-loop promoter though its ability to bind to ssDNA as well as RNAs species such as telomere repeat containing RNA (TERRA) [180] to mediate R-loop formation *in trans* and to initiate recombination events downstream of R-loop mediated DNA damage. Detailed information on function of FA proteins in R-loop metabolism has been reviewed by [167], [172], [175], [181].

1.4.5 RNA binding proteins

RNA metabolism has been found to be a crucial modulator of R-loop. The yeast and vertebrate THO/TREX complex was the first player to be identified in such a context [182]–[184]. It is formed by the four nuclear proteins Hrp1, Tho2, Mft1 and Thp2, and the three proteins of the TREX complex (Tex1 and the mRNA export factors Sub2 and Yra1): loss of the THO/TREX complex function leads to hyper-recombination effects due to accumulation of R-loop that obstacle RNA Pol II elongation, in addition to defects in transcriptional elongation, mRNA export and recombination [88], [183], [185], [186]. Moreover, it was demonstrated that depletion of the splicing factors ASF1 (also known as ASF/SF2) leads to R-loop formation and subsequent DNA rearrangements mediated by double-strand breaks (DSBs) and genome instability in chicken DT40 cells and human HeLa cells. Indeed, it interacts with the phosphorylated CTD of Pol II to promote RNA splicing but also to prevent R-loop formation and subsequent genomic instability [23], [187].

1.5 Novel R-loop suppressor

1.5.1 Chromatin remodelers / histone chaperon

R-loops cannot be arranged in a nucleosome template and are characterized by increased DNA accessibility of the local chromatin template [188] [34].

To counteract R-loop mediated chromatin aberrations, the cell evolved many chromatin remodelers, such as the SWI/SNF complex, the FACT complex, the INO80 complex, and the SIN3 deacetylase, actively playing a role in R-loop managements and resolution [189]–[193]. These findings postulate chaperon activity containing proteins as strong R-loop suppressors by maintaining a correct chromatin template with normal nucleosome phasing. A novel actor in this scenario is the Death Domain-Associated Protein (DAXX), a multifunctional protein that plays a crucial role in various cellular processes, including transcriptional regulation, apoptosis, and DNA damage response [194]. In recent studies, DAXX has been identified as a specialized histone chaperone that specifically interacts with histone variant H3.3, acting as a DNA replication independent histone chaperon [195]–[198]. DAXX forms a complex with the ATP-dependent helicase and SWI/SNF family chromatin remodeler ATRX (Alpha Thalassemia/Mental Retardation Syndrome X-Linked) that supports the deposition of the H3.3 histone variant containing nucleosomes at specific genomic loci, characterized by constitutive heterochromatic structure, including telomeres and pericentromeric regions, in G1/G2 phases of the cell cycle [195], [197]–[202]. This process is crucial for maintaining chromatin structure and function. In terms of functionality, ATRX is responsible for directing DAXX-dependent deposition of H3.3 at chromatin regions enriched with H3K9me3 through a conserved ATRX–DNMT3–DNMT3L (ADD) domain that directly binds such modifications [195], [203], [204]. Importantly, loss of ATRX was associated with a significant increase in R-loops, anticipating its role in suppressing such structures, with particular relevance for repetitive sequences [200]. Interestingly, ATRX has also been found to bind and suppress G4 structures, known to form R-loop, particularly at telomeres [205], [206]. Loss of DAXX/ATRX has been linked with the collapse of replication fork, mediating DSBs and downstream HR at telomeres and other target sites [200]. Accordingly, ATRX, and less frequently DAXX, mutations strongly associate with ALT phenotype cancers, most frequently observed in PanNET, liposarcomas, adult gliomas, and osteosarcomas [207]–[211], as well as in 90 % of ALT+ cancer cell lines [212]. Strikingly, re-introduction of WT ATRX in ATRXmut ALT+ cells was able to revert the ATL phenotype [213].

Concordantly, expression of WT DAXX in cell expressing a mutated form resulted in suppression of the ALT phenotype [214], underling the interplay between ATRX and DAXX is crucial for maintaining telomere stability, preventing abnormal R-loop structures.

1.5.2 DBHS proteins

SFPQ is an RNA binding protein originally identified as splicing regulator [215] that together with NONO, (Non-POU domain-containing octamer-binding protein), and PSPC1, (Paraspeckle Protein Component 1) belongs to the family of the Drosophila-Behaviour Human Splicing family (DBHS) [216]. In human cells SFPQ plays a crucial role in RNA metabolism by regulating splicing and RNA transport. Interestingly, SFPQ has also been shown to support DNA damage repair[217]. Specifically, SFPQ has the ability to directly bind to DNA double-strand breaks (DSBs) and interacts with RAD51D, a key player in maintaining genomic stability by stimulating homologous recombination-mediated repair [218], [219]. As a consequence, a deficiency of SFPQ results in defects in sister chromatid cohesion, chromosome instability, and heightened sensitivity to DNA-damaging agents [219].

SFPQ and other DBHS family proteins act as homo or heterodimers. Structural analysis show that SFPQ's NOPS, RRMs and coiled coil domains are involved in the formation of heterodimers with NONO [220] as well as homodimers [221]. In fact, it has been reported that SFPQ/NONO heterodimeric complex plays a crucial role in many cellular processes such as transcriptional regulation and mRNA processing [222], [223]. SFPQ, together with other DBHS members aggregate to subnuclear, membrane-less compartments also known as Paraspeckles. Paraspeckles have been shown to contribute to various cellular functions, including cellular stress responses but also the regulation of gene expression [219], [224], [225]. SFPQ/NONO can recruiting the exonuclease XRN2, facilitating the processing of pre-mRNA at the 3'-end and aiding in transcription termination [226]. It has also been observed that the complex localizes to paraspeckles and interacts with other proteins to regulate RNA processing and nuclear retention of specific transcripts [222]. SFPQ and NONO act as transcription regulators and are involved in both repression and activation of gene expression, specifically, SFPQ/NONO can form a bridge between Pol II transcription machinery and other splicing or polyadenylation factors [223]. Concerning the formation of R-loops, it has been demonstrated that in H1299 cells NONO and SFPQ locate at telomeres and have a

common role in suppressing RNA:DNA hybrids and replication defects at telomeres blocking recombination [227]. Since SFPQ and NONO lack direct enzymatic activity, it has been postulated that their role is to facilitate the recruitment of specific factors. This recruitment aims to trigger a response against telomere fragility and to suppress the formation of TERRA:DNA hybrids.

1.6 SFPQ Structure

All DBHS family proteins share a common structure defined as DBHS core, which consists of two RNA Recognition Motifs (RRM1 and RRM2), the DBHS protein family unique sequence NOPS, and a coiled coil domain responsible for polymerization.
Besides these conserved domain, SFPQ possess other extremely important ones, schematically reported in **Figure 3**.

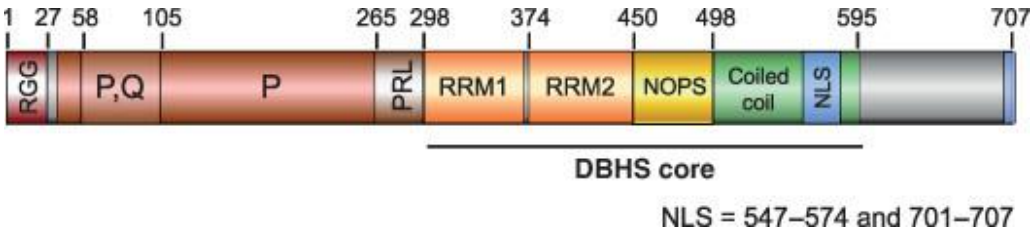


Figure 3. Representation of SFPQ domains

SFPQ N-terminus contains the RGG(arginine/glycine rich), P/Q (proline/glutamine rich) and the P (proline rich) domains. PRL domain separates the P domain from RRM1. The DBHS core consists of the RNA binding motifs RRM1 and RRM2, the NONA/Paraspeckles (NOPS) domain and the oligomerization coiled-coil domain. At the C-terminal region there are two nuclear localization sequences. Picture taken from Yarosh et. al., 2015 [214].

1.6.1 RGG domain

The 27 amino acids SFPQ RGG domain contains a high percentage of arginine and glycine. RGG domains are considered intrinsically disordered, are enriched in RNA-binding proteins and influence RNA substrate specificity [228]. RGG domains are also target sites for arginine methylation, generating novel binding site for other proteins [229]. Notably, RRG domain was demonstrated to be important for 3' end cleavage of mRNA, identifying SFPQ as pre-m-RNA processing factor [230].

1.6.2 Proline rich domains (P/Q and P domains)

The P/Q domain is a 47 amino acids containing more than 40 proline or glutamine residues. The 150 amino acids P domain consist of 66% proline residues. Proline rich domains have been associated with a protein binding function, interacting with strong binding but with low specificity [231].

The P domain of SFPQ has been reported to be directly involved in protein binding, as is the case of Rad51 [218]. Rad51 is involved in homologous recombination, and it promotes the invasion of the homologous template [232]. However, it has been reported that high levels of Rad51 are actually inhibited by SFPQ, preventing deleterious recombination events, thus showing SFPQ dual role in regulating Rad51 activity and subsequent DNA damage [218], [219]. Moreover, the P domain appears to be important for localization of SFPQ to sites of DNA damage. Together SFPQ modulates DRR, non-homologous end joining (NHEJ), homologous recombination, sister chromatid cohesion, and telomere stability [219], [227], [233].

Finally, the SFPQ N-terminus was reported to provide liquid-liquid phase separation (LLPS). It was shown that SFPQ can undergo LLPS at DNA damage sites together with FUS, a RNA binding protein involved in RNA splicing and export, allowing the accumulation of factors related to DNA damage response directly at DNA damage sites [234]. It is interesting to note that the FUS N-terminus, crucial for LLPS, has an RGG and a proline-rich domain like SFPQ, pointing out the possible functional analogy between the two proteins.

1.6.3 RNA Recognition Motifs

RNA Recognition Motifs (RRM) are widespread domains present in RNA-binding proteins. SFPQ contains two consecutive RRMs, separated by a linker. The 70-90 amino acids RRMs domains are folded into two α-helixes that are separated by four antiparallel β-sheets [235]. The canonical protein-RNA-binding relies on aromatic residues that stack together on one β-sheets to favour the interaction with nucleotides. SFPQ RRM1 contains aromatic residues in a proper position to allow canonical RNA binding, while the RRM2 does not display a canonical positioning of aromatic residues [236]. Despite its non-canonical binding, RRM2 activity appears to be more relevant for RNA-binding than RRM1[237]. In particular, SFPQ RRMs have been demonstrated to bind NEAT1, a lncRNA that characterizes paraspeckles subnuclear bodies, forming a scaffold for the binding of SFPQ and other DBHS proteins (namely, NONO and PSPC1). The actual

function of paraspeckles is yet to be fully understood, however there is increase evidence for their role in gene regulation and protein retention upon stress response [222], [225], [238]. SFPQ has been demonstrated to repress Interleukin 8 (IL8) transcription by physically binding to its promoter. However, upon stress induction (e.g. viral infection), NEAT1 overexpression causes the concentration of SFPQ in paraspeckles formation. Lack of SFPQ at the IL8 promoter results in gene activation and inflammatory response. Notably, SFPQ version lacking RRM1 or RMM2 maintain IL8 suppression [239]. SFPQ RRM domain have been also identified as crucial in mediating an anti-apoptotic alternative splicing of caspase 9 mRNA [227].

1.6.4 NOPS and Coiled-coil Domain

The NONA/Paraspeckles (NOPS) domain is unique for all three members of the DBHS family and is essential for homo and heterodimerization by binding to the RRM2 domain of another DBHS family member [240]. A recent study has revealed that SFPQ prefers to heterodimerize rather than binding to another SFPQ monomer [241]. However, DBHS proteins have a differential expression in human tissues, eventually influencing complex formation.

The Coiled-Coil Domain is shared between all the proteins of the family as well, and in contrast to the NOPS domain allows protein homo or hetero polymerization [236] and positioning into paraspeckles [242]. Thus, this domain is important for higher level structural organization targeting to paraspeckle.

1.6.5 C-terminal domain

The C-terminal domain contains two nuclear localization sequences and represents a flexible region due to the enrichment of glycine residues. Moreover, the C-terminus can be a site for post translational modifications (PTMs), such as phosphorylation, that influence SFPQ binding to other proteins [243].

1.7 R-loop at repetitive regions

The human genome contains up to 75% of repetitive DNA elements, predominantly represented by transposable elements, (peri)centromeric repeats, telomere repeats and rDNA arrays[244], [245]. The majority of these elements have the capacity to undergo transcription, anticipating a role for R-loop regulation in controlling genome stability, transcription and chromatin status at such regions. This is supported by increasing evidence that involve R-loops in human diseases characterized by genetic or epigenetic alterations in repetitive DNA elements [246]

DNA repeats can be divided in two major groups. Variable Number of Tandem Repeats (VNTRs) that contain microsatellite (1-10bp) and minisatellite (10-100bp) repeats and Small-Scale Repetitive Elements (SSREs) ranging from 0,1 – 8kb. Hallmark examples for VNTRs are SatI, SatII and SatII satellite DNAs, ranging from 5bp (unit length of SatIII repeats) to 25 bp (unit length of SatI repeats), pericentromeric β-satellite (68-69 bp) and telomere (6 bp) repeats, but also disease related tri-nucleotide repeat expansions [246], [247]. The group of SSREs DNA comprise centromeric α-satellite (171bp) repeats, γ-satellite (220 bp) repeats, transposons (ranging from 100 to 10.000 bp in length, composing 2,8% of human genome), LTR retroelements (0.2-3 kb unit length) that compose the 8,3% of the human genome, and non-LTR retroelements such as LINEs (6–8 kb unit length) and SINEs (0.1–0.4 kb unit length) elements that together make up approximately 34% of the human genome[248], [249]. The indicated types of satellite repeats build eukaryotic centromeres and pericentromeres , but can also be located at subtelomeric sites and at interstitial regions[250], [251]. In addition, the human genome contains 200-600 copies of ribosomal DNA (rDNA, 43kb unit length) and 500 interspersed tRNA genes [246], [247].

Pioneering RNA:DNA hybrid mapping experiments in yeast using ChIP-seq or DRIP followed by hybridization to tiling microarrays provided first evidence for R-loop formation at tandem or dispersed repetitive elements, including tRNA genes, retrotransposons and telomere repeats [252], [253]. Genome-wide DRIP-seq combined with detection of nascent transcripts by Global Run-On sequencing (GRO-seq) allowed to investigate transcriptional activity and R-loop formation at repetitive elements in different model systems [38]. In U-2 OS cells, R-loops were found enriched at telomeres and centromeres, as well as at simple/low complexity repeats, and slightly enriched at satellite repeats and rDNA. Notably, no enrichment was detected at retroelements.

In contrast, studies using *A. thaliana* showed high R-loops enriched at all types of transposable elements, with moderate enrichment at centromeres and simple/low complexity repeats. Curiously, both in human cells and *A. thaliana*, a great overlap between R-loops and nascent transcript (ranging from 60% to 70%) was found, indicating R-loop generation *in cis* [38].

Moreover, data from *D. melanogaster* revealed that R-loop pattern varies during differentiation. In *Drosophila* embryos, transposable elements and satellite repeats where enriched for R-loop compared to Schneider 2 (S2) cells [38]. In contrast, S2 cells presented R-loop accumulation at simple/low complexity repeats. In line with this, preventing R-loop degradation by overexpression of catalytically death RNase H1 resulted in hatching defects in Drosophila embryos [254].

These results strongly indicate that R-loop expression at repeat regions depend on genome sequence features but also on species, developmental status and biological context, such as the cancer setting in U-2 OS cells [38].

1.7.1 R loop at (peri)centromeres

Human (peri)centromeric regions are characterized by many satellite repeats distributed in different tandem repeats.

All human centromeric regions are enriched in repetitive sequences named α-satellites, where the CENP-A nucleosomes resides. These A-T rich regions are organized in ahead-to-tail tandem repeat pattern of 171 bp monomer unit [255] giving rise to higher order repeat (HOR) unit [256]. Each HOR is repeated hundreds-to-thousands of times, producing 2-5 Mb-long arrays with chromosome specific characteristics, varying in number and sequence of the α-satellites monomers as well as HOR size [256], [257].

The flanking pericentromere region is formed by β, γ, I, II, and III satellites tandem repeats (5-200bp) located on different chromosomes [258]–[266], also containing LINE, SINE and retroelements in a more relaxed arrangement [267].

During eukaryotic mitosis, Aurora kinase A (AURKA) promotes centrosome maturation and spindle assemble, while the Aurora kinase B (AURKB) and C (AURKC) regulate chromosome condensation, attachment to kinetochores and the correct alignment of metaphase chromosomes [268]. A key step during mitosis is Aurora B kinase mediated phosphorylation of histone H3 Serine 10 (H3S10P), facilitating chromosome condensation [269]–[271].

Human centromeres and pericentromeres were demonstrated to give rise to non-coding RNAs (ncRNAs) in an RNAPII dependent manner throughout the cell cycle, involved in regulating proper centromere structure and function[272]–[275].

The high rate of transcription and repetitive nature identify (peri)centromeres as candidate hot-spots for R-loops. In line with this, human R-loop mapping revealed that up to 50% of (peri)centromeric sequences are enriched for R-loops [38]. Remarkably, forced resolution of R-loops by ectopic overexpression of RNase H1 leads to defects in centromere cohesion and chromosome segregation during mitosis [276].

First in yeast and then in human, R-loop were found increased at centromeres of mitotic chromosomes, and their dysregulation lead to defects in kinetochore function and chromosome segregation [45].

On the mechanistic level, R-loops were shown to stimulate the recruitment of AURKB to centromeric and pericentromeric sequences, facilitating the phosphorylation of H3S10, promoting chromatin condensation and mitotic fidelity. Accordingly, overexpressing RNase H1 reduces H3S10 phosphorylation, driving genome instability [45]. In human cancer cells, centromeric R-loops were linked to a function for ATR in maintaining mitotic fidelity. Centromeric R-loop were shown to recruit the ss-DNA binding RPA32 in mitosis, activating ATR that in turn stimulates AURKB via Chk1, preventing the formation of lagging chromosomes [103]. Moreover, an $ATR^{-/-}$ cell model presented reduced H3S10 phosphorylation, further demonstrating the requirement of ATR for full activation of AURKB at centromeres [103]. This regulatory pathway suggests that R-loops play an integral part in mitotic control and must by subjected to tight regulation to avoid genome instability.

1.7.2 Telomeres

Telomeres are heterochromatic nucleoprotein structures located at the end of each chromosome that ensure genome stability by protecting these regions from end-to-end fusions and uncontrolled recombination events [277], [278]. Vertebrate telomeres are composed of 5'-TTAGGG-3' microsatellite tandem repeats, flanked by upstream subtelomeric regions that contain different arrangements of heterogeneous sequence repetitions [279]. Telomeres are bound by the Shelterin protein complex that ensures telomere function and telomere length homeostasis [280], [281]. Even though they present constitutive heterochromatin that protect telomeres from un-licensed recombination, rearrangements and controls repeat length [282], [283], vertebrate

subtelomeric promoters drive the expression of a long, (UUAGGG)n telomere tandem repeat containing non-coding RNA (TERRA) [131], [132], [284]–[287]. Importantly, the G-rich content of TERRA transcripts promotes the formation of R-loops in cis or in trans at eukaryotic telomeres[133], [180], [288]–[290]. In line with this, genome wide mapping analysis based on DRIP-seq combined to GRO-seq revealed that telomeres are characterized by a remarkable high R-loop enrichment [38].

R-loops at telomeres are kept under control by redundant pathways to reduce the risk for replication-transcription conflicts, DNA breaks and uncontrolled homologous recombination that may drive genomic instability [180], [291]. Remarkably, telomeric R-loops were demonstrated to fuel recombinogenic substrates that engage in homologous homology-directed repair (HDR) to maintain telomere repeats[133], [180], [292]. Telomerase negative tumors are indeed characterized by elevated R-loop levels at telomeres to drive the BIR dependent, Alternative Lengthening of Telomeres (ALT) pathway, assuring telomere maintenance and replicative immortality[291], [293]–[296]. In this context, RAD51 has been demonstrated to specifically bind TERRA RNA, mediating its insertion into the double-stranded telomeric DNA, forming telomeric R-loop [180], [297], [298]

1.7.3 R-Loops at transposable elements

Transposable elements (TEs) are present almost half of the human genome and , when activated, hold the ability to duplicate and insert in a new position within a genome [299]. Class I TEs, named retrotransposons, transpose by using a RNA intermediate, while class II TEs, called DNA transposons, move directly as DNA sequences [300]. Class I TE are mostly represented by two major type of interspersed element, called LINE (Long Interspersed Nuclear Elements) and SINE (Short Interspersed Nuclear Elements) [301]. To ensure genomic stability, redundant mechanisms ensure the suppression of TE transposition, such as RNA interference (RNAi), the Piwi-interacting RNA (piRNA) pathway, and epigenetic silencing [302]–[304]. LINE and SINE activity support genetic diversity and drive evolution by generating novel mutations and genome rearrangements. However, upon defective TE suppression, drives transposition events that can cause DNA damage, mutations, and genomic instability, being ultimately involved in the formation of cancer and other diseases, such as Aicardi-Goutières syndrome (AGS) and Amyotrophic lateral sclerosis (ALS) [305]–[310]. On top of this, LINEs and SINEs can

be transcribed, give rise to lncRNAs that interfere with a wide range of cellular processes, including gene regulation and chromatin organization [311].

Recently, transcriptional activation of TEs was found to coincide with R-loop enrichment in different organisms [38]. *S. cerevisiae* strains lacking RNH1 or Top1 showed increased R-loop levels at TY1 retrotransposons, promoting transposition frequency [252]. In *S. pombe*, R-loop at retroelements were linked to chromatin regulation, recruiting the RNA-induced transcriptional silencing (RITS) complex, leading to H3K9me3 deposition and constitutive heterochromatin formation [46], [312].

To this end, direct functional evidence of R-loop at transposable elements in vertebrates is limited. However, various factors such as ATRX, DAXX, BRCA1 and FANCD2 were recently demonstrated to control the activity of transposable elements, providing a link to R-loops [201], [313]–[317]. Interestingly, recent works linked LINE activation, mutations in R-loop modulators (e.g. RNaseH2) and activation of the cGAS/STING pathway, further supporting the involvement of retroelements in maintaining genomic stability [318]–[326].

1.8 R-loops connect to innate immunity via the cGAS-STING pathway

1.8.1 Inflammation counteracts or promotes cancer progression

Inflammation is a highly conserved biological process that encompasses the activation, recruitment, and functional involvement of both innate and adaptive immune cells [327]. Innate immunity is the first line of defence of the immune system, providing immediate and non-specific protection against pathogens, involving the recruitment and activation of cellular components such as macrophages and neutrophils that clear infected cells. The innate immune system relies on a limited set of receptors to detect pathogens but compensates by targeting a wide variety microbial components (such as viral nucleic acids), initiating a protective inflammatory response within minutes of pathogen exposure. Furthermore, innate immunity plays a central role in activating the adaptive immune response [328], [329].

While initially recognized for its critical function in defending the host against pathogens, inflammation holds equal significance in facilitating tissue repair, regeneration, and remodelling, thus delicately balancing tissue homeostasis [330], [331]. In recent decades, there has been a significant resurgence of interest for the link between inflammation and cancer, with inflammation playing a significant role in various stages of cancer development and progression by supporting cancer cells growth, survival ,and angiogenesis [332], [333]. Conversely, the inflammatory response can activate immune system to recognize and eliminate cancer cells. However, cancer cells can develop mechanisms to evade immune recognition and attack, leading to immune escape [332]. These findings clearly show that inflammation to act both as anti or pro tumorigenic factor, according to the environment and cancer type [334].

Nevertheless, the idea to use the immune system to combat cancer, with the advent of immunotherapy, has emerged as a promising avenue in cancer treatment [335]. This has led to a substantial surge in research efforts focusing on unravelling the molecular mechanisms that link inflammation and cancer. By understanding these intricate mechanisms, scientists aim to refine and optimize immunotherapeutic approaches for improved cancer management.

1.8.2 Innate immunity is controlled by the cGAS/STING pathway

The cGAS-STING pathway is central for the activation of innate immunity (**Figure 4**) [336]. cGAS serves as cytoplasmatic sensor for various forms of cytoplasmic dsDNA species originating from virus, bacteria, mitochondria, or other types of genomic DNA [337].

Notably, the C-terminal region of cGAS possesses nucleotide transferase activity, which enables it to bind to dsDNA and induce significant conformational changes, particularly in its catalytic pockets [338], [339]. This binding event allows the utilization of ATP and GTP as substrates for the synthesis of the cyclic dinucleotide molecule cyclic guanosine monophosphate-adenosine monophosphate (cGAMP) [340]. Acting as a secondary messenger, cGAMP is then detected by STING, a transmembrane protein that is primarily located in the endoplasmic reticulum (ER) [339], [341]. Upon activation by cGAMP, STING translocates to the Golgi apparatus via the ER-Golgi intermediate compartment [342], [343]. After reaching the Golgi, STING is palmitoylated at two cysteine residues (Cys88 and Cys91) and recruits TANK-binding kinase 1 (TBK1), which in turn phosphorylates the C-terminal of STING, allowing the recruitment of IFN regulatory factor 3 (IRF3). Furthermore, STING activation can bind and active IkB kinase IKK to trigger NF-kB transcriptional activation [341]. The activation of the two transcription factors IRF3 and NF-kB (p50/p65) subsequently induce the expression and secretion of pro-inflammatory cytokines such as IFN-β and IL-1 [344]. Pro-inflammatory cytokines have a broad spectrum of actions, including expression of vascular endothelial receptors necessary for immune cell migration, and activation of macrophages and neutrophils to aid in the process of destruction [345].

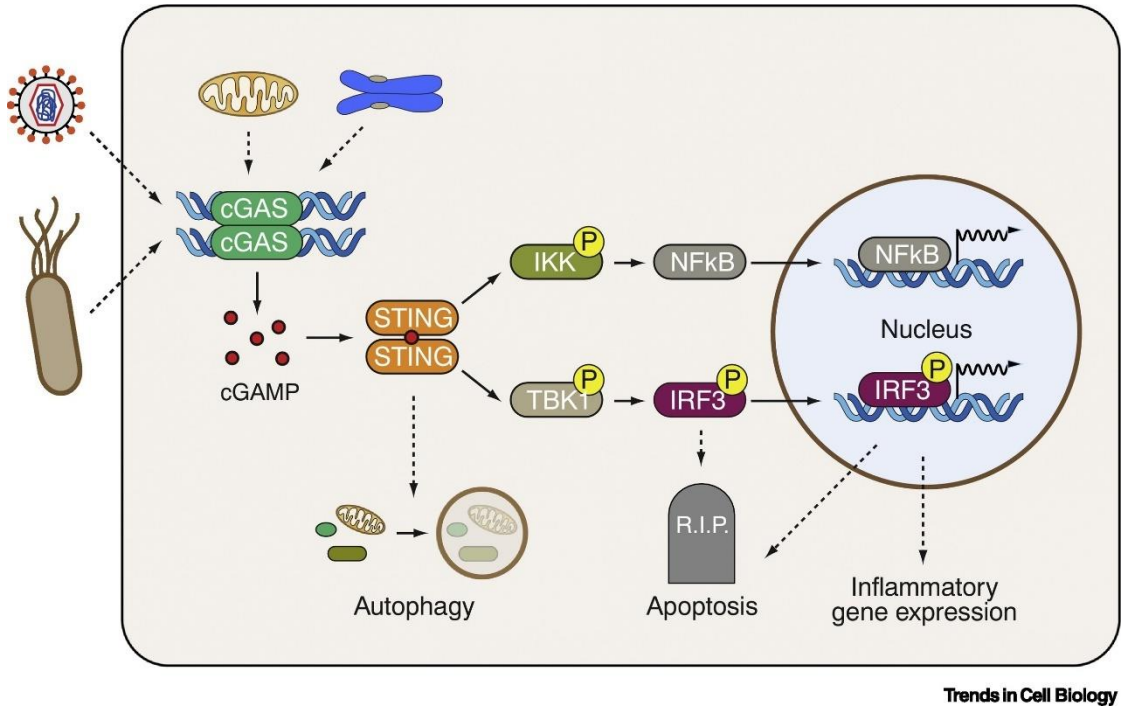


Figure 4. Schematic representation of the cGAS/STING pathway.
cGAS is traditionally activated by viral or bacterial DNA, but it can also respond to self-DNA such as chromosomal DNA or mitochondrial DNA. Upon DNA binding, cGAS becomes enzymatically active and produces cyclic GMP-AMP (cGAMP) synthesis. cGAMP then binds to and activates STING, which in turn activates inflammatory protein kinases (IKK and TBK1) and stimulate nuclear factor kappa B (NF-kB) and interferon regulatory factor 3 (IRF3) by TBK1, activating transcription of inflammatory genes. Image taken from Zierhut et. al., 2020 [332]

1.8.3 Factors contributing to the accumulation of DNA in the cytoplasm

In the context of tumour cells, exposure to various stress, such as metabolic stress, mitotic stress, oxidative stress, and DNA damage stress renders both nuclear and mitochondrial DNA vulnerable to damage and deletion of DNA elements. These conditions were subsequently reported to lead to the production of cytoplasmic DNA accumulation and cGAS/STING pathway activation [336]. A classic source of cytoplasmic DNA comes are resented by so-called "micronuclei", small nuclear-like bodies composed of chromosome fragments wrapped in fragile nuclear membranes. Micronuclei are traditional biomarkers of DNA damage and chromosome instability that can activate the cGAS/STING pathway after the loss of their membrane component [346]. Micronuclei can originate during anaphase from lagging acentric chromosome or chromatid fragments caused by mis-repair of DNA breaks or unrepaired DNA breaks [347]. Additionally, several key players in the control of genome stability are involved in the activation of the cGAS/STING pathway by cytoplasmatic DNA species. For example, abnormal DNA structure at stalled replication forks can be cleave by

endonuclease MUS81 resulting in the build-up of DNA in the cytoplasm [348]. Induction of DNA DSBs following ionizing radiation (IR) or treatment with chemotherapeutic can also activate the cGAS/STING signalling pathway [336]. Interestingly, recent studies have demonstrated a link between unscheduled R-loop formation and cytoplasmic DNA accumulation resulting in activation of the cGAS/STING pathway. For instance, the disruption of R-loop homeostasis due to the loss of SETX or BRCA1 gene has been associated with cytoplasmic R-loops accumulation that bind to the pattern recognition receptors cGAS and activating IRF3 that inducing apoptosis [349]. TOP1 poisons increase micronuclei levels with a mechanism involving R-loops and activate the cGAS/STING pathway leading to increased expression of immune genes in HeLa cells. Overexpression of RNaseH1 markedly reduces micronuclei level [350].
These results establish that RNA-DNA hybrids are immunogenic species that aberrantly accumulate in the cytoplasm after R-loop processing, linking R-loop accumulation to the innate immune response, a new insight that could have positive implications for cancer therapies.

2 Aim of the thesis

Unscheduled R-loop formation is one of the major sources of genomic instability, thus the underling mechanisms should be widely unveiled. The RNA-binding protein SFPQ has been reported to suppress R-loop at telomeres, preventing DNA damage and recombination. However, his action is not limited to these sites, but expands genome-wide. With this work I aimed to investigate the R-loop management function of SFPQ and its novel interacting partner DAXX on a genome-wide level.

In particular, my Ph.D. thesis aims:

- To validate and dissect the interaction between SFPQ and its novel interactor DAXX
- To investigate R-loop repressing action of SFPQ and DAXX at telomeres and at other repetitive regions genome-wide
- To investigate the downstream effect of un-managed R-loop formation upon SFPQ depletion, in terms of genomic instability and innate immunity

By performing such investigations I will provide evidence for an SFPQ mediated, R-loop dependent activation of innate immunity that shall play an important role in sarcoma treatment.

3 Material and methods

3.1 cell lines and culture

Human cell lines used were obtained from ATCC and have not been cultured for longer than 6 months. All cells were cultivated at 37°C, 5 % CO2. U-2 OS (osteosarcoma) cells were cultured in low glucose Dulbecco's modified Eagle's (DMEM) medium (Euroclone) with 10% foetal bovine serum (Corning), 1% L- glutamine (Gibco), 1% penicillin/streptomycin (Gibco). H1299 (carcinoma; non–small cell lung cancer) cells were cultured in Roswell Park Memorial Institute (RPMI) medium (Euroclone) supplemented with 10% foetal bovine serum (Euroclone), 1% L-glutamine (Gibco), 1% penicillin/streptomycin (Gibco). OVACR4 (ovarian; non–small cell lung cancer) cells were cultured in Roswell Park Memorial Institute (RPMI) medium (Euroclone) supplemented with 10% foetal bovine serum (Gibco), 1% L-glutamine (Gibco), 1% penicillin/streptomycin (Gibco). HEK293 (human embryonic kidney) cells were cultured in high glucose Dulbecco's modified Eagle's (DMEM) medium (Euroclone) supplemented with 10% foetal bovine serum (Euroclone), 1% L-glutamine (Gibco), 1% penicillin/streptomycin (Gibco).

U-2 OS inducible cell lines were gently provided by Prof. Sébastien Britton (IPBS, Toulouse) and cultivated as canonical U-2 OS cells, with the addition of Puromycin (Sigma) 0.25 µg/ml. Doxycycline induction was performed with 2 µg/ml doxycycline (Sigma) in complete medium for 24 h.

U2-OS cells stably expressing mutated forms of SFPQ or the single P domain were cultivated as canonical U-2 OS cells, with the addition of Puromycin (Sigma) 0.5 µg/ml.

3.2 siRNAs and vectors transfection, retroviral transduction, stable cell lines production

For siRNAs transfection, RNAi-MAX Lipofectamine (Invitrogen) was used according to the manufacturer's suggestions.

siRNAs have been transfected at a final concentration of 30 nM for 72 hours.

siRNA used to transiently transfect cells are the following:

Human, ON-TARGETplus CONTROL siRNAs (Dharmacon)

ON-TARGET plus smartpool SFPQ siRNAs (Dharmacon)

SFPQ 5'UTR siRNA (MWG) 5'- CCACGUUUCCUGAGCGUCU(TT)-3'

RNase H1 siRNA (MWG) 5'- ACAAACCAAAGAGCGGAAAUUCAUG(TT)-3'

DAXX siRNA (MWG) 5'- GGAGUUGGAUCUCUCAGAA(TT)-3'

To generate stable cell lines stably expressing SFPQ mutants, U 2-OS cells were co-transfected with 1 µg of linearised desired vector carrying Myc-tagged SFPQ, related mutants or Myc empty control vector, and 0.1 µg of linearised pLPC vector carrying puromycin resistance. Cells were transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer's suggestions. 48 h after transfection, Puromycin 0.25 µg/ml was added, and subsequently increased to 0.5 µg/µl to completely select stable cells.

For lentiviral transduction HEK 293 were used as packaging cell line.

Reaction mix containing DNA was prepared adding:

10 µg Backbone

7.5 µg psPAX2

2.5 µg pMD2-env

50 µl Polyethylenimine (PEI, 1mg/ml)

OptiMEM to 950 µl

After 10 minutes incubation RT, the mix was added drop by drop on a 10 cm^2 dish pre-filled with 9 ml complete medium and left overnight in incubator.

The day after, medium was replaced by 6 ml fresh medium and left 72 hours.

Subsequently, the supernatant containing the virus was collected, centrifuges and filtered.

1.5 ml aliquots were stored at -80° for future use or immediately used for cell transduction. For this purpose, 15 µg of 10 mg/ml Polybrene (Sigma) was added, then the mixture was put on a 3 cm^2 dish with recipient cells. After 24 hours, Puromycin 0.25 µg/ml was added, and subsequently increased to 0.5 µg/µl to completely select stable cells.

3.3 Immunofluorescence

Cells were washed with 1X PBS and fixed in 4% paraformaldehyde (PFA) for 15 minutes. Subsequently, cells were washed with 1X PBS and then treated with Citrate Buffer (0.1% sodium citrate, 0.05% Triton X-100) for 5 minutes at room temperature and washed with 1X PBS, 0,1% Tween-20. Cells were blocked for an hour in Blocking Solution (3% BSA/1X PBS, 0.1% Tween-20). Primary antibodies (antibodies table) were diluted in blocking solution and slides were incubated for 2 hours at room temperature in a wet chamber or overnight at 4°C for pIRF3 antibody. Cells were washed three times in washing solution (0.3% BSA/1X PBS, 0.1% Tween-20) for 5

minutes and incubated with Alexa Fluor secondary antibodies (Invitrogen, 1:500) diluted in washing solution for an hour in a wet dark chamber. After incubation, slides were washed two times in washing solution for 10 minutes. After that, DAPI (Sigma) was added in order to stain the nuclei and slides were incubated for 5 minutes. Slides were mounted with ProLong (Invitrogen). For S9.6 immunofluorescence, cells were fixed and permeabilized with ice-cold methanol for 10 minutes and ice-cold acetone for 1 minute, respectively, on ice as previously described (Bhatia et al., 2014). Blocking, antibody and washing solutions were performed in 4X SSC. Images were captured using classic immunofluorescence microscope (Leica DM4000B).

For classic immunofluorescence analysis, the number of co-localizations was counted by manual inspection. Quantitative immunofluorescence analysis were performed using ImageJ. The Student's t-test was used to calculate the statistical significance. For signals intensity of interphase nuclei was analysed using ImageJ. Student's t-test was used to calculate the statistical significance.

3.4 Co-Immunoprecipitation

Cells were lysed in Lysis Buffer (50 mM Tris pH 8, 150 mM NaCl, 1% NP-40, 5 mM EDTA, 5% glycerol) supplemented with 1mM PMSF, 1X PIC and 1mM NaF (proteases /phosphatases inhibitors) and passed through a small siring several time to ensure proper nuclear disruptor. Between 800 ng and 1000 ng of toral protein material was used per IP. Desired antibody were added to the protein extract. The mixture was left rocking O/N at 4°. The following day, 25 µl of Dynabeads Protein A (Invitrogen) were used for each IP to capture ab-antigen interaction 2 h at 4°C, rocking. 4 washes in Lysis Buffer were performed. The sample was then eluted in Sample Buffer 2X (125 mM Tris pH 6.8, 0.5% SDS, 10% glycerol, 5% 2-mercapto-ethanol) and the protein-protein interaction was investigated by Western Blot.

3.5 Chromatin immunoprecipitation (ChIP)

U-2 OS cells transiently transfected for 72 hours with siRNAs for SFPQ ad control were fixed by adding Formaldehyde Solution (Sigma) in PBS 1X, to a final concentration of 1%. Cells were shaken for 15 minutes at room temperature. Cross-linking was stopped by incubation with glycine to a final concentration of 0.125 M while shaking for 5 minutes. Cells were washed twice with cold 1X PBS and collected by scraping with ice-cold PBS 1X, 1mM PMSF, 1X PIC and 1mM NaF. Cells were

centrifugated at 4000 rpm-4C for 5 minutes and resuspended in Lysis Buffer I (50 mM HEPES pH 7.5, 10 mM NaCl, 1mM EDTA, 10% Glycerol, 0.5% NP-40, 0.25% Triton X-100), 1mM PMSF, 1X PIC and 1mM NaF, then rocked for 10 minutes at 4°C and pellet was resuspended in Lysis Buffer II (10 mM Tris-HCL pH 8.0, 200mM NaCl, 1mM EDTA, 0.5 mM EGTA) 1mM PMSF, 1X PIC and 1mM NaF and centrifugated 5 minutes at 4°C. Pellet was then resuspended in Lysis Buffer III (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1mM EDTA, 0.5 mM EGTA, 0.1% Na-deoxycholate, 0.5% N-lauroylsarcosine) 1mM PMSF, 1X PIC and 1mM NaF. Sonication was performed with BioRuptor (Diagenode) to obtain DNA fragments between 150-300 bp. Sonicated lysates were centrifugated for 15 minutes at 4°C and supernatant contain sonicated DNA was collected into a new Eppendorf tube. Sample was diluted with Equilibration Buffer (10mM Tris-HCl pH 8.0, 100mM NaCl, 1mM EDTA, 1.66% Triton X, 0.166% Na-deoxycholate) 1mM PMSF, 1X PIC and 1mM NaF to reach a final volume of 900 µl for each IP and Antibody was added to the solution and rocked O/N at 4°C. 25 µl of Dynabeads Protein A (Invitrogen) were used for each IP upon O/N blocking solution at 4°C with Beads Blocking Solution (0.5 % BSA, 5 mM EDTA in 1X PBS). Beads antibody conjugation has been performed washing and resuspending beads in equilibration buffer with inhibitor, next antibodies conjugated with beads were mixed with samples and incubated at 4°C in rock for 1 hours. Elution of the immunocomplex was performed by adding 250 µl of Elution Buffer (10mM Tris pH 8.0, 300 mM NaCl, 5 mM EDTA, 0.5% SDS) and incubated at 37°C under shaking at 800 rpm for 15 minutes. Magnetic rack was used to collect the eluted material, elution was performed twice. Revert Crosslink was performed adding 20 µl of 5M NaCl at 65°C O/N. DNA purification was performed adding 10 µl of 0.5M EDTA, 20 µl Tris-HCl pH 6.5-7.4, 2 µl RNase A (Invitrogen, 10 µg/µl), 2 µl Protease K (Invitrogen, 20 µg/µl) to each sample and incubated 1 hour at 45°C. After incubation PCI was performed and samples were resuspended in 30 µl of TE 1X. Samples were analysed by q-PCR.

