Transcription-Replication Conflicts and R-loops disrupt the local chromatin landscape in human model cell lines

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES

DER FAKULTÄT FÜR BIOLOGIE

DER LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Marcel Werner

April 2025

Diese Dissertation wurde angefertigt
unter der Leitung von Dr. Stephan Hamperl
am Institut für Epigenetik und Stammzellen
des Helmholtz Zentrum München

Erstgutachter:	Prof. Dr. Maria-Elena Torres-Padilla
Zweitgutachter:	Prof. Dr. Christof Osman
Tag der Abgabe:	02.04.2025
Tag der mündlichen Prüfung	05.08.2025

Eigenständigkeitserklärung

Hiermit versichere ich an Eides statt, dass die vorliegende Dissertation mit dem Titel

“Transcription-Replication Conflicts and R-loops disrupt the local chromatin landscape in human model cell lines”

von mir selbstständig verfasst wurde und dass keine anderen als die angegebenen Quellen und Hilfsmittel benutzt wurden. Die Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinne nach entnommen sind, wurden in jedem Fall unter Angabe der Quellen (einschließlich des World Wide Web und anderer elektronischer Text- und Datensammlungen) kenntlich gemacht. Weiterhin wurden alle Teile der Arbeit, die mit Hilfe von Werkzeugen der künstlichen Intelligenz de novo generiert wurden, durch Fußnote/Anmerkung an den entsprechenden Stellen kenntlich gemacht und die verwendeten Werkzeuge der künstlichen Intelligenz gelistet. Die genutzten Prompts befinden sich im Anhang. Diese Erklärung gilt für alle in der Arbeit enthaltenen Texte, Graphiken, Zeichnungen, Kartenskizzen und bildliche Darstellungen.

München, den 01.04.2025

(Ort / Datum)

MARCEL WERNER

(Vor- und Nachname in Druckbuchstaben)

Marcel Werner

(Unterschrift)

Erklärung

Hiermit erkläre ich,

dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist.

dass ich mich anderweitig einer Doktorprüfung ohne Erfolg **nicht** unterzogen habe.

München, den 01.04.2025	Marcel Werner
Ort, Datum	Unterschrift

Wesentliche Teile dieser Arbeit sind in der folgenden Publikation veröffentlicht und zusammengefasst: Werner et al., 2025

Werner, M., Trauner, M., Schauer, T., Ummethum, H., Márquez-Gómez, E., Lalonde, M., Lee, C. S. K., Tsirkas, I., Sajid, A., Muriello, A. C., Längst, G., & Hamperl, S.
Transcription-replication conflicts drive R-loop-dependent nucleosome eviction and require DOT1L activity for transcription recovery.
Nucleic Acids Research, 53(4), 13–14. https://doi.org/10.1093/NAR/GKAF109, (2025).

Table of Contents

1. Summary

Transcription and DNA replication are fundamental nuclear processes that control the expression of genes and ensure the correct duplication of the genome, respectively. Tight regulation of these molecular machines is of paramount importance as both dysregulated transcription and replication are associated with numerous diseases including cancer. Importantly, both processes require extensive remodeling of the underlying chromatin template to perform their function. Transcription-replication conflicts (TRCs) are nuclear events during which transcription and replication collide on the same DNA template during the S-phase of the cell cycle. Despite their rare occurrence, TRCs are potent inducers of DNA damage, mutations, and genomic instability. Given that transcription and replication extensively remodel chromatin during their progression on the genome, one hypothesis suggests that TRC sites might be particularly vulnerable to chromatin and epigenome alterations that could provoke genome instability. In this work, I engineered an inducible TRC reporter system by genomically integrating an R-loop-prone sequence previously used for conflict induction and characterized the dynamic changes of the local chromatin structure inflicted by TRCs. Importantly, the new reporter system was not only able to induce chromosomal TRCs with high efficiency but also induced drastic local replication impairments, which coincided with an activation of the DNA damage response in a local but also global manner. TRC-inducing cells were also particularly sensitive to inhibition of the Ataxia Telangiectasia and Rad3 related (ATR), a crucial kinase of the replication stress response. Analyzing the impact of TRCs on chromatin organization, I found a striking loss of nucleosome occupancy at the activated TRC reporter that was highly dependent on the stability of the associated R-loops. Furthermore, I investigated numerous histone modifications for an association with TRCs and found an increase in H3K79 methylation specifically at the R-loop forming TRC site. Through inhibition of H3K79 methylation writer enzyme disruptor of telomeric silencing-1-like (DOT1L), I showed that H3K79 methylation is actively deposited at TRC sites. Further, lack of DOT1L activity resulted in reduced transcriptional output and exacerbated DNA damage response, suggesting that deposition of this modification is required for effective transcription recovery and resolution of TRCs. Beyond the work on the reporter system, I characterized the potential of several oncogene overexpression cell lines to induce TRCs and identified CDC25 and CDC6 overexpression as potent disruptors of transcription-replication coordination. Ultimately, my work establishes a powerful new reporter system and cancer model cell lines to study TRC biology, while also uncovering novel chromatin dynamics at TRC sites and revealing a specific epigenetic modifier bookmarking TRCs, that are relevant to cancer and other diseases.

2. Zusammenfassung

Transkription und DNA-Replikation sind fundamentale Prozesse im Nukleus der Zelle, welche jeweils die Genexpression und korrekte Duplikation des Genoms kontrollieren. Eine strenge Regulierung dieser Prozesse ist von größter Bedeutung, da sowohl gestörte Transkription als auch gestörte Replikation mit zahlreichen Krankheiten einschließlich Krebs in Verbindung gebracht werden. Beide Prozesse erfordern eine umfassende Umgestaltung des zugrunde liegenden Chromatins, damit sie ihre Funktion erfüllen können. Transkriptions-Replikations-Konflikte (TRC) sind nukleare Ereignisse, bei denen Transkription und Replikation während der S-Phase des Zellzyklus auf derselben DNA kollidieren. Trotz ihres seltenen Auftretens haben sich TRCs als potente Auslöser von DNA-Schäden, Mutationen und genomischer Instabilität erwiesen. Da Transkription und Replikation das Chromatin während ihres Fortschreitens auf dem Genom umfassend umgestalten, legt eine Hypothese nahe, dass TRC-Stellen besonders anfällig für Chromatin- und Epigenomveränderungen sein könnten, die wiederum eine Instabilität des Genoms begünstigen könnten. In dieser Studie habe ich ein induzierbares TRC-Reportersystem entwickelt, indem ich eine zuvor zur Konfliktinduktion verwendete R-Loop-anfällige Sequenz genomisch integriert und die durch TRCs verursachten dynamischen Veränderungen der lokalen Chromatinstruktur charakterisiert habe. Das neue Reportersystem induziert nicht nur chromosomale TRCs mit hoher Effizienz, sondern erzeugt auch drastische lokale Replikationsbeeinträchtigungen, die mit lokalen und globalen DNA-Schäden einhergingen. TRC-induzierende Zellen reagierten außerdem besonders empfindlich auf die Inhibierung von Ataxia Telangiectasia and Rad3 related (ATR), einer entscheidenden Kinase in der Reaktion auf Replikationsstress. Bei der Analyse der Auswirkungen von TRCs auf die Chromatinorganisation stellte ich einen deutlichen Verlust in der Dichte von Nukleosomen am aktivierten TRC-Reporter fest, der stark von der Stabilität der zugehörigen R-Loops abhängig war. Darüber hinaus habe ich zahlreiche Histon Modifikationen auf eine Assoziation mit TRCs untersucht und einen Anstieg der H3K79-Methylierung speziell an den R-Loop-bildenden TRC-Stellen festgestellt. Durch Hemmung des H3K79-Methylierungs-Writer-Enzyms Disruptor of Telomeric Silencing-1-like (DOT1L) konnte ich zeigen, dass H3K79-Methylierung aktiv an TRC-Stellen positioniert wird. Weiterhin führte das Fehlen von DOT1L-Aktivität zu einer verringerten Transkriptionsleistung und verstärkten DNA-Schäden, was darauf schließen lässt, dass die aktive Positionierung dieser Histonmodifikation für einen effektiven Neustart der Transkription und Auflösung von TRCs erforderlich ist. Abgesehen von der Arbeit am Reportersystem habe ich das Potenzial zur Auslösung von TRC in verschiedener Zelllinien mit Überexpression von Onkogenen

charakterisiert. Dabei habe ich die Überexpression von CDC25 und CDC6 als potente Störfaktoren der Transkriptions-Replikations-Koordination identifiziert. Letztendlich etabliert meine Arbeit ein leistungsstarkes neues Reportersystem und Krebszellen basierte Modelle zur Untersuchung der TRC-Biologie, deckt neue Chromatindynamiken an TRC-Stellen auf und charakterisiert eine spezifische epigenetische Modifikation, die TRCs markiert und für Krebs und andere Krankheiten relevant ist.

3. Introduction

3.1 Chromatin function and organization

Chromatin is a complex polymer of DNA, proteins, and associated RNAs located in the nucleus of eukaryotic cells. Its primary function is to organize and compact the DNA, thereby allowing it to fit within the confines of the nucleus while also playing a fundamental role in regulating gene expression and DNA replication. The basic organization unit of chromatin is the nucleosome, which consists of a segment of about 147 base pairs (bp) of DNA wound around an octameric core of histone proteins (two molecules each of H2A, H2B, H3, and H4) (**Fig. 1**) (Luger et al., 1997). Adjacent nucleosomes are linked to each other by small segments of 20-50 bp linker DNA, which is frequently bound by histone H1 to stabilize the nucleosome core particle and enable the formation of higher-order chromatin structures (Hergeth & Schneider, 2015; Thoma et al., 1979). This "beads on a string" nucleosome configuration has been proposed to further organize in compacted and fiber-like structures with a diameter of ~30 nm (**Fig. 1**) (Philip JJ Robinson & Rhodes, 2006; Thoma et al., 1979). Nevertheless, the function, existence *in vivo,* and exact shape of this higher-order chromatin fiber remains a matter of active research and debate (Maeshima et al., 2019; Quénet et al., 2012). The action of chromatin architectural proteins such as CCCTC-binding factor (CTCF) and cohesin complexes further enable the formation of chromatin loops and topologically associated domains (TADs) (**Fig. 1**) (Beagan & Phillips-Cremins, 2020; Dixon et al., 2012; Gabriele et al., 2022; Hongshan Zhang et al., 2023). TADs are megabase-scale genomic regions whose DNA sequences preferentially contact each other (Beagan & Phillips-Cremins, 2020; Dixon et al., 2012). They have been shown to regulate gene expression through enhancer-promoter interaction or restriction and aid the compartmentalization of chromatin (Beagan & Phillips-Cremins, 2020; Cavalheiro et al., 2021). Finally, individual chromosomes are compartmentalized and occupy discrete territories within the nucleus (Misteli, 2020), thereby providing the basic scaffold for nuclear organization (**Fig. 1**).

From a simplified point of view, chromatin can be functionally divided into two primary forms: euchromatin and heterochromatin. Euchromatin, is generally less compacted, transcriptionally active, early replicating, and located towards the center of the nucleus (Hildebrand & Dekker, 2020; Julienne et al., 2013; Lieberman-Aiden et al., 2009; Pope et al., 2014). Euchromatic regions are typically gene-rich and contain many active genes required for cell homeostasis (Hildebrand & Dekker, 2020). Moreover, the more open state of euchromatin further enables replication-dependent and independent DNA repair processes (Lorat et al., 2012). Heterochromatin, on the

other hand, is more densely packed, transcriptionally inactive, late-replicating, peripheral, and associated with the nuclear lamina (Hildebrand & Dekker, 2020; Julienne et al., 2013; Pope et al., 2014). Heterochromatin is frequently further subdivided into constitutive heterochromatin, a permanently condensed state found for example at centromeres, telomeres, and other repetitive sequences (Saksouk et al., 2015), and facultative heterochromatin, which can more easily switch from an inactive to an active state upon adequate stimulation (Grewal, 2023). The maintenance and balance between euchromatin and heterochromatin are fundamental for normal cellular function and integrity (Mirabella et al., 2015).

The structure-function relationship of chromatin is regulated by numerous highly interdependent processes, including histone modification (such as acetylation, methylation, ubiquitination, and phosphorylation), the incorporation of histone variants, chromatin remodeling, and topological rearrangements (Bannister & Kouzarides, 2011; Millán-Zambrano et al., 2022). These modifications can either condense or decondense chromatin or alter its nuclear localization, ultimately, influencing gene expression. In addition to its role in gene regulation, chromatin is also involved in DNA repair and replication (Alabert et al., 2017; Alabert & Groth, 2012; Lorat et al., 2012). During cell division, chromatin compaction ensures that the DNA is accurately replicated and properly segregated into the two daughter cells (Schneider et al., 2022). Deregulation of chromatin structure or function can lead to various diseases, including cancer, thereby highlighting its critical role in numerous cellular processes such as transcription or DNA replication (Mirabella et al., 2015).

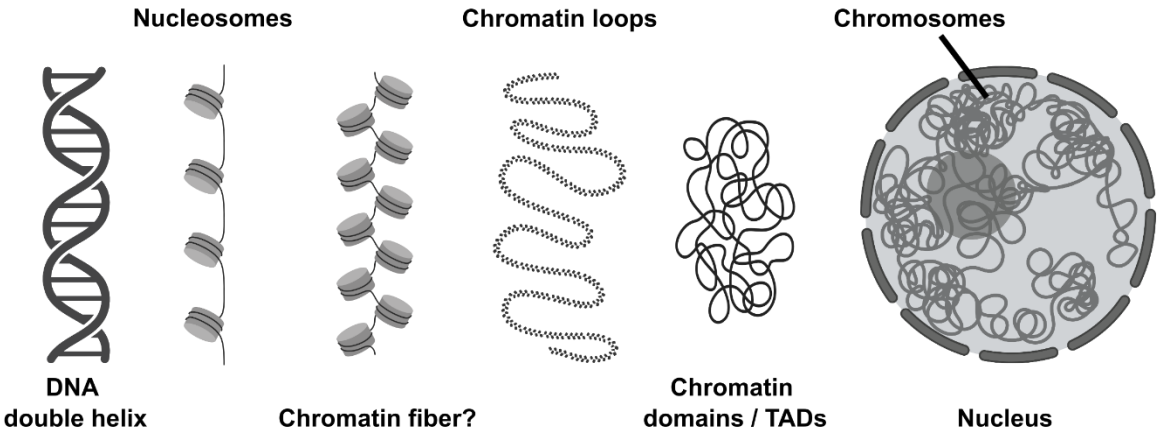

Figure 1: The organization levels of DNA in the eukaryotic genome

The DNA double helix is wrapped around the nucleosome, which is made up of an octameric core histones. Although debated, nucleosomes might organize into chromatin fibers, which fold into chromatin loops. These loops can be further grouped into chromatin domains, also called TADs, which associate with each other to form chromatin compartments. The DNA of each chromosome occupies a distinct chromosome territory within the cell nucleus. Created in BioRender.com.

3.2 Transcription regulation and co-transcriptional secondary structures

During the development and life of an organism, specific gene expression programs need to be established and maintained in different cell types. This is achieved by the process of transcription, where RNA polymerase enzymes synthesize RNA in a DNA-dependent manner. In eukaryotic cells, three different RNA polymerases have been identified that transcribe different classes of genes (Cramer, 2019; Roeder & Rutter, 1969; Sentenac, 1985). RNA polymerase I (RNAPI) transcribes large ribosomal RNA precursors, RNA polymerase II (RNAPII) synthesizes messenger RNAs (mRNAs) and several non-coding RNAs, and RNA polymerase III (RNAPIII) produces transfer RNAs (tRNAs) and small ribosomal RNAs (Sentenac, 1985). Besides their target genes, the three RNA polymerase complexes differ in their regulation and associated cofactors (Cramer, 2019).

For the initiation of transcription, an RNA polymerase first needs to recognize the promoter of its target gene. Efficient promoter recognition and binding requires the concerted action of an RNA polymerase with the basal transcriptional machinery or general transcription factors and is aided by numerous coactivators such as transcription factors as well as chromatin remodelers that open up chromatin to create an environment permissive for transcription (Knezetic & Luse, 1986; Lambert et al., 2018; Lorch & Kornberg, 2017; Utley et al., 1998; Zhu et al., 2018). Further sequence features of promoters like CpG islands and TATA box elements contribute to promoter recognition, activity, and transcriptional output (Deaton & Bird, 2011; Müller & Tora, 2014; Schübeler, 2015). As RNAPII transcribes the vast majority of eukaryotic genes, transcription initiation, elongation, and termination has been extensively characterized for this enzyme. Transcription initiation begins with the formation of the pre-initiation complex (PIC) that includes several general class II initiation factors that bind upstream DNA sequences and aid in the correct positioning of RNAPII at the promoter (He et al., 2016; Kostrewa et al., 2009; Louder et al., 2016; Philip J. Robinson et al., 2016). The local chromatin context such as nucleosome positioning and DNA topology have been suggested in aiding PIC establishment (Andersen et al., 2017; Levens et al., 2016; Vermeulen et al., 2007). Particularly, a well-positioned +1 nucleosome downstream of the PIC and an upstream -1 nucleosome demarcate a nucleosome-free region of around 200 bp and contribute to transcription regulation (Abril-Garrido et al., 2023; Jiang & Pugh, 2009). Following recognition, the promoter will be opened via the action of TFIIH subunit XPB. XPB acts as a DNA translocase that unwinds DNA and propels it into the active center of RNAPII (Grünberg et al., 2012; Holstege et al., 1996). RNAPII initiation is enabled by the action of the Mediator

complex, which stimulates the phosphorylation of the RNAPII C-terminal domain (CTD), a tail-like extension of the largest RNAPII subunit RPB1. CTD phosphorylation on Serine 5 (Ser5P) and Serine 7 (Ser7P) by TFIIH kinase subunit cyclin-dependent kinase 7 (CDK7) results in promoter escape and transition into transcription elongation (**Fig. 2**) (Ebmeier et al., 2017; Eick & Geyer, 2013; Kornberg, 2005; Wong et al., 2014). Especially Ser5P was shown to reduce the affinity of RNAPII to the Mediator complex and facilitate promoter escape (Singh et al., 2022; Velychko et al., 2024). In metazoan cells transcription is also regulated at the elongation step. Frequently, RNAPII engages in promoter-proximal pausing, during which RNAPII pauses ~50 bp downstream of the transcription start site (TSS) (Core & Adelman, 2019; Eick & Bornkamm, 1986; Rougvie & Lis, 1988; Strobl & Eick, 1992). Pausing is mediated by negative elongation factor (NELF) and DRB sensitivity inducing factor (DSIF) (Bernecky et al., 2017; Vos, Farnung, Urlaub, et al., 2018; Yamaguchi et al., 2013). RNAPII pausing is thought to act as an intermediate step to recruit additional elongation factors to form a stable elongation complex (Vos, Farnung, Boehning, et al., 2018). Positive elongation factor b (P-TEFb) subunit cyclin-dependent kinase 9 (CDK9) phosphorylates NELF, DSIF, and the RNAPII CTD to release RNAPII from pausing into processive elongation (Vos, Farnung, Boehning, et al., 2018; Yamada et al., 2006; Q. Zhou et al., 2012). Actively elongating RNAPII is marked by Tyrosine 1 phosphorylation (Tyr1P), Serine 2 phosphorylation (Ser2P), and Threonine 4 phosphorylation (Thr4P), which are deposited by numerous CTD kinases (Buratowski, 2009; Tellier et al., 2020). These marks accumulate throughout the gene body but disappear towards transcription termination (**Fig. 2**). Whereas Ser2P and Thr4P have been shown to recruit elongation factors for RNAPII processivity, Tyr1P prevents binding of termination factors and is removed before termination occurs (Singh et al., 2022). In contrast to initiation and elongation, the details of RNAPII termination are less well understood since termination can occur at many stages of the transcription cycle (Proudfoot, 2016). Termination of a productive transcription cycle occurs at the polyadenylation signal (PAS) (Huimin Zhang et al., 2015). Here a torpedo model of nucleolytic RNA degradation by exonuclease XRN2 (Luo et al., 2006) and an allosteric switch model with protein phosphatase 1 (PP1)-dependent dephosphorylation of the CTD and RNA cleavage have been proposed (Parua et al., 2018).

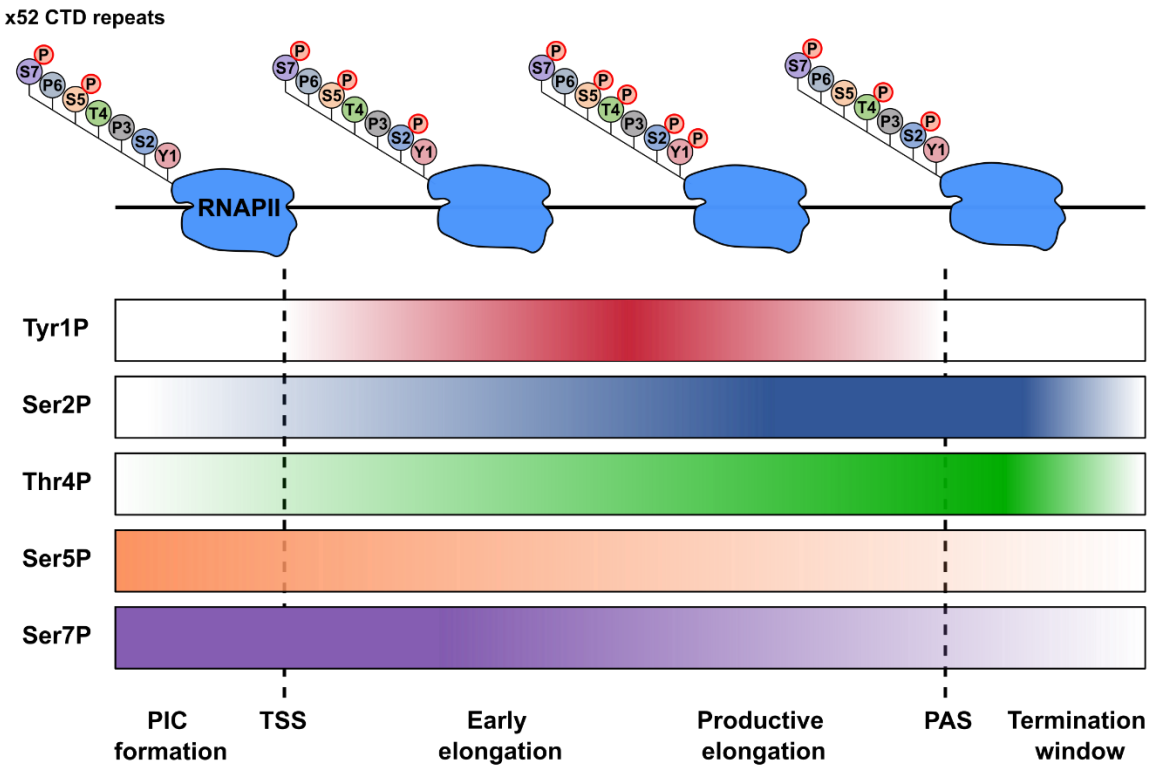

Figure 2: The phosphorylation state of the RNAPII CTD is regulated during transcription.

During the transcription cycle of human RNAPII and its progression through the transcribed gene, different phosphorylation modifications are added or removed on the 52 heptad repeats of the CTD, thereby promoting unique functions such as promoter escape, processive elongation, and termination.

The protective chromatin structure is necessarily disrupted during the transcription cycle, providing a window of opportunity for secondary structures to form between transcribed RNA molecules and DNA as well as on DNA alone. Thus, many of these non-canonical structures described below arise frequently during transcription of all RNA polymerases. Strikingly, elevated levels of these structures have been shown to cause DNA damage and constitute an endogenous threat to genomic integrity (García-Muse & Aguilera, 2019). For example, R-loops comprise RNA:DNA hybrids that form co-transcriptionally when the transcribed RNA molecule invades double-stranded DNA (dsDNA) behind the polymerase and hybridizes to the template strand (Brickner et al., 2022; García-Muse & Aguilera, 2019; Petermann et al., 2022). Coincidingly, the non-template strand is displaced as single-stranded DNA (ssDNA) (Subin Kim et al., 2024). Quantitative analysis of R-loops using imaging and sequencing techniques has demonstrated that RNA:DNA hybrids occupy 5-10 % of the genome and reach sequence lengths reaching from less than 60 bp up to several kilobases (Brickner et al., 2022; Chédin, 2016; Crossley et al., 2020, 2021; Ginno et al., 2012). Due to their transcriptional origin, R-loops are predominantly genic

features but have also been found at enhancers, centromeres, telomers, and other repetitive genomic regions (Q. Liu et al., 2021; Niehrs & Luke, 2020; Tan & Lan, 2020). Moreover, R-loop formation is favored in regions with high GC-content and GC-skew, which describes a strong strand asymmetry in the distribution of G and C residues (Castillo-Guzman & Chédin, 2021). Importantly, the exposed ssDNA opposing the R-loop is susceptible to secondary structure formation like G-quadruplexes (G4) and DNA damage (Subin Kim et al., 2024). Transcription with simultaneous R-loop formation has been shown to slow the progression of RNA polymerases (Belotserkovskii et al., 2020; Tous & Aguilera, 2007), thereby causing potential discoordination with other genomic processes such as DNA replication (Huertas & Aguilera, 2003). Stable R-loop formation frequently impairs the progression of replication forks, thereby inducing DNA replication stress and genomic instability (García-Muse & Aguilera, 2019; Hamperl et al., 2017). Consequently, the cell employs numerous enzymes to remove R-loops. R-loop helicases such as Aquarius (AQR) or Senataxin (SETX) can unwind hybrids (Hasanova et al., 2023; Skourti-Stathaki et al., 2011; Yang et al., 2023), whereas the R-loop specific ribonucleases (RNases) RNase H1 and RNase H2 degrade the RNA part of the R-loop for removal (Hyjek et al., 2019; H. Zhao et al., 2018). Despite being a threat to genomic integrity, R-loops also have physiological roles in mediating class-switch recombination in B-cells (Yu et al., 2003), contributing to gene regulation (Niehrs & Luke, 2020) and promoting DNA repair processes in transcribed regions as well as telomeres (Marnef & Legube, 2021; Tan & Lan, 2020). Consequently, tight regulation of R-loop transcription is essential to preserve nuclear function, gene regulation, and genomic integrity.

3.3 DNA replication and replication stress

The accurate duplication of the genetic information contained in the DNA double helix is essential for the inheritance and maintenance of cell identity in an organism. DNA replication occurs during all stages of life starting with the initial division of the fertilized egg, continuing throughout development, and maintaining tissue integrity in adult life. The human genome consists of roughly 3 billion base pairs separated into 24 different chromosomes in males (Nurk et al., 2022). The vast majority of cells contain two copies of the genome, amounting to a total of 6 billion base pairs in a single cell. Considering that trillions of new cells are made in a human lifetime; accurate and highly regulated duplication of the genome is required to maintain cell function and life. Failure to preserve faithful DNA replication quickly leads to elevated mutation rates, frequently associated with cancer and other diseases (Mertz et al., 2017).

DNA replication initiates at chromosomal positions called replication origins. Origins are recognized by the origin recognition complex (ORC) (Bell et al., 1993; Foss et al., 1993). ORC binds DNA in a ring-like fashion together with Cell Division Cycle 6 (CDC6), thereby bending the DNA for further loading of the minichromosome maintenance complex (MCM2-7) that acts as the replicative helicase (Feng et al., 2021). Via a number of intermediate steps, ORC, CDC6, and Chromatin licensing and DNA replication factor 1 (CDT1) assist in loading the replicative MCM2-7 double hexamer for bidirectional DNA replication (Remus et al., 2009; Weissmann et al., 2024). Next, numerous firing factors enable the recruitment of CDC45 and GINS that can convert the loaded MCM2-7 hexamers into the activated replicative helicase CMG (CDC45-MCM2-7-GINS) (Ilves et al., 2010). DNA is synthesized in a continuous fashion by DNA polymerase ε (POLE) on the leading strand and in a discontinuous manner by DNA polymerase δ (POLD) on the lagging strand (Garg & Burgers, 2005; Nick McElhinny et al., 2008; Roske & Yeeles, 2024). Lagging strand synthesis is aided by DNA polymerase α/primase synthesizing an RNA primer to initiate DNA synthesis by POLD (Yuan et al., 2023). Discontinuous synthesis results in short (100-200 bp) Okazaki fragments that are later matured by the concerted action of nucleases, DNA polymerases, and ligases (H. Sun et al., 2023).

Unlike in simple eukaryotes such as *Saccharomyces cerevisiae* (yeast), human DNA replication origins are not defined by a consensus sequence. Instead, a combination of sequence features and local chromatin composition was proposed as the determinant of DNA replication initiation in human cells (Hu & Stillman, 2023; Hyrien, 2016), thus leading to the definition of initiation zones (Petryk et al., 2016). Furthermore, DNA replication occurs in a highly coordinated and defined temporal manner, resulting in a replication timing program that organizes the genome into early- or late-replicating regions (Fragkos et al., 2015; Hansen et al., 2010; Rhind & Gilbert, 2013). Replication timing has been linked to 3D genome organization and shown to be cell type-specific (Emerson et al., 2022; Marchal et al., 2019; Sima et al., 2019).

Interference with DNA replication by a number of cellular stresses is collectively termed replication stress. DNA replication stress is defined by the slowing or stalling of replication forks and/or DNA synthesis (Saxena & Zou, 2022; Zeman & Cimprich, 2013). Importantly, replication stress threatens genome stability in normal and cancer cells and is considered a hallmark of cancer (Hanahan & Weinberg, 2000, 2011). Major sources of replication stress are limited nucleotide pools, unrepaired DNA lesions, DNA secondary structures, fragile sites, R-loops, and transcription complexes (Saxena & Zou, 2022; Zeman & Cimprich, 2013). Additionally, the expression of numerous oncogenes can cause replication stress by promoting increased unscheduled

replication initiation, dysregulation of the replication checkpoint, and induction of fork reversal (Kotsantis et al., 2016; Macheret & Halazonetis, 2018; Neelsen et al., 2013). To deal with replication stress and maintain genomic integrity, cells have evolved a complex network of stress response pathways to prevent, overcome, and tolerate replication impairments. A replication stress response includes several interconnected responses at the replication fork including the activation of the replication checkpoint (Recolin et al., 2014), fork remodeling (W. Liu et al., 2020), and initiation of DNA repair (Oh & Myung, 2022) or tolerance pathways (Ashour & Mosammaparast, 2021; Buoninfante et al., 2023). (Ashour & Mosammaparast, 2021; Buoninfante et al., 2023). If replication stress is not adequately resolved, it can lead to replication fork collapse, which results in DNA breaks that prevent proper duplication of the genome (Kondratick et al., 2021). This can trigger further genomic instability processes in mitosis, such as micronuclei formation, and anaphase bridges (Saxena & Zou, 2022), which are frequently observed during the malignant transformation of cancer cells (Bignell et al., 2010; Hanahan & Weinberg, 2000, 2011)

Ataxia Telangiectasia and Rad3 related (ATR) and Ataxia Telangiectasia Mutated (ATM) are two key kinases that play essential roles in the cellular response to DNA damage and replication stress. ATR is primarily activated in response to stalled replication forks that frequently accumulate ssDNA (Choi et al., 2010; Zou & Elledge, 2003). This ssDNA is rapidly coated by Replication Protein A (RPA), which recruits ATR via its partner ATR-Interacting Protein (ATRIP) (Cortez et al., 2001; Zou & Elledge, 2003). Alternatively, ATR can be activated via Ewing's tumor-associated antigen 1 (ETAA1) that binds RPA directly and contains an ATR-activating domain (Haahr et al., 2016; Saldivar et al., 2018). Both pathways appear to be working in an independent but complementary manner (Haahr et al., 2016). Upon activation, ATR phosphorylates several downstream targets, most notably the checkpoint kinase 1 (CHK1) (Matsuoka et al., 2007). This phosphorylation triggers a cascade of events aimed at stabilizing the replication fork, thus, preventing its collapse (Toledo et al., 2013). Additionally, ATR activation leads to the inhibition of origin firing, slowing down the overall replication process and providing time for the replication checkpoint to overcome the stress (Moiseeva et al., 2019; Toledo et al., 2013). Crucially, ATR function has been shown to not only impair the progression of a stalled fork locally but instead extends to adjacent forks in larger chromatin domains (Collins et al., 2020; Mutreja et al., 2018). In contrast, ATM is predominantly activated by DNA double-strand breaks (DSBs) (Falck et al., 2005; Ji Hoon Lee & Paull, 2005). Upon detecting a DSB, ATM is rapidly activated through autophosphorylation and consequently phosphorylates a wide range of target proteins, including the checkpoint kinase CHK2 (Bakkenist & Kastan, 2003; Matsuoka et al., 2007). Activation of the

ATM kinase ultimately halts the cell cycle at the G1/S and G2/M checkpoints (Kastan et al., 1992; Khanna et al., 2001) and further regulates repair via homologous recombination (HR) or non-homologous end joining (NHEJ) pathway choice (Britton et al., 2020; Chanut et al., 2016). Despite their specialization, both kinases frequently crosstalk to enable an integrated DNA damage response.

3.4 Transcription replication-conflicts

The genome serves as a template for a large number of nuclear processes such as transcription, epigenetic regulation, DNA replication, and DNA repair. Many processes require intensive crosstalk and cooperation with each other in order to maintain cellular function and integrity (García-Muse & Aguilera, 2016). Nevertheless, the co-occurrence of two machineries at the same genomic location and time can lead to interference. One endogenous cause of replication stress are TRCs, during which transcribing RNA polymerases collide with active replication forks on the same DNA strand (Bermejo et al., 2012; Brewer, 1988; García-Muse & Aguilera, 2016; Lalonde et al., 2021). These conflicts have been shown to give rise to genomic instability (Helmrich et al., 2011), mutations (Lang et al., 2017; Sankar et al., 2016), DNA breaks (Patel et al., 2023; St Germain et al., 2022), and perturbed DNA replication (Groelly et al., 2022; Stoy et al., 2023), highlighting the importance of tightly controlled transcription-replication coordination in all dividing cells. TRCs can occur in two different forms depending on the relative orientation of transcription and replication towards each other. Transcription and replication complexes moving towards each other will result in head-on (HO) TRCs, whereas movement in the same direction will induce co-directional (CD) TRCs (**Fig. 3**) (Hamperl et al., 2017). Importantly, HO-TRCs have been shown to be the more deleterious type of TRCs in various model organisms due to their ability to induce pausing and blockage of the replication fork, which in turn may lead to its collapse and the formation of DNA breaks (Hamperl et al., 2017; Lang et al., 2017). In contrast, CD-TRCs are considered less harmful and are more easily overcome by displacement of the RNA polymerase (Brüning & Marians, 2020; Bruno et al., 2024; Hamperl et al., 2017; Lang et al., 2017). Additionally, HO conflicts are particularly susceptible to topological problems as both transcription and replication accumulate positive supercoiling in their direction of movement, leading to a drastic buildup of positive supercoiling at HO but not CD conflicts (Lang & Merrikh, 2021). Thus, the removal of topological stress through topoisomerases is required to prevent genomic instability at TRC sites (Lang & Merrikh, 2021; Y. Liu et al., 2021; Promonet et al., 2020). The more detrimental role of HO-TRCs is further supported by an evolutionary-driven co-directional orientation bias of

transcription and replication in bacteria and eukaryotes (Marsolier-Kergoat & Goldar, 2012; Petryk et al., 2016; J. D. Wang et al., 2007). Efficiently firing replication origins frequently overlap with TSSs of highly transcribed genes, thereby preventing HO-collisions (Marsolier-Kergoat & Goldar, 2012; Sequeira-Mendes et al., 2009). Moreover, evidence from plasmid-based TRC reporter systems suggests that cells induce a conflict orientation-dependent DNA damage response. Particularly, HO conflicts have been shown to induce the ATR signaling cascade, whereas CD-TRCs rely on the activation of the ATM pathway (Hamperl et al., 2017). Consistent with results from bacteria, the plasmid-based HO/CD constructs showed elevated R-loops levels in the HO orientation, supporting a role for these structures in promoting TRC-induced genomic instability (García-Muse & Aguilera, 2019; Hamperl et al., 2017; Lang et al., 2017).

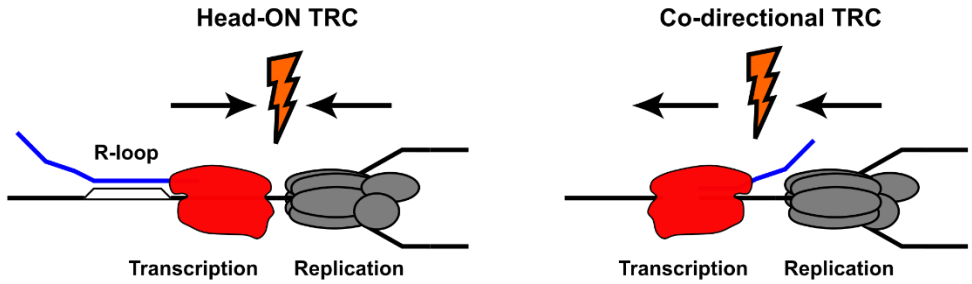

Figure 3: TRCs can occur in head-on and co-directional orientation

Transcription and replication complexes moving towards each other cause head-on TRCs, whereas movement in the same direction induces co-directional TRCs. Particularly, head-on conflicts favor the formation of stable R-loops.

3.4.1 Fragile sites are genomic TRC hotspots

While TRCs have the potential to threaten genome stability in any dividing cell, it is important to note that in the majority of genomic locations, transcription and replication occur in a coordinated manner, thereby preventing TRCs. Additionally, the DNA damage response signaling, TRC resolution factors, and R-loop processing enzymes ensure that TRCs are rapidly overcome, and genome integrity is maintained (Lalonde et al., 2021). Nevertheless, a small subset of genomic regions, termed common fragile sites (CFSs) (Debatisse et al., 2012; Durkin & Glover, 2007; T. W. Glover, 1981; Thomas W. Glover et al., 1984), remains particularly susceptible to replication stress and TRC-driven instability (Helmrich et al., 2011). Breakages in these regions can lead to chromosomal rearrangements, deletions, or amplifications, thereby contributing to cancer and other diseases (Bignell et al., 2010; Hellman et al., 2002). Many tumor suppressor genes are

found in CFSs, making them vulnerable to oncogenic mutations (Bignell et al., 2010; S. Li & Wu, 2020). CFSs are cell type-specific, located in gene-rich regions, and exceptionally large, spanning hundreds of kilobases to megabases (Durkin & Glover, 2007; S. Li & Wu, 2020). The enormous length of CFS genes and their elevated propensity to form DNA secondary structures make them difficult templates for both transcription and replication (Helmrich et al., 2011). In fact, transcription of CFSs can require more than one cell cycle to be completed (Helmrich et al., 2011). This inherently creates a problem for DNA replication since the underlying DNA sequence must be faithfully replicated during S-phase to ensure that no genetic information is lost in daughter cells. Consequently, CFSs are likely hotspots for TRCs and TRC-driven genomic instability (Helmrich et al., 2011). Besides CFSs, also other types of fragile sites such as early replicating fragile sites (ERFS) in mouse B-cells (Barlow et al., 2013; St Germain et al., 2022) and recurrent DNA break clusters (RDCs) (Corazzi et al., 2024; Wei et al., 2016) have been described to underlie TRC-dependent genomic instability. ERFSs in contrast to CFS are located in gene- and replication origin-rich regions with high transcription levels and early replication timing (Barlow et al., 2013; St Germain et al., 2022). While the exact mechanisms of ERFSs instability remain unclear, high transcription and putative TRCs have been demonstrated as crucial contributors (Barlow et al., 2013; St Germain et al., 2022). RDCs have been primarily described in the replication timing transition regions of neuronal progenitor cells, in which sparse replication origins connect unidirectional forks (Corazzi et al., 2024). RDCs show a particularly high density of R-loops and enrichment for HO-TRC-driven DSBs (Corazzi et al., 2024), demonstrating that TRCs can threaten genomic integrity in different genomic contexts and cell types.

3.4.2 Oncogenes as drivers of transcription-replication interference

Genome instability and epigenetic dysregulation are hallmarks of cancer and contribute to tumorigenesis and malignant behavior of many tumors (Hanahan & Weinberg, 2000, 2011). Cancer cells frequently show aberrant replication and transcription programs that could favor the occurrence of TRCs and genome destabilization. Additionally, oncogene-driven hypertranscription has recently been connected to replication stress since unscheduled excessive transcription could result in exacerbated discoordination with replication and high TRC levels (Bowry et al., 2021). Evidence from numerous recent studies suggests that several commonly amplified or overexpressed oncogenes can cause TRCs and genomic instability in different cancer models. Particularly, inducible overexpression of oncogenes *CCNE1* and *MYC,* two of the most frequently amplified genes in human cancers (Beroukhim et al., 2010; Bignell et al., 2010;

Zack et al., 2013), were shown to induce firing of novel replication origins within highly transcribed genes in U-2-OS osteosarcoma cells (**Fig. 4A**) (Macheret & Halazonetis, 2018). These oncogene-induced replication forks showed high levels of fork collapse and DSBs due to TRCs, supporting the concept of TRCs as drivers of tumor progression. In contrast, neuronal MYC isoform MYCN expression in neuroblastoma was shown to cooperate with the nuclear exosome to prevent DSBs and maintain transcription elongation, thereby mitigating TRCs in these highly proliferating cells (Papadopoulos et al., 2022). Studies using the overexpression of glycine to valine mutated HRAS on amino acid 12 ($HRAS^{V12}$) in human fibroblasts, further demonstrate that oncogene-perturbed transcription dynamics can cause replication stress via hypertranscription and resulting TRCs (**Fig. 4B**), thus driving genomic instability in cancer (Kotsantis et al., 2016). For other oncogenes such as phosphatase CDC25A (Neelsen et al., 2013) or origin licensing factor CDC6 (Zampetidis et al., 2021), S-phase and/or transcription-associated replication stress has been described but a direct connection to TRCs has not been established (**Fig. 4C**). However, increased DNA DSBs at TSSs imply TRCs as potentially essential contributors (Zampetidis et al., 2021). Collectively, many oncogenes are involved or likely involved in transcription and replication deregulation that directly or indirectly can lead to TRCs. Considering the complex dysregulation of several oncogenes and tumor suppressor proteins in cancer, a more comprehensive and cancer-type-specific analysis will be required to understand oncogenic TRC formation at a high level of detail.

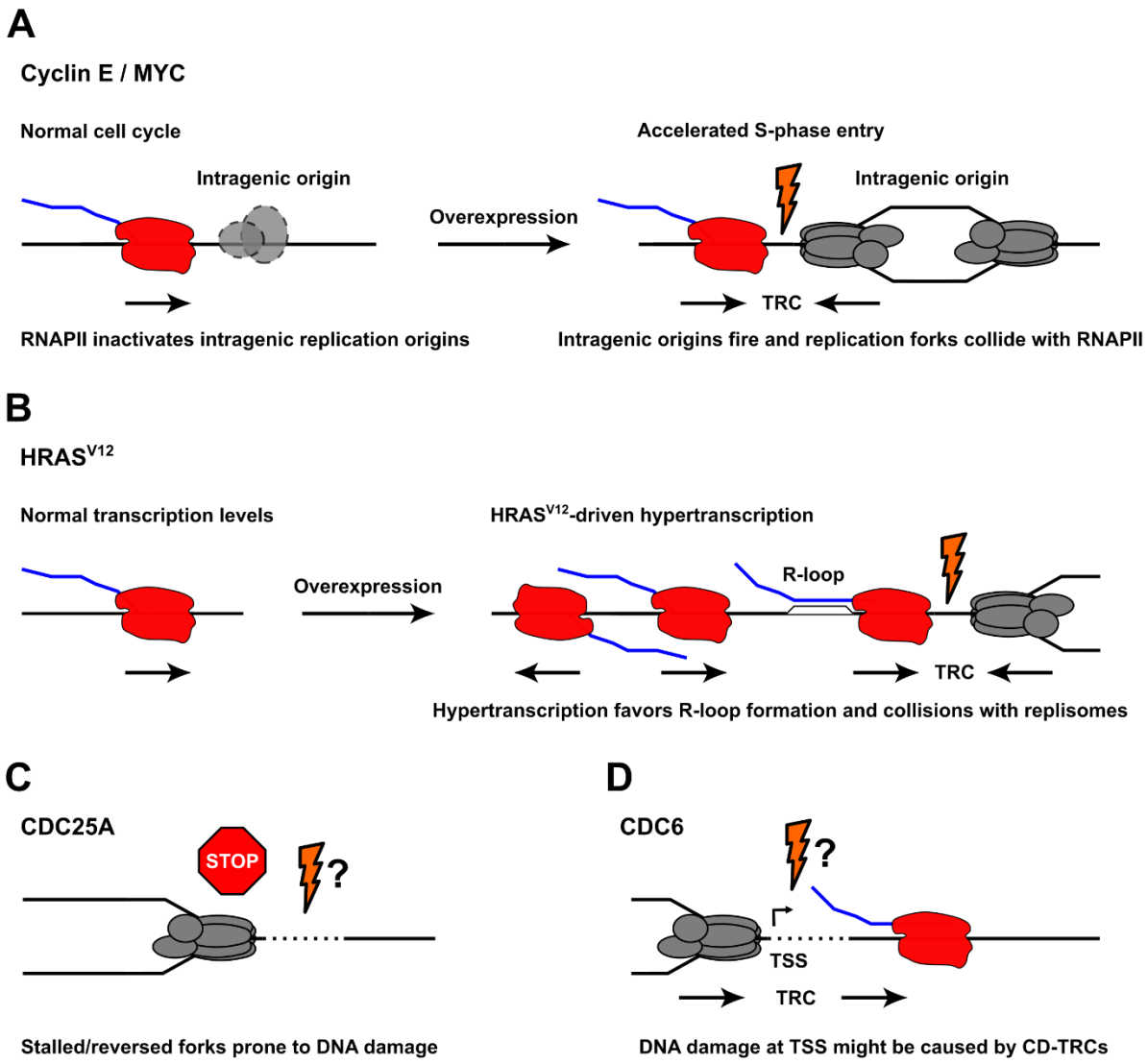

Figure 4: Current view on how oncogenes could drive TRC formation

A) In a normal cell transcribing RNAPII inactivates intragenic origins in G1-phase, thereby preventing their firing in subsequent S-phase. Upon overexpression of oncogenes Cyclin E and MYC S-phase entry is accelerated and RNAPII does not have sufficient time to clear intragenic origins. These cryptically firing origins give rise to replication forks that clash with still ongoing transcription and create TRCs. (Macheret & Halazonetis, 2018)

B) Overexpression of oncogenic $HRAS^{V12}$ mutant leads to hypertranscription. This unscheduled transcription is more prone to stalling and thereby favors R-loop formation and collisions with the replisome, thereby giving rise to TRCs (Kotsantis et al., 2016)

C) CDC25A controls the S-phase and M-phase entry checkpoints and works upstream of Cyclin E. So far, no direct evidence for TRCs has been shown upon overexpression. Nevertheless, increased CDC25A levels favor fork stalling and reversal as well as DNA damage, processes frequently occurring at TRC sites (Neelsen et al., 2013).

D) CDC6 is crucial for the formation of the pre-replicative complex and can induce replication stress upon overexpression. Elevated CDC6 levels were shown to cause DNA damage at TSS sites, implying an induction of CD TRCs (Zampetidis et al., 2021).

3.4.3 Model systems to study TRCs

The rare nature of TRCs depending on two pervasive nuclear machineries makes them a difficult event to study. While previous work on fragile sites and oncogenic dysregulation of transcription and replication strongly point towards TRCs as inducers of genomic instability (Groelly et al., 2022; Hamperl et al., 2017; Helmrich et al., 2011; Kotsantis et al., 2016; Lang et al., 2017; Macheret & Halazonetis, 2018), many questions including TRC frequency, duration, genomic location, and local proteome remain unanswered. This problem is partially due to the limitation in methodology to detect and quantify TRCs. Proximity ligation assay (PLA) with antibodies targeting RNAPII and proliferation cell nuclear antigen (PCNA) has been established as a proxy for visualization and quantification of TRC levels (Hamperl et al., 2017; Lalonde et al., 2021). It remains the only currently available method for probing direct vicinity of transcription and replication machineries implying TRCs. Furthermore, specific induction and characterization of TRC-dependent molecular changes remain difficult since nearly all current methods to induce TRCs in eukaryotic cells rely on drug treatments or aberrant pathway activation via oncogenes. For example, short-term transcription elongation inhibition by 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), a CDK9 inhibitor, induces a spike of TRC levels within 30 min (Shao et al., 2020), whereas long-term DRB treatment reduced TRC occurrence, thus highlighting the complexity of TRC dynamics (Petropoulos et al., 2024). Similarly, elevated topological stress imposed by topoisomerase inhibitors can provoke an elevated TRC burden (Lang & Merrikh, 2021; Y. Liu et al., 2021). Beyond the previously described action of several oncogenes (see 3.4.2), exposure of breast cancer cells to estrogen can induce transcriptional bursting of several hundred target genes that impose a high R-loop burden (Stork et al., 2016), likely giving rise to coinciding TRCs. Unfortunately, all these approaches are highly prone to induce genomic disturbances and dysregulation independent of TRCs, making it impossible to directly link a given downstream effect to TRCs or R-loops exclusively.

To overcome this limitation, researchers have constructed several model systems to more precisely study TRCs and their consequences. While model systems are artificial by design and might not fully recapitulate the endogenous cellular processes, they offer specific inducibility of TRCs and R-loops and allow for mechanistic investigations. Initial studies took advantage of prokaryotes, which usually harbor one circular chromosome replicated by a single origin of replication, thus enabling the study of TRCs in a precise and localized manner. Insertion of reporter genes such as lacZ or luxABCDE next to the replication origin in *B. Subtilis*, demonstrated that conflicts in HO but not CD orientation can induce pervasive R-loop formation in a topological

stress-dependent manner (**Fig. 5A**) (Lang et al., 2017; Lang & Merrikh, 2021). If the conflicts cannot be resolved by RNase HIII processing, they can lead to a persistent replication fork block and coinciding mutagenesis (Lang et al., 2017). Independent work also in *B. Subtilis* leveraged the fact that loss of function mutations in the thymidylate synthetase reporter gene thyP3 can be used for selection via trimethoprim resistance (**Fig. 5A**) (Sankar et al., 2016). By chromosomal integration of the thyP3 gene under an inducible promoter either in HO or CD direction, the mutational consequences could be addressed. Both conflict types caused insertion and deletions (indels) reflective of the collision location, at which transcription and replication first encountered each other (Sankar et al., 2016). CD conflicts showed indels more towards the 5' region of the reporter genes, whereas HO mutations were distributed throughout the entire gene but also led to particularly high mutation rates at the respective promoter residues.

Similar to the work in prokaryotes, researchers have taken advantage of constructs harboring two repeats of a 0.6 kb internal fragment of the LEU2 gene positioned either in HO or CD orientation in front of an early firing replication origin in yeast (Prado & Aguilera, 2005). Crucially, HO but not CD conflicts could trigger genomic instability resulting in efficient recombination of the fragments into a functional LEU2 gene (**Fig. 5B**). This phenotype could be rescued by RNase H overexpression, suggesting a HO-conflict dependent stabilization of R-loops as a central contributor to recombination (Prado & Aguilera, 2005). In analogous approaches, a tetracycline-regulated LYS2 reporter system (N. Kim et al., 2007) or a galactose-inducible LEU2 gene (García-Rubio et al., 2018) were inserted close to efficient origins on yeast chromosome III either in HO or CD direction. Resulting TRCs gave rise to -2 frameshift mutations, insertions, and deletions as the result of DNA DSBs. Mutational phenotypes were generally more frequent in the HO conflict setup (García-Rubio et al., 2018; N. Kim et al., 2007). Finally, recent work established a single live cell microscopy system allowing to track the transcription and replication behavior of a reporter gene upon HO or CD collisions with the replisome (Tsirkas et al., 2022). While in this system replisomes showed remarkable robustness in unperturbed cells, deletion of topoisomerase I led to replisome stalling events, particularly for HO TRCs (Tsirkas et al., 2022).