For ChIP-seq experiment, sample were purified with PCR purification kit (Qiagen) and sent to Macrogen for library preparation and sequencing.

3.6 RNA:DNA hybrids immunoprecipitation (DRIP)

U-2 OS cells transiently transfected for 72 hours with siRNAs for SFPQ ad control were scraped with cold PBS 1X and centrifuged at 4000 rpm for 5 minutes at 4°C. Samples were lysed in 1:1 of Lysis Buffer (1% SDS, 20 mM Tris-HCl pH 7.5, 40 mM EDTA pH 8.0, 100mM NaCl, ddH_2O) and TE buffer (100 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0) supplemented with Proteinase K (150 ug/ml final concentration, Invitrogen) at 37°C O/N and purified by phenol-chloroform extraction, ethanol precipitation and resuspend in Elution buffer (5 mM Tris-HCl pH 8.5). The purified nucleic acid preps were sonicated to obtain chromatin fragments size of 300-500 pb and quantified. Part of samples was kept 1 ug was taken as Input (IN), 4 µg were stored for immunoprecipitation while other 4 µg were incubated with RNase H (5000U/mL NEB) at 37°C overnight as a negative control. Magnetic Beads were pre-blocked with 5mM EDTA/PBS containing 0.5% BSA O/N rocking at 4°C. Magnetic beads were washed two times for 5 minutes rocking at 4°C with IP buffer (50 mM Hepes/KOH pH 7.5, 0.14 M NaCl, 5 mM EDTA pH 8.0, 1% Triton X-100, 0,1% Na-Deoxycholate) and resuspended in IP buffer. Magnetic beads were incubated with S9.6 antibody (Sabbioneda (Homemade) 1µg for every 2 µg DNA) for 4 hours rocking at 4°C this sample was added the sample and incubated overnight rocking at 4°C. Afterwards, beads were recovered and washed once with following buffer: IP buffer, HIGH SALT buffer (50 mM Hepes/KOH pH 7.5, 0.5 M NaCl, 5 mM EDTA pH 8.0, 1% Triton X-100, 0,1% Na-Deoxycholate), Wash buffer (10 mM Tris-HCl pH 8.0, 0.25M LiCl, 1 mM EDTA pH 8.0, 0.5% NP-40, 0,5% Na-Deoxycholate), TE (100 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0) at 4°C with rotation. Elution was performed in 500µl of Elution Buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 1% SDS) for 15 minutes at 65°C. The samples were then purified using standard phenol/chloroform and ethanol precipitation. Samples were analysed by q-PCR.

3.7 RNA extraction and retro-transcription

Cells were washed with 1X PBS, then 500 µl of Trifast (Euroclone) were added directly on 3 cm^2 cell dish , pipetted up and down until smooth and left 5' at room temperature. After transferring the sample into an eppendorf, 100 µl of chloroform (Sigma) were added, the sample were then shaken for 15'' and left 5' at room temperature. After 10 minutes centrifugation 13,000 rpm, 4°C, the transparent upper phase has been collected in a new Eppendorf tube. An equal amount of Isopropanol

(Sigma) was added and the solution was mixed until combined, then put at -20° for at least 30 minutes. After 30 minutes centrifugation at 13,000 rpm, 4°C, supernatant has been discarded and two washes with 500 µl of 70% EtOH has been performed. After discarding supernatant, samples were left drying for 40-45 minutes under chemical hood. Samples were resuspended in 20 µl of H_2O. Reverse transcription was performed by using QuantiTect® Reverse Transcription Kit (Qiagen) following manufacturer's instructions. For RNA-seq experiment, samples were collected by using Norgen Biotech RNA extraction Kit and sequenced by Macrogen.

3.8 RealtTime-PCR

Quantitative PCR (RT-qPCR) was performed using 500 nM of specific primers in iTaq™ Universal SYBR® Green Supermix (BioRad) with CFX Connect Real-Time PCR Detection System (BioRad) following manufacture instruction. Specificity of PCR products was routinely checked with melting curves and agarose gel electrophoresis. PCR primers are reported in Table 2.

3.9 IF-DNA FISH

Cells were washed with 1X PBS and fixed in 4% paraformaldehyde (PFA) for 20 minutes at 4°C. Subsequently, cells were washed with 1X PBS 0.1% Triton X-100 at room temperature for 7 minutes. Cells were blocked with PBS1X BSA 5% at 37° for 20 minutes. Primary antibodies (antibodies table) were diluted in blocking solution (PBS1X BSA 5%) and slides were incubated O/N at 4°C in a wet chamber. Cells were washed twice in PBS 1X, 0,05% Triton-X-100 for 5 minutes and incubated with Alexa Fluor secondary antibodies (Invitrogen, 1:500) diluted in blocking solution for an hour in a wet dark chamber at room temperature.

After incubation, slides were washed twice in PBS 1X solution for 7 minutes. For the DNA-FISH step, cells were re-fixed in PBS 1X, 4% formaldehyde for 2 minutes at room temperature and were washed three times with PBS 1X, this step was repeated twice. Serial dehydration was performed, 5 minutes with 70%, 90% and 100% in EtOH. Hybridisation solution (1M Tris pH 7.4, MgCl2 buffer pH 7.0 (25mM MgCl, 9mM citric acid, 82 mM Na2HPO4), deionize formamide 70%, Probe 1 ug/ul, blocking reagent 0.25% final concentration) was added on top of every coverslip. Samples were denaturated at 80°C for 3 minutes and incubated for 2 hours at room temperature in wet dark chamber. Cells were washed in FISH washing (Formamide 50%, 10mM

TrisHCl pH 7.2, 0.1% BSA) solution at room temperature three time under shaking and subsequently washed in TBS1x, 0.08% Tween twice for 5 minutes. After that, DAPI (Sigma), TBS 1X, 0.08% Tween was added for 5 minutes and was performed serial dehydration in 70%, 90% and 100% EtOH for 5 minutes. Slides were mounted with ProLong (Invitrogen). Images were captured using classic immunofluorescence (Leica DM4000B) microscope. Co-localization counting were quantified by visual inspection. Pericentromeric probe for Sat2D:
5'-5TYE563-GATCGAATGGAATCTGAATGGAA-3'.

3.10 Vector cloning

Vectors expressing GFP tagged version of SFPQ mutants were produced in home.
Briefly, we first cloned a NLS-MCS-Myc tag construct into pLV-eGFP vector (Addgene, #36083) by using BsrgI rescriction site, producing an entry vectory for downstream application, as well as being an empty control vector.
The entry vector was the digested with SpeI and XhoI restriction enzyme to insert inside the desired construct (SFPQ mutant or the P domain/P domain nFS alone).
Ligation was carried out using 50 ng of vector and a 1:7 ration of insert by using Hi-T4 DNA Ligase (NEB) following manufacture instructions.
Competent Stable *E.coli* (NEB) were transformed by heat shock and left growing at 30° O/N. Obtained colonies were inspected for correct construct insertion by MiniPrep (homemade) and validated by Sanger Sequencing (Eurofins Genomics).
Expression vectors for recombinant proteins were produced by LIC cloning by inserting Myc-SFPQ into pNIC-CTHF (Addgene, #26105) digested with BfuAI, and HA-DAXX into pGTVL1 (Addgene, #39188) using BseRI restriction sites, giving rise to Myc-SFPQ-His-FLAG and HA-DAXX-GST, respectively.
Cloning oligos are listed in table Primer Table (3.13).

3.11 Protein extraction and Western Blotting

Whole-cell lysates were prepared using a modified RIPA buffer (50 mmol/L Tris-HCl (pH 7.5) 250 mmol/L NaCl, 1% Triton X-100, 1% deoxycholic acid, 1% SDS) supplemented with serine protease inhibitor phenylmethylsulfony fluoiride (PMSF) 1mM, phosphoprotein phosphatase inhibitors sodium fluoride (NaF) 1mM and protease inhibitor cocktail (PIC) 1X (SIGMA). After scraping, cells were incubated for 60 minutes on ice and sonicated (Fisher scientific). After centrifugation, 15 minutes,

13,000 rpm, 4°C, supernatants have been recovered and used for Western blotting according to standard procedures. Primary antibodies used are listed in Table 1. Membranes were incubated with the specific secondary antibodies bound to the HRP enzyme (horseradish peroxidase-conjugated antibody) (Sigma) and the levels of protein expression were detected by chemiluminescence using the ECL system with subsequent exposure on ChemiDoC (Bio-Rad).

3.12 Recombinant proteins production

For producing recombinant SFPQ and DAXX proteins, *E.coli* Rosetta 2 (DE3) cells were transformed with expressing vectors. Upon transformation, cells were grown in LB medium under shaking until reaching OD600=0,6. 0.1 mM IPTG was then added to induce protein expression, leaving cells growing O/N at 18°C under shaking. After centrifugation, bacterial cells were resuspended in Lysis buffer (50 mM Tris pH 7.5, 300 mM NaCl, 5% glycerol, 5 mM Imidazole) + inhibitors (protease inhibitor cocktail Roche, 1mM PMSF, 1mM TCEP, lysozyme 1mg/ml) and disrupted using homogenizer and clarified by centrifugation at 30.000g for 1 hour at 4°C. Supernatant was applied to nickel (SFPQ) or glutathione (DAXX) resin for 1 hour and 30 minutes under rotation at 4°C. SFPQ bound resin was pelleted, washed with 50mM Imidazole in Lysis buffer, then eluted in 300 mM Imidazole in Lysis buffer. DAXX bound resin was pelleted, washed using a gradient of KCl, ranging from 0 M to 1M, then back to 0 M, in Lysis buffer, then eluted in 300 mM Glutathione in Lysis buffer. The most enriched fraction(s) was loaded into a size exclusion chromatography column. Collected fractions were inspected by SDS-PAGE Giotto ProBlue safe stain to identify the fraction(s) containing more pure protein.

3.13 Immunofluorescence on Chromosome Spreads

For chromosome spreads, asynchronous cells were treated with 0.2 μg/mL Colcemid for 4 hours and mitotic cells were collected my mitotic shake-off as described [103]. Cells were subsequently washed once in 1X PBS at room temperature. Cells were then resuspended in 65 mM KCl buffer for 10 minutes and spun onto glass slides by Cytospin at 1400 rpm for 5 minutes. Slides were incubated in KCM buffer (10 mM KCl, 20 mM NaCl, 10 mM Tris-HCl pH 8, 0.5 mM EDTA, 0.1% Triton X-100) for 10 minutes. Primary antibodies were diluted in 1% BSA/KCM buffer and slides were incubated with primary antibodies for 2 hours in a wet chamber. The slides were

washed with KCM buffer three times for 5 minutes each, and then were incubated with secondary antibodies in 1% BSA/KCM buffer for an hour. Cells were subsequently washed three times with KCM buffer (5 minutes each time) and fixed with 4% paraformaldehyde for 15 minute and mounted in ProLong (Invitrogen). All buffers used were supplemented with NaF and sodium orthovanadate (Na2VO4) to prevent dephosphorylation of proteins. Images were captured using classic immunofluorescence microscope (Leica DM4000B) and were quantified by visual inspection. The Student t-test was used to calculate the statistical significance.

3.14 Bioinformatics methods

RNA sequencing was performed in triplicate for both conditions. Raw sequences files' quality was checked via FastQC (v 0.11.9, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Transcript quantification was conducted with STAR (version v.2.7.9a) using Ensembl GRCh38 GTF file and genome version (accessed on September 2022). The generated gene counts were consequently analysed using R (2) package DESeq2 (3). The normalized count matrix was obtained from variance stabilizing transformation (VST) method as implemented in DESeq2 package. In order to explore high-dimensional data property, among the available algorithms, Principal Component Analysis (PCA) coupled with a dimensionality reduction algorithm was used to scale data. In order to select only the statistically significant changing genes between comparisons of interest, a differential gene expression analysis was performed. The differentially expressed genes (DEGs) were selected with a *p-adjusted* cut off of 0.05 and a log2 Fold Change value greater than 1.5 (up-regulated DEGs) or lower than -1.5 (down-regulated DEGs). P-value was adjusted for multiple testing using the Benjamini–Hochberg (BH) correction with a false discovery rate (FDR) ≤ 0.05. To be able to identify features biological identities and the pathways they belong to, a functional annotation analysis was performed for all the comparisons and for feature list of interest with gprofiler2 package (4). Different databases were used to annotate the DEGs: Gene Ontology (Molecular Functions – MF, Biological Processes – BP, Cellular Component – CC), Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, WikiPathways (WP), Transfac (TF), miRTarBase (MIRNA), Human Protein Atlas (HPA), CORUM (CORUM protein complexes), Human Phenotype Ontology (HP), RNA central. Functional annotation results were visualized

with ggpubr package (5) via Balloon plots. The Gene Ontology network visualization was obtained with the cnetplot function, belonging to enrichplot package (6).

ChIP-seq analysis was performed with the ChiP-AP pipeline, performing quality control (FastQC), read trimming and filtering (Bbduk and Trimmomatic), genome mapping (BWA-MEM), peak calling (Genrich, HOMER, MACS2, SICER2), peak merging and annotation (HOMER). Using R/Bioconductor environment, peaks that were called by three out of four peak calling methods were considered as present. Peaks were considered differentially expressed when having p-value < 0.05 for at least three methods. In particular, for HOMER, MACS2, and SICER2, the Fold Change must be greater than 1 or for Genrich the enrichment score must be positive (up-regulated peaks). Overlaps between peaks were calculated with findOverlaps function (IRanges package) with a minimum overlap of 100bp. Density plots were generated with plotKaryotype and kpPlotDensity functions (karyoploteR package). The distribution of the width of the peaks was investigated across all differentially expressed peaks for every class of annotation, and were then log2 transformed for plotting convenience. Outliers for width values, whether falling below the fifth percentile or exceeding the ninety-fifth percentile, were treated as having values equivalent to the fifth or ninety-fifth percentile, respectively. The average Fold Change for differentially expressed peaks was determined based on Fold Change values from HOMER, MACS2, and SICER2. Genrich was omitted from the analysis due to having an enrichment score. Barplots and boxplots were created with ggpubr package. Piecharts were generated with the PieDonut function (webr package). Lineplots were generated with the newggslopegraph function (CGPfunctions package).

To analyse patients overall survival, The Cancer Genome Atlas (TCGA) Pancancer Sarcoma dataset (accessed September 2023, n = 250) was retrieved from cBioPortal, jointly with patient metadata. The analysis was performed in R/Bioconductor environment. Patients were categorized into high-expression and low-expression groups using the surv_cutpoint function from the survminer package. The p-value was computed using the Survdiff function from the survival package, and the hazard ratio was determined using the coxph function from the same package. These results were visualized through Kaplan-Meier plots. Censoring was applied for cases starting at 60 months (equivalent to 5 years) onward, and patients who passed away after that point were treated as though they were still alive before that time.

3.15 Primer Table

Primer	Sequence
CCL5 FW	CCTCCCCATATTCCTCGGAC
CCL5 RV	CACACTTGGCGGTTCTTTC
DDX60 FW	CCCAGGGTCCAGGATTTTAT
DDX60 RV	GAACAGTTGCTGCCACTTGA
IFI44 FW	AGCCTGTGAGGTCCAAGCTA
IFI44 RV	TTTGCTCAAAAGGCAAATCC
IL1α FW	TAAGCTGCCAGCCAGAGAGGGA
IL1α RV	AGCCTTCATGGAGTGGGCCATAGC
IL6 FW	AAAGCAGCAAAGAGGCACTGGCA
IL6 REV	CTGCACAGCTCTGGCTTGTTCCT
IL8 FW	GGCAGCCTTCCTGATTTCTG
IL8 RV	CTTGCCAAAACTGCACCTTCA
IL12 FW	CCTTCACCACTCCCAAAACCT
IL12 RV	TGTCTGGCCTTCTGGAGCAT
CGAS FW	AAG GAT AGC CGC CAT GTT TCT
CGAS REV	TGG CTT TCA GCA AAA GTT AGG
STING FW	AGC ATT ACA ACA ACC TGC TAC G
STING REV	GTT GGG GTC AGC CAT ACT CAG
IRF3 FW	CTC GTG ATG GTC AAG GTT GTG
IRF3 REV	AGT TTA TTG GTT GAG GTG GTG G
MAD1L1 FW	TCACCACCATTCTCACGTCA

MAD1L1 RV	TGATGGTGTGGAGAGTGGTC
Sat II FW	TCGATGTTGATTCCATTAGTTTCCA
Sat II RV	AAATGTGATCTTCATTGAATGGACT
Sat III FW	AATCAACCCGAGTGCAATCG
Sat III RV	TTCCCTTCCATTCCATTATTATCCATG
αSat FW	CTCACAGAGTTGAACCTTCC
αSat RV	GAAGTTTCTGAGAATGCTTCTG
Telomere FW	ACACTAAGGTTTGGGTTTGGGTTTGGGTTTGGGTTAGTGT
Telomere RV	TGTTAGGTATCCCTATCCCTATCCCTATCCCTATCCCTAACA
DRIP Telo FW	CGGTTTGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTT
DRIP Telo RV	GGCTTGCCTTACCCTTACCCTTACCCTTACCCTTACCCT
Subtel 16p FW	TGCAACCGGGAAAGATTTTATT
Subtel 16p RV	GCCTGGCTTTGGGACAACT
LINE1_FW	AGGCCACTGTGTGCGCGC
LINE1_RV	CCAGGTGTGGGATATAGTCTC
AluY_FW	AGATCGAGACCATCCTGGCT
AluY_RV	CCGCCTCCCGGGTTCACGCC
AluS_FW	GCCGAGGCGGGCGGATCACC
AluS_RV	GCCTCCCGAGTAGCTGGGAT
SYCP2_FW	TCCCTTCCTGCTCGCCTGTCC
SYCP2_RV	TCCTCTCGCGCTATGCTCTCCT

PDE3A_FW	GACAGCGATGAGTCAGGAGA
PDE3A_RV	TCTGAAGAGTGCGACTGAGG
SFPQ FW	TGCCATTCATGCTTCTATGCA
SFPQ RV	GGCCTAGACACTCTCATGCTTTC
DAXX FW	AAGCCTCCTTGGATTCTGGT
DAXX RV	ATCATCCTCCTGACCCTCCT
RNaseH1 FW	CCTGTACTTACTGGTGTGGAAAATAGC
RNaseH1 RV	CCGTGTGAAAGACGCATCTG
ACTIN_FW	AGCACTGTGTTGGCGTACAG
ACTIN_RV	TCCCTGGAGAAGAGCTACGA
Cloning Oligos	
NLS-MCS-MYC TOP	GTACAAGGATCCAAAAAAGAAGAGAAAGG TAACTAGTGCTAGCTTAATTAAGCTCGAG GAACAAAAACTCATCTCAGAAGAGGATCTGTAAT
NLS-MCS-MYC BOT	GTACATTACAGATCCTCTTCTGAGATGAGTTTTTGTT CCTCGAGCTTAATTAAGCTAGC ACTAGTTACCTTTCTCTTCTTTTTTGGATCCTT
P dom FW	GTTAATACTAGTCCGCAGGACTCTTCCAAG
P dom RV	GTATCCTCGAGGATCTTCTCCTCGCTGCG
SFPQ FW	GTTAATACTAGTATGAAGCTCATGTCTC
SFPQ RV	GTATCCTCGAGGCTAAAATCGGGGTT
SFPQ LIC FW	TTAAGAAGGAGATATACTATGGAACAAAAACTC
SFPQ LIC RV	ATTGGAAGTAGAGGTTCTCTGCAAATCGGGGTTTTTT
DAXX LIC FW	TACTTCCAATCCATGTACCCATACGATGTT
DAXX LIC RV	TATCCACCTTTACTGCTAATCAGAGTCTGA

3.16 Antibody Table

Antibody	Company	Code	WB	IF	ChIP
ATR	CELL SGNALING	2790S			5 μg
cGAS	AbClonal			1:100	
CREST	Antibodies Incorporated	15-235		1:400	
DAXX	BETHYL	A301-353A			2.5 μg
H3	ABCAM	AB1791			1 μg
H3.3	MILLIPORE	09-838			1.5 μg
IKBα	CELL SGNALING	#4814	1:2000		
p65	CELL SGNALING	#7956	1:500		
pATR	GENETEX	GTX128145	1:500		
pIKBα	CELL SGNALING	#2859	1:2000		
pIRF3	AbClonal	AP1412		1:100	
p-p65	CELL SGNALING	#3033	1:1000		
pSTAT3	CELL SGNALING	#9145	14000		
SFPQ/PSF	Bethyl	A301-321A	1:5000		
STAT3	CELL SGNALING	#9139	1:4000		
TRF2	MILLIPORE	05-521		1:200	
Vinculin	CELL SGNALING	#13901	1:1000		
γH2AX	MILLIPORE	07-164			4 μg
Actin	SIGMA	A5441	1:20000		

4 RESULTS

4.1. SFPQ interacts with DAXX to suppress R-loops at repetitive regions genome-wide

Our group has recently identified two RNA-binding proteins that localize to telomeres and interact with TERRA lncRNA, polypyrimidine tract-binding (PTB) protein associated splicing factor proline/glutamine-rich (SFPQ) and the nuclear RNA-binding protein 54 kDa (NONO) [227]. SFPQ suppresses R-loop levels at telomeres and modulates ALT activity in cell models for Alternative lengthening of telomeres [227]. However, the mechanism SFPQ display remains ununderstood. Thus, we decided to focus our experiments on the role of SFPQ in regulating genome wide R-loop formation and unveiling its interaction with a putative novel partner.

4.1.1 SFPQ and DAXX direct interact with each other and have a common role in suppressing R-loop

Knockdown of SFPQ results in increased RNA:DNA hybrids in both telomerase positive non-small cell lung cancer H1299 cells, and telomerase negative, ALT positive U-2 OS osteosarcoma cells, that is paralleled by increased recombination events at telomeres of both cell lines. However, SFPQ does not possesses any known catalytic domain, suggesting that SFPQ may act as recruiter for R-loop resolution activity.

Previous mass spectrometry experiments on SFPQ immunoprecipitates in our laboratory revealed DAXX as a novel SFPQ interacting protein in H1299 cells (A. Zappone, unpublished work). DAXX is an H3.3 specific histone chaperon that, along with the ATP-dependent helicase ATRX, has been extensively reported to suppress R-loops and counteract ALT phenotype [207], [208], [213], [351]. This information led to the hypothesis that DAXX and SFPQ, by interacting, can for a previously unknown complex able to suppress telomeric R-loop. To test this hypothesis, co-immunoprecipitation (Co-IP) essay have been conducted to validate SFPQ-DAXX interaction.

To investigate SFPQ-DAXX interaction, we decided to use both H1299 cells, WT for ATRX, and U2OS cells, that carry a deletion of exon 2 to 19 in the ATRX gene, resulting in absence of ATRX expression [212], providing the unique opportunity to investigate SFPQ-DAXX interaction in a ATRX clean system.

H1299 cells were transiently transfected with vectors expressing HA-tagged DAXX or Myc-tagged SFPQ to validate the interaction in both directions. Empty vectors expressing only HA tag or Myc tag were used as negative controls. 48 h post transfection, cells were used for immunoprecipitation followed by anti-HA or anti-Myc antibody (according to the overexpressed tagged protein) western blotting (WB).

Western blot showed that ectopically expressed Myc-SFPQ successfully immunoprecipitated endogenous DAXX (**Figure 5A**). Moreover, ectopic HA-DAXX was able to immunoprecipitate endogenous SFPQ (**Figure 5B**).

Notably, while SFPQ was found to co-immunoprecipitate also with its canonical partner NONO, no interaction was detected for ATRX in H1299 cells. Moreover, HA-DAXX immunoprecipitated ATRX, but not NONO (**Figure 5A and B**). This indicates that SFPQ-DAXX interaction is independent for their canonical partners, proposing SFPQ-DAXX as a new complex. This experiment confirmed IP-MS data, validating a specific SFPQ-DAXX interaction.

To further corroborate this result, endogenous immunoprecipitation of both SFPQ or DAXX was carried out in U-2 OS cells. Consistently with previous results, endogenous proteins successfully interacted with each other (**Figure 5C**). Both SFPQ and DAXX carry out their canonical functions on chromatin substrates [195], [198], [211], [216], [236]. U-2 OS cells were fixed in 1% Formaldehyde, collected and lysed. Chromatin was sonicated to obtain 300-800 bp fragments and subsequently incubated with anti-SFPQ antibody. After immunoprecipitation, eluted proteins were loaded on SDS-PAGE gel and WB was performed. In this experimental setting, DAXX was successfully co-immunoprecipitated (**Figure 5D**), suggesting interaction on the chromatin level.

Finally, SFPQ-DAXX interaction was confirmed using recombinant proteins. HA-DAXX-GST and Myc-SFPQ-His were purified form *E.coli* by affinity chromatography.

Recombinant proteins co-incubed and protein pull-down experiments using anti-Myc or glutathione conjugated beads were performed. WB analysis on eluates revealed specific interaction between the two proteins, confirming SFPQ-DAXX interaction on the molecular level (**Figure 5E and F**). All together these results show that SFPQ interacts with the novel, binding partner DAXX, in protein extracts and chromatin preparations, postulating a common action in modulating R-loops.

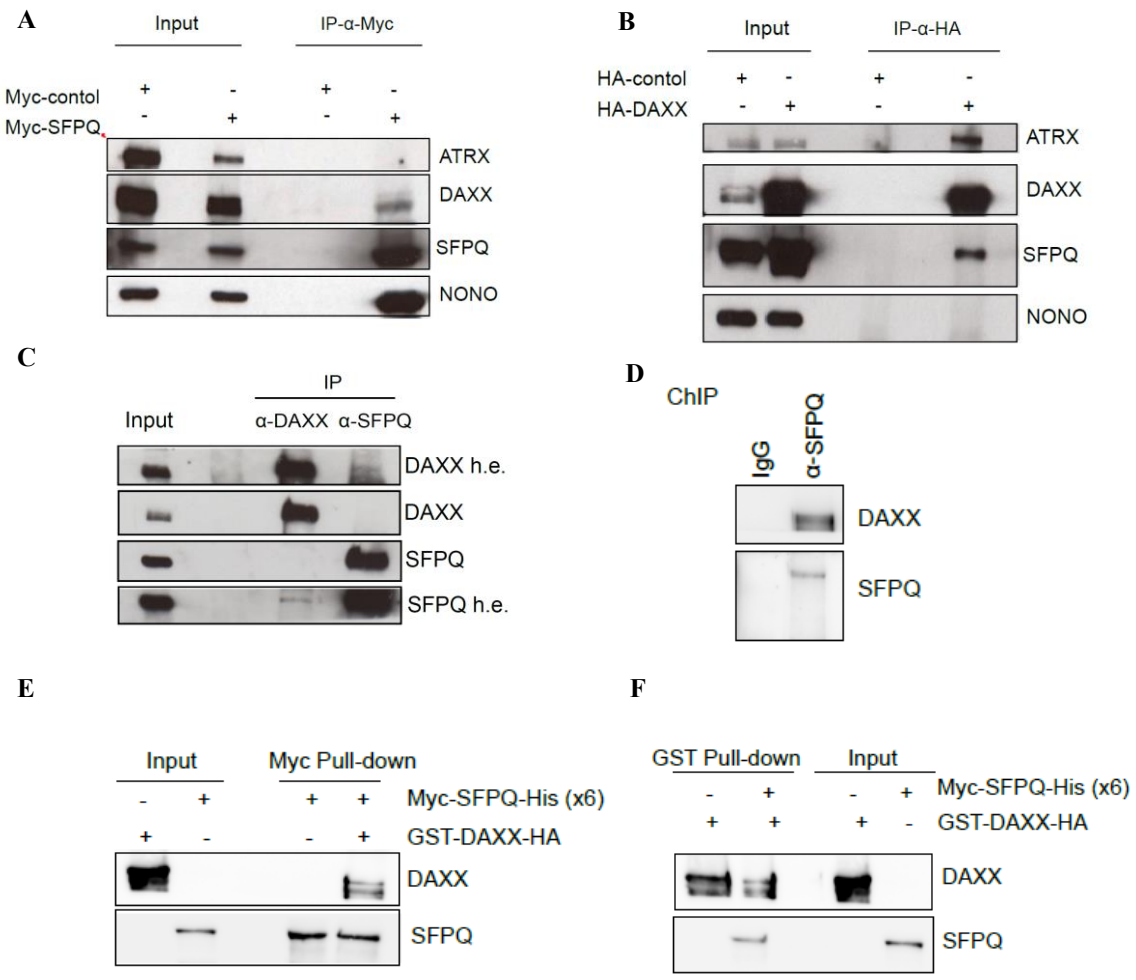


Figure 5. The RNA binding protein SFPQ directly interacts with the H3.3 histone chaperon DAXX

H1299 cells were transfected with Myc-SFPQ (**A**), HA-DAXX (**B**), or empty expressing vector and immunoprecipitated using anti-Myc or anti-HA antibody, respectively. Membranes were stained using anti-ATRX, DAXX, SFPQ, NONO antibodies. Input 20 µg. **C**) Immunoprecipitation of endogenous SFPQ or DAXX, using respective antibodies, was performed in U-2 OS cells. Membrane was stained using anti-SFPQ or anti-DAXX antibodies. **D**) Cross-linked chromatin of U-2 OS cells was immunoprecipitated using anti-SFPQ antibody, and stained for DAXX and SFPQ. IgG was used as negative control. Recombinant Myc-SFPQ-His and GST-DAXX-HA proteins were incubated together and subsequently pulled-down using Myc or GST conjugated beads (**E** and **F**, respectively). Membranes were stained for SFPQ and DAXX. Input 5 µg.

To evaluate whether the two proteins work in the same process, immunofluorescence (IF) experiment using anti-RNA:DNA hybrid specific monoclonal antibody (S9.6 antibody) combined with anti-TRF1 antibody was carried out in U-2 OS cells. U-2 OS cells are a useful model to study RNA:DNA hybrid biology as they bear elevated levels of RNA:DNA hybrids when compared to telomerase positive cells [352]. Cells were transient transfected with Control, SFPQ or DAXX specific SiRNAs for 72 h, fixed and IF was performed. Fluorescent microscopy analysis revealed that loss of SFPQ or DAXX leads to RNA:DNA hybrid formation at telomeres, as well as overall increased signal for pan-nuclear RNA:DNA hybrids, thus suggests that both SFPQ and DAXX proteins suppress R-loops, upon silencing of either SFPQ or DAXX, indicating that both proteins act at telomeres, suppressing R-loops at telomeres and non-telomeric sites. (**Figure 6A**).

To validate this finding, DNA:RNA immunoprecipitation followed by q-PCR (DRIP-qPCR) was performed. After SFPQ or DAXX knock-down (**Figure 6B**), DNA was extracted and RNA:DNA hybrids were immunoprecipitated using of S9.6 antibody. Quantitative PCR was performed using primers amplifying telomeric repeats. Silencing of either SFPQ or DAXX causes a drastic increase in R-loop levels at telomeres compare to the control condition, confirming their role in suppressing these atypical DNA structures at telomeres.(**Figure 6C**).

All together, these results indicate that both SFPQ and DAXX act as R-loop suppressors at telomeres, a classic model system for studying R-loop biology, but also point towards a more general role of SFPQ and DAXX in R-loop suppression at the entire genome level.

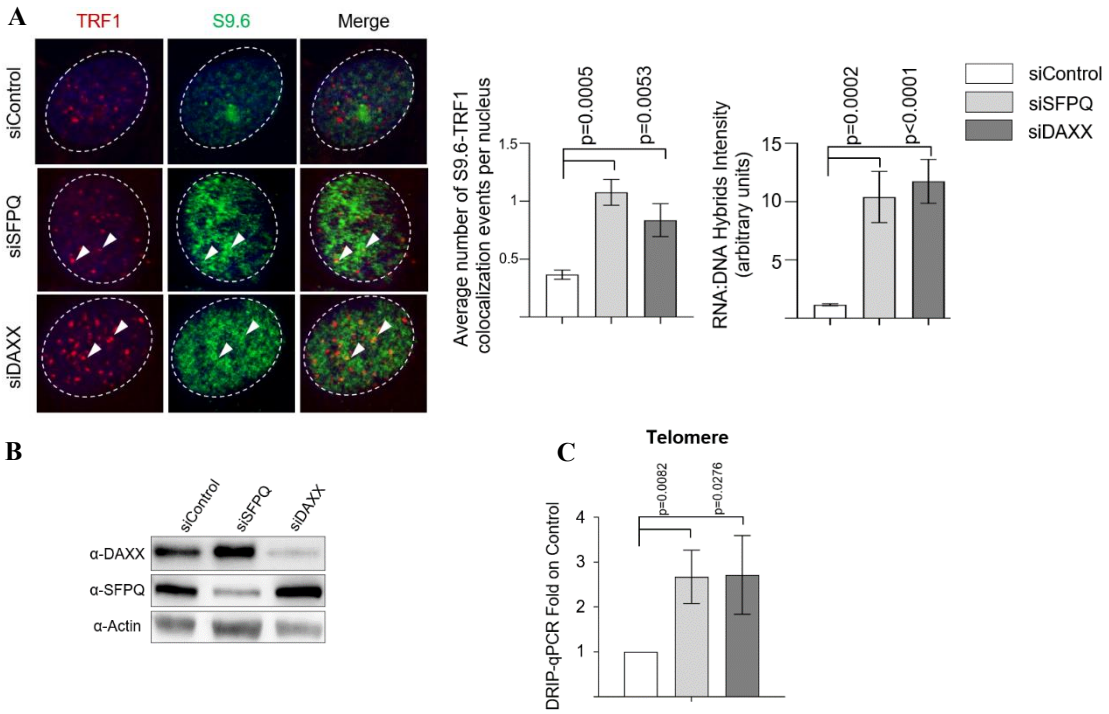


Figure 6. SFPQ and DAXX act on the same process supressing R-loops

A) Left, representative images of immunofluorescence experiment performed on U-2 OS cells transfected with indicated SiRNAs and stained using anti-TRF2 telomeric marker and anti-S9.6 RNA:DNA hybrids specific antibody. Right, quantification of TRF2-S9.6 colocalizations and pan-nuclear intensity of S9.6 antibody. **B**) WB experiment blotting SFPQ or DAXX under reported silencing conditions. **C**) DRIP-qPCR experiment amplifying telomeric repeats under reported silencing conditions in U-2 OS cell lines. RNaseH1 treatment was performed as negative control to digest RNA:DNA hybrids. IF data show means ± SEM of four biological replicates for a total of 90 nuclei. RT-PCR data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.

4.1.2 SFPQ is responsible for DAXX localization and deposition of H3.3 histone variant

Data from our laboratory shows that SFPQ is able to bind R-loop structures, as demonstrated by EMSA assays (A. Gambelli, unpublished work).

We therefore speculated that SFPQ may be responsible for DAXX recruitment and localization, thus directing its histone chaperone function.

To test this hypothesis, we transfected U-2 OS cells with SFPQ specific or Control SiRNAs and subsequently proceeded with immunofluorescence experiments using anti-DAXX and anti-TRF2 antibodies. Fluorescence microscopy analysis pointed out a significant decrease in DAXX/TRF2 colocalization events upon SFPQ knock-down, providing first line of evidence for SFPQ mediated DAXX localisation at telomeres (**Figure 7A**). Notably, DAXX staining pattern was severely altered upon removal of SFPQ, changing from a well-defined, pointed staining to a more disperse one, presumably indicating a global de-localization of DAXX due to the lack of SFPQ (**Figure 7A**).

To strengthen this result, a second IF experiment was conducted by investigating DAXX colocalization with centromeres, a reported DAXX binding region [195], [201]. Centromeres were detected by using a Centromere specific antibody from human CREST serum and colocalization with DAXX was tested under control and loss of SFPQ conditions. As expected, loss of SFPQ mediates DAXX delocalization from centromeres. This indicates that SFPQ action is required for proper localisation of DAXX not only at telomeres, but also at centromeres (**Figure 7B**).

To fully demonstrate that SFPQ is responsible for DAXX localisation and not *vice versa*, two other immunofluorescence experiment were performed by transiently knocking-down DAXX in U-2 OS cells. Loss of DAXX results in increased R-loops at telomeres and centromeres, recognised by SFPQ. Cells we stained by using an anti-SFPQ antibody combined with either anti-TRF2 or anti-CREST antibodies. In both experiments, SFPQ colocalization events were increased upon silencing of DAXX (**Figure 7C and D**).

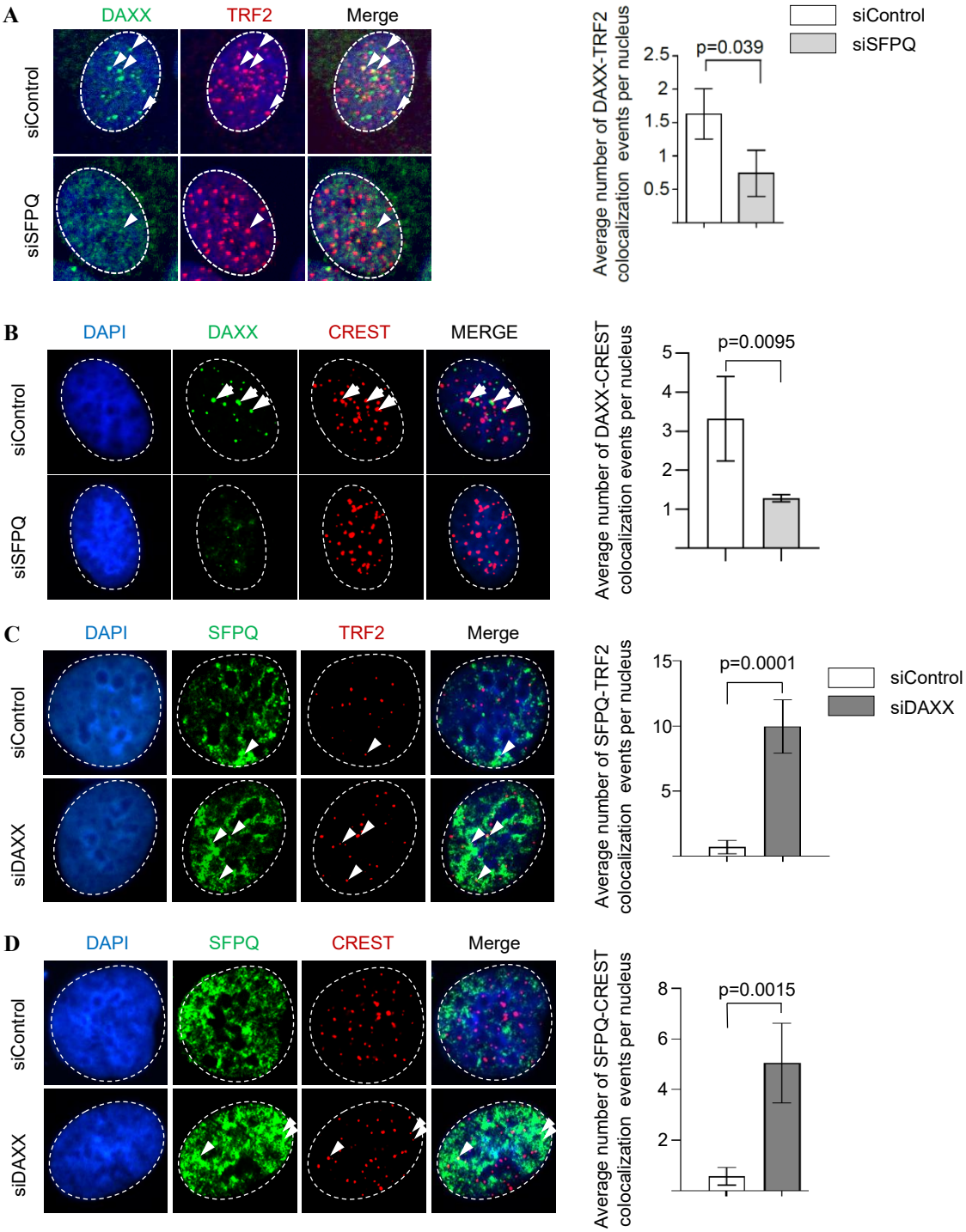
DAXX was reported to have replication independent H3.3 chaperon activity to preserve the compacted chromatin structure of telomeres and other repeat elements [195], [198], [207]. To validate a functional relevance for SFPQ-DAXX interaction in controlling chromatin structure, Chromatin Immunoprecipitation (ChIP) followed by q-PCR (ChIP-qPCR) experiment was conducted. U-2 OS cells were transiently transfected with Control or SFPQ specific SiRNA, fixed, lysed and sonicated to get 300-800 bp DNA fragments. Chromatin was then precipitated using anti-DAXX, anti-H3.3 and global anti-H3 antibodies; non-specific IgG were used as negative control. Immunoprecipitated DNA was

purified and analysed by qPCR by amplifying telomeric or centromeric repeats. We found that, upon loss of SFPQ both DAXX and H3.3 localization at telomere and centromere was strongly decreased. This indicates that SFPQ is responsible for proper DAXX localisation at these sites, prerequisite for DAXX mediated deposition of histone H3.3. Notably, levels of total histone H3 did not result in appreciable modification, confirming the specific action of DAXX in depositing histone H3.3 variant proteins (**Figure 7E**).

Collectively, these results indicate that SFPQ in able to interact with DAXX and localise it to specific genomic loci (i.e. Telomeres and Centromeres), mediating DAXX dependent deposition of histone variant H3.3, thus ensuring proper chromatin structure. Notably, reduced incorporation of H3.3 by the absence of DAXX has been reported to increase RNA:DNA hybrids [353], reinforcing our working hypothesis. Together, these changings might strongly impact on chromatin structure, resulting in its destabilization, possibly leading to DNA damage and genomic instability.

Figure 7. SFPQ recruits DAXX at telomere and centromeres, mediating proper H3.3 deposition

A) Left, IF representative images of DAXX-TRF2 co-stained U-2 OS cells silenced for SFPQ. Right, average number of DAXX-TRF2 colocalization events per nucleus under reported silencing conditions. **B**) Left, IF representative images of DAXX-CREST co-stained U-2 OS cells silenced for SFPQ. Right, average number of DAXX-CREST colocalization events per nucleus under reported silencing conditions. **C**) Left, IF representative images of SFPQ-TRF2 co-stained U-2 OS cells for DAXX. Right, average number of SFPQ-TRF2 colocalization events per nucleus under reported silencing conditions. **D**) Left, IF representative images of SFPQ-CREST co-stained U-2 OS cells silenced for DAXX. Right, average number of SFPQ-CREST colocalization events per nucleus under reported silencing conditions. **E**) ChIP-qPCR experiment amplifying telomeric and centromeric regions was performed in U-2 OS cells silenced for SFPQ and immunoprecipitated for DAXX, H3 and H3.3. IF data show means ± SEM of four biological replicates for a total of 90 nuclei. RT-PCR data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.


A
DAXX
TRF2
Merge
siControl
siSFPQ
Average number of DAXX-TRF2 colocalization events per nucleus
p=0.039
siControl
siSFPQ
B
DAPI
DAXX
CREST
MERGE
siControl
siSFPQ
Average number of DAXX-CREST colocalization events per nucleus
p=0.0095
C
DAPI
SFPQ
TRF2
Merge
siControl
siDAXX
Average number of SFPQ-TRF2 colocalization events per nucleus
p=0.0001
siControl
siDAXX
D
DAPI
SFPQ
CREST
Merge
siControl
siDAXX
Average number of SFPQ-CREST colocalization events per nucleus
p=0.0015
E
Telomere
α Sat
siControl
siSFPQ
ChIP-qPCR Fold on Control
p<0.0001
p=0.0017
p=0.0001
p=0.0004
p=0.0014
p=0.0009
DAXX
H3.3
H3

4.2 SFPQ suppresses replication stress at repetitive elements

4.2.1 SFPQ directs DAXX dependent H3.3 chaperon activity at repetitive regions

We then investigated whether SFPQ may act on the genome-wide level, directing DAXX to different types of non-coding, repetitive regions (supported by the fact that DAXX has been identified as H3.3 histone chaperon at these sites [197], [198], [201], [313]).

To investigate whether SFPQ has an actual effect on regulating repetitive regions genome wide, we took advantage of commercially available COT-1 DNA, a mixture of DNA fragments composed of repetitive regions of the human genome (comprising Alu and LINE sequences) to generate fluorescent labelled FISH probes.

We performed COT-1 RNA-FISH combined with immunofluorescence (IF RNA-FISH) using SFPQ or RNaseH1 knock-down U-2 OS cells. RNaseH1 was previously shown to suppress R-loops at repetitive DNA [8], [46], [252], [312], [354].

The experiment showed a substantial increase in COT-1 staining intensity as well as number of COT-1 foci per nucleus, together indicating an increase in repetitive elements increase transcription upon SFPQ depletion (**Figure 8A**). Moreover, an increased number of colocalization events of COT-1 RNAs with the replication stress marker pATR were detected, indicative for the presence of replication stress at repeats, presumably due to increased R-loop levels (**Figure 8A**). Finally, knock-down of RNaseH1 increases SFPQ localization at COT-1 sites, suggesting a SFPQ recruitment to these sites in order to call for DAXX activity (**Figure 8B**).

To confirm of increased transcription of COT-1 DNA labelling (i.e. repetitive regions) in RNaseH1 or SFPQ loss of function cells, Realtime PCR experiments were conducted. After 72 hours of transient knock-down for SFPQ or RNaseH1, U-2 OS cells were harvested and total RNA was extracted. RT-PCR experiment revealed a slight, yet significant increase in Alu, LINE, pericentromeric SatIII and Telomeric transcripts upon loss of SFPQ (**Figure 8C**). This experiment pointed out the ability of SFPQ in suppressing repetitive regions transcription to suppress replication stress at repetitive elements.

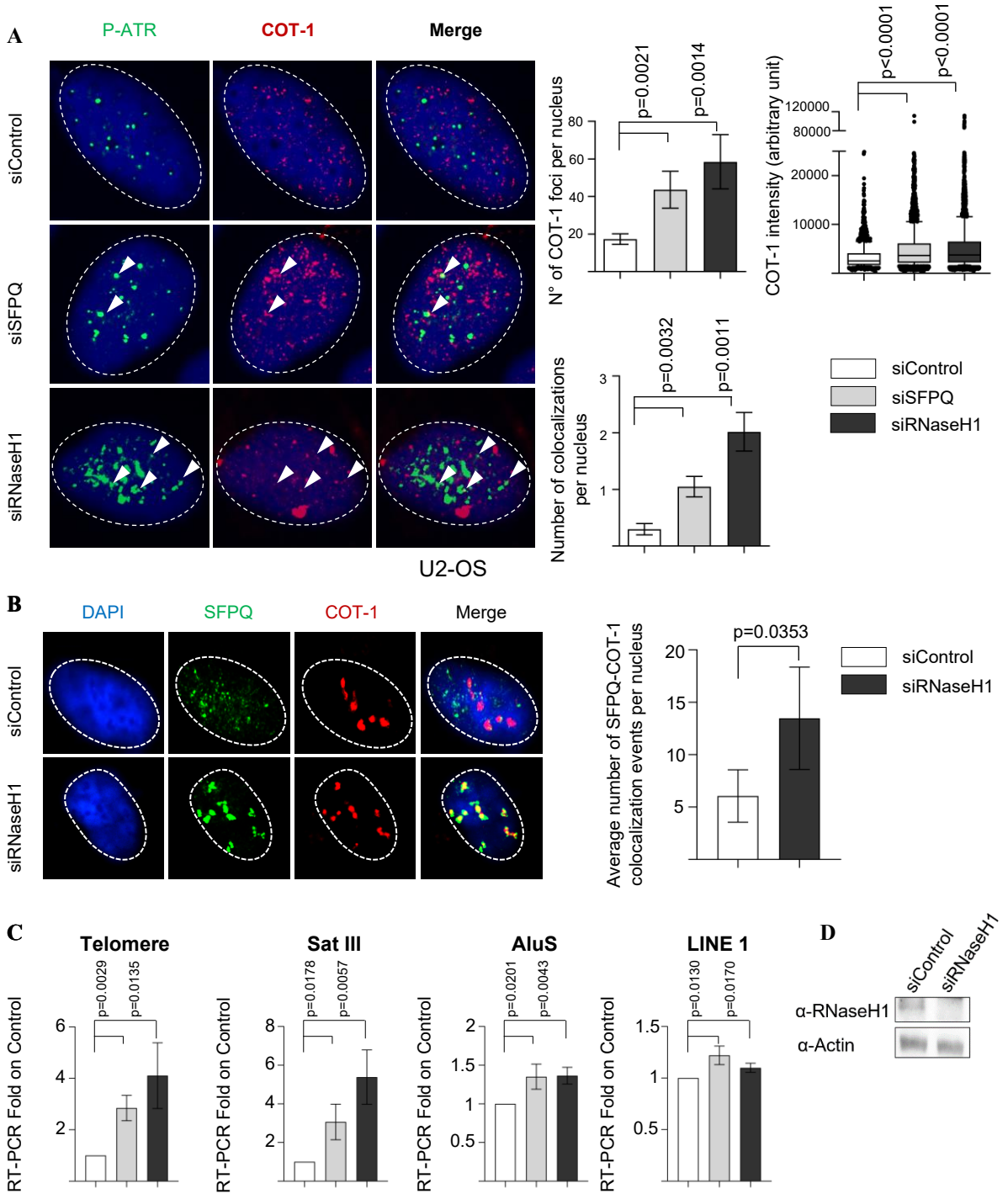


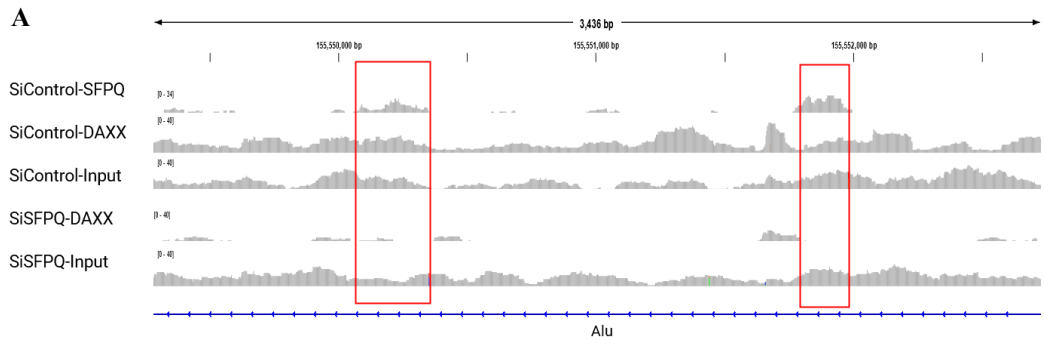
Figure 8. SFPQ prevents replication stress due to unscheduled transcription of repetitive regions

A) Left, representative images of pATR/Cot-1 IF RNA-FISH performed in U-2 OS cells transiently transfected with reported siRNAs. Right top, quantification of Cot-1 signals intensity and percentage of total number of Cot-1 foci per nucleus under reported silencing conditions. Right bottom, quantification of pATR/Cot-1 colocalization events per nucleus under reported silencing conditions. **B**) Up, representative images of SFPQ/Cot-1 IF RNA-FISH performed in U-2 OS cells transiently transfected with reported siRNAs. Down, quantification of SFPQ/Cot-1 colocalization events per nucleus under reported silencing conditions. **C**) Gene expression of indicated repetitive regions transcripts under reported silencing conditions determined by RT-qPCR. **D**) WB experiment blotting RNaseH1 under reported silencing conditions. IF RNA FISH data show means ± SEM of four biological replicates for a total of 90 nuclei. Arrowheads indicate co-localization events. RT-PCR data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.

It still remains to be uncovered whether DAXX can also be directed at repetitive regions, and to determine which regions are actually bound by the SFPQ-DAXX proteins.

To address this question, a ChIP-seq experiment was conducted. U-2 OS cells transiently transfected with Si-Control or Si-SFPQ and samples were immunoprecipitated using anti-SFPQ and anti-DAXX antibody, followed by Illumina massive parallel sequencing (**Figure 9A**).

SFPQ and DAXX strongly co-localize in control condition, binding to non-coding regions, mostly enriched for ALU and LINE elements, satellite repeats, and repeats elements (**Figure 9B left and centre**). Strikingly, upon depletion of SFPQ, DAXX localization at these sites was reduced (**Figure 9B, right**), indicating that SFPQ is mediating DAXX localization at non-coding, repetitive elements genome-wide.



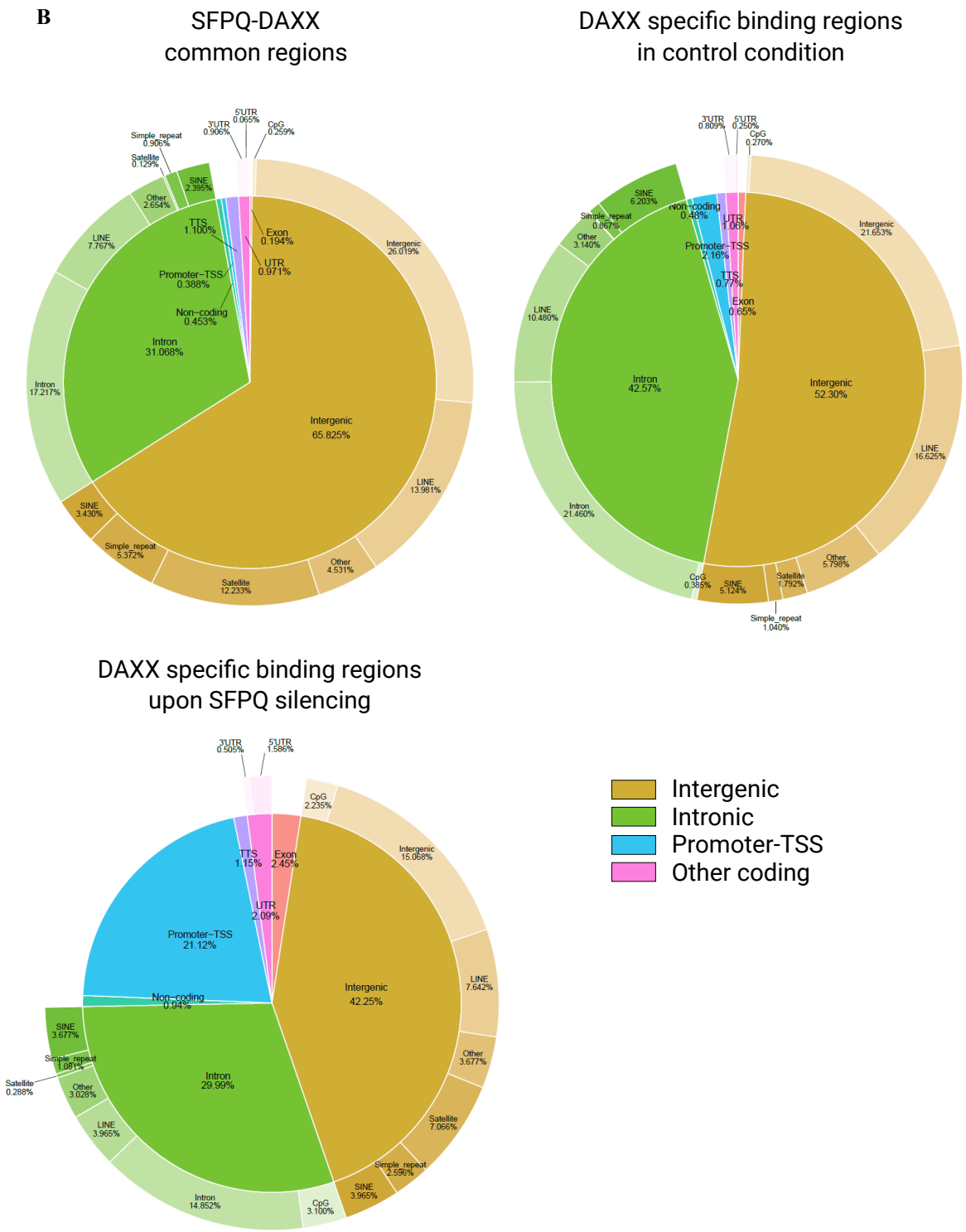


Figure 9. ChIP-seq experiment identify SFPQ-DAXX binding sites

A) Genome browser snapshot showing SFPQ and DAXX peaks, in control and siSFPQ conditions. **B**) Pie chart showing SFPQ-DAXX common binding sites (left), DAXX specific binding sites in control condition (centre), and DAXX binding sites upon loss of SFPQ (right). ChIP-seq experiment was performed in duplicate.

To validate ChIP-seq experiment, ChIP-qPCR was performed using primers amplifying classic repetitive regions, comprising centromeres (αsat SatII and SatIII), pericentromeric repeats, subtelomere, Alu and LINE repeats. Two additional regions, SYCP2 and PD3A, are used as positive control regions, as they have been reported as DAXX or SFPQ target sites, respectively. As expected, upon loss of SFPQ both DAXX and H3.3, but not total H3, were strongly decreased at all type of repetitive elements (**Figure 10A**). As a control, SFPQ immunoprecipitation upon its silencing revealed a decrease in its occupancy at reported loci, indicative of an actual occupancy reduction upon silencing (**Figure 10B**). This data confirm our hypothesis that SFPQ is able to direct DAXX and mediates H3.3 deposition at repetitive regions at the genome-wide level.

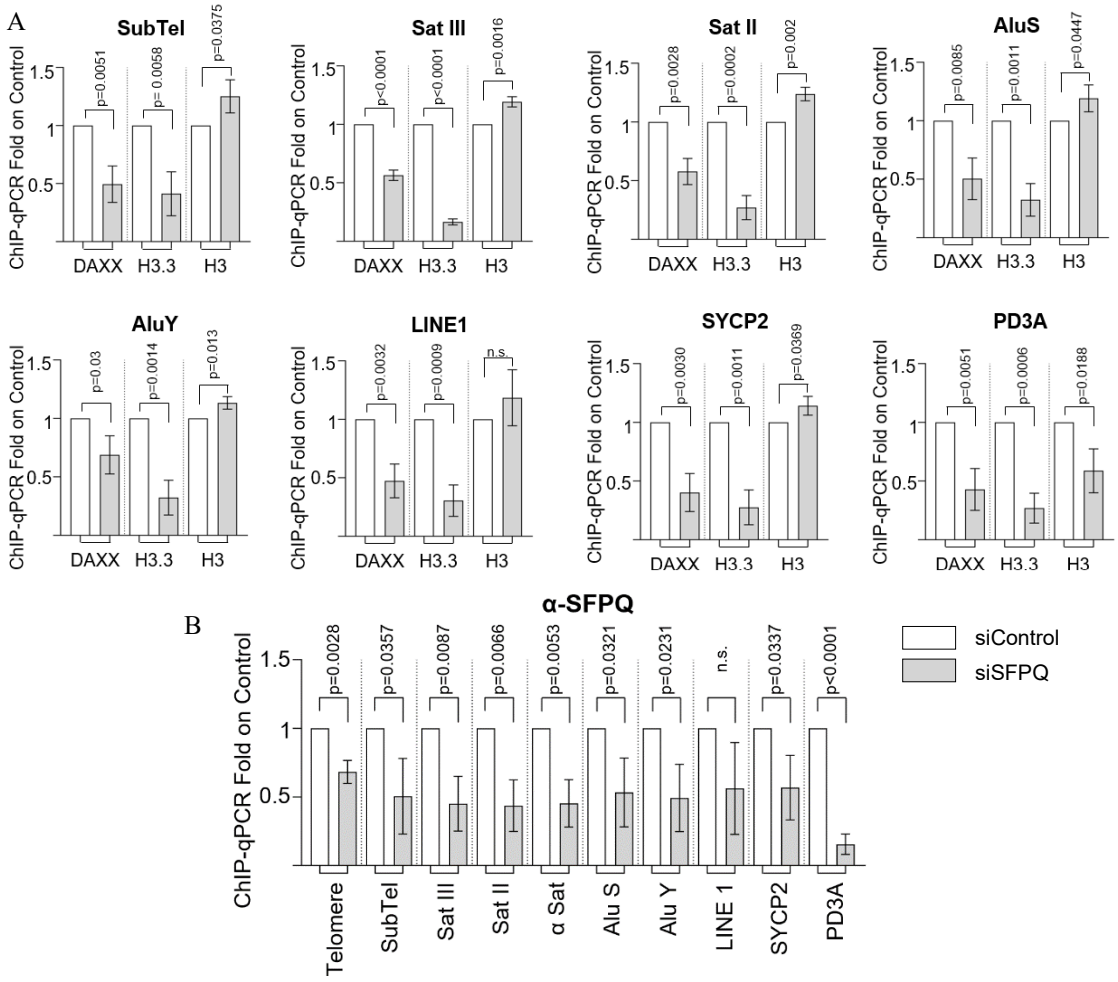


Figure 10. SFPQ mediates DAXX localization and H3.3 deposition at repetitive regions genome-wide

A) ChIP-qPCR experiment amplifying indicated repetitive regions was performed in U-2 OS cells silenced for SFPQ and immunoprecipitated for DAXX, H3, and H3.3. **B**) ChIP-qPCR experiment amplifying indicated repetitive regions was performed in U-2 OS cells silenced for SFPQ and immunoprecipitated for SFPQ itself, as positive control for its delocalisation. RT-PCR data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.

4.2.2 SFPQ and DAXX prevent replication stress and DNA damage by suppressing R-loop at repetitive regions

We have previously demonstrated that SFPQ act as potent strong R-loop suppressor at telomeres. DAXX, in collaboration with ATRX, can resolve R-loops, thus ensuring proper nucleosome architecture, at telomeres [201], [208]. To test whether SFPQ mediated DAXX localisation at telomeres and other repetitive regions results in R-loop suppression, DRIP followed by qPCR assay was performed. U-2 OS cells depleted for SFPQ for 72 hours, were lysed, genomic DNA containing R-loops was purified and sonicated to obtain 300-800 bp fragments. Immunoprecipitation was carried out using S9.6 antibody. Negative control samples were treated with recombinant RNaseH1 overnight prior immunoprecipitation, as RNaseH1 treatment degrades RNA:DNA hybrids, thus allowing to distinguish between specific or aspecific precipitated regions. Immunoprecipitated DNA was analysed through qPCR, amplifying repetitive region reported above.

DRIP-qPCR revealed that loss of SFPQ causes a strong increase in R-loops at repetitive regions, comprising telomeres, (peri)centromeres and transposable elements (Alu, LINE) (**Figure 11A**). This indicates that SFPQ suppresses R-loop not only at telomeres, but also at diverse types of repetitive elements. This finding is in line with reports linking DAXX/ATRX mutation or H3.3 impaired incorporation to ALT phenotype, increased G-quadruplex and increased R-loop levels [205], [208], [213], [214], [353].

Unscheduled R-loop are known to cause transcription-replication collision, replication stress, and DNA damage, ultimately leading to genome instability [64], [66], [77], [84], [86], [312], [355].

To test whether loss of SFPQ causes these unfavourable conditions, ChIP-qPCR experiment was performed, using antibody specific for RPA32, ATR, γH2AX, RAD51, and FANCD2. RPA32, ATR and γH2AX are proteins involved in sensing and signalling ssDNA (both derived from the displaced strand of R-loops and DNA damage), leading to cell cycle arrest and cell death[90], [93], [142], [356]–[358]. RAD51 and FANCD2, by contrast, have a strong role in DNA damage repair, being involved in DNA strand invasion during HR the first, and in replication stress resolution and R-loop suppression the latter [176], [180], [232], [298], [359]–[363].

ChIP-qPCR analysis revealed increased levels of all proteins of interest at all repeat element tested under loss of SFPQ, indication that SFPQ is a potent suppressor of R-loop mediating genomic stability (**Figure 11B**).

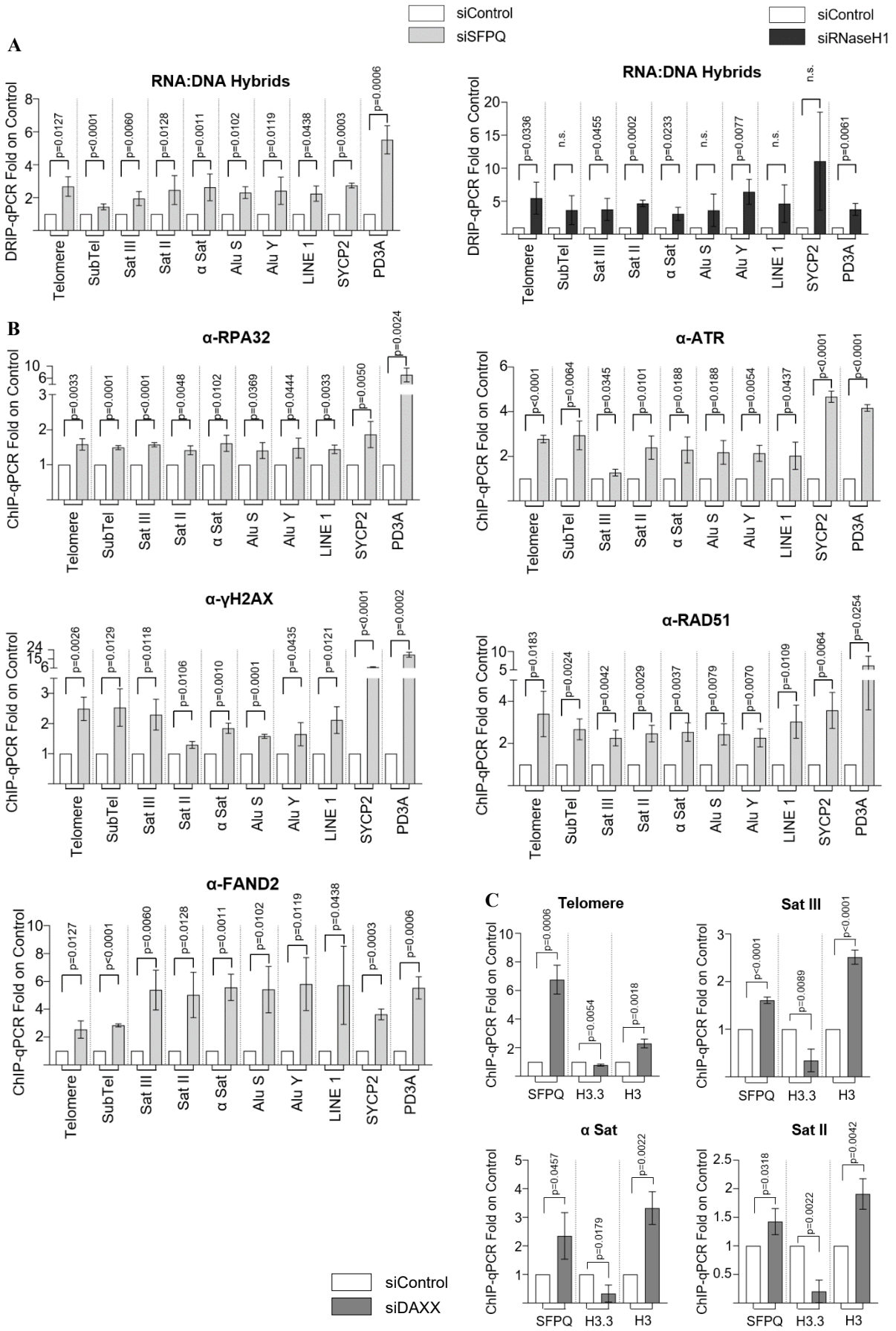
All together our experiments highlight the role of SFPQ in recruiting DAXX at repetitive region genome-wide to suppress R-loop, as depicted from DRIP-qPCR experiment under loss of SFPQ condition. Moreover, R-loop prevention by the complex maintains genome integrity by preventing DNA damage and HR, as shown from the latter ChIP-qPCR essay. Therefore, SFPQ and DAXX are important for R-loop suppression, mediating proper chromatin status by H3.3 incorporation, avoiding DNA damage and excessive recombination events, promoting genome stability.

To validate our work model, according to which SFPQ senses R-loops to recruit DAXX at repetitive elements, a ChIP experiment using chromatin from DAXX depleted U-2 OS cells was performed, revealing a strong increase in SFPQ occupancy of telomeric repeats, and other non-coding, repetitive regions (**Figure 11C**). This is paralleled by a decrease in H3.3 deposition, as expected by a reduction of DAXX expression [195], [198]. Surprisingly, also total H3 levels changed upon silencing of DAXX, suggesting that loss of DAXX causes a more severe overall defect in chromatin structure.

Overall, these experiments provides new evidence for SFPQ centrality in directing DAXX at repetitive elements. In particular SFPQ increases at reported sites underlines a role for SFPQ in responding to increased R-loop accumulation.

Figure 11. SFPQ prevents replication stress, DNA damage, and harmful recombination at repetitive elements

A) DRIP-qPCR experiment amplifying indicated repetitive regions was performed in U-2 OS cells silenced for SFPQ (left) or RNaseH1 (right) **B**) ChIP-qPCR experiments amplifying indicated repetitive regions was performed in U-2 OS cells silenced for SFPQ and immunoprecipitated using DNA damage markers (RPA32, ATR, γH2AX) and HR (RAD51, FANCD2) antibodies. **C**) Validating ChIP-qPCR experiment amplifying indicated repetitive regions was performed in U-2 OS cells silenced for DAXX and immunoprecipitated for SFPQ, H3, and H3.3. RT-PCR data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.


A
siControl
siSFPQ
siControl
siRNaseH1
RNA:DNA Hybrids
DRIP-qPCR Fold on Control
Telomere
SubTel
Sat III
Sat II
α Sat
Alu S
Alu Y
LINE 1
SYCP2
PD3A
B
α-RPA32
α-ATR
α-γH2AX
α-RAD51
α-FAND2
ChIP-qPCR Fold on Control
C
Telomere
Sat III
α Sat
Sat II
SFPQ
H3.3
H3
siControl
siDAXX

4.2.3 R-loop drive SFPQ-DAXX recruitment at telomere

To prove that the increase in R-loops level is detected by SFPQ that successfully calls DAXX chaperon activity, we transiently knock-down RNaseH1. This is reported to increase overall R-loop levels, thus allowing us to follow SFPQ and DAXX recruitment. We used anti-SFPQ or anti-DAXX antibody in combination with anti-TRF2 antibody, in order to evaluate the colocalization events of these proteins with telomeric repeats, known spot for R-loops formation. As depicted from **Figure 12A and 12B**, silencing of RNaseH1 causes a strong increase in both SFPQ and DAXX localisation at telomeres, indicating that the augmented R-loops might recruit SFPQ and DAXX for their accumulation. To strengthen our result, ChIP-qPCR experiment upon loss of RNaseH1 was performed, immunoprecipitating SFPQ, DAXX, H3 and H3.3. Telomeric occupancy of each factor was increased upon removal of RNaseH1 (**Figure 12C**), indicating that the R-loop increase is responsible for factors accumulation and action at telomeres. Altogether, these results indicate that R-loop accumulation cause SFPQ and DAXX localization at telomers, mediating insertion of H3.3 histone variant.

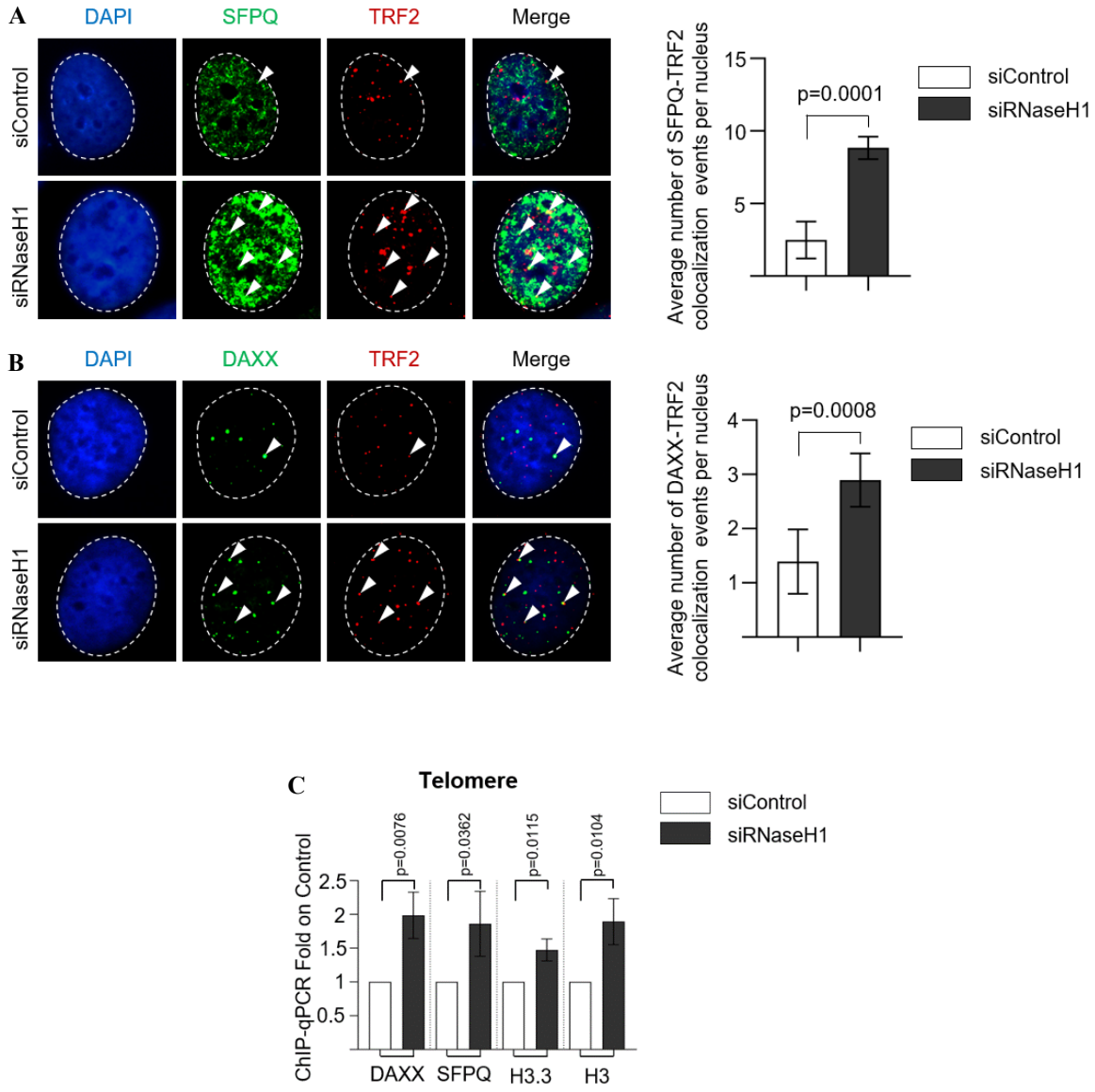


Figure 12. SFPQ is recruited by R-loop at telomeric regions

A) Left, IF representative images of SFPQ-TRF2 co-stained U-2 OS cells silenced for RNaseH1. Right, average number of SFPQ-TRF2 colocalization events per nucleus under reported silencing conditions. **B**) Left, IF representative images of DAXX-TRF2 co-stained U-2 OS cells silenced for RNaseH1. Right, average number of DAXX-TRF2 colocalization events per nucleus under reported silencing conditions. **C**) ChIP-qPCR experiments amplifying telomeric repeats was performed in U-2 OS cells silenced for RNaseH1 and immunoprecipitated anti-DAXX, SFPQ, H3, H3.3 antibodies. IF data show means ± SEM of four biological replicates for a total of 90 nuclei. RT-PCR data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance. Arrowheads indicate co-localization events.

4.3 Mapping of SFPQ domains interacting with DAXX

4.3.1 SFPQ deletion mutants revealed the proline-rich domain as the one involved in DAXX recruitment

SFPQ is known to interacts with many different binding partner (such as NONO, PSFPC1,SRSF2) to correctly carry out RNA splicing, gene regulation and R-loop suppression [227], [364], [365].

To better characterize SFPQ-DAXX interaction, we wanted to investigate with SFPQ domain is biochemically able to bind DAXX. To this extent, a series of SFPQ deletion mutants (Δ) were generated and cloned into expression vector. In addition, a patient derived non-frameshift mutation located on the P domain (a trinucleotide insertion into a previous existing codon that leads to T150delinsKP, named SFPQ nFS) [366] was generated by site specific mutagenesis (**Figure 13A**).

DAXX co-immunoprecipitations were performed to identify domain(s) crucial for DAXX interaction. U-2 OS cells were co-transfected with HA-DAXX and each Myc-SFPQ mutant. 48 hours post transfection cells were lysed and anti-Myc immunoprecipitation was performed.

We found that DAXX was able to co-immunoprecipitate with the majority of mutants, except constructs lacking a portion of the N terminus region. More detailed mapping identified the proline riche P domain as essential for DAXX binding. Remarkably, the patient derived mutant SFPQ version (nFS) is no longer able to bind DAXX. All together these finding indicate that the P domain is responsible for SFPQ-DAXX interaction (**Figure 13B**).

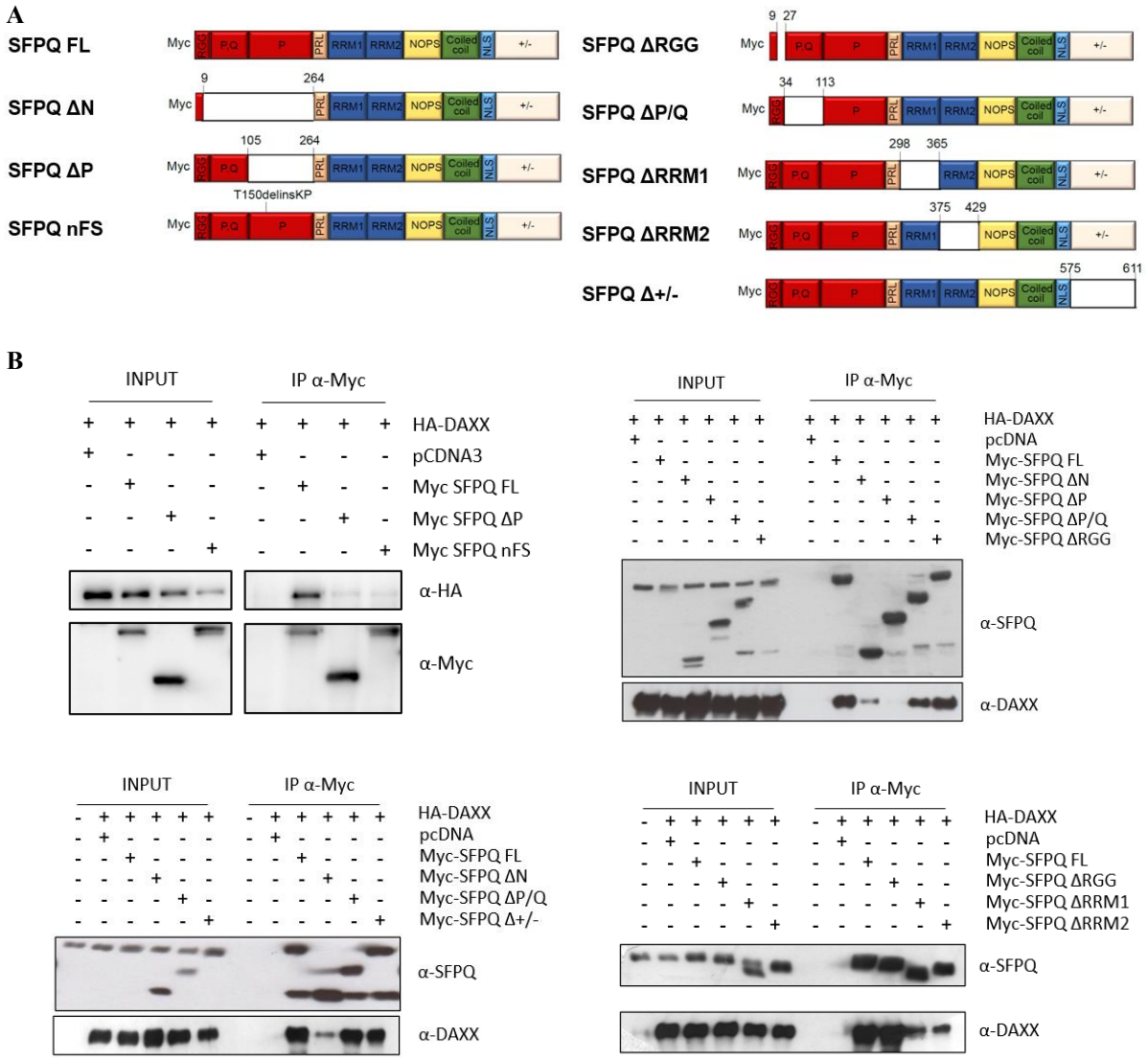


Figure 13. SFPQ-DAXX interaction is mediated by the P domain

A) Schematic representation of SFPQ deletion mutants. **B**) Co-immunoprecipitation experiments, immunoprecipitating Myc-tagged SFPQ mutants, blotting for SFPQ and DAXX. Overexpressed proteins are indicate on top of the blot.