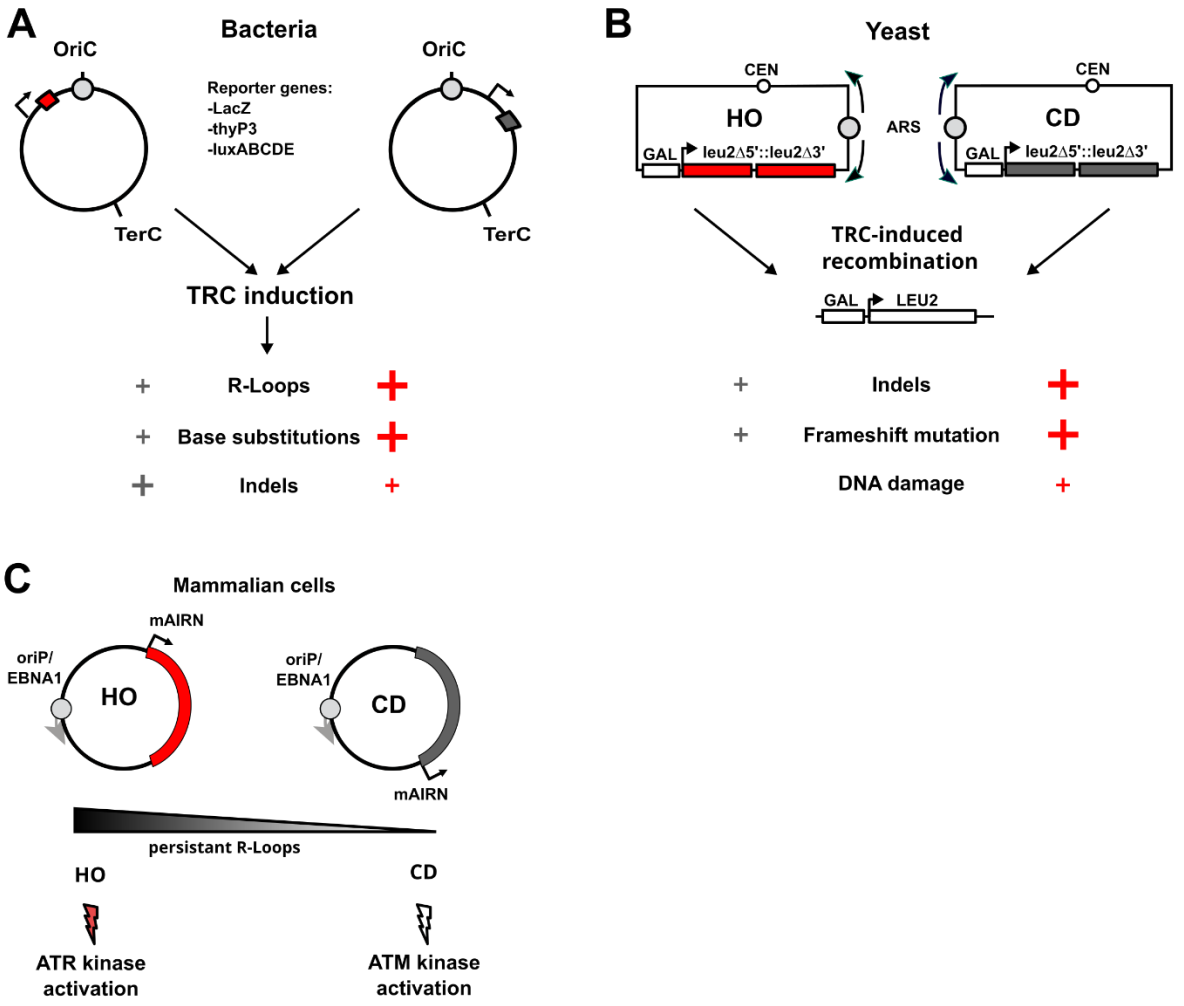

Figure 5: Model systems currently available to study transcription replication conflicts

A) In bacterial cells, reporter genes such as LacZ, thyP3, or luxABCDE were positioned in CD or HO orientation towards the chromosomal replication origin (OriC) to induce TRCs and study their mutational outcome. Although HO conflicts are more deleterious, both orientations cause R-loops and base substitutions. Additionally, insertions and deletions (indels) are more frequent at CD-TRCs, preferentially at the promoter region.

B) In yeast, LEU2 reporter constructs were engineered in the CD or HO orientation to an autonomous replicating sequence (ARS) to study the recombinogenic outcome of TRCs. HO-TRCs displayed a stronger R-loop-dependent induction of recombination and an increased frequency of indels, frameshift mutations, and DNA damage.

C) In mammalian cells, plasmids containing the R-loop prone murine Antisense Of IGF2R Non-Protein Coding RNA (mAIRN) gene in CD or HO orientation relative to the viral unidirectional replication origin P/Epstein-Barr nuclear antigen 1 (oriP/EBNA1) were used to study TRCs and orientation dependent behavior. HO-oriented TRCs were characterized by persistent R-loop formation and activation of the ATR kinase, whereas CD-oriented TRCs displayed low levels of R-loop formation and showed activation of the ATM kinase.

In contrast to yeast, mammalian cells do not contain strictly defined replication origins but rather initiate DNA replication within broad replication initiation zones (Aze & Maiorano, 2018; Petryk et al., 2016). Moreover, a large excess of available replication origins allows for adaptive and stochastic origin usage, creating replication patterns with limited predictability (Macheret & Halazonetis, 2018; Petryk et al., 2016). Thus, direct positioning of an inducible reporter gene in front of an efficient origin is not possible, which complicates the study of TRCs in mammalian cells. To overcome this issue, researchers have designed plasmid systems to investigate TRCs in mammalian cells. The respective constructs contain different doxycycline (DOX) inducible transcription units, which are susceptible to R-loop formation like the promoter region of the murine Antisense of IGF2R Non-Protein Coding RNA (mAIRN) gene (Ginno et al., 2012) or control regions that do not form R-loops such as the enhanced cyan fluorescent protein (ECFP) gene (**Fig. 5C**) (Hamperl et al., 2017). The plasmids also include a single unidirectional replication origin (oriP/EBNA1) (Hamperl et al., 2017). This origin can recruit endogenous replication complexes and is activated only once in S-phase (Hodin et al., 2013; Moriyama et al., 2012; Yates et al., 2000). Cloning of the reporter gene constructs in different orientations towards the unidirectional origin enables the specific induction of either HO or CD TRCs (**Fig. 5C**). Consistent with previous results from prokaryotes and yeast, HO but not CD conflicts contributed to a stabilization of R-loop levels (Hamperl et al., 2017). Interestingly, induced R-loop forming reporter units in both orientations could cause genomic instability, albeit with different downstream signaling. HO conflicts specifically provoked an ATR-dependent DNA damage response, whereas CD conflicts predominantly activated the ATM kinase (**Fig. 5C**) (Hamperl et al., 2017). Although the plasmid system offers a controlled approach for TRC induction in mammalian cells, it likely does not capture the full complexity of TRCs in the endogenous genomic context. Particularly, these short ~10 kb plasmids might not accurately mimic topological constraints or recapitulate the native chromatin context including its highly complex regulatory dynamics.

Taken together, TRC model systems have been crucial contributors to the study of TRCs across different species and have shaped our understanding of conflict occurrence, orientation, and consequences. Importantly, all model systems are subject to specific limitations and the respective results require careful evaluation and if possible independent validation. Especially, with respect to understanding TRCs in highly complex and dynamic mammalian genomes, additional approaches and model systems are urgently needed to comprehensively explore TRC and R-loop biology.

3.4.4 TRCs can impact the chromatin landscape

Transcription and replication are both nuclear processes that involve substantial chromatin remodeling as they travel on the genome (Lalonde et al., 2021). Therefore, collisions between both protein complexes are likely to disturb local chromatin organization, possibly leading to genetic and epigenetic instability. Indeed, previous studies uncovered various roles of specific histone modifications and chromatin proteins in mitigating TRCs and resolving R-loops. Early studies in yeast identified histone H3 serine 10 phosphorylation (H3S10P) as a chromatin mark essential for R-loop-induced chromatin compaction and signaling of genomic instability (García-Pichardo et al., 2017). Subsequent work in mammals uncovered that H3S10P accumulates in large upstream domains (~1 megabase (Mb)) around HO-TRCs prone to forming R-loops (Bayona-Feliu et al., 2023) (**Fig. 6A**). Additionally, H3 lysine 4 (H3K4) methylation has been shown to act as a protective “speed bump”, which is deposited during transcription and slows replication forks, thereby reducing the chance of TRCs under replication stress (**Fig. 6B**) (Chong et al., 2020). Independently, a replication stress-triggered switch from H2AK119 crotonylation to H2AK119 ubiquitylation (H2AK119cr to H2AK119ub) was found to release RNAPII from chromatin and suppress transcription near stalled replication forks, thus decreasing TRC frequency, along with associated R-loops and DSBs (**Fig. 6C**) (Hao et al., 2022).

Apart from histone modifications, numerous other chromatin proteins have been identified as key players in TRC and R-loop resolution. The transcription activator BRG1 (BRG1) subunit of the mammalian SWI/SNF chromatin remodeling complex cooperates with the Fanconi Anemia (FA) pathway and thus resolves R-loop dependent TRCs (**Fig. 6D**) (Bayona-Feliu et al., 2021). Moreover, recent work defined putative genomic TRC hotspot regions and correlated TRC-prone genomic locations with public ENCODE ChIP-Seq data, thereby identifying a role for chromatin remodelers SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5 (SMARCA5), INO80 Complex ATPase Subunit (INO80), and Metastasis-associated protein MTA2 (MTA2) in R-loop driven genomic instability (Bayona-Feliu et al., 2023). Interestingly, the Integrator complex, predominantly known to act in transcription termination of RNAPII, was shown to attenuate CD TRCs and genomic instability by removing stalled RNAPII to preserve productive replication fork progression (Bhowmick et al., 2023). Independent of transcription regulation, Integrator subunits also interact with the MCM2-7-helicase, suggesting alternative functions of chromatin factors in transcription-replication coordination. Finally, MYCN-dependent RNA exosome recruitment was demonstrated to be essential to avoid TRCs in MYCN-driven neuroblastoma (Papadopoulos et al., 2022).

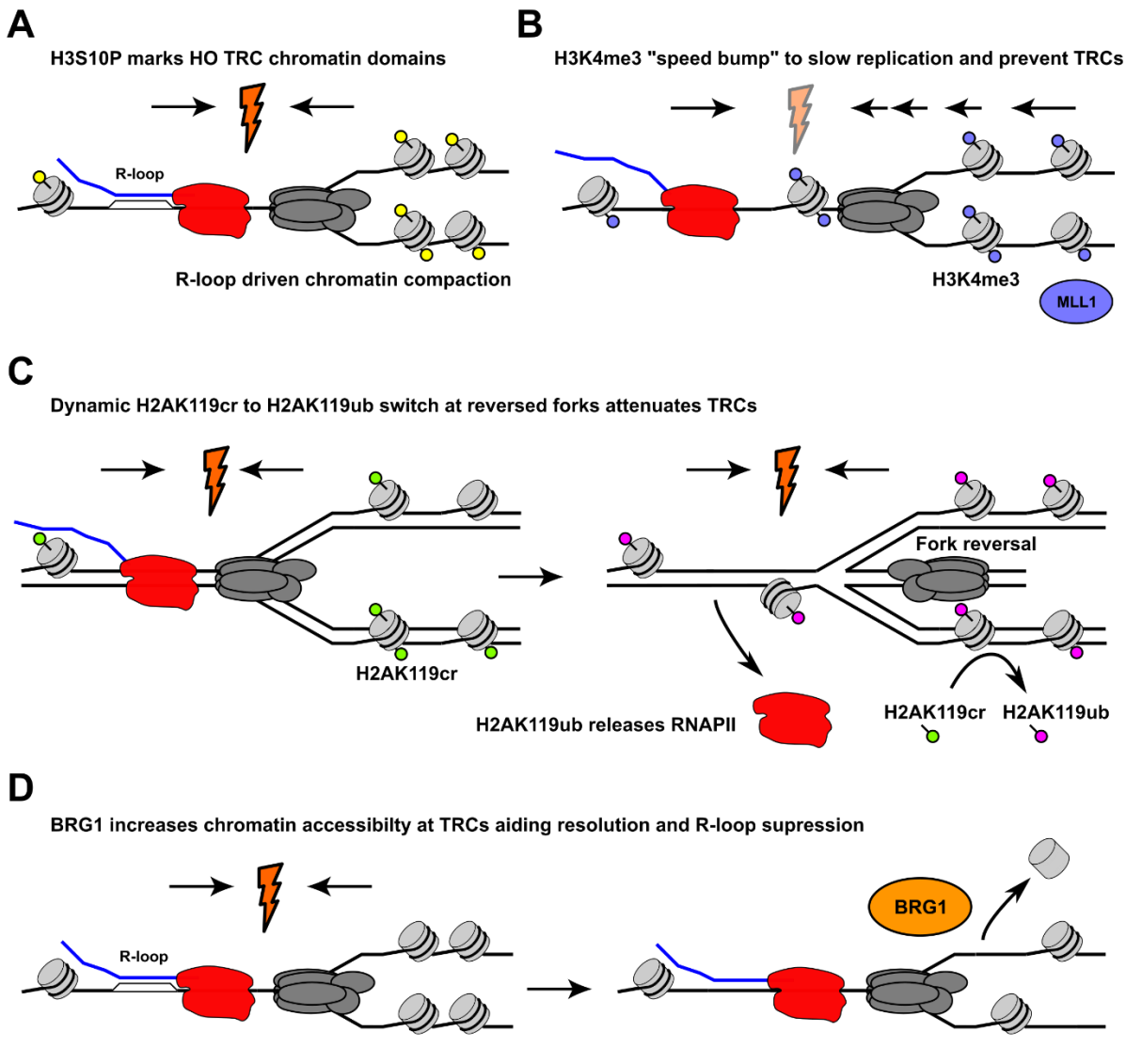

Figure 6: Chromatin processes impacting TRC occurrence and resolution

A) Evidence from yeast and human cells suggest that H3S10P is deposited at HO TRCs and contributes to R-loop driven compaction (Bayona-Feliu et al., 2023; García-Pichardo et al., 2017).

B) In yeast, MLL1-deposited H3K4me3 was shown to act as a “speed bump” for the replisome, thereby slowing it down and helping to avoid collisions with ongoing transcription (Chong et al., 2020).

C) A dynamic switch from H2AK119cr to H2AK119ub at reversed forks was demonstrated to attenuate TRCs. Specifically, H2AK119ub deposition releases RNAPII from chromatin, thereby removing the obstacle for the replisome (Hao et al., 2022).

D) In mammalian cells the SWI/SNF subunit BRG1 suppresses R-loop formation and remodels chromatin at TRC sites creating a more open chromatin environment permissive to resolution and repair processes (Bayona-Feliu et al., 2021).

3.4.5 TRC prevention and resolution

To circumvent harmful effects driven by TRCs, cells must quickly sense and resolve the conflicts. Numerous TRC resolution pathways working at different levels have been identified. Considering the complexity of mammalian replisomes and DNA replication impairments as detrimental threats to cell function, degradation and/or removal of RNA polymerases from chromatin appears as the simplest solution to resolve TRCs (**Fig. 7A**). Indeed, several RNAPII removal pathways in the context of DNA damage and TRCs have been described (Hobson et al., 2012; Wilson, Harreman, & Svejstrup, 2013). Importantly, targeting the phosphorylation state of RNAPII can aid TRC resolution. Recent work showed that the PP1 nuclear targeting subunit (PNUTS) together with WD repeat-containing protein 82 (WDR82) reduces replication stress by removing Ser5P on the RNAPII CTD, thereby promoting RNAP II removal and preventing TRCs (**Fig. 7A**) (Landsverk et al., 2020; Jeong Heon Lee et al., 2010). Moreover, RNAPII stalling at sites of DNA lesions has been thoroughly studied in the context of transcription-coupled nucleotide excision repair (TC-NER). TC-NER entails two sequential mono- and poly-ubiquitination steps, followed by removal of the stalled RNAPII by proteasomal degradation (**Fig. 7A**), thus giving access to the underlying DNA lesion (Hobson et al., 2012; Wilson, Harreman, & Svejstrup, 2013; Wilson, Harreman, Taschner, et al., 2013). While it remains unclear if TC-NER can directly engage in TRC resolution, it serves as a crucial pathway to remove trapped RNAPII from chromatin, thereby avoiding TRC occurrence in S-phase. Beyond RNAPII degradation, two recent studies provide evidence that RNAPII can be transiently disengaged from chromatin and retained in proximity of the replisome, before reengaging onto chromatin shortly after replisome passage (Bruno et al., 2024; Fenstermaker et al., 2023). Although the loss of a RNA transcript is generally less consequential than DNA replication stress, RNAPII removal as the only resolution pathway would lead to constant transcription abortion at CFS genes that require more than one cell cycle to transcribe (see 3.4.1), providing a rationale for the existence of multiple alternative pathways.

As mentioned above, R-loops frequently coincide with TRCs, increase their lifetime, and impair both transcription and replication resumption. Consequently, efficient R-loop removal is an important step in TRC resolution (**Fig. 7A**). The most well characterized R-loop processing factors are the RNase H enzymes, RNase H1 and RNase H2 in eukaryotic cells. RNase H enzymes have been characterized as specific endonucleases cleaving the RNA moiety in RNA:DNA hybrids (Hyjek et al., 2019). Whereas RNase H1 primarily functions in the removal of R-loops, RNase H2 is involved in R-loop removal and ribonucleotide excision repair (Sparks et al., 2012; H. Zhao et al., 2018). Loss of either RNase H1 or H2 are embryonic lethal in mice and mutations in RNase H2

can lead to severe development defects such as the Aicardi-Goutières syndrome (Cristini et al., 2022; Reijns et al., 2012). Recent work has also identified RNase H resistant R-loops throughout the genome (Crossley et al., 2020), suggesting that RNase H action alone is insufficient for full R-loop removal at a subset of genomic loci. Intriguingly, TC-NER machinery nucleases XPG and XPF have also been demonstrated to recognize and cleave R-loops as non-canonical targets (Sollier et al., 2014). Alternative to nucleolytic cleavage, R-loops can be removed via unwinding through numerous RNA:DNA helicases including SETX, Bloom (BLM), AQR, and Fanconi anemia complementation group M (FANCM) (Alzu et al., 2012; Chang et al., 2017; Silva et al., 2019; Yang et al., 2023). Particularly, SETX and the yeast homolog Sen1 have been characterized to travel with the replisome and act as an essential protein to avoid replication fork impairments by HO-TRCs and dormant replication origin activation at R-loop prone loci (Aiello et al., 2022). Furthermore, recent studies identified numerous DEAD-box RNA helicases as unwinders of RNA:DNA hybrids including DHX9 (Ren et al., 2024), DDX1 (de Amorim et al., 2024), DDX5 (Mersaoui et al., 2019), DDX17 (Polenkowski et al., 2023), DDX18 (W. L. Lin et al., 2022), DDX19 (Hodroj et al., 2017), DDX21 (Song et al., 2017), DDX23 (Sridhara et al., 2017), DDX39B (Pérez-Calero et al., 2020), DDX41 (Mosler et al., 2021), DDX43 (Talwar et al., 2017), and DDX47 (Marchena-Cruz et al., 2023). While these helicases have demonstrated to contain RNA:DNA hybrid unwinding activity either *in vitro* or *in vivo*, the regulatory complexity and specificity of the individual enzymes is not yet understood. With respect to TRC-specific R-loop processing, the DNA interstrand crosslink repair Fanconi anemia group D2 protein (FANCD2) was shown to be required for R-loop removal at CFSs (Madireddy et al., 2016). FANCD2 acts together with the BLM helicase and the homologous recombination factor breast cancer type 2 susceptibility protein (BRCA2) in an early response to CD-TRCs and R-loops (Shao et al., 2020). Although the exact mechanism of FANCD2 action at TRCs still remains unclear, the interactions with BLM and DDX47 strongly imply a function in R-loop removal that is further bridged to TRC resolution by the ATR-dependency of their recruitment (Okamoto et al., 2019). Similar to FANCD2, PCNA unloader ATPase family AAA domain-containing protein 5 (ATAD5) is needed to prevent replication fork slow down and TRCs via the recruitment of DEAD-box helicases for R-loop resolution (Sangin Kim et al., 2020). ATAD5 was also proposed to mitigate PCNA accumulation behind the fork to reduce additional interference, thus emphasizing a tight regulatory interplay between R-loop processing and replication fork remodeling (Sangin Kim et al., 2020). Collectively, the existence of numerous independent and redundant R-loop processing pathways indicates that R-loop removal, despite the physiological functions of R-loops in certain processes, is essential to prevent transcription induced genomic instability and replication stress including TRCs.

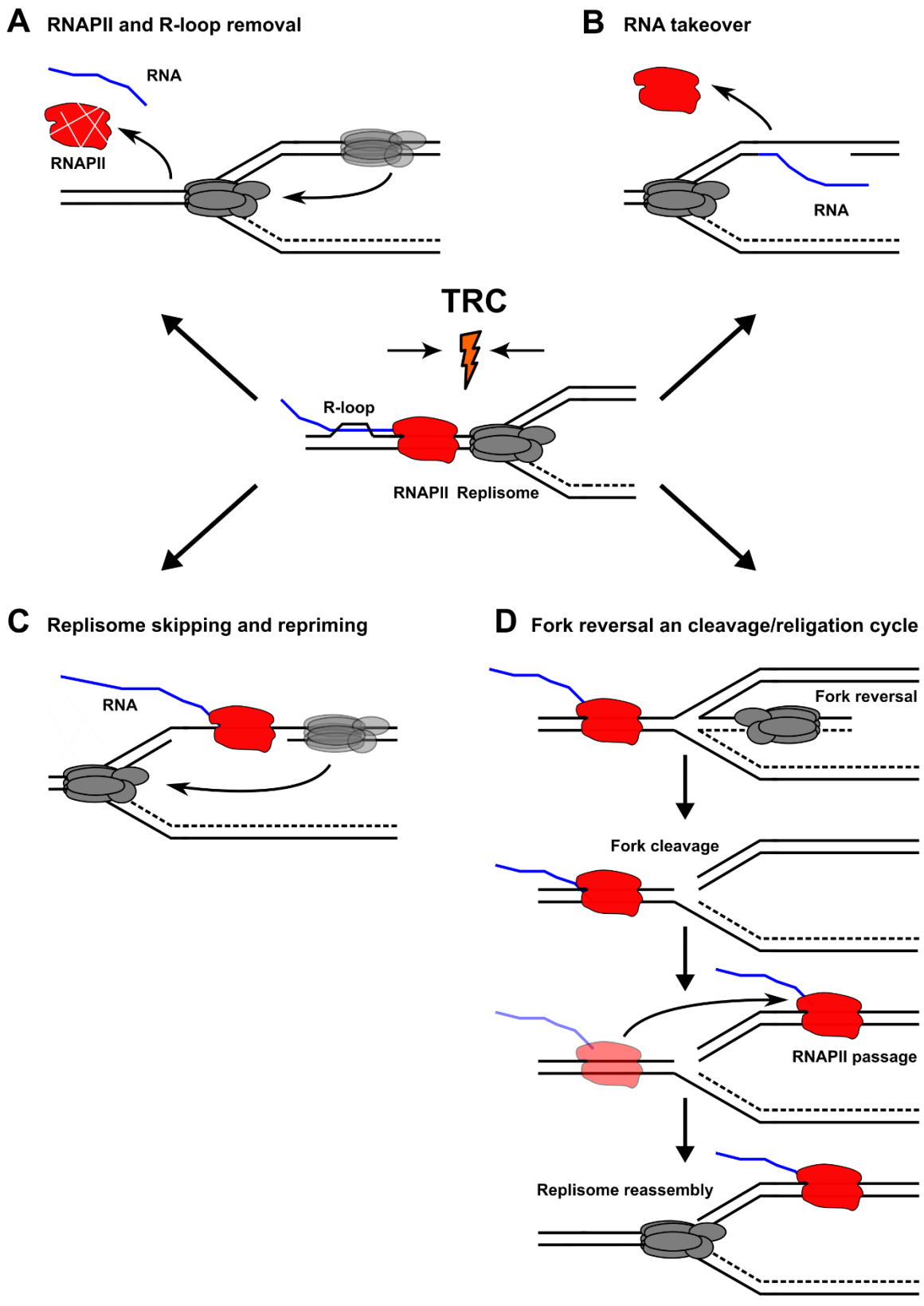
A RNAPII and R-loop removal
RNA
RNAPII
B RNA takeover
RNA
TRC
R-loop
RNAPII Replisome
C Replisome skipping and repriming
RNA
D Fork reversal an cleavage/religation cycle
Fork reversal
Fork cleavage
RNAPII passage
Replisome reassembly

Figure 7: Overview of several TRC resolution pathways

A) The cell employs numerous pathways to remove or degrade RNAPII and remove associated R-loops from chromatin, allowing the continuation of DNA synthesis.
B) In CD conflicts, the replisome can displace the RNAP and use the hybridized RNA as a primer to reinitiate DNA replication.
C) When facing a transcription block, the replisome may skip the RNAP and reprime downstream. Depending on the affected strand, PRIM1 primase or PRIMPOL1 polymerase are used for lagging or leading strand blocks, respectively.
D) In the case of persistent RNAPII complexes, the replication fork can undergo fork reversal to stabilize the fork and initiate repair mechanisms. Reversed forks were shown to undergo a cycle of fork cleavage and re-ligation. This allows the passage of RNAPII and requires replisome reassembly downstream of the conflict.

In case of severe replication fork impairments, RNA polymerase removal and R-loop processing alone might not suffice to resolve a TRC, thus demanding replication fork remodeling pathways to solve the conflict. Evidence, predominantly from TRC studies in *Escherichia coli* (*E. coli)*, suggests that the replisome can bypass the RNAP complex by repriming downstream of the obstruction (**Fig. 7C**) (Brüning & Marians, 2020; Gómez-González & Aguilera, 2019). Repriming requires either the newly made RNA to serve as a primer for replication restart, as demonstrated for co-directional TRCs, or the synthesis of a new Okazaki fragment (**Fig. 7B**) (Conti & Smogorzewska, 2020; Gómez-González & Aguilera, 2019). While the synthesis of a new Okazaki fragment is a convenient solution based on the inherently discontinuous synthesis of the lagging strand, leading strand repriming is more complex and needs a specialized DNA polymerase. In human cells, the PRIMPOL enzyme contains primase and DNA polymerase activities to enable leading strand skipping and repriming (García-Gómez et al., 2013; Mourón et al., 2013). Interestingly, PRIMPOL was shown to be actively recruited to stalled forks upon replication stress and aid in repriming downstream of DNA secondary structures such as R-loop and G4, which are frequently associated with TRCs (Straka et al., 2024; Šviković et al., 2019). For remodeling of the replication fork, it has been demonstrated that the fork can be transiently cleaved and subsequently re-ligated, allowing RNA polymerases to restart transcription and progress past the replication fork (**Fig. 7D**). This multistep process depends on several key proteins, including the ATP-dependent DNA helicases Q1 and Q5 (RECQ1 and RECQ5), the structure-specific endonuclease subunit SLX4 (SLX4), the crossover junction endonucleases MUS81 and EME1 (MUS81/EME1), DNA repair protein RAD52 homolog (RAD52), DNA ligase IV, the DNA polymerase δ subunit POLD3, and the transcription elongation factor ELL for transcriptional recovery (Chappidi et al., 2020). Finally, in many of these pathways, the replisome may undergo

an intermediate remodeling process called fork reversal. This mechanism stabilizes and protects the replication fork (Stoy et al., 2023), providing the cell with additional time to resolve the conflict.

3.5 Aims of this study

Aim 1: Establishment of a genomically integrated TRC reporter system and characterization of TRC-induced chromatin changes

As described in the introduction, research on TRCs is currently limited by only very few methods and cellular systems available to study the mechanisms of TRC resolution *in vivo*. Therefore, one goal of this thesis was to develop a genomically integrated TRC reporter system, which would enable me to specifically induce TRCs in the mammalian genome and analyze their downstream consequences with unprecedented resolution. With a system like this at hand, it would be possible to further expand on the limited understanding of TRCs and R-loops as disruptors of the chromatin landscape. Consequently, the main objective of this study is to determine local and global characteristics of TRC and R-loop-driven chromatin alterations with a special focus on nucleosome dynamics, histone modifications, and DNA replication impairments.

Aim 2 Systematic comparison of distinct oncogene activation systems and their consequences on TRC levels and genomic instability

As previously highlighted, oncogenes are potent disruptors of transcription and DNA replication processes, thereby leading to elevated TRC levels in some cancer cell types. While certain oncogenes have been directly implicated in TRC induction, the evidence for other cellular models suggests a potential involvement. Nevertheless, a comprehensive analysis of different oncogenes as inducers of TRCs with consistent methodology has not yet been conducted. Thus, another goal of this work was to systematically characterize and compare the ability of different oncogenes to induce TRCs and drive transcription-associated genomic instability.

4. Results

4.1 RNA:DNA hybrids are resistant to nucleosome assembly and form a nucleosome-depleted chromatin structure *in vitro*

As the impact of TRCs and R-loops on chromatin remains insufficiently understood, my first aim was to address to what extent these conflicts and associated secondary structures might disrupt nucleosome organization and thereby drive genome instability. Due to the high complexity of the chromatin polymer and nucleosome organization in cells, I initially focused on an *in vitro* assay to probe the interplay of R-loops and their potential for nucleosome assembly. To this end, an RNA:DNA hybrid was reconstituted by *in vitro* transcription of the R-loop-prone mAIRN DNA sequence and subsequent hybridization of the DNA strand with its complementary RNA. Expectedly, the RNA:DNA intermediate exhibited reduced mobility in agarose gel electrophoresis, which is likely a result of altered topology of the two strands. This shift could be rescued by RNase H treatment, thereby confirming specific RNA:DNA hybrid reconstitution (**Fig. 8A**). To address the capability of R-loops to incorporate nucleosomes, a competition assay for nucleosome formation was performed between mAIRN dsDNA and mAIRN RNA:DNA hybrids present in the same reaction tube. Increasing addition of the four core histone proteins resulted in stable nucleosome formation for dsDNA as seen by reduced mobility in the gel. Strikingly, RNA:DNA hybrids were unable to form nucleosomes despite increasing concentrations of supplied histone proteins. Even at histone protein concentrations that induced precipitation of dsDNA, the signal intensity of the mAIRN RNA:DNA hybrids remained constant, indicating no nucleosome formation on the hybrids (**Fig. 8B**). Strikingly, the results suggest that mAIRN RNA:DNA hybrids are incompatible with nucleosome formation *in vitro*. This observation might be due to stronger rigidity of the RNA strand and a previously suggested intermediate state between an A-DNA and B-DNA conformation for RNA:DNA hybrids (J. H. Liu et al., 2019).

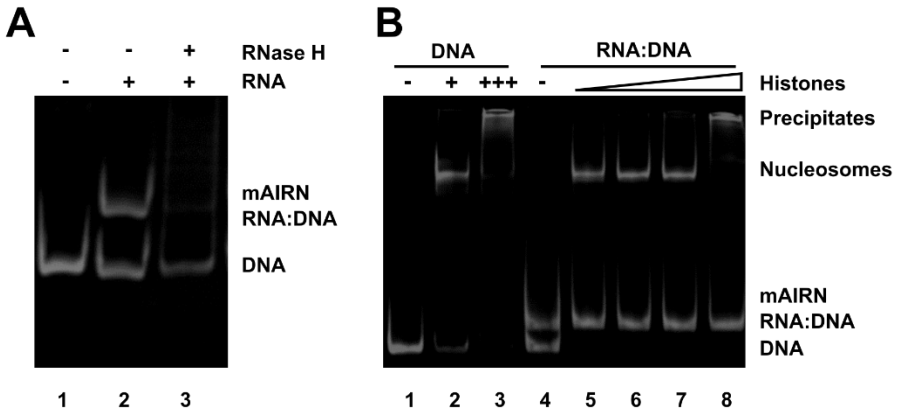

Figure 8: RNA:DNA hybrid structures are incompatible with nucleosome formation *in vitro*

A) Native 6 % polyacrylamide gel showing the reconstituted RNA:DNA hybrids (lane 2) compared to dsDNA alone (lane 1). RNA:DNA hybrids were treated with RNase H (5 U RNaseH, 30 min at 37 °C) (lane 3).

B) Native 6 % polyacrylamide gel of the *in vitro* nucleosome assembly assay on mAIRN dsDNA alone (lanes 1-3) versus mAIRN RNA:DNA hybrids in competition with mAIRN dsDNA (lanes 4-8) using increasing amounts of histone octamers, as indicated. The positions of the DNA, RNA:DNA hybrid, nucleosomes, and DNA-histone precipitates are marked on the right side of the gel picture.

To provide complementary evidence in an *in vivo* system, I took advantage of a previously established episomal reporter system for R-loop and TRC induction in human cells (Hamperl et al., 2017). This construct contains a doxycycline (DOX)-inducible (Tet-ON), R-loop-forming mAIRN transcription unit as well as the unidirectional origin of replication (oriP) from the Epstein-Barr virus (EBV) (Hodin et al., 2013; Moriyama et al., 2012). Depending on the positioning of the reporter unit towards oriP, either HO or CD TRCs can be created in an inducible manner. Quantification of R-loop levels using DNA-RNA immunoprecipitation (DRIP) and quantitative PCR (qPCR) readout indicated that HO TRC induction favors stable R-loop formation, whereas in CD TRC orientation R-loop levels are reduced (Hamperl et al., 2017). Taking into consideration this orientation-dependent bias of R-loop formation on the HO episomes, I next asked how this difference would affect local nucleosome organization. Thus, Micrococcal nuclease (MNase) digestion was performed followed by Southern Blot readout on HO and CD TRC episomes with and without DOX-dependent transcriptional induction of the mAIRN reporter sequence. Without transcriptional induction, both HO and CD episomes showed digestion products corresponding to nucleosome monomers, dimers, and trimers, demonstrating that the plasmids are chromatinized in cells (**Fig. 9**). Crucially, transcriptional induction on the HO plasmid but not on the CD construct induced a drastic loss of nucleosome levels and/or positioning as indicated by reduced

nucleosome band intensity and poorly defined banding pattern (**Fig. 9**). These findings suggest that particularly HO TRCs but not CD TRCs interfere with nucleosome architecture at conflict sites.

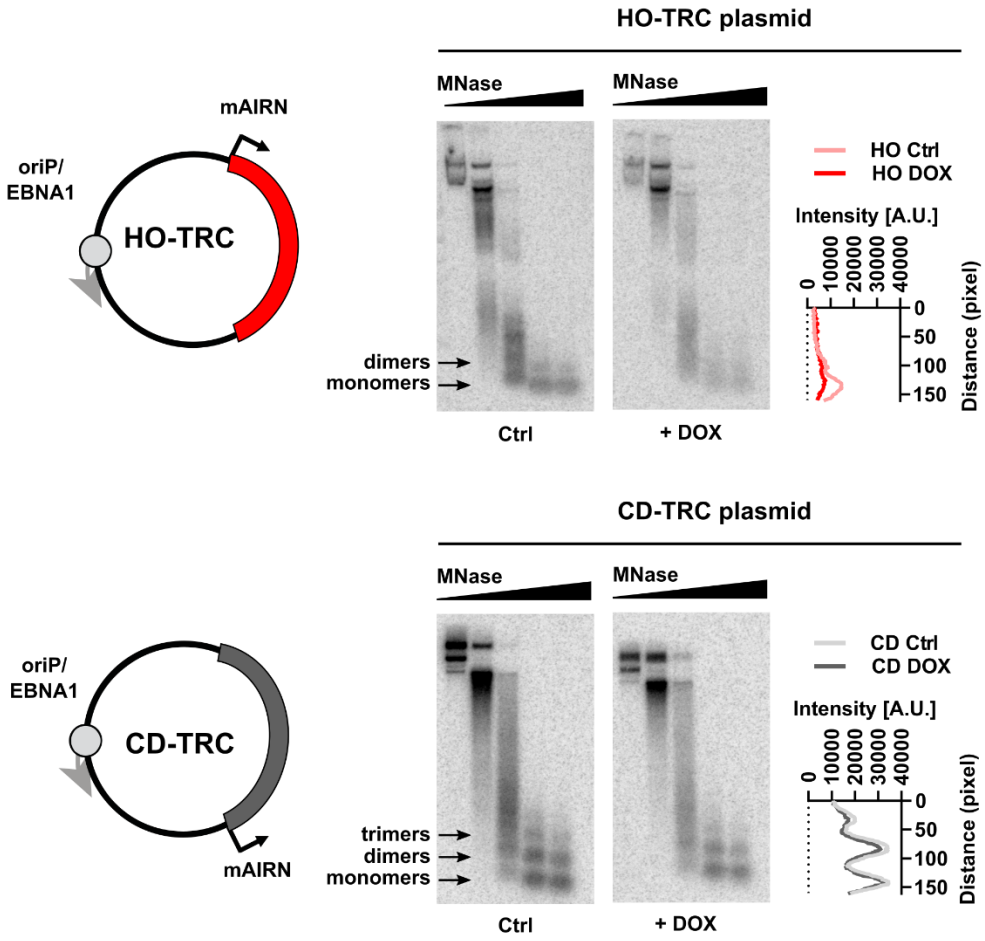

Figure 9: MNase assay reveals that R-loops at HO TRC episomes reduce nucleosome occupancy

Plasmids contain the mAIRN reporter sequence in either HO or CD orientation towards the unidirectional origin of replication (oriP/EBNA1). Southern blot images of mAIRN HO or CD TRC plasmids following treatment with 0 or 1 µg/ml DOX for 24 h. Samples were treated with increasing concentrations of MNase (0, 2.5, 25, 100, or 250 gel units). Arrows show nucleosome monomers, dimers, and trimers. Quantification of the Southern blot signal from nucleosome bands in arbitrary units (A.U.) for the 250 gel units MNase lane is shown next to the blots.

To further validate these findings, an independent chromatin immunoprecipitation (ChIP) against canonical histone H3 followed by qPCR analysis was performed. Confirming the MNase-assay results, episomes with induced HO conflicts showed a loss of H3 levels, while CD conflicts did

not disrupt H3 occupancy (**Fig. 10A**). As a control, I also analyzed RNAPII occupancy at the reporter sequences using ChIP, which confirmed identical RNAPII recruitment to HO and CD episomes upon DOX-dependent transcriptional induction (**Fig. 10B**). Moreover, unrelated genomic locus ACTB as well as an intergenic control site did not show significantly altered H3 or increased RNAPII levels upon DOX treatment. Collectively, the data from both *in vitro* and plasmid reporter systems supports the concept that the increased stability of mAIRN RNA:DNA hybrids at HO TRCs can displace nucleosomes and thereby cause a local chromatin disruption, potentially associated with elevated susceptibility to DNA damage.

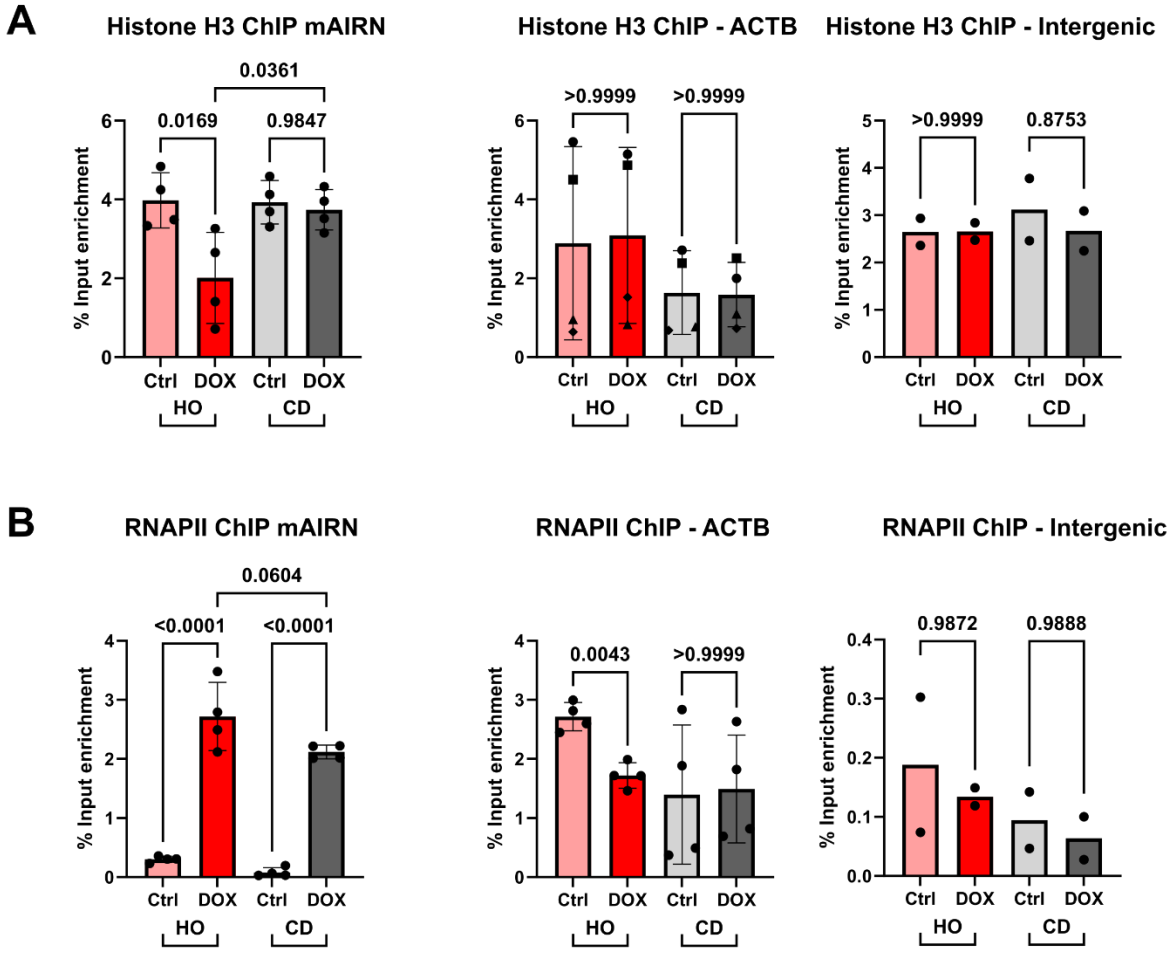

Figure 10: ChIP-qPCR analysis confirms that R-loops at HO TRC episomes reduce nucleosome occupancy

A) ChIP-qPCR analysis for histone H3 at the mAIRN gene (n=4) or ACTB (n=4) and Intergenic control loci (n=2) in HEK293 containing plasmids for either HO or CD TRC induction. Cells were treated with 0 or 1 µg/mL DOX for 24 h. Error bars show SD. Ordinary one-way analysis of variance (ANOVA) with Tukey's multiple comparison test.

B) Same as in **A)** for ChIP-qPCR analysis of RNAPII levels.

4.2 Chromosomal integration of the R-loop forming mAIRN gene increases cellular TRC levels

Although the plasmid-based reporter system has yielded novel insight into the relationship of R-loop formation, TRC directionality, and chromatin disruption, it remains debatable to what extent the episomes are chromatinized in a cell. Given this limitation and the inherent artificiality of plasmid reporters, I sought to construct an advanced TRC reporter system that would be able to more accurately reflect the dynamics of TRCs in the native chromatin environment. First, the previously described mAIRN sequence together with its DOX-inducible promoter was cloned into a Sleeping Beauty transposase vector. Sleeping Beauty transposase-based genomic integration allowed us to integrate the reporter construct into the genome of the U-2 OS osteosarcoma cell line (**Fig. 11**). Through this approach, multiple insertion sites at different locations are expected, thus positioning the reporter sequence in diverse chromatin contexts. In contrast to the episomal plasmid system (Hamperl et al., 2017), the unidirectional origin oriP was omitted from the integration construct. This origin had previously been shown to be late replicating in a plasmid context and thus, would likely not provide sufficient levels of replication initiation in a chromosomal context necessary for TRC induction (Moriyama et al., 2012). Instead, endogenous origins with early replication timing and high efficiency are expected to replicate the integration sequence before oriP activation. This scenario is particularly likely as Sleeping Beauty transposase preferentially targets genic and early replicating regions of the genome (Kowarz et al., 2015; G. Liu et al., 2005). Despite the lack of oriP, I hypothesized that the transcriptional activation of the strongly R-loop forming mAIRN sequence would result in profound stalling of transcription complexes as previously described (Tous & Aguilera, 2007). These impaired RNAPII complexes are a potent genomic obstacle, which in turn would give rise to collisions with incoming replication forks and thus lead to TRCs (**Fig. 11**).

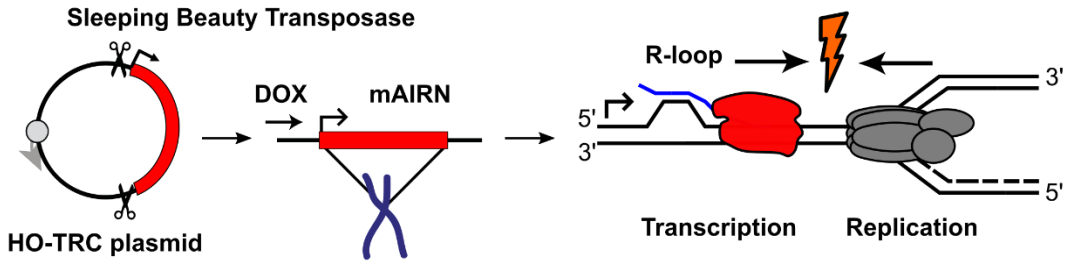

Figure 11: Schematic showing the generation of the chromosomal TRC reporter cell lines

Genomic integration of the R-loop forming mAIRN sequence of the episomal reporter plasmids was performed using Sleeping Beauty transposase. Inducible R-loop formation at the mAIRN locus with DOX stalls RNAPII progression and creates a potent obstacle for collisions with replication forks.

Following the co-transfection of a U-2 OS Tet-ON cell line with the mAIRN integration plasmid and a vector expressing Sleeping Beauty transposase, we generated monoclonal U-2 OS cell lines containing the insertions in varying genomic positions. For initial characterization of the resulting cell lines, we performed copy number analysis by qPCR on six clones that showed robust growth under antibiotic selection, thus indicating successful integration events. The analyzed cell lines harbored between one to five integrations (**Fig. 12A**). Interestingly, Clone#12 showed the highest copy number and was further characterized by whole genome sequencing (WGS), which unveiled five integrations across four chromosomes (**Fig. 12B, Table S1**). WGS results were confirmed for an exemplary site on chromosome 10 using genotyping PCR followed by Sanger sequencing (**Fig. 12C, D)**. To test the functionality of our reporter cell line, I conducted RT-qPCR for analysis of mAIRN gene expression. In contrast to the parental U-2 OS cell line without integrations, mAIRN expression was induced 50 to 100-fold after 4 h and 24 h of DOX treatment, respectively (**Fig. 12E**). Clone#12 was consequently chosen as the main working model for all following experiments unless stated otherwise.

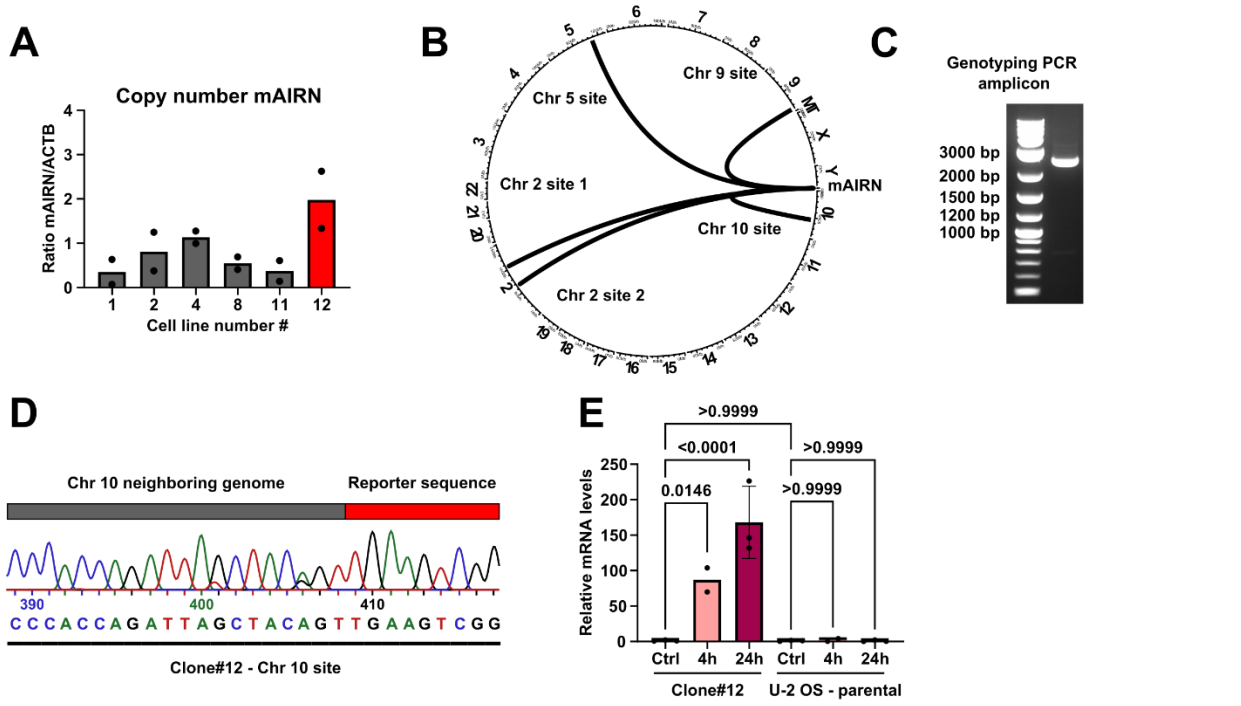

Figure 12: Characterization of the Clone#12 chromosomal TRC reporter integrations

A) Copy number analysis of several monoclonal cell lines generated by Sleeping Beauty integration of the mAIRN reporter construct (n=2). Copy number compared to endogenous ACTB locus.
B) Circus plot visualization of the position of the five integration sites of the mAIRN reporter construct in the monoclonal U-2 OS cell line Clone#12.
C) Agarose gel showing the genotyping PCR amplicon for the Chr 10 integration site using primers mAIRN#1_REV and Intsite chr10 FWD2. Fragment size (bp) of DNA ladder shown for selected bands.
D) The exemplary site on Chr 10 was confirmed using Sanger sequencing using the PCR product from **C)**. The obtained sequencing read shows a short region with the beginning of the reporter sequence and the neighboring genome of Chr 10.
E) RT-qPCR analysis of mAIRN RNA expression with primer pair mAIRN#1 in Clone#12 or parental U-2 OS cells exposed to 0 or 1 µg/mL DOX for 4 h or 24 h (n=3). Error bars indicate mean values with standard deviations (SD). Welch ANOVA with Tukey's multiple comparison test.

To test the functionality of our newly developed reporter system, I first aimed to verify whether transcription induction of the R-loop prone gene would indeed cause elevated TRC levels in this cell line. Here, I performed PLA with antibodies targeting the actively elongating form of RNAPII (RNAPII Ser2P) and DNA clamp and replication processivity factor PCNA as a part of the replication fork complex. Crucially, elevated numbers of TRC-PLA foci were detectable in S-phase cells as soon as 4 h after DOX induction and showed significant increases after 8 h and 12 h of DOX exposure (**Fig. 13A**). Extended treatment of 24 h induced a moderate decline in TRC-PLA foci, likely a result of compensatory mechanisms being activated to alleviate the high TRC burden (**Fig. 13A**). Comparably, Clone#1, another independent cell line generated with the Sleeping Beauty approach but harboring the mAIRN reporter in different genomic locations (see **Fig. 11A**), also contained a higher number of TRC upon 4 h DOX treatment (**Fig. 13B**). In contrast, parental U-2 OS cells did not show a DOX-dependent increase of TRC-PLA foci (**Fig. 13C**), thereby demonstrating that the TRC induction is indeed a specific result of the transcriptional activation of the mAIRN reporter sequence.

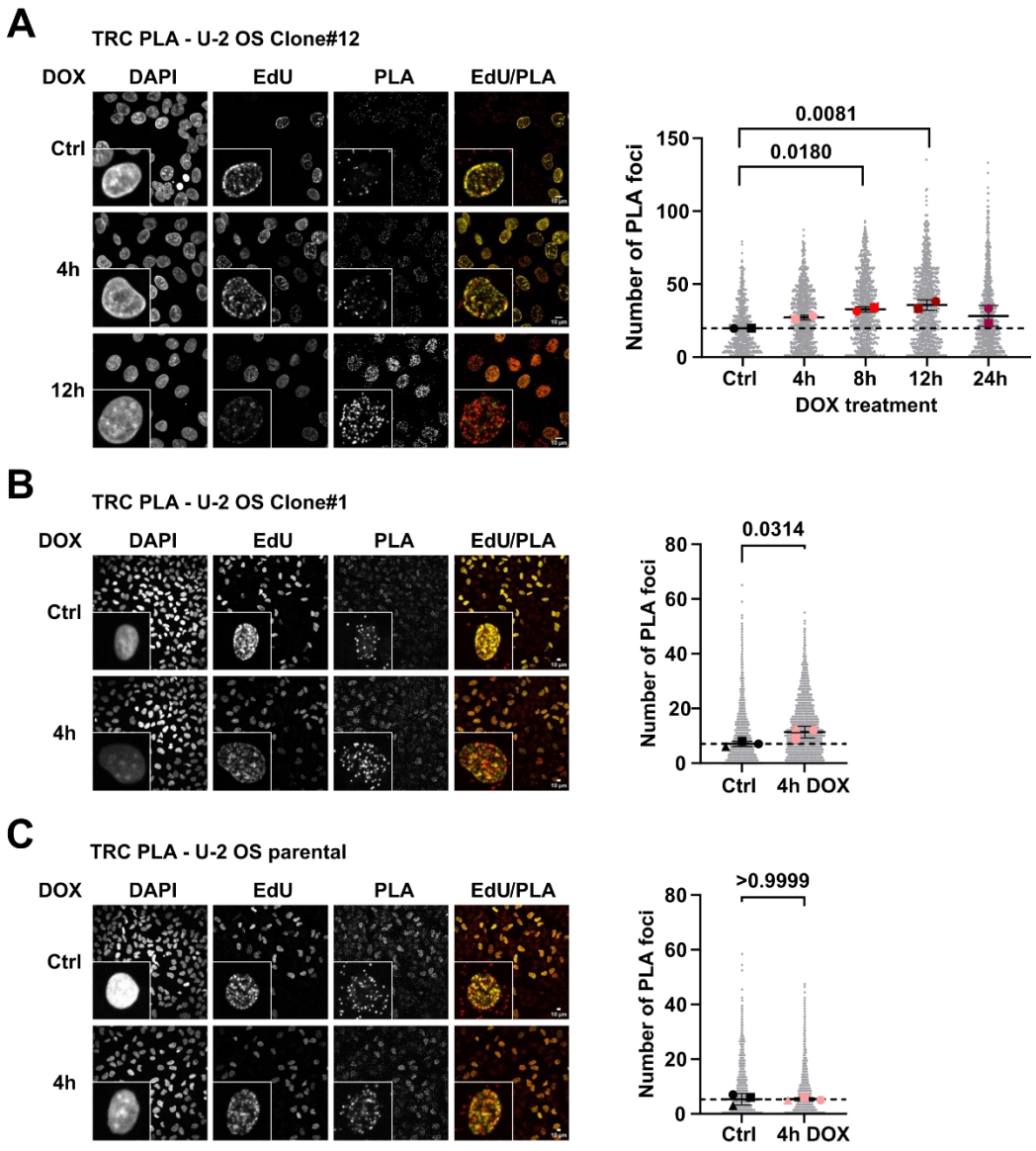

Figure 13: Induction of genomically integrated mAIRN reporters causes an increased TRC burden

A) Representative images of TRC PLA assays in Clone#12 with RNAPII Ser2P and PCNA antibodies (Ctrl, 4 h, and 12 h time points). 5-Ethynyl-2′-deoxyuridine (EdU) click-it staining was performed to label S-phase cells. Cells were treated with 0 or 1 µg/mL DOX for TRC induction for the indicated time points. Scale bar 10 µm. Quantification of TRC PLA foci number in S-phase cells (n=2, mean foci values per biological replicate as colored dots). Bars indicate mean values with standard deviations (SD). Ordinary one-way ANOVA with Tukey's multiple comparison test.