4.3.2 P-domain is essential for the suppression of replication stress

To fully confirm a functional relevance of the P-domain in suppressing replication stress, we generated U-2 OS cell lines stably expressing Myc-tagged SFPQ deletion mutants.
We next took advantage of a specific SiRNA targeting the 5' UTR of SFPQ to knockdown the endogenous SFPQ without affecting the expression of the exogenous, mutated SFPQ. By using this system, two functional experiments were conducted to evaluate the impact of the mutated version of SFPQ.
First, an immunofluorescence experiment was performed staining cells with anti-Myc and anti-pATR specific antibodies. As depicted in **Figure 14**, after transient knock-down of endogenous SFPQ silencing, FL SFPQ was able to counteract the loss of the endogenous protein, avoiding activation of pATR. By contrast, ΔN, ΔP and nFS cells showed strong pATR activation upon depletion of endogenous SFPQ. This indicates that SFPQ with an impaired P-domain function is not able to suppress pATR activation, presumably due to a loss in its capacity to bind DAXX.
Notably, ΔP and nFS cells showed pATR activation also in control condition, suggesting that these SFPQ mutants may have dominant negative function. This underlines the relevance of the P domain in the suppression of replication stress.

4.3.3 The P-domain is important to direct DAXX localization

To demonstrate that the P-domain is central to control DAXX function, a second immunofluorescence experiment using U-2 OS cells stably expressing SFPQ mutants was performed . After depletion of endogenous SFPQ, cells were stained with anti-Myc and anti-DAXX antibodies. We found that silencing of the endogenous SFPQ causes DAXX delocalisation in U-2 OS cells stably expressing ΔN, ΔP and nFS mutants (**Figure 15**). This result recapitulates delocalization of DAXX observed upon depletion of SFPQ (see **Figure 7A**) and indicates that the P-domain is fundamental for proper DAXX interaction.

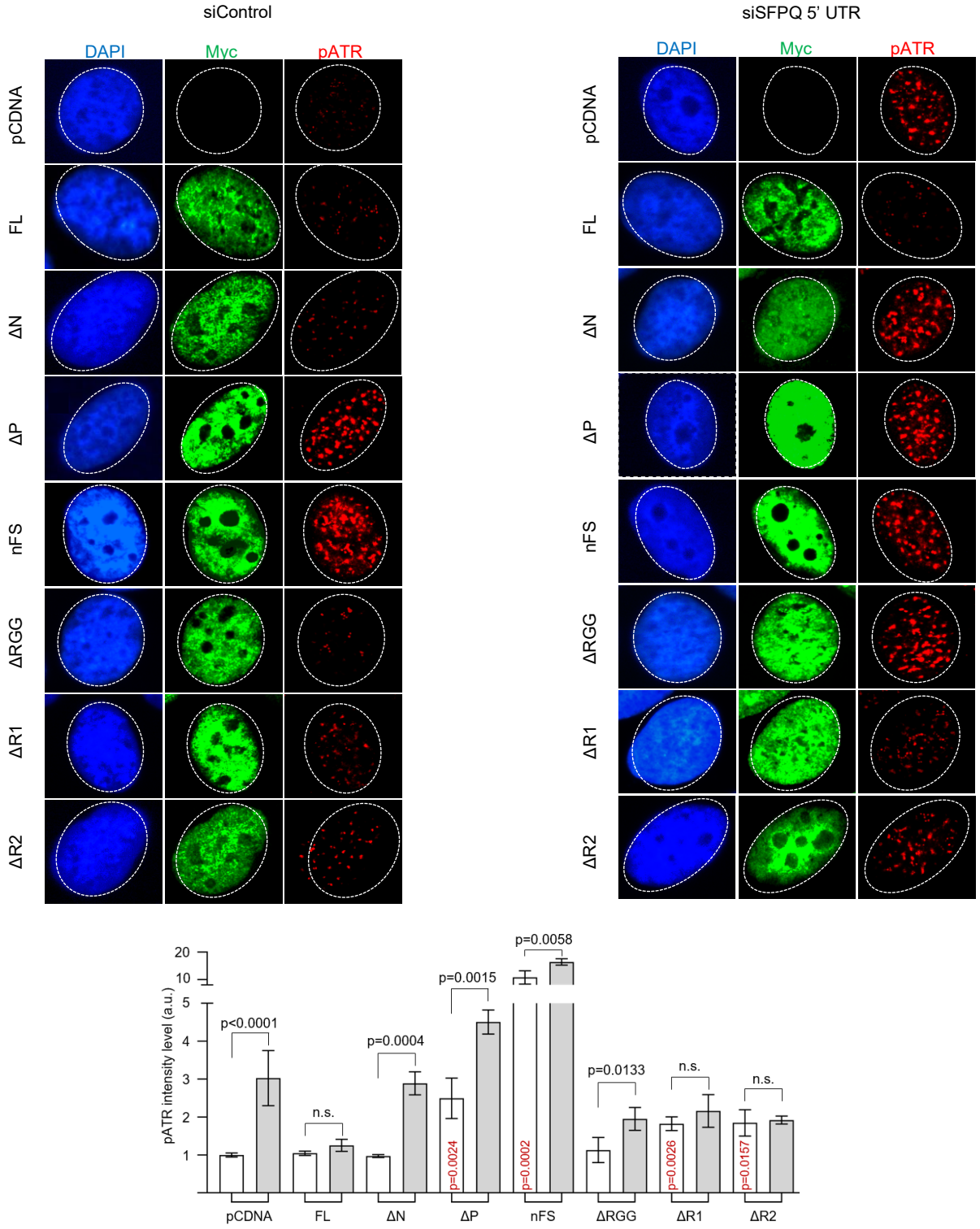


Figure 14. SFPQ lacking the P domain is not able to rescue the loss of the endogenous one

Top, IF representative images of Myc-pATR co-stained U-2 OS cells expressing different SFPQ deletion mutants (Δ) silenced for endogenous SPFQ. Bottom, average pATR signals intensity under reported silencing conditions. In black, p values of each endogenous SFPQ silenced mutant cell line with respect to its own control. In red, p values of endogenous SFPQ silenced mutant cell lines with respect to SiControl pCDNA cells. Data show means ± SEM of four biological replicates for a total of 90 nuclei. Student t-test was used to calculate statistical significance.

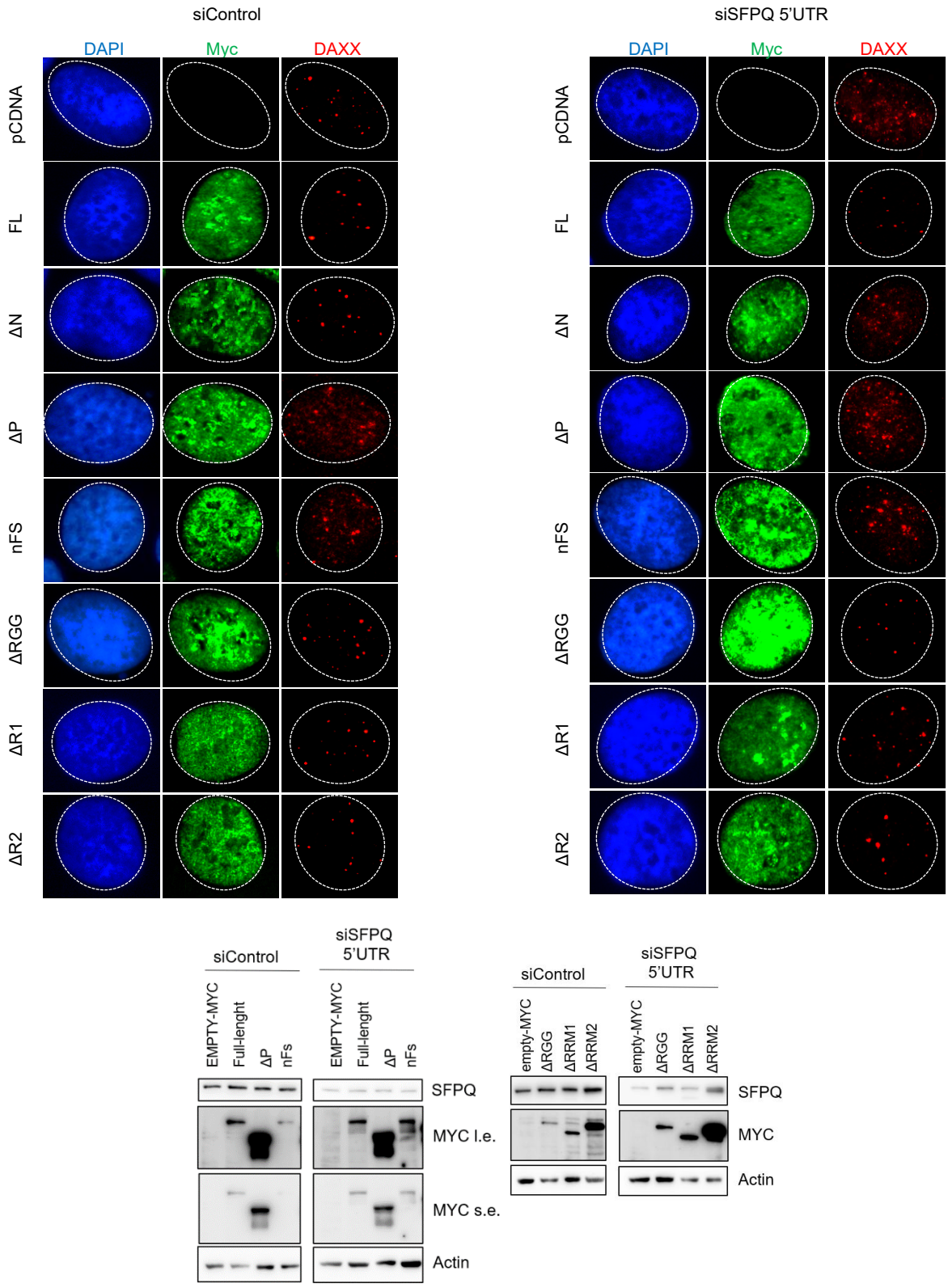


Figure 15. SFPQ lacking the P domain causes DAXX delocalization upon loss of the endogenous one

Top, IF representative images of Myc-DAXX co-stained U-2 OS cells expressing different SFPQ deletion mutants (Δ) silenced for endogenous SPFQ. Bottom, WB analysis to verify silencing of endogenous SFPQ. Endogenous SFPQ was detected using a SFPQ specific antibody; Myc-tagged deletion mutant were detected using a Myc-tag antibody. l.e. stands for long exposure. s.e. stands for short exposure.

Finally, to highlight the relevance of the P domain in mediating SFPQ-DAXX interaction, we generated expression constructs encoding a Myc-tagged wildtype or mutated P-domain fused to GFP. Localization of the fusion protein to the nucleus was directed by a NLS signal into the fusion protein.

U-2 OS cells were co-transfected with GFP-P dom/ GFP-P dom nFS and HA-DAXX and, upon 48 h of transfection, co-immunoprecipitation assay was performed using either anti-Myc or anti-HA antibodies. Western blot analysis revealed that, while the WT GFP-P dom was able to strongly bind DAXX, the P dom nFS presented decreased interaction with the protein, strengthening previous results (**Figure 16**).

All together, these experiments provided clear evidenced for the P domain mediated SFPQ-DAXX interaction, as mutant SFPQ lacking this domain, or harbouring a mutation on it, was not able to bind DAXX anymore. Moreover, cells expressing these mutants presented delocalised DAXX and activated pATR, symptoms of non-functional SFPQ-DAXX interaction, localization and function. Lastly, the P domain alone, but not its mutated version, was able to directly bind DAXX protein, thus confirming its centrality in the complex formation.

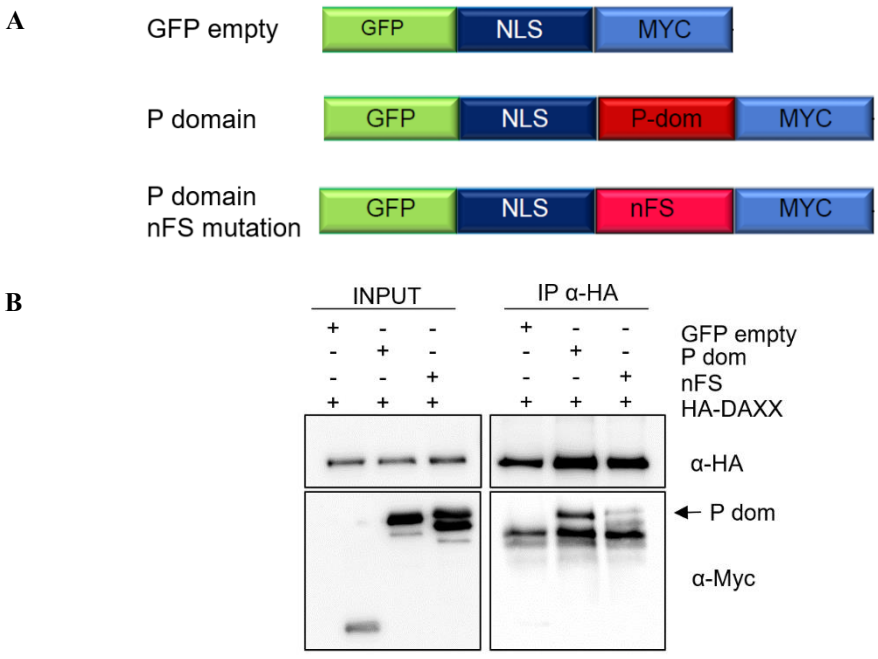


Figure 16. The P domain alone is sufficient to co-immunoprecipitate DAXX

A) Schematic representation of cloned P domain constructs. **B**) U-2 OS cells were co-transfected with HA-DAXX and GFP-P dom-Myc, its mutated version P dom nFS, or an empty GFP-Myc expression vector and immunoprecipitated using anti-HA (top) or anti-Myc (bottom) antibody. Membranes were stained using anti-HA and anti-Myc antibody. Input 20 µg. Membrane was stained using anti-Myc or HA tag antibodies.

4.4 Loss of SFPQ causes activation of interferon and inflammation response through the cGAS-STING pathway

4.4.1 RNA-seq revealed activation of antiviral response and innate immunity pathways upon SFPQ depletion

To fully characterize the impact of the loss of SFPQ on the cell physiology, an RNA-seq experiment was conducted to detect gene expression alterations after SFPQ depletion.
SFPQ was transiently knock-down and total RNA was extracted 72 hours post-transfection.
Differentially regulated genes (DEG) analysis with log fold change (FC) of >1.5 and a P value < 0.05 revealed a total of 2431 modulated genes, 1268 upregulated and 1163 downregulated, indicative of a robust response of the cells to the loss of SFPQ. In support of this, heatmap representation of DEGs showed a clear division between control and silenced conditions, with a striking opposite trend in the two reported conditions (**Figure 17A**). A principal component analysis (PCA) confirmed aforementioned results, positioning control and silenced samples away from each other, but clustered together within the single group, indicating similarity of replicates, and differences between conditions (**Figure 17B**). Moreover, volcano plot representation of DEGs strengthened previously obtained results, showing numerous genes with differential expression upon SFPQ knock-down, both in terms of upregulation and downregulation (**Figure 17C**).
To understand the biological response of U-2 OS cells to the loss of SFPQ, a GeneOntology (GO) pathway enrichment analysis was performed. Balloon plot representations (**Figure 17D**) showing top 25 up-regulated pathways revealed enriched biological terms related to cytokine activity, innate immunity and response to virus. This indicates indicating that loss of SFPQ programs gene expression towards inflammatory and immune response. Genes belonging to the top 5 up-regulated pathway are shown in **Figure 17E** as a network exhibiting interactions between them.

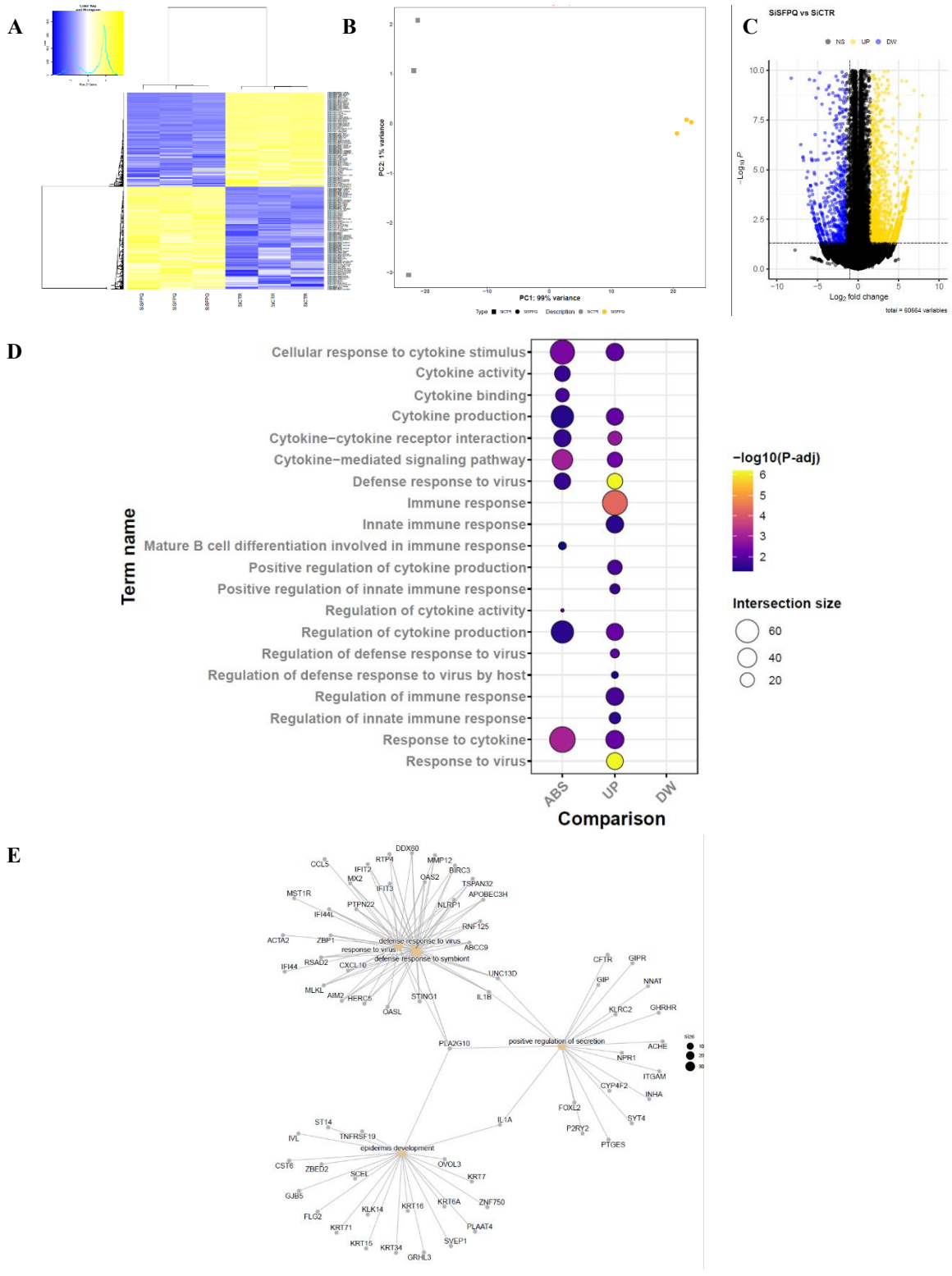


Figure 17. RNA-seq experiment revealed great difference between the control and loss of SFPQ condition, indicating activation of immune response and viral defence.

A) Heat map, **B**) PCA analysis and **C**) Volcano plot showing DEG in loss of SFPQ compared to control U-2 OS cells. **D**) Balloon plot showing top 25 upregulated pathways in loss of SFPQ compared to control U-2 OS cells. **E**) network representation of the top 25 upregulated genes in loss SFPQ condition, indicating their reciprocal connection and specific pathway membership. Log FC>1.5, p<0.05. Three independent replicates were analysed per condition.

4.4.2 Loss of SFPQ triggers activation of the cGAS-STING-IRF3 pathway in cancer cells

Viral infections activate innate immunity pathway thorough viral cytosolic DNA sensed by the cGAS-STING pathway [336], [337]. The dsDNA binding protein cGAS is known to bind to cytosolic DNA that can originate from different sources, ranging from pathogen to cell-derived DNA, such as mitochondrial DNA or genomic DNA. In particular, micronuclei (MN) originated by exacerbated genomic instability has been demonstrated to strongly activate cGAS and the downstream pathways, leading to the activation of interferon stimulated genes (ISG) and NF-kB mediated inflammation [346], [350], [367]. Hence, we wanted to test whether loss of SFPQ may causes accumulation of micronuclei that lead to activation of the cGAS-STING pathway.

First, we investigate the impact of the loss of SFPQ on the genomic stability, known source of cytoplasmatic DNA [105]. Chromosomal studies were conducted to evaluate the macroscopic effects of unproper R-loop resolution by analysing mitotic chromosome structure defects, as SFPQ has been described to have a role in chromosome architecture [219]. Indeed, we found that loss of SFPQ caused increased number of recombination events and chromosome breaks in Giemsa stained metaphase chromosomes spreads of U-2 OS, thus confirming a genomic stability protecting function of SFPQ (P. Veneziano Broccia, unpublished work).

We have previously demonstrated that SFPQ prevents pATR activation, DNA damage, and recombination at telomeres of metaphase chromosomes [227]. Along with acting at telomeres, SFPQ and DAXX work also at centromere, as described in section 4.1.3. Recent work demonstrated that the R-loop driven, ATR-Chk1-AurB pathway is essential for faithful chromosome segregation, as an increase in ATR activation causes DNA damage and mitotic defects [103].

As loss of SFPQ causes increased R-loop levels and subsequent over-activation of ATR, we decided to test whether cells lacking SFPQ present increased R-loop activated pathways at centromeres, providing novel evidence for SFPQ function in controlling chromosome segregation in an R-loop mediated manner.

We decided to perform immunofluorescence on metaphase chromosomes spreads of U-2 OS cells silenced for SFPQ or RNaseH1, used ad positive control. After 72 hours of silencing, cells were treated with Colcemid (200 μg/μl), metaphase cells were harvested, splashed by the use of a cytospin and immunofluorescence was performed by using CREST antibody combined with anti-pATR(Thr1989) or anti-pRPA32Ser33.

Upon loss of SFPQ, an increased localisation of pATR at the pericentromere (adjacent to CREST centromere signal) was detected (**Figure 18A**), as well as an increased pRPA32 colocalization with the centromere (**Figure 18B**), thus indicating over-activation of the ATR signalling pathway. As the S9.6 antibody was unsuccessful in detecting R-loop on metaphase chromosomes, we decided to perform an IF RNA-FISH on metaphase spreads combining CREST antibody and a SatIID pericentromeric probe, using U-2 OS cells as in the previous experiment. As expected, an increased signal of pericentromeric transcripts was detected upon silencing of SFPQ (**Figure 18C**).

Altogether, these findings provide strong evidence for SFPQ protecting genomic instability by suppressing unscheduled R-loop at (peri)centromeres, protecting the cell from unproper ATR activation, allowing faithful chromosome segregation.

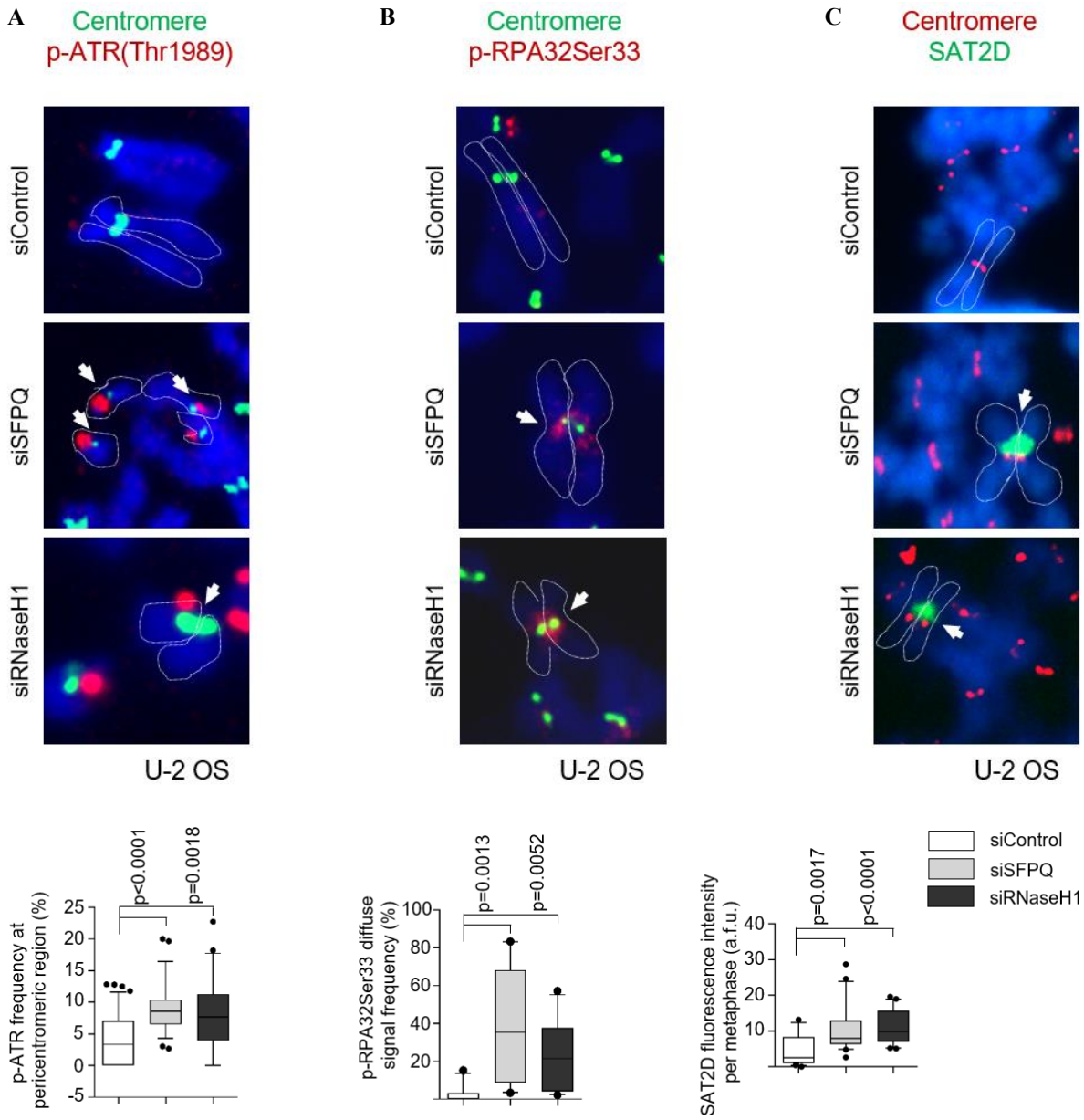


Figure 18. SFPQ suppresses unproper R-loop – ATR signalling to ensure chromosome stability

A) Up, IF representative images of pATR-CREST co-stained U-2 OS cells metaphase spread silenced with indicated SiRNAs. Down, percentage of pATR localization at pericentromeric (near CREST) measured as the percentage of p-ATR signal per number of chromosome in one metaphase, under reported silencing conditions. **B)** Up, IF representative images of pRPA32-CREST co-stained U-2 OS cells metaphase spread silenced with indicated SiRNAs. Down, quantification of diffuse p-RPA32Ser33 frequency on mitotic chromosomes, measured as the percentage of p-RPA32Ser33 signal per number of chromosome in one metaphase, under reported silencing conditions. **C**) Up, IF RNA-FISH representative images of Sat2D probe-CREST co-stained U-2 OS cells metaphase spread silenced with indicated SiRNAs. Down, quantification Sat2D intensity levels per metaphase chromosome, under reported silencing conditions. Data show means ± SEM of three biological replicates, for a minimum of 30 metaphase and 500 chromosomes. Student t-test was used to calculate statistical significance. Arrowheads indicate co-localization events.

We then investigated micronuclei formation upon SFPQ knock-down U-2 OS cell, by staining them with DAPI and inspected by fluorescence microscopy to individuate the presence of micronuclei located in the cytoplasm.

SFPQ removal caused a strong increase in the number of micronuclei compared to the control condition (**Figure 19A**). This is consistent with previous results showing that loss of SFPQ causes increased R-loop levels, activation of DNA damage markers and increased recruitment of HR proteins at repetitive regions, indicative of genomic instability onset, presumably leading to micronuclei formation.

We next wanted to test whether the increased number of micronuclei was paralleled by increased cGAS activation. To test this hypothesis, immunofluorescence analysis on U-2 OS cells silenced for SPFQ was performed using anti-cGAS antibody, and the number of cGAS positive micronuclei was counted. As expected, cells silenced for SFPQ showed a strong increased in the percentage of cGAS positive micronuclei compared the control ones, indicating the activation of the cGAS-STING pathway (**Figure 19B**).

We earlier demonstrated that SFPQ is protecting repetitive regions from harmful R-loop formation and DNA damage. We thus hypothesised that, when removing SFPQ from cells, repetitive regions should be more prone to damage accumulation and rupture, thus being over-represented in micronuclei. To prove this, we performed an immunofluorescence combined with DNA-FISH (IF DNA-FISH), using anti-TRF2 and anti-CREST antibody, as well as a PNA FISH probe for SatII sequences, that allow the localization of telomeres, centromeres and pericentromeres, respectively. U 2-OS cells were transiently transfected with control or SFPQ targeting SiRNA. 72 hours post-transfection, cells were fixed and the immunofluorescence experiment, followed by DNA-FISH was performed. We then inspected DAPI visible micronuclei for the presence or absence of signals from the desired regions. We found that cells lacking SFPQ presented a greater number of micronuclei enriched for telomeric, centromeric and pericentromeric regions (**Figure 19C**), reinforcing the idea that SFPQ is important to prevent repetitive sequences rupture, thus protecting genomic integrity.

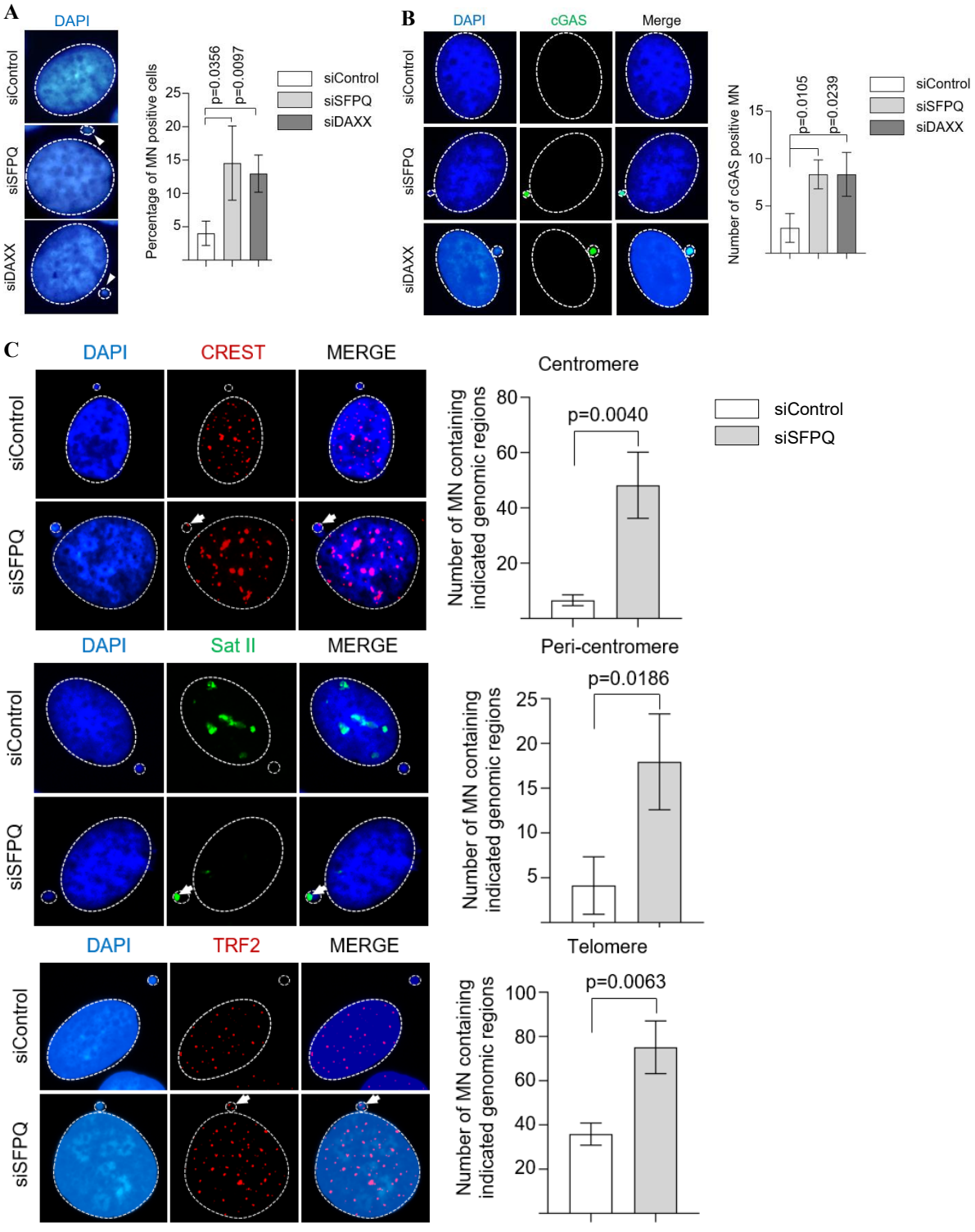


Figure 19. Silencing of SFPQ causes micronuclei formation and cGAS activation

A) Left, representative images of DAPI stained U-2 OS cells silenced for SFPQ or DAXX. Right, percentage of micronuclei (MN) positive cells. Part of experiment shown in **Figure 19A**. **B**) Top, IF representative images of anti-cGAS stained U-2 OS cells silenced for SFPQ or DAXX. Bottom, percentage of MN positive cells. **C**) Representative imagines and quantification of the number of MN containing stained genomic regions of U-2 OS cells silenced for SFPQ, stained for CREST (centromere, IF, top), SatIID (pericentromere, DNA FISH, middle), and TRF2 (telomere, IF, bottom). Data show means ± SEM of three biological replicates for a total of 90 nuclei. Student t-test was used to calculate statistical significance.

To confirm that cGAS activation of downstream signalling events occurs via STING we monitored expression and activation of IRF3 and NF-kB transcription factors [336], [337], [368].

We firstly inspected by immunofluorescence the activation of IRF3 by staining cells with an antibody specifically recognising its phosphorylated form (pIRF3 S396), known to be sufficient to activate downstream signalling [336], [337]. U-2 OS cells knocked down for either SFPQ or DAXX showed enriched pIRF3 signal compared to the control (**Figure 20A**), confirming activation of the transcription factor. In parallel, the same experiment was carried out using an antibody recognising total IRF3. Again, increased levels of total protein were detected in silenced cells, reinforcing previous results (**Figure 20B**).

Moreover, western blot analysis revealed the activation of the other branch of STING signalling, driven by NF-kB. Silencing of SFPQ triggered strong increase in p-p65, a core component of the NF-kB complex, p-IKBα, transcriptional target of NF-kB, and p-STAT3, a critical transcription factor involved in response to inflammatory stimuli (**Figure 20C**), indicating activation also of this axis. p-ATR was shown as control for loss of SFPQ mediated replication stress marker.

We repeated the immunofluorescence experiment staining two other cell lines (H1299 and OVCAR2) using an anti-pIRF3 antibody. Upon loss of SFPQ IRF3 phosphorylation increased, indicating that the pathway activation is a cell-independent, common feature (**Figure 20D**).

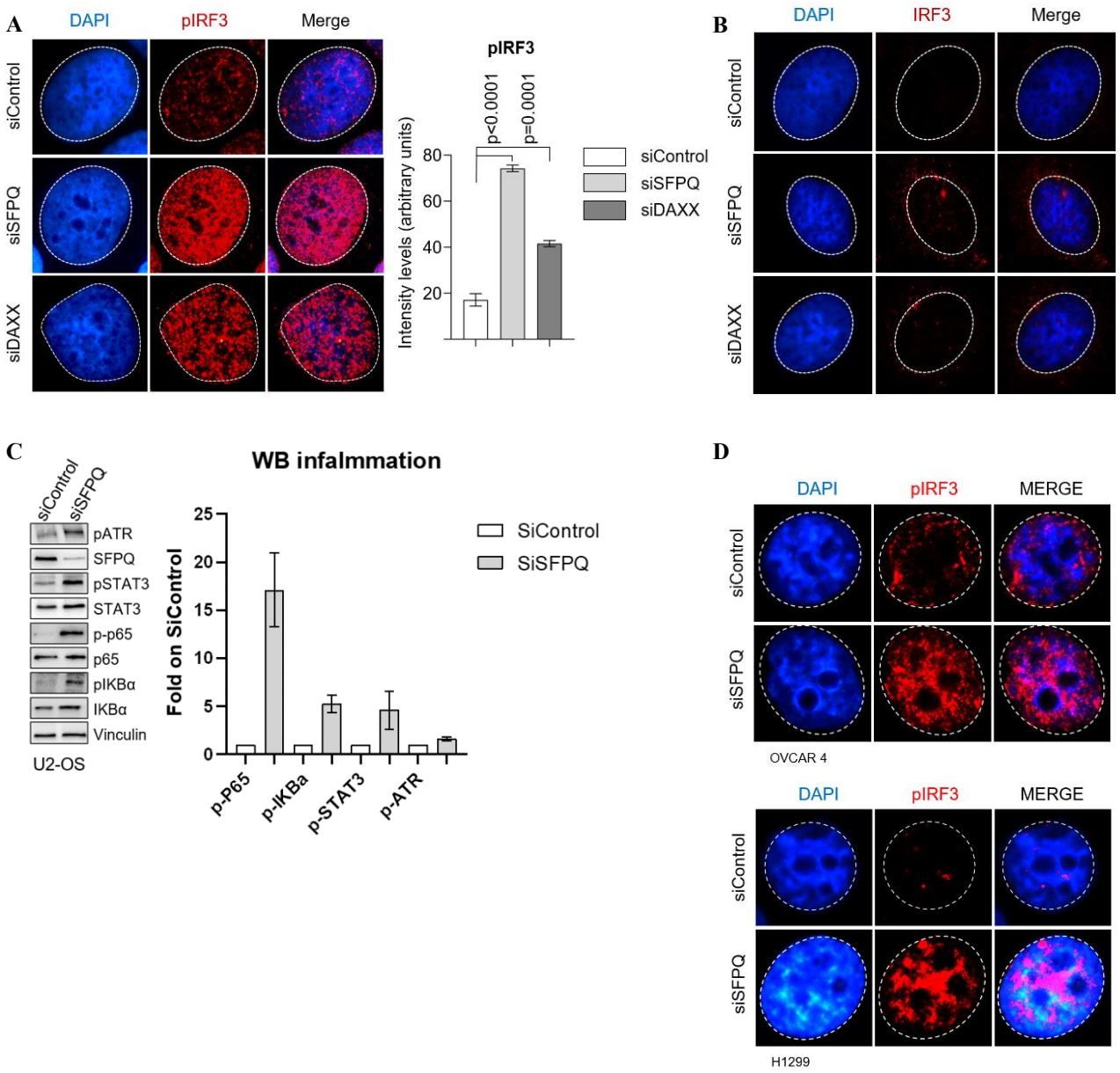


Figure 20. Silencing of SFPQ causes activation of pIRF3

A) Left, IF representative images of pIRF3 stained U-2 OS cells silenced for SFPQ or DAXX. Right, average pIRF3 intensity levels under silencing conditions. **B**) IF representative images of total IRF3 stained U-2 OS cells silenced for SFPQ or DAXX. **C**) WB analysis (left) and relative quantification (right) of NF-kB inflammatory pathway in U-2 OS cells silenced for SFPQ. **D**) Validating IF representative images of pIRF3 stained OVCAR4 (top) or H1299 (bottom) cells silenced for SFPQ. IF data show means ± SEM of four biological replicates for a total of 90 nuclei. WB data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.