B) Same TRC PLA readout as in **A)** for Clone#1 cells treated with DOX for 4 h. Quantification of TRC PLA foci number in S-phase cells (n=3). Student's t-test.

C) Same TRC PLA readout as in **A)** for parental U 2-OS cells treated with DOX for 4 h. Quantification of TRC PLA foci number in S-phase cells (n=3). Student's t-test.

4.3 Chromosomal reporter-driven TRCs impose a local and global replication stress response

Previous work has shown that TRCs can induce genome instability in mammalian cells (Chappidi et al., 2020; Hamperl et al., 2017; Helmrich et al., 2011). Therefore, I wondered if the TRCs created by our mAIRN-driven genomic reporter can generate replication stress and associated DNA damage. Initially, I focused my investigations on FANCD2 as a marker of stalled and damaged replication forks that had previously been connected to R-loops and TRCs (Bayona-Feliu et al., 2021, 2023; Okamoto et al., 2019). Remarkably, ChIP-qPCR analysis for FANCD2 revealed significantly increased binding at the mAIRN reporter sequence upon DOX treatment for TRC induction, thus providing direct evidence for locus-specific replication stress response (**Fig. 14A**). Importantly, we also benchmarked FANCD2 ChIP by treatment with the DNA polymerase inhibitor Aphidicolin (APH), which had previously been shown to increase FANCD2 levels at CFSs (Okamoto et al., 2018). Indeed, FANCD2 occupancy increased at two CFS genes *NRG3* and *WWOX* upon APH treatment (**Fig. 14B**). In contrast, DOX treatment did not alter FANCD2 levels at CFS genes (**Fig. 14B**), thereby supporting a DOX-dependent replication stress increase only at the TRC reporter sites.

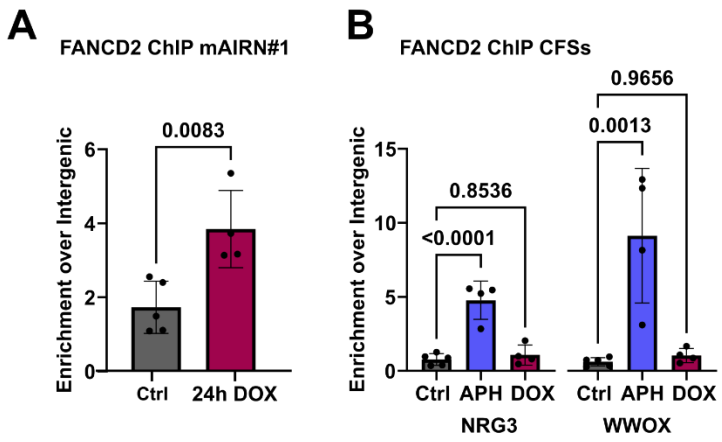

Figure 14: TRC induction at the mAIRN reporter increases occupancy of FANCD2

A) ChIP-qPCR analysis showing FANCD2 levels in asynchronous cells at the TRC reporter sequence using the mAIRN#1 primer pair. Cells were treated 0 or 1 µg/mL DOX for 24 h (n≥4). Error bars indicate SD. Student's T-test.

B) ChIP-qPCR analysis showing FANCD2 levels at CFS genes *WWOX* and *NRG3*. Cells were treated 0 or 1 µg/mL DOX for 24 h (n≥4). For CFS analysis, cells were additionally treated with dimethyl sulfoxide (DMSO) or 0.4 µM APH for 24 h. Ordinary one-way ANOVA with Tukey's multiple comparison test.

To further validate whether this phenotype is driven by the R-loop formation at the mAIRN sequence, I also constructed a control cell line containing a DOX-inducible ECFP sequence in analogy to the original plasmid system (Hamperl et al., 2017). In contrast to mAIRN, ECFP expression does not form R-loops and would be expected to induce little to no genomic instability. Here, I selected a monoclonal ECFP cell line (Clone#2) with similar number of integration sites and transcriptional inducibility, as demonstrated by copy number and expression analysis by qPCR and RT-qPCR, respectively (**Fig. 15A, B**). In comparison to the mAIRN Clone#12 cell line (**Fig. 12A**) the ECFP Clone#2 cells displayed a comparable number of 4-5 integration sites and ~850-fold increase in expression level upon DOX treatment, which is about 5 times higher than in the mAIRN construct (~160-fold, see **Fig. 12D**). Crucially, this high level of ECFP expression did not increase local FANCD2 occupancy as determined by FANCD2 ChIP in both cell lines (**Fig. 15C**), demonstrating that the observed genomic instability phenotype is specific to the R-loop forming mAIRN TRC reporter loci.

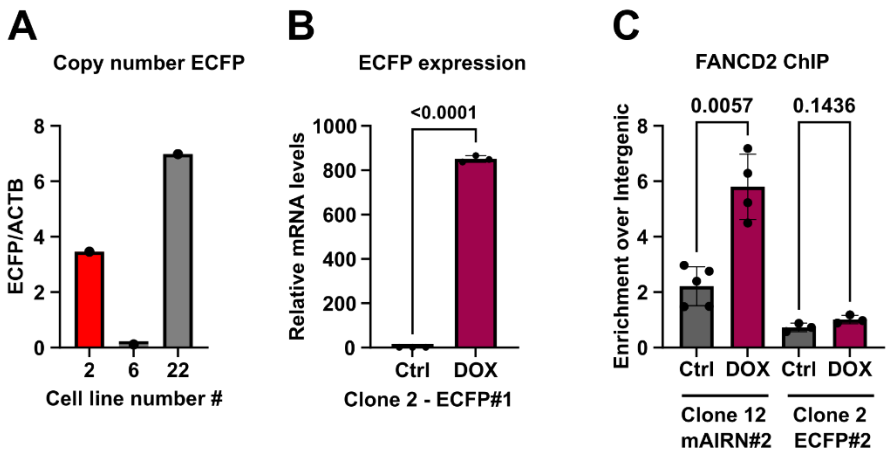

Figure 15: A reporter cell line with inducible ECFP expression does not increase FANCD2 occupancy upon +DOX induction

A) Copy number analysis of several monoclonal cell lines generated by Sleeping Beauty integration of the ECFP reporter construct (n=1). The copy number of ECFP is compared to the endogenous ACTB locus using specific primers amplifying the two genomic loci.
B) RT-qPCR analysis of ECFP RNA expression with primer pair ECFP#1 in Clone#2 cells exposed to 0 or 1 µg/mL DOX for 24 h (n=3).
C) ChIP-qPCR analysis showing FANCD2 levels in asynchronous cells at the TRC reporter sequence using the mAIRN#2 (n≥4) and ECFP#2 (n=3) primer pairs. Error bars indicate mean values with SD. Student's T-test or Welch ANOVA with Tukey's multiple comparison test.

Aiming to more comprehensively explore the replication stress phenotype, I conducted immunofluorescence (IF) staining against FANCD2 combined with EdU-Click-it labeling for

analysis of S-phase cells. Remarkably, the number of FANCD2 foci associated with chromatin increased upon DOX induction in a time-dependent manner. The highest levels were observed after 24 h of induction in the time course (**Fig. 16**). Elevated FANCD2 foci were detectable in both S-phase and non-S-phase cells but showed at least twofold higher numbers in S-phase (**Fig. 16**), suggesting that replication interference via TRCs in S-phase cells is the primary driver of the observed phenotype.

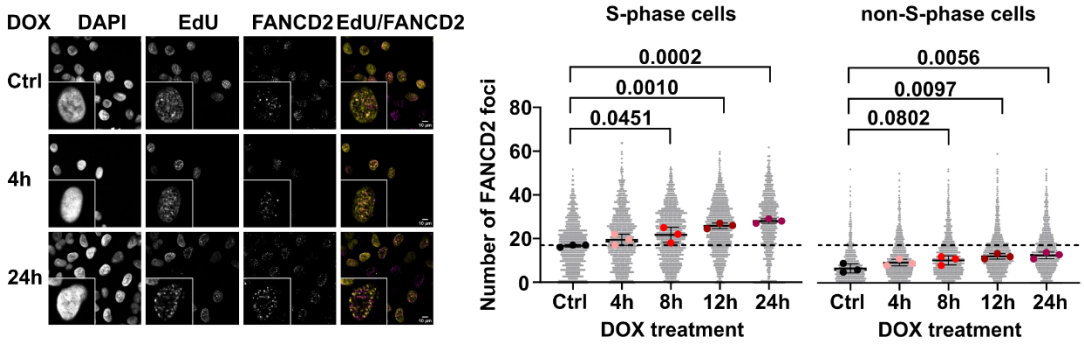

Figure 16: TRC reporter induction causes a global FANCD2 replication stress phenotype

Representative images of FANCD2 immunofluorescence staining (Ctrl, 4 h, and 24 h). EdU click-it staining was performed for labeling of S-phase cells. Cells were exposed to 0 or 1 µg/mL DOX for TRC induction. Scale bar 10 µm. Quantification of FANCD2 foci number in S-phase and non-S-phase cells (n=3). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

Considering the strong global FANCD2 response, I wondered if reporter induction would cause global transcription deregulation of RNAPII that could be indirectly responsible for the observed stress phenotype. Importantly, transcription levels determined by RNAPII Ser2P IF remained unchanged (**Fig. 17**), underscoring that the replication stress response is unlikely a result of globally altered transcription dynamics but rather a specific effect triggered by the induced mAIRN TRC reporter sites.

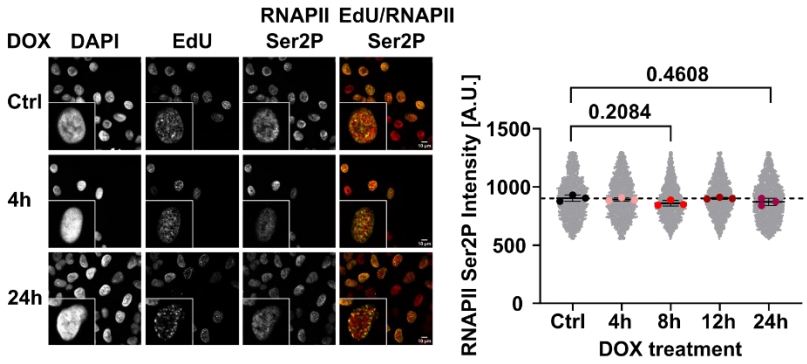

Figure 17: TRC reporter induction does not alter global transcription dynamics

Representative images of RNAPII Ser2P immunofluorescence staining on the same cells as in **Fig. 16** (Ctrl, 4 h, and 24 h). EdU click-it staining was performed for labeling of S-phase cells. Cells were exposed to 0 or 1 µg/mL DOX for TRC induction. Scale bar 10 µm (n=3). Measurement in absolute units (A.U.). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

As a complementary approach, I used Western blot analysis of whole cell lysates to quantify additional markers of the replication stress response (**Fig. 18A**). TRC induction with DOX led to globally elevated levels of DNA damage markers histone H2AX phosphorylation (γH2AX) and Serine 33 phosphorylated ssDNA binding protein RPA32 (RPA32pS33) when compared to the nuclear loading control ORC2. Interestingly, γH2AX levels increased earlier in the time course when compared to RPA32pS33, which is consistent with previous literature that RPA phosphorylation is needed to initiate DNA repair and dissolves γH2AX foci, thereby positioning RPA32pS33 downstream of H2AX phosphorylation (Anantha et al., 2007). Since γH2AX and RPA32pS33 are targets of the ATR kinase in the DNA damage response (Anantha et al., 2007; Toledo et al., 2013), I subsequentially exposed the cells to the specific ATR inhibitor VE-821 (ATRi). Importantly, increased ATR activity had previously been connected to DNA damage inflicted by HO TRCs in mammalian cells (Hamperl et al., 2017), further supporting a potential involvement in the observed global DNA damage response. Intriguingly, cells subjected to ATRi demonstrated sensitivity to TRC induction as seen by a significantly decreased fraction of S-phase cells after 24 h of treatment with DOX. In contrast, cells exposed to TRC induction by DOX alone (DMSO control) displayed an unaltered fraction of S-phase cells (**Fig. 18B**).

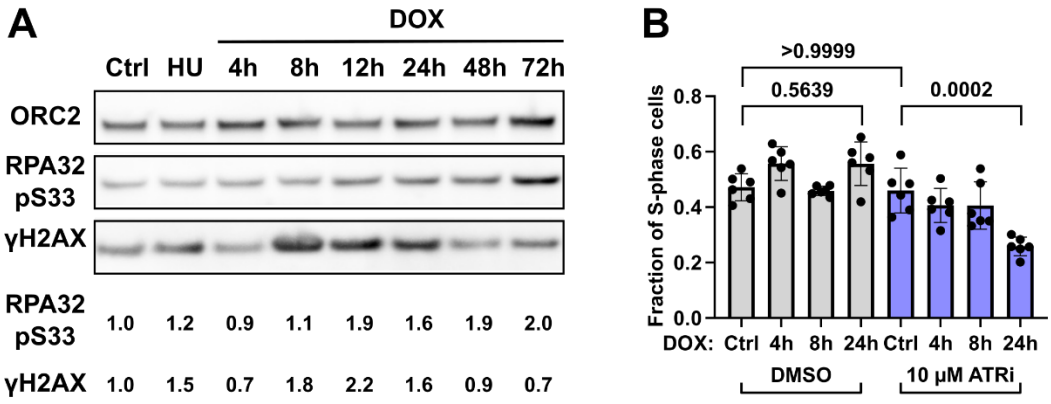

Figure 18: TRC induction in reporter cells causes elevated γH2AX and RPA32pS33 levels and delays S-phase in dependence on ATR activity

A) Representative Western blot analysis of DNA damage markers γH2AX and RPA32pS33 using cell lysates from mAIRN reporter cells exposed to 0 or 1 µg/µL DOX for the indicated time points (0-72 h).

Hydroxy urea (HU) treatment was used as a positive control for DNA damage induction. ORC2 served as the loading control. Quantifications of γH2AX and RPA32pS33 signals normalized to ORC2 are provided below.

B) Quantification of the fraction of cells in S-phase (n=6). Cells were treated with 0 or 1 µg/mL DOX for 4 h, 8 h, or 24 h. 10 µM ATR inhibitor VE-821 (ATRi) or DMSO were added respectively. Control (Ctrl) cells were exposed to ATRi or DMSO for 24h.

To validate these results and more thoroughly characterize cell growth impairments, we conducted proliferation analysis of cells challenged by DOX, ATRi inhibition or both combined over 7 days. TRC induction via DOX alone resulted in a modest but significant growth reduction, which was strongly exacerbated upon additional ATRi treatment with 1 µM VE-821 (**Fig. 19A, B**). Importantly, DOX and ATRi showed a synergistic behavior in halting proliferation that cannot be explained by just the individual effects of the treatments. Interestingly, the inhibition of the ATM kinase (1 µM KU-60019), which has been shown to work in DNA double-strand break repair and becomes activated upon CD TRCs in mammalian cells (Hamperl et al., 2017), also further reduced cell proliferation in DOX-treated cells but did not demonstrate obvious synergistic behavior as seen for ATRi (**Fig. 19A, B**). Higher doses of 10 µM ATRi or ATMi completely abolished cell growth independent of DOX induction. Taken together, our results highlight that TRC induction by the reporter system creates DNA damage and fork stalling that in turn cause proliferation defects in an ATR- and ATM-dependent manner. Further, our data indicates that an elevated TRC burden can be tolerated by the cells, but only if the DNA damage response remains fully functional.

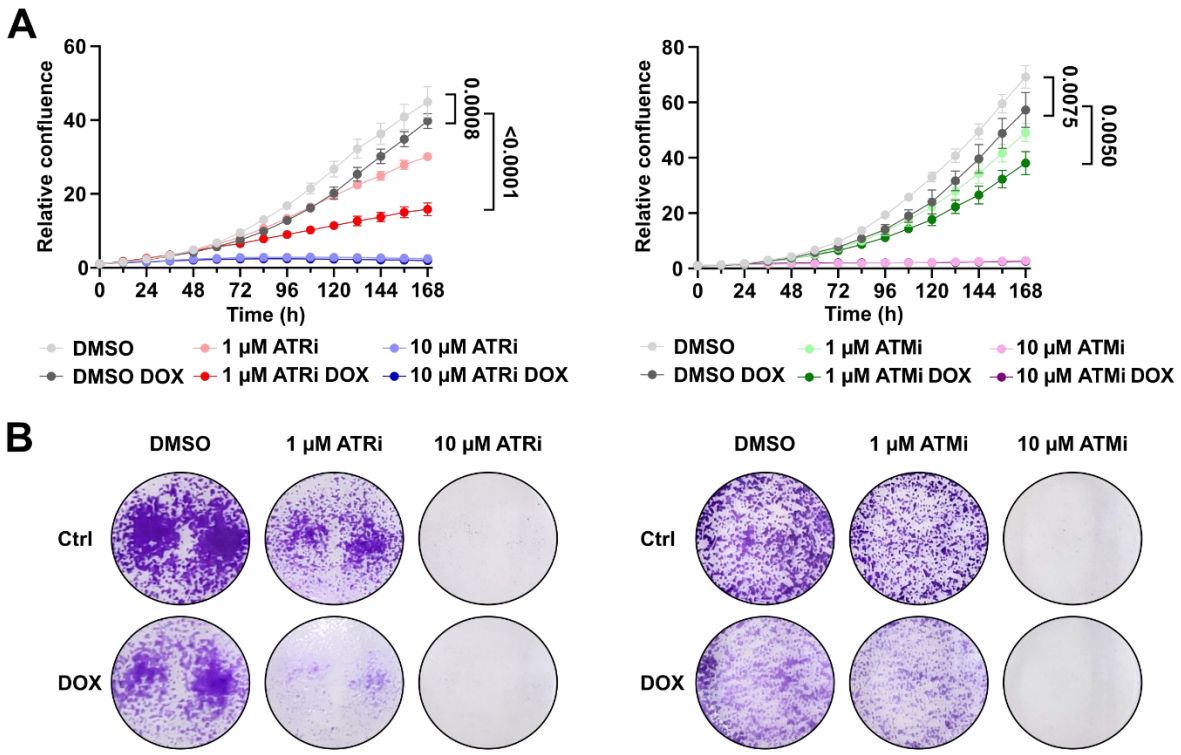

Figure 19: ATR inhibition exacerbates TRC-induced cell proliferation impairments

A) Proliferation assay of TRC reporter cells upon treatment with 0 or 1 µg/mL DOX for a duration of 168 h in 12 h intervals using Incucyte S3 Live-Cell Analysis System. Cells were additionally challenged with 1 or 10 µM ATRi with VE-821 or DMSO control treatment. Alternatively, cells were exposed to 1 or 10 µM ATMi with KU-60019 or DMSO control treatment. Data points represent the mean of three replicates (n=3) with error bars indicating SD. Area under the curve (AUC) measurements for each replicate were performed. Statistical analysis with Ordinary one-way ANOVA with Tukey's multiple comparison test was applied to the AUC measurements.

B) Corresponding crystal violet staining at the end of the time courses shown in (**A**).

4.4 R-loop mediated TRCs impair DNA replication fork progression

Since elevated FANCD2 levels have been connected to replication fork stalling events and have simultaneously been shown in R-loop-associated genomic instability processes (Bayona-Feliu et al., 2023; Okamoto et al., 2019; Shao et al., 2020), I wondered if the induction of R-loops and TRC at the mAIRN reporter will result in impaired DNA synthesis. Here, I first chose to investigate the interplay of DNA replication with TRCs and R-loops in a cell cycle-resolved manner, since impairments might be dependent on the local replication timing program. To this end, I performed a double thymidine block to synchronize the cells at the G1/S border and subsequently released the cells into S-phase for a time course between 0-8 h. First, I confirmed successful cell

synchronization by flow cytometry analysis using BrdU incorporation as a proxy for nascent DNA synthesis. Asynchronous control conditions showed the expected distribution of around 50 % in G1-phase, 45 % in S-phase, and 5 % in G2/M-phase (**Fig. 20**). Double thymidine block resulted in 98 % of cells synchronized at the G1/S border and allowed for efficient release with 91 % of cells being in S-phase after 4h.

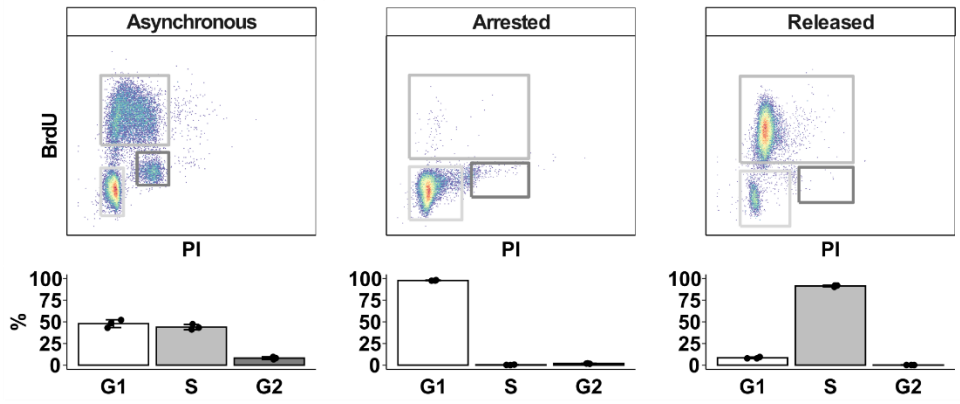

Figure 20: Flow cytometry analysis of asynchronous, arrested, and released cells

Flow cytometry analysis of asynchronous, double thymidine block arrested, and 4 h S-phase released cells. BrdU incorporation was used to detect nascent DNA synthesis. Frames in the cell distribution graphs show the gating applied for separation into G1, S, and G2/M cell populations. The percentage of cells in each cell cycle phase is shown below (n=3).

Using this efficient synchronization strategy, I next analyzed R-loop levels in G1 synchronized and 4 h released S-phase cells using DRIP-qPCR readout. R-loop levels at the induced mAIRN reporter site were about 2-fold increased in both G1 and S-phase conditions. Strikingly, a direct comparison of G1 and S-phase conditions revealed a modest but significant increase in S-phase cells when compared to G1 cells (**Fig. 21A**). As an important control, treating samples with *E. coli* RNase H reduced DRIP signal close to baseline levels, indicating the specificity of our DRIP-qPCR readout. Importantly, this effect was not the result of differences in transcriptional induction strength as measured by RT-qPCR (**Fig. 21B**), thus arguing for a DNA replication and potentially TRC-dependent stabilization of R-loops at the reporter sequences in S-phase cells. In contrast, unrelated R-loop forming control locus RPL13A did not show inducibility by DOX treatment and similar R-loop levels when comparing G1 arrested with S-phase released cells.

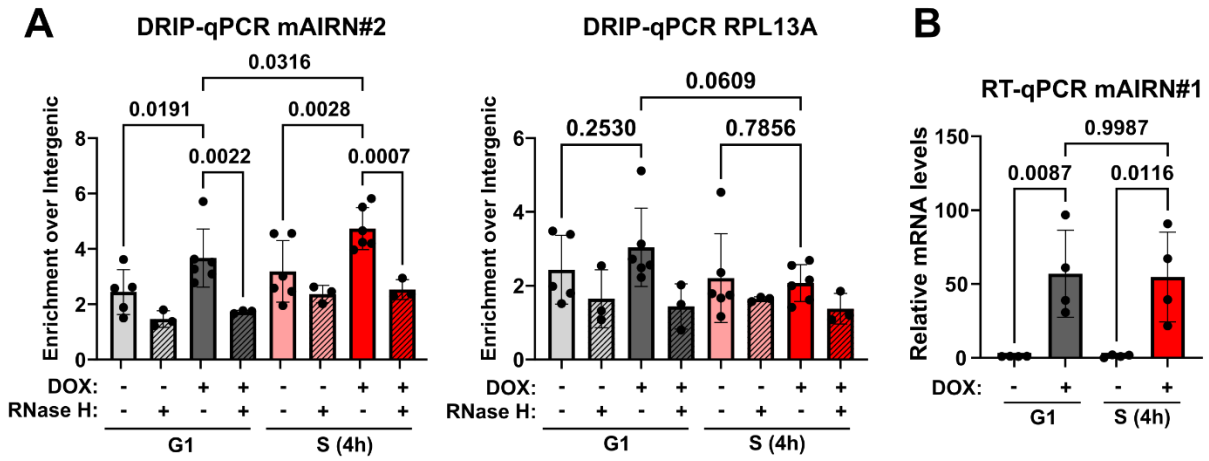

Figure 21: mAIRN reporter induction results in stable R-loop formation

A) R-loop levels measured by DRIP-qPCR at the mAIRN reporter or RPL13A locus in G1 or 4 h released S-phase cell treated with 0 or 1 µg/mL DOX using the mAIRN#2 primer pair. For RNase H control conditions, isolated genomic DNA from cells was incubated with *E. coli* RNase H1 overnight to degrade R-loops (n≥3).

B) RT-qPCR analysis of mAIRN gene expression levels in G1 or 4 h released S-phase cell treated with 0 or 1 µg/mL DOX using the mAIRN#1 primer pair (n=4). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

To precisely analyze the impact of induced TRC and R-loops on DNA replication through different stages of S-phase, I chose to perform a BrdU-seq experiment (J. Wang et al., 2021) at different time points across S-phase. Following the arrest by a double thymidine block for cell synchronization, cells were released in S-phase for a duration of 2 h, 4 h, 6 h, and 8 h and simultaneously exposed to DOX for TRC induction or respective control conditions. Additionally, samples of G1/S cells were kept in the presence of thymidine as negative controls without replication activity and therefore BrdU incorporation. 30 min before harvesting, cells were pulsed with BrdU for labeling of nascent DNA (**Fig. 22A**). BrdU-containing DNA was then immunoprecipitated from the resulting genomic DNA with a BrdU-specific antibody and further processed into BrdU-seq libraries. On the global level, BrdU-seq data clustered according to the experimental time point and displayed high levels of correlation between the biological replicates (**Fig. 22B**). Further, early time points (2-4 h) showed higher similarity than late replicating regions (6-8 h). G1 samples did not correlate, indicating an expected unspecific distribution of reads.

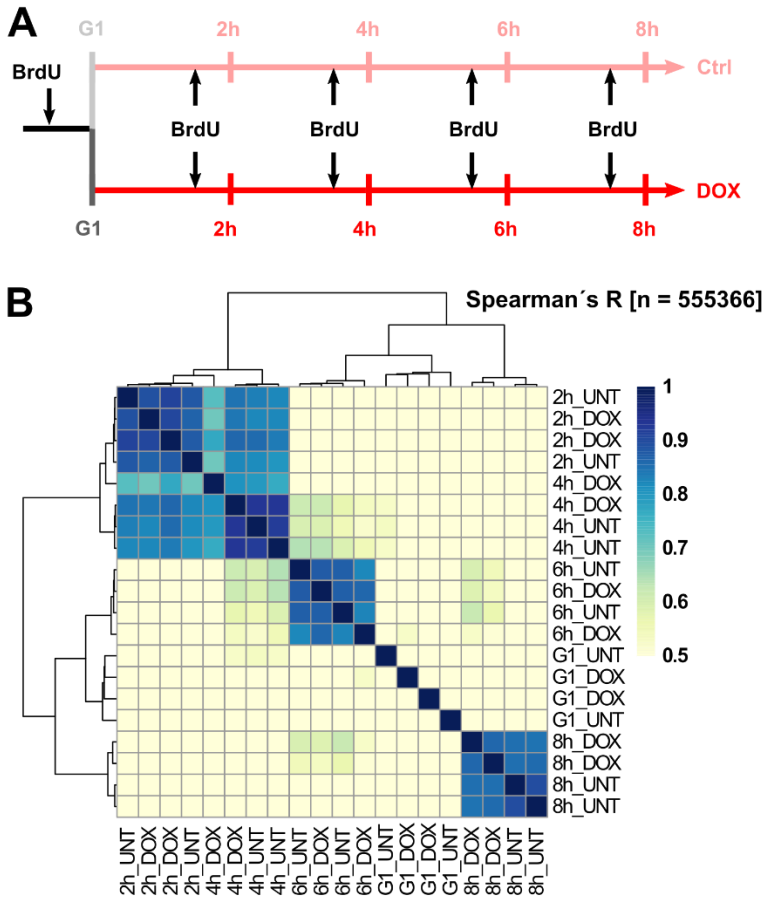

Figure 22: BrdU-seq data is highly correlated and clusters according to S-phase progression

A) Schematic depiction of the treatments in the BrdU-seq time course experiment. Cells were synchronized at the G1/S border with double thymidine block. Cells were released into S-phase for durations of 2 h, 4 h, 6 h or 8 h. Cells were pulsed with 25 µM BrdU for 30 min to label nascent DNA. Upon release into S-phase, cells were treated with 0 or 1 µg/mL DOX. G1 control cells were continuously cultured in thymidine-containing medium and treated with DOX for 8 h.

B) Heatmap displaying replicate correlation (Spearman's R) of both BrdU-seq replicates for all time points and treatment conditions. ´n´ denotes the number of genomic bins with a size of 5 kb.

Before asking if R-loops and TRC could directly impair DNA replication fork progression at the reporter sites, I first used the time resolved BrdU-seq data to define the replication timing of the individual integration sites. As highlighted by an example region of ca. 10 Mb on chromosome 10, the BrdU-seq data could successfully distinguish early-, mid-, and late-replicating regions of the genome, as indicated by preferential BrdU incorporation at 2h-4h, 4h-6h, or 6h-8h after release into S-phase, respectively (**Fig. 23A**). This genome-wide data allowed me to zoom into the replication timing dynamics of the five reporter sites by quantification of the BrdU-seq signal in

5 kb bins in a +/-100 kb region around the five integration sites (**Fig. 23B**). Intriguingly, the Chr 5, Chr 9, and Chr 10 sites all showed the strongest enrichment of BrdU-seq signal at 2 h into S-phase (**Fig. 23B**). Subsequent time points displayed little to no signal, defining these sites as early replicating. In contrast, Chr 2 site 2 exhibited an early to mid-replication behavior with the strongest signal at 2-4 h into S-phase. Chr 2 site 1 was the only late-replicating region with BrdU-seq signal only appearing 6-8 h after release into S-phase. Collectively, the integration sites exhibit a predominantly early S-phase replication timing, which is consistent with the reported preferential targeting of open and active chromatin regions by Sleeping Beauty transposase (G. Liu et al., 2005; Miskey et al., 2022), a chromatin state typically associated with early replication timing (Rhind & Gilbert, 2013; P. A. Zhao et al., 2020).

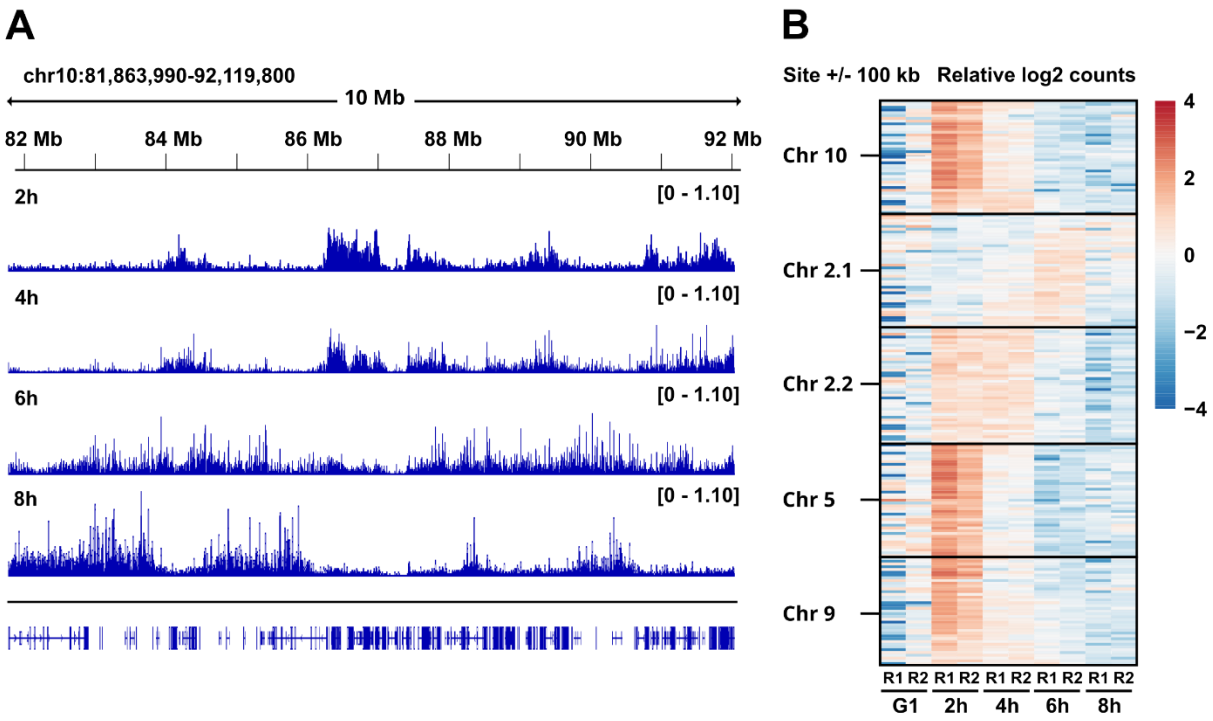

Figure 23: Replication timing analysis identifies the majority of integrated reporter sites in early replicating domains of the genome

A) Genome browser snapshot of a representative 10 Mb region on chromosome 10 demonstrating that BrdU-seq time course analysis can successfully track DNA replication timing and detect early, mid, and late replicating domains.

B) Heatmap of BrdU-seq signal in +/-100 kb regions around the mAIRN integration sites with 5 kb bin size in synchronized G1 cells and 2 h, 4 h, 6 h, and 8 h released S-phase cells. BrdU-seq signal is displayed as log2 normalized read counts relative to the mean of all samples. The signal of both biological replicates is shown side by side (R1 and R2).

DOX-treated versus control samples were highly similar when compared from a global perspective, implying that transcriptional activation of the mAIRN reporter does not perturb DNA replication fork progression on a genome-wide scale with the exception of one non-specific bin (**Fig. 24A**). This is consistent with previous results where TRC induction in unchallenged cells did not impair DNA replication within a single cell cycle (**Fig. 18B**). Instead, a drastic reduction of BrdU-seq signal was observed in the DOX sample compared to the control sample, which exclusively spanned the 5 kb genomic bin harboring the mAIRN reporter sites (**Fig. 24B, C**). Crucially, this significant disruption of BrdU incorporation was only detectable at the 2 h release time point that matches the previously defined replication timing of 4 out of the 5 integration locations. Quantification of the BrdU-seq signal in the same +/-100 kb regions used for replication timing analysis also revealed the mAIRN reporter as the only 5 kb bin with significantly reduced BrdU incorporation at 2 h into S-phase (**Fig. 24B**).

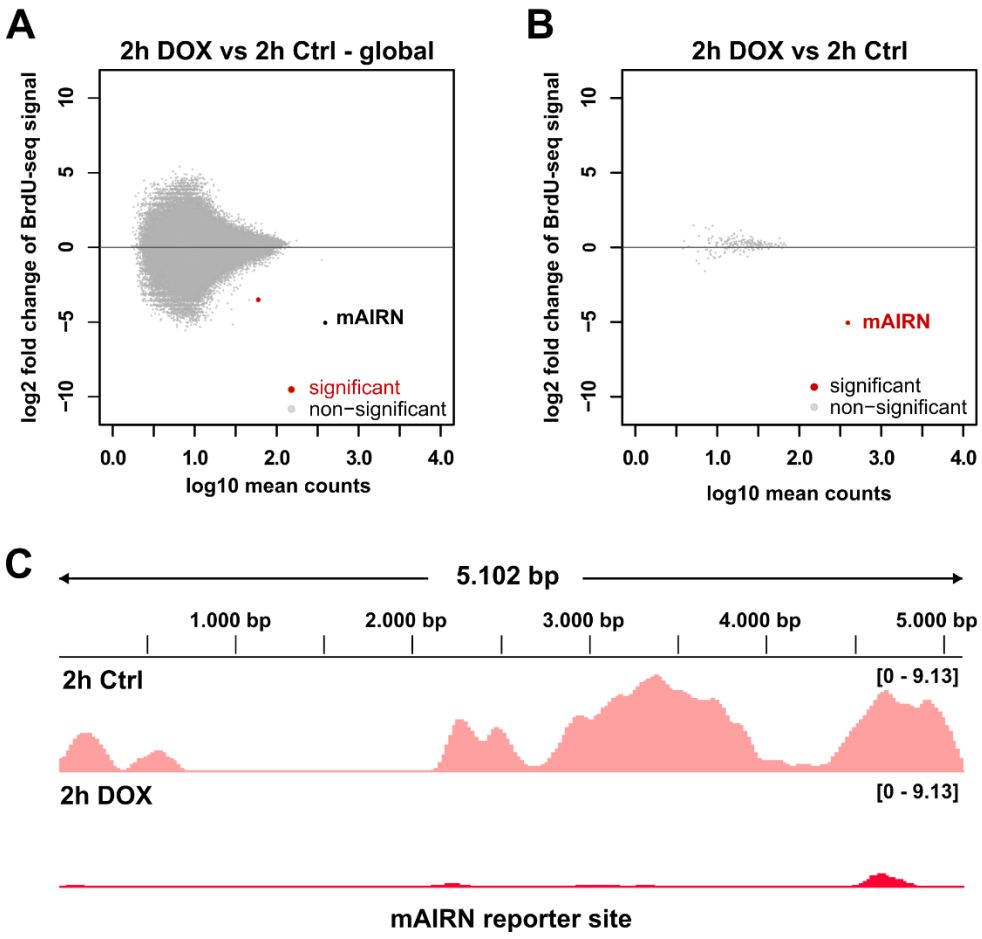

Figure 24: TRC induction at the mAIRN reporter induces drastic replication fork impairments

A) MA plots showing differential regulation of BrdU-seq signal at the 2 h S-phase release time point comparing DOX vs Ctrl conditions globally. 5 kb bin size. Significant bins are highlighted in red, and not significantly changed bins are displayed in grey.

B) MA plots showing differential regulation of BrdU-seq signal at the 2 h S-phase release time point comparing DOX vs Ctrl conditions in the +/- 100 kb regions around the integration sites shown in **Fig. 23B**. 5 kb bin size. Significant bins are highlighted in red, and not significantly changed bins are displayed in grey.

C) Representative genome browser snapshot of averaged BrdU-seq signal across all mAIRN reporter integrations at the 2 h S-phase time point in 0 or 1 µg/mL DOX treated cells.

In contrast, no major differences could be seen for the 4 h, 6 h, and 8 h samples (**Fig. 25**). Interestingly, although insignificant, slightly elevated BrdU-signal could be detected for the 4 h time point, suggesting a delayed and potentially compensatory DNA synthesis at the mAIRN reporter following the previously incomplete replication at 2 h (**Fig. 24B, C**). In summary, our data demonstrates that TRCs driven by mAIRN R-loop formation can lead to direct and strong DNA replication fork impediments that could cause the earlier identified DNA damage, fork stalling, and growth phenotypes.

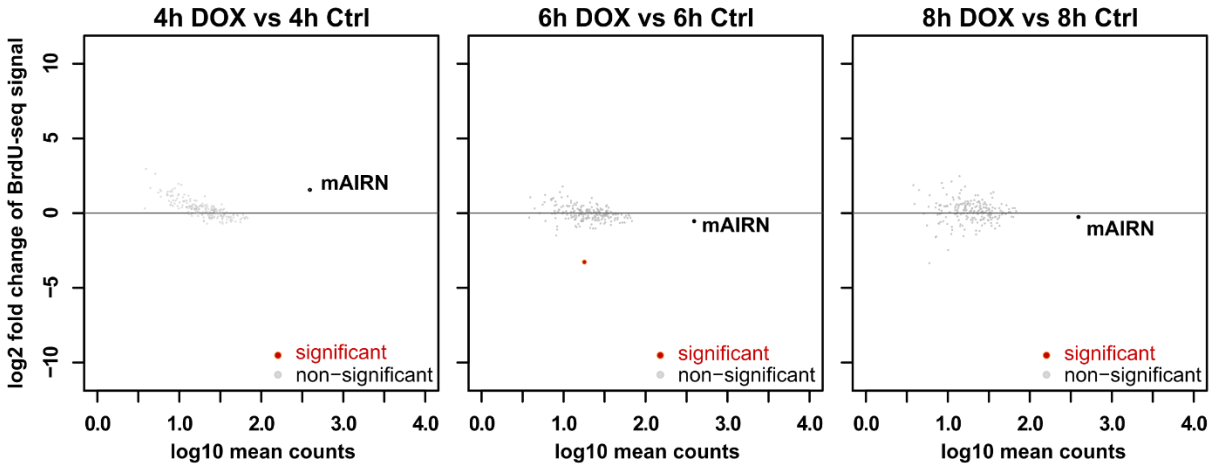

Figure 25: TRC induction at the mAIRN reporter does not affect BrdU incorporation in mid-to-late S-phase

MA plots showing the differential regulation of BrdU-seq signal at the 4 h, 6 h, and 8 h S-phase release time points comparing DOX vs Ctrl conditions in the +/- 100 kb regions around the integration sites. Bin size 5 kb. Significant bins are highlighted in red, and not significantly changed bins are displayed in grey.

4.5 TRC induction disrupts the local chromatin structure on integrated R-loop reporter sites

Having discovered that reporter system-driven TRCs can lead to replication impairments and DNA damage, I next asked whether this local interference could also disrupt the underlying chromatin template. Analogous to the approach for the plasmid TRC reporters, my initial goal was to characterize the behavior of RNAPII and histone H3 levels, allowing me to obtain insight into the dynamics of gene transcription and nucleosome occupancy upon TRC induction. For this purpose, I conducted ChIP-qPCR in synchronized cells 4 h after S-phase release with antibodies targeting RNAPII Ser2P and histone H3. As anticipated, RNAPII Ser2P occupancy was strongly increased at the mAIRN reporter site upon the addition of DOX, thus confirming successful RNAPII recruitment consistent with transcriptional activation (**Fig. 26A**). Remarkably, activated reporter transcription led to a strong reduction of nucleosome occupancy as indicated by significantly decreased H3 enrichment at the mAIRN loci (**Fig. 26B**). DOX treatment did not alter RNAPII Ser2P at the highly transcribed housekeeping *ACTB* gene, nor affected histone H3 occupancy at the non-transcribed, nucleosome-dense *NRXN2* gene (**Fig. 26A, B**), serving as specificity controls for the DOX treatment.

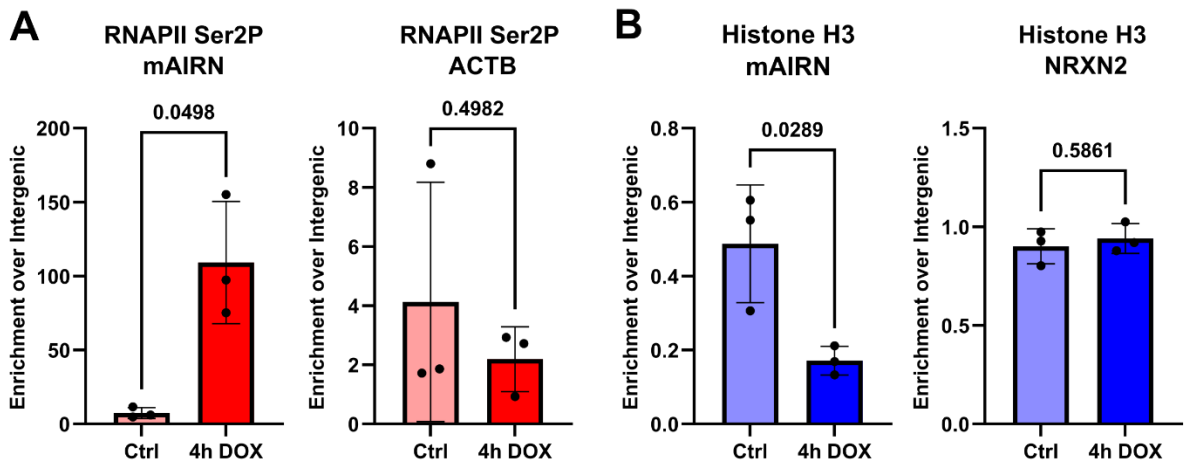

Figure 26: TRC and R-loop induction at the mAIRN reporter reduce nucleosome occupancy

A) ChIP-qPCR analysis depicting RNAPII Ser2P at the reporter site (mAIRN#1 primers) or ACTB control locus in synchronized S-phase cells 4 h after release from double thymidine block. Cells were treated with 0 or 1 µg/mL DOX for 4 h (n=3). Error bars indicate SD. Welch's t-test.

B) Same as in **A)** for ChIP-qPCR of histone H3 at the mAIRN#1 or NRXN2 sites.

Next, I wondered if the observed H3 loss is a specific result of R-loop formation, as suggested by the previously obtained *in vitro* data (**Fig. 8B**), or occurs in a non-R-loop dependent way, for

example as a by-product of the strong transcriptional activation and RNAPII load at the mAIRN sequence. To this end, I attempted to rescue the depletion of nucleosomes by overexpression of human FLAG-tagged RNase H1, an enzyme responsible for the degradation of R-loops and an accepted approach to confirm R-loop dependency of a phenotype (Hyjek et al., 2019; H. Zhao et al., 2018). In addition to H3 levels, I also quantified FANCD2 levels by ChIP-qPCR upon RNase H1 overexpression, as FANCD2 had been shown to be recruited to R-loop sites (Liang et al., 2019). Despite Western blot analysis showing highly efficient overexpression of RNase H1 after 24 h of transfection (**Fig. 27A**), RNase H1 could neither rescue the loss of H3 nor the previously observed increase of FANCD2 (**Fig. 27B, C**).

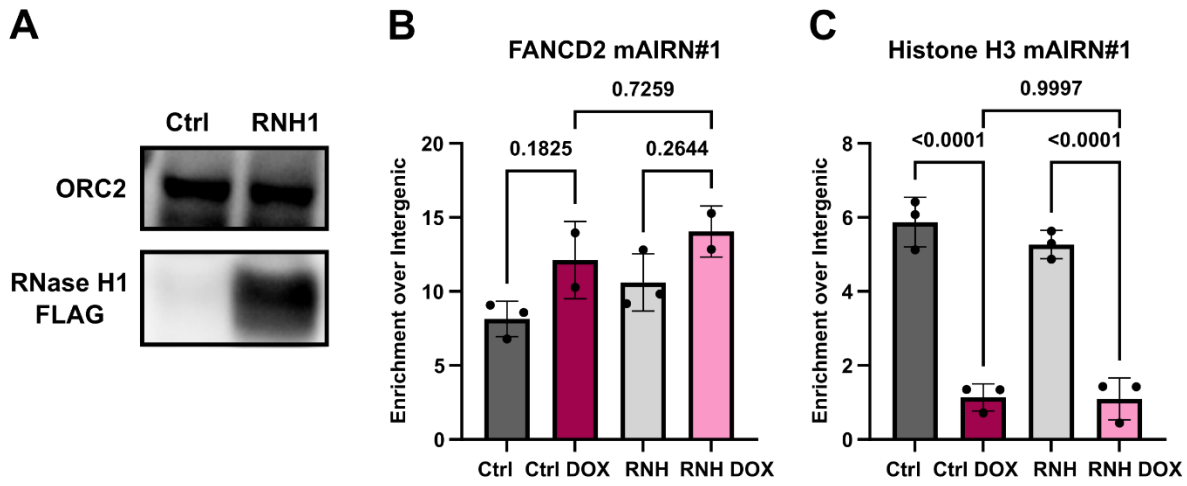

Figure 27: Loss of nucleosome occupancy and FANCD2 recruitment cannot be rescued by RNase H1 overexpression

A) Representative Western blot images displaying RNase H1-FLAG signal in Clone#12 upon overexpression of RNase H1 or an empty vector plasmid for 24 h. ORC2 serves as a loading control.
B) ChIP-qPCR data of FANCD2 in Clone#12 cells at the mAIRN#1 locus upon overexpression of RNase H1 (RNH) or an empty vector plasmid for 24 h. Cells were treated with 0 or 1 µg/mL DOX 24 h (n=3). Error bars show SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.
C) Same as in **B)** for ChIP-qPCR data for histone H3.

While this result was initially unexpected, particularly strong R-loops resistant to processing by RNase H1 had been described before (Crossley et al., 2020). Given the artificially strong and stable nature of the induced mAIRN R-loops, limited degradation efficiency by RNaseH1 could be a possible explanation for the lack of rescue. Alternatively, the nucleosome loss or FANCD2 accumulation at induced mAIRN sites could also be R-loop independent. To discriminate between these two possibilities, I took advantage of the matching non-R-loop forming ECFP control cell line as an independent and complementary approach (**Fig. 28A**). Strikingly, H3 levels were not

significantly decreased at the induced ECFP gene or the unrelated *NRXN2* control locus (**Fig. 28B**), providing evidence that the nucleosome loss at the mAIRN reporter is likely a consequence of its strong R-loop formation. Furthermore, this data is agreement with the previously observed inability of DOX-induced ECFP cells to create a local DNA damage response in FANCD2 ChIP-qPCR (**Fig. 15C**), suggesting that both DNA damage response and nucleosome loss induction require an R-loop forming sequence.

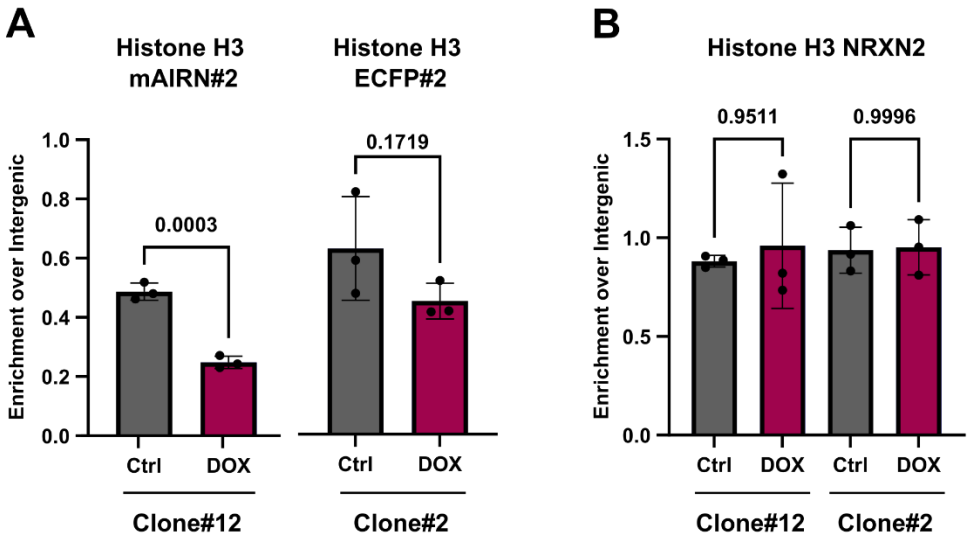

Figure 28: Nucleosome loss occurs at R-loop forming mAIRN sites but not at ECFP control sites

A) ChIP-qPCR analysis of Histone H3 levels in Clone#12 and Clone#2 cells at the mAIRN#2 and ECFP#2 loci. Cells were exposed to 0 or 1 µg/mL DOX for 24 h (n=3). Error bars indicate SD. Student's test.

B) Same as in **A)** for the NRXN2 control region.

To obtain a more global overview of the transcriptional and nucleosome changes upon DOX induction, the ChIP-qPCR samples were subsequently used for preparation of Next Generation Sequencing (NGS) libraries to perform ChIP-seq analysis. Averaging the obtained ChIP-Seq signal across the five mAIRN integration sites demonstrated that the transcription activation and nucleosome depletion phenotypes are not restricted to the mAIRN sequence itself but rather extend throughout the entire reporter construct of 5 kb (**Fig. 29A**). As this analysis depicts the average signal of all five integration sites due to their identical sequence, I wondered if differential behavior could be detected in the close vicinity of each integration site. To this end, I analyzed RNAPII Ser2P and histone H3 ChIP-seq data in 100 bp bins -/+ 5kb upstream and downstream of the reporter sites. Surprisingly, RNAPII Ser2P signal did not remain confined to the integrated sequence but frequently extended beyond its genomic location with heterogenous extend

(**Fig. 29B**). For example, Chr 2 site 1 displays increased levels of RNAPII Ser2P upon DOX treatment, which extended into the upstream intronic sequence of the MRPS9-AS2 gene for several thousand base pairs (**Fig. 29C**).