We next used Real-Time PCR experiment to evaluate the activation of IRF3 and NF-kB target genes. It has been reported that IRF3 activates CCL5, CXCL10 and IFNβ gene expression, that in turn will activate ISGs (DDX60, IFIT1 and IFI44). By contrast, NF-kB is known to activate interleukin production (e.g. IL6, IL8, IL12, IL23).

RT-PCR experiment confirmed that, upon loss of SFPQ, the majority of the genes analysed resulted up-regulated (with the exception of IL6 and IL23), validating simultaneously both RNA-seq analysis and activation of the pathways (**Figure 21A**).

Of note, loss of DAXX condition didn't recapitulate what observed under SFPQ knock-down (**Figure 21B**). Western blot analysis showed that, upon loss of DAXX, pSTAT3 levels didn't change (**Figure 21C**). This indicates that, even if the upstream phenotype are shared by SFPQ and DAXX knock-down conditions (increased R-loops formation, micronuclei number, and IRF3 phosphorylation), the downstream transcriptional landscape seems to vary between the two conditions, suggesting DAXX involved in signal transduction, thus SFPQ as the most promising target to activate a potent inflammatory response.

We repeated the Real-time experiment in SFPQ knock-down OVCAR4 and H1299 cell lines. Upon loss of SFPQ both cell lines activated innate immune signalling resembling what observed in U-2 OS, but pointing out interesting differences. Indeed, ovarian cancer cells showed a strong interferon response and a drastic increase in IL6, in opposite directions with osteosarcoma, yet a negligible up-regulation of cGAS-STING components. By contrast, lung cancer cells shows a mild activation of interferon and inflammation, not even reaching a 2 fold increase, with the exception of CXCL10, which is actually down-regulated (**Figure 21D**).

Together, this results confirmed the cell independent activation of SFPQ in preventing cGAS-sTING-IRF3 pathway activation, but pointed out the cell dependency of the inflammatory response upon its loss, suggesting that in different contexts SFPQ removal might result in different outcomes and proposing osteosarcoma as the most effective in terms of desired response.

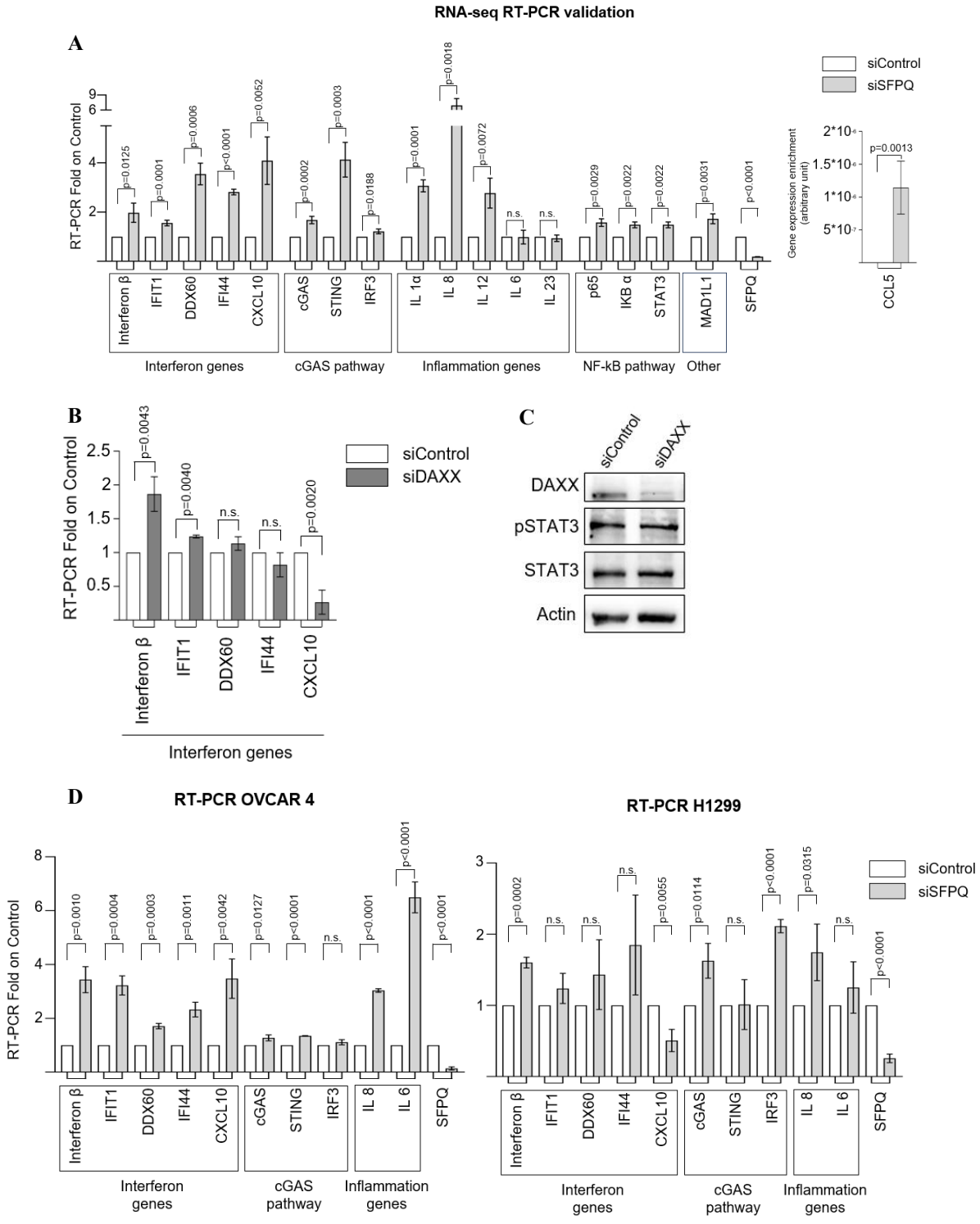


Figure 21. SFPQ prevents activation of innate immunity and interferon genes in different cell lines

A) Gene expression of indicated inflammatory genes transcripts under reported silencing of SFPQ determined by RT-qPCR. CCL5 enrichment is reported as arbitrary unit as no signal was detectable in control condition **B**) Gene expression of indicated interferon genes transcripts under reported silencing of DAXX determined by RT-qPCR. **C**) WB analysis of STAT3 in U-2 OS cells silenced for DAXX. **D**) Validation experiment of gene expression of indicated inflammatory genes transcripts under reported silencing of SFPQ in OVCAR4 (left) and H1299 (right) determined by RT-qPCR. Data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.

4.4.3 Activation of the cGAS-STING pathway is R-loop dependent

To this end so we showed that loss of SFPQ caused both increased in R-loop and activation of the cGAS-STING pathway. We wondered whether unscheduled R-loops mediated by silencing of SFPQ are causative for the activation of the cGAS-STING pathway.

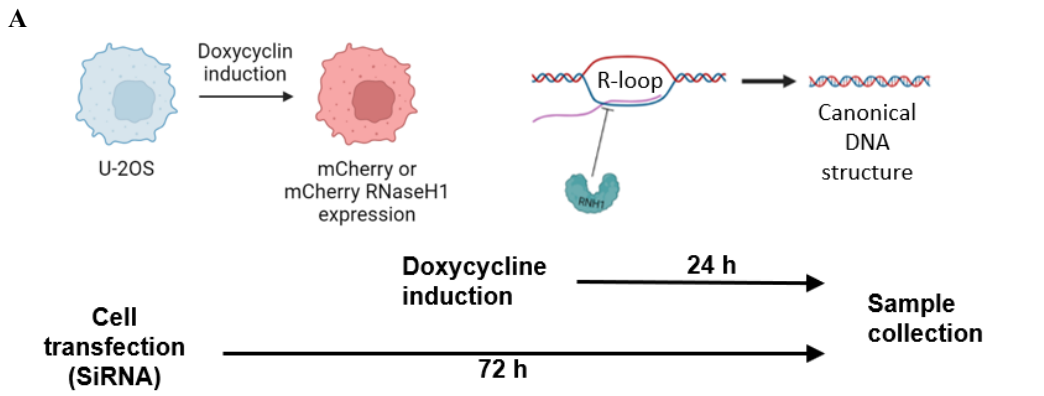
To verify this hypothesis, we took advantage of two U-2 OS cell lines able to express, in a doxycycline inducible manner, a mCherry fluorescent protein or a mCherry protein fused with RNaseH1 (mCherry-RNH1). RNaseH1 degrades the RNA component of a RNA:DNA hybrid, thus resolving R-loops.

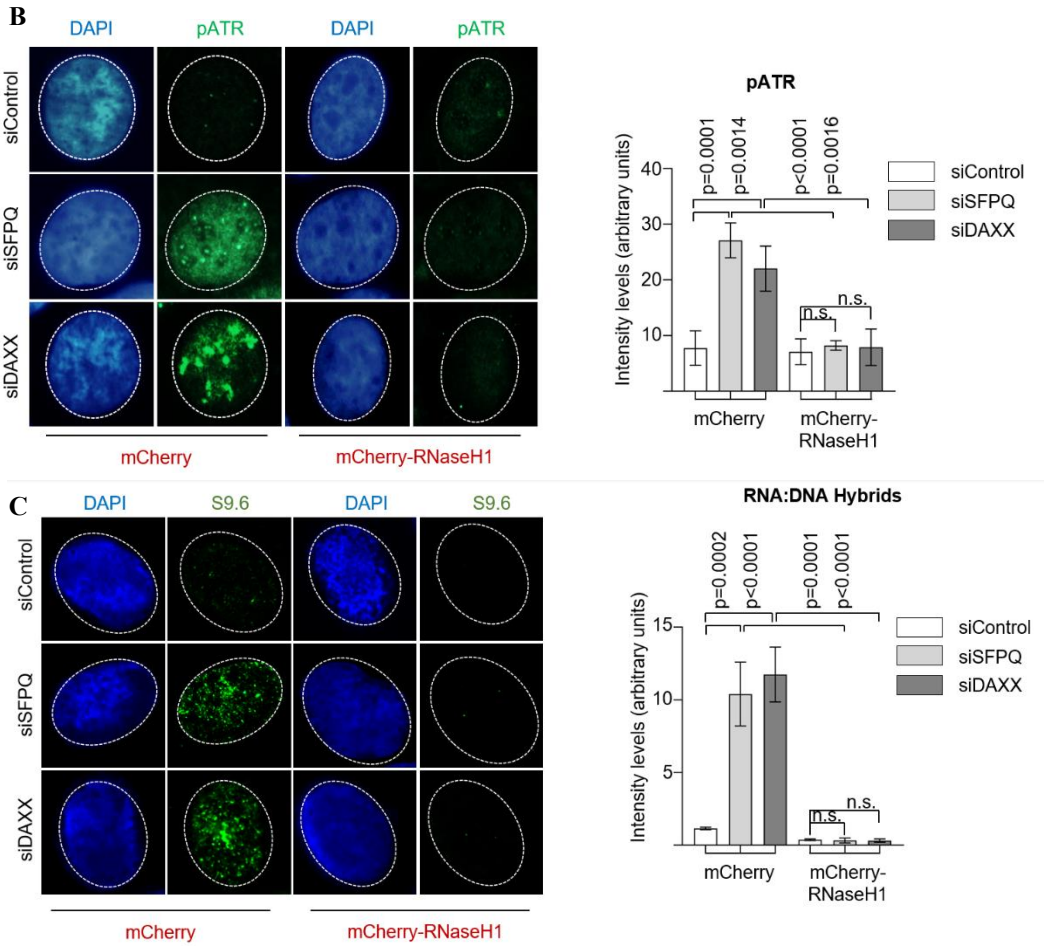
We decided to perform a rescue of phenotype experiment to demonstrate that the removal of the R-loops in SFPQ loss of function cells is sufficient to attenuate an innate immune response.

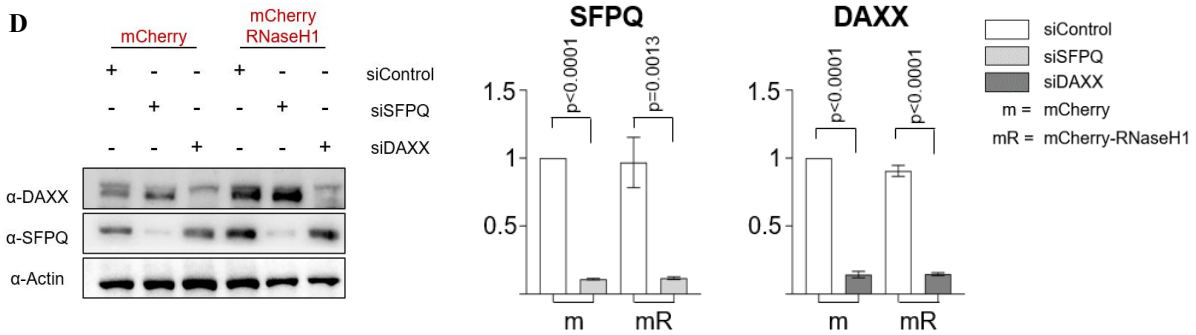
mCherry and mChrry-RNH1 U-2 OS cells were transiently silenced for SFPQ or DAXX for 72 hours. 24 h before stopping the experiment (**Figure 22A**), Doxycycline was added to induce expression of the mCherry or mCherry-RNH1 proteins.

First, we tested whether the system is actually able to suppress R-loop formation and downstream signalling (via ATR). SFPQ or DAXX knock-down in mCherry expressing cells resulted in increased RNA:DNA hybrid levels and phosphorylation of ATR. In contrast, mCherry-RNaseH1 overexpression was able to abrogate R-loop formation and ATR phosphorylation in SFPQ or DAXX loss of function cells (**Figure 22B and C**). Of note, overexpression of RNaseH1 in experimental cells did not cause an increase in SFPQ or DAXX protein, indicating that the R-loop suppression was not caused by altered SFPQ or DAXX levels (**Figure 22D**). Remarkably, RNaseH1 overexpression was able to rescue increased numbers of micronuclei upon loss of SFPQ or DAXX (**Figure 22E**), consistent with current literature [350].

We then performed RT-PCR to validate the expression of marker genes for innate immunity, that resulted reduced upon knock-down of SFPQ and concomitant expression of mCherry-RNH1, when compared to SFPQ knock-down in mCherry cells (**Figure 22F**). This experiment allowed us to identify R-loops as driving force for the activation of innate immunity pathways upon removal of SFPQ.


A
U-2OS
Doxycyclin induction
mCherry or mCherry RNaseH1 expression
R-loop
Canonical DNA structure
Doxycycline induction
24 h
Sample collection
Cell transfection (SiRNA)
72 h


B
DAPI
pATR
DAPI
pATR
siControl
siSFPQ
siDAXX
mCherry
mCherry-RNaseH1
pATR
Intensity levels (arbitrary units)
p=0.0001
p=0.0014
p<0.0001
p=0.0016
n.s.
n.s.
siControl
siSFPQ
siDAXX
mCherry
mCherry-RNaseH1
C
DAPI
S9.6
DAPI
S9.6
siControl
siSFPQ
siDAXX
mCherry
mCherry-RNaseH1
RNA:DNA Hybrids
Intensity levels (arbitrary units)
p=0.0002
p<0.0001
p=0.0001
p<0.0001
n.s.
n.s.
siControl
siSFPQ
siDAXX
mCherry
mCherry-RNaseH1


D
mCherry
mCherry RNaseH1
siControl
siSFPQ
siDAXX
α-DAXX
α-SFPQ
α-Actin
SFPQ
p<0.0001
p=0.0013
DAXX
p<0.0001
p<0.0001
m
mR
siControl
siSFPQ
siDAXX
m = mCherry
mR = mCherry-RNaseH1

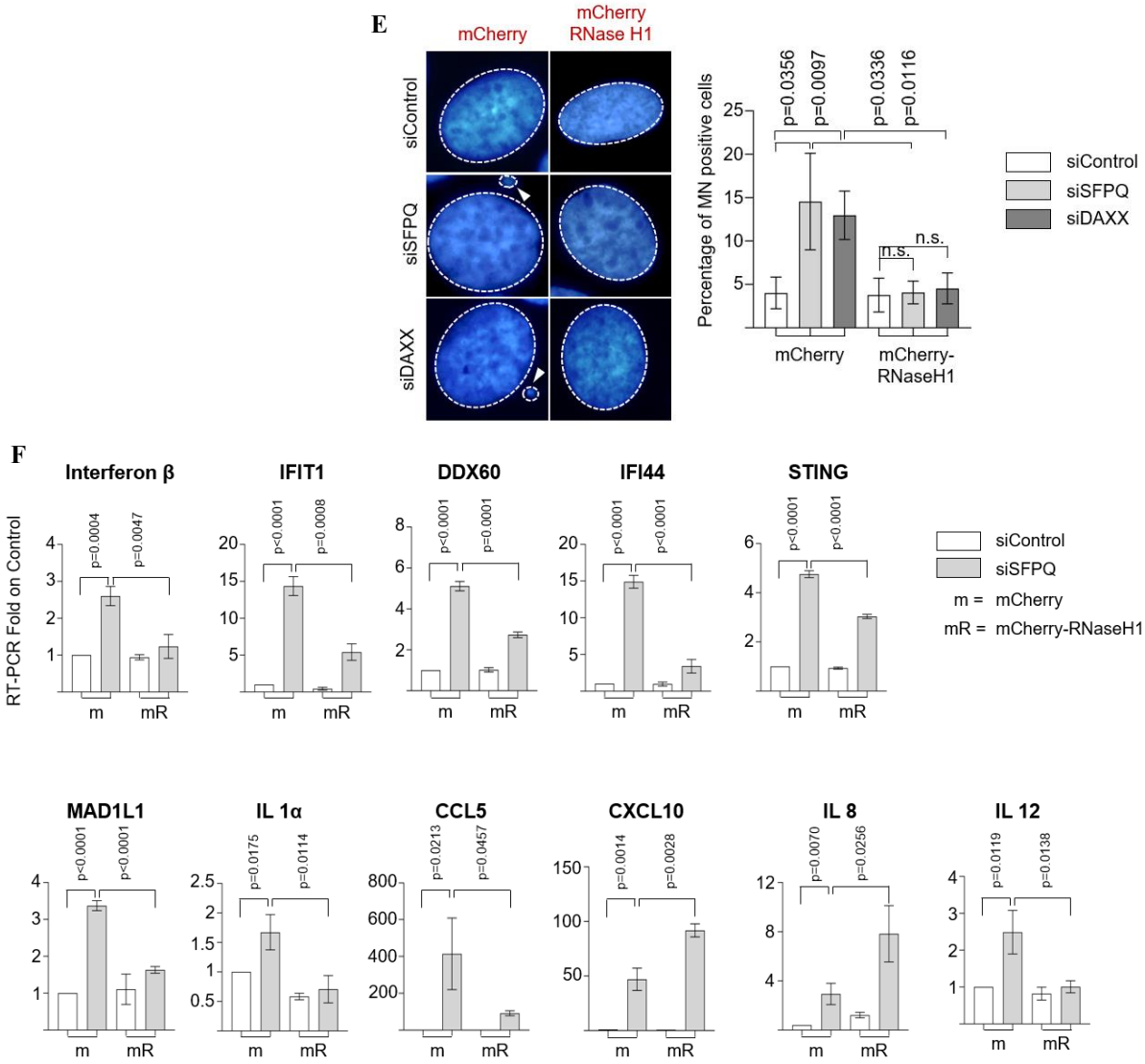


Figure 22. RNaseH1 overexpression is sufficient to revert innate immunity activation upon silencing of SFPQ

A) Schematic representation of rescue of phenotype experimental set-up. **B**) Left, IF representative images of pATR stained U-2 OS cells silenced for SFPQ or DAXX upon doxycycline induction. Right, average pATR intensity levels per nucleus upon silencing and doxycycline induction. **C**) Left, IF representative images of S9.6 stained U-2 OS cells silenced for SFPQ or DAXX upon doxycycline induction. Right, average S9.6 intensity levels per nucleus upon silencing and doxycycline induction. **D**) Western Blot (left) and RT-PCR (right) assays to verify expression levels of indicated proteins under reported silencing conditions. **E**) Left, representative images of DAPI stained U-2 OS cells silenced for SFPQ or DAXX, in combination with overexpression of RNaseH1. Right, percentage of micronuclei (MN) positive cells. Complete experiment of **Figure19 A**. **F**) Gene expression of indicated inflammatory genes transcripts under silencing of SFPQ in combination with doxycycline induced RNaseH1 overexpression determined by RT-qPCR. IF data show means ± SEM of four biological replicates for a total of 90 nuclei. RT-PCR Data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.

Remarkably, two genes (IL8 and CXCL10) showed divergent regulation in our experimental set up related to SFPQ loss of function condition.

IL8 has been previously reported to be directly regulated by SFPQ in a negative manner by binding to its promoter. Thus, removing SPFQ *per se* may allow IL8 expressions bypassing an R-loop dependent regulatory mechanism [239].

For CXCL10, we hypothesise a possible involvement of a DAXX dependent regulatory step . This may explain that silencing of DAXX causes increase in R-loop levels but at the same time suppress CXCL10 expression (**Figure 21B**). Recent studies report that CXCL10 is up regulated by activated STAT3 [369]–[371]. We found that loss of SFPQ activates STAT3 explaining CXCL10 up-regulation (**Figure 20C**). In contrast, DAXX knock-down does not activate STAT3 (**Figure 22E**), suggesting a lack of signal transduction required for CXCL10 up-regulation. Whether this action is controlled directly or indirectly by DAXX still has to be demonstrated, but this result provides a first line of evidence for a possible DAXX-mediated modulation of innate immunity and inflammatory response.

To finally demonstrate the involvement of the cGAS-STING pathway in modulating interferon response and immunity genes activation, we decided to take advantage of the STING inhibitor H151, reported to prevent STING palmitoylation and clustering, blocking its trafficking from the ER to the Golgi apparatus and impeding signal transduction [372], [373]. U 2-OS cells, knocked down for SFPQ for 72h, were treated with 2 μM H151 for 24 h before stopping the experiment (**Figure 23A, top**). Cells were then fixed and stained with anti-pIRF3 antibody, to ensure pathway activation activation. As previously described, knock-down of SFPQ causes activation of pIRF3 that, in turn, is rescued by H151 mediated STING inhibition, confirming the STING mediated activation of the pathway (**Figure 23A, bottom**). Real-time PCR experiment performed by treating cells with H151 STING inhibitor in combination with the knock-down of SFPQ resulted in decreased transcription of target genes, indicating that STING signal transduction in required for gene activation (**Figure 23B**) .

To note, IL8 and STING didn't followed this trend. However, we have already described IL8 transcription regulation as SFPQ dependent but R-loop independent, indicating that concordantly, neither STING is involved in its regulation. STING upregulation, by contrast, might be due to a compensatory mechanism acted by the cell to counteract its inhibition.

Altogether, these results indicate that the inflammatory response caused by the loss of SFPQ is actually R-loop dependent and is signalled by STING, thus providing evidences for a R-loop mediated activation of the cGAS-STING pathway.

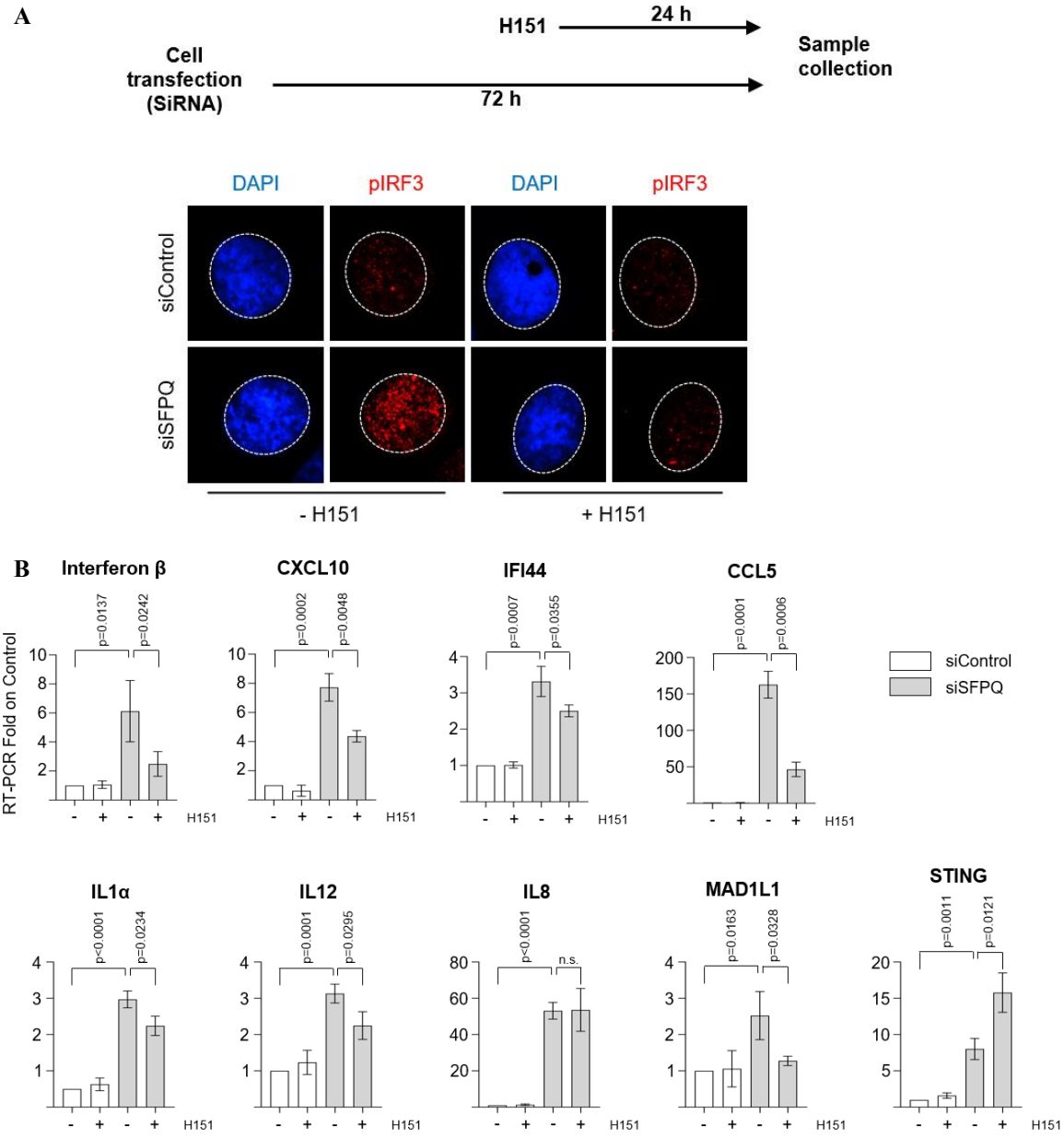


Figure 23. STING inhibition by H151 impairs activation of interferon response upon loss of SFPQ

A) Top, schematic representation of experimental set-up. Bottom, IF representative images of pIRF3 stained U-2 OS cells silenced for SFPQ in combination with 2 µm H151 STING inhibitor treatment. DMSO used as negative control. **B**) Gene expression of indicated inflammatory genes transcripts under silencing of SFPQ in combination with H151 STING inhibitor treatment determined by RT-qPCR. Data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.

4.4.4 SFPQ-DAXX interaction destabilisation recapitulates loss of SFPQ phenotypes

A possible therapeutic approach proposed for cancer treatment relies on the usage of little molecules (e.g. aptamers) that can disrupt protein-protein interaction, inhibiting downstream effects [374]–[376].

As a proof of concept for the applicability of this tool to our system, we wanted to investigate whether impairing SFPQ-DAXX interaction by using the P-domain can result in DAXX tethering, impaired complex functions, and activation of immunity pathways.

To this extent, U-2 OS cells were transiently transfected with a vector expressing the P-domain or its mutated version (nFS). 48 hours post transfection, cells were fixed and an immunofluorescence was performed, staining cells using I) anti-pATR, II) anti-pIRF3 and III) anti-DAXX antibodies.

Upon overexpression of the P domain, but not its mutated forms, we detected an appreciable increase in pATR and pIRF3 levels (**Figure 24A, left and centre**), concomitant with DAXX delocalisation (**Figure 24A, right**). This confirms a similarity between SFPQ silencing and overexpression of the P domain, suggesting that the P domain alone is able to compete with SFPQ for DAXX binding, disrupting their interaction and activating downstream signalling. To prove innate immunity activation upon P domain overexpression, Real-time experiment was performed in cells overexpressing the P domain. As expected, overexpression of the P domain, but not its mutated form, resulted in increased transcription of ISGs, resembling loss of SFPQ condition. To note, once again CXCL10 didn't follow the seen trend, supporting the hypothesis that its regulation is DAXX dependent, thus differing from the others (**Figure 24B**).

These experiments indicate that P domain overexpression recapitulates loss of SFPQ phenotype, postulating the use of interfering molecules for destabilizing the SPFQ-DAXX interaction as a putative therapeutic approach able to indue activation of innate immunity and inflammation.

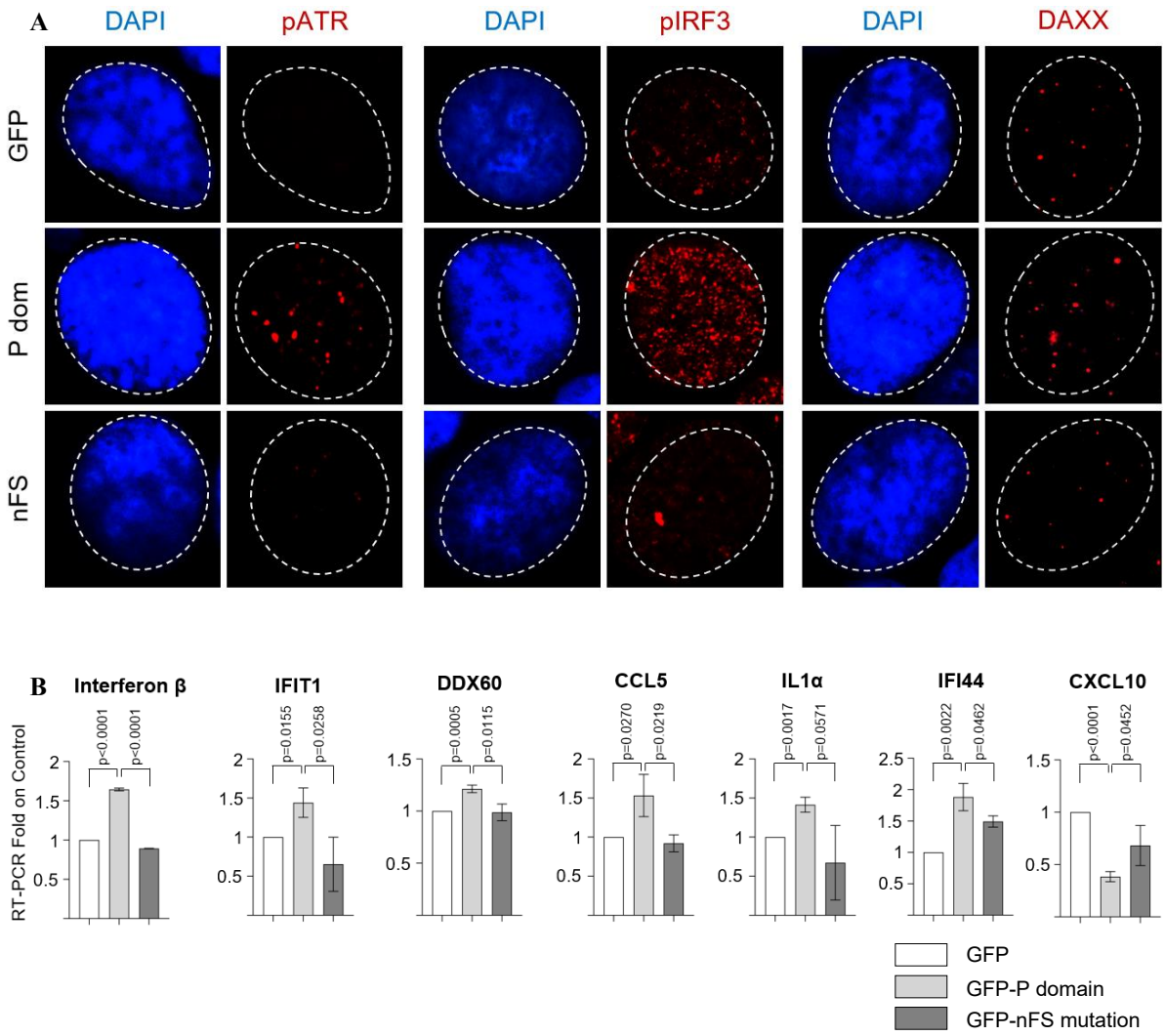


Figure 24. P domain overexpression destabilises SFPQ-DAXX interaction, recapitulating loss of SFPQ condition

A) Gene expression of indicated inflammatory genes transcripts upon overexpression of GFP, GFP-P dom or its mutated form (GFP-nFS P dom), determined by RT-qPCR. **B**) IF representative images of pATR (left), DAXX (right) and pIRF3 (bottom) stained U-2 OS cells overexpressing GFP, GFP-P dom or its mutated form (GFP-nFS P dom). Data show means ± SEM of three biological replicates. Student t-test was used to calculate statistical significance.

4.4.5 Clinical relevance of impairing SFPQ function in sarcoma

Performed experiments provide strong evidence for SFPQ-DAXX interaction destabilization as putative therapeutical target in osteosarcoma due to innate immunity and antiviral response activation.

To test whether the innate immune response due to the loss of SFPQ can have a pro or anti-tumoral activity, we investigated a TCGA dataset of sarcoma patients building Kaplan-Meier plots by stratifying the patients according to the expression of single genes. Strikingly, sarcoma patients with low expression of SFPQ showed a higher survival rate compared to high expression SFPQ patients. Conversely, patients with high levels of ISGs (CCL5, DDX60, STING) showed longer survival time compared to the low expression ones. By contrast, DAXX does not stratify the patients (**Figure 25A**). Moreover, Kaplan Meier plots analysing combination of SFPQ and ISGs, provided surprising results. While high SFPQ/low ISG resulted in an increased patient death, low SFPQ/high ISG showed a prolonger parents survival (**Figure 25B**).

This analysis show that patient survival curves recapitulates our experimental results.

Together, these analysis provided strongly evidences for a putative clinical application of SFPQ downregulation or SFPQ-DAXX interaction destabilization, as this would translate into a better survival rate of sarcoma patients.

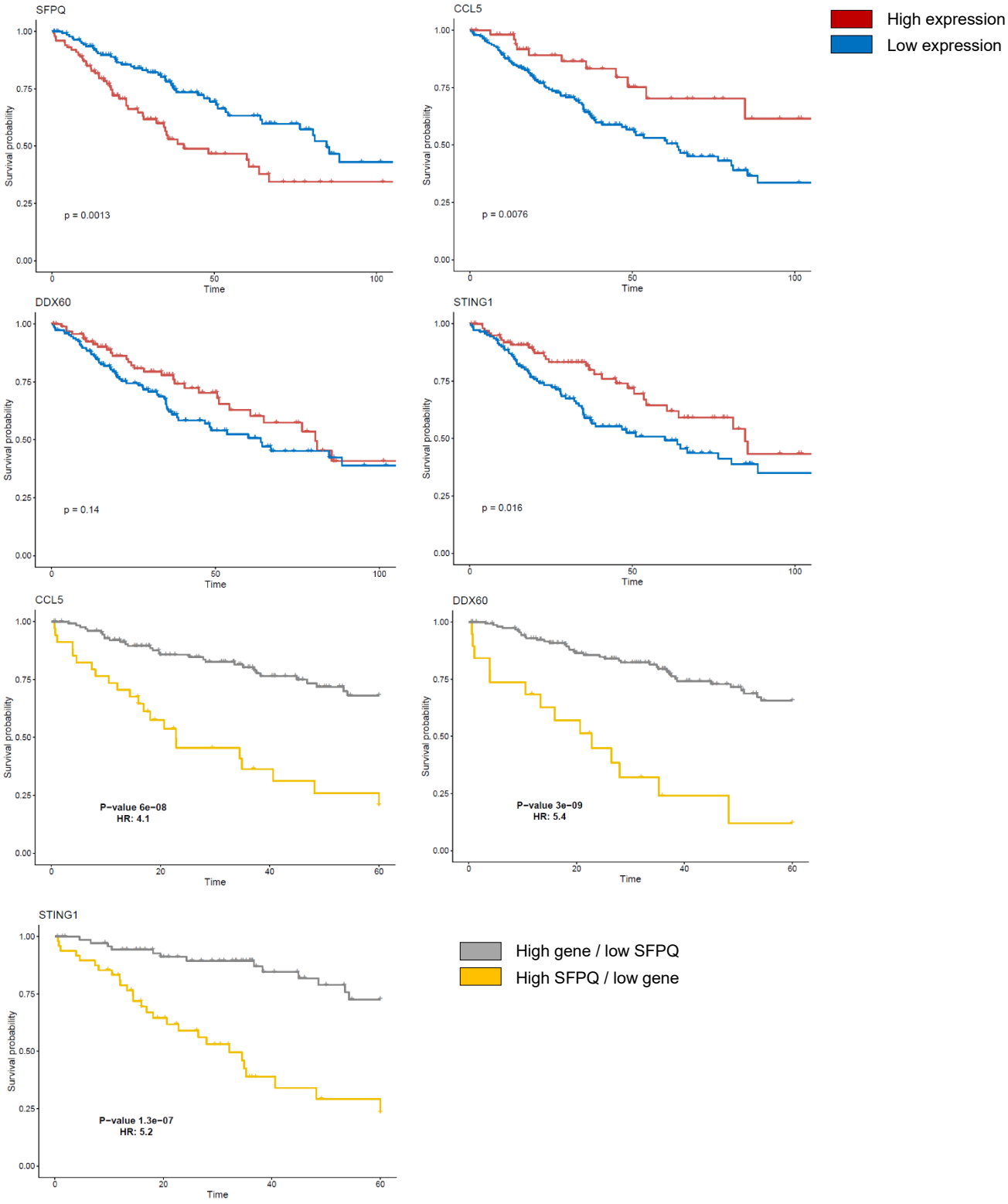


Figure 25. Kaplan Meier plots show better survival of sarcoma patients with high innate immunity genes

A) Kaplan Meier plots of sarcoma patients indicating their survival time based on the expression of indicted genes. **B**) Kaplan Meier plots of sarcoma patients indicating their survival time based on the expression of indicted genes combined with SFPQ levels. Time is in months.

5 DISCUSSION

Genomic instability is a hallmark of human cancer, with a critical role in tumour formation and progression [105]. The failure in contrasting replication stress and repairing damaged DNA are key sources of genomic instability. Persistent, uncontrolled R-loops are a known cause of replication stress, DNA breakage and mutagenesis, but also control chromatin condensation, gene silencing and mitotic stability [35], [100], [355]. Thus, R-loops are potent modulators of genomic instability and epigenetics in human cancer. R-loops form with high frequency at hotspots, defined by negative supercoiling, G-richness, CG or AT-skew, and presence of tandem repeats across the entire genome, also in non-coding and repetitive regions [8]. However, the mechanisms of R-loop management and the direct contribution of R-loops to genomic instability and cancer progression remain poorly understood.

Recently, our laboratory identified SFPQ as novel TERRA interacting protein, playing a central role in controlling telomere function by suppressing telomeric R-loops. Loss of SFPQ leads to increased levels of R-loops, replication stress, DNA damage, chromosome fragility, and homologous recombination at telomeres [227]. However, the exact mechanism employed by SFPQ in suppressing R-loops in still unknown, as well as whether its action is restricted to the telomere.

With this Ph.D. work we investigated SFPQ function at the genome-wide level and characterise its interaction with the histone chaperon DAXX. Moreover, the impact of the loss of SFPQ was investigated to unveil its impact on cell physiology.

Among a panel of proteins found to interact with SFPQ, the H3.3 specific histone chaperon DAXX represented a promising candidate. Indeed, DAXX has being already described as R-loop regulator together with its canonical partner, the ATP dependent helicase ATRX, impacting on chromatin structure and accessibility [201], [208], [209], [377]–[379].

Using co-immunoprecipitation and protein pulldown assays we were able to demonstrate the specific interaction between SFPQ and DAXX. Notably, neither SFPQ interacts with ATRX, nor DAXX interacts with NONO, indicating the SFPQ-DAXX interaction as independent of their canonical partners. Moreover, SFPQ-DAXX interaction was maintained in ATRX null U-2 OS cells, supporting our hypothesis. However, Additional experiments involving gel-filtration will be necessary to demonstrate the formation of physiologically relevant SFPQ-DAXX complexes in cells.