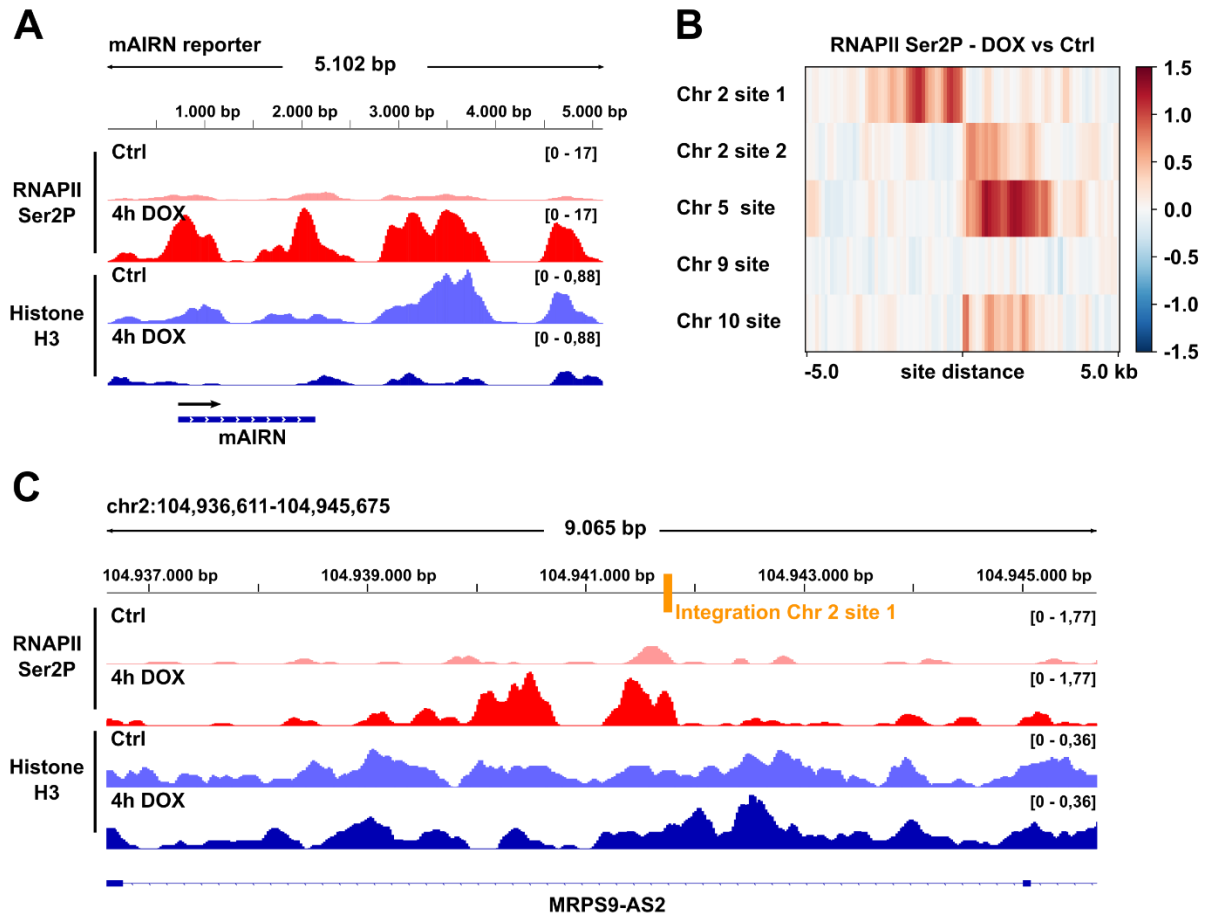

Figure 29: RNAPII Ser2P levels are increased in genomic bins neighboring the induced reporter sites

A) Representative genome browser snapshot of RNAPII Ser2P and histone H3 signal along the full mAIRN reporter construct in DOX treated or untreated conditions in synchronized S-phase cells.

B) Heatmap showing log2 fold change of RNAPII Ser2P signal upon DOX treatment over control in a +/- 5 kb region surrounding the integration site locations (the 5 kb bin containing the mAIRN reporter itself is not included), 100 bp bin size.

C) Representative genome browser snapshot displaying RNAPII Ser2P and histone H3 signal at the MRPS9-AS2 locus which contains the Chr 2 site 1 integration site (orange bar) in the same conditions as in **A)**. Sequencing libraries were derived from samples shown in **Fig. 26A** and **26B**.

To test whether elevated RNAPII Ser2P occupancy would result in gene expression changes, I performed RT-qPCR analysis. Interestingly, *MRPS9-AS2* as a gene in direct vicinity to the mAIRN reporter was 50-fold upregulated upon DOX treatment (**Fig. 30A**). In contrast, neighboring genes of the Chr 2 site 2 (*CERS6* and *STK9*) or Chr 10 site (*FAM25A* and *GLUD1*) located further away from the integration site were not affected in their expression (**Fig. 30A**). Due to the predominantly unidirectional extension of the RNAPII Ser2P signal beyond the reporter sites, this is most likely the result of readthrough transcription that is determined by the orientation of the integrated Tet-ON promoter controlling mAIRN transcription. While all integration sites showed varying degrees of readthrough transcription (**Fig. 29B, C, and Fig. 30B**), the Chr 9 site did not exhibit transcription outside the reporter (**Fig.30 C**).

Reporter induction did not induce any global deregulation of transcription apart from the affected genomic bins in direct vicinity of the integration sites (**Fig. 31A**). Most importantly, these neighboring genomic bins (the reporter site of 5 kb is excluded from the visualization) displaying RNAPII readthrough were unaltered in their H3 occupancy (**Fig. 31B, C**). This data suggests that the H3 loss phenotype cannot be explained by high levels of RNAPII occupancy and transcription but is rather specifically driven by the R-loop and TRC formation of the mAIRN reporter. Global histone H3 levels remained unchanged (**Fig. 31B**). In summary, these findings demonstrate that R-loop-driven TRCs can cause a reduction of local nucleosome occupancy, thus disturbing the local chromatin environment and increasing its vulnerability to DNA breakage and genomic destabilization.

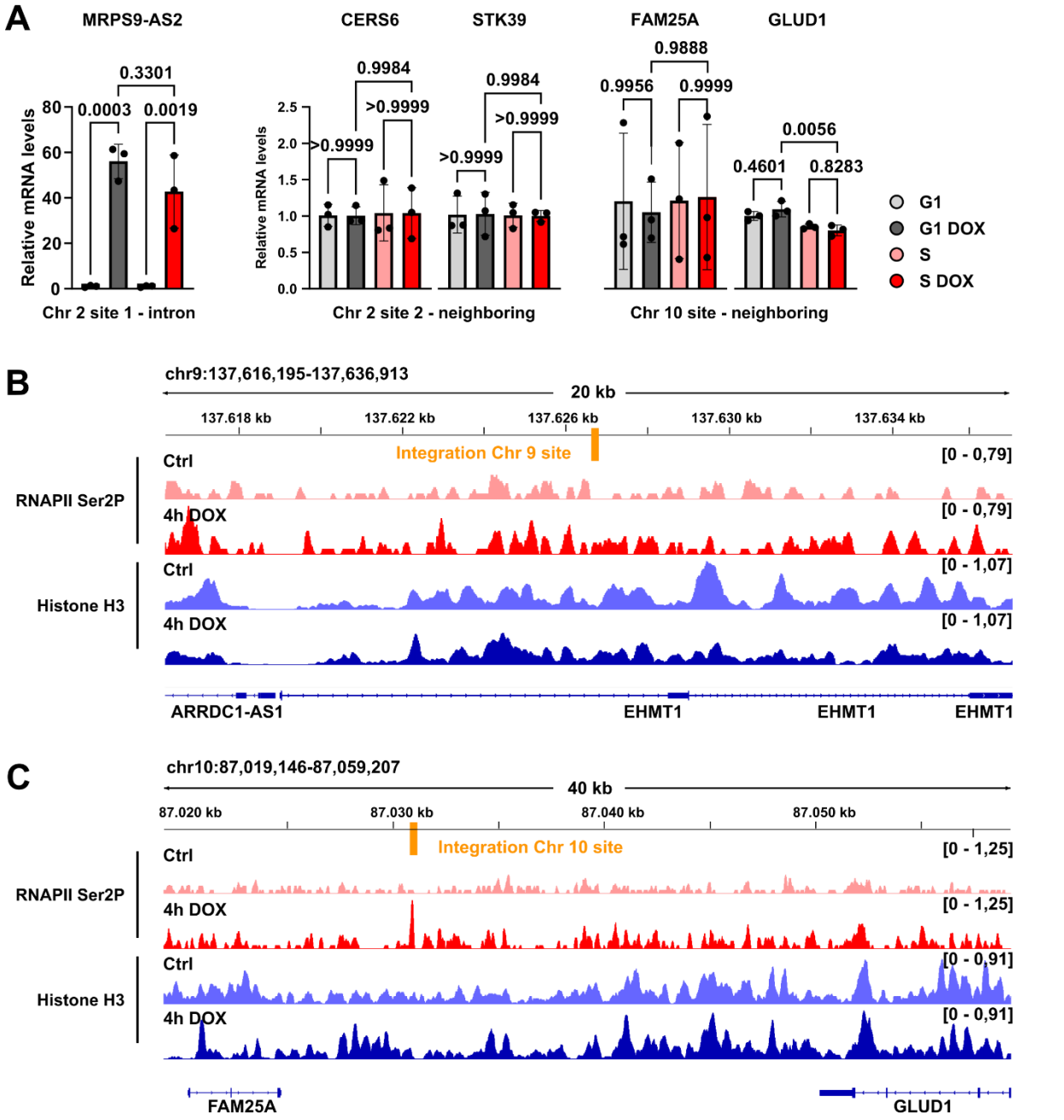

Figure 30: Readthrough transcription of the activated reporter affects the surrounding genome in a site-specific manner

A) RT-qPCR quantification of expression levels of genes containing or neighboring different reporter integration sites. Synchronized G1 or 4 h released S-phase cells were treated with 0 or 1 µg/mL DOX for 4 h (n=3). Error bars show SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

B) Representative genome browser snapshot of a 20 kb region surrounding the Chr 9 integration site region (orange bar) showing RNAPII Ser2P and histone H3 occupancy in DOX-treated or untreated conditions in synchronized S-phase cells.

C) Same as in B) for a 40 kb region around the integration site on Chr 10, which also contains the *FAM25A* and *GLUD1* genes.

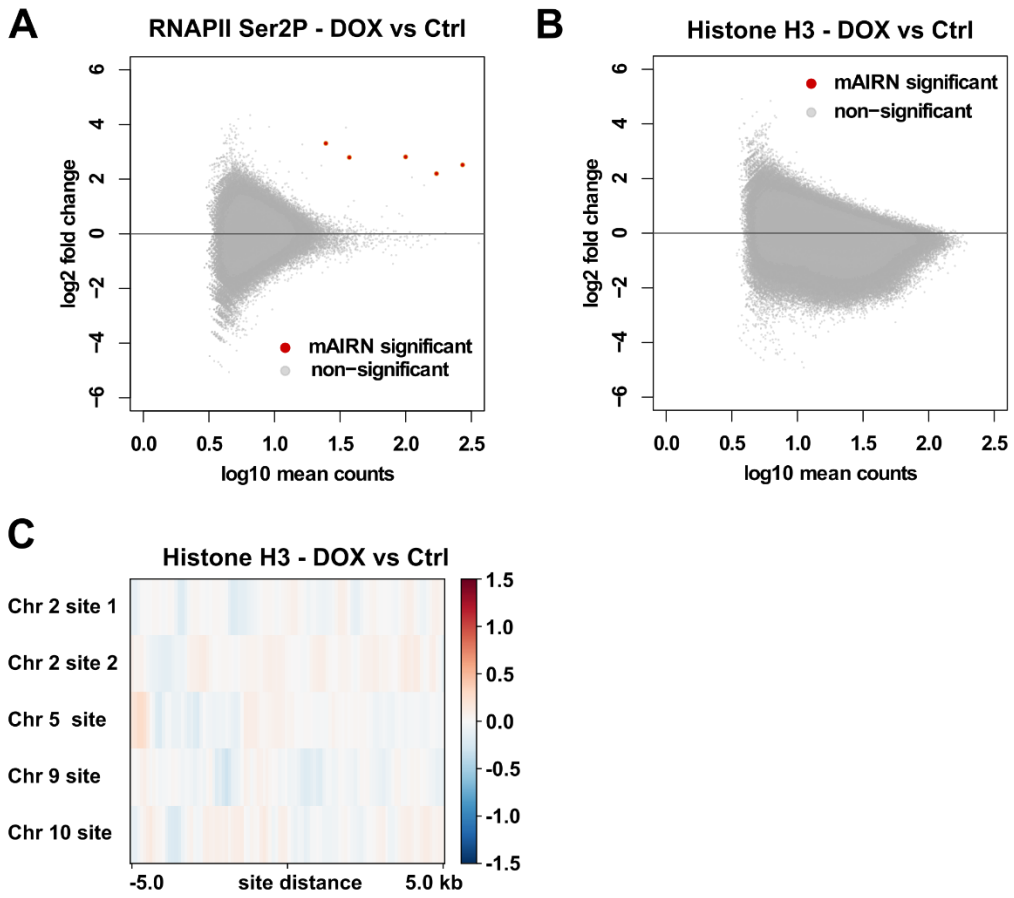

Figure 31: Loss of nucleosome occupancy remains confined to the R-loop reporter sequence and does not occur genome-wide

A) MA plots showing differential regulation of global RNAPII Ser2P ChIP signal, 1 kb bins. Significantly changed mAIRN bins are highlighted in red, not significantly affected bins are displayed in grey.
B) MA plot of histone H3 ChIP signal in the same parameters as in **A)**.
C) Heatmap illustrating the log2 fold change of histone H3 levels upon DOX treatment over control in a +/- 5 kb region surrounding the integration site locations, 100 bp bin size.

4.6 H3K79 methylation is a TRC-enriched histone modification at the R-loop reporter and genome-wide

Beyond the increase in chromatin accessibility resulting from nucleosome eviction, I wondered if TRC and R-loop induction could impact other chromatin features. As several papers have reported altered levels of histone posttranslational modifications (PTMs) such as H3K4 methylation (H3K4me), H3S10 phosphorylation (H3S10P), and H2AK119ub at TRCs in different model systems (Bayona-Feliu et al., 2023; Chong et al., 2020; Hao et al., 2022), I decided to comprehensively analyze histone PTMs in mAIRN reporter cells using these previously identified

marks as a candidate list for ChIP-qPCR experiments. Importantly, for each PTM, I included a side-by-side histone H3 ChIP from the identical nuclear lysate, thus allowing me to account for the loss of histone H3 occupancy at the induced mAIRN reporter and perform normalization to H3 levels accordingly (**Fig. 32A**). Interestingly, the loss of histone H3 was already noticeable upon 4 h of DOX treatment in arrested G1 cells and further decreased upon release of cells into S-phase with simultaneous DOX exposure. This observation was consistent for several tested primer pairs along the reporter sequence (**Fig. 32B, C**), suggesting that active DNA replication could further reduce nucleosome occupancy, an effect likely resulting from enhanced R-loop stabilization by TRCs in S-phase (see **Fig. 21A** and discussion). No differential H3 occupancy was observed at the unrelated NRXN2 control site, suggesting that this effect is specific to the mAIRN TRC reporter (**Fig. 32D**).

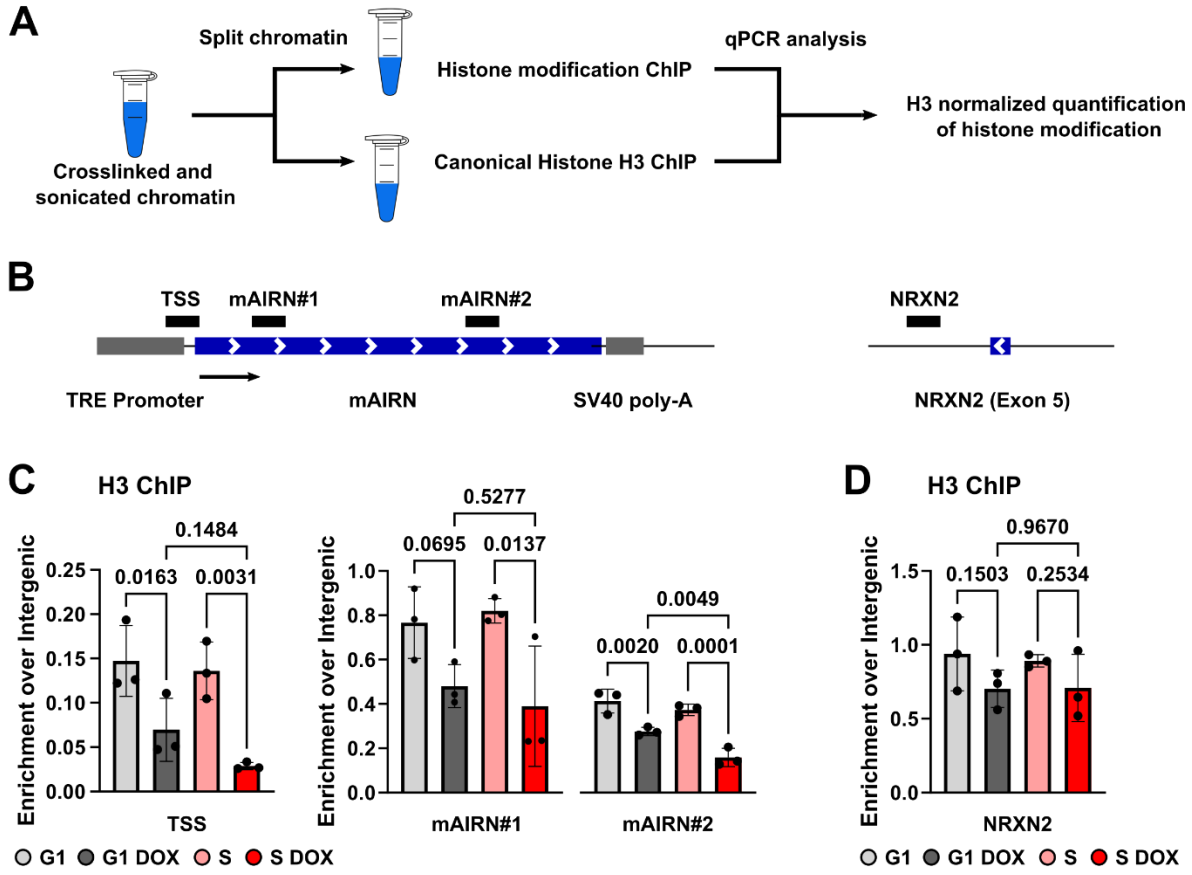

Figure 32: Loss of nucleosome occupancy is exacerbated at the reporter in S-phase cells

A) Cartoon of the ChIP workflow used for screening TRC-dependent histone modifications. For each tested modification, a matched ChIP of canonical histone H3 from identical nuclear lysates was conducted, allowing for subsequent H3 normalization.

B) Schematic representation of the mAIRN locus of the integrated reporter construct and control region near NRXN2 exon 5. Locations of the tested primer pairs (TSS, mAIRN#1, mAIRN#2, and NRXN2) in subsequent ChIP experiments are indicated as black bars.

C) Histone H3 ChIP in G1 or 4 h released S-phase cell treated with 0 or 1 µg/mL DOX for 4 h. H3 levels were quantified at the reporter sequence with primers at TSS, mAIRN#1, and mAIRN#2 or the NRXN2 control locus (n=3). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

D) Same as in **C)** for the NRXN2 locus.

H3K4 trimethylation (H3K4me3) is frequently deposited at transcription start sites and shows a decreasing 5' to 3' gradient throughout the gene body. H3K4 methylation was recently characterized as a mitigator of TRCs by acting as a "speed bump" to decelerate active replication forks (Chong et al., 2020). Consistently, I found a DOX-dependent increase of H3K4me3 in S-phase cells at primer locations mAIRN#1 and mAIRN#2 in the gene body of the mAIRN reporter, but a region just upstream of the TSS or the unrelated NRXN2 locus remained unchanged. (**Fig. 33A)**. Moreover, recent work demonstrated that a dynamic switch from H2AK119cr to H2AK119ub leads to transcriptional repression and eviction of RNAPII for TRC prevention (Hao et al., 2022). In agreement, I found H2AK119ub to be enriched in DOX-treated S-phase cells particularly at the TSS and 5' region of the mAIRN reporter gene, whereas little to no differential occupancy between the conditions could be detected towards the 3' end and at the NRXN2 control site (**Fig. 33B**). Lastly, the behavior of H3K79 di- and trimethylation (H3K79me2/3) upon TRC induction was monitored at the mAIRN reporter. H3K79me2/3 was shown to be enriched at the promoters and gene bodies of actively transcribed genes (Steger et al., 2008; Veloso et al., 2014) but also to aid DNA replication at a subset of origins and prevent genomic instability through recruitment of TP53-binding protein 1 (53BP1) (Fu et al., 2013; Huyen et al., 2004), making it an interesting target to study in the context of TRCs. Strikingly, both H3K79me2 and H3K79me3 were strongly enriched at all tested locations of the induced mAIRN TRC reporter and showed particularly high levels in DOX-induced S-phase cells (**Fig. 33C, D**). In contrast, no changes were seen at the NRXN2 control locus in all conditions tested. (**Fig. 33C, D**).

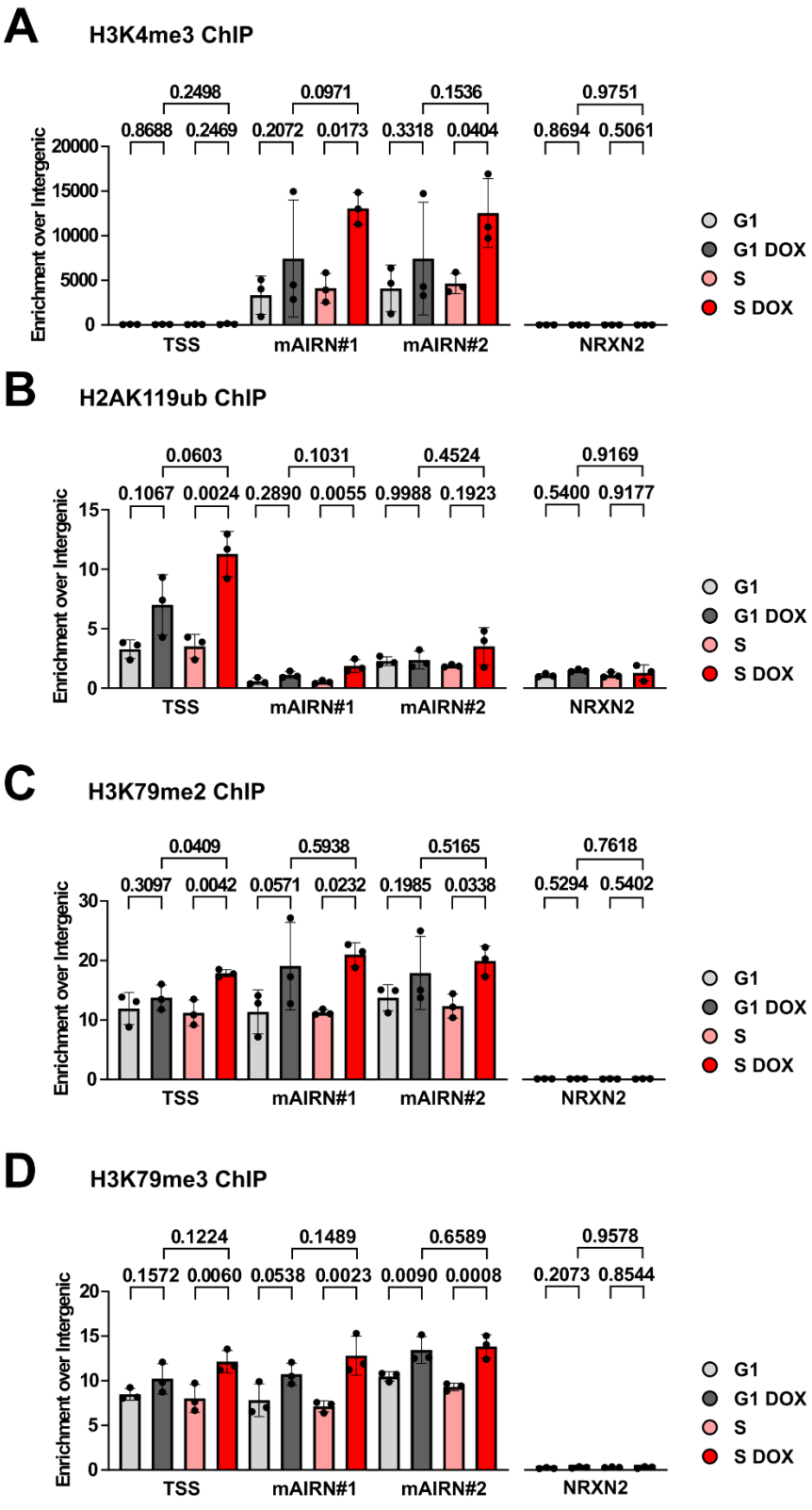
A
H3K4me3 ChIP
0.2498
0.0971
0.1536
0.9751
0.8688
0.2469
0.2072
0.0173
0.3318
0.0404
0.8694
0.5061
Enrichment over Intergenic
20000
15000
10000
5000
0
TSS
mAIRN#1
mAIRN#2
NRXN2
G1
G1 DOX
S
S DOX
B
H2AK119ub ChIP
0.0603
0.1031
0.4524
0.9169
0.1067
0.0024
0.2890
0.0055
0.9988
0.1923
0.5400
0.9177
Enrichment over Intergenic
15
10
5
0
TSS
mAIRN#1
mAIRN#2
NRXN2
G1
G1 DOX
S
S DOX
C
H3K79me2 ChIP
0.0409
0.5938
0.5165
0.7618
0.3097
0.0042
0.0571
0.0232
0.1985
0.0338
0.5294
0.5402
Enrichment over Intergenic
30
20
10
0
TSS
mAIRN#1
mAIRN#2
NRXN2
G1
G1 DOX
S
S DOX
D
H3K79me3 ChIP
0.1224
0.1489
0.6589
0.9578
0.1572
0.0060
0.0538
0.0023
0.0090
0.0008
0.2073
0.8544
Enrichment over Intergenic
20
15
10
5
0
TSS
mAIRN#1
mAIRN#2
NRXN2
G1
G1 DOX
S
S DOX

Figure 33: H3K4me3, H2AK119ub, and H3K79 methylation are TRC-enriched chromatin modifications at the R-loop reporter

A) H3 normalized H3K4me3 ChIP-qPCR in G1 or 4 h released S-phase cell treated with 0 or 1 µg/mL DOX for 4 h. Occupancy levels of the respective histone modification were quantified at the reporter sequence with primers at TSS, mAIRN#1, mAIRN#2, or the NRXN2 control site (n=3). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.
B) Same as in **A)** for ChIP-qPCR of H2AK119ub.
C) Same as in **A)** for ChIP-qPCR of H3K79me2.
D) Same as in **A)** for ChIP-qPCR of H3K79me3.

To independently validate these findings, I chose to leverage a previously established bioinformatic approach to test for the enrichment of different histone modifications at putative genomic HO or CD TRC regions (Hamperl et al., 2017). In brief, putative genomic regions were defined by overlapping DRIP-seq data for mapping of R-loops, Global run-on sequencing (GRO-seq) data for quantifying the immediate transcriptional activity of a gene, and Okazaki fragment sequencing (OK-seq) data for mapping of replication origins and fork directionality. By integrating these three publicly available genomic datasets in HeLa cells, genome-wide TRC hotspots were defined based on genomic regions, in which replication origins were residing within R-loop forming genes that are actively transcribed in S-phase (**Fig. 34A**). Furthermore, transcribing RNA polymerase at these hotspots have the chance to collide with bidirectionally progressing replication forks either in HO or CD direction, allowing for the distinction between HO or CD-oriented TRCs. After the identification of these putative TRC sites, publicly available ENCODE ChIP-seq data were obtained for different histone modifications in HeLa cells and their profiles overlayed over the TRC hotspot regions. H3K4me3, H3K27me3, and H3K36me3 displayed symmetric distributions between HO and CD sides with varying enrichment levels (**Fig. 34B-D**). In strong contrast, H3K79me2 and H3K79me3 showed a drastic asymmetry towards the HO regions, implying that these modifications might be a marker of the more detrimental HO collisions (**Fig. 34E, F**). H2AK119ub ChIP-Seq data were only publicly available in MCF7 breast cancer cells and did not show meaningful enrichment of this modification over the HeLa TRC hotspot regions, suggesting that the disparity between the two cell lines may cause this analysis to be inconclusive (**Fig. 34G**).

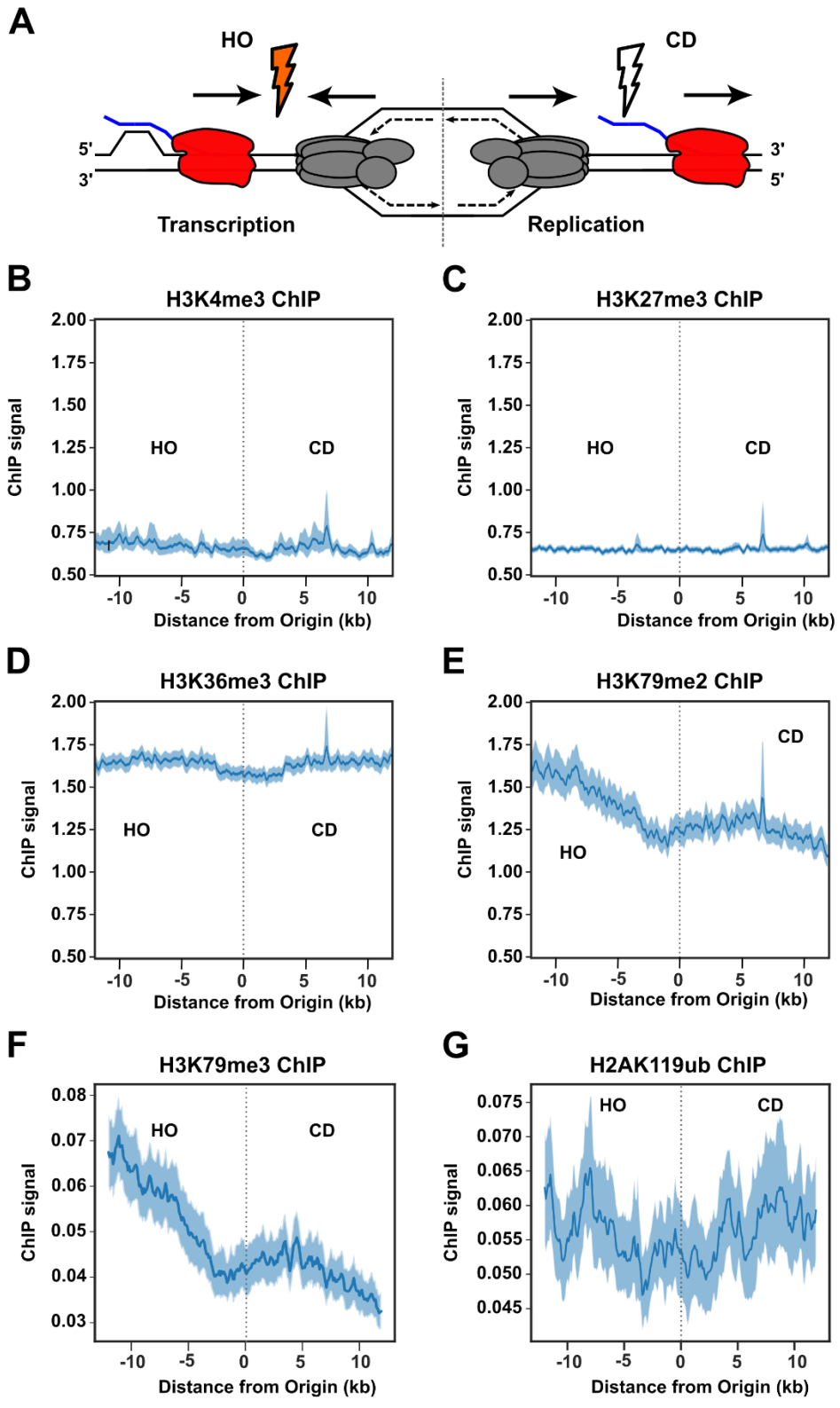
A
HO
CD
5'
3'
3'
5'
Transcription
Replication
B
H3K4me3 ChIP
C
H3K27me3 ChIP
D
H3K36me3 ChIP
E
H3K79me2 ChIP
F
H3K79me3 ChIP
G
H2AK119ub ChIP
ChIP signal
Distance from Origin (kb)

Figure 34: H3K79 methylation is enriched at genomic TRC and R-loop sites

A) Cartoon for the selection of genomic regions prone to HO versus CD collisions by the identification of intragenic origins of replication within actively transcribed genes (Hamperl et al., Cell 2017).

B) Analysis of H3K4me3, **C)** H3K27me3, **D)** H3K36me3, **E)** H3K79me2; and **F)** H3K79me3 ChIP-seq signal from HeLa cells at intragenic origins residing in actively transcribed genes. The analysis windows surrounding the regions span 24 kb in size and are positioned at least 5 kb away from promoters and terminators. Error bands show a 95 % confidence interval as determined by a bootstrap of the mean.

G) Same as in **B-F)** for H2AK119ub ChIP signal in MCF7 cells.

Building on this initial evidence of H3K79me2/3 as a potentially important player at TRC and R-loop sites, I was curious whether a global correlation between the presence of R-loops and H3K79me2/3 occupancy exists. Following peak calling on quantitative DRIP (qDRIP) data to define R-loop prone genomic regions in HeLa cells, I split the genome into R-loop forming and non-R-loop forming regions and quantified H3K79me2 and H3K79me3 signal in both groups. Crucially, both H3K79me2 and H3K79me3 showed high signal in R-loop forming regions when compared to the control set, further establishing a genome wide positive correlation between R-loops and H3K79me2/3 (**Fig. 35A**), which could also be observed at the R-loop forming example genes *ACTB* and *RPL13A (***Fig. 35B-C**). Collectively, these findings demonstrate that the TRC reporter system can recapitulate the behavioral dynamics of previously characterized TRC-related histone PTMs like H3K4me3 and H2AK119 and further provides crucial evidence for H3K79 methylation as a modification relevant to TRC and R-loop biology.

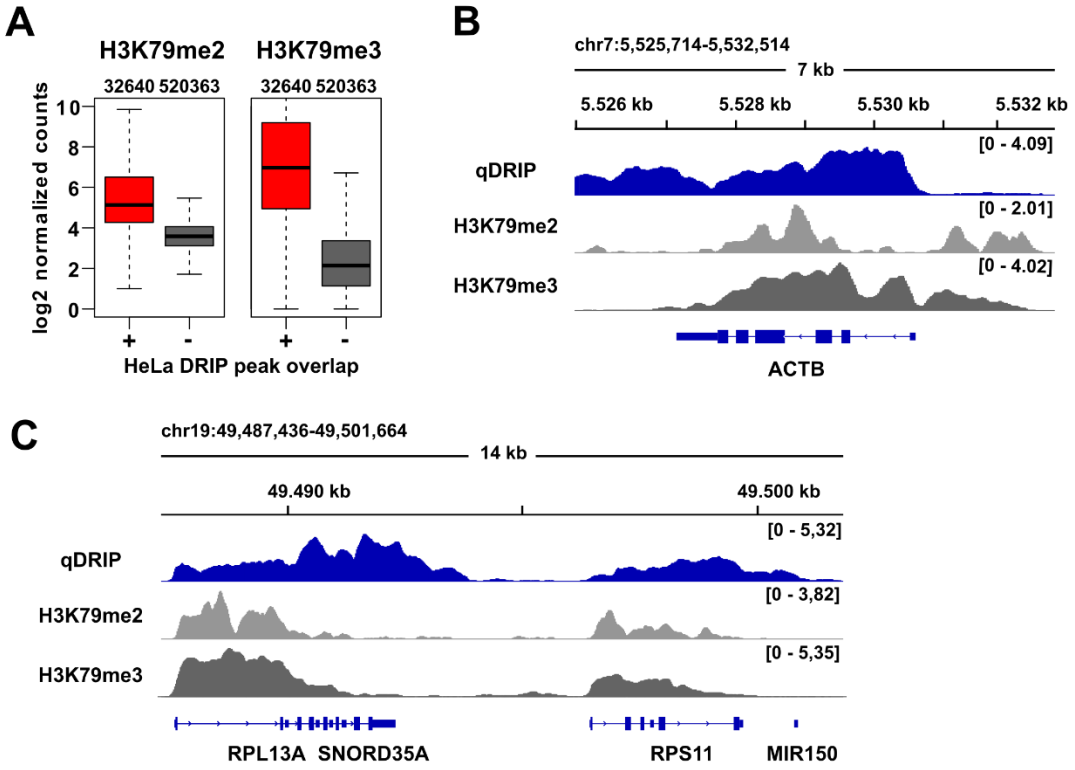

Figure 35: R-loop and H3K79 methylation overlap genome-wide

A) Box-plot comparing H3K79me2 and H3K79me3 ChIP signal at sites overlapping (+) or not overlapping (-) with R-loops. Bin size 5 kb. ´n´ indicates number of genomic bins.
B) Genome browser snapshot of a genomic example region with R-loop-prone ACTB locus from chr 7 showing overlapping qDRIP-seq, H3K79me2, and H3K79me3 ChIP-seq signal in HeLa cells.
C) Same as in **B)** for a similar region containing the R-loop forming RPL13A locus on chr 19.

4.7 Evaluation of the role of H3K79 methylation at TRC sites

Following the identification of H3K79 methylation as a strong candidate histone modification involved in the regulation of TRCs and R-loops, I was eager to elucidate whether this histone modification has any functional relevance at TRC sites. Importantly, DOT1L is the only known histone methyltransferase capable of depositing H3K79me2/3 (Nguyen & Zhang, 2011). Consequently, I inhibited DOT1L activity in the TRC reporter cell line using the small molecule inhibitor EPZ-5676 (Pinometostat, DOT1Li), which has been previously shown to be highly specific in reducing DOT1L activity (Daigle et al., 2013). Western blot analysis of cells treated with 5 µM DOT1Li for 72 h resulted in a global loss of about 40% of the H3K79me2 signal when compared to loading controls GAPDH and ORC2. In contrast, an acute DOT1Li treatment of 8 h was insufficient to alter H3K79me2 (**Fig. 36**). In agreement with existing literature that no active H3K79 demethylase has been identified to date, these results demonstrate that the drug treatment can successfully prevent the new deposition of the histone modification, but requires 2-3 cell division to dilute the pre-existing mark from total chromatin (Kari et al., 2019).

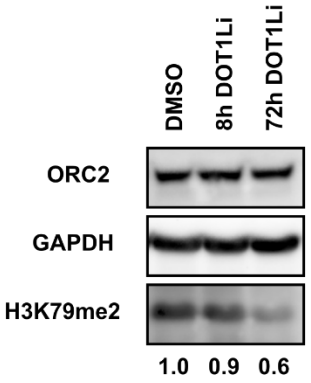

Figure 36: Prolonged DOT1Li inhibition reduces global H3K79me2 levels

Representative Western blot of H3K79me2 levels in cells treated with 5 µM DOT1L inhibition (EPZ-5676) or DMSO for 8 h or 72 h. GAPDH and ORC2 serve as loading controls. Quantifications of H3K79me2 signal intensities were normalized to ORC2 signal and are shown relative to the DMSO condition displayed below.

For specific analysis of H3K79me2 signal at the mAIRN reporter site, I next performed H3 normalized H3K79me2 ChIP. Expectedly, cells in DMSO conditions showed a clear DOX-dependent increase of H3K79me2 levels at the mAIRN locus but not at the unrelated NRXN2 or ACTB control sites (**Fig. 37A**). Crucially, both acute (8 h) and prolonged (72 h) treatment with DOT1Li entirely abolished the previously observed increase of H3K79me2 upon DOX (**Fig. 37A**). These results suggest a specific and DOX-dependent deposition of H3K79 methylation at the mAIRN reporter sequences via activated DOT1L. As the H3K79me2 ChIP-qPCR was conducted with histone H3 normalization, I wondered if DOT1L activity could also affect nucleosome occupancy and maybe promote the observed nucleosome loss at the TRC reporter site. Inhibition of DOT1L for both 8 h and 72 h did not prevent or alleviate the nucleosome loss phenotype at the mAIRN sites seen upon DOX induction (**Fig. 37B**), thus suggesting that the nucleosome loss occurs independently of H3K79 methylation and DOT1L action.

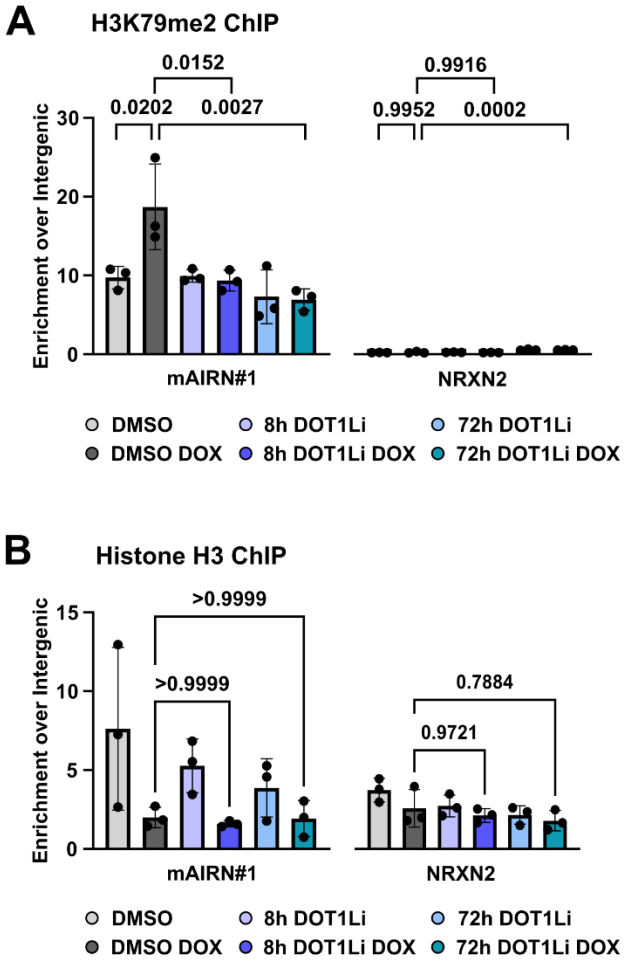

Figure 37: H3K79me2 is actively deposited at the induced TRC reporter sites

A) H3 normalized H3K79me2 ChIP-qPCR in cells treated with 5 µM DOT1L inhibition (EPZ-5676) for 8 h and 72 h. Cells were also treated with 0 or 1 µg/mL DOX. H3K79me2 levels were analyzed at the mAIRN#1 or the NRXN2 and ACTB control sites (n=3). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

B) Corresponding histone H3 ChIP for the mAIRN#1 and NRXN2 loci used for normalization in **A)**.

To obtain more evidence for the role of DOT1L in TRC biology, I conducted PLA for RNAPII Ser2P and PCNA to test whether DOT1Li can increase TRC frequency. Consistent with my previous results, a short 4 h DOX treatment did not cause changes in overall TRC PLA levels in DMSO conditions (**Fig. 38**). In contrast, DOT1Li for 72 h reduced global TRC burden in the basal state without TRC reporter activation. This is most likely the result of globally impaired transcription elongation and thus lowered levels of RNAPII Ser2P on chromatin, a phenotype previously reported for DOT1Li (Steger et al., 2008; Wu et al., 2021). Strikingly, combined treatment of 72 h DOT1Li and 4 h DOX induction caused an increase of TRC PLA foci number (**Fig. 38)**, implying that DOT1L activity is required for proper transcription-replication coordination and helps the cell to overcome acutely induced TRCs in our model reporter cell line.

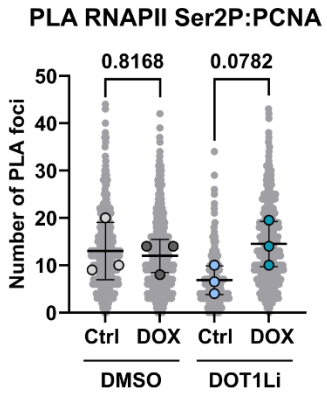

Figure 38: Combined DOT1L and DOX treatment exacerbates TRC levels in mAIRN reporter cells

Quantification of TRC PLA foci (RNAPII Ser2P and PCNA antibodies) in S-phase cells. Cells were treated with DMSO or DOT1Li for 72 h. Cells were also treated with 0 or 1 µg/mL DOX for 4 h (n=3). Bars indicate mean values with SD. Unpaired t-tests.

These data suggest that DOT1L is specifically recruited to the integrated reporter sites upon TRC induction. To gain more mechanistic insights into how this recruitment can occur specifically at the reporter sites, I considered the possibility that DOT1L might be passively recruited by continuous association with either the transcription or replication machinery. To test this, I

performed PLA assays using primary antibodies against DOT1L and the replication fork component PCNA. Interestingly, a robust number of 4-5 DOT1L-PCNA PLA foci could be detected particularly in EdU-positive S-phase cells, suggesting that DOT1L can associate with a specific fraction of active replication forks (**Fig. 39A, B**). The specificity of this PLA combination was verified via siRNA knockdown of DOT1L for 72 h (siDOT1L). 72 h siDOT1L treatment reduced DOT1L protein levels by 50 % compared to a non-targeting siRNA control (siCtrl) and strongly reduced the H3K79me2 signal by 80 % (**Fig. 39C**). Correspondingly, the number of DOT1L-PCNA PLA foci in siDOT1L conditions was significantly reduced when compared to siCtrl cells (**Fig. 39 D**). Moreover, single antibody controls for PCNA and DOT1L showed little to no detectable foci, indicating a low background signal (**Fig. 39D**). Most interestingly, TRC induction by DOX demonstrated a reduction of DOT1L-PCNA interaction (**Fig. 39A, B**), thereby indicating that DOT1L dissociates from active replication forks upon TRC induction. These data support the notion that DOT1L may be released from the replisome and therefore engage with replicating or post-replicative chromatin to deposit H3K79me2/me3.

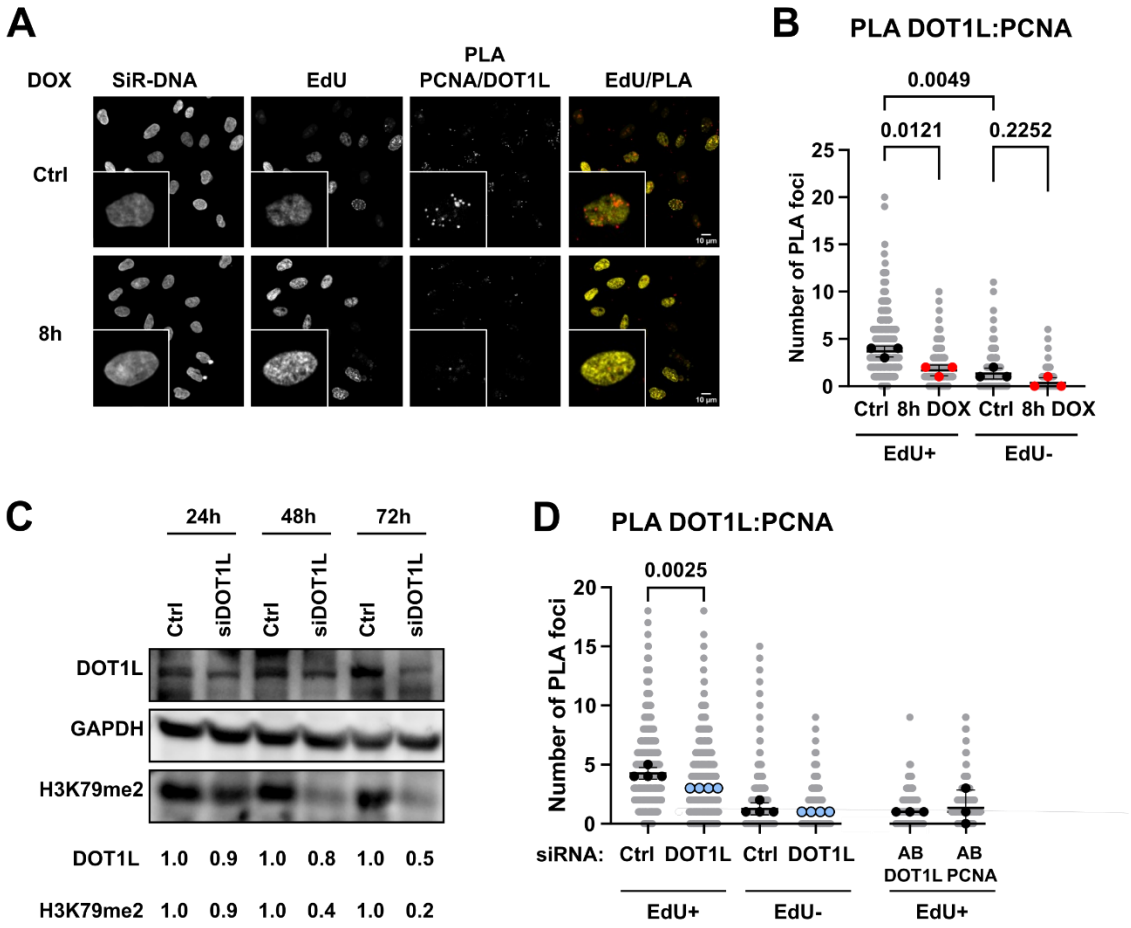

Figure 39: DOT1L dissociates from the replisome upon TRC induction

A) Representative images of PLA assay with DOT1L and PCNA antibodies. EdU click-it staining was performed to label S-phase cells. Cells were treated with 0 or 1 µg/mL DOX for TRC induction. Scale bar 10 µm.

B) Quantification of **B)** in EdU positive and negative cells (n=3). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

C) Representative Western Blot showing DOT1L and H3K79me2 signal upon knockdown of DOT1L (siDOT1L) for 24 h, 48 h, and 72 h compared to non-targeting siRNA control (Ctrl). GAPDH loading control. Quantification of GAPDH normalized DOT1L and H3K79me2 signal relative to the respective siCtrl treatment is shown below.

D) Quantification of PLA foci (DOT1L and PCNA antibodies) in EdU+ or EdU- cells. Cells were treated with siCtrl or siDOT1L (n=4). Single antibody controls for PCNA and DOT1L show the level of background signal (n=3). Bars indicate mean values with standard deviations (SD). Unpaired t-test.

Given the proposed role of H3K79me2/3 in aiding transcription elongation, I was wondering if DOT1L might act to rescue or restart transcription at TRC sites. This hypothesis is in agreement with two recent studies, providing evidence for rapid reengagement and transcription restart of RNAPII complexes on post-replicative chromatin (Bruno et al., 2024; Fenstermaker et al., 2023). For initial support of this hypothesis, I reanalyzed available ChIP-seq data of RNAPII, H3K79me2, and DOT1L from a leukemia cell line and overlapped their enrichment with the previously defined genomic HO and CD TRC regions (see **Fig. 34)**. Consistent with the data from HeLa cells, H3K79me2 again was strongly enriched at HO TRC sites (**Fig. 40A**). Strikingly, RNAPII and DOT1L followed a very similar pattern and showed a slight enrichment at the HO side of the plots, suggesting that DOT1L activity is likely required for efficient transcription elongation at such HO TRC regions (**Fig. 40B, C**).

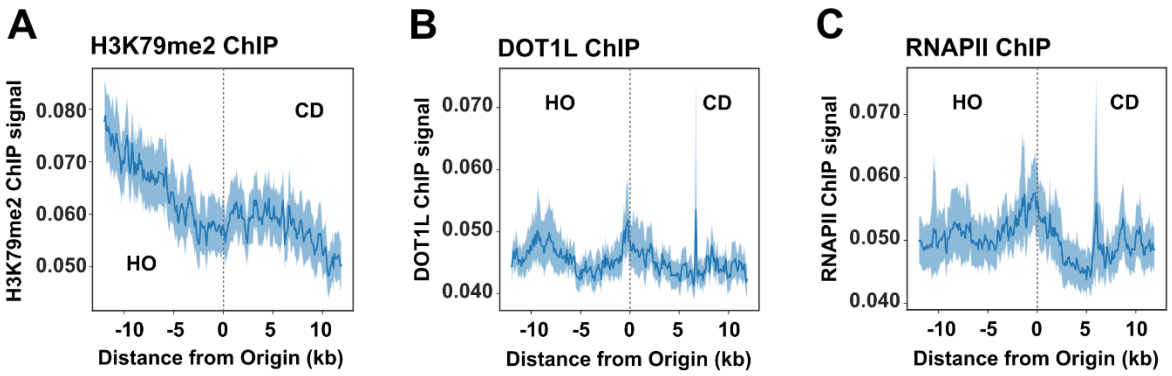

Figure 40: RNAPII and DOT1l are enriched at genome-wide TRC sites marked by H3K79me2

Analysis of H3K79me2 (**A**), DOT1L (**B**), and RNAPII ChIP signal (**C**) from MOLM13 cells at intragenic replication origins within active genes previously defined in HeLa cells. The analysis windows surrounding the regions span 24 kb in size and are positioned at least 5 kb away from promoters and terminators. Error bands show a 95 % confidence interval as determined by a bootstrap of the mean.