We showed that both SFPQ and DAXX have a common role in suppressing R-loops at telomeres. To test whether SFPQ was responsible for DAXX localization, both immunofluorescence and chromatin immunoprecipitation (ChIP) indicated that loss of

SFPQ resulted in DAXX delocalization from telomeres. The concept of SFPQ as recruiting factor is supported by the fact that SFPQ possesses RNA binding domains (RRM1 and 2), as well as an RGG domain, a type of domain reported to bind both RNA or DNA [221], [228], [229], [236]. Moreover, EMSA demonstrated that SFPQ is capable of binding R-loop structures. However, SFPQ does not hold any catalytic activity *per se*. Instead, DAXX bears histone chaperon activity, but lacks nucleotide recognition domain(s) [351]. SFPQ was previously described to suppress R-loops at telomeres, whereas DAXX was reported to act at telomeres, centromeres and other heterochromatic regions[201], [227], [377]. RNA-FISH demonstrated that also SFPQ localises to different types of repetitive regions, especially in context of increased R-loop formation by knock-down of RNaseH1. These experiments postulate a genome-wide action of SFPQ in suppressing R-loop formation, mediating DAXX localization.

To prove this hypothesis, and to identify common binding sites of the SFPQ and DAXX, we performed a ChIP-Seq experiment in control and loss of SFPQ condition. We were able to confirm SFPQ and DAXX co-occupancy of many binding sites, commonly classifies as non-coding, repetitive regions, such as (peri)centromeres, satellites, and transposable elements (ALU, LINE). When investigating the impact on these binding sites upon knock-down of SFPQ, we found out that loss of SFPQ causes DAXX delocalization paralleled by reduced H3.3 abundance, and increase occupancy of DDR proteins, such as RPA, ATR, H2AX, RAD51, and FANCD2. Moreover, DNA:RNA hybrid immunoprecipitation (DRIP) confirmed increased hybrids level at analysed sites.

Altogether, these findings indicate that SFPQ is protecting the genome from unscheduled R-loops at repetitive elements that would cause DNA damage and genomic instability. Indeed, metaphase chromosomes spreads of cells lacking SFPQ showed increased activation of the ATR pathway at the centromeric regions and elevated chromosomes arms ruptures and recombination events.

Exacerbating R-loop formation to induce genomic instability is a powerful tool for treating cancer patients. Indeed, R-loop bursting agents such as ATR inhibitors VE-822 or topoisomerase II inhibitor Etoposide are currently on clinical trials or used in lung cancer treatment, respectively [380], [381]. Thus, the possibility to target SFPQ seems a promising option for cancer treatment. Currently, antisense-oligonucleotides (ASOs), morpholinos or PNA, are under clinical trials for different diseases [382]–[384]. A putative cancer treatment would require the use of an ASO targeting SFPQ, resembling siRNA mediated SFPQ knock-down.

Alternatively, the use of molecules that destabilize SFPQ-DAXX interaction may result in similar outcomes to loss of SFPQ, indicating an alternative therapeutic strategy. We consequently decided to dissect SFPQ-DAXX interaction to identify SFPQ domain(s) involved in DAXX recruitment. A panel of SFPQ deletion mutants were generated and co-immunoprecipitation experiments identified the proline rich P domain critical for DAXX interaction. Moreover, stable cell lines showed that ectopic SFPQ lacking the P domain is neither able to prevent ATR activation nor to control DAXX localization upon depletion of endogenous SFPQ.

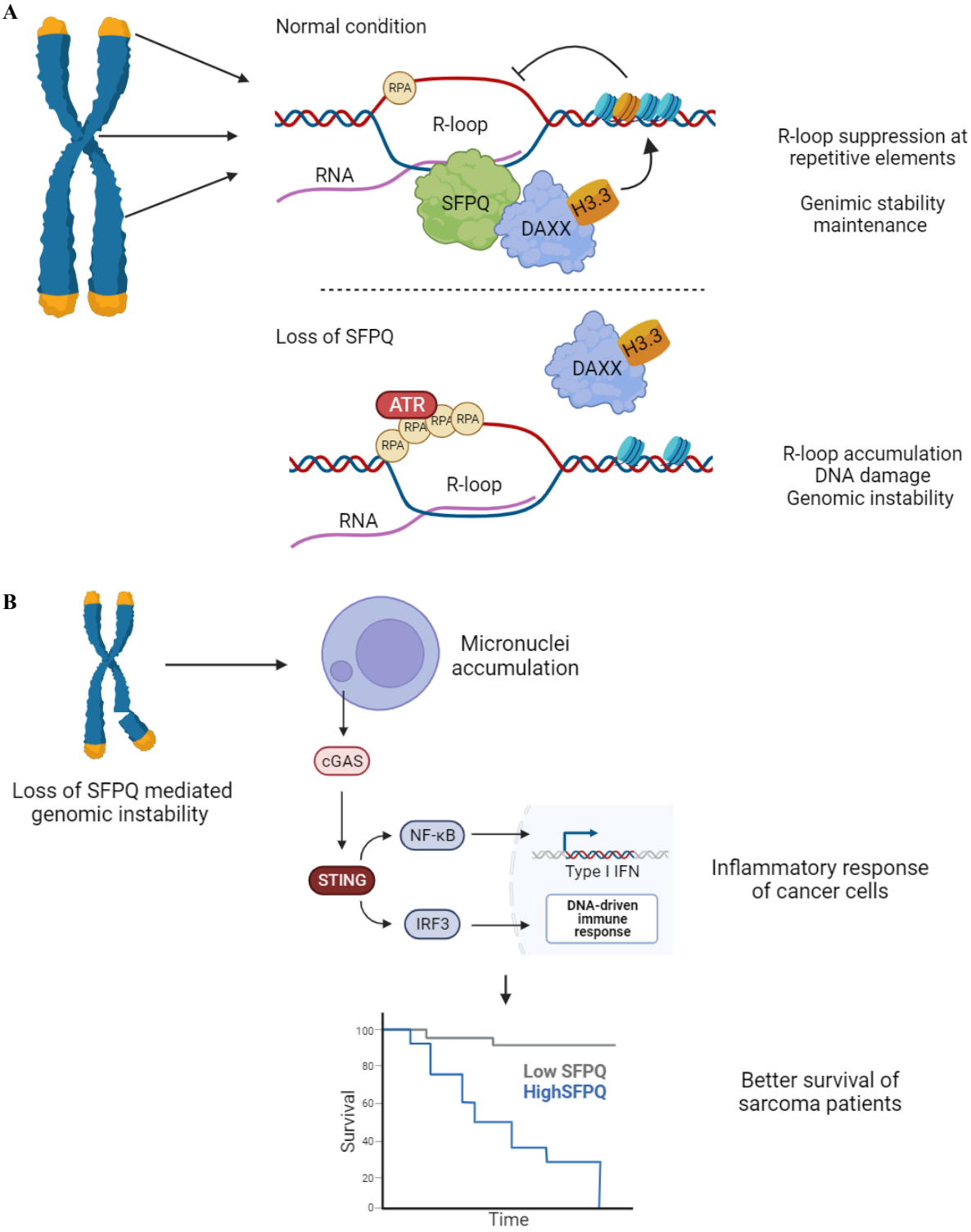
In osteosarcoma patients, two SFPQ mutations have been identified, characterized by the insertions of four (AGCG) or three (AGC) nucleotides, both located in the P domain [366]. The first mutation causes a frameshift of the protein sequence, leading to the formation of a premature stop codon in position 200 of the amino acid sequence of SFPQ. By contrast, the tri-nucleotide insertion involves the loss of threonine in position 150 of the amino acid sequence, and the gain of a lysine and a proline (T150delinsKP). Notably, the P domain tri-nucleotides insertion (named SFPQ non-frameshift mutation, nFS) recapitulates effects observed in SFPQ ΔP version. The P domain has been previously escribed to play an important role in mediating protein–protein interactions, particularly with non-DBHS proteins. As reported by Rosonina et al., deletion of the proline-rich domain abolished the ability of SFPQ to associate with the transcriptional enhancer VP16 [230]. Altogether, this confirms the central role of the P domain in regulating protein-protein interaction, including SFPQ and DAXX.

To explore the biological response linked to the loss of SFPQ, we performed RNA-seq experiment using U-2 OS cells. We found the activation of unexpected pathways, including the activation of innate immunity and response to virus. Is has been extensively reported that these pathways are triggered by the cytosolic DNA activated cGAS-STING pathway [337], [339]. Besides activation by pathogen infection, cGAS is able to sense endogenous cytoplasmatic DNA, especially under the form of micronuclei [346]. We confirmed that, upon loss of SFPQ, increased micronuclei in the cytoplasm that were able to stimulate cGAS-STING signalling, inducing interferon expression and activation of proinflammatory pathways. Interestingly, knock-down of DAXX recapitulates pathway activation (increased micronuclei, activated pIRF3), confirming that SFPQ and DAXX are acting in the same pathway. However, the downstream signalling (ISGs transcription) upon loss of DAXX didn't recapitulate effects observed with loss of SFPQ. In particular, ISGs were not upregulated upon knock-down of DAXX, with the exception of CXCL10 that was downregulated under the same experimental conditions. These differences may

be due to different mechanisms. Removing DAXX still allows SFPQ presence and retention of its additional functions, such as RNA splicing and maturation, strongly impacting the transcriptional profile. On the other hand, DAXX is also known to be involved in alternative pathways, such as signal transduction during apoptosis or regulation of transcription [194], [385]–[388]. We propose that SFPQ depletion would leave unaltered DAXX functions, but permits the activation of an inflammatory response. In contrast, depletion of DAXX will affect multiple pathways that promote alternative outcomes. In line with this, loss of SFPQ causes STAT3 activation, while loss of DAXX does not, reinforcing our hypothesis in which DAXX is involved in inflammatory signal transduction. This substantial difference let us assume that loss of SFPQ may have a more potent outcome in modulating innate immunity activation and, presumably, counteracting cancer growth. For that reason, we decided to focus our attention on the involvement of SFPQ in modulating innate immunity, considering DAXX as an essential component of response activation. By using inducible U-2 OS cell lines overexpressing RNaseH1, that resolves R-loops and prevents the activation of ATR, we were able to demonstrate that the inflammatory response occurring upon knock-down of SFPQ was R-loop dependent. This indicates that the nuclear accumulation of unscheduled R-loops drives an accumulation of cytoplasmatic DNA that activates the cGAS-STING pathway, confirmed by H151 STING inhibitor treatment. By using R-loop busting drugs, it has been demonstrated that Pyridostatin provokes micronuclei accumulation and activate an interferon response, while Camptothecin drives R-loop mediated activation of the cGAS-STING pathway [350], [367], indicating that our data are consistent with and mirrored by literature data.

We found that loss of SFPQ mediated activation of the cGAS-STING pathway is not specific for U-2 OS osteosarcoma cells, but is shared between different cell types, including high grade ovarian cancer OVCAR4 and non-small cell lung cancer H1299 cells. However, we observed that downstream signalling via cGAS-STING varies between different cancer cell types. Compared to U-2 OS cells, innate immunity activation in OVCAR4 was different in terms of magnitude of activation (e.g. IL6 was highly up-regulated), while in H1299 the activation less potent. This indicates that the outcome of the loss of SFPQ is strongly cancer cell type dependent, hypothesising differently impact on cancer progression and patient survival. By inspecting patients survival using TCGA datasets, we were able to observe that sarcoma patients with low expression of SFPQ have a better survival compared to high level patients. By contrast, patients with high ISGs expression (e.g. CCL5, DDX60, STING) present better survival compared to cohorts with low expression. These finding were not reproduced in ovarian or lung cancers. This

indicates that SFPQ-DAXX-cGAS-STING axis impairment may have particular relevance for sarcoma patients (schematic representation of the mechanism of action at **Figure 26**).



A) At repetitive regions (arrows), SFPQ acts with DAXX by suppressing R-loops, ensuring genomic stability. Upon SFPQ depletion, DAXX delocalization causes an increase in R-loops, leading to DNA damage and genomic instability. **B)** Loss of SPFQ mediated genomic instability causes micronuclei accumulation that, in turn, stimulate the cGAS-STING pathway, activating an immune response. The signature of such inflammatory response stratifies sarcoma patients, showing better survival with higher expression of innate immunity genes, postulating a therapeutic relevance of our finding on such tumors. Image generated with BioRender.com.

Sarcomas represent a highly heterogeneous group of tumors originating from connecting tissues, such as bones, cartilages, muscles, and vessels [389]. This heterogeneity is reflected on the high diversity in the genetic background, with common mutations found only on few oncogenes, such as p53 and RB. In fact, cancer treatment relies mostly on multidrug chemotherapy using canonical, unspecific drugs, such as PARP inhibitors, PD-L1 antagonists, and cytotoxic agents (Doxorubicin, Methotrexate, Cisplatin), and very limited target therapy, mostly represented by growth factor receptor inhibitors (Nilotinib, Bevacizumab), thus resulting in low precision treatments [390], [391]. The same scenario is true for osteosarcoma, confirming an unmet clinical need for better cancer treatments [392], [393]. Interestingly, mutations of paediatric osteosarcoma patients are enriched in genes involved in DNA damage repair (BRCA1/2, ATM, WRN), chromatin modification (ATRX, FANCE, RECQL4), and control of cell cycle (p53, RB, MDM2, MYC), that have been demonstrated to be involved in R-loop management [147], [177], [200], [354], [394], [395]. This suggests that R-loop busting therapies might represent a new line of treatment for osteosarcoma patients. Of note, interferon treatment in combination with chemotherapy such as Cisplatin or Etoposide has been reported to contrast osteosarcoma cells growth [396], [397]. Additionally, cells depleted for SFPQ were demonstrated to be more sensitive to Cisplatin treatments [365], together suggesting that mediating SFPQ function paired with Interferon and chemotherapy treatment may result in better patients outcome. Further experiments using more advanced preclinical models are required to support this hypothesis.

Acute loss of SFPQ causes cell death and is lethal in animal models [398]. Thus, depleting SFPQ in patients by ASOs may result in severe side effects. To circumvent this problem, alternative therapeutic approaches may consist in disrupting SFPQ-DAXX interaction, without abolishing individual protein levels. Complex destabilization should result in DAXX delocalization, increased R-loop levels, and activation of the innate immune response, without impairing alternative DAXX functions.

To produce first evidence for SFPQ-DAXX interaction destabilisation as trigger of an innate immune response, we took advantage of generated P domain variants. Ectopic overexpression of the P domain was able to recapitulate the loss of SFPQ (DAXX delocalization, ATR and IRF3 activation, ISGs increased expression). This phenotypes were not observed upon overexpression of the nFS mutated variant. These experiments suggesting that complex disruption may have therapeutic applicability.

Aptamers, synthetically modified ingle-stranded oligonucleotides, have emerged as putative cancer drugs that bind and block protein activity, including protein-protein

interaction. Aptamers hold advantages when compared to classical chemotherapy agents [399]. Aptamers have demonstrated to be stable, non-immunogenic, and capable of cell internalization. Aptamers may be used to specifically disrupt SFPQ-DAXX interaction, affecting functions only related to this couple (e.g. R-loop resolution), leaving other functions related to DAXX or SFPQ unaltered. This may result in a less severe impact healthy cell and, potentially, on the patient. Further investigations are needed to verify the applicability of such systems and their actual contribution on cancer defeat. Based on the fact that our hypothetical treatment would be focused on the increased R-loop formation, a combinatorial treatment of cancer cell with R-loop inducing drugs might represent a powerful tool to exacerbate R-loop formation and subsequent inflammatory response. For instance, VE-822, Etoposide, Camptothecin derivatives (Topoisomerase I inhibitor) or Pyridostatin (G-4 stabiliser) are drugs reported to increase R-loop accumulation, and are currently in clinics (Etoposide, Camptothecin derivatives) or in clinical trials (VE-822, Pyridostatin). Cisplatin has been reported as well to have stronger effect when simultaneously altering R-loop metabolism [400], suggesting that loss of SFPQ might act synergistically with clinical drugs to defeat osteosarcoma.

With this work, we identified SFPQ as strong suppressor of R-loops at repetitive regions at the genome-wide level, preventing the activation of the cGAS-STING pathway, thus blocking innate immunity. SFPQ is able to directly bind and suppress R-loops by mediating DAXX deposition of H3.3, impacting on the chromatin structure. When cells are depleted for SFPQ, boosted levels of R-loops are detected at repetitive regions, causing replication stress and activation of the HR dependent DNA damage repair. By suppressing R-loops, SFPQ prevents activation of the cGAS-STING pathway, avoiding innate immunity activation. This pathway may have relevance for sarcoma cancer treatment, by wiping out R-loops suppression by DAXX. Targeting SFPQ-DAXX interaction is also expected to increase inflammation that may modulate the tumour micro-environment (TME) and improve survival.

The next step of our research will be to use interfering molecules (e.g. aptamers resembling portions of the P domain or ASOs) and evaluate their role in activating innate immunity. *In vivo* studies should be performed as well, first to evaluate whether loss of SFPQ has an impact on carcer growth, and later to apply small molecules for cancer treatment.

In parallel, samples form patients with SFPQ mutations shall be analysed to determine the impact of these mutations on genomic stability, ALT, DDR, and the impact on TME (cell infiltrate, inflammation markers).

This will provide new therapeutic strategies for treating osteosarcoma patients.

6. BIBLIOGRAPHY

[1] J. Santos-Pereira and A. Aguilera, 'R loops : new modulators of genome dynamics and function .', *Nat. Rev. Genet.*, vol. 16, no. 10, p. 583.97, 2015, doi: 10.1038/nrg3961.

[2] N. N. Shaw and D. P. Arya, 'Recognition of the unique structure of DNA:RNA hybrids', *Biochimie*, vol. 90, no. 7, pp. 1026–1039, Jul. 2008, doi: 10.1016/j.biochi.2008.04.011.

[3] M. Drolet, P. Phoenix, R. Menzel, E. Massé, L. F. Liu, and R. J. Crouch, 'Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I', *Proc. Natl. Acad. Sci. U. S. A.*, vol. 92, no. 8, pp. 3526–3530, Apr. 1995, doi: 10.1073/pnas.92.8.3526.

[4] M. Tresini *et al.*, 'The core spliceosome as target and effector of non-canonical ATM signalling', *Nature*, vol. 523, no. 7558, pp. 53–58, Jul. 2015, doi: 10.1038/nature14512.

[5] R. E. Wellinger, F. Prado, and A. Aguilera, 'Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex', *Mol. Cell. Biol.*, vol. 26, no. 8, pp. 3327–3334, Apr. 2006, doi: 10.1128/MCB.26.8.3327-3334.2006.

[6] C. T. Stork *et al.*, 'Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage', *eLife*, vol. 5, p. e17548, Aug. 2016, doi: 10.7554/eLife.17548.

[7] K. S. Lang *et al.*, 'Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis', *Cell*, vol. 170, no. 4, pp. 787-799.e18, Aug. 2017, doi: 10.1016/j.cell.2017.07.044.

[8] A. Gambelli, A. Ferrando, C. Boncristiani, and S. Schoeftner, 'Regulation and function of R-loops at repetitive elements', *Biochimie*, Aug. 2023, doi: 10.1016/j.biochi.2023.08.013.

[9] A. Aguilera and T. García-Muse, 'R Loops: From Transcription Byproducts to Threats to Genome Stability', *Mol. Cell*, 2012, doi: 10.1016/j.molcel.2012.04.009.

[10] C. Niehrs and B. Luke, 'Regulatory R-loops as facilitators of gene expression and genome stability', *Nat. Rev. Mol. Cell Biol.*, vol. 21, no. 3, pp. 167–178, 2020, doi: 10.1038/s41580-019-0206-3.

[11] J. Lafuente-Barquero, M. L. García-Rubio, M. S. Martin-Alonso, B. Gómez-González, and A. Aguilera, 'Harmful DNA:RNA hybrids are formed in cis and in a Rad51-independent manner', *eLife*, vol. 9, p. e56674, Aug. 2020, doi: 10.7554/eLife.56674.

[12] L. Wahba, S. K. Gore, and D. Koshland, 'The homologous recombination machinery modulates the formation of RNA–DNA hybrids and associated chromosome instability', *eLife*, vol. 2, p. e00505, 2013, doi: 10.7554/eLife.00505.

[13] K. Toriumi, T. Tsukahara, and R. Hanai, 'R-Loop Formation In Trans at an AGGAG Repeat', *J. Nucleic Acids*, vol. 2013, p. 629218, 2013, doi: 10.1155/2013/629218.

[14] J. L. Huppert, A. Bugaut, S. Kumari, and S. Balasubramanian, 'G-quadruplexes: the beginning and end of UTRs', *Nucleic Acids Res.*, vol. 36, no. 19, pp. 6260–6268, Nov. 2008, doi: 10.1093/nar/gkn511.

[15] L. Ratmeyer, R. Vinayak, Y. Y. Zhong, G. Zon, and W. D. Wilson, 'Sequence Specific Thermodynamic and Structural Properties for DNA. cntdot. RNA Duplexes', *Biochemistry*, vol. 33, no. 17, pp. 5298–5304, 1994.

[16] P. A. Ginno, P. L. Lott, H. C. Christensen, I. Korf, and F. Chédin, 'R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters', *Mol. Cell*, 2012, doi: 10.1016/j.molcel.2012.01.017.

[17] P. A. Ginno, Y. W. Lim, P. L. Lott, I. Korf, and F. Chédin, 'GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination', *Genome Res.*, vol. 23, pp. 1590–1600, 2013, doi: 10.1101/gr.158436.113.

[18] L. Wahba, L. Costantino, F. J. Tan, A. Zimmer, and D. Koshland, 'S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation.', *Genes Dev.*, vol. 30, no. 11, pp. 1327–1338, Jun. 2016, doi: 10.1101/gad.280834.116.

[19] A. De Magis *et al.*, 'DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells', *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 3, pp. 816–825, 2019, doi: 10.1073/pnas.1810409116.

[20] K. Rippe and B. Luke, 'TERRA and the state of the telomere', *Nat. Struct. Mol. Biol.*, vol. 22, no. 11, pp. 853–858, 2015, doi: 10.1038/nsmb.3078.

[21] Y. Liu *et al.*, 'Genome-wide mapping reveals R-loops associated with centromeric repeats in maize', *Genome Res.*, vol. 31, no. 8, pp. 1409–1418, 2021.

[22] M. Groh, M. M. P. Lufino, R. Wade-Martins, and N. Gromak, 'R-loops Associated with Triplet Repeat Expansions Promote Gene Silencing in Friedreich Ataxia and Fragile X Syndrome', *PLoS Genet.*, vol. 10, no. 5, 2014, doi: 10.1371/journal.pgen.1004318.

[23] X. Li and J. L. Manley, 'Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability', *Cell*, vol. 122, no. 3, pp. 365–378, 2005, doi: 10.1016/j.cell.2005.06.008.

[24] F. Chédin, 'Nascent Connections: R-Loops and Chromatin Patterning.', *Trends Genet. TIG*, vol. 32, no. 12, pp. 828–838, Dec. 2016, doi: 10.1016/j.tig.2016.10.002.

[25] G. Miglietta, M. Russo, and G. Capranico, 'G-quadruplex-R-loop interactions and the mechanism of anticancer G-quadruplex binders.', *Nucleic Acids Res.*, vol. 48, no. 21, pp. 11942–11957, Dec. 2020, doi: 10.1093/nar/gkaa944.

[26] D. Yang, 'G-Quadruplex DNA and RNA.', *Methods Mol. Biol. Clifton NJ*, vol. 2035, pp. 1–24, 2019, doi: 10.1007/978-1-4939-9666-7_1.

[27] G. Miglietta, M. Russo, and G. Capranico, 'Correction to article "G-quadruplex–R-loop interactions and the mechanism of anticancer G-quadruplex binders"', *Nucleic Acids Res.*, vol. 49, no. 10, p. 6000, Jun. 2021, doi: 10.1093/nar/gkab483.

[28] D. Roy and M. R. Lieber, 'G Clustering Is Important for the Initiation of Transcription-Induced R-Loops In Vitro, whereas High G Density without Clustering Is Sufficient Thereafter', *Mol. Cell. Biol.*, vol. 29, no. 11, pp. 3124–3133, 2009, doi: 10.1128/MCB.00139-09.

[29] T. García-Muse and A. Aguilera, 'R Loops: From Physiological to Pathological Roles', *Cell*, vol. 179, no. 3, pp. 604–618, Oct. 2019, doi: 10.1016/j.cell.2019.08.055.

[30] J. P. Wells, J. White, and P. C. Stirling, 'R Loops and Their Composite Cancer Connections', *Trends Cancer*, vol. 5, no. 10, 2019, doi: 10.1016/j.trecan.2019.08.006.

[31] J. Sollier and K. A. Cimprich, ‘Breaking bad: R-loops and genome integrity.’, *Trends Cell Biol.*, vol. 25, no. 9, pp. 514–522, Sep. 2015, doi: 10.1016/j.tcb.2015.05.003.

[32] F. Chédin, ‘Nascent Connections: R-Loops and Chromatin Patterning.’, *Trends Genet. TIG*, vol. 32, no. 12, pp. 828–838, Dec. 2016, doi: 10.1016/j.tig.2016.10.002.

[33] P. A. Ginno, P. L. Lott, H. C. Christensen, I. Korf, and F. Chédin, ‘R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters’, *Mol. Cell*, 2012, doi: 10.1016/j.molcel.2012.01.017.

[34] L. A. Sanz *et al.*, ‘Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals’, *Mol. Cell*, vol. 63, no. 1, pp. 167–178, 2016.

[35] M. P. Crossley, M. Bocek, and K. A. Cimprich, ‘R-Loops as Cellular Regulators and Genomic Threats.’, *Mol. Cell*, vol. 73, no. 3, pp. 398–411, Feb. 2019, doi: 10.1016/j.molcel.2019.01.024.

[36] L. Wahba, L. Costantino, F. J. Tan, A. Zimmer, and D. Koshland, ‘S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation.’, *Genes Dev.*, vol. 30, no. 11, pp. 1327–1338, Jun. 2016, doi: 10.1101/gad.280834.116.

[37] W. Xu *et al.*, ‘The R-loop is a common chromatin feature of the Arabidopsis genome’, *Nat. Plants*, vol. 3, no. 9, pp. 704–714, 2017, doi: 10.1038/s41477-017-0004-x.

[38] C. Zeng, M. Onoguchi, and M. Hamada, ‘Association analysis of repetitive elements and R-loop formation across species’, *Mob. DNA*, vol. 12, no. 1, p. 3, 2021, doi: 10.1186/s13100-021-00231-5.

[39] K. Yu, F. Chedin, C. L. Hsieh, T. E. Wilson, and M. R. Lieber, ‘R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells’, *Nat. Immunol.*, 2003, doi: 10.1038/ni919.

[40] M. Muramatsu, K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai, and T. Honjo, ‘Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme.’, *Cell*, vol. 102, no. 5, pp. 553–563, Sep. 2000, doi: 10.1016/s0092-8674(00)00078-7.

[41] P. Revy *et al.*, ‘Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2).’, *Cell*, vol. 102, no. 5, pp. 565–575, Sep. 2000, doi: 10.1016/s0092-8674(00)00079-9.

[42] K. N. Kreuzer and J. R. Brister, ‘Initiation of bacteriophage T4 DNA replication and replication fork dynamics: a review in the Virology Journal series on bacteriophage T4 and its relatives’, *Virol. J.*, vol. 7, no. 1, p. 358, Dec. 2010, doi: 10.1186/1743-422X-7-358.

[43] T. Itoh and J. Tomizawa, ‘Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H.’, *Proc. Natl. Acad. Sci. U. S. A.*, vol. 77, no. 5, pp. 2450–2454, May 1980, doi: 10.1073/pnas.77.5.2450.

[44] J. L. O. Pohjoismäki *et al.*, ‘Mammalian mitochondrial DNA replication intermediates are essentially duplex, but contain extensive tracts of RNA/DNA hybrid’, *J. Mol. Biol.*, vol. 397, no. 5, pp. 11144–1155, 2010, doi: 10.1007/s10817-010-9194-x.

[45] M. Castellano-Pozo *et al.*, ‘R loops are linked to histone H3 S10 phosphorylation and chromatin condensation’, *Mol. Cell*, 2013, doi: 10.1016/j.molcel.2013.10.006.

[46] M. Nakama, K. Kawakami, T. Kajitani, T. Urano, and Y. Murakami, 'DNA-RNA hybrid formation mediates RNAi-directed heterochromatin formation', *Genes Cells*, 2012, doi: 10.1111/j.1365-2443.2012.01583.x.

[47] M. S. Ivaldi, C. S. Karam, and V. G. Corces, 'Phosphorylation of histone H3 at Ser10 facilitates RNA polymerase II release from promoter-proximal pausing in Drosophila', *Genes Dev.*, vol. 21, no. 21, pp. 2818–2831, Nov. 2007, doi: 10.1101/gad.1604007.

[48] A. Zippo, A. De Robertis, R. Serafini, and S. Oliviero, 'PIM1-dependent phosphorylation of histone H3 at serine 10 is required for MYC-dependent transcriptional activation and oncogenic transformation', *Nat. Cell Biol.*, vol. 9, no. 8, pp. 932–944, Aug. 2007, doi: 10.1038/ncb1618.

[49] R. Boque-Sastre *et al.*, 'Head-to-head antisense transcription and R-loop formation promotes transcriptional activation', *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 18, pp. 5785–5790, 2015, doi: 10.1073/pnas.1421197112.

[50] K. Skourti-Stathaki, E. Torlai Triglia, M. Warburton, P. Voigt, A. Bird, and A. Pombo, 'R-Loops Enhance Polycomb Repression at a Subset of Developmental Regulator Genes', *Mol. Cell*, vol. 73, no. 5, pp. 930-945.e4, Mar. 2019, doi: 10.1016/j.molcel.2018.12.016.

[51] J. Nadel *et al.*, 'RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships', *Epigenetics Chromatin*, vol. 8, p. 46, 2015, doi: 10.1186/s13072-015-0040-6.

[52] C. L *et al.*, 'R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters', *Mol. Cell*, vol. 68, no. 4, pp. 745-757.e5, Nov. 2017, doi: 10.1016/J.MOLCEL.2017.10.008.

[53] J. G. Dumelie and S. R. Jaffrey, 'Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq', *eLife*, vol. 6, p. e28306, Oct. 2017, doi: 10.7554/eLife.28306.

[54] S. M. Tan-Wong, S. Dhir, and N. J. Proudfoot, 'R-Loops Promote Antisense Transcription across the Mammalian Genome', *Mol. Cell*, vol. 76, no. 4, pp. 600-616.e6, Nov. 2019, doi: 10.1016/j.molcel.2019.10.002.

[55] C. Grunseich *et al.*, 'Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters', *Mol. Cell*, vol. 69, no. 3, pp. 426–437, Feb. 2018, doi: 10.1016/j.molcel.2017.12.030.

[56] L. Zardoni *et al.*, 'Elongating RNA polymerase II and RNA:DNA hybrids hinder fork progression and gene expression at sites of head-on replication-transcription collisions', *Nucleic Acids Res.*, vol. 49, no. 22, pp. 12769–12784, Dec. 2021, doi: 10.1093/nar/gkab1146.

[57] C. St Germain, H. Zhao, and J. H. Barlow, 'Transcription-Replication Collisions-A Series of Unfortunate Events', *Biomolecules*, vol. 11, no. 8, p. 1249, Aug. 2021, doi: 10.3390/biom11081249.

[58] C. Tous and A. Aguilera, 'Impairment of transcription elongation by R-loops in vitro', *Biochem. Biophys. Res. Commun.*, 2007, doi: 10.1016/j.bbrc.2007.06.098.

[59] B. P. Belotserkovskii, J. H. Soo Shin, and P. C. Hanawalt, 'Strong transcription blockage mediated by R-loop formation within a G-rich homopurine–homopyrimidine sequence localized in the vicinity of the promoter', *Nucleic Acids Res.*, vol. 45, no. 11, pp. 6589–6599, Jun. 2017, doi: 10.1093/nar/gkx403.

[60] K. Skourti-Stathaki, N. J. Proudfoot, and N. Gromak, 'Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination', *Mol. Cell*, vol. 42, no. 6, pp. 794–805, Jun. 2011, doi: 10.1016/j.molcel.2011.04.026.

[61] J. L. Huppert, A. Bugaut, S. Kumari, and S. Balasubramanian, 'G-quadruplexes: the beginning and end of UTRs', *Nucleic Acids Res.*, vol. 36, no. 19, pp. 6260–6268, Nov. 2008, doi: 10.1093/nar/gkn511.

[62] J. C. Morales *et al.*, 'XRN2 Links Transcription Termination to DNA Damage and Replication Stress', *PLOS Genet.*, vol. 12, no. 7, p. e1006107, Jul. 2016, doi: 10.1371/journal.pgen.1006107.

[63] A. Cristini, M. Groh, M. S. Kristiansen, and N. Gromak, 'RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage', *Cell Rep.*, vol. 23, no. 6, pp. 1891–1905, 2018, doi: 10.1016/j.celrep.2018.04.025.

[64] T. García-Muse and A. Aguilera, 'Transcription-replication conflicts: how they occur and how they are resolved.', *Nat. Rev. Mol. Cell Biol.*, vol. 17, no. 9, pp. 553–563, Sep. 2016, doi: 10.1038/nrm.2016.88.

[65] H. Gaillard, T. García-Muse, and A. Aguilera, 'Replication stress and cancer', *Nat. Rev. Cancer*, vol. 15, no. 5, pp. 276–289, 2015, doi: 10.1038/nrc3916.

[66] S. Kemiha, J. Poli, Y.-L. Lin, A. Lengronne, and P. Pasero, 'Toxic R-loops: Cause or consequence of replication stress?', *DNA Repair*, vol. 107, p. 103199, 2021, doi: https://doi.org/10.1016/j.dnarep.2021.103199.

[67] N. Saini and D. A. Gordenin, 'Hypermutation in single-stranded DNA', *DNA Repair*, vol. 91–92, p. 102868, 2020, doi: 10.1016/j.dnarep.2020.102868.

[68] D. Roy, K. Yu, and M. R. Lieber, 'Mechanism of R-Loop Formation at Immunoglobulin Class Switch Sequences', *Mol. Cell. Biol.*, vol. 28, no. 1, pp. 50–60, Jan. 2008, doi: 10.1128/MCB.01251-07.

[69] U. Basu *et al.*, 'The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates', *Cell*, vol. 144, no. 3, pp. 353–363, Feb. 2011, doi: 10.1016/j.cell.2011.01.001.

[70] S. S. Wallace, 'Base excision repair: a critical player in many games', *DNA Repair*, vol. 19, pp. 14–26, Jul. 2014, doi: 10.1016/j.dnarep.2014.03.030.

[71] X. A. Su and C. H. Freudenreich, 'Cytosine deamination and base excision repair cause R-loop–induced CAG repeat fragility and instability in Saccharomyces cerevisiae', *Proc. Natl. Acad. Sci.*, vol. 114, no. 40, pp. E8392–E8401, Oct. 2017, doi: 10.1073/pnas.1711283114.

[72] Y. Teng *et al.*, 'ROS-induced R loops trigger a transcription-coupled but BRCA1/2-independent homologous recombination pathway through CSB', *Nat. Commun.*, vol. 9, no. 1, p. 4115, Oct. 2018, doi: 10.1038/s41467-018-06586-3.

[73] K. W. Caldecott, 'Single-strand break repair and genetic disease', *Nat. Rev. Genet.*, vol. 9, no. 8, Art. no. 8, Aug. 2008, doi: 10.1038/nrg2380.

[74] R. Anindya, 'Single-stranded DNA damage: Protecting the single-stranded DNA from chemical attack', *DNA Repair*, vol. 87, p. 102804, Mar. 2020, doi: 10.1016/j.dnarep.2020.102804.

[75] H. Merrikh, C. Machón, W. H. Grainger, A. D. Grossman, and P. Soultanas, 'Co-directional replication-transcription conflicts lead to replication restart', *Nature*, vol. 470, no. 7335, pp. 554–557, Feb. 2011, doi: 10.1038/nature09758.

[76] A. Azvolinsky, P. G. Giresi, J. D. Lieb, and V. A. Zakian, 'Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae', *Mol. Cell*, vol. 34, no. 6, pp. 722–734, Jun. 2009, doi: 10.1016/j.molcel.2009.05.022.

[77] L. Costantino and D. Koshland, 'The Yin and Yang of R-loop biology', *Curr. Opin. Cell Biol.*, vol. 34, pp. 39–45, 2015, doi: https://doi.org/10.1016/j.ceb.2015.04.008.

[78] K. Skourti-Stathaki and N. J. Proudfoot, 'A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression', *Genes Dev.*, 2014, doi: 10.1101/gad.242990.114.

[79] C. L and K. D, 'Genome-wide Map of R-Loop-Induced Damage Reveals How a Subset of R-Loops Contributes to Genomic Instability', *Mol. Cell*, vol. 71, no. 4, pp. 487-497.e3, Aug. 2018, doi: 10.1016/J.MOLCEL.2018.06.037.

[80] H. R. Gibbons, G. Shaginurova, L. C. Kim, N. Chapman, C. F. Spurlock, and T. M. Aune, 'Divergent lncRNA GATA3-AS1 Regulates GATA3 Transcription in T-Helper 2 Cells', *Front. Immunol.*, vol. 9, p. 2512, Oct. 2018, doi: 10.3389/fimmu.2018.02512.

[81] F. Prado and A. Aguilera, 'Impairment of replication fork progression mediates RNA polII transcription-associated recombination', *EMBO J.*, vol. 24, no. 6, pp. 1267–1276, Mar. 2005, doi: 10.1038/sj.emboj.7600602.

[82] M. K. Zeman and K. A. Cimprich, 'Causes and Consequences of Replication Stress', vol. 16, no. 1, pp. 2–9, 2015, doi: 10.1038/ncb2897.Causes.

[83] W. Gan *et al.*, 'R-loop-mediated genomic instability is caused by impairment of replication fork progression', *Genes Dev.*, vol. 25, no. 19, pp. 2041–2056, Jan. 2011, doi: 10.1101/gad.17010011.

[84] S. Hamperl, M. J. Bocek, J. C. Saldivar, T. Swigut, and K. A. Cimprich, 'Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses', *Cell*, vol. 170, no. 4, pp. 774-786.e19, Aug. 2017, doi: 10.1016/j.cell.2017.07.043.

[85] H. Merrikh, 'Spatial and Temporal Control of Evolution through Replication-Transcription Conflicts', *Trends Microbiol.*, vol. 25, no. 7, pp. 515–521, Jul. 2017, doi: 10.1016/j.tim.2017.01.008.

[86] R. Bermejo, M. S. Lai, and M. Foiani, 'Preventing Replication Stress to Maintain Genome Stability : Resolving Conflicts between Replication and Transcription', *Mol. Cell*, vol. 45, pp. 710–718, 2012, doi: 10.1016/j.molcel.2012.03.001.

[87] C. Rinaldi, P. Pizzul, M. P. Longhese, and D. Bonetti, 'Sensing R-Loop-Associated DNA Damage to Safeguard Genome Stability', *Front. Cell Dev. Biol.*, vol. 8, 2021, doi: 10.3389/fcell.2020.618157.

[88] P. Huertas and A. Aguilera, 'Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination', *Mol. Cell*, vol. 12, no. 3, pp. 711–721, Sep. 2003, doi: 10.1016/j.molcel.2003.08.010.

[89] P. Kotsantis *et al.*, 'Increased global transcription activity as a mechanism of replication stress in cancer', *Nat. Commun.*, vol. 7, p. 13087, Oct. 2016, doi: 10.1038/ncomms13087.

[90] A. N. Blackford and S. P. Jackson, 'ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response', *Mol. Cell*, vol. 66, no. 6, pp. 801–817, Jun. 2017, doi: 10.1016/j.molcel.2017.05.015.

[91] R. M. Williams, L. A. Yates, and X. Zhang, 'Structures and regulations of ATM and ATR, master kinases in genome integrity', *Curr. Opin. Struct. Biol.*, vol. 61, pp. 98–105, Apr. 2020, doi: 10.1016/j.sbi.2019.12.010.

[92] A. Maréchal and L. Zou, 'DNA damage sensing by the ATM and ATR kinases', *Cold Spring Harb. Perspect. Biol.*, vol. 5, no. 9, p. a012716, Sep. 2013, doi: 10.1101/cshperspect.a012716.

[93] K. A. Cimprich and D. Cortez, 'ATR: an essential regulator of genome integrity', *Nat. Rev. Mol. Cell Biol.*, vol. 9, no. 8, pp. 616–627, Aug. 2008, doi: 10.1038/nrm2450.

[94] J. R. Brickner, J. L. Garzon, and K. A. Cimprich, 'Walking a tightrope: The complex balancing act of R-loops in genome stability', *Mol. Cell*, vol. 82, no. 12, pp. 2267–2297, 2022, doi: https://doi.org/10.1016/j.molcel.2022.04.014.

[95] S. Barroso, E. Herrera-Moyano, S. Muñoz, M. García-Rubio, B. Gómez-González, and A. Aguilera, 'The DNA damage response acts as a safeguard against harmful DNA-RNA hybrids of different origins', *EMBO Rep.*, vol. 20, no. 9, p. e47250, Sep. 2019, doi: 10.15252/embr.201847250.

[96] S. Liu *et al.*, 'ATR Autophosphorylation as a Molecular Switch for Checkpoint Activation', *Mol. Cell*, vol. 43, no. 2, pp. 192–202, Jul. 2011, doi: 10.1016/j.molcel.2011.06.019.

[97] D. Hodroj *et al.*, 'An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism', *EMBO J.*, vol. 36, no. 9, pp. 1182–1198, May 2017, doi: https://doi.org/10.15252/embj.201695131.

[98] Ö. Yüce and S. C. West, 'Senataxin, Defective in the Neurodegenerative Disorder Ataxia with Oculomotor Apraxia 2, Lies at the Interface of Transcription and the DNA Damage Response', *Mol. Cell. Biol.*, vol. 33, no. 2, pp. 406–417, Jan. 2013, doi: 10.1128/MCB.01195-12.

[99] E. Petermann, L. Lan, and L. Zou, 'Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids', *Nat. Rev. Mol. Cell Biol.*, vol. 23, no. 8, pp. 521–540, 2022, doi: 10.1038/s41580-022-00474-x.

[100] M. P. Crossley, M. Bocek, and K. A. Cimprich, 'R-Loops as Cellular Regulators and Genomic Threats.', *Mol. Cell*, vol. 73, no. 3, pp. 398–411, Feb. 2019, doi: 10.1016/j.molcel.2019.01.024.

[101] E. Petermann, L. Lan, and L. Zou, 'Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids', *Nat. Rev. Mol. Cell Biol.*, vol. 23, no. 8, pp. 521–540, 2022, doi: 10.1038/s41580-022-00474-x.

[102] D. A. Matos, J.-M. Zhang, J. Ouyang, H. D. Nguyen, M.-M. Genois, and L. Zou, 'ATR protects the genome against R loops through a MUS81-triggered feedback loop', *Mol. Cell*, vol. 77, no. 3, pp. 514–527, 2020.

[103] L. Kabeche, H. D. Nguyen, R. Buisson, and L. Zou, 'A mitosis-specific and R loop–driven ATR pathway promotes faithful chromosome segregation', *Science*, vol. 359, no. 6371, pp. 108–114, Jan. 2018, doi: 10.1126/science.aan6490.

[104] P. A. Jeggo, L. H. Pearl, and A. M. Carr, 'DNA repair, genome stability and cancer: a historical perspective', *Nat. Rev. Cancer*, vol. 16, no. 1, pp. 35–42, 2016, doi: 10.1038/nrc.2015.4.

[105] Y. Yao and W. Dai, ‘Genomic Instability and Cancer.’, *J. Carcinog. Mutagen.*, vol. 5, 2014, doi: 10.4172/2157-2518.1000165.

[106] I. Kovalchuk, ‘Chapter 25 - Human diseases associated with genome instability☆☆Original article in the 1st edition was prepared by Bruno César Feltes, Joice de Faria Poloni, Kendi Nishino Miyamoto, and Diego Bonatto.’, in *Translational Epigenetics*, vol. 26, I. Kovalchuk and O. B. T.-G. S. (Second E. Kovalchuk, Eds., Boston: Academic Press, 2021, pp. 479–493. doi: https://doi.org/10.1016/B978-0-323-85679-9.00025-8.

[107] J. P. Wells, J. White, and P. C. Stirling, ‘R Loops and Their Composite Cancer Connections’, *Trends in Cancer*, vol. 5, no. 10. 2019. doi: 10.1016/j.trecan.2019.08.006.

[108] E. C. Friedberg, ‘DNA damage and repair’, *Nature*, vol. 421, no. 6921, pp. 436–440, 2003, doi: 10.1038/nature01408.

[109] S. Cohen *et al.*, ‘Senataxin resolves RNA: DNA hybrids forming at DNA double-strand breaks to prevent translocations’, *Nat. Commun.*, vol. 9, no. 1, p. 533, 2018.

[110] S. K. Sotiriou *et al.*, ‘Mammalian RAD52 Functions in Break-Induced Replication Repair of Collapsed DNA Replication Forks’, *Mol. Cell*, vol. 64, no. 6, pp. 1127–1134, Dec. 2016, doi: 10.1016/j.molcel.2016.10.038.

[111] J. Kramara, B. Osia, and A. Malkova, ‘Break-Induced Replication: The Where, The Why, and The How’, *Trends Genet.*, vol. 34, no. 7, pp. 518–531, 2018, doi: https://doi.org/10.1016/j.tig.2018.04.002.

[112] M. J. Rossi, S. F. DiDomenico, M. Patel, and A. V. Mazin, ‘RAD52: Paradigm of Synthetic Lethality and New Developments’, *Front. Genet.*, vol. 12, 2021, Accessed: Dec. 19, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fgene.2021.780293

[113] J. Sollier, C. T. Stork, M. L. García-Rubio, R. D. Paulsen, A. Aguilera, and K. A. Cimprich, ‘Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability’, *Mol. Cell*, vol. 56, no. 6, pp. 777–785, 2014, doi: https://doi.org/10.1016/j.molcel.2014.10.020.

[114] M. K. K. Shivji, X. Renaudin, Ç. H. Williams, and A. R. Venkitaraman, ‘BRCA2 Regulates Transcription Elongation by RNA Polymerase II to Prevent R-Loop Accumulation’, *Cell Rep.*, vol. 22, no. 4, pp. 1031–1039, 2018, doi: https://doi.org/10.1016/j.celrep.2017.12.086.

[115] T. Yasuhara *et al.*, ‘Human Rad52 Promotes XPG-Mediated R-loop Processing to Initiate Transcription-Associated Homologous Recombination Repair’, *Cell*, vol. 175, no. 2, pp. 558-570.e11, Oct. 2018, doi: 10.1016/j.cell.2018.08.056.

[116] F. d’Adda di Fagagna, ‘Site-specific DICER and DROSHA RNA products control the DNA-damage response.’, *Cancer Res.*, vol. 73, no. 8_Supplement, p. 1124, 2013.

[117] F. Storici and A. E. Tichon, ‘RNA takes over control of DNA break repair’, *Nat. Cell Biol.*, vol. 19, no. 12, pp. 1382–1384, 2017, doi: 10.1038/ncb3645.

[118] G. D’Alessandro *et al.*, ‘BRCA2 controls DNA: RNA hybrid level at DSBs by mediating RNase H2 recruitment’, *Nat. Commun.*, vol. 9, no. 1, p. 5376, 2018.

[119] S. Sharma *et al.*, 'MRE11-RAD50-NBS1 Complex Is Sufficient to Promote Transcription by RNA Polymerase II at Double-Strand Breaks by Melting DNA Ends', *Cell Rep.*, vol. 34, no. 1, Jan. 2021, doi: 10.1016/j.celrep.2020.108565.

[120] F. Pessina *et al.*, 'Functional transcription promoters at DNA double-strand breaks mediate RNA-driven phase separation of damage-response factors', *Nat. Cell Biol.*, vol. 21, no. 10, Art. no. 10, Oct. 2019, doi: 10.1038/s41556-019-0392-4.

[121] F. Michelini *et al.*, 'Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks', *Nat. Cell Biol.*, vol. 19, no. 12, pp. 1400–1411, 2017.

[122] M. Gao *et al.*, 'Ago2 facilitates Rad51 recruitment and DNA double-strand break repair by homologous recombination', *Cell Res.*, vol. 24, no. 5, pp. 532–541, 2014.

[123] Y. Zhao *et al.*, 'Telomere extension occurs at most chromosome ends and is uncoupled from fill-in in human cancer cells', *Cell*, vol. 138, no. 3, pp. 463–475, Aug. 2009, doi: 10.1016/j.cell.2009.05.026.

[124] L. Hayflick, 'Mortality and immortality at the cellular level. A review', *Biochem. Biokhimiia*, vol. 62, no. 11, pp. 1180–1190, Nov. 1997.

[125] E. Gilson and V. Géli, 'How telomeres are replicated', *Nat. Rev. Mol. Cell Biol.*, vol. 8, no. 10, pp. 825–838, Oct. 2007, doi: 10.1038/nrm2259.

[126] M. A. Dunham, A. A. Neumann, C. L. Fasching, and R. R. Reddel, 'Telomere maintenance by recombination in human cells', *Nat. Genet.*, vol. 26, no. 4, pp. 447–450, Dec. 2000, doi: 10.1038/82586.

[127] T. M. Bryan, A. Englezou, L. Dalla-Pozza, M. A. Dunham, and R. R. Reddel, 'Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines', *Nat. Med.*, 1997, doi: 10.1038/nm1197-1271.

[128] M. J. McEachern and J. E. Haber, 'Break-induced replication and recombinational telomere elongation in yeast', *Annu. Rev. Biochem.*, vol. 75, pp. 111–135, 2006, doi: 10.1146/annurev.biochem.74.082803.133234.

[129] A. Sommer and N. J. Royle, 'ALT: A Multi-Faceted Phenomenon', *Genes*, vol. 11, no. 2, p. 133, Jan. 2020, doi: 10.3390/genes11020133.

[130] K. Rippe and B. Luke, 'TERRA and the state of the telomere', *Nat. Struct. Mol. Biol.*, vol. 22, no. 11, pp. 853–858, 2015, doi: 10.1038/nsmb.3078.

[131] S. Schoeftner and B. Ma, 'Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II .', *Nat. Cell Biol.*, vol. 10, no. 2, pp. 228–36, 2008, doi: 10.1038/ncb1685.

[132] C. M. Azzalin, P. Reichenbach, L. Khoriauli, E. Giulotto, and J. Lingner, 'Telomeric repeat–containing RNA and RNA surveillance factors at mammalian chromosome ends', *science*, vol. 318, no. 5851, pp. 798–801, 2007.

[133] B. Balk *et al.*, 'Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence', *Nat. Struct. Mol. Biol.*, vol. 20, no. 10, pp. 1199–1206, 2013, doi: 10.1038/nsmb.2662.

[134] T. Kent, D. Gracias, S. Shepherd, and D. Clynes, 'Alternative Lengthening of Telomeres in Pediatric Cancer: Mechanisms to Therapies', *Front. Oncol.*, vol. 9, 2020, Accessed: Aug. 18, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fonc.2019.01518

[135] A. P. Sobinoff and H. A. Pickett, 'Mechanisms that drive telomere maintenance and recombination in human cancers', *Curr. Opin. Genet. Dev.*, vol. 60, pp. 25–30, Feb. 2020, doi: 10.1016/j.gde.2020.02.006.

[136] A. P. Sobinoff and H. A. Pickett, 'Alternative Lengthening of Telomeres: DNA Repair Pathways Converge', *Trends Genet.*, vol. 33, no. 12, pp. 921–932, Dec. 2017, doi: 10.1016/J.TIG.2017.09.003.

[137] R. P. Mackay, Q. Xu, and P. M. Weinberger, 'R-Loop Physiology and Pathology: A Brief Review', *DNA Cell Biol.*, vol. 39, no. 11, pp. 1914–1925, Oct. 2020, doi: 10.1089/dna.2020.5906.

[138] S. M. Cerritelli and R. J. Crouch, 'Ribonuclease H: The enzymes in eukaryotes', *FEBS J.*, vol. 276, no. 6, pp. 1494–1505, 2009, doi: 10.1111/j.1742-4658.2009.06908.x.

[139] N. M. Shaban, S. Harvey, F. W. Perrino, and T. Hollis, 'The structure of the mammalian RNase H2 complex provides insight into RNA·DNA hybrid processing to prevent immune dysfunction', *J. Biol. Chem.*, vol. 285, no. 6, pp. 3617–3624, 2010, doi: 10.1074/jbc.M109.059048.

[140] S. M. Cerritelli, E. G. Frolova, C. Feng, A. Grinberg, P. E. Love, and R. J. Crouch, 'Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice', *Mol. Cell*, vol. 11, no. 3, pp. 807–815, Mar. 2003, doi: 10.1016/S1097-2765(03)00088-1.

[141] D. A. Cornelio, H. N. C. Sedam, J. A. Ferrarezi, N. M. V. Sampaio, and J. L. Argueso, 'Both R-loop removal and ribonucleotide excision repair activities of RNase H2 contribute substantially to chromosome stability', *DNA Repair*, vol. 52, pp. 110–114, 2017, doi: 10.1016/j.dnarep.2017.02.012.

[142] S. Liu *et al.*, 'Distinct roles for DNA-PK, ATM and ATR in RPA phosphorylation and checkpoint activation in response to replication stress', *Nucleic Acids Res.*, vol. 40, no. 21, pp. 10780–10794, Nov. 2012, doi: 10.1093/nar/gks849.

[143] H. D. Nguyen, T. Yadav, S. Giri, B. Saez, T. A. Graubert, and L. Zou, 'Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1', *Mol. Cell*, vol. 65, no. 5, pp. 832-847.e4, Mar. 2017, doi: 10.1016/j.molcel.2017.01.029.

[144] E. Jankowsky, 'RNA helicases at work: Binding and rearranging', *Trends Biochem. Sci.*, vol. 36, no. 1, pp. 19–29, 2011, doi: 10.1016/j.tibs.2010.07.008.

[145] H. E. Mischo *et al.*, 'Yeast Sen1 helicase protects the genome from transcription-associated instability', *Mol. Cell*, vol. 41, no. 1, pp. 21–32, Jan. 2011, doi: 10.1016/j.molcel.2010.12.007.

[146] O. J. Becherel *et al.*, 'Senataxin Plays an Essential Role with DNA Damage Response Proteins in Meiotic Recombination and Gene Silencing', *PLoS Genet.*, vol. 9, no. 4, p. e1003435, Apr. 2013, doi: 10.1371/journal.pgen.1003435.

[147] E. Y. C. Chang *et al.*, 'RECQ-like helicases Sgs1 and BLM regulate R-loop- associated genome instabil', *J. Cell Biol.*, vol. 216, no. 12, pp. 3991–4005, Dec. 2017, doi: 10.1083/jcb.201703168.

[148] V. Marabitti *et al.*, 'R-Loop-Associated Genomic Instability and Implication of WRN and WRNIP1', *Int. J. Mol. Sci.*, vol. 23, no. 3, 2022, doi: 10.3390/ijms23031547.

[149] C. Ribeiro de Almeida *et al.*, 'RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination', *Mol. Cell*, vol. 70, no. 4, pp. 650-662.e8, May 2018, doi: 10.1016/j.molcel.2018.04.001.

[150] S. Yang, L. Winstone, S. Mondal, and Y. Wu, 'Helicases in R-loop Formation and Resolution', *J. Biol. Chem.*, vol. 299, no. 11, Nov. 2023, doi: 10.1016/j.jbc.2023.105307.

[151] C. F. Bourgeois, F. Mortreux, and D. Auboeuf, 'The multiple functions of RNA helicases as drivers and regulators of gene expression', *Nat. Rev. Mol. Cell Biol.*, vol. 17, no. 7, Art. no. 7, Jul. 2016, doi: 10.1038/nrm.2016.50.

[152] A. Promonet *et al.*, 'Topoisomerase 1 prevents replication stress at R-loop-enriched transcription termination sites', *Nat. Commun.*, vol. 11, no. 1, pp. 1–12, 2020, doi: 10.1038/s41467-020-17858-2.

[153] S. Tuduri *et al.*, 'Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription', *Nat. Cell Biol.*, vol. 11, no. 11, pp. 1315–1324, 2009, doi: 10.1038/ncb1984.

[154] J. E. Cho and S. Jinks-Robertson, 'Topoisomerase I and genome stability: The good and the bad', in *Methods in Molecular Biology*, vol. 1703, Humana Press Inc., 2018, pp. 21–45. doi: 10.1007/978-1-4939-7459-7_2.

[155] S. G. Manzo *et al.*, 'DNA Topoisomerase I differentially modulates R-loops across the human genome', *Genome Biol.*, vol. 19, no. 1, pp. 1–18, Jul. 2018, doi: 10.1186/s13059-018-1478-1.

[156] O. Sordet *et al.*, 'Ataxia telangiectasia mutated activation by transcription- and topoisomerase I-induced DNA double-strand breaks', *EMBO Rep.*, vol. 10, no. 8, pp. 887–893, Aug. 2009, doi: 10.1038/embor.2009.97.

[157] L. Baranello, F. Kouzine, and D. Levens, 'DNA Topoisomerases', *Transcription*, vol. 4, no. 5, pp. 232–237, Sep. 2013, doi: 10.4161/trns.26598.

[158] C. Hraiky, M. A. Raymond, and M. Drolet, 'RNase H overproduction corrects a defect at the level of transcription elongation during rRNA synthesis in the absence of DNA topoisomerase I in Escherichia coli', *J. Biol. Chem.*, vol. 275, no. 15, pp. 11257–11263, Apr. 2000, doi: 10.1074/jbc.275.15.11257.

[159] J. J. Champoux, 'DNA Topoisomerases: Structure, Function, and Mechanism', *Annu. Rev. Biochem.*, vol. 70, no. 1, pp. 369–413, Jun. 2001, doi: 10.1146/annurev.biochem.70.1.369.

[160] A. El Hage, S. L. French, A. L. Beyer, and D. Tollervey, 'Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synEl Hage, A., French, S. L., Beyer, A. L., & Tollervey, D. (2010). Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA', *Genes Dev.*, vol. 24, no. 14, pp. 1546–1558, Jul. 2010, doi: 10.1101/gad.573310.

[161] S. K. Calderwood, 'A critical role for topoisomerase IIb and DNA double strand breaks in transcription', *Transcription*, vol. 7, no. 3, pp. 75–83, May 2016, doi: 10.1080/21541264.2016.1181142.

[162] A. Canela *et al.*, 'Genome Organization Drives Chromosome Fragility', *Cell*, vol. 170, no. 3, pp. 507-521.e18, Jul. 2017, doi: 10.1016/j.cell.2017.06.034.

[163] S. Salerno *et al.*, 'Recent Advances in the Development of Dual Topoisomerase I and II Inhibitors as Anticancer Drugs', *Curr. Med. Chem.*, vol. 17, no. 35, pp. 4270–4290, Nov. 2010, doi: 10.2174/092986710793361252.

[164] E. L. Baldwin and N. Osheroff, 'Etoposide, topoisomerase II and cancer', *Curr. Med. Chem. - Anti-Cancer Agents*, vol. 5, no. 4, pp. 363–372, Jul. 2005, doi: 10.2174/1568011054222364.

[165] P. Pilati, D. Nitti, and S. Mocellin, 'Cancer Resistance to Type II Topoisomerase Inhibitors', *Curr. Med. Chem.*, vol. 19, no. 23, pp. 3900–3906, Jul. 2012, doi: 10.2174/092986712802002473.

[166] Y. Yang, K. M. McBride, S. Hensley, Y. Lu, F. Chedin, and M. T. Bedford, 'Arginine Methylation Facilitates the Recruitment of TOP3B to Chromatin to Prevent R Loop Accumulation', *Mol. Cell*, vol. 53, no. 3, pp. 484–497, Feb. 2014, doi: 10.1016/j.molcel.2014.01.011.

[167] A. Rodríguez and A. D'Andrea, 'Fanconi anemia pathway', *Curr. Biol.*, vol. 27, no. 18, pp. R986–R988, 2017.

[168] A. Gueiderikh, F. Maczkowiak-Chartois, and F. Rosselli, 'A new frontier in Fanconi anemia: From DNA repair to ribosome biogenesis.', *Blood Rev.*, vol. 52, p. 100904, Mar. 2022, doi: 10.1016/j.blre.2021.100904.

[169] M. Nepal, R. Che, J. Zhang, C. Ma, and P. Fei, 'Fanconi Anemia Signaling and Cancer', *Trends Cancer*, vol. 3, no. 12, pp. 840–856, Dec. 2017, doi: 10.1016/j.trecan.2017.10.005.

[170] R. Roy, J. Chun, and S. N. Powell, 'BRCA1 and BRCA2: Different roles in a common pathway of genome protection', *Nat. Rev. Cancer*, vol. 12, no. 1, pp. 68–78, Jan. 2012, doi: 10.1038/nrc3181.

[171] K. D. Yu and Z. M. Shao, 'Initiation, evolution, phenotype and outcome of BRCA1 and BRCA2 mutation-associated breast cancer', *Nat. Rev. Cancer*, vol. 12, no. 5, pp. 372–373, May 2012, doi: 10.1038/nrc3181-c1.

[172] Y. Okamoto, J. Hejna, and M. Takata, 'Regulation of R-loops and genome instability in Fanconi anemia', *J. Biochem. (Tokyo)*, vol. 165, no. 6, pp. 465–470, Jun. 2019, doi: 10.1093/jb/mvz019.

[173] E. Hatchi *et al.*, 'BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair', *Mol. Cell*, vol. 57, no. 4, pp. 636–647, Feb. 2015, doi: 10.1016/j.molcel.2015.01.011.

[174] V. Bhatia, S. I. Barroso, M. L. García-Rubio, E. Tumini, E. Herrera-Moyano, and A. Aguilera, 'BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2', *Nature*, vol. 511, no. 7509, pp. 362–365, 2014, doi: 10.1038/nature13374.

[175] V. Bhatia, S. I. Barroso, M. L. García-Rubio, E. Tumini, E. Herrera-Moyano, and A. Aguilera, 'BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2', *Nature*, vol. 511, no. 7509, pp. 362–365, 2014, doi: 10.1038/nature13374.

[176] Z. Liang *et al.*, 'Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination', *Cell Rep.*, vol. 26, no. 3, pp. 564-572.e5, 2019, doi: 10.1016/j.celrep.2018.12.084.

[177] M. L. García-Rubio *et al.*, 'The Fanconi Anemia Pathway Protects Genome Integrity from R-loops', *PLoS Genet.*, vol. 11, no. 11, pp. 1–17, 2015, doi: 10.1371/journal.pgen.1005674.

[178] Z. Liang *et al.*, 'Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination', *Cell Rep.*, vol. 26, no. 3, pp. 564-572.e5, 2019, doi: 10.1016/j.celrep.2018.12.084.

[179] B. Silva *et al.*, 'FANCM limits ALT activity by restricting telomeric replication stress induced by deregulated BLM and R-loops', *Nat. Commun.*, vol. 10, no. 1, Dec. 2019, doi: 10.1038/s41467-019-10179-z.

[180] M. Feretzaki, M. Pospisilova, R. Valador Fernandes, T. Lunardi, L. Krejci, and J. Lingner, 'RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops', *Nature*, vol. 587, no. 7833, pp. 303–308, Nov. 2020, doi: 10.1038/s41586-020-2815-6.

[181] M. Thomas, C. Dubacq, E. Rabut, B. S. Lopez, and J. Guirouilh-Barbat, 'Noncanonical Roles of RAD51', *Cells*, vol. 12, no. 8, 2023, doi: 10.3390/cells12081169.

[182] S. Chávez *et al.*, 'A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae', *EMBO J.*, vol. 19, no. 21, pp. 5824–5834, Nov. 2000, doi: 10.1093/emboj/19.21.5824.

[183] K. Stäßer *et al.*, 'TREX is a conserved complex coupling transcription with messenger RNA export', *Nature*, vol. 417, no. 6886, pp. 304–308, Apr. 2002, doi: 10.1038/nature746.

[184] J. Katahira and Y. Yoneda, 'Roles of the TREX complex in nuclear export of mRNA', *RNA Biol.*, vol. 6, no. 2, pp. 149–152, 2009, doi: 10.4161/rna.6.2.8046.

[185] M. S. Domínguez-Sánchez, S. Barroso, B. Gómez-González, R. Luna, and A. Aguilera, 'Genome instability and transcription elongation impairment in human cells depleted of THO/TREX', *PLoS Genet.*, vol. 7, no. 12, p. 1002386, Dec. 2011, doi: 10.1371/journal.pgen.1002386.

[186] A. G. Rondón, S. Jimeno, M. García-Rubio, and A. Aguilera, 'Molecular evidence that the eukaryotic THO/TREX complex is required for efficient transcription elongation', *J. Biol. Chem.*, vol. 278, no. 40, pp. 39037–39043, Oct. 2003, doi: 10.1074/jbc.M305718200.

[187] A. Aguilera, 'mRNA processing and genomic instability', *Nat. Struct. Mol. Biol.*, vol. 12, no. 9, pp. 737–738, Sep. 2005, doi: 10.1038/nsmb0905-737.

[188] K. Dunn and J. D. Griffith, 'The presence of RNA in a double helix inhibits its interaction with histone protein', *Nucleic Acids Res.*, vol. 8, no. 3, pp. 555–566, 1980.

[189] C. Davó-Martínez *et al.*, 'Different SWI/SNF complexes coordinately promote R-loop- and RAD52-dependent transcription-coupled homologous recombination', *Nucleic Acids Res.*, vol. 51, no. 17, pp. 9055–9074, Sep. 2023, doi: 10.1093/nar/gkad609.

[190] E. Herrera-Moyano, X. Mergui, M. L. García-Rubio, S. Barroso, and A. Aguilera, 'The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-replication conflicts', *Genes Dev.*, vol. 28, no. 7, pp. 735–748, Apr. 2014, doi: 10.1101/gad.234070.113.

[191] L. Prendergast *et al.*, 'Resolution of R-loops by INO80 promotes DNA replication and maintains cancer cell proliferation and viability', *Nat. Commun.*, vol. 11, no. 1, Art. no. 1, Sep. 2020, doi: 10.1038/s41467-020-18306-x.

[192] A. Bayona-Feliu, S. Barroso, S. Muñoz, and A. Aguilera, 'The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription-replication conflicts', *Nat. Genet.*, vol. 53, no. 7, pp. 1050–1063, Jul. 2021, doi: 10.1038/s41588-021-00867-2.

[193] I. Salas-Armenteros, C. Pérez-Calero, A. Bayona-Feliu, E. Tumini, R. Luna, and A. Aguilera, 'Human <scp>THO</scp> –Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability', *EMBO J.*, vol. 36, no. 23, pp. 3532–3547, Dec. 2017, doi: 10.15252/embj.201797208.

[194] I. Mahmud and D. Liao, 'DAXX in cancer: phenomena, processes, mechanisms and regulation', *Nucleic Acids Res.*, vol. 47, no. 15, pp. 7734–7752, Sep. 2019, doi: 10.1093/nar/gkz634.

[195] P. W. Lewis, S. J. Elsaesser, K.-M. Noh, S. C. Stadler, and C. D. Allis, 'Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres', *Proc. Natl. Acad. Sci.*, 2010, doi: 10.1073/pnas.1008850107.

[196] P. Salomoni, 'The PML-Interacting Protein DAXX: Histone Loading Gets into the Picture', *Front. Oncol.*, vol. 3, p. 152, 2013, doi: 10.3389/fonc.2013.00152.

[197] A. D. Goldberg *et al.*, 'Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions', *Cell*, vol. 140, no. 5, pp. 678–691, 2010, doi: https://doi.org/10.1016/j.cell.2010.01.003.

[198] P. Drané, K. Ouararhni, A. Depaux, M. Shuaib, and A. Hamiche, 'The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3', *Genes Dev.*, vol. 24, no. 12, pp. 1253–1265, Jun. 2010, doi: 10.1101/gad.566910.

[199] V. M. Morozov, S. Giovinazzi, and A. M. Ishov, 'CENP-B protects centromere chromatin integrity by facilitating histone deposition via the H3.3-specific chaperone Daxx', *Epigenetics Chromatin*, vol. 10, no. 1, pp. 1–18, 2017, doi: 10.1186/s13072-017-0164-y.

[200] D. T. Nguyen *et al.*, 'The chromatin remodelling factor <scp>ATRX</scp> suppresses R-loops in transcribed telomeric repeats', *EMBO Rep.*, vol. 18, no. 6, pp. 914–928, Jun. 2017, doi: 10.15252/embr.201643078.

[201] Q. He *et al.*, 'The Daxx/Atrx complex protects tandem repetitive elements during DNA hypomethylation by promoting H3K9 trimethylation', *Cell Stem Cell*, vol. 17, no. 3, pp. 273–286, 2015.

[202] F. Li *et al.*, '<scp>ATRX</scp> loss induces telomere dysfunction and necessitates induction of alternative lengthening of telomeres during human cell immortalization', *EMBO J.*, vol. 38, no. 19, p. e96659, Oct. 2019, doi: 10.15252/embj.201796659.

[203] A. Dhayalan *et al.*, 'The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9', *Hum. Mol. Genet.*, vol. 20, no. 11, pp. 2195–2203, Jun. 2011, doi: 10.1093/hmg/ddr107.

[204] S. Iwase *et al.*, 'ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome', *Nat. Struct. Mol. Biol.*, vol. 18, no. 7, pp. 769–776, Jun. 2011, doi: 10.1038/nsmb.2062.

[205] Y.-C. Teng *et al.*, 'ATRX promotes heterochromatin formation to protect cells from G-quadruplex DNA-mediated stress', *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–14, Jun. 2021, doi: 10.1038/s41467-021-24206-5.

[206] R.-X. Tsai *et al.*, 'TERRA regulates DNA G-quadruplex formation and ATRX recruitment to chromatin', *Nucleic Acids Res.*, vol. 50, no. 21, pp. 12217–12234, Nov. 2022, doi: 10.1093/nar/gkac1114.

[207] Y. Hu *et al.*, 'Switch telomerase to ALT mechanism by inducing telomeric DNA damages and dysfunction of ATRX and DAXX', *Sci. Rep.*, vol. 6, no. 1, pp. 1–10, Aug. 2016, doi: 10.1038/srep32280.

[208] M. A. Dyer, Z. A. Qadeer, D. Valle-Garcia, and E. Bernstein, 'ATRX and DAXX: Mechanisms and mutations', *Cold Spring Harb. Perspect. Med.*, vol. 7, no. 3, p. a026567, Mar. 2017, doi: 10.1101/cshperspect.a026567.

[209] C. M. Heaphy *et al.*, 'Altered telomeres in tumors with ATRX and DAXX mutations', *Science*, vol. 333, no. 6041, p. 425, Jul. 2011, doi: 10.1126/science.1207313.

[210] E. Mason-Osann *et al.*, 'Identification of a novel gene fusion in ALT positive osteosarcoma', *Oncotarget*, vol. 9, no. 67, pp. 32868–32880, Aug. 2018, doi: 10.18632/oncotarget.26029.

[211] A. de Nonneville and R. R. Reddel, 'Alternative lengthening of telomeres is not synonymous with mutations in ATRX/DAXX', *Nat. Commun.*, vol. 12, no. 1, p. 1552, Mar. 2021, doi: 10.1038/s41467-021-21794-0.

[212] C. A. Lovejoy *et al.*, 'Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway', *PLoS Genet.*, vol. 8, no. 7, p. e1002772, 2012, doi: 10.1371/journal.pgen.1002772.

[213] D. Clynes *et al.*, 'Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX', *Nat. Commun.*, 2015, doi: 10.1038/ncomms8538.

[214] K. E. Yost *et al.*, 'Rapid and reversible suppression of ALT by DAXX in osteosarcoma cells', *Sci. Rep.*, vol. 9, no. 1, pp. 1–11, Dec. 2019, doi: 10.1038/s41598-019-41058-8.

[215] J. G. Patton, E. B. Porro, J. Galceran, P. Tempst, and B. Nadal-Ginard, 'Cloning and characterization of PSF, a novel pre-mRNA splicing factor', *Genes Dev.*, vol. 7, no. 3, pp. 393–406, Mar. 1993, doi: 10.1101/gad.7.3.393.

[216] G. J. Knott, C. S. Bond, and A. H. Fox, 'The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold', *Nucleic Acids Res.*, vol. 44, no. 9, pp. 3989–4004, May 2016, doi: 10.1093/nar/gkw271.

[217] M. Salton, Y. Lerenthal, S.-Y. Wang, D. J. Chen, and Y. Shiloh, 'Involvement of Matrin 3 and SFPQ/NONO in the DNA damage response', *Cell Cycle*, vol. 9, no. 8, pp. 1568–1576, Apr. 2010, doi: 10.4161/cc.9.8.11298.

[218] Y. Morozumi, Y. Takizawa, M. Takaku, and H. Kurumizaka, 'Human PSF binds to RAD51 and modulates its homologous-pairing and strand-exchange activities', *Nucleic Acids Res.*, vol. 37, no. 13, pp. 4296–4307, Jul. 2009, doi: 10.1093/nar/gkp298.

[219] C. Rajesh, D. K. Baker, A. J. Pierce, and D. L. Pittman, 'The splicing-factor related protein SFPQ/PSF interacts with RAD51D and is necessary for homology-directed repair and sister chromatid cohesion', *Nucleic Acids Res.*, vol. 39, no. 1, pp. 132–145, Jan. 2011, doi: 10.1093/nar/gkq738.

[220] B. Schell, P. Legrand, and S. Fribourg, 'Crystal structure of SFPQ-NONO heterodimer', *Biochimie*, vol. 198, pp. 1–7, Jul. 2022, doi: 10.1016/j.biochi.2022.02.011.

[221] M. Lee *et al.*, 'The structure of human SFPQ reveals a coiled-coil mediated polymer essential for functional aggregation in gene regulation', *Nucleic Acids Res.*, vol. 43, no. 7, pp. 3826–3840, Apr. 2015, doi: 10.1093/nar/gkv156.

[222] A. H. Fox and A. I. Lamond, 'Paraspeckles', *Cold Spring Harb. Perspect. Biol.*, vol. 2, no. 7, p. a000687, Jul. 2010, doi: 10.1101/cshperspect.a000687.

[223] A. Emili *et al.*, 'Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD', *RNA N. Y. N*, vol. 8, no. 9, pp. 1102–1111, Sep. 2002, doi: 10.1017/s1355838202025037.

[224] D. K. Rhee *et al.*, 'SFPQ, a multifunctional nuclear protein, regulates the transcription of PDE3A', *Biosci. Rep.*, vol. 37, no. 4, p. BSR20170975, Aug. 2017, doi: 10.1042/BSR20170975.

[225] A. H. Fox, S. Nakagawa, T. Hirose, and C. S. Bond, 'Paraspeckles: Where Long Noncoding RNA Meets Phase Separation', *Trends Biochem. Sci.*, vol. 43, no. 2, pp. 124–135, Feb. 2018, doi: 10.1016/j.tibs.2017.12.001.

[226] L. Jaafar, Z. Li, S. Li, and W. S. Dynan, 'SFPQ•NONO and XLF function separately and together to promote DNA double-strand break repair via canonical nonhomologous end joining', *Nucleic Acids Res.*, vol. 45, no. 4, pp. 1848–1859, Feb. 2017, doi: 10.1093/nar/gkw1209.

[227] E. Petti *et al.*, 'SFPQ and NONO suppress RNA:DNA-hybrid-related telomere instability', *Nat. Commun.*, vol. 10, no. 1, 2019, doi: 10.1038/s41467-019-08863-1.

[228] B. A. Ozdilek, V. F. Thompson, N. S. Ahmed, C. I. White, R. T. Batey, and J. C. Schwartz, 'Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding', *Nucleic Acids Res.*, vol. 45, no. 13, pp. 7984–7996, Jul. 2017, doi: 10.1093/nar/gkx460.

[229] P. Thandapani, T. R. O'Connor, T. L. Bailey, and S. Richard, 'Defining the RGG/RG motif', *Mol. Cell*, vol. 50, no. 5, pp. 613–623, Jun. 2013, doi: 10.1016/j.molcel.2013.05.021.

[230] E. Rosonina *et al.*, 'Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo', *Mol. Cell. Biol.*, vol. 25, no. 15, pp. 6734–6746, Aug. 2005, doi: 10.1128/MCB.25.15.6734-6746.2005.

[231] M. P. Williamson, 'The structure and function of proline-rich regions in proteins', *Biochem. J.*, vol. 297 (Pt 2), no. Pt 2, pp. 249–260, Jan. 1994, doi: 10.1042/bj2970249.

[232] P. Baumann and S. C. West, 'Role of the human RAD51 protein in homologous recombination and double-stranded-break repair', *Trends Biochem. Sci.*, vol. 23, no. 7, pp. 247–251, Jul. 1998, doi: 10.1016/s0968-0004(98)01232-8.

[233] K. Ha, Y. Takeda, and W. S. Dynan, 'Sequences in PSF/SFPQ mediate radioresistance and recruitment of PSF/SFPQ-containing complexes to DNA damage sites in human cells', *DNA Repair*, vol. 10, no. 3, pp. 252–259, Mar. 2011, doi: 10.1016/j.dnarep.2010.11.009.

[234] B. R. Levone *et al.*, 'FUS-dependent liquid-liquid phase separation is important for DNA repair initiation', *J. Cell Biol.*, vol. 220, no. 5, p. e202008030, May 2021, doi: 10.1083/jcb.202008030.

[235] C. Maris, C. Dominguez, and F. H.-T. Allain, 'The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression', *FEBS J.*, vol. 272, no. 9, pp. 2118–2131, May 2005, doi: 10.1111/j.1742-4658.2005.04653.x.

[236] C. A. Yarosh, J. R. Iacona, C. S. Lutz, and K. W. Lynch, 'PSF: nuclear busy-body or nuclear facilitator?', *Wiley Interdiscip. Rev. RNA*, vol. 6, no. 4, pp. 351–367, 2015, doi: 10.1002/wrna.1280.

[237] C. A. Yarosh *et al.*, 'TRAP150 interacts with the RNA-binding domain of PSF and antagonizes splicing of numerous PSF-target genes in T cells', *Nucleic Acids Res.*, vol. 43, no. 18, pp. 9006–9016, Oct. 2015, doi: 10.1093/nar/gkv816.

[238] Y. Wang and L.-L. Chen, 'Organization and function of paraspeckles', *Essays Biochem.*, vol. 64, no. 6, pp. 875–882, Dec. 2020, doi: 10.1042/EBC20200010.

[239] K. Imamura *et al.*, 'Long Noncoding RNA NEAT1-Dependent SFPQ Relocation from Promoter Region to Paraspeckle Mediates IL8 Expression upon Immune Stimuli', *Mol. Cell*, vol. 54, no. 6, p. 1055, Jun. 2014, doi: 10.1016/j.molcel.2014.06.013.

[240] T. W. Hewage, S. Caria, and M. Lee, 'A new crystal structure and small-angle X-ray scattering analysis of the homodimer of human SFPQ', *Acta Crystallogr. Sect. F Struct. Biol. Commun.*, vol. 75, no. Pt 6, pp. 439–449, Jun. 2019, doi: 10.1107/S2053230X19006599.

[241] J. Huang, G. P. Casas Garcia, M. A. Perugini, A. H. Fox, C. S. Bond, and M. Lee, 'Crystal structure of a SFPQ/PSPC1 heterodimer provides insights into preferential heterodimerization of human DBHS family proteins', *J. Biol. Chem.*, vol. 293, no. 17, pp. 6593–6602, Apr. 2018, doi: 10.1074/jbc.RA117.001451.

[242] D. M. Passon *et al.*, 'Structure of the heterodimer of human NONO and paraspeckle protein component 1 and analysis of its role in subnuclear body formation', *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 13, pp. 4846–4850, Mar. 2012, doi: 10.1073/pnas.1120792109.

[243] F. Heyd and K. W. Lynch, 'Phosphorylation-dependent regulation of PSF by GSK3 controls CD45 alternative splicing', *Mol. Cell*, vol. 40, no. 1, pp. 126–137, Oct. 2010, doi: 10.1016/j.molcel.2010.09.013.

[244] H. Lee, Z. Zhang, and H. M. Krause, 'Long Noncoding RNAs and Repetitive Elements: Junk or Intimate Evolutionary Partners?', *Trends Genet.*, vol. 35, no. 12, pp. 892–902, Dec. 2019, doi: 10.1016/j.tig.2019.09.006.

[245] T. Liehr, *Benign and pathological chromosomal imbalances: microscopic and submicroscopic copy number variations (CNVs) in genetics and counseling*. Academic Press, 2013.

[246] T. Liehr, 'Repetitive Elements in Humans', *Int. J. Mol. Sci.*, vol. 22, no. 4, p. 2072, Feb. 2021, doi: 10.3390/ijms22042072.

[247] G.-F. Richard, A. Kerrest, and B. Dujon, 'Comparative genomics and molecular dynamics of DNA repeats in eukaryotes', *Microbiol. Mol. Biol. Rev. MMBR*, vol. 72, no. 4, pp. 686–727, Dec. 2008, doi: 10.1128/MMBR.00011-08.