To further experimentally verify the role of DOT1L in RNAPII transcription at TRC sites, I tested the transcriptional output of the reporter gene upon DOT1Li. Crucially, acute local H3K79me2/3 depletion at the reporter site by 8 h DOT1Li significantly reduced the transcriptional output of the induced mAIRN gene as observed by RT-qPCR (**Fig. 41A**). Unrelated housekeeping genes PUM1 and ALAS1 were instead not affected (**Fig. 41A**). In further support, treatment of the reporter cells with a structurally distinct DOT1L inhibitor (EPZ004777) showed an identical small but significant decrease in mAIRN transcription output after a short 8 h DOT1Li treatment (**Fig. 41B**). Additionally, I also tested the effects of RNAi-mediated depletion of DOT1L on mAIRN transcription activity as an alternative to chemical inhibition. Despite only an approximately reduction of DOT1L levels by 50 % and H3K79me2 levels by 80 % following 72 h of knockdown (**Fig. 39C**), I could still observe a consistent significant reduction of mAIRN transcription levels upon DOX treatment (**Fig. 41C**). Unrelated housekeeping genes PUM1 and ALAS1 remained largely unaffected (**Fig. 41B, C**). Collectively, these data provide compelling evidence that DOT1L enzyme activity is required to maintain the full transcriptional potential of the mAIRN reporter gene upon TRC induction.

Finally, I asked if the presence of H3K79 methylation is required to mitigate DNA damage at TRC sites and thus performed FANCD2 ChIP-qPCR after TRC induction with and without DOT1Li treatment. Crucially, FANCD2 levels were further increased upon 8 h DOT1L inhibition at the induced mAIRN reporter but not the NRXN2 control locus (**Fig. 42**), indicating that the DNA damage inflicted at the TRC site is exacerbated without active H3K79 methylation deposition by DOT1L. In summary, these findings establish DOT1L as the crucial writer of H3K79me2/3 at the TRC reporter sites and suggest an active methylation deposition upon conflict induction. DOT1L is likely recruited to TRCs in a replisome dependent manner via PCNA interaction. Finally, DOT1L activity is crucial for maintenance of active transcription and prevention of excessive DNA damage at TRC sites.

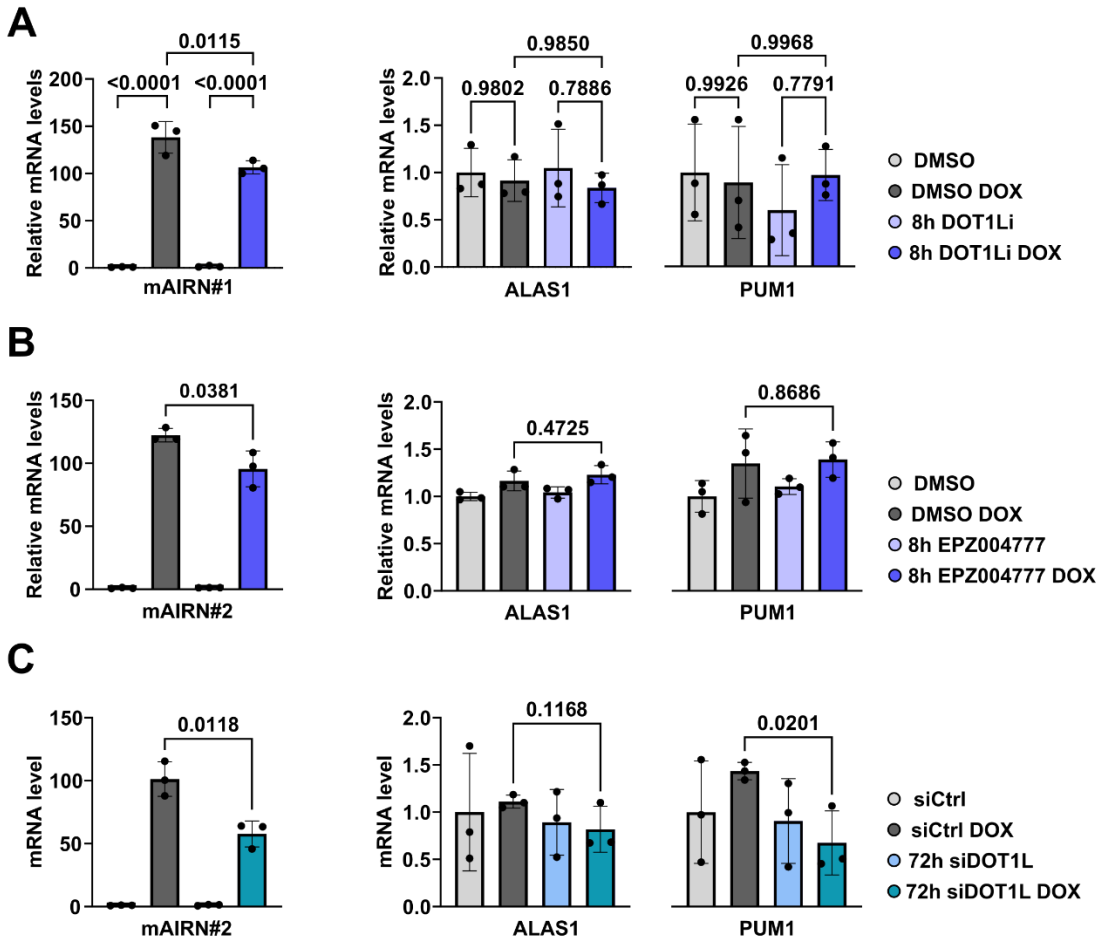

Figure 41: Interfering with DOT1L activity reduces the transcriptional output of the mAIRN reporter gene

A) RT-qPCR analysis of RNA levels using primer pairs mAIRN#1, ALAS1, and PUM1 in cells treated with 5 µM DOT1L inhibition (EPZ-5676) or DMSO control treatment. Moreover, 0 or 1 µg/mL DOX were added for 4 h (n=3). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

B) RT-qPCR analysis of RNA levels using primer pairs mAIRN#2, ALAS1 and PUM1 in cells treated with 5 µM DOT1L inhibition (EPZ000477) or DMSO control treatment. Moreover, 0 or 1 µg/mL DOX were added for 4 h (n=3). Error bars indicate SD. Student's t-test.

C) RT-qPCR analysis of RNA levels at mAIRN#2, ALAS1, and PUM1 in cells treated with siCtrl or siDOT1L for 72 h. Additionally, 0 or 1 µg/mL DOX were added for 4 h (n=3). Error bars indicate SD. Student's t-test.

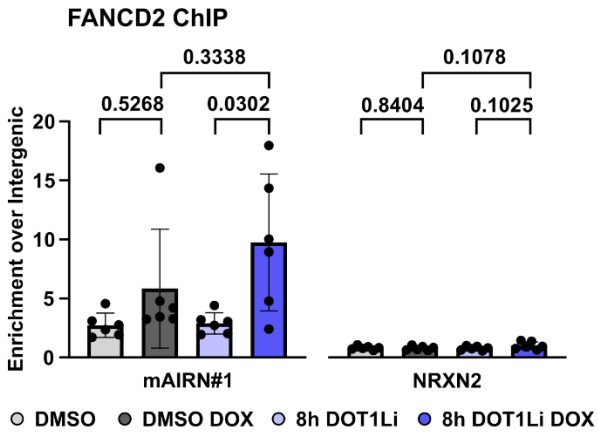

Figure 42: Loss of DOT1L activity exacerbates fork stalling and DNA damage at the TRC reporter sites

FANCD2 ChIP-qPCR analysis with primer pairs mAIRN#1 or NRXN2 in cells treated with 5 µM DOT1L inhibition (EPZ-5676) or DMSO control. In addition, cells were exposed to 0 or 1 µg/mL DOX for 24 h (n=3). Error bars indicate SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

4.8 Addressing the potential of oncogenes to induce TRCs

Evidence from numerous recent studies suggests that several commonly amplified or overexpressed oncogenes can cause TRC-driven genomic instability in various cancer cell models (Kotsantis et al., 2016; Macheret & Halazonetis, 2018). Nevertheless, a systematic and comprehensive side-by-side analysis of the overexpression effect of multiple oncogenes on TRC biology with consistent methodology has not yet been performed. To this end, I chose to determine the potential of the oncogenes Cyclin E, CDC25A, and CDC6 to induce TRCs and associated genomic instability.

4.8.1 Overexpression of Cyclin E in U 2-OS cells fails to reproducibly elevate TRC levels

The gene *CCNE1* encoding Cyclin E is among the most frequently amplified genes in human cancers (Beroukhim et al., 2010; Bignell et al., 2010; Zack et al., 2013). Overexpression of Cyclin E was shown to induce firing of novel replication origins within highly transcribed genes in U-2 OS osteosarcoma cells (Macheret & Halazonetis, 2018). These oncogene-induced replication initiation events were suggested to lead to collisions with ongoing transcription, thereby giving rise to an elevated TRC burden and DNA damage (Macheret & Halazonetis, 2018; Neelsen et al., 2013). While the original work proposed an occurrence of TRCs based on overlapping EU-seq and EdU-seq profiles as well as DRB-sensitivity of the DNA damage (Macheret & Halazonetis, 2018), direct experimental validation of an increased TRC burden has not been conducted. Consequently, I chose to address this gap in our understanding using a PLA-based readout. To begin, I tested the ability of U-2 OS Cyclin E Tet-OFF cells (Santoni-Rugiu et al., 2000) to induce Cyclin E expression upon removal of tetracycline by western blot analysis. Cyclin E was moderately (1.6-fold) overexpressed within 8 h and maintained high protein levels throughout the time course up to 72 h (**Fig. 43A**). Next, I performed PLA against RNAPII Ser2P and PCNA following Cyclin E overexpression. As novel origin firing was reported to be occurring within only a few hours after overexpression (Macheret & Halazonetis, 2018), I chose to test the number of TRC-PLA foci throughout a time course of 4 h, 8 h, and 24 h. Unexpectedly, Cyclin E overexpression could not increase the burden of TRC PLA foci at any of the observed timepoints (**Fig. 43B, C**). Interestingly, further analysis of EdU-signal in S-phase cells as a marker of replication efficiency showed strong albeit not significant reduction after 4 h and 8 h (**Fig. 43D**). In summary, these pilot experiments could not verify overexpressed Cyclin E as an efficient inducer of TRCs. This could be in part due the limited (1.6-fold) overexpression efficiency and a certain level of leakiness of the Tet-OFF system. Consequently, I did not further pursue the

investigation of Cyclin E in the TRC context but rather focused on the two other model systems overexpressing CDC25A and CDC6, respectively.

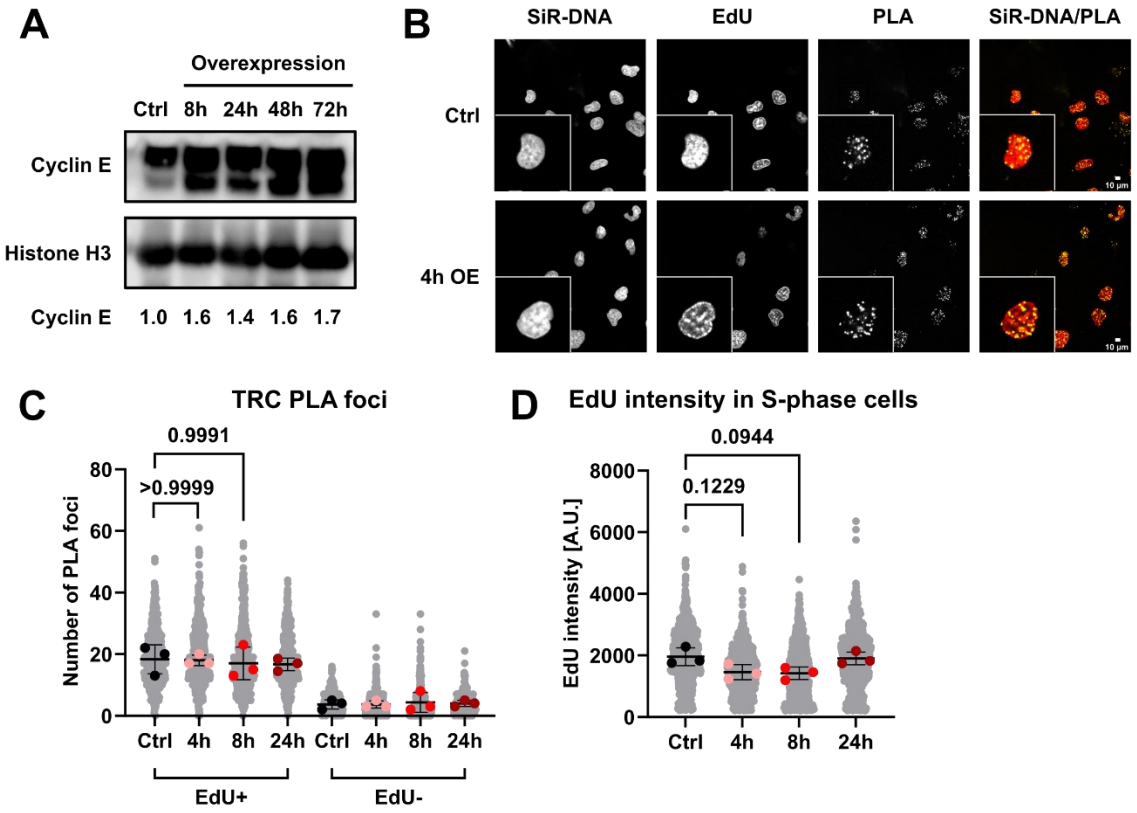

Figure 43: Cyclin E overexpression fails to increase global TRC levels in U-2 OS cells

A) Representative Western Blot images showing Cyclin E levels upon Cyclin E overexpression by removal of tetracycline from the culturing medium for 8 h, 24 h, 48 h, and 72 h compared to the control condition (Ctrl) without overexpression. Histone H3 serves as a loading control. Quantification of Cyclin E signal normalized to H3 is provided below.

B) Representative images of TRC PLA assay with RNAPII Ser2P and PCNA antibodies (Ctrl and 4 h time points). EdU click-it staining was performed to label S-phase cells. Cyclin E expression was induced by the removal of tetracycline from the culturing medium. Scale bar 10 µm.

C) Quantification of TRC PLA foci number in S-phase (EdU+) and non-S-phase cells (EdU-) from **B)** as well as additional time points (n=3). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

D) Quantification of mean EdU signal intensity in S-phase cells from **B)** as well as additional time points (n=3). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

4.8.2 Overexpression of CDC25A in U 2-OS cells causes TRCs and associated genomic instability

The dual specific-phosphatase CDC25A has been shown to remove inhibitory phosphorylation on Cyclin E/CDK2 and Cyclin B/CDK1 complexes, thereby promoting S-phase and M-phase entry, respectively (Donzelli & Draetta, 2003). Moreover, activation of the DNA damage response has been demonstrated to converge on CDC25A. Phosphorylation of CDC25A via CHK1 or CHK2 induces proteasomal degradation of CDC25A, thereby blocking S-phase entry and DNA replication (Donzelli & Draetta, 2003; Mailand et al., 2000). Importantly, overexpression of CDC25A enables it to bypass this mechanism, resulting in elevated DNA damage and fork stalling (Mailand et al., 2000; Neelsen et al., 2013). Given its role in DNA damage and convergence with the Cyclin E pathway (Donzelli & Draetta, 2003), I wondered if CDC25A overexpression could give rise to an elevated TRC burden. In analogy to the experiments for Cyclin E, I verified a successful overexpression in U-2 OS CDC25A Tet-OFF cells (Mailand et al., 2000) upon tetracycline removal. Interestingly, short overexpression for 8 h showcased an induction of 6.3-fold, followed by a moderate reduction across the time course of 24 h, 48 h, and 72 h (**Fig. 44A**). Strikingly, CDC25A significantly increased global TRC PLA foci levels in S-phase cells after 4 h overexpression (**Fig. 44B, C**). In contrast, subsequent time points at 8 h or 24 h did not show this phenotype as the TRC-PLA levels dropped back to similar levels as in the control samples without CDC25A overexpression, suggesting that CDC25A overexpression triggers an acute but rather transient response to increased TRC levels. Interestingly, quantifying the EdU-signal in S-phase cells as a marker of replication efficiency was not affected after 4 h of CDC25A overexpression but showed a moderate although non-significant decline at the later 8 h and 24 h timepoints, implying that the TRC response might translate into replication impairments in a sequential manner (**Fig. 44D**).

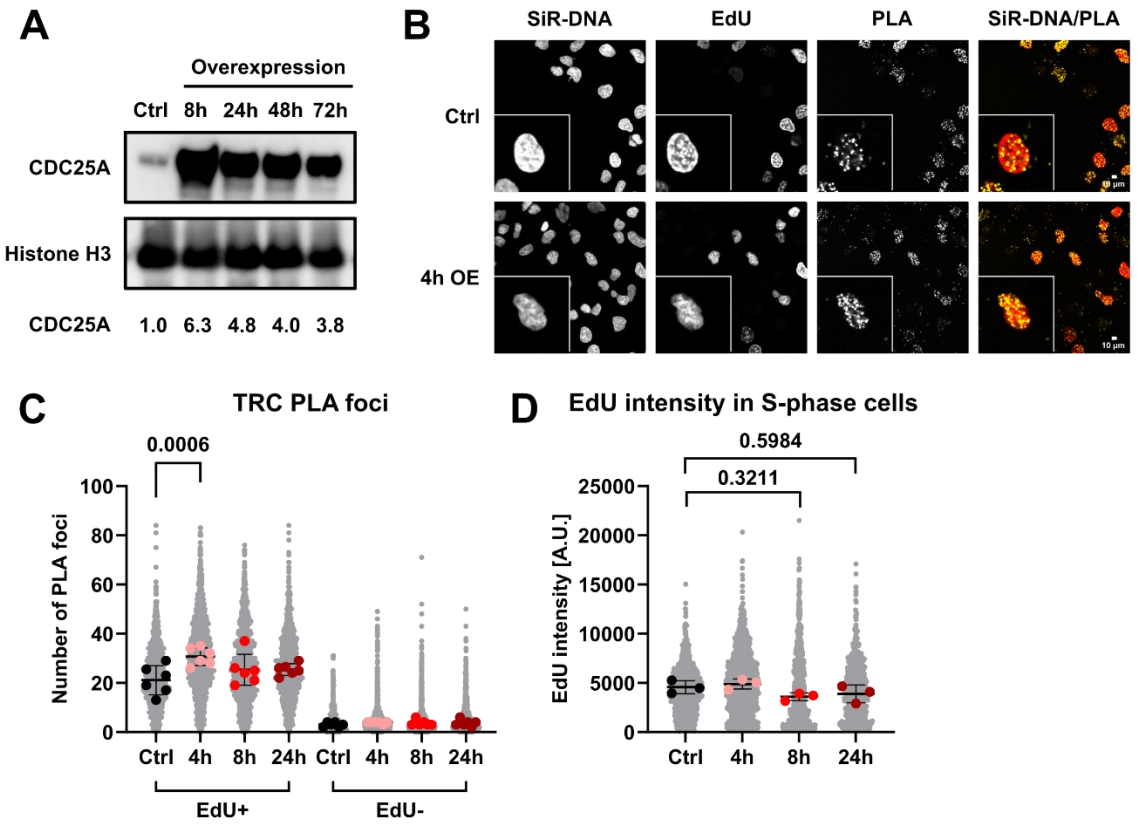

Figure 44: CDC25A overexpression shows an acute but transient increase of global TRC levels in U-2 OS cells

A) Representative Western Blot images showing CDC25A levels upon Cyclin E overexpression for 8 h, 24 h, 48 h, and 72 h by removal of tetracycline from the culturing medium compared to a control condition (Ctrl) without overexpression. Histone H3 serves as a loading control. Quantification of the CDC25A signal normalized to H3 is provided below.

B) Representative images of TRC PLA assay with RNAPII Ser2P and PCNA antibodies (Ctrl and 4 h time points). EdU click-it staining was performed to label S-phase cells. CDC25A expression was induced by the removal of tetracycline from the culturing medium. Scale bar 10 µm.

C) Quantification of TRC PLA foci number in S-phase (EdU+) and non-S-phase cells (EdU-) from **B)** as well as additional time points (n=3). Bars indicate mean values with SD. Ordinary one-way ANOVA with multiple comparison test.

D) Quantification of mean EdU signal intensity in S-phase cells from **B)** as well as additional time points (n=3). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

Next, I wanted to test if the increased TRC burden imposed by CDC25A would lead to elevated DNA damage and genomic instability. To this end, I conducted IF staining against the DNA damage marker γH2AX. Crucially, the γH2AX foci number and intensity were unchanged after 4 h of CDC25A overexpression, the time point with elevated TRC burden but significantly increased at the 8 h time point (**Fig. 45A-C**), suggesting a delayed DNA damage response consistent with

reduced EdU signal at this time point (**Fig. 44D**). Interestingly, γH2AX foci number reached back to baseline levels after 24 h of CDC25A overexpression, indicating that the cells were able to overcome or compensate the acute spike in TRC levels at 4 h and associated DNA damage at 8 h over the course of one cell cycle. In summary, these results identified the CDC25A overexpression model system in U-2 OS cells as a potent inducer of TRCs and genomic instability. Short overexpression resulted in the highest protein and TRC levels, implying that CDC25A overexpression rapidly induces discoordination between transcription and replication. Finally, elevating CDC25A levels lead to genomic instability as seen by increased γH2AX signal, which occurred slightly delayed in comparison to the spike in TRC-PLA foci. Further research will be required to understand the molecular mechanisms of how exactly CDC25A induces TRCs and DNA damage.

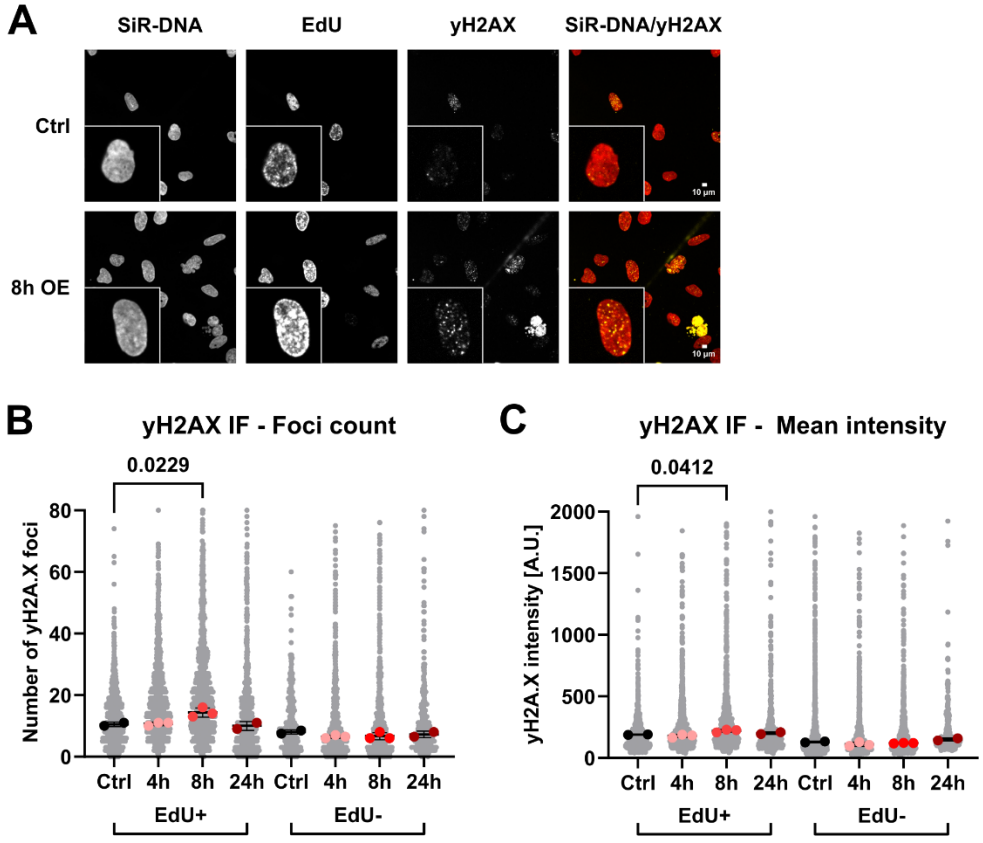

Figure 45: CDC25A overexpression increases DNA damage

A) Representative images of IF staining of γH2AX (Ctrl and 24 h time points). EdU click-it staining was performed to label S-phase cells. CDC25A expression was induced by the removal of tetracycline from the culturing medium. Scale bar 10 µm.

B) Quantification of γH2AX foci number in S-phase cells (EdU+) from **B)** as well as additional time points (n≥2). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

C) Same as in **B)** for γH2AX mean intensity in absolute units (A.U.).

4.8.3 Overexpression of CDC6 in Human Bronchial Epithelial Cells (HBEC) causes TRCs and associated genomic instability

CDC6 is an essential regulator of DNA replication in human cells primarily known for its function in the assembly of the pre-replicative complex at replication origins (Borlado & Méndez, 2008; Feng et al., 2021; Weissmann et al., 2024) (also see introduction). Additionally, CDC6 plays an important role in the regulation of the S-phase and M-phase checkpoints (Borlado & Méndez, 2008). CDC6 has been demonstrated to have proto-oncogenic function (Borlado & Méndez, 2008; Komseli et al., 2018). Particularly, a recently established model of human bronchial epithelial cells (HBEC) overexpressing CDC6 was demonstrated to undergo a short period of hyperreplication, which was accompanied by high levels of genomic instability (Komseli et al., 2018; Zampetidis et al., 2021). Subsequently, cells rapidly succumb to oncogene-induced replicative senescence. Following prolonged overexpression of CDCD6 of around 30-40 days, a subpopulation of cells was shown to escape from senescence, undergo a malignant oncogenic transformation, and display epithelial-mesenchymal transition phenotypes (Komseli et al., 2018; Zampetidis et al., 2021). Interestingly, the initially observed DNA damage during the hyperreplication phase was predominantly located in TSS regions, suggesting that TRCs could be a highly relevant contributor (Zampetidis et al., 2021).

To test the function of oncogenic CDC6 activity in HBECs, I first overexpressed CDC6 by the addition of DOX for 8 h, 24 h, 48 h, and 72 h and quantified its protein levels by western blot. CDC6 was clearly overexpressed at all observed time points (**Fig. 46A**). The strongest inducibility of CDC6 expression was seen after 8 h showing an induction of approximately 26-fold, which was followed by a gradual decline through the time course. Next, I conducted quantification of TRC PLA foci following 4 h, 8 h, and 24 h of CDC6 overexpression. Interestingly, the 4 h and 8 h time points both showed a moderate reduction of TRC PLA foci (**Fig. 46C**). In contrast, a strong increase of the TRC burden was observed after 24 h (**Fig. 46B, C**), suggesting that at least one cell division cycle is required to manifest the effects of CDC6 overexpression on transcription-replication coordination. Analysis of EdU signal as a marker of DNA replication activity also showed a drastic reduction after 24 h of CDC6 overexpression consistent with the expected induction of senescence (**Fig. 46D**). As the highest numbers of TRCs could be counted only after

24 h of CDC6 overexpression, I wondered if this TRC burden would increase further in a prolonged time course over the course of two or three additional cell cycles. Consequently, I performed TRC PLA following 24 h, 48 h, and 72 h of CDC6 overexpression. While the 24 h time point showed results highly similar to the previous time course analysis (**Fig. 46C, E**), the TRC burden increased even further in a time-dependent manner at 48 h and 72 h (**Fig. 46E**). This data suggests that cells resisting the induction of replicative senescence the longest, become highly susceptible to TRCs.

To test whether this subpopulation of cells highly prone to TRC induction also shows a concomitant increase in DNA damage, I performed IF staining against γH2AX. γH2AX foci number gradually increased in S-phase cells with the highest level at 48 h, providing evidence that TRC-prone cells experience high levels of replication stress, DNA damage, and genomic instability(**Fig. 47A, B**). Interestingly, this change was not reflected in the overall γH2AX intensity, suggesting that the level of DNA damage signaling at individual foci might be limited. Finally, I also wanted to confirm the high γH2AX levels by complementary Western blot analysis. To this end, I repeated the time course of CDC6 overexpression up to 48 h and analyzed the γH2AX signal in whole cell lysates. Interestingly, I only observed an increased γH2AX signal after 8 h of CDC6 overexpression and no further induction at 24 h and 48 h (**Fig. 47C**). One potential explanation for this discrepancy could be that the Western blot analysis included all cells independent of their cell cycle stage, whereas the IF analysis distinguished S-phase and non-S-phase cells. In the IF analysis the latter population showed much reduced γH2AX signal (**Fig. 47B**). Thus, the effect seen in S-phase cells is likely masked by non-S-phase cells that dominate the population as a result of senescence induction at the later timepoints. As part of this experiment, I also analyzed the levels of total RNAPII and RNAPII Ser2P by Western blot and found a strong increase in both levels at 8 h upon CDC6 overexpression. Subsequent time points displayed a decrease of both RNAPII and RNAPII Ser2P levels compared to 8 h but still remained above the baseline state (**Fig. 47C**). This initial spike of RNAPII levels might result in a hypertranscription phenotype that could explain the subsequent induction of TRCs seen in the PLA data.

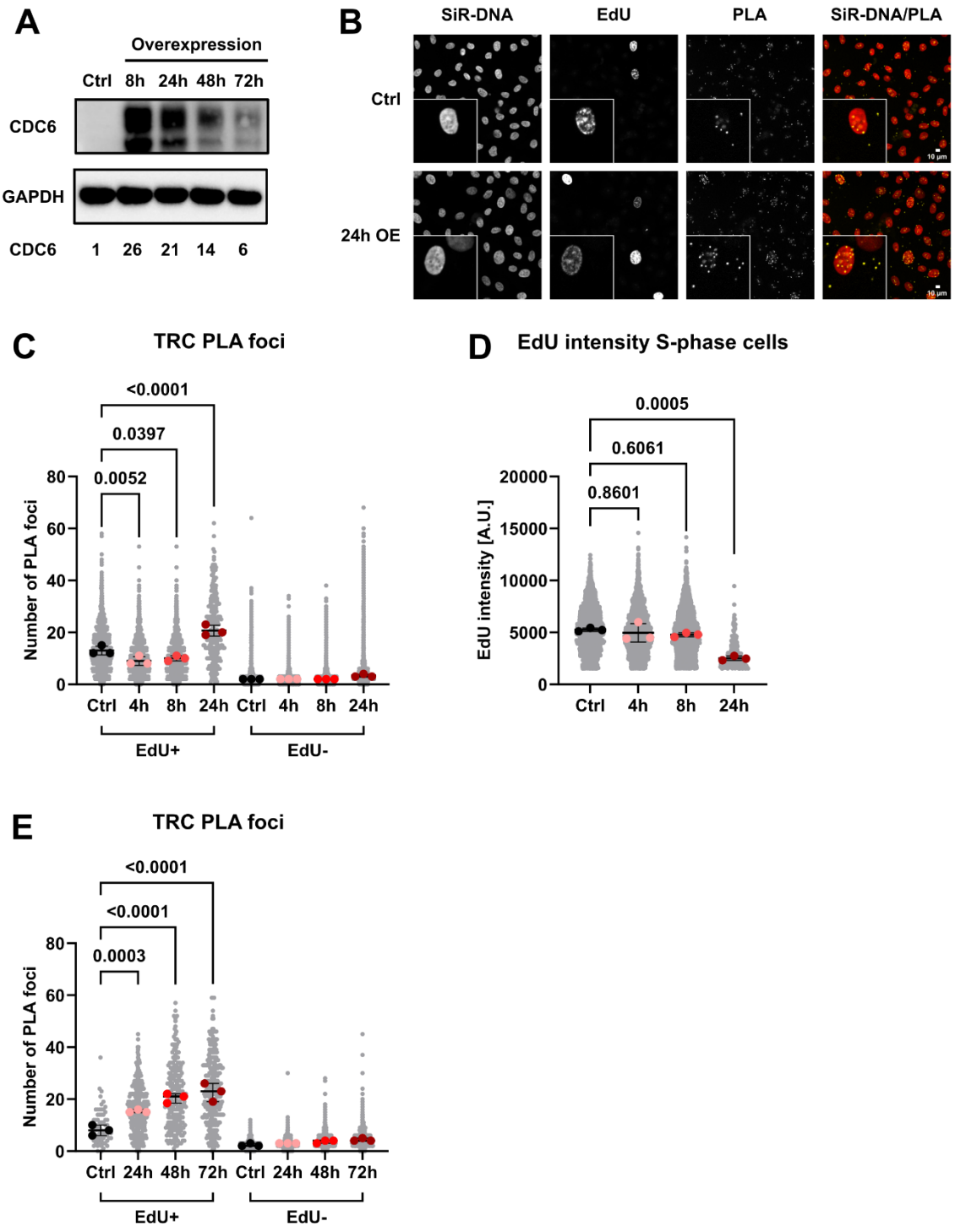

Figure 46: CDC6 overexpression increases global TRC levels in HBECs during the transition to senescence

A) Representative Western Blot images showing CDC6 levels upon CDC6 overexpression induced by the addition of DOX for 8 h, 24 h, 48 h, and 72 h compared to a control condition (Ctrl) without overexpression. GAPDH serves as a loading control. Quantification of CDC6 signal normalized to GAPDH is provided below.

B) Representative images of TRC PLA assay with RNAPII Ser2P and PCNA antibodies (Ctrl and 24 h time points). EdU click-it staining was performed to label S-phase cells. CDC6 expression was induced by the addition of DOX. Scale bar 10 µm.

C) Quantification of TRC PLA foci number in S-phase (EdU+) and non-S-phase cells (EdU-) from **B)** as well as additional time points (n=3). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

D) Quantification of mean EdU signal intensity in S-phase cells from **B)** as well as additional time points (n=3). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

E) Same as in **C)** for cells overexpressing CDC6 for 24 h, 48 h, and 72 h (n=3)

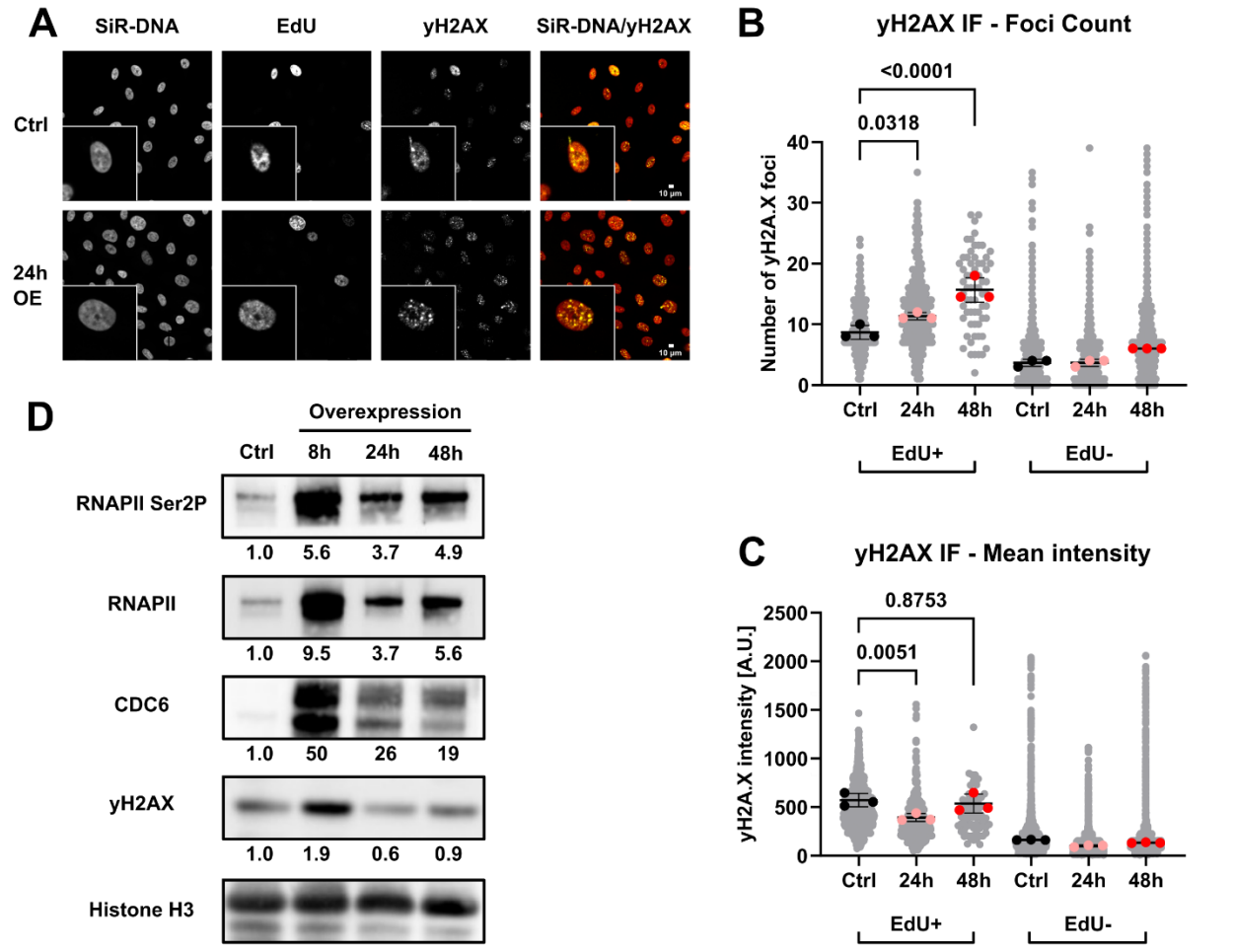

Figure 47: CDC6 overexpression in HBECs increases DNA damage and RNAPII levels

A) Representative images of IF staining of γH2AX (Ctrl and 24 h time points). EdU click-it staining was performed to label S-phase cells. CDC6 expression was induced by the addition of DOX. Scale bar 10 µm.

B) Quantification of γH2AX foci number in S-phase cells (EdU+) and non-S-phase cells (EdU-) from **A)** as well as additional time points (n=3). Bars indicate mean values with SD. Ordinary one-way ANOVA with Tukey's multiple comparison test.

C) Same as in **B)** for γH2AX mean intensity in absolute units (A.U.).

D) Representative Western Blot images showing RNAPII Ser2P, RNAPII, CDC6, and γH2AX levels upon overexpression of CDC6 for 8 h, 24 h, and 48 h compared to a control condition without overexpression. Histone H3 serves as the loading control. Quantifications of RNAPII Ser2P, RNAPII, CDC6, and γH2AX signals relative to H3 are provided below the respective bands.

Taken together, I have demonstrated that different oncogenes have varying capabilities to induce TRCs and associated genomic instability. Surprisingly, Cyclin E overexpression that was previously connected to TRC induction did not demonstrate an elevated TRC burden based on the TRC PLA assay. Instead, CDC25A and CDC6 overexpression induced TRCs and DNA damage, albeit with vastly different extend and timing, highlighting the demand for more comprehensive and systematic analysis when studying oncogene-driven TRCs.

5. Discussion

In this work, a novel system to robustly induce TRCs in human cancer cells at defined chromosomal sites was developed and used to analyze the impact of TRCs on the local chromatin environment. Strikingly, I observed a TRC-driven loss of nucleosome occupancy at the reporter sites that coincided with local R-loop formation and DNA replication fork impairment. As a result, I detected site-specific induction of DNA damage marked by FANCD2 but also a global DNA damage response as seen by the increased number of FANCD2 foci and γH2AX signal throughout the whole nucleus in these cells.

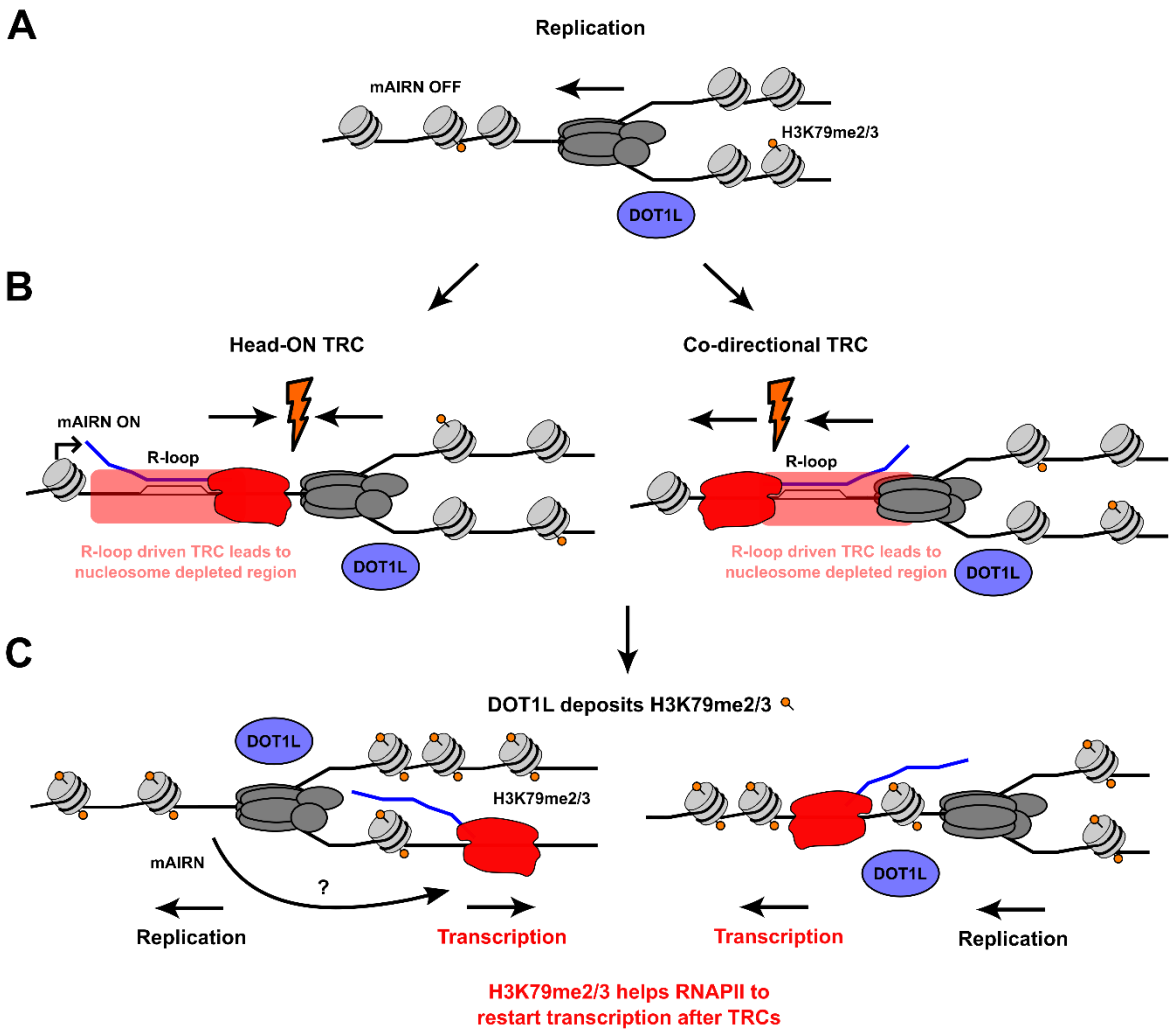

Figure 48: Model for the functional role of H3K79 methylation at TRC sites

A) Cells not transcribing the mAIRN reporter (OFF) can replicate normally without disruption of chromatin organization.

B) Upon transcriptional activation, the mAIRN reporter (ON) creates R-loops and impairs replication fork progression. The resulting HO or CD TRCs with an associated R-loop cause a local reduction of nucleosome occupancy. At the same time, DOT1L dissociates from the replication machinery and deposits H3K79me2/3.

C) H3K79me2/3 aids efficient transcription recovery at the mAIRN reporter site. Figure from (Werner et al., 2025)

Moreover, screening of potential TRC-associated histone modifications confirmed the enrichment of chromatin marks previously connected to TRCs but also identified H3K79 methylation as a novel and important histone modification at TRC sites (**Fig. 48**). Perturbing the deposition of H3K79me2/3 by DOT1L inhibition led to an exacerbated DNA damage response and impaired efficient transcription progression at the reporter site. Collectively, my work provides mechanistic insight into how TRCs and associated R-loops threaten the local chromatin landscape through disrupted nucleosome occupancy and DNA replication impairments, while also defining H3K79me2/3 as a crucial histone modification at chromosomal TRC sites.

5.1 Persistent R-loop formation at TRC sites is incompatible with nucleosome incorporation on the DNA

In this work, multiple different approaches and model systems have shown that R-loop formation on the mAIRN reporter sequence is incompatible with nucleosome incorporation on the resulting three-stranded nucleic acid structure. First, *in vitro* data showed that reconstituted mAIRN RNA:DNA hybrids are resistant to nucleosome formation when competing with dsDNA (**Fig. 8B**). Second, HO-TRC but not CD-TRC plasmids were subject to a profound loss of nucleosome occupancy seen in both MNase and histone H3 ChIP-qPCR assays (**Fig. 9, 10A**). Importantly, it was shown that only HO-TRC plasmids are capable of stable R-loop formation at the mAIRN locus, implying that high R-loop stability might underlie this defect in nucleosome incorporation (Hamperl et al., 2017). Third, the activated chromosomal mAIRN reporter was subject to significant H3 loss. While this phenotype occurred in both G1 and S-phase cells, it was significantly more pronounced in S-phase cells correlating with their higher R-loop burden (**Fig.26B, 32C**). These lines of evidence suggest that TRCs, and especially more detrimental HO-TRCs, can lead to the formation of stable R-loops that disrupt nucleosome organization. Importantly, nucleosome depletion was exclusively observed at the R-loop forming TRC reporter

sites but not at adjacent genomic regions that show readthrough transcription with similar levels as the mAIRN reporter itself (**Fig. 29B, 31C**), thereby demonstrating that the phenotype is not a cause of transcription activation but additionally requires an R-loop prone sequence.

Notably, overexpression of human RNase H1 was not able to rescue the histone H3 loss and FANCD2 accumulation phenotypes at the mAIRN loci (**Fig. 27C**). Although these results contradict the known function of RNase H1 as a *bona fide* R-loop removing enzyme (Hyjek et al., 2019; H. Zhao et al., 2018), I do not attribute the observations to a lack of R-loop dependency but rather to technical limitations of reporter activation and RNase H1 overexpression that prevented successful R-loop removal. In my experimental design, RNase H1 was overexpressed for a period of 24 h, during which the transcription of the mAIRN reporter was continuously activated by DOX treatment to induce R-loop formation. In contrast to treatment with exogenous *E. coli* RNase H (**Fig. 21A**), endogenously overexpressed human RNase H1 may not efficiently remove or degrade the exceptionally stable R-loops on the mAIRN sequence. Thus, reporter activation would lead to the persistent formation of R-loops and create a steady-state equilibrium, in which RNase H1 removes the R-loop but its turnover remains insufficient to efficiently counteract the continuous DOX-induced R-loop formation. Indeed, experiments from collaborators in the Vincent Vanoosthuyse lab showed that RNase H overexpression cannot reduce R-loop levels at the mAIRN reporter cell line (data not shown). Additionally, the mAIRN sequence contains a very high GC-skew, a feature that correlates with R-loop formation but has also been found in RNase H-resistant R-loops (Crossley et al., 2020). Finally, the mAIRN locus was also shown to induce distinct but yet-to-be-defined other DNA secondary structures on the template strand in an RNase H-independent manner *in vitro* . These DNA secondary structures could contribute to the replication fork stalling and nucleosome depletion phenotypes. To distinguish these possibilities in future experiments, point mutations that were previously described to impair the formation of these DNA secondary structures, but not R-loop formation at the mAIRN sequence could be introduced into the reporter. Finally, the construction of TRC and R-loop reporter systems with various R-loop-prone sequences would enable a more comprehensive and comparative analysis.As an independent approach to RNase H overexpression, we leveraged a reporter cell line overexpressing a chromosomally integrated ECFP sequence with highly comparable copy number and transcriptional output as the mAIRN reporter cell line (**Fig. 15A, B**). Importantly, the ECFP sequence does not permit R-loop formation (Hamperl et al., 2017). ECFP activation did not result in H3 loss or DNA damage accumulation (**Fig. 15C, 28A**), providing independent evidence that the observed phenotypes are likely a cause of R-loops at the mAIRN TRC reporter loci. My observations are consistent with early observations describing that RNA presence in double-

helical DNA prevents its interaction with histone proteins (Dunn & Griffith, 1980). Furthermore, previous structural characterization of R-loops had shown that RNA:DNA hybrids adopt an intermediate conformation between A-DNA and B-DNA (J. H. Liu et al., 2019). The resulting atypical secondary structure is less flexible and thereby prevented from wrapping around the histone octamer, providing a possible explanation for the nucleosome loss. In analogy to the exposure of ssDNA in the non-template strand at R-loops (García-Muse & Aguilera, 2019), reduced protection of the DNA through nucleosome loss and thereby elevated accessibility could offer another layer of explanation as to why R-loop and TRC regions are more prone to DNA damage (Brambilla et al., 2020). In summary, my work establishes a closely connected relationship between TRCs, R-loop formation, and chromatin disruption, while also emphasizing the need for proper R-loop resolution and removal to maintain genomic and epigenomic integrity. In future work, genome-wide R-loop stabilization approaches, such as RNase H1/2 inactivation or topoisomerase inhibition (P. Lin et al., 2023; Y. Liu et al., 2020), will be crucial to extrapolate our findings from the reporter system to a global genome-scale, thereby offering a chance for unprecedented insights into TRC, R-loop, and chromatin biology.

5.2 Genomic integration of an inducible mAIRN R-loop sequence allows locus-specific TRC analysis

Previous studies in prokaryotic and eukaryotic model systems have characterized TRCs as dangerous genomic events giving rise to mutations and DNA damage, thereby threatening genome integrity (García-Rubio et al., 2018; Hamperl et al., 2017; Lang et al., 2017; Prado & Aguilera, 2005; Sankar et al., 2016). Despite many efforts, TRCs remain a particularly difficult genomic event to study given their stochastic occurrence and dependence on two highly complex nuclear machineries. Hence, researchers have commonly employed plasmid-based reporter systems to model the dynamics of TRCs in bacteria, yeast, and mammalian cells (Hamperl et al., 2017; Lang et al., 2017; Prado & Aguilera, 2005). Additionally, these studies are complemented by correlative approaches, in which drug treatments (Shao et al., 2020) or oncogene overexpression are used to generate DNA damage that is then correlated to co-occurring TRCs (Kotsantis et al., 2016; Macheret & Halazonetis, 2018). Despite their value in helping to understand numerous TRC-associated processes, plasmid reporter systems may not capture the full complexity and dynamics of TRC induction and resolution on the genome. Particularly, chromatin-related processes that are highly relevant for the control of both transcription and replication in complex mammalian genomes cannot be accurately modeled with plasmids. On the

contrary, global TRC induction approaches rely on disrupting cellular processes on a global scale such as topoisomerase depletion or oncogenes helping to bypass the G1/S cell cycle checkpoints (Kotsantis et al., 2016; Macheret & Halazonetis, 2018; Promonet et al., 2020). These global disruptions induce TRCs with varying efficiencies but are unavoidably accompanied by unspecific responses and side effects, making it impossible to link cellular responses such as DNA damage to TRCs exclusively. In summary, these problems prevent true mechanistic insight into how TRCs affect chromatin and get resolved.

To overcome these limitations, this work constructed a chromosomal TRC reporter system by genomic integration of the strongly R-loop forming mAIRN sequence that had been used in the human TRC plasmid reporters before (**Fig. 11**). In contrast to the plasmid system (**Fig. 9**), I chose not to integrate the unidirectional origin of replication oriP/EBNA due to its late and inefficient replication activity in human cells (Carroll et al., 1991; J. Zhou et al., 2009). Endogenous replication origins would likely be activated before oriP, thereby already replicating the integrated sequence and obscuring the orientational information compared to plasmid reporters. R-loop formation has been shown to slow down RNAPII progression (Tous & Aguilera, 2007). Thus, I speculated that stable R-loop formation on the mAIRN sequence would impede RNAPII progression sufficiently to create a potent obstacle for incoming replication forks, ultimately resulting in a TRC. For chromosomal integration of the mAIRN reporter sequence, I chose the Sleeping Beauty transposase due to its ability to generate integrations at multiple random locations (Izsvák & Ivics, 2004; Kowarz et al., 2015). Sleeping Beauty transposase integrates sequences into the genome with a bias to euchromatic and actively transcribed regions (G. Liu et al., 2005; Miskey et al., 2022), which are likely the sites that predominantly encounter TRCs. Subsequent analysis allowed me to detect the signal from an average of five integration sites in the monoclonal mAIRN Clone#12 cell line (**Fig. 12A, B**), providing robustness and reducing locus-specific biases. While the current TRC reporter system remains agnostic to TRC orientation, site-specific single integration via CRISPR/Cas9 could position the reporter in a specific orientation next to a replication initiation zone that shows highly robust replication fork directionality (Bayona-Feliu et al., 2023; Petryk et al., 2016), thus inducing HO or CD TRC selectively. Taken together, the chromosomal TRC reporter cells offer a controlled environment for TRC induction residing in endogenous chromatin and thereby serving as a powerful tool for future mechanistic investigations of TRCs and their interplay with R-loops, mutational burden, and (epi)genome stability.