[248] N. Bannert and R. Kurth, 'Retroelements and the human genome: new perspectives on an old relation', *Proc. Natl. Acad. Sci.*, vol. 101, no. suppl_2, pp. 14572–14579, 2004.

[249] J. N. Wells and C. Feschotte, 'A Field Guide to Eukaryotic Transposable Elements.', *Annu. Rev. Genet.*, vol. 54, pp. 539–561, Nov. 2020, doi: 10.1146/annurev-genet-040620-022145.

[250] F. Brändle, B. Frühbauer, and M. Jagannathan, 'Principles and functions of pericentromeric satellite DNA clustering into chromocenters', *Semin. Cell Dev. Biol.*, vol. 128, pp. 26–39, 2022, doi: https://doi.org/10.1016/j.semcdb.2022.02.005.

[251] L. P. Jenner, V. Peska, J. Fulnečková, and E. Sýkorová, 'Telomeres and Their Neighbors', *Genes*, vol. 13, no. 9, 2022, doi: 10.3390/genes13091663.

[252] A. El Hage, S. Webb, A. Kerr, and D. Tollervey, 'Genome-Wide Distribution of RNA-DNA Hybrids Identifies RNase H Targets in tRNA Genes, Retrotransposons and Mitochondria', *PLoS Genet.*, vol. 10, no. 10, 2014, doi: 10.1371/journal.pgen.1004716.

[253] Y. A. Chan *et al.*, 'Genome-Wide Profiling of Yeast DNA:RNA Hybrid Prone Sites with DRIP-Chip', *PLoS Genet.*, 2014, doi: 10.1371/journal.pgen.1004288.

[254] A. Munden, M. L. Benton, J. A. Capra, and J. Nordman, 'R-loop mapping and characterization during Drosophila embryogenesis reveals developmental plasticity in R-loop signatures', *bioRxiv*, p. 2021.10.29.465954, Jan. 2021, doi: 10.1101/2021.10.29.465954.

[255] J. S. Waye and H. F. Willard, 'Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome', *Mol. Cell. Biol.*, vol. 6, no. 9, pp. 3156–3165, Sep. 1986, doi: 10.1128/mcb.6.9.3156-3165.1986.

[256] H. F. Willard, 'Chromosome-specific organization of human alpha satellite DNA.', *Am. J. Hum. Genet.*, vol. 37, no. 3, pp. 524–532, May 1985.

[257] B. Vissel and K. H. Choo, 'Human alpha satellite DNA--consensus sequence and conserved regions.', *Nucleic Acids Res.*, vol. 15, no. 16, pp. 6751–6752, Aug. 1987.

[258] G. M. Greig and H. F. Willard, 'β satellite DNA: Characterization and localization of two subfamilies from the distal and proximal short arms of the human acrocentric chromosomes', *Genomics*, vol. 12, no. 3, pp. 573–580, Mar. 1992, doi: 10.1016/0888-7543(92)90450-7.

[259] M. F. Cardone *et al.*, 'Evolution of beta satellite DNA sequences: evidence for duplication-mediated repeat amplification and spreading', *Mol. Biol. Evol.*, vol. 21, no. 9, pp. 1792–1799, Sep. 2004, doi: 10.1093/molbev/msh190.

[260] C. C. Lin, R. Sasi, Y. S. Fan, and D. Court, 'Isolation and identification of a novel tandemly repeated DNA sequence in the centromeric region of human chromosome 8', *Chromosoma*, vol. 102, no. 5, pp. 333–339, May 1993, doi: 10.1007/BF00661276.

[261] C. Lee, R. Critcher, J. G. Zhang, W. Mills, and C. J. Farr, 'Distribution of gamma satellite DNA on the human X and Y chromosomes suggests that it is not required for mitotic centromere function', *Chromosoma*, vol. 109, no. 6, pp. 381–389, Sep. 2000, doi: 10.1007/s004120000095.

[262] M. G. Schueler *et al.*, 'Progressive proximal expansion of the primate X chromosome centromere', *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 30, pp. 10563–10568, Jul. 2005, doi: 10.1073/pnas.0503346102.

[263] B. Vissel, A. Nagy, and K. H. Choo, 'A satellite III sequence shared by human chromosomes 13, 14, and 21 that is contiguous with alpha satellite DNA', *Cytogenet. Cell Genet.*, vol. 61, no. 2, pp. 81–86, 1992, doi: 10.1159/000133374.

[264] J. Prosser, M. Frommer, C. Paul, and P. C. Vincent, 'Sequence relationships of three human satellite DNAs.', *J. Mol. Biol.*, vol. 187, no. 2, pp. 145–155, Jan. 1986, doi: 10.1016/0022-2836(86)90224-x.

[265] I. Tagarro, A. M. Fernández-Peralta, and J. J. González-Aguilera, 'Chromosomal localization of human satellites 2 and 3 by a FISH method using oligonucleotides as probes', *Hum. Genet.*, vol. 93, no. 4, pp. 383–388, Apr. 1994, doi: 10.1007/BF00201662.

[266] N. Altemose, K. H. Miga, M. Maggioni, and H. F. Willard, 'Genomic characterization of large heterochromatic gaps in the human genome assembly', *PLoS Comput. Biol.*, vol. 10, no. 5, p. e1003628, May 2014, doi: 10.1371/journal.pcbi.1003628.

[267] K. Smurova and P. De Wulf, 'Centromere and Pericentromere Transcription: Roles and Regulation … in Sickness and in Health', *Front. Genet.*, vol. 9, p. 674, Dec. 2018, doi: 10.3389/fgene.2018.00674.

[268] V. M. Bolanos-Garcia, 'Aurora kinases', *Int. J. Biochem. Cell Biol.*, vol. 37, no. 8, pp. 1572–1577, 2005.

[269] G. Wang, Q. Jiang, and C. Zhang, 'The role of mitotic kinases in coupling the centrosome cycle with the assembly of the mitotic spindle', *J. Cell Sci.*, vol. 127, no. 19, pp. 4111–4122, Oct. 2014, doi: 10.1242/jcs.151753.

[270] M. Carmena, S. Ruchaud, and W. C. Earnshaw, 'Making the Auroras glow: regulation of Aurora A and B kinase function by interacting proteins', *Curr. Opin. Cell Biol.*, vol. 21, no. 6, pp. 796–805, 2009, doi: https://doi.org/10.1016/j.ceb.2009.09.008.

[271] A. P. Damodaran, L. Vaufrey, O. Gavard, and C. Prigent, 'Aurora A Kinase Is a Priority Pharmaceutical Target for the Treatment of Cancers', *Trends Pharmacol. Sci.*, vol. 38, no. 8, pp. 687–700, 2017, doi: https://doi.org/10.1016/j.tips.2017.05.003.

[272] D. Quénet and Y. Dalal, 'A long non-coding RNA is required for targeting centromeric protein A to the human centromere', *eLife*, vol. 3, pp. e03254–e03254, Aug. 2014, doi: 10.7554/eLife.03254.

[273] F. Lyn Chan and L. H. Wong, 'Transcription in the maintenance of centromere chromatin identity', *Nucleic Acids Res.*, vol. 40, no. 22, pp. 11178–11188, 2012, doi: 10.1093/nar/gks921.

[274] T. Ideue, Y. Cho, K. Nishimura, and T. Tani, 'Involvement of satellite I noncoding RNA in regulation of chromosome segregation', *Genes Cells*, vol. 19, no. 6, pp. 528–538, Jun. 2014, doi: https://doi.org/10.1111/gtc.12149.

[275] S. M. McNulty, L. L. Sullivan, and B. A. Sullivan, 'Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are Complexed with CENP-A and CENP-C', *Dev. Cell*, vol. 42, no. 3, pp. 226-240.e6, 2017, doi: 10.1016/j.devcel.2017.07.001.

[276] E. C. Moran *et al.*, 'Mitotic R-loops direct Aurora B kinase to maintain centromeric cohesion', *bioRxiv*, p. 2021.01.14.426738, Jan. 2021, doi: 10.1101/2021.01.14.426738.

[277] T. De Lange *et al.*, 'Structure and Variability of Human Chromosome Ends', 1990.

[278] M. L. DuBois, Z. W. Haimberger, M. W. McIntosh, and D. E. Gottschling, 'A quantitative assay for telomere protection in Saccharomyces cerevisiae', *Genetics*, vol. 161, no. 3, pp. 995–1013, 2002.

[279] J. Flint *et al.*, 'Sequence comparison of human and yeast telomeres identifies structurally distinct subtelomeric domains', *Hum. Mol. Genet.*, vol. 6, no. 8, pp. 1305–1314, 1997.

[280] T. De Lange, 'Shelterin: The protein complex that shapes and safeguards human telomeres', *Genes Dev.*, vol. 19, no. 18, pp. 2100–2110, 2005, doi: 10.1101/gad.1346005.

[281] K. Cleal, K. Norris, and D. Baird, 'Telomere length dynamics and the evolution of cancer genome architecture', *Int. J. Mol. Sci.*, vol. 19, no. 2, p. 482, 2018.

[282] S. Schoeftner and M. A. Blasco, 'Chromatin regulation and non-coding RNAs at mammalian telomeres', *Semin. Cell Dev. Biol.*, vol. 21, no. 2, pp. 186–193, 2010, doi: 10.1016/j.semcdb.2009.09.015.

[283] M. Achrem, I. Szućko, and A. Kalinka, 'The epigenetic regulation of centromeres and telomeres in plants and animals', *Comp. Cytogenet.*, vol. 14, no. 2, p. 265, 2020.

[284] A. Diman and A. Decottignies, 'Genomic origin and nuclear localization of TERRA telomeric repeat-containing RNA: from Darkness to Dawn', *FEBS J.*, vol. 285, no. 8, pp. 1389–1398, 2018.

[285] E. Cusanelli and P. Chartrand, 'Telomeric repeat-containing RNA TERRA : a noncoding RNA connecting telomere biology to genome integrity', *Front. Genet.*, vol. 6, pp. 1–9, 2015, doi: 10.3389/fgene.2015.00143.

[286] N. Bettin, C. Oss Pegorar, and E. Cusanelli, 'The emerging roles of TERRA in telomere maintenance and genome stability', *Cells*, vol. 8, no. 3, p. 246, 2019.

[287] A. Chebly *et al.*, 'Telomeric repeat-containing RNA (TERRA): A review of the literature and first assessment in cutaneous T-cell lymphomas', *Genes*, vol. 13, no. 3, p. 539, 2022.

[288] R. Arora, Y. Lee, H. Wischnewski, C. M. Brun, T. Schwarz, and C. M. Azzalin, 'RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells', *Nat. Commun.*, vol. 5, pp. 1–11, 2014, doi: 10.1038/ncomms6220.

[289] V. Pfeiffer, J. Crittin, L. Grolimund, and J. Lingner, 'The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening', *EMBO J.*, vol. 32, no. 21, pp. 2861–2871, 2013.

[290] R. V. Fernandes, M. Feretzaki, and J. Lingner, 'The makings of TERRA R-loops at chromosome ends', *Cell Cycle*, vol. 20, no. 18, pp. 1745–1759, 2021.

[291] Y. Gong and Y. Liu, 'R-Loops at Chromosome Ends: From Formation, Regulation, and Cellular Consequence', *Cancers*, vol. 15, no. 7, p. 2178, 2023.

[292] D. Hanahan and R. A. Weinberg, 'Chapter 2 : Hallmarks of Cancer : An Organizing Principle for Cancer Medicine', 2015.

[293] G. Chin and C. S. Lansing, 'The biological sciences collaboratory', Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2004.

[294] M. De Vitis, F. Berardinelli, and A. Sgura, 'Molecular Sciences Telomere Length Maintenance in Cancer: At the Crossroad between Telomerase and Alternative Lengthening of Telomeres (ALT)', 2018, doi: 10.3390/ijms19020606.

[295] R. J. O'Sullivan and J. Karlseder, 'Telomeres: Protecting chromosomes against genome instability', *Nat. Rev. Mol. Cell Biol.*, vol. 11, no. 3, pp. 171–181, 2010, doi: 10.1038/nrm2848.

[296] J. P. Amorim, G. Santos, J. Vinagre, and P. Soares, 'The Role of ATRX in the Alternative Lengthening of Telomeres (ALT) Phenotype', *Genes*, vol. 7, no. 9, 2016, doi: 10.3390/genes7090066.

[297] T. Yadav, J.-M. Zhang, J. Ouyang, W. Leung, A. Simoneau, and L. Zou, 'TERRA and RAD51AP1 promote alternative lengthening of telomeres through an R- to D-loop switch', *Mol. Cell*, vol. 82, no. 21, pp. 3985-4000.e4, 2022, doi: https://doi.org/10.1016/j.molcel.2022.09.026.

[298] N. Kaminski *et al.*, 'RAD51AP1 regulates ALT-HDR through chromatin-directed homeostasis of TERRA', *Mol. Cell*, vol. 82, no. 21, pp. 4001-4017.e7, 2022, doi: https://doi.org/10.1016/j.molcel.2022.09.025.

[299] C. R. L. Huang, K. H. Burns, and J. D. Boeke, 'Active transposition in genomes', *Annu. Rev. Genet.*, vol. 46, pp. 651–675, 2012, doi: 10.1146/annurev-genet-110711-155616.

[300] C. Feschotte and E. J. Pritham, 'DNA transposons and the evolution of eukaryotic genomes', *Annu. Rev. Genet.*, vol. 41, pp. 331–368, 2007, doi: 10.1146/annurev.genet.40.110405.090448.

[301] T. Wicker *et al.*, 'A unified classification system for eukaryotic transposable elements', *Nat. Rev. Genet.*, vol. 8, no. 12, pp. 973–982, 2007, doi: 10.1038/nrg2165.

[302] M. Ravi, S. Ramanathan, and K. Krishna, 'Factors, mechanisms and implications of chromatin condensation and chromosomal structural maintenance through the cell cycle', *J. Cell. Physiol.*, vol. 235, no. 2, pp. 758–775, Feb. 2020, doi: https://doi.org/10.1002/jcp.29038.

[303] K. Saito, 'The epigenetic regulation of transposable elements by PIWI-interacting RNAs in *Drosophila*', *Genes Genet. Syst.*, vol. 88, no. 1, pp. 9–17, 2013, doi: 10.1266/ggs.88.9.

[304] D. Holoch and D. Moazed, 'RNA-mediated epigenetic regulation of gene expression.', *Nat. Rev. Genet.*, vol. 16, no. 2, pp. 71–84, Feb. 2015, doi: 10.1038/nrg3863.

[305] H. H. Kazazian Jr and J. V. Moran, 'Mobile DNA in Health and Disease', *N. Engl. J. Med.*, vol. 377, no. 4, pp. 361–370, Jul. 2017, doi: 10.1056/NEJMra1510092.

[306] H. H. Kazazian, C. Wong, H. Youssoufian, A. F. Scott, D. G. Phillips, and S. E. Antonarakis, 'Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man', *Nature*, vol. 332, no. 6160, pp. 164–166, 1988, doi: 10.1038/332164a0.

[307] K. H. Burns, 'Transposable elements in cancer', *Nat. Rev. Cancer*, vol. 17, no. 7, pp. 415–424, 2017, doi: 10.1038/nrc.2017.35.

[308] F. Yu, N. Zingler, G. Schumann, and W. H. Strätling, 'Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription', *Nucleic Acids Res.*, vol. 29, no. 21, pp. 4493–4501, Nov. 2001, doi: 10.1093/nar/29.21.4493.

[309] Y. J. Crow and N. Manel, 'Aicardi–Goutières syndrome and the type I interferonopathies', *Nat. Rev. Immunol.*, vol. 15, no. 7, pp. 429–440, 2015, doi: 10.1038/nri3850.

[310] A. Saleh, A. Macia, and A. R. Muotri, 'Transposable elements, inflammation, and neurological disease', *Front. Neurol.*, vol. 10, p. 894, 2019.

[311] V. Fort, G. Khelifi, and S. M. I. Hussein, 'Long non-coding RNAs and transposable elements: A functional relationship', *Biochim. Biophys. Acta BBA - Mol. Cell Res.*, vol. 1868, no. 1, p. 118837, 2021, doi: https://doi.org/10.1016/j.bbamcr.2020.118837.

[312] Q. Al-Hadid and Y. Yang, 'R-loop: an emerging regulator of chromatin dynamics', *Acta Biochim. Biophys. Sin.*, vol. 48, no. 7, pp. 623–631, Jul. 2016, doi: 10.1093/abbs/gmw052.

[313] S. J. Elsässer, K.-M. Noh, N. Diaz, C. D. Allis, and L. A. Banaszynski, 'Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells.', *Nature*, vol. 522, no. 7555, pp. 240–244, Jun. 2015, doi: 10.1038/nature14345.

[314] M. Alkailani *et al.*, 'A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer.', *NAR Cancer*, vol. 3, no. 1, p. zcaa040, Mar. 2021, doi: 10.1093/narcan/zcaa040.

[315] Y. Nie, A. F. Wilson, T. DeFalco, A. R. Meetei, S. H. Namekawa, and Q. Pang, 'FANCD2 is required for the repression of germline transposable elements.', *Reprod. Camb. Engl.*, vol. 159, no. 6, pp. 659–668, Jun. 2020, doi: 10.1530/REP-19-0436.

[316] D. Filipponi, J. Muller, A. Emelyanov, and D. V. Bulavin, 'Wip1 controls global heterochromatin silencing via ATM/BRCA1-dependent DNA methylation.', *Cancer Cell*, vol. 24, no. 4, pp. 528–541, Oct. 2013, doi: 10.1016/j.ccr.2013.08.022.

[317] P. Mita *et al.*, 'BRCA1 and S phase DNA repair pathways restrict LINE-1 retrotransposition in human cells.', *Nat. Struct. Mol. Biol.*, vol. 27, no. 2, pp. 179–191, Feb. 2020, doi: 10.1038/s41594-020-0374-z.

[318] K. Bartsch *et al.*, 'Absence of RNase H2 triggers generation of immunogenic micronuclei removed by autophagy', *Hum. Mol. Genet.*, vol. 26, no. 20, pp. 3960–3972, Oct. 2017, doi: 10.1093/hmg/ddx283.

[319] M. Benitez-Guijarro *et al.*, 'RNase H2, mutated in Aicardi-Goutières syndrome, promotes LINE-1 retrotransposition.', *EMBO J.*, vol. 37, no. 15, Aug. 2018, doi: 10.15252/embj.201798506.

[320] V. Pokatayev *et al.*, 'RNase H2 catalytic core Aicardi-Goutières syndrome-related mutant invokes cGAS-STING innate immune-sensing pathway in mice.', *J. Exp. Med.*, vol. 213, no. 3, pp. 329–336, Mar. 2016, doi: 10.1084/jem.20151464.

[321] P. Li *et al.*, 'Aicardi-Goutières syndrome protein TREX1 suppresses L1 and maintains genome integrity through exonuclease-independent ORF1p depletion.', *Nucleic Acids Res.*, vol. 45, no. 8, pp. 4619–4631, May 2017, doi: 10.1093/nar/gkx178.

[322] E. Orecchini *et al.*, 'ADAR1 restricts LINE-1 retrotransposition', *Nucleic Acids Res.*, vol. 45, no. 1, pp. 155–168, Jan. 2017, doi: 10.1093/nar/gkw834.

[323] C. A. Thomas *et al.*, 'Modeling of TREX1-Dependent Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation.', *Cell Stem Cell*, vol. 21, no. 3, pp. 319-331.e8, Sep. 2017, doi: 10.1016/j.stem.2017.07.009.

[324] Y. W. Lim, L. A. Sanz, X. Xu, S. R. Hartono, and F. Chédin, 'Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi–Goutières syndrome', *eLife*, vol. 4, p. e08007, 2015, doi: 10.7554/eLife.08007.

[325] K. Zhao *et al.*, 'Modulation of LINE-1 and Alu/SVA Retrotransposition by Aicardi-Goutières Syndrome-Related SAMHD1', *Cell Rep.*, vol. 4, no. 6, pp. 1108–1115, 2013, doi: https://doi.org/10.1016/j.celrep.2013.08.019.

[326] H. E. Volkman and D. B. Stetson, 'The enemy within: endogenous retroelements and autoimmune disease', *Nat. Immunol.*, vol. 15, no. 5, pp. 415–422, 2014, doi: 10.1038/ni.2872.

[327] G. W. Schmid-Schönbein, 'Analysis of inflammation', *Annu. Rev. Biomed. Eng.*, vol. 8, pp. 93–131, 2006, doi: 10.1146/annurev.bioeng.8.061505.095708.

[328] S. E. Turvey and D. H. Broide, 'Innate immunity', *J. Allergy Clin. Immunol.*, vol. 125, no. 2, Supplement 2, pp. S24–S32, Feb. 2010, doi: 10.1016/j.jaci.2009.07.016.

[329] F. A. Bonilla and H. C. Oettgen, 'Adaptive immunity', *J. Allergy Clin. Immunol.*, vol. 125, no. 2, Supplement 2, pp. S33–S40, Feb. 2010, doi: 10.1016/j.jaci.2009.09.017.

[330] F. R. Greten and S. I. Grivennikov, 'Inflammation and Cancer: Triggers, Mechanisms, and Consequences', *Immunity*, vol. 51, no. 1, pp. 27–41, Jul. 2019, doi: 10.1016/j.immuni.2019.06.025.

[331] R. Medzhitov, 'Origin and physiological roles of inflammation', *Nature*, vol. 454, no. 7203, pp. 428–435, Jul. 2008, doi: 10.1038/nature07201.

[332] N. Singh, D. Baby, J. P. Rajguru, P. B. Patil, S. S. Thakkannavar, and V. B. Pujari, 'Inflammation and Cancer', *Ann. Afr. Med.*, vol. 18, no. 3, pp. 121–126, 2019, doi: 10.4103/aam.aam_56_18.

[333] L. M. Coussens and Z. Werb, 'Inflammation and cancer', *Nature*, vol. 420, no. 6917, pp. 860–867, Dec. 2002, doi: 10.1038/nature01322.

[334] S. I. Grivennikov, F. R. Greten, and M. Karin, 'Immunity, Inflammation, and Cancer', *Cell*, vol. 140, no. 6, pp. 883–899, Mar. 2010, doi: 10.1016/j.cell.2010.01.025.

[335] K. Esfahani, L. Roudaia, N. Buhlaiga, S. V. Del Rincon, N. Papneja, and W. H. Miller, 'A review of cancer immunotherapy: from the past, to the present, to the future', *Curr. Oncol. Tor. Ont*, vol. 27, no. Suppl 2, pp. S87–S97, Apr. 2020, doi: 10.3747/co.27.5223.

[336] Y. Gan *et al.*, 'The cGAS/STING Pathway: A Novel Target for Cancer Therapy', *Front. Immunol.*, vol. 12, p. 795401, 2021, doi: 10.3389/fimmu.2021.795401.

[337] M. Motwani, S. Pesiridis, and K. A. Fitzgerald, 'DNA sensing by the cGAS–STING pathway in health and disease', *Nat. Rev. Genet.*, vol. 20, no. 11, Art. no. 11, Nov. 2019, doi: 10.1038/s41576-019-0151-1.

[338] L. Sun, J. Wu, F. Du, X. Chen, and Z. J. Chen, 'Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway', *Science*, vol. 339, no. 6121, pp. 786–791, Feb. 2013, doi: 10.1126/science.1232458.

[339] X. Zhang *et al.*, 'The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop', *Cell Rep.*, vol. 6, no. 3, pp. 421–430, Feb. 2014, doi: 10.1016/j.celrep.2014.01.003.

[340] P. Gao *et al.*, 'Cyclic [G(2',5')pA(3',5')p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase', *Cell*, vol. 153, no. 5, pp. 1094–1107, May 2013, doi: 10.1016/j.cell.2013.04.046.

[341] H. Ishikawa and G. N. Barber, 'STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling', *Nature*, vol. 455, no. 7213, pp. 674–678, Oct. 2008, doi: 10.1038/nature07317.

[342] K. Mukai *et al.*, 'Activation of STING requires palmitoylation at the Golgi', *Nat. Commun.*, vol. 7, p. 11932, Jun. 2016, doi: 10.1038/ncomms11932.

[343] S. Srikanth *et al.*, 'The Ca2+ sensor STIM1 regulates the type I interferon response by retaining the signaling adaptor STING at the endoplasmic reticulum', *Nat. Immunol.*, vol. 20, no. 2, pp. 152–162, Feb. 2019, doi: 10.1038/s41590-018-0287-8.

[344] T. Agalioti, S. Lomvardas, B. Parekh, J. Yie, T. Maniatis, and D. Thanos, 'Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter', *Cell*, vol. 103, no. 4, pp. 667–678, Nov. 2000, doi: 10.1016/s0092-8674(00)00169-0.

[345] L. R. Watkins, S. F. Maier, and L. E. Goehler, 'Immune activation: the role of pro-inflammatory cytokines in inflammation, illness responses and pathological pain states', *Pain*, vol. 63, no. 3, pp. 289–302, Dec. 1995, doi: 10.1016/0304-3959(95)00186-7.

[346] K. J. Mackenzie *et al.*, 'cGAS surveillance of micronuclei links genome instability to innate immunity', *Nature*, vol. 548, no. 7668, Art. no. 7668, Aug. 2017, doi: 10.1038/nature23449.

[347] M. Fenech *et al.*, 'Molecular mechanisms of micronucleus, nucleoplasmic bridge and nuclear bud formation in mammalian and human cells', *Mutagenesis*, vol. 26, no. 1, pp. 125–132, Jan. 2011, doi: 10.1093/mutage/geq052.

[348] S. S. W. Ho *et al.*, 'The DNA Structure-Specific Endonuclease MUS81 Mediates DNA Sensor STING-Dependent Host Rejection of Prostate Cancer Cells', *Immunity*, vol. 44, no. 5, pp. 1177–1189, May 2016, doi: 10.1016/j.immuni.2016.04.010.

[349] M. P. Crossley *et al.*, 'R-loop-derived cytoplasmic RNA–DNA hybrids activate an immune response', *Nature*, vol. 613, no. 7942, pp. 187–194, 2023, doi: 10.1038/s41586-022-05545-9.

[350] J. Marinello *et al.*, 'Topoisomerase I poison-triggered immune gene activation is markedly reduced in human small-cell lung cancers by impairment of the cGAS/STING pathway', *Br. J. Cancer*, vol. 127, no. 7, pp. 1214–1225, Oct. 2022, doi: 10.1038/s41416-022-01894-4.

[351] I. Bogolyubova and D. Bogolyubov, 'DAXX Is a Crucial Factor for Proper Development of Mammalian Oocytes and Early Embryos', *Int. J. Mol. Sci.*, vol. 22, no. 3, p. 1313, Jan. 2021, doi: 10.3390/ijms22031313.

[352] I. Rosso and F. d'Adda di Fagagna, 'Detection of Telomeric DNA:RNA Hybrids Using TeloDRIP-qPCR', *Int. J. Mol. Sci.*, vol. 21, no. 24, p. 9774, Dec. 2020, doi: 10.3390/ijms21249774.

[353] I. Roeschert *et al.*, 'Combined inhibition of Aurora-A and ATR kinase results in regression of MYCN-amplified neuroblastoma', *Nat. Cancer*, vol. 2, no. 3, pp. 312–326, Mar. 2021, doi: 10.1038/s43018-020-00171-8.

[354] C. W. Templeton and L. A. Laimins, 'p53-dependent R-loop formation and HPV pathogenesis', *Proc. Natl. Acad. Sci. U. S. A.*, vol. 120, no. 35, p. e2305907120, Aug. 2023, doi: 10.1073/pnas.2305907120.

[355] D. F. Allison and G. G. Wang, 'R-loops: formation, function, and relevance to cell stress.', *Cell Stress*, vol. 3, no. 2, pp. 38–46, Jan. 2019, doi: 10.15698/cst2019.02.175.

[356] E. Bochkareva, L. Frappier, A. M. Edwards, and A. Bochkarev, 'The RPA32 subunit of human replication protein A contains a single-stranded DNA-binding domain', *J. Biol. Chem.*, vol. 273, no. 7, pp. 3932–3936, 1998.

[357] P. Haahr *et al.*, 'Activation of the ATR kinase by the RPA-binding protein ETAA1', *Nat. Cell Biol.*, vol. 18, no. 11, pp. 1196–1207, 2016.

[358] P. Awasthi, M. Foiani, and A. Kumar, 'ATM and ATR signaling at a glance.', *J. Cell Sci.*, vol. 128, no. 23, pp. 4255–4262, Dec. 2015, doi: 10.1242/jcs.169730.

[359] E. Laurini, D. Marson, A. Fermeglia, S. Aulic, M. Fermeglia, and S. Pricl, 'Role of Rad51 and DNA repair in cancer: A molecular perspective', *Pharmacol. Ther.*, vol. 208, p. 107492, 2020, doi: https://doi.org/10.1016/j.pharmthera.2020.107492.

[360] I. Chaudhury, A. Sareen, M. Raghunandan, and A. Sobeck, 'FANCD2 regulates BLM complex functions independently of FANCI to promote replication fork recovery', *Nucleic Acids Res.*, vol. 41, no. 13, pp. 6444–6459, 2013.

[361] X. Chen *et al.*, 'The Fanconi anemia proteins FANCD2 and FANCJ interact and regulate each other's chromatin localization.', *J. Biol. Chem.*, vol. 289, no. 37, pp. 25774–25782, Sep. 2014, doi: 10.1074/jbc.M114.552570.

[362] Q. Fan, F. Zhang, B. Barrett, K. Ren, and P. R. Andreassen, 'A role for monoubiquitinated FANCD2 at telomeres in ALT cells', vol. 37, no. 6, pp. 1740–1754, 2009, doi: 10.1093/nar/gkn995.

[363] J. Michl, J. Zimmer, F. M. Buffa, U. McDermott, and M. Tarsounas, 'FANCD2 limits replication stress and genome instability in cells lacking BRCA2.', *Nat. Struct. Mol. Biol.*, vol. 23, no. 8, pp. 755–757, Aug. 2016, doi: 10.1038/nsmb.3252.

[364] Y. W. Lim, D. James, J. Huang, and M. Lee, 'The Emerging Role of the RNA-Binding Protein SFPQ in Neuronal Function and Neurodegeneration', *Int. J. Mol. Sci.*, vol. 21, no. 19, p. 7151, Sep. 2020, doi: 10.3390/ijms21197151.

[365] I. Pellarin *et al.*, 'Splicing factor proline- and glutamine-rich (SFPQ) protein regulates platinum response in ovarian cancer-modulating SRSF2 activity', *Oncogene*, vol. 39, no. 22, Art. no. 22, May 2020, doi: 10.1038/s41388-020-1292-6.

[366] M. Kovac *et al.*, 'Exome sequencing of osteosarcoma reveals mutation signatures reminiscent of BRCA deficiency', *Nat. Commun.*, vol. 6, p. 8940, Dec. 2015, doi: 10.1038/ncomms9940.

[367] G. Miglietta, M. Russo, R. C. Duardo, and G. Capranico, 'G-quadruplex binders as cytostatic modulators of innate immune genes in cancer cells', *Nucleic Acids Res.*, vol. 49, no. 12, pp. 6673–6686, Jul. 2021, doi: 10.1093/nar/gkab500.

[368] S. Yum, M. Li, Y. Fang, and Z. J. Chen, 'TBK1 recruitment to STING activates both IRF3 and NF-κB that mediate immune defense against tumors and viral infections', *Proc. Natl. Acad. Sci. U. S. A.*, vol. 118, no. 14, p. e2100225118, Apr. 2021, doi: 10.1073/pnas.2100225118.

[369] S. Lee *et al.*, 'Lipocalin-2 Is a chemokine inducer in the central nervous system: role of chemokine ligand 10 (CXCL10) in lipocalin-2-induced cell migration', *J. Biol. Chem.*, vol. 286, no. 51, pp. 43855–43870, Dec. 2011, doi: 10.1074/jbc.M111.299248.

[370] X. Liu, Y. Tian, N. Lu, T. Gin, C. H. K. Cheng, and M. T. V. Chan, 'Stat3 inhibition attenuates mechanical allodynia through transcriptional regulation of chemokine expression in spinal astrocytes', *PloS One*, vol. 8, no. 10, p. e75804, 2013, doi: 10.1371/journal.pone.0075804.

[371] E. Makuch, I. Jasyk, A. Kula, T. Lipiński, and J. Siednienko, 'IFNβ-Induced CXCL10 Chemokine Expression Is Regulated by Pellino3 Ligase in Monocytes and Macrophages', *Int. J. Mol. Sci.*, vol. 23, no. 23, p. 14915, Nov. 2022, doi: 10.3390/ijms232314915.

[372] S. M. Haag *et al.*, 'Targeting STING with covalent small-molecule inhibitors', *Nature*, vol. 559, no. 7713, pp. 269–273, Jul. 2018, doi: 10.1038/s41586-018-0287-8.

[373] Y. Pan *et al.*, 'The STING antagonist H-151 ameliorates psoriasis via suppression of STING/NF-κB-mediated inflammation', *Br. J. Pharmacol.*, vol. 178, no. 24, pp. 4907–4922, Dec. 2021, doi: 10.1111/bph.15673.

[374] E. Guida *et al.*, 'Peptide aptamers targeting mutant p53 induce apoptosis in tumor cells', *Cancer Res.*, vol. 68, no. 16, pp. 6550–6558, Aug. 2008, doi: 10.1158/0008-5472.CAN-08-0137.

[375] M. R. Dunn, R. M. Jimenez, and J. C. Chaput, 'Analysis of aptamer discovery and technology', *Nat. Rev. Chem.*, vol. 1, no. 10, Art. no. 10, Oct. 2017, doi: 10.1038/s41570-017-0076.

[376] Y. Zhang, B. S. Lai, and M. Juhas, 'Recent Advances in Aptamer Discovery and Applications', *Molecules*, vol. 24, no. 5, p. 941, Mar. 2019, doi: 10.3390/molecules24050941.

[377] H. P. J. Voon and L. H. Wong, 'New players in heterochromatin silencing: histone variant H3.3 and the ATRX/DAXX chaperone', *Nucleic Acids Res.*, vol. 44, no. 4, pp. 1496–1501, Feb. 2016, doi: 10.1093/nar/gkw012.

[378] A. Tafessu *et al.*, 'H3.3 contributes to chromatin accessibility and transcription factor binding at promoter-proximal regulatory elements in embryonic stem cells', *Genome Biol.*, vol. 24, no. 1, p. 25, Feb. 2023, doi: 10.1186/s13059-023-02867-3.

[379] J. P. Gerber *et al.*, 'Aberrant chromatin landscape following loss of the H3.3 chaperone Daxx in haematopoietic precursors leads to Pu.1-mediated neutrophilia and inflammation', *Nat. Cell Biol.*, vol. 23, no. 12, Art. no. 12, Dec. 2021, doi: 10.1038/s41556-021-00774-y.

[380] M. R. Middleton *et al.*, 'Phase 1 study of the ATR inhibitor berzosertib (formerly M6620, VX-970) combined with gemcitabine ± cisplatin in patients with advanced solid tumours', *Br. J. Cancer*, vol. 125, no. 4, Art. no. 4, Aug. 2021, doi: 10.1038/s41416-021-01405-x.

[381] G. Reyhanoglu and P. Tadi, 'Etoposide', in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2023. Accessed: Oct. 24, 2023. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK557864/

[382] D. Bartolucci, A. Pession, P. Hrelia, and R. Tonelli, 'Precision Anti-Cancer Medicines by Oligonucleotide Therapeutics in Clinical Research Targeting Undruggable Proteins and Non-Coding RNAs', *Pharmaceutics*, vol. 14, no. 7, Jul. 2022, doi: 10.3390/pharmaceutics14071453.

[383] H. Xiong, R. N. Veedu, and S. D. Diermeier, 'Recent Advances in Oligonucleotide Therapeutics in Oncology', *Int. J. Mol. Sci.*, vol. 22, no. 7, p. 3295, Mar. 2021, doi: 10.3390/ijms22073295.

[384] Y. Zhang *et al.*, 'Research progress on non-protein-targeted drugs for cancer therapy', *J. Exp. Clin. Cancer Res.*, vol. 42, no. 1, p. 62, Mar. 2023, doi: 10.1186/s13046-023-02635-y.

[385] L.-Y. Chen and J. D. Chen, 'Daxx silencing sensitizes cells to multiple apoptotic pathways', *Mol. Cell. Biol.*, vol. 23, no. 20, pp. 7108–7121, Oct. 2003, doi: 10.1128/MCB.23.20.7108-7121.2003.

[386] P. Salomoni and A. F. Khelifi, 'Daxx: death or survival protein?', *Trends Cell Biol.*, vol. 16, no. 2, pp. 97–104, Feb. 2006, doi: 10.1016/j.tcb.2005.12.002.

[387] D. Wang *et al.*, 'Acetylation-regulated interaction between p53 and SET reveals a widespread regulatory mode', *Nature*, vol. 538, no. 7623, Art. no. 7623, Oct. 2016, doi: 10.1038/nature19759.

[388] C.-C. Chang, D.-Y. Lin, H.-I. Fang, R.-H. Chen, and H.-M. Shih, 'Daxx Mediates the Small Ubiquitin-like Modifier-dependent Transcriptional Repression of Smad4*', *J. Biol. Chem.*, vol. 280, no. 11, pp. 10164–10173, Mar. 2005, doi: 10.1074/jbc.M409161200.

[389] L. J. Helman and P. Meltzer, 'Mechanisms of sarcoma development', *Nat. Rev. Cancer*, vol. 3, no. 9, Art. no. 9, Sep. 2003, doi: 10.1038/nrc1168.

[390] T. G. Grünewald *et al.*, 'Sarcoma treatment in the era of molecular medicine', *EMBO Mol. Med.*, vol. 12, no. 11, p. e11131, Nov. 2020, doi: 10.15252/emmm.201911131.

[391] J. W. Potter, K. B. Jones, and J. J. Barrott, 'Sarcoma-The standard-bearer in cancer discovery', *Crit. Rev. Oncol. Hematol.*, vol. 126, pp. 1–5, Jun. 2018, doi: 10.1016/j.critrevonc.2018.03.007.

[392] B. A. Lindsey, J. E. Markel, and E. S. Kleinerman, 'Osteosarcoma Overview', *Rheumatol. Ther.*, vol. 4, no. 1, pp. 25–43, Jun. 2017, doi: 10.1007/s40744-016-0050-2.

[393] H. T. Ta, C. R. Dass, P. F. M. Choong, and D. E. Dunstan, 'Osteosarcoma treatment: state of the art', *Cancer Metastasis Rev.*, vol. 28, no. 1–2, pp. 247–263, Jun. 2009, doi: 10.1007/s10555-009-9186-7.

[394] A. M. Czarnecka *et al.*, 'Molecular Biology of Osteosarcoma', *Cancers*, vol. 12, no. 8, p. 2130, Jul. 2020, doi: 10.3390/cancers12082130.

[395] E. Panatta *et al.*, 'Metabolic regulation by p53 prevents R-loop-associated genomic instability', *Cell Rep.*, vol. 41, no. 5, p. 111568, Nov. 2022, doi: 10.1016/j.celrep.2022.111568.

[396] S. F. Jia, T. An, L. Worth, and E. S. Kleinerman, 'Interferon-alpha enhances the sensitivity of human osteosarcoma cells to etoposide', *J. Interferon Cytokine Res. Off. J. Int. Soc. Interferon Cytokine Res.*, vol. 19, no. 6, pp. 617–624, Jun. 1999, doi: 10.1089/107999099313758.

[397] J. Zhao *et al.*, 'Interferon-α suppresses invasion and enhances cisplatin-mediated apoptosis and autophagy in human osteosarcoma cells', *Oncol. Lett.*, vol. 7, no. 3, pp. 827–833, Mar. 2014, doi: 10.3892/ol.2013.1762.

[398] A. Takeuchi *et al.*, 'Loss of Sfpq Causes Long-Gene Transcriptopathy in the Brain', *Cell Rep.*, vol. 23, no. 5, pp. 1326–1341, May 2018, doi: 10.1016/j.celrep.2018.03.141.

[399] A. D. Keefe, S. Pai, and A. Ellington, 'Aptamers as therapeutics', *Nat. Rev. Drug Discov.*, vol. 9, no. 7, Art. no. 7, Jul. 2010, doi: 10.1038/nrd3141.

[400] L. Chen *et al.*, 'Knockdown of ZBTB11 impedes R-loop elimination and increases the sensitivity to cisplatin by inhibiting DDX1 transcription in bladder cancer', *Cell Prolif.*, vol. 55, no. 12, p. e13325, Dec. 2022, doi: 10.1111/cpr.13325.