5.3 H3K79 methylation is deposited at TRC sites and provides a chromatin environment for effective transcription recovery

Numerous histone modifications including H3K4me3 and H2AK119ub have been recently proposed to play a role in TRC biology (Chong et al., 2020; Hao et al., 2022). By investigating multiple candidates to identify TRC-associated histone modifications, I could confirm enrichment of both H3K4me3 and H2AK119ub at our reporter sites (**Fig. 33A, B**). This suggests that the reporter system can successfully recapitulate the chromatin dynamics previously observed at endogenous TRC sites. Most interestingly, I also detected a specific enrichment of H3K79me2 and H3K79me3 upon TRC induction at the reporter as well as genomic regions prone to R-loop forming HO-TRCs (**Fig. 33C, D and 34E, F**). H3K79me2/3 is deposited by DOT1L, the only known writer enzyme for this modification in mammalian cells (Daigle et al., 2013; Nguyen & Zhang, 2011). H3K79me2/3 has primarily been described in the context of RNAPII transcription serving as a mark of RNAPII elongation complexes consistent with its location in the gene body of actively transcribed genes (Steger et al., 2008; Wu et al., 2021). Nevertheless, the exact function of the modification in transcription regulation remains unknown. Furthermore, H3K79me2/3 has been linked to the DNA damage response with H3K79me3 serving as a binding platform for the Tudor domain of 53BP1 (Huyen et al., 2004), thereby aiding in the response and repair of DNA double-strand breaks (DSBs). Recent work has also identified Menin as a reader of H3K79me2 (J. Lin et al., 2023). Menin binding was proposed to be involved in transcriptional activation through binding to H3K79me2 at potential intragenic enhancers (J. Lin et al., 2023). Beyond its putative role in transcription regulation, Menin was previously shown to interact with FANCD2, a marker of fork stalling and DNA damage (Jin et al., 2003), thereby linking the functions of H3K79me2/3 in transcription and DNA damage control. Finally, H3K79me2/3 was shown to interact with PCNA, implying a role in DNA replication fork progression (Kang et al., 2020). Taken together, H3K79me2/3 appears as a prime candidate for a TRC-associated chromatin mark bringing transcription and replication control to prevent and/or overcome TRCs. Crucially, our study shows that H3K79me2 is actively deposited at TRC sites (**Fig. 37A**), suggesting a dynamic behavior of the modification and its writer DOT1L at TRC sites. This response might be distinct from globally deposited H3K79me2/3 that serves as a processivity signal for transcription activation and elongation. Notably, acute inhibition of DOT1L reduces the RNA output of the mAIRN reporter gene (**Fig. 41**) and results in increased fork stalling and DNA damage (**Fig. 42**). Moreover, H3K79me2/3 marks were enriched at sites of R-loop formation genome-wide (**Fig. 35**),

implying a role of the chromatin mark in transcription maintenance or recovery at genomic sites prone to R-loop formation, replication fork stalling, and DNA damage.

While my work uncovered H3K79me2/3 as a crucial histone modification being deposited at R-loop forming TRC sites, it remains unclear how exactly DOT1L is initially recruited to TRC sites. My PLA data showing an interaction of DOT1L with PCNA points towards DOT1L as a protein associated or potentially traveling with the replication fork (**Fig. 39**). Following TRC occurrence, DOT1L may either have increased residence time on the chromatin in proximity of the TRC site due to the stalling of the replication fork or alternatively, fork stalling may introduce a conformational change and remodel the fork that allows displacement of DOT1L from the replisome and deposition of H3K79me2/3 (**Fig. 37 and 39**). While my results suggests replication-dependent recruitment of DOT1L, recent structural studies have identified the transcription-linked ubiquitination of histone H2B at lysine 120 (H2BK120ub) as a recruitment platform for DOT1L (Valencia-Sánchez et al., 2019). Furthermore, H2BK120ub writer RNF20/40 E3 ubiquitin-protein ligase complex (RNF20/40) has been shown to play crucial roles in the DNA damage repair (So et al., 2019) and H3K79me2/3 levels are reduced upon their depletion (E. Wang et al., 2013). Consequently, a transcription and DNA damage-based pathway via RNF20/40 and H2BK120ub could serve as a second and complementary axis of DOT1L recruitment and H3K79me2/3 deposition at sites of transcription-dependent DNA damage such as TRCs. Converging pathways would offer additional robustness for the cell to ensure TRC-dependent DNA damage is repaired appropriately. Independent of the preceding signaling, Menin as a reader of H3K79me2/3 could then recruit FANCD2 to sites of transcription and replication-dependent DNA damage (J. Lin et al., 2023), stalling replication forks to enable proper repair and transcription recovery. Nevertheless, this putative pathway is in conflict with the observed increased FANCD2 recruitment upon DOT1Li (**Fig. 42**), demanding for more time-resolved analysis. Future research will be needed to clarify the exact molecular mechanisms and pathways of H3K79me2/3 regulation at TRC sites. Collectively, my study sheds light on a novel role of H3K79me2/3 in TRC biology, which likely acts as a platform to recruit repair factors and enable TRC resolution by aiding the restart of transcription after collisions between transcription and replication machineries (**Fig. 48**).

5.4 TRCs impair DNA replication fork progression and sensitize the cells to ATR kinase inhibition

TRCs and R-loops have been demonstrated to act as obstacles to DNA replication fork progression, thereby inducing replication stress in different model systems (Stoy et al., 2023; Tsirkas et al., 2022). Importantly, the chromosomal TRC reporter system allowed me to show local TRC-dependent replication fork progression impairments on endogenous mammalian chromatin (**Fig. 24**). Stalled DNA replication at the induced TRC sites precisely matched the expected replication timing of the superordinate replication domain (**Fig. 23B**). In contrast, no replication impairments outside of the TRC reporter could be detected (**Fig. 24A**). Based on these findings, I conclude that TRCs are no global disruptors of the replication timing program but rather induce highly localized delays of replication fork progression. As most of the integration sites are located in early replication domains (**Fig. 23B**), the replication delay was strongest at the earliest 2 h S-phase time point (**Fig. 24B, C**). Interestingly, slightly elevated replication signal at the subsequent 4 h and 6 h time points suggests that the conflicts can eventually be overcome, and the sequence is replicated (**Fig. 25**). This delay of DNA replication to a later cell cycle stage is reminiscent of recently described sites of G2/M DNA synthesis (G-MiDS) (J. Wang et al., 2021). G-MiDS are highly transcribed throughout S-phase and require the removal of RNAPII in G2/M-phase to complete DNA replication, particularly at their TSS regions. Notably, however, G-MiDS sites are significantly smaller, up to several hundred base pairs in size, and do not exhibit any detectable levels of DNA damage (J. Wang et al., 2021). In contrast, the TRC reporter sites are larger (approximately 5 kb) (**Table S1**), appear to be resolved during S-phase (**Fig. 24 and 25**), and induce local and global DNA damage (**Fig. 14 and 16**), implying that they act as more potent disruptors of the replication program compared to the natural occurring G-MiDS. While sharing similarities with respect to locally delayed DNA replication, the underlying resolution mechanisms at induced TRCs appear to be fundamentally different from G-MiDS as they involve the activation of the DNA damage checkpoints. However, in future work, it will be important to place TRC reporters into late replicating domains to address whether TRC-induced replication delays can also lead to incomplete DNA synthesis in G2/M-phase.

In agreement with the identified local replication delay, I found that persistent DOX treatment across multiple cell divisions leads to mild cell growth impairment (**Fig. 19**). Strikingly, this proliferation defect was strongly exacerbated upon treatment with an ATR inhibitor, highlighting the crucial role of the ATR kinase in the TRC response (**Fig. 19**). A previous study had connected activated ATR signaling to the more detrimental HO conflicts, whereas CD conflicts responded by

activation of the ATM kinase (Hamperl et al., 2017). In contrast to ATR, ATM inhibition did not show a synergistic reduction of cell proliferation when combined with TRC induction (**Fig. 19**). Based on these findings, it is tempting to speculate that particularly HO-TRCs give rise to replication delays and act as the main drivers of the ATR inhibition-dependent sensitivity in the reporter cells. Importantly, several ATR inhibitors have shown promising outcomes as anti-cancer drugs in clinical trials (Salguero et al., 2023). Thus, a stratification of patients based on the TRC burden of their tumors could further improve the efficiency of ATR inhibitor treatment in cancer therapy.

5.5 Oncogenic TRCs are associated with genomic instability and cell state transitions

Previous studies have shown that overexpression of numerous oncogenes can increase the TRC burden in human cells. Particularly, inducible overexpression of oncogenes *CCNE1* and *MYC* was shown to induce firing of novel replication origins within highly transcribed genes in U-2 OS osteosarcoma cells (Macheret & Halazonetis, 2018). These oncogene-induced replication initiation events were shown to result in higher TRC levels and associated genomic instability. For the oncogenic phosphatase CDC25A (Neelsen et al., 2013) and proto-oncogene CDC6 (Zampetidis et al., 2021), S-phase or transcription-associated replication stress have been described but a direct connection to TRCs has not yet been established. In this work, I have performed a side-by-side comparison of overexpression of the oncogenes Cyclin E, CDC25A, and CDC6 to confirm/determine their potential to induce TRCs and associated DNA damage.

Using the identical U-2 OS Cyclin E overexpression model system that has previously been described to give rise to TRCs at cryptically activated intragenic origins (Macheret & Halazonetis, 2018; Neelsen et al., 2013), I could confirm successful Cyclin E overexpression (**Fig. 43A**). However, Cyclin E overexpression was clearly less efficient compared to the other oncogenes. Strikingly and in contrast to published findings, Cyclin E overexpression was unable to increase the number of TRC PLA foci across numerous tested time points (**Fig. 43B, C**). Nevertheless, Cyclin E overexpression resulted in impaired DNA replication (**Fig. 43D**), suggesting that the firing of intragenic origins causes replication stress. Importantly, the original study never quantified the TRC burden by PLA but showed that fork collapse at oncogene-induced intragenic origins correlated with high transcription levels and was rescued by transcription inhibition with DRB (Macheret & Halazonetis, 2018). While these genomic data strongly point towards transcription-dependent replication stress, the conclusion is based on corelative evidence rather than direct

quantification of TRC levels. The PLA-based analysis of TRC remains the only currently available technique to directly visualize and quantify TRCs, as it detects events in which actively elongating RNAPII (RNAPII Ser2P) and replication fork component PCNA are within approximately 40 nm distance (Hegazy et al., 2020). Interestingly, an independent study used Cyclin E overexpression in synchronized RPE-1 cells could demonstrate increased levels of RNAPII-PCNA TRC PLA foci (Bhowmick et al., 2023). The differences to my observations could result from the use of a different cell line background and the application of cell synchronization. Despite their immortalization with p53 knockout and human telomerase reverse transcriptase (hTERT) overexpression, RPE-1 cells are considered as a non-transformed alternative to cancer cell lines (Bodnar et al., 1998; Scott et al., 2020). Given their more stable genome, Cyclin E overexpression might exert more pronounced effects on transcription-replication coordination. Furthermore, the authors applied cell synchronization via nocodazole and thymidine treatments (Bhowmick et al., 2023), which induce moderate replication stress independent of Cyclin E overexpression (Halicka et al., 2017). Taken together, my results cannot confirm Cyclin E overexpression as an inducer of TRCs in a U-2 OS cell line model but rather suggest that additional research will be needed to clarify the role of Cyclin E in TRC biology.

The dual specific-phosphatase CDC25A has been demonstrated to remove inhibitory phosphorylation on Cyclin E/CDK2 and Cyclin B/CDK1 complexes, thereby inducing S-phase and M-phase entry, respectively (Donzelli & Draetta, 2003). As CDC25A and Cyclin E functions converge on the S-phase checkpoint, I tested the effect of CDC25A overexpression on cellular TRC levels in an identical U-2 OS model cell line. Strikingly, acute overexpression for 4 h induced a strong increase of TRC PLA foci, uncovering CDC25A as potent inducer of TRCs (**Fig. 44A-C**). Coincidingly, reduced DNA replication efficiency and elevated γH2AX foci number were observed after 8 h, suggesting a delayed DNA damage response (**Fig. 44D and 45A, B**). My data is consistent with previous work suggesting that CDC25A overexpression can promote fork stalling and reversal in the context of DNA damage (Donzelli & Draetta, 2003; Neelsen et al., 2013). While my work has uncovered CDC25A overexpression as a potent approach to increase the TRC burden in cells, it remains unclear why CDC25A but not its downstream target Cyclin E could induce this phenotype. In addition to the S-phase checkpoint, CDC25A also controls M-phase entry (Mailand et al., 2000, 2002). However, no increased TRC or DNA damage burden could be detected in non-S-phase cells, suggesting that control of S-phase entry is primarily responsible for the phenotypes. Whether CDC25A overexpression also results in cryptic firing on intragenic origins, has not been addressed. To understand the molecular mechanisms driving CDC25A-

dependent increases in TRC burden, further research will be required. In summary, I identified CDC25A overexpression as a novel and potent approach to induce TRCs.

A recently established model of human bronchial epithelial cells (HBEC) overexpressing CDC6 was shown to undergo a short period of hyperreplication, which was accompanied by high levels of genomic instability (Komseli et al., 2018; Zampetidis et al., 2021). Overexpressing cells rapidly succumbed to oncogene-induced replicative senescence, from which a subpopulation escaped after 30-40 days by undergoing malignant oncogenic transformation. Interestingly, the initially observed DNA damage during the hyperreplication phase was predominantly located at TSS regions, indicating that TRCs could drive genome instability upon CDC6 overexpression (Zampetidis et al., 2021). Crucially, CDC6 overexpression induced a drastic increase in TRC PLA foci after 24 h, whereas acute overexpression of 4 h and 8 h did not alter TRC levels (**Fig. 46A-C**). This behavior suggests that TRC induction via CDC6 overexpression requires completion of at least one cell cycle and is likely fundamentally different from the processes driving TRC induction in the context of CDC25A. Subsequent extended time course analysis revealed even higher TRC PLA foci numbers following 48 h and 72 h of overexpression (**Fig. 46E**). Importantly, extended overexpression also led to strong reduction of EdU incorporation, in agreement with the onset of senescence (**Fig. 46D**). High TRC burden at 24 h and 48 h also coincided with an elevated number of γH2AX foci indicative of DNA damage (**Fig. 47A, B**). Finally, I also tested transcriptional changes after CDC6 overexpression by determining RNAPII levels and CTD phosphorylation changes and found that acute overexpression drastically increases total RNAPII and RNAPII Ser2P levels (**Fig. 47C**). This data suggests a global upregulation of RNAPII expression at early time points, a phenomenon similar to a recently described increase in RNAPII levels in aged mouse livers (Gyenis et al., 2023). Importantly, increased RNAPII levels in aged mouse livers did not increase overall transcription output but coincided with reduced RNA synthesis. Complementary work also suggested that RNAPII elongation speeds increase during aging, consistent with increased RNAPII Ser2P signal (Debès et al., 2023). Interestingly, the 8 h overexpression time point also showed the highest level of γH2AX signal in the western blot readout (**Fig. 47C**). This high level of DNA damage matches the period of highest RNAPII levels, implying transcription-dependent genomic instability. In contrast, no increase in TRC PLA foci was observed at 8 h of overexpression (**Fig. 46C**). However, recent work suggests that RNAPII stalling and enrichment on chromatin in aged mouse liver cells occurs due to preceding endogenous DNA lesions that block transcription progression (Gyenis et al., 2023). Based on current evidence it remains unclear whether this observation is identical to the early DNA damage in the HBEC CDC6 model system. Here, further research is needed. A possible way to reconcile these observations

is that initial DNA damage causes RNAPII complexes to be trapped on chromatin, increasing their residence time. Upon start of a new cell cycle, overexpressed CDC6 also leads to aberrant DNA replication that favors collisions with still trapped RNAPII, thereby giving rise to TRCs in a delayed manner. Arising TRCs could then induce a second wave of genome destabilization. Taken together, I established that CDC6 overexpression in HBEC cells is a potent model system to induce TRCs and study their contribution to genomic instability in senescence and aging.

5.6 Outlook and future directions

In my work I have established a chromosomal TRC reporter system that allows me to create conflicts directly on the genome of human cancer cells. While I have used this system to address TRC-dependent chromatin alterations with respect to nucleosome occupancy and histone modifications, it is applicable to many questions related to TRC biology. Particularly, the timing of nearly all TRC-related events in the cell remains insufficiently explored. Precise time course analysis of TRC induction with the reporter system would allow for a better understanding of when a TRC occurs, if it translates into DNA damage, when the DNA damage response occurs, and when the conflict has eventually been resolved. So far, researchers have addressed individual aspects of this cascade, yet a holistic view and understanding are urgently needed. With respect to my own research, I have only explored the chromatin dynamics of TRCs and R-loops with single time points in synchronized S-phase cells. Time-resolved analysis of nucleosome or histone PTM dynamics has the potential to uncover various unexpected findings at TRC sites, thereby providing significant insight into TRC occurrence and resolution. Especially with respect to the newly identified deposition of H3K79me2/3 as a TRC associated histone PTM, future research will be required to characterize additional upstream and downstream events, thus building a potential biological pathway. Current evidence points towards RNF20/40 and H2BK120ub as crucial mediators of H3K79me2/3 deposition (**Fig. 49A**), whereas Menin, 53BP1, and FANCD2 are likely involved in the ensuing DNA damage response activation (**Fig. 49B**). Additionally, I have demonstrated that H3K79me2/3 is required for efficient transcriptional output at the mAIRN TRC reporter. Two recent independent studies have suggested modes of RNAPII recovery or restart at genomic TRC sites (Bruno et al., 2024; Fenstermaker et al., 2023). Nevertheless, the underlying molecular mechanisms will need to be clarified. Given the direct control over TRC formation and transcription dynamics, I expect the mAIRN TRC reporter system to be perfectly suited to characterize this process. Apart from the current mAIRN reporter system in U-2 OS cells, future evolutions of the same system have the potential to aid researchers to

better understand R-loops and TRCs. Instead of a Sleeping Beauty based integration, it would be possible to perform site-specific integration by CRISPR/Cas9 genome editing. Thus, reports placed in front of replication origins in different orientations would allow one to study the effect of HO and CD TRCs specifically. Finally, it will be possible to introduce an R-loop/TRC inducing cassette into different cell lines of interest to study the contribution of conflicts to various biological processes such as development, aging, and cancer.

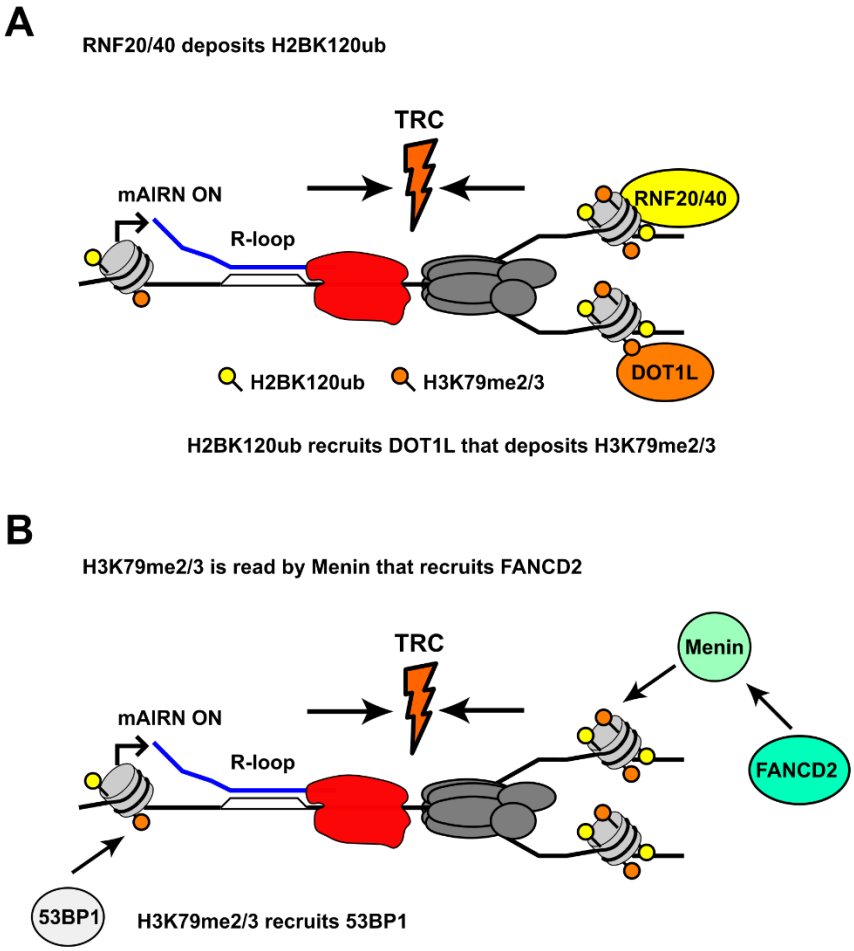

Figure 49: Putative pathways upstream and downstream of H3K79me2/3 deposition at TRCs

A) RNF20/40 might sense DNA damage at TRCs and deposit H2BK120ub which was previously shown to recruit DOT1L for deposition of H3K79me2/3.
B) H3K79me2/3 then recruits 53BP1 for damage repair. Independently, H3K79me2/3 is read by Menin that has been shown to recruit FANCD2.

Beyond the scope of my study and future applications of the reporter system, the current findings will enable researchers to more comprehensively understand TRC biology. Thereby it will be

possible to extrapolate numerous findings from model systems to endogenous TRC sites, helping to define a molecular signature of these conflicts. This signature will be particularly important to address the persistent limitations in genome-wide mapping of TRCs, allowing researchers to test and benchmark future TRC mapping technologies. Finally, it will be crucial to extrapolate these findings into disease-related contexts such as cancer. My work together with the efforts of many colleagues has started to shed light on how oncogenes can cause TRCs and TRC-driven genomic instability. Future studies could translate this knowledge on TRC biology into the clinics, enabling the stratification of patient tumors based on TRC-burden, and thereby offering more precise and effective treatment options.

6. Materials and Methods

6.1 Materials

6.1.1 Experimental model systems

REAGENT or RESOURCE	SOURCE	IDENTIFIER
U-2 OS Tet-ON	Takara Clontech	Michael Kirsch; Cat# 630919
HEK293 Tet-ON	Takara Clontech	Cat# 631182

6.1.2 Antibodies

Antibody Name	SOURCE	IDENTIFIER
DYKDDDDK (FLAG) Tag Polyclonal Antibody	Thermo Fisher Scientific	Cat#PA1-984B, RRID:AB_347227
Goat Anti-mouse IgG (H+L), Horseradish peroxidase conjugate	Invitrogen	Cat#G21040; RRID:AB_2536527
Goat Anti-Rabbit IgG (H+L), Horseradish peroxidase conjugate	Invitrogen	Cat#G21234; RRID:AB_2536530
Monoclonal mouse IgG anti DNA-RNA-hybrid (S9.6)	Merck	Cat#MABE1095; RRID:AB_2861387
Monoclonal mouse IgG anti GAPDH (6C5)	Merck	Cat#CB1001-500UG; RRID:AB_2107426
Monoclonal mouse IgG Anti-RNA polymerase II, clone CTD4H8	Merck Millipore	Cat# 05-623; RRID:AB_309852
Monoclonal mouse IgG Cyclin E1 (HE12)	Cell signaling	Cat#4129; RRID:AB_2071200
Monoclonal mouse IgG1 anti BrdU (B44)	BD bioscience	Cat#347580, RRID:AB_400326
Monoclonal mouse IgM anti RNA polymerase II RPB1 (H5)	BioLegend	Cat#920204; RRID:AB_2616695

Monoclonal rabbit anti Ubiquityl-Histone H2AK119ub (D27C4)	Cell Signaling Technology	Cat#8240S; RRID:AB_10891618
Monoclonal rabbit IgG anti DOT1L (D1W4Z)	Cell Signaling Technology	Cat#77087; RRID:AB_2799889
Monoclonal rabbit IgG anti yH2A.X (Ser139) (20E3)	Cell Signaling Technology	Cat#9718S; RRID:AB_2118009
Mouse monoclonal anti-RNA Polymerase II CTD Antibody (8WG16)	Merck	Cat#05-952; RRID: AB_11213782
Polyclonal goat anti mouse IgG Alexa Fluor Plus 488 labeled	Invitrogen	Cat#A-32723, RRID:AB_2633275
Polyclonal goat anti mouse IgG Alexa Fluor Plus 594 labeled	Invitrogen	Cat#A-11032; RRID:AB_2534091
Polyclonal goat anti mouse IgG Alexa Fluor Plus 647 labeled	Invitrogen	Cat#A-32728; RRID:AB_2866490
Polyclonal goat anti rabbit IgG Alexa Fluor Plus 488 labeled	Invitrogen	Cat#A-11008; RRID:AB_143165
Polyclonal goat anti rabbit IgG Alexa Fluor Plus 594 labeled	Invitrogen	Cat#A-11037; RRID:AB_2534095
Polyclonal goat anti rabbit IgG Alexa Fluor Plus 647 labeled	Invitrogen	Cat#A-21245; RRID:AB_2535813
Polyclonal IgG CDC6	Thermo Fisher Scientific	Cat#PA5-77901; RRID:AB_2735620
Polyclonal rabbit anti Histone H3	Abcam	Cat#ab1791; RRID:AB_302613
Polyclonal rabbit anti Histone H3K4me3	Abcam	Cat#ab8580; RRID:AB_306649
Polyclonal rabbit anti Histone H3K79me2	Active Motif	Cat#39143; RRID:AB_2561018
Polyclonal rabbit anti Histone H3K79me3	Abcam	Cat#ab195500; RRID:AB_2888917
Polyclonal rabbit anti human FANCD2	NovusBio	Cat#NB100-182; RRID:AB_10002867
Polyclonal rabbit IgG anti ORC2	Thermo Fisher Scientific	Cat#PA5-67313; RRID:AB_2663245
Polyclonal rabbit IgG anti PCNA	Abcam	Cat#ab18197; RRID:AB_444313
Polyclonal rabbit IgG CDC25A	Cell signaling	Cat#3652S; RRID:AB_2275795
Polyclonal RNA polymerase II CTD repeat YSPTSPS (phospho S2)	Abcam	Cat#ab5095; RRID:AB_304749
Rabbit polyclonal IgG anti human pRPA32 (Ser33)	Bethyl laboratories Inc.	Cat#A300-246A; RRID:AB_526488

6.1.3 Bacterial Strains

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Escherichia coli DH5alpha	Thermo Fisher Scientific	Cat#18265017

6.1.4 Chemicals and recombinant proteins

[α−32P] dATP	Hartmann Analytik	Cat#SCP-203
0.25 % Trypsin-EDTA (1x)	Gibco	Cat# 25200-072
4′,6-Diamidino-2-phenylindole dihydrochloride (DAPI)	Sigma Aldrich	Cat# 32670
5-Bromo-2'-Deoxy-Uridine (BrdU)	Sigma Aldrich	Cat#B5002-1G
5-Ethynyl-2'-deoxyuridine (EdU)	Carl Roth	Cat#7845.3
AlexaFluor 488 azide	ThermoFisher Scientific	Cat# A10266
AlexaFluor 594 azide	ThermoFisher Scientific	Cat# A10270
Ampure XP beads	Beckman Coulter	Cat#A63880
Aphidicolin (APH)	Santa Cruz	Cat#SC-201535
Ascorbic acid	Sigma-Aldrich	Cat#A7506-100G
BSA Fraction V	Sigma-Aldrich	Cat#10735078001
Deoxycholic acid	Santa Cruz	Cat#sc-214865A
DNase I	New England Biolabs	Cat# M0303S
Doxycycline hydrochloride (DOX)	Sigma-Aldrich	Cat#D3447
Dulbecco's Modified Eagle Medium (DMEM)	Gibco	Cat# 41966-029
Dulbecco's phosphate-buffered saline (PBS)	Gibco	Cat# 14190-094
E. coli RNase H	New England Biolabs	Cat#M0297S
EPZ 004777	Tocris	Cat#5567
EPZ-5676 (DOT1L inhibitor)	BPS Bioscience	Cat#27625
Fetal bovine serum for cell culture (tetracycline-free) (FBS tet-)	Takara	Cat# 631106
Glycogen	Thermo Fisher Scientific	Cat#AM9510
Halt Protease and Phosphatase Inhibitor Cocktail (100x)	Thermo Fisher Scientific	Cat# 78446
Halt Protease-Inhibitor-Cocktail (100x)	Thermo Fisher Scientific	Cat#78429
Hydroxyurea (HU)	Biomol	Cat#H9120.10
Igepal CA-630	Sigma-Aldrich	Cat#I3021-100ML
KU-60019 (ATM inhibitor)	Sigma Aldrich	Cat#SML1416-5MG
Lipofectamine 3000	Thermo Fisher Scientific	Cat#L3000001
MNase	New England Biolabs	Cat#M0247S
OmniPur Phenol:Chloroform:Isoamyl Alcohol, 25:24:1	Millipore	Cat#6805-100ML
Opti-MEM I (1X) + GlutaMAX -I	Gibco	Cat# 51985-034
Penicillin-Streptomycin-Glutamin (100x) (PSG)	Gibco	Cat# 10378-016
phenol:chlorophorm:isoamyl alcohol (25:24:1; v/v)	Invitrogen	Cat# 15593-031
Pierce Protein G Magnetic Beats	Thermo Fisher Scientific	Cat# 88847

Propidiumiodid	Invitrogen	Cat# P3566
Protein A/G agarose	Thermo Fisher Scientific	Cat#20421
Proteinase K	SERVA	Cat#33756
Puromycin	Sigma Aldrich	Cat#P9620-10ML
RNase A	Thermo Fisher Scientific	Cat#EN0531
SiR-DNA	Tebu-bio	Cat#SC007
Thymidine	Sigma-Aldrich	Cat#T9250-5G
Triton X-100	Sigma-Aldrich	Cat#X100-100ML
TRIzol	Invitrogen	Cat#15596-026
Trizol reagent	Invitrogen	Cat# 15596018
Tween-20	Kraft	Cat#21440.2000
VE-821 (ATR inhibitor)	Biomol	Cat#Cay17587-5

6.1.5 Critical commercial assays

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
Duolink detection reagents (ligation/amplification) - FarRed	Sigma Aldrich	Cat#DUO92013
Duolink detection reagents (ligation/amplification)- Green	Sigma Aldrich	Cat# DUO92014
Duolink *In Situ* PLA Probe Anti-Mouse MINUS	Sigma-Aldrich	Cat#DUO92004
Duolink *In Situ* PLA Probe Anti-Mouse PLUS	Sigma Aldrich	Cat#DUO92001
Duolink *In Situ* PLA Probe Anti-Rabbit MINUS	Sigma Aldrich	Cat#DUO92005
Duolink *In Situ* PLA Probe Anti-Rabbit PLUS	Sigma Aldrich	Cat#DUO92002
Gibson Assembly Cloning Kit	New England Biolabs	Cat# E5510S
Invitrogen™ Click-iT™ EdU Cell Proliferation Kit for Imaging, Alexa Fluor 488 dye	Invitrogen	Cat#C10337
Invitrogen™ Click-iT™ EdU Cell Proliferation Kit for Imaging, Alexa Fluor 594 dye	Invitrogen	Cat#C10339
iTaq Universal SYBR Green Supermix	Bio-Rad; REF	Cat#1725121
NEBNext Ultra II DNA Library Kit with Purification Beads	New England Biolabs	Cat#E7103S
Proteinase K	SERVA	Cat#33756
Qubit 1X dsDNA HS Kit	Thermo Fisher Scientific	Cat#Q33230
RadPrime DNA labeling Kit	Thermo Fisher Scientific	Cat#18428011
SuperScript III First-Strand Synthesis System	Thermo Fisher Scientific	Cat#18080-051

SuperSignal™ West Pico PLUS Chemiluminescent Substrate	Thermo Fisher Scientific	Cat#34580

6.1.6 Oligonucleotides

PRIMER	SOURCE	SEQUENCE 5'-3'	APPLICATION
ACTB FWD	This study	CGGGGTCTTTGTCTGAGC	ChIP-qPCR, Copy number analysis
ACTB REV	This study	CAGTTAGCGCCCAAAGGAC	ChIP-qPCR, Copy number analysis
ACTB RT-qPCR	This study	CCTGGCACCCAGCACAAT	RT-qPCR
ACTB RT-qPCR	This study	GGGCCGGACTCGTCATACT	RT-qPCR
ACTB#in3 FWD	(Skourti-Stathaki et al., 2014)	TAACACTGGCTCGTGTGACAA	ChIP-qPCR
ACTB#in3 REV	(Skourti-Stathaki et al., 2014)	AAGTGCAAAGAACACGGCTAA	ChIP-qPCR
ALAS1 FWD	(Dong et al., 2022)	GGCAGCACAGATGAATCAGA	RT-qPCR
ALAS1 REV	(Dong et al., 2022)	CCTCCATCGGTTTTCACACT	RT-qPCR
CERS6 FWD	This study	AAGCTGGGAGATCGTTGGAC	RT-qPCR
CERS6 REV	This study	CATCCTTGGACACCTTGCCT	RT-qPCR
ECFP#1 FWD	This study	ACGTAAACGGCCACAAGTTC	ChIP-qPCR, RT-qPCR, Copy number analysis
ECFP#1 REV	This study	AAGTCGTGCTGCTTCATGTG	ChIP-qPCR, RT-qPCR, Copy number analysis
ECFP#2 FWD	This study	TGGTTTGTCCAAACTCATCAA	ChIP-qPCR, RT-qPCR, Copy number analysis
FAM25A FWD	This study	ATCCTAGTTCACCACTGTCTGC	RT-qPCR
FAM25A REV	This study	CTTCCACGGCATGAATGGCTC	RT-qPCR
FHIT FWD	This study	GCACCTGATGTTAAGCCGGA	BrdU-IP
FHIT REV	This study	CTGCCCGACGAGAAACAAGA	BrdU-IP
GLUD1 FWD	This study	TGCAAGGGAGGTATCCGTTA	RT-qPCR
GLUD1 REV	This study	CAAACGGCACATCAACCACT	RT-qPCR
Int_site chr10 FWD	This study	CCACCCACATCCTGCTGATT	Sanger sequencing

Intergenic FWD	This study	CCAGGTGGGTCTCGAACTTC	ChIP-qPCR, DRIP
Intergenic REV	This study	CAGGCTGGGCAACATACTGA	ChIP-qPCR, DRIP
Intsite chr10 FWD2	This study	ACATCCTGCTGATTTGCCCA	Genotyping PCR
M13 FWD	This study	GTAAAACGACGGCCAGT	ChIP-qPCR, DRIP
M13 REV	This study	CAGGAAACAGCTATGAC	ChIP-qPCR, DRIP
mAIRN#1 FWD	This study	TAGAGGATTCCGCAAAGGAA	BrdU-IP, ChIP-qPCR, DRIP, RT-qPCR, Copy number analysis
mAIRN#1 REV	This study	TTCACCCTAGCGCTGAATCT	BrdU-IP, ChIP-qPCR, DRIP, RT-qPCR, Copy number analysis
mAIRN#2 FWD	This study	CGAGAGAGGCTAAGGGTGAA	ChIP-qPCR, DRIP, RT-qPCR, Copy number analysis
mAIRN#2/ ECFP#2 REV	This study	ACATGGTCCTGCTGGAGTTC	ChIP-qPCR, DRIP, RT-qPCR, Copy number analysis, Southern Blot, Genotyping PCR
MRSP9-AS2	This study	CGGCATCTCGTTAGCTCTGA	RT-qPCR
MRSP9-AS2 FWD	This study	GCTTGTGAGCAACCCAAGAA	RT-qPCR
MT073_gib_bb_FWD	This study	AAGTAAAACCTAACAACAACAATTGCATTCATTTTATGTTTCAGGT	Cloning
MT073_gib_bb_TEV	This study	ACTAATATGGCGTCTAGATAGCGGACCCC	Cloning
MT073_gib_ins_FWD	This study	CTATCTAGACGCCATATTAGTCATTGGTTATATAGCATAAATCAATATTGGCT	Cloning
MT073_gib_ins_REV	This study	ATTGTTGTTGTTAGGTTTTACTTGCTTTAAAAAACCTCCCAC	Cloning
NRG3 FWD	This study	GAGGCCCAGGACACATAGAA	ChIP-qPCR
NRG3 REV	This study	GCAACAGGCTAACATGCAGA	ChIP-qPCR
NRXN2 FWD	This study	CGCAAAGCCCAGTTGTTCTG	ChIP-qPCR
NRXN2 REV	This study	TTAAATTGGGGTTGCCGTGC	ChIP-qPCR
oriP5' REV	This study	TCGCTGTTCCTTAGGACCCT	Southern Blot
PUM1 FWD	(Dong et al., 2022)	CAGGCTGCCTACCAACTCAT	RT-qPCR
PUM1 REV	(Dong et al., 2022)	GTTCCCGAACCATCTCATTC	RT-qPCR

RPL13A FWD	This study	AGGTGCCTTGCTCACAGAG T	DRIP
RPL13A REV	This study	GGTTGCATTGCCCTCATTAC	DRIP
siPOOL 2 -negative control	siTOOLs Biotech	Cat# si-C002	siRNA KD
siPOOL 2 -siDOT1L	siTOOLs Biotech	Cat#si-G020-84444	siRNA KD
STK39 FWD	This study	AGGAGGTTATCGGCAGTGG A	RT-qPCR
STK39 REV	This study	CTGGTCTGGCATTTTTCCAA GT	RT-qPCR
TSS FWD	This study	ATGTCGAGGTAGGCGTGTA C	ChIP-qPCR
TSS REV	This study	TGAAGCCTCGGGTACCGAG CTCGAATTC	ChIP-qPCR
WWOX FWD	This study	CAGCCAGCACTCCTTCTCAA	ChIP-qPCR
WWOX REV	This study	CTCTGTGGAGAAGCCAAGC A	ChIP-qPCR

6.1.7 Plasmids

PLASMID	INSERT	CONSTRUCTION/SOURCE
K016_pcDNA3.1(+) ΔN1-27 RNAseH1-FLAG	-	(Hamperl et al., 2017)
K031_pSH26 1x LEXA	-	(Hamperl et al., 2017)
K069_pSH36 1xLEXA	-	(Hamperl et al., 2017)
K072_pSH37 1xLEXA	-	(Hamperl et al., 2017)
K191_pSB100	-	Tomas Zikmund
K192_pSBtet_DNMT3A_P2A_ds Red2	-	Tomas Zikmund
K206_ pcDNA3.1(+)	-	(Hamperl et al., 2017)
K275_pMT03_sb_mAIRN_puroR	Tight-TRE promoter_mAIR N-reporter_SV40-polyA	Tight-TRE promoter_mAIRN-reporter_SV40-polyA insert was amplified from K069_pSH36 1x LEXA with primers MT073_gib_ins_FWD and MT073_gib_ins_REV, Sleeping Beauty backbone from K192_pSBet_DNMT3A_P2A-dsRed2 was amplified using primers MT073_gib_bb_FWD and MT073_gib_bb_REV, insert and backbone were joined using Gibson assamebly
K276_pMT04_sb_ECFP_puroR	Tight-TRE promoter_ECFP	Tight-TRE promoter_ECFP-reporter_SV40-polyA insert was

	-reporter_SV40-polyA	amplified from K031_pSH26 1x LEXA with primers MT073_gib_ins_FWD and MT073_gib_ins_REV, Sleeping Beauty backbone from K192_pSBet_DNMT3A_P2A-dsRed2 was amplified using primers MT073_gib_bb_FWD and MT073_gib_bb_REV, insert and backbone were joined using Gibson assamebly

6.1.8 Equipment

Device	**Manufacturer**
2100 Bioanalyzer	Agilent
BD FACSCanto	BD Bioscience
Bioruptor UCD-200	Diagenode
Centrifuge 5424R	Eppendorf
Centrifuge 5424	Eppendorf
Centrifuge 5810 R	Eppendorf
ChemiDoc Touch	Bio-Rad
E220 evolution	Covaris
HX-2 Block heater	Peqlab
Kern EMB precision balance	Kern
Lab 850 pH meter	SI Analytics
LightCycler 480 II	Roche
Mastercycler nexus	Eppendorf
Mastercycler nexus gradient	Eppendorf
ThermoMixer F1.5	Thermo Fisher Scientific
NanoDrop 2000c	Thermo Fisher Scientific
Qubit 4 Fluorometer	Invitrogen
Typhoon FLA 7000	GE Healthcare
UVP Crosslinker	Analytic Jena

6.1.9 Software and algorithms

REAGENT or RESOURCE	**SOURCE**	**WEB LINK**
ImageJ	NIH	https://imagej.net/ij/index.html
FlowJo (version 10)	BD Bioscience	
bwa (version 0.7.17)	(H. Li & Durbin, 2009)	https://github.com/lh3/bwa
samtools (version 1.16.1 and 1.17)	(Danecek et al., 2021)	https://www.htslib.org/doc/samtools.html

tiddit (version 3.3.2)	(Lindstrand et al., 2017)	https://github.com/SciLifeLab/TIDDIT/releases/tag/TIDDIT-3.3.2
bedtools (version 2.31.0)	(Quinlan & Hall, 2010)	https://github.com/arq5x/bedtools2/releases/tag/v2.31.0
rtracklayer (version 1.54.0)	R Core Team	https://bioconductor.org/packages/3.18/bioc/html/rtracklayer.html
GenomicRanges (version 1.46.1)	(Lawrence et al., 2013)	https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html
R v4.1.2	R Core Team	https://www.r-project.org
trim_galore (version 0.6.10)	(Krueger et al., 2023)	https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
bowtie2 (version 2.5.1)	(Danecek et al., 2021)	https://bowtie-bio.sourceforge.net/bowtie2/index.shtml
picard MarkDuplicates (version 3.0.0)	Broad Institute	https://broadinstitute.github.io/picard/
deeptools (version 3.5.2)	(Ramírez et al., 2016)	https://deeptools.readthedocs.io/en/develop/index.html
DESeq2 package (version 1.34.0)	(Love et al., 2014)	https://github.com/thelovelab/DESeq2
sva package (version 3.42.0)	(Leek et al., 2012)	https://bioconductor.org/packages/release/bioc/html/sva.html
pheatmap (version 1.0.12)	Raivo Kolde	https://CRAN.R-project.org/package=pheatmap
macs2 (version 2.2.9.1)	(Y. Zhang et al., 2008)	https://pypi.org/project/MACS2/
Fusion (version 2.3)	Andor Oxford Instruments	https://andor.oxinst.com/downloads/view/fusion-release-2.3

6.2 Methods

6.2.1 Cell lines and cell culture

HEK293 Tet-ON and U-2 OS Tet-ON cell lines (Table S2) were cultured in Dulbeccos Modified Eagle Medium (DMEM) (GIBCO) containing 10 % Tet-approved fetal bovine serum (FBS), 2 mM L-glutamine, and penicillin/streptomycin, in a 5 % CO_2 environment at 37°C. U-2 OS Cyclin E Tet-OFF and U-2 OS CDC25A Tet-OFF cells were additionally supplemented with 4 µg/mL tetracycline to suppress oncogene expression. HBEC cells were cultured in Keratinocyte SFM (GIBCO) supplemented with 5 mg/L human recombinant Epidermal Growth Factor (rEGF) and 50 mg/L Bovine Pituitary Extract (BPE) (GIBCO).

HEK293 cells containing TRC reporter plasmids were constructed in previous work (Hamperl, 2017). U-2 OS Cyclin E Tet-OFF and U-2 OS CDC25A Tet-OFF cells were a kind gift from Prof.

Dr. Massimo Lopes (University of Zürich). HBEC cells containing a CDC6 Tet-ON overexpression cassette were kindly provided by Agyris Papantonis (University Medical Center Göttingen).

6.2.2 Antibodies and Reagents

A detailed list of all cell lines (6.1.1), antibodies (6.1.2), bacterial strains (6.1.3), chemicals and recombinant proteins (6.1.4), critical commercial assays (6.1.5), oligonucleotides (6.1.6), and plasmids (6.1.7) used is provided in the Materials section.

6.2.3 Construction of Sleeping Beauty plasmids and genomic integration of reporter sequence

In the first step of the construction of mAIRN sequence containing Sleeping Beauty vector (plasmid K275), the plasmid K192 pSBtet_DNMT3A_P2A_dsRed2 was amplified by PCR using primers MT073_gib_bb_fwd and MT073_gib_bb_rev. The resulting linearized vector of the Sleeping Beauty backbone was then used for Gibson assembly together with a fragment harboring the Tet-On promoter, mAIRN reporter sequence, and the SV40 poly-A signal, which was produced by PCR from plasmid K069_pSH36_1xLEXA using primers MT073_gib_ins_fwd and MT073_gib_ins_rev. The mAIRN fragment and the Sleeping Beauty vector backbone were combined by Gibson Assembly utilizing the Gibson Assembly Cloning Kit (NEB, REF: E5510S) according to the manufacturer's recommendations. An identical strategy was used for constructing an ECFP containing Sleeping Beauty vector (plasmid K276). The ECFP sequence originated from plasmid K031_pSH26_1xLEXA and was amplified by PCR using identical primers as for the mAIRN sequence integration.

200,000 U-2 OS cells per well were seeded in a 6-well plate. 200 ng of the mAIRN or ECFP reporter sequence-containing sleeping beauty shuttle plasmid (K275 or K276) and 1800 ng of the transposase expression plasmid (K191) (Mátés et al., 2009) were diluted in 100 µL of OptiMEM, while 3.5 µL of Lipofectamine 2000 were also diluted in 100 µL of OptiMEM. Both solutions were kept for 5 min at room temperature, then mixed and incubated for another 20 min at room temperature. The cell culture medium of each 6-well was then replaced with 1.8 mL of fresh DMEM, before gently adding the mix to the cells in dropwise fashion. Following a growth period of 24 h, the medium was changed to 2 mL of fresh DMEM containing 1 µg/mL Puromycin. After selection for 10 days, the puromycin-resistant polyclonal cell population was diluted to a concentration of 0.5 cells/100 µL per well on multiple 96-well plates. Hence, monoclonal cell lines were derived, growing from a single cell under 1 µg/mL Puromycin selection. Clones were

selected based on integration number and successful expression of the reporter gene, which were determined using qPCR and RT-qPCR (see 6.2.13 and 6.2.19)

6.2.4 Plasmid and siRNA transfections

A suitable number of cells was seeded into the respective 96-well (5,000 cells), 6-well (200,000 cells) or 15 cm plate (1.5 Mio) to reach 40-60% confluency on the day of transfection. For transfection of plasmids, the respective DNA and Lipofectamine 2000 were diluted in OptiMEM in a ratio of 1:25. Both solutions were incubated for 5 min at room temperature, then combined and incubated for an additional 20 min at room temperature. The cell culture medium was replaced with fresh DMEM including FBS and antibiotics. Finally, the mix was added to the cells in a dropwise manner. For transfections of siRNA, 5 nM siRNA and Lipofectamine RNAiMAX were diluted in OptiMEM in ratios of 1:6.25 (siRNA) and 1:62.5 (RNAiMAX). Both solutions were incubated for 5 min at room temperature, mixed by vortexing and then incubated for another 20 min at room temperature. The reaction mix was added to the cells as described above.

6.2.5 MNase Assay and DNA isolation

HEK293 cells were grown in 6-well plates to 70-90 % confluency (one well for each sample/condition). For harvesting, the cells were rinsed with 2 mL of 1x PBS, trypsinized with 200 µL of 0.25 % trypsin, and incubated at 37°C for 5 min. Following resuspension in 0.5 mL of 1x PBS, the cell suspension was transferred to microcentrifuge tubes. The respective wells were washed again with 0.5 mL of 1x PBS, which was added to the cell suspension. The combined suspension was centrifuged at 500 g for 5 min. After discarding the supernatant, the cell pellet was dissolved in 1 mL of MNase lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.5 % NP-40, 0.15 mM spermine, 0.5 mM spermidine) containing PMSF and benzamidine. After 5 min incubation on ice, nuclei were collected by centrifugation at 500 g for 5 min and rinsed with fresh MNase lysis buffer also supplemented with PMSF and benzamidine. The nuclei were again centrifuged at 500 g for 5 min and resuspended in 600 µL of MNase digestion buffer (10 mM Tris-HCl, pH 7.4, 15 mM NaCl, 60 mM KCl, 0.15 mM spermine, 0.5 mM spermidine, 1 mM $CaCl_2$). The nuclei preparation was split into aliquots of 100 µL by transferring it to separate microcentrifuge tubes containing diluted MNase (0, 2.5, 25, 100, and 250 gel units). For MNase digestion, the samples were incubated for 5 min at 30°C. Next, 100 µL of IRN (50 mM TRIS-HCl pH 8, 20 mM EDTA, 500 mM NaCl) buffer, 7.5 µL of Proteinase K (10 mg/ml), and 20 µL of 10 % sodium dodecyl sulfate (SDS) were added to each sample to stop the MNase reaction. The samples were incubated overnight at 37°C. MNase-digested DNA was purified by phenol/chloroform extraction. To this end, 200 µL phenol/chloroform/isoamylalcohol (25:24:1, v/v) were added to all samples.

The mixture was thoroughly vortexed to obtain a homogenous suspension and afterwards centrifuged for 10 min at 20,000 g. The upper (aqueous) phase was transferred to a new tube. For removal of remaining RNA, 1 µL RNase A (10mg/mL) was added to the samples, followed by an incubation for 2 h at 37°C. The DNA extraction was repeated using 200 µL of chloroform. DNA precipitation was conducted by adding 2.5 volumes of 100 % ethanol and 1 µL of glycogen (10 mg/mL) to each sample. After incubation at -20°C for at least 30 min, the DNA was precipitated at 20,000 g for at least 30 min at 4°C. The supernatant was discarded, and the pellet was rinsed with 200 µL 70 % ethanol. The samples were centrifuged again for 10 min at 20,000 g at 4°C. The DNA pellet was dried for 15-20 min at RT and resuspended in 30 µL of H_2O or TE buffer. Finally, 10-15 µL of each sample were visualized on a 1.2 % agarose gel.

6.2.6 Southern Blot

Following agarose gel electrophoresis on a 1.2 % gel TBE gel for 2 h, the DNA fragments were transferred from the gel onto a nylon membrane for subsequent hybridization and detection with a Southern blot probe. To denature the DNA the gel was submerged in denaturing solution (0.5 M NaOH, 1.5 M NaCl)for 2x 15 min, followed by 2x 15 min neutralization with transfer buffer (1 M NH_4OAc). During the gel incubations, a nylon membrane was soaked in water, and a large Whatman paper was soaked in transfer buffer (1 M NH_4OAc). For the assembly of the Southern blot transfer, a platform was placed in a tray filled with transfer buffer. A fully wet Whatman paper devoid of air bubbles was placed on the platform. This paper worked as a wick to draw the transfer buffer from the reservoir upwards through the gel. Afterwards, the gel was carefully positioned face-down on the wet Whatman paper. The pre-wet nylon membrane was placed on top of the gel, while making certain no air bubbles were trapped. Finally, two more Whatman paper sheets were placed on the membrane, followed by a stack of paper towels and a light weight (approximately 0.5 kg), thus generating a suction force to draw the transfer buffer upwards. Blotting was done overnight to allow complete transfer of the DNA fragments, as fragments up to 15 kb require approximately 18 h transfer time. Next, the membrane was crosslinked by UV utilizing a UVP Crosslinker on automatic setting to covalently bind the DNA fragments to the membrane, thereby improving hybridization signals during subsequent detection steps. The membrane was dried and stored at room temperature for subsequent analysis. DNA probes for hybridization were produced from PCR-amplified DNA fragments containing the mAIRN reporter sequence using primers oriP 5' REV and mAIRN#2 REV and body-labeled with the RadPrime DNA labeling system (Invitrogen, 18428-011) with the incorporation of [α−32P] dATP (Hartmann Analytik) according to the manufacturer’s recommendations. Membranes were prehybridized for

1 h at 65°C with 10-15 mL of hybridization buffer (2x SSC, 0.5 M sodium phosphate buffer pH 7.2, 7 % SDS). After prehybridization, the buffer was discarded and replaced with 15 mL of fresh prewarmed hybridization buffer. Before adding the DNA probe to the hybridization solution, it was mixed with salmon sperm DNA (final concentration 100 µg/ml) and denatured for 5 min. Hybridization was conducted overnight at 65 °C with gentle rotation in a hybridization oven. Blots were rinsed once with 30 mL of 3x SSC and 0.1 % SDS after hybridization. Additional washes for stringency were done at hybridization temperature with rotation, using three buffers in sequential order: 0.3x SSC with 0.1 % SDS, 0.1x SSC with 0.1 % SDS, and lastly 0.1x SSC with 1.5 % SDS. Each wash step was repeated twice for 15 min. Lastly, blots were dried, stored at room temperature, placed on phosphorimaging screens and read out using a Typhoon scanner.

6.2.7 *In vitro* reconstitution of RNA:DNA hybrids

A T7 promoter was added to the mAIRN DNA sequence by PCR, using a primer including the extended T7 promoter sequence (TAA TAC GAC TCA CTA TAG GGA GA). The corresponding PCR fragment was purified by gel extraction and used for *in vitro* transcription of the mAIRN sequence. Preparative *in vitro* transcription reactions (1 mL reaction volume, 5 h at 37 °C) were set up with 100 µg of recombinant T7 Polymerase, 2 µg of PCR-template, and 5 mM NTPs in the reaction buffer (30 mM Tris-HCl pH 8.0, 25 mM $MgCl_2$, 2 µg Spermidin, 10 mM DTT, 0.01 % Triton X-100). For RNA separation, the samples were run on denaturing polyacrylamide gels (8 %) and the RNA was cut out from the gel. Elution from the gel fragment was conducted in 400 mM NaCl solution for 3 h at 4°C. The resulting supernatant was collected and purified by isopropanol precipitation, dissolved, and quantified. RNA:DNA hybrids were reconstituted by supplementing the fluorescently labeled PCR fragment with the purified RNA at a 1:2 ratio. The mixture was denatured (5 min at 95 °C) and slowly cooled down. Successful assembly of RNA:DNA hybrids was confirmed via restriction enzyme and *E. coli* RNase H (NEB, Cat#M0297S) digestion.

6.2.8 *In vitro* nucleosome formation assay

In vitro nucleosome assembly through slat dialysis was done as described before (Maldonado et al., 2019). In brief, recombinant human histone octamers were mixed with the respective DNA and 100 ng/µL BSA in high salt buffer (10 mM Tris-HCl pH 7.6, 2 M NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.05 % Igepal CA-630) and transferred into a small dialysis chamber. The dialysis chambers were then placed into a beaker containing 300 mL high salt buffer. To slowly reduce the NaCl concentration, 3 L of low salt buffer (10 mM Tris-HCl pH 7.6, 50 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.05 % Igepal CA-630) were pumped into the beaker overnight.

The resulting reconstituted nucleosomes were characterized by gel electrophoresis using 6 % PAA 0.4x TBE gels and visualized by ethidium bromide staining or fluorescence imaging.

6.2.9 Cell proliferation assay and Crystal Violet Staining

U-2 OS cells were seeded at 3,000 cells per well in a 6-well plate. Following 24 h of unperturbed growth, cells were treated with the respective drugs and transferred to an Incucyte S3 Live-Cell Analysis System. Cell proliferation was tracked for seven days. The obtained confluency measurements for each timepoint were normalized to the confluency at the 0 h measurement timepoint.

Crystal violet staining was conducted with cells at the endpoint of the proliferation assay. Cells were rinsed with PBS and fixed with 4 % paraformaldehyde for 15 min in the dark. Next, the cells were stained using crystal violet solution for 20 min with gentle agitation. To remove excess staining, cells were washed three times with PBS for 5 min with gentle agitation. Finally, the plates were dried at room temperature and imaged using a commercial camera.

6.2.10 Cell cycle synchronization by double thymidine block

U-2 OS cells were plated at 1 million cells per 10 cm dish and grown under standard culture conditions (see 6.2.1). The following day, the culture medium was changed to fresh DMEM medium supplemented with 2 mM thymidine. Following 18 h of thymidine block, the cells were washed with PBS and cultured in fresh medium without thymidine. Nine hours later, the medium was again exchanged to fresh medium containing 2 mM thymidine. After 17 h of thymidine block duration, the cell cycle arrest could either be stopped by adding fresh medium, thereby releasing cells into S-phase, or continued by switching to fresh medium with 2 mM of thymidine. At the time point and cell cycle stage of interest, cell pellets were collected by trypsinization and centrifugation (see 6.2.1).

6.2.11 Western blot

SDS-polyacrylamide gel electrophoresis (PAGE) was conducted using a 1.5mm NuPAGE 4 to 12 % Bis-Tris Mini Protein Gel (Thermo Fisher Scientific) with 15 wells run with 1X MES at 120 V for 1.5 h. The size-separated proteins were transferred to a methanol-activated PVDF membrane (Millipore) using a tank blot setup. Blotting was performed with transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) at 0.3 A for 3 h at 4°C. Next, the membranes were blocked with 3 % BSA in PBS for at least 45 min. The primary antibody was added to the membrane and incubated in 3 % BSA at 4°C overnight. Antibodies against GAPDH (1:10000), ORC2 (1:1000), RPA32 pS33

(1:1000), yH2A.X (1:200), H3K79me2 (1:2500), H3K79me3 (1:2000), H3 (1:4000), DOT1L (1:1000), Cyclin E (1:2000), CDC25A (1:1000), CDC6 (1:1000), RNAPII (1:1000), and RNAPII Ser2P (1:1000) were used. The membrane was washed twice with PBS-T (PBS with 0.05 % Tween-20) for 5 min each with slight agitation. The appropriate HRP-conjugated secondary antibody (1:10000) was then applied in 3 % BSA in PBS for 1 h. Following two washes with PBS-T for 5 min, the membrane was developed on a ChemiDoc Touch Imaging System (BioRad) using the SuperSignal Kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

6.2.12 Flow Cytometry

Approximately 2 million cells were treated with a 30 min pulse of 25 µM 5-Bromo-2′-deoxyuridine (BrdU). Next, cells were harvested by washing with PBS, trypsinization, and centrifugation. The samples were fixed with ice-cold 70 % ethanol and permeabilized with 0.25 % Triton X-100/PBS for 15 min on ice. To generate ssDNA accessible to the anti-BrdU antibody, the cells were denatured by incubation in 2 M HCl for 15 min at 25°C and washed with 100 mM sodium borate (pH 8.5). After a blocking step in 1 % BSA/PBS containing 0.1 % Tween-20 for 15 min, the cells were incubated with a primary BrdU antibody (1:100; BD bioscience) for 2 h at room temperature. After three washes with PBS, the cells were incubated with AlexaFluoro-488 secondary antibody (1:500; Invitrogen) for 1 h at room temperature and then washed three times with PBS. Propidium iodide (0.01 mg/mL) and RNase A (0.02 mg/mL) were added to determine DNA content and remove residual RNA, respectively. Finally, flow cytometry analysis was performed on a FACSMelody device. The cell cycle profiles were analyzed with FlowJo software.

6.2.13 DNA copy number analysis by qPCR

About 2-4 x 10^5 cells per sample were harvested by trypsinization and pelleted by centrifugation. The cell pellet was dissolved in 100 µL of TE buffer, before adding the following reagents for DNA extraction: 100 µL IRN buffer, 10 µL 10 % SDS (final concentration: 0.50 %), and 10 µL Proteinase K (10 mg/mL, final concentration: 45 µg/mL). The mix was incubated at 37°C for at least 2 h or overnight. Afterwards, 200 µL phenol:chloroform:isoamyl alcohol (25:24:1, v/v) were added, the sample was vortexed, and centrifuged at 20,000 g for 5 min. The upper aqueous phase of ca. 180 µL containing the DNA was transferred into a new 1.5 mL tube. Next, 5 µL RNase A was added for RNA removal, and the sample was incubated at 37 °C for 1-2 h. Finally, 200 µL chloroform was added, and the sample was again vortexed and centrifuged again at 20,000 g for 5 min to separate the phases. The upper aqueous phase was moved to a fresh 1.5 mL tube, DNA was precipitated with 2.5 volumes of 100 % ethanol and 1 µL glycogen (10 mg/mL), and the DNA pellet was resuspended in DNase-free water. The DNA was then mixed with iTaq SYBR Green

Supermix according to manufacturer's instructions and analyzed on a Roche Light Cycler 480 Instrument II. The copy number of target regions of interest were determined by comparison to a genomic control locus with known copy number like *ACTB*. All primers used for copy number analysis are listed in 6.1.6 Oligonucleotides.

6.2.14 Chromatin Immunoprecipitation (ChIP)

Approximately 5 million cells per sample were washed with 1x PBS and harvested by trypsinization. The cells were fixed in 9.5 mL of 1 % formaldehyde solution in PBS and agitated for 10 min. The formaldehyde was neutralized by addition of 0.5 mL of 2.5 mM glycine and incubation for an additional 5 min with agitation. The cells were pelleted at 450 g at 4 °C for 5 min and resuspended in 900 µL of PBS containing a protease and phosphatase inhibitor cocktail (Thermo Fisher Scientific, Cat#1861282). The cell pellets were quickly frozen in liquid nitrogen and stored at -80 °C until further use. Pierce Protein G Magnetic Beads (Thermo Fisher Scientific, Cat#88847) were prepared by adding 50 µL of bead slurry to a new 1.5 mL tube containing ice-cold 0.02% PBS-T. Next, the beads were washed twice with 0.02% PBS-T, twice with IP dilution buffer (20 mM TRIS-HCl pH 8, 2 mM EDTA, 150 mM NaCl, 1 % Triton-X 100, 0.01 % SDS) and stored at 4 °C until further use. For cell lysis, cell pellets were resuspended in 0.5 mL of cell lysis buffer (10 mM TRIS-HCl pH 8, 10 mM NaCl, 0.2 % Igepal CA-630) supplemented with protease and phosphatase inhibitor cocktail, and incubated at 4°C for 10 min. The nuclei were separated by centrifugation at 1700 g for 5 min at 4°C and then lysed in 550 µL of nuclei lysis buffer (50 mM TRIS-HCl pH 8, 10 mM EDTA, 1 % SDS) with phosphatase and protease inhibitors (Thermo Fisher Scientific, CAT#1861282), before adding 550 µL of IP dilution buffer. Lysed nuclei were sonicated to a target size of 200-300 bp using the Covaris E220 sonicator (10 min sonication with 140 Peak Incident Power, 5 % Duty Factor, 200 Bursts per Cycle). Unfragmented chromatin was removed by centrifugation at 15,000 g for 5 min at 4 °C. 10 % of the total IP volume (100 µL) was saved as input sample, and the remaining 900 µL were transferred to a new tube containing the previously prepared magnetic beads. Chromatin fragments were incubated overnight with beads and the antibody of interest at 4°C using gentle agitation. Antibodies (amount/sample) against RNAPII (8WG16) (5 µg), Histone H3 (4 µg), FANCD2 (5 µg), RNAPII Ser2P (10 µg), H3K4me3 (2 µg), H2AK119ub (2.5 µg), H3K79me2 (7.5 µg), and H3K79me3 (2 µg) were used in ChIP and ChIP-seq experiments. On the following day, multiple washes of 5 min each were performed at 4°C with agitation: one wash with IP wash 1 buffer (20 mM Tris-HCl pH 8, 2 mM EDTA, 50 mM NaCl, 1 % Triton-X100, 0.1 % SDS), two washes with high-salt buffer (20 mM Tris-HCl pH 8, 2 mM EDTA, 500 mM NaCl, 1 % Triton-X 100, 0.1 % SDS), one wash with IP wash 2 buffer

(10 mM Tris-HCl pH 8, 1 mM EDTA, 250 mM LiCl, 1 % NP-40, 1 % deoxycholic acid), and two washes with TE (10 mM TRIS-HCl pH 8, 2 mM EDTA). For elution of antibody bound chromatin fragments, the beads were resuspended in 47 µL of EB (50 mM TRIS-HCl pH 8, 10 mM EDTA, 1 % SDS) and 3 µL of Protease K (10 mg/mL) and incubated at 56 °C for 20 min. The supernatant was collected, and the elution process was repeated, combining the respective supernatants. Additionally, 6 µL of Proteinase K (10 mg/mL) were added to the IN samples for protein digestion. For decrosslinking, 100 µL of IRN buffer was added to all samples and then incubated at 65 °C overnight. DNA fragments were purified by phenol/chloroform extraction followed by ethanol precipitation (see 6.2.5) and qPCR was performed as described previously (see 6.1.13). A detailed list of all primers used in ChIP-qPCR is provided in 6.1.6 Oligonucleotides. The samples of selected experiments were also subjected to library preparation for ChIP-seq-

6.2.15 DNA-RNA immunoprecipitation (DRIP)

Approximately 2 million cells were lysed in 1.6 mL TE buffer supplemented with 82 µL of 10% SDS and 10 µL of 10 mg/mL Proteinase K and incubated at 37°C overnight. Isolation of DNA and associated RNAs was performed by gentle phenol:chloroform:isoamyl alcohol (25:24:1, v/v) extraction and isopropanol precipitation. Following precipitation, the aggregated DNA was gently transferred to a new 1.5 mL tube using a cut 1 mL pipette tip, thereby avoiding disruptive centrifugation. Next, the DNA was washed three times with 70 % ethanol. Isolated DNA was resuspended in 130 µL TE buffer and transferred to an AFA microTube with Snap Cap. The samples were sonicated for 4 min using a Covaris E220 sonicator (140 Peak Incident Power, 10 % Duty Factor, 200 Bursts per Cycle) resulting in DNA fragments of around 300 bp. Sonicated DNA was quantified on a NanoDrop 2000c spectrophotometer. In the case of RNase H-treatment, which was conducted as a specificity control for RNA:DNA hybrid detection, 4 µg of DNA were treated with 3 µL of 5000 U/µL *E. coli* RNase H (NEB, Cat#M0297S) overnight at 37°C. For each sample, 10 % of the total DNA was set aside as an input control. For immunoprecipitation, 4 µg of DNA were incubated with 6 µg of S9.6 antibody in 1X binding buffer (10 mM Na_3PO_4 pH 7, 140 mM NaCl, 0.05% Triton X-100) overnight at 4°C with rotation. Next, Protein A/G agarose beads (Thermo Fisher Scientific, Cat#20421) were added for 2 h to precipitate the RNA:DNA hybrid-antibody complexes. The beads were rinsed three times in 1X binding buffer for 10 min at 4°C. Elution was conducted in elution buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 0.5 % SDS, Proteinase K) for 45 min at 55°C with agitation. The eluates, including the beads, were subjected to a single round of phenol-chloroform extraction for DNA purification (see 6.2.5 MNase and DNA

isolation). RNA:DNA hybrid levels were quantified from input and eluate samples using qPCR analysis (see 6.2.13). Primers for DRIP-qPCR are listed in 6.1.6 Oligonucleotides.

6.2.16 Whole genome sequencing and identification of genomic integration sites

Genomic DNA of U-2 OS Clone#12 cells was isolated by phenol:chloroform:isoamyl alcohol (25:24:1, v/v) extraction and ethanol precipitation (see 6.2.5 MNase assay and DNA isolation). The extracted genomic DNA was fragmented by sonication on a Covaris E220 sonicator (140 Peak Incident Power, 10% Duty Factor, 200 Bursts per Cycle) resulting in DNA fragments of 200-300 bp size. The DNA fragments were used to prepare libraries for whole genome sequencing (see 6.2.18 library preparation and sequencing). For integration site identification, libraries were sequenced to a depth of 300 million reads (paired-end) per sample.

6.2.17 BrdU-seq

BrdU-seq was adapted from a previously published protocol (J. Wang et al., 2021). To label nascent DNA, cells were treated with a 50 µM 5-bromo-2'-deoxyuridine (BrdU) pulse for 30 min before harvesting. Cells were lysed in a mix of 100 µL TE buffer, 100 µL IRN buffer, 0.5 % SDS, and 10 µL Proteinase K (10 mg/mL) and incubated at 37 °C with 500 rpm agitation for at least 90 min. Genomic DNA was isolated by phenol-chloroform extraction and ethanol precipitation (see 6.2.5). Reconstituted DNA was quantified by Qubit dsDNA BR kit (Thermo Fisher Scientific) on a Qubit 4 Fluorometer (Thermo Fisher Scientific) according to the manufacturer's instructions. 20 µg of genomic DNA in 130 µL water were transferred to AFA microtubes with Snap Cap and sonicated for 10 min using a Covaris E220 sonicator (140 Peak Incident Power, 10 % Duty Factor, 200 Bursts per Cycle) to create DNA fragments of 200-300 bp size. For immunoprecipitation of BrdU-labeled DNA, 10 µL Dynabeads Protein G were rinsed three times with PBS-T 0.02 % and afterwards incubated with 2 µg anti-BrdU antibody at 4°C overnight with agitation. The fragmented DNA was diluted to reach a total volume of 200 µL, then denatured at 100°C for 10 min and chased on ice to keep the DNA as ssDNA accessible to BrdU antibody binding. 15 µL of denatured DNA were kept as an input control and combined with 35 µL elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1 % SDS). Next, 170 µL of denatured DNA were mixed with 180 µL of 2X blocking solution (2 % BSA, 2X PBS, 0.2 % Tween 20) and added to the previously prepared beads for immunoprecipitation at 4°C overnight with agitation. On the following day, BrdU-labelled DNA bound to beads was washed twice with 1 mL of lysis buffer 1 (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 1 % Triton X-100, 0.1 % Na-Deoxycholate). Subsequently, the samples were rinsed twice with 1 mL of lysis buffer 2 (50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 1 % Triton X-100, 0.1 % Na-Deoxycholate), followed by two washes with 1 mL of

wash buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl,1 mM EDTA, 0.5 % Na-Deoxycholate, 0.5 % Igepal CA-630). All washes were conducted for a duration of 5 min at 4 °C with agitation. To remove residual wash buffer components, beads were resuspended in 1 mL of TE and centrifuged for 3 min, 400 g at 4 °C. For elution of BrdU labeled DNA, the beads were mixed with 100 µL elution buffer containing 0.5 mg/µL Proteinase K and incubated at 65°C for 10 min shaking at 400 rpm. Following a repeat of the previous elution step, eluates of the respective samples were combined and incubated at 37 °C for 1 h for further protein digestion. Input samples were supplemented with 150 µL TE buffer containing 0.5 mg/ml Proteinase K and also incubated at 37 °C for 1 h. Both BrdU-labelled DNA and input DNA were purified by a single round of phenol-chloroform extraction and ethanol precipitation and reconstituted in 30 µL of water (see 6.2.5). Successful BrdU enrichment was verified by qPCR analysis (see 6.2.13) using primers mAIRN#1 FWD, mAIRN#1 REV, FHIT FWD, and FHIT REV (see 6.1.6). To reconstitute the dsDNA required for library preparation, second strand synthesis was performed with the RadPrime labelling kit according to the manufacturer's instruction (Thermo Fisher Scientific). The obtained dsDNA was purified using the GeneJET PCR Purification Kit (Thermo Fisher Scientific). Purified DNA was subjected to library preparation (see 6.2.18).

6.2.18 Library Preparation and sequencing

First, DNA samples were quantified with Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). Depending on the available material, 1 ng to 1 µg of DNA was used as starting material. The library preparation for Whole Genome Sequencing, ChIP-seq, and BrdU-seq samples was conducted according to the manufacturer's instructions utilizing the NEBNext Ultra II DNA Library Kit with Purification Beads (E7103S, E7103L) and NEBNext Multiplex Oligos for Illumina (E6440S). Libraries were again quantified and checked for integrity using a Bioanalyzer High Sensitivity DNA Analysis kit (Agilent 5067-4627). In case of remaining adapter or primer contamination, a purification step using 1X AMPure XP Beads (Beckmann Coulter A63880) was performed according to the manufacturer's instructions. Libraries were pooled and sequenced on an Illumina NovaSeq 6000 machine with 150 bp paired-end reads at the Helmholtz Munich Genomics Core Facility.

6.2.19 RNA isolation and RT-qPCR

About 1 million cells per sample were harvested by trypsinization and pelleted by centrifugation. The pellet was resuspended in 1 mL TRIzol reagent (Invitrogen) and incubated at RT for 5 min. Next, 200 µL chloroform was added, following vortexing for 10 s and incubating again at RT for 2 min. Phase separation was achieved by centrifugation at 15,000 g at 4°C for 10 min. The upper

aqueous phase (ca. 550 µL, containing nucleic acids) was moved to a new tube. Afterwards, 500 µL chloroform was added and the sample was vortexed for 10 seconds and centrifuged at 15,000 g at 4°C for 10 min. The upper aqueous phase was transferred to a new tube. For RNA precipitation, 500 µL isopropanol was added, followed by vortexing for 10 s, and incubation at RT for 10 min. After centrifugation at 15,000 g at 4 °C for 10 min, the supernatant was discarded. The RNA pellet was washed with 1 mL 75 % ethanol, centrifuged at 15,000 g at 4°C for 5 min, and air dried for 10-15 min. The RNA was reconstituted in 30 µL RNase-free water. RNA amounts were quantified on a Nanodrop spectral photometer (Thermo Fisher Scientific). Equal amounts of RNA ranging from 0.5-2 µg were used as the starting material for DNase digestion and cDNA synthesis. DNA was removed by the addition of 1 µL DNase (2,000 units/mL, NEB M0303S) and incubation at 37°C for 30 min. Subsequently, cDNA synthesis was performed using the RT-SuperScriptIII kit, following the manufacturer's instructions (Thermo Fisher Scientific; Cat#18080051). qPCR analysis was performed as described above (see 6.2.13). Gene expression levels were normalized to housekeeping gene *ACTB*. Primers used in RT-qPCR analysis are listed in 6.1.6 Oligonucleotides.

6.2.20 Immunofluorescence (IF) and EdU-Click chemistry staining

For IF, approximately 5,000-20,000 cells were seeded in a 96-well plate (Ibidi, Cat#89626) and grown for a minimum of 24 h. In case labeling of S-phase cells with EdU-Click chemistry was desired, cells were treated with 10 µM EdU for 30 min. For specific analysis of chromatin-bound proteins, pre-extraction to remove other cytoplasmic and nuclear components was applied. Pre-extraction was conducted by incubating the cells in CSK buffer (100 mM NaCl, 300 mM Sucrose, 3 mM $MgCl_2$, 10 mM 3-(N-morpholino) propanesulfonic acid (MOPS), 0.5 % Triton X-100) for 5 min. Then cells were rinsed with PBS and fixed with 4 % PFA in PBS for 15 min in the dark. After fixation, cells were washed with 150 µl PBS, permeabilized with 0.2 % Triton-X in PBS for 4 min, and washed with 150 µL PBS again. An EdU click-it reaction cocktail (100 mM Tris-HCl pH 8.5, 1 mM $CuSO_4$, 100 mM Ascorbic acid, 0.9 µg Alexa Fluor 488/595 azide (Thermo Fisher Scientific, A10266 and A10270)) was freshly prepared and added to the cells for a 30 min incubation period at room temperature in the dark. Cells were then washed with 150 µL PBS and blocked in 100 µL 3 % BSA in PBS for at least 30 min. Primary antibodies were applied in 100 µL 3 % BSA in PBS and the cells were incubated at 4 °C overnight. For immunofluorescence staining, antibodies against FANCD2 (1:1000), RPA32 pS33 (1:500), γH2AX (1:200), and RNAPII RPB1 (H5) (1:2000) were used. On the following day, cells were washed twice with PBS for 5 min and the appropriate fluorophore-conjugated secondary antibodies and DAPI (1:1000) or SiR-DNA

(1:5000) were applied in 3 % BSA in PBS. The plate was wrapped in tin foil to protect it from light and incubated for 1 h. Cells were then washed twice with PBS for 5 min and stored at 4 °C until imaging.

6.2.21 Proximity Ligation Assay (PLA)

For PLA approximately 5,000-20,000 cells were seeded in a 96-well plate (Ibidi) and grown for a minimum of 24 h. Cells were EdU treated, pre-extracted, fixed, and subjected to EdU-click labeling as previously described (see 6.2.20 Immunofluorescence and EdU-Click chemistry staining). Next, the cells were blocked in 5 % BSA in PBS for at least 45 min, followed by application of primary antibodies in 5 % BSA at 4°C overnight. Antibodies against RNAPII RPB1 (H5) (1:2000), PCNA (PC10) (1:200), PCNA (1:2000), and DOT1L (1:200 and 1:1000) were used in PLA. On the following day, cells were washed twice with PBS for 5 min. Afterwards, Duolink PLUS (Sigma Aldrich, Cat#DUO92001 and Cat#DUO92002) and MINUS (Sigma Aldrich Cat#DUO92004 and Cat#DUO92005) probes were 1:10 diluted in Duolink Antibody Diluent and added for 1 h at 37°C. For the removal of residual probes, the cells were washed twice with PLA Wash buffer A (Sigma Aldrich). The ligation solution (1x Duolink ligation buffer diluted in water, Ligase at 1:70 dilution) (Sigma Aldrich) was applied for 30 min at 37°C. Following two additional washes with PLA Wash buffer A, the amplification solution (1x Amplification buffer diluted in water, Polymerase at 1:140 dilution) (Sigma Aldrich) was applied and cells were incubated at 37°C for 100 min in the dark. The cells were then rinsed twice with PLA Wash buffer B (Sigma Aldrich) for 10 min and stained with DAPI (1:1000) or SiR-DNA (1:5000) in PBS for 1 h. Finally, cells were again washed twice with PBS for 5 min, protected from light, and stored at 4°C until microscopy analysis.

6.2.22 Microscopy setup and image acquisition

Images were acquired on a Nikon T2 inverted microscope system equipped with an Andor Dragonfly confocal spinning disk, a 40X air objective, and an iXon Life 888 EMCCD camera while running the Fusion 2.2 software suite. Per condition 67-81 positions were imaged at which a Z-stack of 7 images across 10 µm was acquired.

6.2.23 Image Analysis

A custom ImageJ (Lalonde et al., 2023; Schindelin et al., 2012) macro was used to identify DAPI or SiR-DNA stained nuclei and perform intensity measurements as well as blob detection. In summary, the z-stack in focus was identified based on maximum normalized variance (Y. Sun et al., 2004) in the DAPI/SiR-DNA channel. Afterwards, nuclei were detected using a custom trained

neural network model, StarDist (Schmidt et al., 2018; Weigert et al., 2020). Depending on the research question, different measurements of signal intensity and foci number were obtained. Identification of cells in S-phase was performed based on their nuclear mean EdU intensity. A fixed threshold for a given experiment was established by measurement of mean EdU intensity levels in cell nuclei and comparison to the background signal. Cells with nuclear mean intensity above this empirical threshold were considered as EdU-positive cells. Foci numbers from PLA and IF stainings were quantified using the ImageJ Find Maxima function with prominence values (commonly 700-900) adjusted to the respective staining and foci type.

6.2.24 Identification of reporter insertion sites

The human genome (version GRCh38 primary assembly) was obtained from the Ensembl database, and the sequence of the reporter construct (plasmid K275) was concatenated with the genome fasta file. Genome alignment of the paired-end sequencing reads to the combined genome was conducted using bwa (version 0.7.17) with the parameter -B 3 (H. Li & Durbin, 2009). Mapped reads were filtered via samtools (version 1.16.1) with the parameter -q 12 (Danecek et al., 2021). Structural variants were identified on the filtered reads and the combined genome file utilizing tiddit (version 3.3.2) with the parameters --sv -p 1 -r 1 --min_contig 4500 (Lindstrand et al., 2017). The resulting genomic positions found by reads split between the reporter and another chromosome were extracted from the variant call file (vcf). Positions that were mapped internally to the reporter sequence were excluded as they might represent incomplete integration events. Only those positions that are within 300 bp of the reporter ends were considered. Identified adjacent genomic positions with a maximum distance of 200 bp were merged using bedtools merge (version 2.31.0) with the parameter -d 200 (Quinlan & Hall, 2010). Consequently, the average of the start and end position of the merged region was used as the consensus insertion site. These processing steps were carried out for the whole genome sequencing, the BrdU-seq, and the RNAPII and H3 ChIP-seq datasets. The final insertion sites were defined as those genomic positions that were present in at least two of the four datasets within a 200 bp window (**Appendix, Table S1**). The latter step was conducted using rtracklayer (version 1.54.0) and GenomicRanges (version 1.46.1) (Lawrence et al., 2013) R packages (version 4.1.2).

6.2.25 BrdU-seq analysis

Paired-end reads of BrdU-seq samples were trimmed using trim_galore (version 0.6.10) (Krueger et al., 2023) with the parameter --quality 28 and mapped to the previously constructed genome containing the reporter sites (see 6.2.24) using bowtie2 (version 2.5.1) with the parameters --end-to-end --very-sensitive --no-unal --no-mixed --no-discordant --dovetail -I 10 -X 700 (Langmead &

Salzberg, 2012). Following the alignment sequencing reads were filtered by samtools (version 1.17) with the parameter -q 12 (Danecek et al., 2021). For removal of duplicate reads, picard MarkDuplicates (version 3.0.0) with the parameter -REMOVE_DUPLICATES TRUE (Broad Institute, 2019) was used. Genome coverages were generated using deeptools bamCoverage (version 3.5.2) (Ramírez et al., 2016) with parameters --blackListFileName hg38-blacklist.v2.bed --ignoreForNormalization MT --binSize 20 --smoothLength 60 –extendReads --maxFragmentLength 700 --normalizeUsing "CPM". A curated list of blacklisted regions was obtained from the Boyle Lab github repository (Amemiya et al., 2019). For analysis and quantification, reads were counted in 5 kb consecutive genomic bins using bedtools makewindows (version 2.31.0) with the parameter -w 5000 and deeptools multiBamSummary with parameters BED-file --BED bins.bed --blackListFileName hg38-blacklist.v2.bed --smartLabels --extendReads --centerReads --samFlagInclude 64 --outRawCounts. Only the main chromosomes as well as the reporter sequence were considered in the counting step. Differential analysis was conducted by DESeq2 package (version 1.34.0) (Love et al., 2014) in R (version 4.1.2). For calculating normalization (size) factors all genomic bins were used by the default DESeq method. Differential testing was carried out by two different approaches, either globally for all genomic bins or in a focused way in the proximity of the insertion sites i.e. only for those bins that are located within 100 kb of the insertion sites and in the reporter sequence itself. For the global analysis, genomic bins with more than 5 reads counted in at least two out of all BrdU-seq samples were included, whereas for the insertion proximal analysis, more than one read count was required. Biological replicates were considered as batch variable in the DESeq model. For subsequent analyses, normalized counts were log2 transformed following the addition of a pseudo-count of one and the application of batch correction using the ComBat function from the sva package (version 3.42.0) (Leek et al., 2012). Global Spearman´s correlation coefficients were calculated on log2 normalized counts of all genomic bins. For visualization of the BrdU-seq signal as heatmaps, log2 normalized counts were centered by the mean across samples for each bin and were visualized by pheatmap (version 1.0.12).

6.2.26 ChIP-seq analysis

Genome and ChIP-seq (RNAPII Ser2P and H3) read processing steps were carried out as described for BrdU-seq (see 6.2.25). Additionally, coverages were averaged by deeptools bigwigAverage --binSize 20 and the log2 ratio between treatment and control was calculated by bigwigCompare --binSize 100 --fixedStep. Insertion site-centered matrices were created by deeptools computeMatrix using parameters --referencePoint 'center' --binSize 100 --downstream

5000 --upstream 5000 and were visualized with deeptools plotHeatmap function. Differential analysis was done as for BrdU-seq, except for the bin size being set to 1 kb and more than 4 read counts per bin being required in at least half of the samples for the respective ChIP antibody.

6.2.27 Public sequencing data analysis

Public datasets were downloaded from GEO with the accession numbers GSE134084 (DRIP-seq) (Crossley et al., 2020), GSE29611 (H3K79me2 ChIP-seq) (Dunham et al., 2012) and GSE110354 (H3K79me3 ChIP-seq) (Reverón-Gómez et al., 2018). Genome and reads were processed as described for BrdU-seq (see 6.2.25). Additionally, peak calling on DRIP-seq replicates was performed using macs2 (version 2.2.9.1) (Y. Zhang et al., 2008) with parameters --gsize 3e9 --format BAMPE --nomodel --nolambda --broad --broad-cutoff 1e-7 --pvalue 1e-7. Peaks showing overlap with blacklist regions were filtered out using bedtools intersect with parameter -v. The final peak set was defined as the intersect of the replicates. ChIP-seq reads were counted analogously to BrdU-seq in 5 kb consecutive genomic bins. Importantly, only bins with more than one read in at least half the samples were used. No batch correction was done on the log2 normalized counts. Bins were grouped based on the overlap with the DRIP-seq peaks and log2 normalized counts were visualized as box plots.

For the analysis of histone modification patterns in putative TRC regions, bigwig files from ENCODE ChIP-seq data were obtained from GenCode (https://ftp.ebi.ac.uk/pub/databases/ensembl/encode/integration_data_jan2011/byDataType/signal/jan2011/bigwig/). Additional bigwig files for ChIP-seq data on H3K79me2, RNAPII, and H3K79me2 in MOLM13 cells (GSE185094 – Samples GSM5606327, GSM5606329, and GSM5606331), H3K79me3 in HeLa cells (GSE116310 – Samples GSM3227896 and GSM3227897), and H2AK119ub in MCF7 cells (GSE201262 – Samples GSM6056895 and GSM6056896) were downloaded from NCBI GEO. Histone modification patterns around HO and CD collision regions identified by (Hamperl et al., 2017) were analyzed. The DeepTools computeMatrix tool was utilized to determine enrichment within 12 kb upstream and downstream of replication origins, using a 100 bp bin size. To evaluate enrichment significance specific to HO vs. CD regions, 2,000 bootstrap replicates were performed, and enrichment signals were visualized with correspondent 95 % confidence intervals.

6.2.28 Quantification and statistical analyses

Error bars on Figures 10A-B, 12D, 13A-C, 14A-B, 15B-C,16, 17, 18B, 19A, 20, 21A-B, 26A-B, 27B-C, 28A-B, 30A, 32C-D, 33A-D, 37A-B, 38, 39B, 39D, 41A-C, 42, 43C-D, 44C-D, 45B, 46C-E

and 47B indicate the standard deviation of the biological replicates. All statistical details of the experiments can be found in the figure legends. All statistical comparisons were done using GraphPad Prism 10.0.2.

Statistical analysis of the BrdU-Seq and ChIP-Seq sequencing results in Figures 24, 25, and 31 was performed as described in the Methods section (see BrdU-seq analysis and ChIP-seq analysis). Detailed information about software and algorithms used for sequencing data including their sources and weblinks can be found in 6.1.9 Software and algorithms.

6.2.29 Data and code availability

BrdU-Seq, ChIP-Seq, and Whole Genome Sequencing data are available on GEO with the accession number GSE267494.

This study does not report any original code.

7. References

Abril-Garrido, J., Dienemann, C., Grabbe, F., Velychko, T., Lidschreiber, M., Wang, H., & Cramer, P. (2023). Structural basis of transcription reduction by a promoter-proximal +1 nucleosome. *Molecular Cell*, *83*(11), 1798-1809.e7. https://doi.org/10.1016/J.MOLCEL.2023.04.011

Aiello, U., Challal, D., Wentzinger, G., Lengronne, A., Appanah, R., Pasero, P., Palancade, B., & Libri, D. (2022). Sen1 is a key regulator of transcription-driven conflicts. *Molecular Cell*, *82*(16), 2952-2966.e6. https://doi.org/10.1016/J.MOLCEL.2022.06.021

Alabert, C., & Groth, A. (2012). Chromatin replication and epigenome maintenance. *Nature Reviews Molecular Cell Biology 2012 13:3*, *13*(3), 153–167. https://doi.org/10.1038/nrm3288

Alabert, C., Jasencakova, Z., & Groth, A. (2017). Chromatin Replication and Histone Dynamics. *Advances in Experimental Medicine and Biology*, *1042*, 311–333. https://doi.org/10.1007/978-981-10-6955-0_15

Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., & Liberi, G. (2012). Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell*, *151*(4), 835–846. https://doi.org/10.1016/J.CELL.2012.09.041/ATTACHMENT/DDC8805A-89F4-4633-AA61-2D366B939FE4/MMC2.PDF

Amemiya, H. M., Kundaje, A., & Boyle, A. P. (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Scientific Reports 2019 9:1*, *9*(1), 1–5. https://doi.org/10.1038/s41598-019-45839-z

Anantha, R. W., Vassin, V. M., & Borowiec, J. A. (2007). Sequential and Synergistic Modification of Human RPA Stimulates Chromosomal DNA Repair. *Journal of Biological Chemistry*, *282*(49), 35910–35923. https://doi.org/10.1074/JBC.M704645200

Andersen, P. R., Tirian, L., Vunjak, M., & Brennecke, J. (2017). A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature 2017 549:7670*, *549*(7670), 54–59. https://doi.org/10.1038/nature23482

Ashour, M. E., & Mosammaparast, N. (2021). Mechanisms of damage tolerance and repair during DNA replication. *Nucleic Acids Research*, *49*(6), 3033–3047. https://doi.org/10.1093/NAR/GKAB101

Aze, A., & Maiorano, D. (2018). Recent advances in understanding DNA replication: cell type–specific adaptation of the DNA replication program. *F1000Research 2018 7:1351*, *7*, 1351. https://doi.org/10.12688/f1000research.15408.1

Bakkenist, C. J., & Kastan, M. B. (2003). DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature 2003 421:6922*, *421*(6922), 499–506. https://doi.org/10.1038/nature01368

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Research 2011 21:3*, *21*(3), 381–395. https://doi.org/10.1038/cr.2011.22

Barlow, J. H., Faryabi, R. B., Callén, E., Wong, N., Malhowski, A., Chen, H. T., Gutierrez-Cruz, G., Sun, H. W., McKinnon, P., Wright, G., Casellas, R., Robbiani, D. F., Staudt, L.,

Fernandez-Capetillo, O., & Nussenzweig, A. (2013). A novel class of early replicating fragile sites that contribute to genome instability in B cell lymphomas. *Cell*, *152*(3), 620. https://doi.org/10.1016/J.CELL.2013.01.006

Bayona-Feliu, A., Barroso, S., Muñoz, S., & Aguilera, A. (2021). The SWI/SNF chromatin remodeling complex helps resolve R-loop-mediated transcription–replication conflicts. *Nature Genetics 2021 53:7*, *53*(7), 1050–1063. https://doi.org/10.1038/s41588-021-00867-2

Bayona-Feliu, A., Herrera-Moyano, E., Badra-Fajardo, N., Galván-Femenía, I., Soler-Oliva, M. E., & Aguilera, A. (2023). The chromatin network helps prevent cancer-associated mutagenesis at transcription-replication conflicts. *Nature Communications 2023 14:1*, *14*(1), 1–16. https://doi.org/10.1038/s41467-023-42653-0

Beagan, J. A., & Phillips-Cremins, J. E. (2020). On the existence and functionality of topologically associating domains. *Nature Genetics 2020 52:1*, *52*(1), 8–16. https://doi.org/10.1038/s41588-019-0561-1

Bell, S. P., Kobayashi, R., & Stillman, B. (1993). Yeast Origin Recognition Complex Functions in Transcription Silencing and DNA Replication. *Science*, *262*(5141), 1844–1849. https://doi.org/10.1126/SCIENCE.8266072

Belotserkovskii, B. P., Ng, S. yan, & Hanawalt, P. C. (2020). Transcription Inhibition by PNA-Induced R-Loops. *Methods in Molecular Biology (Clifton, N.J.)*, *2105*, 141–155. https://doi.org/10.1007/978-1-0716-0243-0_8

Bermejo, R., Lai, M. S., & Foiani, M. (2012). Preventing Replication Stress to Maintain Genome Stability: Resolving Conflicts between Replication and Transcription. *Molecular Cell*, *45*(6), 710–718. https://doi.org/10.1016/J.MOLCEL.2012.03.001/ASSET/92257FD9-809C-4A3F-95C5-A83EB55F0004/MAIN.ASSETS/GR5.JPG

Bernecky, C., Plitzko, J. M., & Cramer, P. (2017). Structure of a transcribing RNA polymerase II–DSIF complex reveals a multidentate DNA–RNA clamp. *Nature Structural & Molecular Biology 2017 24:10*, *24*(10), 809–815. https://doi.org/10.1038/nsmb.3465

Beroukhim, R., Mermel, C. H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J. S., Dobson, J., Urashima, M., McHenry, K. T., Pinchback, R. M., Ligon, A. H., Cho, Y. J., Haery, L., Greulich, H., Reich, M., Winckler, W., Lawrence, M. S., … Meyerson, M. (2010). The landscape of somatic copy-number alteration across human cancers. *Nature 2010 463:7283*, *463*(7283), 899–905. https://doi.org/10.1038/nature08822

Bhowmick, R., Mehta, K. P. M., Lerdrup, M., & Cortez, D. (2023). Integrator facilitates RNAPII removal to prevent transcription-replication collisions and genome instability. *Molecular Cell*, *83*(13), 2357. https://doi.org/10.1016/J.MOLCEL.2023.05.015

Bignell, G. R., Greenman, C. D., Davies, H., Butler, A. P., Edkins, S., Andrews, J. M., Buck, G., Chen, L., Beare, D., Latimer, C., Widaa, S., Hinton, J., Fahey, C., Fu, B., Swamy, S., Dalgliesh, G. L., Teh, B. T., Deloukas, P., Yang, F., … Stratton, M. R. (2010). Signatures of mutation and selection in the cancer genome. *Nature 2010 463:7283*, *463*(7283), 893–898. https://doi.org/10.1038/nature08768

Bodnar, A. G., Ouellette, M., Frolkis, M., Holt, S. E., Chiu, C. P., Morin, G. B., Harley, C. B., Shay, J. W., Lichtsteiner, S., & Wright, W. E. (1998). Extension of life-span by introduction of telomerase into normal human cells. *Science*, *279*(5349), 349–352. https://doi.org/10.1126/SCIENCE.279.5349.349/ASSET/41792167-CBD2-48A3-BCF0-

99DE013E23B0/ASSETS/GRAPHIC/SE038621104A.JPEG

Borlado, L. R., & Méndez, J. (2008). CDC6: from DNA replication to cell cycle checkpoints and oncogenesis. *Carcinogenesis*, *29*(2), 237–243. https://doi.org/10.1093/CARCIN/BGM268

Bowry, A., Kelly, R. D. W., & Petermann, E. (2021). Hypertranscription and replication stress in cancer. *Trends in Cancer*, *7*(9), 863–877. https://doi.org/10.1016/J.TRECAN.2021.04.006/ASSET/882B14A5-436B-4FE7-8308-D9048EC6BC96/MAIN.ASSETS/GR4.JPG

Brambilla, F., Garcia-Manteiga, J. M., Monteleone, E., Hoelzen, L., Zocchi, A., Agresti, A., & Bianchi, M. E. (2020). Nucleosomes effectively shield DNA from radiation damage in living cells. *Nucleic Acids Research*, *48*(16), 8993–9006. https://doi.org/10.1093/NAR/GKAA613

Brewer, B. J. (1988). When polymerases collide: Replication and the transcriptional organization of the E. coli chromosome. *Cell*, *53*(5), 679–686. https://doi.org/10.1016/0092-8674(88)90086-4

Brickner, J. R., Garzon, J. L., & Cimprich, K. A. (2022). Walking a tightrope: The complex balancing act of R-loops in genome stability. *Molecular Cell*, *82*(12), 2267–2297. https://doi.org/10.1016/J.MOLCEL.2022.04.014

Britton, S., Chanut, P., Delteil, C., Barboule, N., Frit, P., & Calsou, P. (2020). ATM antagonizes NHEJ proteins assembly and DNA-ends synapsis at single-ended DNA double strand breaks. *Nucleic Acids Research*, *48*(17), 9710–9723. https://doi.org/10.1093/NAR/GKAA723

Broad Institute. (2019). *Picard Toolkit* (p. GitHub Repository). https://broadinstitute.github.io/picard/

Brüning, J. G., & Marians, K. J. (2020). Replisome bypass of transcription complexes and R-loops. *Nucleic Acids Research*, *48*(18), 10353–10367. https://doi.org/10.1093/NAR/GKAA741

Bruno, F., Coronel-Guisado, C., & González-Aguilera, C. (2024). Collisions of RNA polymerases behind the replication fork promote alternative RNA splicing in newly replicated chromatin. *Molecular Cell*, *84*(2), 221-233.e6. https://doi.org/10.1016/J.MOLCEL.2023.11.036/ATTACHMENT/65D95DD0-92BF-482F-8B6D-8CD406D852F5/MMC6.PDF

Buoninfante, O. A., Pilzecker, B., Spanjaard, A., de Groot, D., Prekovic, S., Song, J. Y., Lieftink, C., Ayidah, M., Pritchard, C. E. J., Vivié, J., McGrath, K. E., Huijbers, I. J., Philipsen, S., von Lindern, M., Zwart, W., Beijersbergen, R. L., Palis, J., van den Berk, P. C. M., & Jacobs, H. (2023). Mammalian life depends on two distinct pathways of DNA damage tolerance. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(4), e2216055120. https://doi.org/10.1073/PNAS.2216055120/SUPPL_FILE/PNAS.2216055120.SAPP.PDF

Buratowski, S. (2009). Progression through the RNA Polymerase II CTD Cycle. *Molecular Cell*, *36*(4), 541–546. https://doi.org/10.1016/J.MOLCEL.2009.10.019

Carrasco-Salas, Y., Malapert, A., Sulthana, S., Molcrette, B., Chazot-Franguiadakis, L., Bernard, P., Chédin, F., Faivre-Moskalenko, C., & Vanoosthuyse, V. (2019). The extruded non-template strand determines the architecture of R-loops. *Nucleic Acids Research*, *47*(13), 6783–6795. https://doi.org/10.1093/NAR/GKZ341

Carroll, S. M., Trotter, J., & Wahl, G. M. (1991). Replication timing control can be maintained in extrachromosomally amplified genes. *Molecular and Cellular Biology*, *11*(9), 4779–4785. https://doi.org/10.1128/MCB.11.9.4779-4785.1991

Castillo-Guzman, D., & Chédin, F. (2021). Defining R-loop classes and their contributions to genome instability. *DNA Repair*, *106*, 103182. https://doi.org/10.1016/J.DNAREP.2021.103182

Cavalheiro, G. R., Pollex, T., & Furlong, E. E. (2021). To loop or not to loop: what is the role of TADs in enhancer function and gene regulation? *Current Opinion in Genetics & Development*, *67*, 119–129. https://doi.org/10.1016/J.GDE.2020.12.015

Chang, E. Y. C., Novoa, C. A., Aristizabal, M. J., Coulombe, Y., Segovia, R., Chaturvedi, R., Shen, Y., Keong, C., Tam, A. S., Jones, S. J. M., Masson, J. Y., Kobor, M. S., & Stirling, P. C. (2017). RECQ-like helicases Sgs1 and BLM regulate R-loop–associated genome instability. *Journal of Cell Biology*, *216*(12), 3991–4005. https://doi.org/10.1083/JCB.201703168

Chanut, P., Britton, S., Coates, J., Jackson, S. P., & Calsou, P. (2016). Coordinated nuclease activities counteract Ku at single-ended DNA double-strand breaks. *Nature Communications 2016 7:1*, *7*(1), 1–12. https://doi.org/10.1038/ncomms12889

Chappidi, N., Nascakova, Z., Boleslavska, B., Zellweger, R., Isik, E., Andrs, M., Menon, S., Dobrovolna, J., Balbo Pogliano, C., Matos, J., Porro, A., Lopes, M., & Janscak, P. (2020). Fork Cleavage-Religation Cycle and Active Transcription Mediate Replication Restart after Fork Stalling at Co-transcriptional R-Loops. *Molecular Cell*, *77*(3), 528-541.e8. https://doi.org/10.1016/J.MOLCEL.2019.10.026

Chédin, F. (2016). Nascent Connections: R-Loops and Chromatin Patterning. *Trends in Genetics*, *32*(12), 828–838. https://doi.org/10.1016/J.TIG.2016.10.002

Choi, J. H., Lindsey-Boltz, L. A., Kemp, M., Mason, A. C., Wold, M. S., & Sancar, A. (2010). Reconstitution of RPA-covered single-stranded DNA-activated ATR-Chk1 signaling. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(31), 13660–13665. https://doi.org/10.1073/PNAS.1007856107/SUPPL_FILE/PNAS.1007856107_SI.PDF

Chong, S. Y., Cutler, S., Lin, J. J., Tsai, C. H., Tsai, H. K., Biggins, S., Tsukiyama, T., Lo, Y. C., & Kao, C. F. (2020). H3K4 methylation at active genes mitigates transcription-replication conflicts during replication stress. *Nature Communications*, *11*(1). https://doi.org/10.1038/S41467-020-14595-4

Collins, P. L., Purman, C., Porter, S. I., Nganga, V., Saini, A., Hayer, K. E., Gurewitz, G. L., Sleckman, B. P., Bednarski, J. J., Bassing, C. H., & Oltz, E. M. (2020). DNA double-strand breaks induce H2Ax phosphorylation domains in a contact-dependent manner. *Nature Communications 2020 11:1*, *11*(1), 1–9. https://doi.org/10.1038/s41467-020-16926-x

Conti, B. A., & Smogorzewska, A. (2020). Mechanisms of direct replication restart at stressed replisomes. *DNA Repair*, *95*. https://doi.org/10.1016/J.DNAREP.2020.102947

Corazzi, L., Ionasz, V. S., Andrejev, S., Wang, L. C., Vouzas, A., Giaisi, M., Di Muzio, G., Ding, B., Marx, A. J. M., Henkenjohann, J., Allers, M. M., Gilbert, D. M., & Wei, P. C. (2024). Linear interaction between replication and transcription shapes DNA break dynamics at recurrent DNA break Clusters. *Nature Communications 2024 15:1*, *15*(1), 1–13. https://doi.org/10.1038/s41467-024-47934-w

Core, L., & Adelman, K. (2019). Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes & Development*, *33*(15–16), 960–982. https://doi.org/10.1101/GAD.325142.119

Cortez, D., Guntuku, S., Qin, J., & Elledge, S. J. (2001). ATR and ATRIP: Partners in checkpoint signaling. *Science*, *294*(5547), 1713–1716. https://doi.org/10.1126/SCIENCE.1065521/SUPPL_FILE/1066521S3_THUMB.GIF

Cramer, P. (2019). Organization and regulation of gene transcription. *Nature 2019 573:7772*, *573*(7772), 45–54. https://doi.org/10.1038/s41586-019-1517-4

Cristini, A., Tellier, M., Constantinescu, F., Accalai, C., Albulescu, L. O., Heiringhoff, R., Bery, N., Sordet, O., Murphy, S., & Gromak, N. (2022). RNase H2, mutated in Aicardi-Goutières syndrome, resolves co-transcriptional R-loops to prevent DNA breaks and inflammation. *Nature Communications*, *13*(1). https://doi.org/10.1038/S41467-022-30604-0

Crossley, M. P., Bocek, M. J., Hamperl, S., Swigut, T., & Cimprich, K. A. (2020). qDRIP: a method to quantitatively assess RNA–DNA hybrid formation genome-wide. *Nucleic Acids Research*, *48*(14), e84–e84. https://doi.org/10.1093/NAR/GKAA500

Crossley, M. P., Brickner, J. R., Song, C., Zar, S. M. T., Maw, S. S., Chédin, F., Tsai, M. S., & Cimprich, K. A. (2021). Catalytically inactive, purified rnase h1: A specific and sensitive probe for rna–dna hybrid imaging. *Journal of Cell Biology*, *220*(9). https://doi.org/10.1083/JCB.202101092/212458

Daigle, S. R., Olhava, E. J., Therkelsen, C. A., Basavapathruni, A., Jin, L., Boriack-Sjodin, P. A., Allain, C. J., Klaus, C. R., Raimondi, A., Scott, M. P., Waters, N. J., Chesworth, R., Moyer, M. P., Copeland, R. A., Richon, V. M., & Pollock, R. M. (2013). Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood*, *122*(6), 1017–1025. https://doi.org/10.1182/BLOOD-2013-04-497644

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., & Davies, R. M. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2). https://doi.org/10.1093/GIGASCIENCE/GIAB008

de Amorim, J. L., Leung, S. W., Haji-Seyed-Javadi, R., Hou, Y., Yu, D. S., Ghalei, H., Khoshnevis, S., Yao, B., & Corbett, A. H. (2024). The putative RNA helicase DDX1 associates with the nuclear RNA exosome and modulates RNA/DNA hybrids (R-loops). *Journal of Biological Chemistry*, *300*(2), 105646–105647. https://doi.org/10.1016/J.JBC.2024.105646/ATTACHMENT/64BCFBAA-9C42-4EEA-90FB-65B5D8522BCA/MMC2.XLSX

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & Development*, *25*(10), 1010–1022. https://doi.org/10.1101/GAD.2037511

Debatisse, M., Le Tallec, B., Letessier, A., Dutrillaux, B., & Brison, O. (2012). Common fragile sites: mechanisms of instability revisited. *Trends in Genetics*, *28*(1), 22–32. https://doi.org/10.1016/J.TIG.2011.10.003

Debès, C., Papadakis, A., Grönke, S., Karalay, Ö., Tain, L. S., Mizi, A., Nakamura, S., Hahn, O., Weigelt, C., Josipovic, N., Zirkel, A., Brusius, I., Sofiadis, K., Lamprousi, M., Lu, Y. X., Huang, W., Esmaillie, R., Kubacki, T., Späth, M. R., … Beyer, A. (2023). Ageing-associated changes in transcriptional elongation influence longevity. *Nature 2023 616:7958*, *616*(7958), 814–821. https://doi.org/10.1038/s41586-023-05922-y

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., & Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature 2012 485:7398*, *485*(7398), 376–380. https://doi.org/10.1038/nature11082

Dong, X., Yang, Q., Du, Z., Zhang, G., Shi, C., Qin, X., & Song, Y. (2022). Identification of optimal reference genes for gene expression normalization in human osteosarcoma cell lines under proliferative conditions. *Frontiers in Genetics*, *13*, 989990. https://doi.org/10.3389/FGENE.2022.989990/BIBTEX

Donzelli, M., & Draetta, G. F. (2003). Regulating mammalian checkpoints through Cdc25 inactivation. *EMBO Reports*, *4*(7), 671–677. https://doi.org/10.1038/SJ.EMBOR.EMBOR887/ASSET/1C30CCD2-4902-4882-8D2E-82A2B3BA7734/ASSETS/GRAPHIC/EMBR887-GRA-0001-M.JPG

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B. K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., … Lochovsky, L. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. https://doi.org/10.1038/NATURE11247

Dunn, K., & Griffith, J. D. (1980). The presence of RNA in a double helix inhibits its interaction with histone protein. *Nucleic Acids Research*, *8*(3), 555–566. https://doi.org/10.1093/NAR/8.3.555

Durkin, S. G., & Glover, T. W. (2007). Chromosome fragile sites. *Annual Review of Genetics*, *41*(Volume 41, 2007), 169–192. https://doi.org/10.1146/ANNUREV.GENET.41.042007.165900/CITE/REFWORKS

Ebmeier, C. C., Erickson, B., Allen, B. L., Allen, M. A., Kim, H., Fong, N., Jacobsen, J. R., Liang, K., Shilatifard, A., Dowell, R. D., Old, W. M., Bentley, D. L., & Taatjes, D. J. (2017). Human TFIIH Kinase CDK7 Regulates Transcription-Associated Chromatin Modifications. *Cell Reports*, *20*(5), 1173–1186. https://doi.org/10.1016/J.CELREP.2017.07.021

Eick, D., & Bornkamm, G. W. (1986). Transcriptional arrest within the first exon is a fast control mechanism in c-myc gene expression. *Nucleic Acids Research*, *14*(21), 8331–8346. https://doi.org/10.1093/NAR/14.21.8331

Eick, D., & Geyer, M. (2013). The RNA polymerase II carboxy-terminal domain (CTD) code. *Chemical Reviews*, *113*(11), 8456–8490. https://doi.org/10.1021/CR400071F/ASSET/IMAGES/LARGE/CR-2013-00071F_0030.JPEG

Emerson, D. J., Zhao, P. A., Cook, A. L., Barnett, R. J., Klein, K. N., Saulebekova, D., Ge, C., Zhou, L., Simandi, Z., Minsk, M. K., Titus, K. R., Wang, W., Gong, W., Zhang, D., Yang, L., Venev, S. V., Gibcus, J. H., Yang, H., Sasaki, T., … Phillips-Cremins, J. E. (2022). Cohesin-mediated loop anchors confine the locations of human replication origins. *Nature 2022 606:7915*, *606*(7915), 812–819. https://doi.org/10.1038/s41586-022-04803-0

Falck, J., Coates, J., & Jackson, S. P. (2005). Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature 2005 434:7033*, *434*(7033), 605–611. https://doi.org/10.1038/nature03442

Feng, X., Noguchi, Y., Barbon, M., Stillman, B., Speck, C., & Li, H. (2021). The structure of ORC–Cdc6 on an origin DNA reveals the mechanism of ORC activation by the replication

initiator Cdc6. *Nature Communications 2021 12:1*, *12*(1), 1–12. https://doi.org/10.1038/s41467-021-24199-1

Fenstermaker, T. K., Petruk, S., Kovermann, S. K., Brock, H. W., & Mazo, A. (2023). RNA polymerase II associates with active genes during DNA replication. *Nature 2023 620:7973*, *620*(7973), 426–433. https://doi.org/10.1038/s41586-023-06341-9

Foss, M., McNally, F. J., Laurenson, P., & Rine, J. (1993). Origin Recognition Complex (ORC) in Transcriptional Silencing and DNA Replication in S. cerevisiae. *Science*, *262*(5141), 1838–1844. https://doi.org/10.1126/SCIENCE.8266071

Fragkos, M., Ganier, O., Coulombe, P., & Méchali, M. (2015). DNA replication origin activation in space and time. *Nature Reviews Molecular Cell Biology 2015 16:6*, *16*(6), 360–374. https://doi.org/10.1038/nrm4002

Fu, H., Maunakea, A. K., Martin, M. M., Huang, L., Zhang, Y., Ryan, M., Kim, R. G., Lin, C. M., Zhao, K., & Aladjem, M. I. (2013). Methylation of Histone H3 on Lysine 79 Associates with a Group of Replication Origins and Helps Limit DNA Replication Once per Cell Cycle. *PLOS Genetics*, *9*(6), e1003542. https://doi.org/10.1371/JOURNAL.PGEN.1003542

Gabriele, M., Brandão, H. B., Grosse-Holz, S., Jha, A., Dailey, G. M., Cattoglio, C., Hsieh, T. H. S., Mirny, L., Zechner, C., & Hansen, A. S. (2022). Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science*, *376*(6592), 476–501. https://doi.org/10.1126/SCIENCE.ABN6583/SUPPL_FILE/SCIENCE.ABN6583_MDAR_REPRODUCIBILITY_CHECKLIST.PDF

García-Gómez, S., Reyes, A., Martínez-Jiménez, M. I., Chocrón, E. S., Mourón, S., Terrados, G., Powell, C., Salido, E., Méndez, J., Holt, I. J., & Blanco, L. (2013). PrimPol, an Archaic Primase/Polymerase Operating in Human Cells. *Molecular Cell*, *52*(4), 541–553. https://doi.org/10.1016/j.molcel.2013.09.025

García-Muse, T., & Aguilera, A. (2016). Transcription–replication conflicts: how they occur and how they are resolved. *Nature Reviews Molecular Cell Biology 2016 17:9*, *17*(9), 553–563. https://doi.org/10.1038/nrm.2016.88

García-Muse, T., & Aguilera, A. (2019). R Loops: From Physiological to Pathological Roles. *Cell*, *179*(3), 604–618. https://doi.org/10.1016/J.CELL.2019.08.055/ASSET/943352A4-2857-41FD-B4D1-27720DE49912/MAIN.ASSETS/GR4.JPG

García-Pichardo, D., Cañas, J. C., García-Rubio, M. L., Gómez-González, B., Rondón, A. G., & Aguilera, A. (2017). Histone Mutants Separate R Loop Formation from Genome Instability Induction. *Molecular Cell*, *66*(5), 597-609.e5. https://doi.org/10.1016/J.MOLCEL.2017.05.014

García-Rubio, M., Aguilera, P., Lafuente-Barquero, J., Ruiz, J. F., Simon, M. N., Geli, V., Rondón, A. G., & Aguilera, A. (2018). Yra1-bound RNA–DNA hybrids cause orientation-independent transcription–replication collisions and telomere instability. *Genes & Development*, *32*(13–14), 965–977. https://doi.org/10.1101/GAD.311274.117

Garg, P., & Burgers, P. M. J. (2005). DNA Polymerases that Propagate the Eukaryotic DNA Replication Fork. *Critical Reviews in Biochemistry and Molecular Biology*, *40*(2), 115–128. https://doi.org/10.1080/10409230590935433

Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I., & Chédin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Molecular Cell*,

45(6), 814. https://doi.org/10.1016/J.MOLCEL.2012.01.017

Glover, T. W. (1981). FUdR induction of the X chromosome fragile site: evidence for the mechanism of folic acid and thymidine inhibition. *American Journal of Human Genetics*, *33*(2), 234. https://pmc.ncbi.nlm.nih.gov/articles/PMC1684940/

Glover, Thomas W., Berger, C., Coyle, J., & Echo, B. (1984). DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Human Genetics*, *67*(2), 136–142. https://doi.org/10.1007/BF00272988

Gómez-González, B., & Aguilera, A. (2019). Transcription-mediated replication hindrance: a major driver of genome instability. *Genes & Development*, *33*(15–16), 1008–1026. https://doi.org/10.1101/GAD.324517.119

Grewal, S. I. S. (2023). The molecular basis of heterochromatin assembly and epigenetic inheritance. *Molecular Cell*, *83*(11), 1767–1785. https://doi.org/10.1016/J.MOLCEL.2023.04.020

Groelly, F. J., Dagg, R. A., Petropoulos, M., Rossetti, G. G., Prasad, B., Panagopoulos, A., Paulsen, T., Karamichali, A., Jones, S. E., Ochs, F., Dionellis, V. S., Puig Lombardi, E., Miossec, M. J., Lockstone, H., Legube, G., Blackford, A. N., Altmeyer, M., Halazonetis, T. D., & Tarsounas, M. (2022). Mitotic DNA synthesis is caused by transcription-replication conflicts in BRCA2-deficient cells. *Molecular Cell*, *82*(18), 3382-3397.e7. https://doi.org/10.1016/J.MOLCEL.2022.07.011/ATTACHMENT/DC964983-D251-41F4-A956-64D754258CF4/MMC6.XLSX

Grünberg, S., Warfield, L., & Hahn, S. (2012). Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. *Nature Structural & Molecular Biology 2012 19:8*, *19*(8), 788–796. https://doi.org/10.1038/nsmb.2334

Gyenis, A., Chang, J., Demmers, J. J. P. G., Bruens, S. T., Barnhoorn, S., Brandt, R. M. C., Baar, M. P., Raseta, M., Derks, K. W. J., Hoeijmakers, J. H. J., & Pothof, J. (2023). Genome-wide RNA polymerase stalling shapes the transcriptome during aging. *Nature Genetics 2023 55:2*, *55*(2), 268–279. https://doi.org/10.1038/s41588-022-01279-6

Haahr, P., Hoffmann, S., Tollenaere, M. A. X., Ho, T., Toledo, L. I., Mann, M., Bekker-Jensen, S., Räschle, M., & Mailand, N. (2016). Activation of the ATR kinase by the RPA-binding protein ETAA1. *Nature Cell Biology 2016 18:11*, *18*(11), 1196–1207. https://doi.org/10.1038/ncb3422

Halicka, D., Zhao, H., Li, J., Garcia, J., Podhorecka, M., & Darzynkiewicz, Z. (2017). DNA Damage Response Resulting from Replication Stress Induced by Synchronization of Cells by Inhibitors of DNA Replication: Analysis by Flow Cytometry. *Methods in Molecular Biology (Clifton, N.J.)*, *1524*, 107–119. https://doi.org/10.1007/978-1-4939-6603-5_7

Hamperl, S., Bocek, M. J., Saldivar, J. C., Swigut, T., & Cimprich, K. A. (2017). Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell*, *170*(4), 774-786.e19. https://doi.org/10.1016/J.CELL.2017.07.043/ATTACHMENT/F6E6E744-1D91-4E98-9A58-7AB2441A7ADA/MMC2.XLSX

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, *100*(1), 57–70. https://doi.org/10.1016/S0092-8674(00)81683-9/ASSET/41DCC706-DFFF-4E15-8B5C-F071F908D8EA/MAIN.ASSETS/GR4.JPG

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, *144*(5), 646–674. https://doi.org/10.1016/J.CELL.2011.02.013/ASSET/2067D218-2368-451A-B4DA-CF9EF4807B23/MAIN.ASSETS/GR1.JPG

Hansen, R. S., Thomas, S., Sandstrom, R., Canfield, T. K., Thurman, R. E., Weaver, M., Dorschner, M. O., Gartler, S. M., & Stamatoyannopoulos, J. A. (2010). Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(1), 139–144. https://doi.org/10.1073/PNAS.0912402107/SUPPL_FILE/ST05.DOC

Hao, S., Wang, Y., Zhao, Y., Gao, W., Cui, W., Li, Y., Cui, J., Liu, Y., Lin, L., Xu, X., & Wang, H. (2022). Dynamic switching of crotonylation to ubiquitination of H2A at lysine 119 attenuates transcription–replication conflicts caused by replication stress. *Nucleic Acids Research*, *50*(17), 9873. https://doi.org/10.1093/NAR/GKAC734

Hasanova, Z., Klapstova, V., Porrua, O., Stefl, R., & Sebesta, M. (2023). Human senataxin is a bona fide R-loop resolving enzyme and transcription termination factor. *Nucleic Acids Research*, *51*(6), 2818–2837. https://doi.org/10.1093/NAR/GKAD092

He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., & Nogales, E. (2016). Near-atomic resolution visualization of human transcription promoter opening. *Nature 2016 533:7603*, *533*(7603), 359–365. https://doi.org/10.1038/nature17970

Hegazy, M., Cohen-Barak, E., Koetsier, J. L., Najor, N. A., Arvanitis, C., Sprecher, E., Green, K. J., & Godsel, L. M. (2020). Proximity Ligation Assay for Detecting Protein-Protein Interactions and Protein Modifications in Cells and Tissues In Situ. *Current Protocols in Cell Biology*, *89*(1), e115. https://doi.org/10.1002/CPCB.115

Hellman, A., Zlotorynski, E., Scherer, S. W., Cheung, J., Vincent, J. B., Smith, D. I., Trakhtenbrot, L., & Kerem, B. (2002). A role for common fragile site induction in amplification of human oncogenes. *Cancer Cell*, *1*(1), 89–97. https://doi.org/10.1016/S1535-6108(02)00017-X

Helmrich, A., Ballarino, M., & Tora, L. (2011). Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Molecular Cell*, *44*(6), 966–977. https://doi.org/10.1016/J.MOLCEL.2011.10.013/ATTACHMENT/0EBAC4C2-34C6-4518-AEE2-EC5EF7366ECB/MMC1.PDF

Hergeth, S. P., & Schneider, R. (2015). The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Reports*, *16*(11), 1439. https://doi.org/10.15252/EMBR.201540749

Hildebrand, E. M., & Dekker, J. (2020). Mechanisms and functions of chromosome compartmentalization. *Trends in Biochemical Sciences*, *45*(5), 385. https://doi.org/10.1016/J.TIBS.2020.01.002

Hobson, D. J., Wei, W., Steinmetz, L. M., & Svejstrup, J. Q. (2012). RNA polymerase II collision interrupts convergent transcription. *Molecular Cell*, *48*(3), 365–374. https://doi.org/10.1016/J.MOLCEL.2012.08.027

Hodin, T. L., Najrana, T., & Yates, J. L. (2013). Efficient Replication of Epstein-Barr Virus-Derived Plasmids Requires Tethering by EBNA1 to Host Chromosomes. *Journal of Virology*, *87*(23), 13020. https://doi.org/10.1128/JVI.01606-13

Hodroj, D., Recolin, B., Serhal, K., Martinez, S., Tsanov, N., Abou Merhi, R., & Maiorano, D. (2017). An ATR -dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism . *The EMBO Journal*, *36*(9), 1182–1198. https://doi.org/10.15252/EMBJ.201695131/SUPPL_FILE/EMBJ201695131-SUP-0017-MOVIEEV15.ZIP

Holstege, F. C. P., Van Der Vliet, P. C., & Timmers, H. T. M. (1996). Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *The EMBO Journal*, *15*(7), 1666–1677. https://doi.org/10.1002/J.1460-2075.1996.TB00512.X

Hu, Y., & Stillman, B. (2023). Origins of DNA replication in eukaryotes. *Molecular Cell*, *83*(3), 352–372. https://doi.org/10.1016/J.MOLCEL.2022.12.024

Huertas, P., & Aguilera, A. (2003). Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Molecular Cell*, *12*(3), 711–721. https://doi.org/10.1016/j.molcel.2003.08.010

Huyen, Y., Zgheib, O., DiTullio, R. A., Gorgoulis, V. G., Zacharatos, P., Petty, T. J., Sheston, E. A., Mellert, H. S., Stavridi, E. S., & Halazonetis, T. D. (2004). Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature 2004 432:7015*, *432*(7015), 406–411. https://doi.org/10.1038/nature03114

Hyjek, M., Figiel, M., & Nowotny, M. (2019). RNases H: Structure and mechanism. *DNA Repair*, *84*, 102672. https://doi.org/10.1016/J.DNAREP.2019.102672

Hyrien, O. (2016). How MCM loading and spreading specify eukaryotic DNA replication initiation sites. *F1000Research*, *5*, F1000 Faculty Rev-2063. https://doi.org/10.12688/F1000RESEARCH.9008.1

Ilves, I., Petojevic, T., Pesavento, J. J., & Botchan, M. R. (2010). Activation of the MCM2-7 Helicase by Association with Cdc45 and GINS proteins. *Molecular Cell*, *37*(2), 247. https://doi.org/10.1016/J.MOLCEL.2009.12.030

Izsvák, Z., & Ivics, Z. (2004). Sleeping Beauty Transposition: Biology and Applications for Molecular Therapy. *Molecular Therapy*, *9*(2), 147–156. https://doi.org/10.1016/J.YMTHE.2003.11.009

Jiang, C., & Pugh, B. F. (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nature Reviews Genetics 2009 10:3*, *10*(3), 161–172. https://doi.org/10.1038/nrg2522

Jin, S., Mao, H., Schnepp, R. W., Sykes, S. M., Silva, A. C., D'andrea, A. D., & Hua, X. (2003). Menin Associates with FANCD2, a Protein Involved in Repair of DNA Damage 1. *CANCER RESEARCH*, *63*, 4204–4210. http://aacrjournals.org/cancerres/article-pdf/63/14/4204/2504866/4204.pdf

Julienne, H., Zoufir, A., Audit, B., & Arneodo, A. (2013). Human Genome Replication Proceeds through Four Chromatin States. *PLOS Computational Biology*, *9*(10), e1003233. https://doi.org/10.1371/JOURNAL.PCBI.1003233

Kang, J. Y., Park, J. W., Hahm, J. Y., Jung, H., & Seo, S. B. (2020). Histone H3K79 demethylation by KDM2B facilitates proper DNA replication through PCNA dissociation from chromatin. *Cell Proliferation*, *53*(11), e12920. https://doi.org/10.1111/CPR.12920

Kari, V., Raul, S. K., Henck, J. M., Kitz, J., Kramer, F., Kosinsky, R. L., Übelmesser, N., Mansour, W. Y., Eggert, J., Spitzner, M., Najafova, Z., Bastians, H., Grade, M., Gaedcke, J., Wegwitz, F., & Johnsen, S. A. (2019). The histone methyltransferase DOT1L is required for proper DNA damage response, DNA repair, and modulates chemotherapy responsiveness. *Clinical Epigenetics*, *11*(1), 1–13. https://doi.org/10.1186/S13148-018-0601-1/TABLES/3

Kastan, M. B., Zhan, Q., El-Deiry, W. S., Carrier, F., Jacks, T., Walsh, W. V., Plunkett, B. S., Vogelstein, B., & Fornace, A. J. (1992). A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia. *Cell*, *71*(4), 587–597. https://doi.org/10.1016/0092-8674(92)90593-2

Khanna, K. K., Lavin, M. F., Jackson, S. P., & Mulhern, T. D. (2001). ATM, a central controller of cellular responses to DNA damage. *Cell Death & Differentiation 2001 8:11*, *8*(11), 1052–1065. https://doi.org/10.1038/sj.cdd.4400874

Kim, N., Abdulovic, A. L., Gealy, R., Lippert, M. J., & Jinks-Robertson, S. (2007). Transcription-associated mutagenesis in yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication. *DNA Repair*, *6*(9), 1285–1296. https://doi.org/10.1016/J.DNAREP.2007.02.023

Kim, Sangin, Kang, N., Park, S. H., Wells, J., Hwang, T., Ryu, E., Kim, B. G., Hwang, S., Kim, S. J., Kang, S., Lee, S., Stirling, P., Myung, K., & Lee, K. Y. (2020). ATAD5 restricts R-loop formation through PCNA unloading and RNA helicase maintenance at the replication fork. *Nucleic Acids Research*, *48*(13), 7218. https://doi.org/10.1093/NAR/GKAA501

Kim, Subin, Shin, W. H., Kang, Y., Kim, H., & Lee, J. Y. (2024). Direct visualization of replication and R-loop collision using single-molecule imaging. *Nucleic Acids Research*, *52*(1), 259–273. https://doi.org/10.1093/NAR/GKAD1101

Knezetic, J. A., & Luse, D. S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell*, *45*(1), 95–104. https://doi.org/10.1016/0092-8674(86)90541-6

Komseli, E. S., Pateras, I. S., Krejsgaard, T., Stawiski, K., Rizou, S. V., Polyzos, A., Roumelioti, F. M., Chiourea, M., Mourkioti, I., Paparouna, E., Zampetidis, C. P., Gumeni, S., Trougakos, I. P., Pefani, D. E., O'Neill, E., Gagos, S., Eliopoulos, A. G., Fendler, W., Chowdhury, D., … Gorgoulis, V. G. (2018). A prototypical non-malignant epithelial model to study genome dynamics and concurrently monitor micro-RNAs and proteins in situ during oncogene-induced senescence. *BMC Genomics 2018 19:1*, *19*(1), 1–22. https://doi.org/10.1186/S12864-017-4375-1

Kondratick, C. M., Washington, M. T., & Spies, M. (2021). Making choices: DNA replication fork recovery mechanisms. *Seminars in Cell & Developmental Biology*, *113*, 27–37. https://doi.org/10.1016/J.SEMCDB.2020.10.001

Kornberg, R. D. (2005). Mediator and the mechanism of transcriptional activation. *Trends in Biochemical Sciences*, *30*(5), 235–239. https://doi.org/10.1016/j.tibs.2005.03.011

Kostrewa, D., Zeller, M. E., Armache, K. J., Seizl, M., Leike, K., Thomm, M., & Cramer, P. (2009). RNA polymerase II–TFIIB structure and mechanism of transcription initiation. *Nature 2009 462:7271*, *462*(7271), 323–330. https://doi.org/10.1038/nature08548

Kotsantis, P., Silva, L. M., Irmscher, S., Jones, R. M., Folkes, L., Gromak, N., & Petermann, E. (2016). Increased global transcription activity as a mechanism of replication stress in

cancer. *Nature Communications 2016 7:1*, *7*(1), 1–13. https://doi.org/10.1038/ncomms13087

Kowarz, E., Löscher, D., & Marschalek, R. (2015). Optimized Sleeping Beauty transposons rapidly generate stable transgenic cell lines. *Biotechnology Journal*, *10*(4), 647–653. https://doi.org/10.1002/BIOT.201400821

Krueger, F., James, F., Ewels, P., Afyounian, E., Weinstein, M., Schuster-Boeckler, B., Hulselmans, G., & Sclamons. (2023). FelixKrueger/TrimGalore: v0.6.10 - add default decompression path. *Zenodo*. https://doi.org/10.5281/ZENODO.7598955

Lalonde, M., Trauner, M., Werner, M., & Hamperl, S. (2021). Consequences and Resolution of Transcription–Replication Conflicts. *Life 2021, Vol. 11, Page 637*, *11*(7), 637. https://doi.org/10.3390/LIFE11070637

Lalonde, M., Ummethum, H., Trauner, M., Ettinger, A., & Hamperl, S. (2023). An automated image analysis pipeline to quantify the coordination and overlap of transcription and replication activity in mammalian genomes. *Methods in Cell Biology*. https://doi.org/10.1016/BS.MCB.2023.05.012

Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T. R., & Weirauch, M. T. (2018). The Human Transcription Factors. *Cell*, *172*(4), 650–665. https://doi.org/10.1016/J.CELL.2018.01.029/ASSET/E832A18A-42E2-435D-8EC9-28F2343EEC8E/MAIN.ASSETS/GR4.JPG

Landsverk, H. B., Sandquist, L. E., Bay, L. T. E., Steurer, B., Campsteijn, C., Landsverk, O. J. B., Marteijn, J. A., Petermann, E., Trinkle-Mulcahy, L., & Syljuåsen, R. G. (2020). WDR82/PNUTS-PP1 Prevents Transcription-Replication Conflicts by Promoting RNA Polymerase II Degradation on Chromatin. *Cell Reports*, *33*(9). https://doi.org/10.1016/J.CELREP.2020.108469/ATTACHMENT/3831AACA-B6E7-418E-B07C-D155BE3D7758/MMC3.PDF

Lang, K. S., Hall, A. N., Merrikh, C. N., Ragheb, M., Tabakh, H., Pollock, A. J., Woodward, J. J., Dreifus, J. E., & Merrikh, H. (2017). Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis. *Cell*, *170*(4), 787-799.e18. https://doi.org/10.1016/J.CELL.2017.07.044/ATTACHMENT/F4B3DC86-6376-4705-9FA2-1CF30B2BB963/MMC1.PDF

Lang, K. S., & Merrikh, H. (2021). Topological stress is responsible for the detrimental outcomes of head-on replication-transcription conflicts. *Cell Reports*, *34*(9), 108797. https://doi.org/10.1016/J.CELREP.2021.108797

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/NMETH.1923

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., & Carey, V. J. (2013). Software for Computing and Annotating Genomic Ranges. *PLOS Computational Biology*, *9*(8), e1003118. https://doi.org/10.1371/JOURNAL.PCBI.1003118

Lee, Jeong Heon, You, J., Dobrota, E., & Skalnik, D. G. (2010). Identification and characterization of a novel human PP1 phosphatase complex. *Journal of Biological Chemistry*, *285*(32), 24466–24476. https://doi.org/10.1074/JBC.M110.109801/ASSET/24E36A22-1011-4139-BE51-842448D1F179/MAIN.ASSETS/GR9.JPG

Lee, Ji Hoon, & Paull, T. T. (2005). ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science*, *308*(5721), 551–554. https://doi.org/10.1126/SCIENCE.1108297/SUPPL_FILE/LEE.SOM.PDF

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics (Oxford, England)*, *28*(6), 882–883. https://doi.org/10.1093/BIOINFORMATICS/BTS034

Levens, D., Baranello, L., & Kouzine, F. (2016). Controlling gene expression by DNA mechanics: emerging insights and challenges. *Biophysical Reviews*, *8*(3), 259–268. https://doi.org/10.1007/S12551-016-0216-8/FIGURES/2

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–1760. https://doi.org/10.1093/BIOINFORMATICS/BTP324

Li, S., & Wu, X. (2020). Common fragile sites: protection and repair. *Cell & Bioscience 2020 10:1*, *10*(1), 1–9. https://doi.org/10.1186/S13578-020-00392-5

Liang, Z., Liang, F., Teng, Y., Chen, X., Liu, J., Longerich, S., Rao, T., Green, A. M., Collins, N. B., Xiong, Y., Lan, L., Sung, P., & Kupfer, G. M. (2019). Binding of FANCI-FANCD2 Complex to RNA and R-Loops Stimulates Robust FANCD2 Monoubiquitination. *Cell Reports*, *26*(3), 564-572.e5. https://doi.org/10.1016/J.CELREP.2018.12.084

Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*(5950), 289–293. https://doi.org/10.1126/SCIENCE.1181369/SUPPL_FILE/LIEBERMAN-AIDEN.SOM.PDF

Lin, J., Wu, Y., Tian, G., Yu, D., Yang, E., Lam, W. H., Liu, Z., Jing, Y., Dang, S., Bao, X., Hon Wong, J. W., Zhai, Y., & Li, X. D. (2023). Menin “reads” H3K79me2 mark in a nucleosomal context. *Science*, *379*(6633), 717–723. https://doi.org/10.1126/SCIENCE.ADC9318/SUPPL_FILE/SCIENCE.ADC9318_MOVIE_S1.ZIP

Lin, P., Lourenco, C., Cruickshank, J., Palomero, L., van Leeuwen, J. E., Tong, A. H. Y., Chan, K., Ghamrasni, S. El, Pujana, M. A., Cescon, D. W., Moffat, J., & Penn, L. Z. (2023). Topoisomerase 1 Inhibition in MYC-Driven Cancer Promotes Aberrant R-Loop Accumulation to Induce Synthetic Lethality. *Cancer Research*, *183*(24), 4015–4029. https://doi.org/10.1158/0008-5472.CAN-22-2948/730183/P/TOPOISOMERASE-1-INHIBITION-IN-MYC-DRIVEN-CANCER

Lin, W. L., Chen, J. K., Wen, X., He, W., Zarceno, G. A., Chen, Y., Chen, S., Paull, T. T., & Liu, H. wen. (2022). DDX18 prevents R-loop-induced DNA damage and genome instability via PARP-1. *Cell Reports*, *40*(3), 111089. https://doi.org/10.1016/J.CELREP.2022.111089

Lindstrand, A., Eisfeldt, J., Vezzi, F., Olason, P., & Nilsson, D. (2017). TIDDIT, an efficient and comprehensive structural variant caller for massive parallel sequencing data. *F1000Research*, *6*. https://doi.org/10.12688/F1000RESEARCH.11168.2/DOI

Liu, G., Geurts, A. M., Yae, K., Srinivasan, A. R., Fahrenkrug, S. C., Largaespada, D. A., Takeda, J., Horie, K., Olson, W. K., & Hackett, P. B. (2005). Target-site Preferences of

Sleeping Beauty Transposons. *Journal of Molecular Biology*, *346*(1), 161–173. https://doi.org/10.1016/J.JMB.2004.09.086

Liu, J. H., Xi, K., Zhang, X., Bao, L., Zhang, X., & Tan, Z. J. (2019). Structural Flexibility of DNA-RNA Hybrid Duplex: Stretching and Twist-Stretch Coupling. *Biophysical Journal*, *117*(1), 74–86. https://doi.org/10.1016/J.BPJ.2019.05.018

Liu, Q., Liu, Y., Shi, Q., Su, H., Wang, C., Birchler, J. A., & Han, F. (2021). Emerging roles of centromeric RNAs in centromere formation and function. *Genes and Genomics*, *43*(3), 217–226. https://doi.org/10.1007/S13258-021-01041-Y/FIGURES/1

Liu, W., Krishnamoorthy, A., Zhao, R., & Cortez, D. (2020). Two replication fork remodeling pathways generate nuclease substrates for distinct fork protection factors. *Science Advances*, *6*(46), 3598–3611. https://doi.org/10.1126/SCIADV.ABC3598/SUPPL_FILE/ABC3598_SM.PDF

Liu, Y., Lin, Y. L., Pasero, P., & Chen, C. L. (2020). Topoisomerase I prevents transcription-replication conflicts at transcription termination sites. *Molecular & Cellular Oncology*, *8*(1), 1843951. https://doi.org/10.1080/23723556.2020.1843951

Liu, Y., Lin, Y. L., Pasero, P., & Chen, C. L. (2021). Topoisomerase I prevents transcription-replication conflicts at transcription termination sites. *Molecular & Cellular Oncology*, *8*(1). https://doi.org/10.1080/23723556.2020.1843951

Lorat, Y., Schanz, S., Schuler, N., Wennemuth, G., Rübe, C., & Rübe, C. E. (2012). Beyond repair foci: DNA double-strand break repair in euchromatic and heterochromatic compartments analyzed by transmission electron microscopy. *PloS One*, *7*(5). https://doi.org/10.1371/JOURNAL.PONE.0038165

Lorch, Y., & Kornberg, R. D. (2017). Chromatin-remodeling for transcription. *Quarterly Reviews of Biophysics*, *50*, e5. https://doi.org/10.1017/S003358351700004X

Louder, R. K., He, Y., López-Blanco, J. R., Fang, J., Chacón, P., & Nogales, E. (2016). Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. *Nature 2016 531:7596*, *531*(7596), 604–609. https://doi.org/10.1038/nature17394

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12). https://doi.org/10.1186/S13059-014-0550-8

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature 1997 389:6648*, *389*(6648), 251–260. https://doi.org/10.1038/38444

Luo, W., Johnson, A. W., & Bentley, D. L. (2006). The role of Rat1 in coupling mRNA 3′-end processing to transcription termination: implications for a unified allosteric–torpedo model. *Genes & Development*, *20*(8), 954. https://doi.org/10.1101/GAD.1409106

Macheret, M., & Halazonetis, T. D. (2018). Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. *Nature 2018 555:7694*, *555*(7694), 112–116. https://doi.org/10.1038/nature25507

Madireddy, A., Kosiyatrakul, S. T., Boisvert, R. A., Herrera-Moyano, E., García-Rubio, M. L., Gerhardt, J., Vuono, E. A., Owen, N., Yan, Z., Olson, S., Aguilera, A., Howlett, N. G., & Schildkraut, C. L. (2016). FANCD2 facilitates replication through common fragile sites. *Molecular Cell*, *64*(2), 388. https://doi.org/10.1016/J.MOLCEL.2016.09.017

Maeshima, K., Ide, S., & Babokhov, M. (2019). Dynamic chromatin organization without the 30-nm fiber. *Current Opinion in Cell Biology*, *58*, 95–104. https://doi.org/10.1016/J.CEB.2019.02.003

Mailand, N., Falck, J., Lukas, C., Syljuåsen, R. G., Welcker, M., Bartek, J., & Lukas, J. (2000). Rapid destruction of human Cdc25A in response to DNA damage. *Science*, *288*(5470), 1425–1429. https://doi.org/10.1126/SCIENCE.288.5470.1425/ASSET/B613CAD8-3365-492C-9E70-6D7533A280AE/ASSETS/GRAPHIC/SE1908536004.JPEG

Mailand, N., Podtelejnikov, A. V., Groth, A., Mann, M., Bartek, J., & Lukas, J. (2002). Regulation of G(2)/M events by Cdc25A through phosphorylation-dependent modulation of its stability. *The EMBO Journal*, *21*(21), 5911–5920. https://doi.org/10.1093/EMBOJ/CDF567

Maldonado, R., Schwartz, U., Silberhorn, E., & Längst, G. (2019). Nucleosomes Stabilize ssRNA-dsDNA Triple Helices in Human Cells. *Molecular Cell*, *73*(6), 1243-1254.e6. https://doi.org/10.1016/J.MOLCEL.2019.01.007

Marchal, C., Sima, J., & Gilbert, D. M. (2019). Control of DNA replication timing in the 3D genome. *Nature Reviews Molecular Cell Biology 2019 20:12*, *20*(12), 721–737. https://doi.org/10.1038/s41580-019-0162-y

Marchena-Cruz, E., Camino, L. P., Bhandari, J., Silva, S., Marqueta-Gracia, J. J., Amdeen, S. A., Guillén-Mendoza, C., García-Rubio, M. L., Calderón-Montaño, J. M., Xue, X., Luna, R., & Aguilera, A. (2023). DDX47, MeCP2, and other functionally heterogeneous factors protect cells from harmful R loops. *Cell Reports*, *42*(3). https://doi.org/10.1016/J.CELREP.2023.112148

Marnef, A., & Legube, G. (2021). R-loops as Janus-faced modulators of DNA repair. *Nature Cell Biology 2021 23:4*, *23*(4), 305–313. https://doi.org/10.1038/s41556-021-00663-4

Marsolier-Kergoat, M. C., & Goldar, A. (2012). DNA Replication Induces Compositional Biases in Yeast. *Molecular Biology and Evolution*, *29*(3), 893–904. https://doi.org/10.1093/MOLBEV/MSR240

Mátés, L., Chuah, M. K. L., Belay, E., Jerchow, B., Manoj, N., Acosta-Sanchez, A., Grzela, D. P., Schmitt, A., Becker, K., Matrai, J., Ma, L., Samara-Kuko, E., Gysemans, C., Pryputniewicz, D., Miskey, C., Fletcher, B., Vandendriessche, T., Ivics, Z., & Izsvák, Z. (2009). Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nature Genetics*, *41*(6), 753–761. https://doi.org/10.1038/NG.343

Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z., Solimini, N., Lerenthal, Y., Shiloh, Y., Gygi, S. P., & Elledge, S. J. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*, *316*(5828), 1160–1166. https://doi.org/10.1126/SCIENCE.1140321/SUPPL_FILE/MATSUOKA.SOM.PDF

Mersaoui, S. Y., Yu, Z., Coulombe, Y., Karam, M., Busatto, F. F., Masson, J., & Richard, S. (2019). Arginine methylation of the DDX 5 helicase RGG / RG motif by PRMT 5 regulates resolution of RNA:DNA hybrids . *The EMBO Journal*, *38*(15). https://doi.org/10.15252/EMBJ.2018100986/SUPPL_FILE/EMBJ2018100986-SUP-0004-SDATAEV.PDF

Mertz, T. M., Harcy, V., & Roberts, S. A. (2017). Risks at the DNA Replication Fork: Effects upon Carcinogenesis and Tumor Heterogeneity. *Genes*, *8*(1), 46.

https://doi.org/10.3390/GENES8010046

Millán-Zambrano, G., Burton, A., Bannister, A. J., & Schneider, R. (2022). Histone post-translational modifications — cause and consequence of genome function. *Nature Reviews Genetics 2022 23:9*, *23*(9), 563–580. https://doi.org/10.1038/s41576-022-00468-7

Mirabella, A. C., Foster, B. M., & Bartke, T. (2015). Chromatin deregulation in disease. *Chromosoma 2015 125:1*, *125*(1), 75–93. https://doi.org/10.1007/S00412-015-0530-0

Miskey, C., Kesselring, L., Querques, I., Abrusán, G., Barabas, O., & Ivics, Z. (2022). Engineered Sleeping Beauty transposase redirects transposon integration away from genes. *Nucleic Acids Research*, *50*(5), 2807. https://doi.org/10.1093/NAR/GKAC092

Misteli, T. (2020). The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell*, *183*(1), 28–45. https://doi.org/10.1016/J.CELL.2020.09.014

Moiseeva, T. N., Yin, Y., Calderon, M. J., Qian, C., Schamus-Haynes, S., Sugitani, N., Osmanbeyoglu, H. U., Rothenberg, E., Watkins, S. C., & Bakkenist, C. J. (2019). An ATR and CHK1 kinase signaling mechanism that limits origin firing during unperturbed DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(27), 13374–13383. https://doi.org/10.1073/PNAS.1903418116/SUPPL_FILE/PNAS.1903418116.SAPP.PDF

Moriyama, K., Yoshizawa-Sugata, N., Obuse, C., Tsurimoto, T., & Masai, H. (2012). Epstein-Barr nuclear antigen 1 (EBNA1)-dependent recruitment of origin recognition complex (Orc) on oriP of Epstein-Barr virus with purified proteins: Stimulation by Cdc6 through its direct interaction with EBNA. *Journal of Biological Chemistry*, *287*(28), 23977–23994. https://doi.org/10.1074/JBC.M112.368456/ATTACHMENT/BFD3355A-986E-44DF-AFEB-2D863E9477A5/MMC1.PDF

Mosler, T., Conte, F., Longo, G. M. C., Mikicic, I., Kreim, N., Möckel, M. M., Petrosino, G., Flach, J., Barau, J., Luke, B., Roukos, V., & Beli, P. (2021). R-loop proximity proteomics identifies a role of DDX41 in transcription-associated genomic instability. *Nature Communications 2021 12:1*, *12*(1), 1–17. https://doi.org/10.1038/s41467-021-27530-y

Mourón, S., Rodriguez-Acebes, S., Martínez-Jiménez, M. I., García-Gómez, S., Chocrón, S., Blanco, L., & Méndez, J. (2013). Repriming of DNA synthesis at stalled replication forks by human PrimPol. *Nature Structural & Molecular Biology 2013 20:12*, *20*(12), 1383–1389. https://doi.org/10.1038/nsmb.2719

Müller, F., & Tora, L. (2014). Chromatin and DNA sequences in defining promoters for transcription initiation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1839*(3), 118–128. https://doi.org/10.1016/J.BBAGRM.2013.11.003

Mutreja, K., Krietsch, J., Hess, J., Ursich, S., Berti, M., Roessler, F. K., Zellweger, R., Patra, M., Gasser, G., & Lopes, M. (2018). ATR-Mediated Global Fork Slowing and Reversal Assist Fork Traverse and Prevent Chromosomal Breakage at DNA Interstrand Cross-Links. *Cell Reports*, *24*(10), 2629-2642.e5. https://doi.org/10.1016/J.CELREP.2018.08.019/ATTACHMENT/632D34D9-81E5-4445-8C43-CAB54E99B475/MMC2.PDF

Neelsen, K. J., Zanini, I. M. Y., Herrador, R., & Lopes, M. (2013). Oncogenes induce genotoxic stress by mitotic processing of unusual replication intermediates. *Journal of Cell Biology*, *200*(6), 699–708. https://doi.org/10.1083/JCB.201212058

Nguyen, A. T., & Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. *Genes & Development*, *25*(13), 1345–1358. https://doi.org/10.1101/GAD.2057811

Nick McElhinny, S. A., Gordenin, D. A., Stith, C. M., Burgers, P. M. J., & Kunkel, T. A. (2008). Division of Labor at the Eukaryotic Replication Fork. *Molecular Cell*, *30*(2), 137–144. https://doi.org/10.1016/j.molcel.2008.02.022

Niehrs, C., & Luke, B. (2020). Regulatory R-loops as facilitators of gene expression and genome stability. *Nature Reviews Molecular Cell Biology 2020 21:3*, *21*(3), 167–178. https://doi.org/10.1038/s41580-019-0206-3

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S. J., Diekhans, M., Logsdon, G. A., Alonge, M., Antonarakis, S. E., Borchers, M., Bouffard, G. G., Brooks, S. Y., … Phillippy, A. M. (2022). The complete sequence of a human genome. *Science*, *376*(6588), 44–53. https://doi.org/10.1126/SCIENCE.ABJ6987/SUPPL_FILE/SCIENCE.ABJ6987_MDAR_REPRODUCIBILITY_CHECKLIST.PDF

Oh, J. M., & Myung, K. (2022). Crosstalk between different DNA repair pathways for DNA double strand break repairs. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *873*, 503438. https://doi.org/10.1016/J.MRGENTOX.2021.503438

Okamoto, Y., Abe, M., Itaya, A., Tomida, J., Ishiai, M., Takaori-Kondo, A., Taoka, M., Isobe, T., & Takata, M. (2019). FANCD2 protects genome stability by recruiting RNA processing enzymes to resolve R-loops during mild replication stress. *The FEBS Journal*, *286*(1), 139–150. https://doi.org/10.1111/FEBS.14700

Okamoto, Y., Iwasaki, W. M., Kugou, K., Takahashi, K. K., Oda, A., Sato, K., Kobayashi, W., Kawai, H., Sakasai, R., Takaori-Kondo, A., Yamamoto, T., Kanemaki, M. T., Taoka, M., Isobe, T., Kurumizaka, H., Innan, H., Ohta, K., Ishiai, M., & Takata, M. (2018). Replication stress induces accumulation of FANCD2 at central region of large fragile genes. *Nucleic Acids Research*, *46*(6), 2932–2944. https://doi.org/10.1093/NAR/GKY058

Papadopoulos, D., Solvie, D., Baluapuri, A., Endres, T., Ha, S. A., Herold, S., Kalb, J., Giansanti, C., Schülein-Völk, C., Ade, C. P., Schneider, C., Gaballa, A., Vos, S., Fischer, U., Dobbelstein, M., Wolf, E., & Eilers, M. (2022). MYCN recruits the nuclear exosome complex to RNA polymerase II to prevent transcription-replication conflicts. *Molecular Cell*, *82*(1), 159-176.e12. https://doi.org/10.1016/J.MOLCEL.2021.11.002

Parua, P. K., Booth, G. T., Sansó, M., Benjamin, B., Tanny, J. C., Lis, J. T., & Fisher, R. P. (2018). A Cdk9–PP1 switch regulates the elongation–termination transition of RNA polymerase II. *Nature 2018 558:7710*, *558*(7710), 460–464. https://doi.org/10.1038/s41586-018-0214-z

Patel, P. S., Algouneh, A., Krishnan, R., Reynolds, J. J., Nixon, K. C. J., Hao, J., Lee, J., Feng, Y., Fozil, C., Stanic, M., Yerlici, T., Su, P., Soares, F., Liedtke, E., Prive, G., Baider, G. D., Pujana, M. A., Mekhail, K., He, H. H., … Hakem, R. (2023). Excessive transcription-replication conflicts are a vulnerability of BRCA1-mutant cancers. *Nucleic Acids Research*, *51*(9), 4341–4362. https://doi.org/10.1093/NAR/GKAD172

Pérez-Calero, C., Bayona-Feliu, A., Xue, X., Barroso, S. I., Muñoz, S., González-Basallote, V. M., Sung, P., & Aguilera, A. (2020). UAP56/DDX39B is a major cotranscriptional RNA–DNA helicase that unwinds harmful R loops genome-wide. *Genes & Development*, *34*(13–

14), 898–912. https://doi.org/10.1101/GAD.336024.119

Petermann, E., Lan, L., & Zou, L. (2022). Sources, resolution and physiological relevance of R-loops and RNA–DNA hybrids. *Nature Reviews Molecular Cell Biology 2022 23:8*, *23*(8), 521–540. https://doi.org/10.1038/s41580-022-00474-x

Petropoulos, M., Karamichali, A., Rossetti, G. G., Freudenmann, A., Iacovino, L. G., Dionellis, V. S., Sotiriou, S. K., & Halazonetis, T. D. (2024). Transcription–replication conflicts underlie sensitivity to PARP inhibitors. *Nature 2024 628:8007*, *628*(8007), 433–441. https://doi.org/10.1038/s41586-024-07217-2

Petryk, N., Kahli, M., D'Aubenton-Carafa, Y., Jaszczyszyn, Y., Shen, Y., Silvain, M., Thermes, C., Chen, C. L., & Hyrien, O. (2016). Replication landscape of the human genome. *Nature Communications 2016 7:1*, *7*(1), 1–13. https://doi.org/10.1038/ncomms10208

Polenkowski, M., Allister, A. B., Burbano de Lara, S., Pierce, A., Geary, B., El Bounkari, O., Wiehlmann, L., Hoffmann, A., Whetton, A. D., Tamura, T., & Tran, D. D. H. (2023). THOC5 complexes with DDX5, DDX17, and CDK12 to regulate R loop structures and transcription elongation rate. *IScience*, *26*(1). https://doi.org/10.1016/j.isci.2022.105784

Pope, B. D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D. L., Wang, Y., Hansen, R. S., Canfield, T. K., Thurman, R. E., Cheng, Y., Gülsoy, G., Dennis, J. H., Snyder, M. P., Stamatoyannopoulos, J. A., Taylor, J., Hardison, R. C., Kahveci, T., … Gilbert, D. M. (2014). Topologically associating domains are stable units of replication-timing regulation. *Nature 2014 515:7527*, *515*(7527), 402–405. https://doi.org/10.1038/nature13986

Prado, F., & Aguilera, A. (2005). Impairment of replication fork progression mediates RNA polII transcription-associated recombination. *EMBO Journal*, *24*(6), 1267–1276. https://doi.org/10.1038/SJ.EMBOJ.7600602/ASSET/874E2597-2BF2-451E-B5C4-7D2779FAA18E/ASSETS/GRAPHIC/EMBJ7600602-FIG-0007-M.JPG

Promonet, A., Padioleau, I., Liu, Y., Sanz, L., Biernacka, A., Schmitz, A. L., Skrzypczak, M., Sarrazin, A., Mettling, C., Rowicka, M., Ginalski, K., Chedin, F., Chen, C. L., Lin, Y. L., & Pasero, P. (2020). Topoisomerase 1 prevents replication stress at R-loop-enriched transcription termination sites. *Nature Communications 2020 11:1*, *11*(1), 1–12. https://doi.org/10.1038/s41467-020-17858-2

Proudfoot, N. J. (2016). Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science*, *352*(6291). https://doi.org/10.1126/SCIENCE.AAD9926/ASSET/4BAC16B8-0F78-40E3-AE87-7EA560D74D36/ASSETS/GRAPHIC/352_AAD9926_FA.JPEG

Quénet, D., McNally, J. G., & Dalal, Y. (2012). Through thick and thin: The conundrum of chromatin fibre folding in vivo. *EMBO Reports*, *13*(11), 943–944. https://doi.org/10.1038/EMBOR.2012.143/ASSET/DA698005-C154-48B2-B6D6-FB7DFA81C76D/ASSETS/GRAPHIC/EMBR2012143-FIG-0001-M.JPG

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, *26*(6), 841–842. https://doi.org/10.1093/BIOINFORMATICS/BTQ033

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., & Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, *44*(W1), W160–W165. https://doi.org/10.1093/NAR/GKW257

Recolin, B., van der Laan, S., Tsanov, N., & Maiorano, D. (2014). Molecular Mechanisms of DNA Replication Checkpoint Activation. *Genes*, *5*(1), 147. https://doi.org/10.3390/GENES5010147

Reijns, M. A. M., Rabe, B., Rigby, R. E., Mill, P., Astell, K. R., Lettice, L. A., Boyle, S., Leitch, A., Keighren, M., Kilanowski, F., Devenney, P. S., Sexton, D., Grimes, G., Holt, I. J., Hill, R. E., Taylor, M. S., Lawson, K. A., Dorin, J. R., & Jackson, A. P. (2012). Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell*, *149*(5), 1008–1022. https://doi.org/10.1016/J.CELL.2012.04.011/ATTACHMENT/1345A4AC-F9CA-48D3-81C6-A69606499A59/MMC2.PDF

Remus, D., Beuron, F., Tolun, G., Griffith, J. D., Morris, E. P., & Diffley, J. F. X. (2009). Concerted Loading of Mcm2-7 Double Hexamers around DNA during DNA Replication Origin Licensing. *Cell*, *139*(4), 719–730. https://doi.org/10.1016/J.CELL.2009.10.015/ATTACHMENT/676151E2-0F10-493D-89EB-2F3950886C31/MMC1.PDF

Ren, X., Liu, Q., Zhou, P., Zhou, T., Wang, D., Mei, Q., Flavell, R. A., Liu, Z., Li, M., Pan, W., & Zhu, S. (2024). DHX9 maintains epithelial homeostasis by restraining R-loop-mediated genomic instability in intestinal stem cells. *Nature Communications*, *15*(1). https://doi.org/10.1038/S41467-024-47235-2

Reverón-Gómez, N., González-Aguilera, C., Stewart-Morgan, K. R., Petryk, N., Flury, V., Graziano, S., Johansen, J. V., Jakobsen, J. S., Alabert, C., & Groth, A. (2018). Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication. *Molecular Cell*, *72*(2), 239-249.e5. https://doi.org/10.1016/J.MOLCEL.2018.08.010

Rhind, N., & Gilbert, D. M. (2013). DNA Replication Timing. *Cold Spring Harbor Perspectives in Biology*, *5*(8), a010132. https://doi.org/10.1101/CSHPERSPECT.A010132

Robinson, Philip J., Trnka, M. J., Bushnell, D. A., Davis, R. E., Mattei, P. J., Burlingame, A. L., & Kornberg, R. D. (2016). Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. *Cell*, *166*(6), 1411-1422.e16. https://doi.org/10.1016/J.CELL.2016.08.050/ATTACHMENT/DF69BB3F-9727-4649-B71B-0EE7F08C8D57/MMC2.XLSX

Robinson, Philip JJ, & Rhodes, D. (2006). Structure of the '30 nm' chromatin fibre: A key role for the linker histone. *Current Opinion in Structural Biology*, *16*(3), 336–343. https://doi.org/10.1016/J.SBI.2006.05.007

Roeder, R. G., & Rutter, W. J. (1969). Multiple Forms of DNA-dependent RNA Polymerase in Eukaryotic Organisms. *Nature 1969 224:5216*, *224*(5216), 234–237. https://doi.org/10.1038/224234a0

Roske, J. J., & Yeeles, J. T. P. (2024). Structural basis for processive daughter-strand synthesis and proofreading by the human leading-strand DNA polymerase Pol ε. *Nature Structural & Molecular Biology 2024 31:12*, *31*(12), 1921–1931. https://doi.org/10.1038/s41594-024-01370-y

Rougvie, A. E., & Lis, J. T. (1988). The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. *Cell*, *54*(6), 795–804. https://doi.org/10.1016/S0092-8674(88)91087-2

Saksouk, N., Simboeck, E., & Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. *Epigenetics and Chromatin*, *8*(1), 1–17. https://doi.org/10.1186/1756-8935-8-3/FIGURES/4

Saldivar, J. C., Hamperl, S., Bocek, M. J., Chung, M., Bass, T. E., Cisneros-Soberanis, F., Samejima, K., Xie, L., Paulson, J. R., Earnshaw, W. C., Cortez, D., Meyer, T., & Cimprich, K. A. (2018). An intrinsic S/G2 checkpoint enforced by ATR. *Science*, *361*(6404), 806–810. https://doi.org/10.1126/SCIENCE.AAP9346/SUPPL_FILE/AAP9346_SALDIVAR_SM.PDF

Salguero, C., Valladolid, C., Robinson, H. M. R., Smith, G. C. M., & Yap, T. A. (2023). Targeting ATR in Cancer Medicine. *Cancer Treatment and Research*, *186*, 239–283. https://doi.org/10.1007/978-3-031-30065-3_14

Sankar, T. S., Wastuwidyaningtyas, B. D., Dong, Y., Lewis, S. A., & Wang, J. D. (2016). The nature of mutations induced by replication–transcription collisions. *Nature 2016 535:7610*, *535*(7610), 178–181. https://doi.org/10.1038/nature18316

Santoni-Rugiu, E., Falck, J., Mailand, N., Bartek, J., & Lukas, J. (2000). Involvement of Myc Activity in a G1/S-Promoting Mechanism Parallel to the pRb/E2F Pathway. *Molecular and Cellular Biology*, *20*(10), 3497–3509. https://doi.org/10.1128/MCB.20.10.3497-3509.2000

Saxena, S., & Zou, L. (2022). Hallmarks of DNA replication stress. *Molecular Cell*, *82*(12), 2298–2314. https://doi.org/10.1016/J.MOLCEL.2022.05.004/ASSET/C1D1C4C2-EC29-4D0A-988D-8B90695E040E/MAIN.ASSETS/GR2_LRG.JPG

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/NMETH.2019

Schmidt, U., Weigert, M., Broaddus, C., & Myers, G. (2018). Cell detection with star-convex polygons. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11071 LNCS*, 265–273. https://doi.org/10.1007/978-3-030-00934-2_30

Schneider, M. W. G., Gibson, B. A., Otsuka, S., Spicer, M. F. D., Petrovic, M., Blaukopf, C., Langer, C. C. H., Batty, P., Nagaraju, T., Doolittle, L. K., Rosen, M. K., & Gerlich, D. W. (2022). A mitotic chromatin phase transition prevents perforation by microtubules. *Nature 2022 609:7925*, *609*(7925), 183–190. https://doi.org/10.1038/s41586-022-05027-y

Schübeler, D. (2015). Function and information content of DNA methylation. *Nature 2015 517:7534*, *517*(7534), 321–326. https://doi.org/10.1038/nature14192

Scott, S. J., Suvarna, K. S., & D'Avino, P. P. (2020). Synchronization of human retinal pigment epithelial-1 cells in mitosis. *Journal of Cell Science*, *133*(18). https://doi.org/10.1242/JCS.247940/266615/AM/SYNCHRONIZATION-OF-HUMAN-RETINAL-PIGMENT

Sentenac, A. (1985). Eukaryotic RNA polymerase. *Critical Reviews in Biochemistry and Molecular Biology*, *18*(1), 31–90. https://doi.org/10.3109/10409238509082539/ASSET//CMS/ASSET/EABCCA17-21BE-4D5F-B692-291136F001E1/10409238509082539.FP.PNG

Sequeira-Mendes, J., Díaz-Uriarte, R., Apedaile, A., Huntley, D., Brockdorff, N., & Gómez, M.

(2009). Transcription Initiation Activity Sets Replication Origin Efficiency in Mammalian Cells. *PLOS Genetics*, *5*(4), e1000446. https://doi.org/10.1371/JOURNAL.PGEN.1000446

Shao, X., Joergensen, A. M., Howlett, N. G., Lisby, M., & Oestergaard, V. H. (2020). A distinct role for recombination repair factors in an early cellular response to transcription–replication conflicts. *Nucleic Acids Research*, *48*(10), 5467–5484. https://doi.org/10.1093/NAR/GKAA268

Silva, B., Pentz, R., Figueira, A. M., Arora, R., Lee, Y. W., Hodson, C., Wischnewski, H., Deans, A. J., & Azzalin, C. M. (2019). FANCM limits ALT activity by restricting telomeric replication stress induced by deregulated BLM and R-loops. *Nature Communications 2019 10:1*, *10*(1), 1–16. https://doi.org/10.1038/s41467-019-10179-z

Sima, J., Chakraborty, A., Dileep, V., Michalski, M., Klein, K. N., Holcomb, N. P., Turner, J. L., Paulsen, M. T., Rivera-Mulia, J. C., Trevilla-Garcia, C., Bartlett, D. A., Zhao, P. A., Washburn, B. K., Nora, E. P., Kraft, K., Mundlos, S., Bruneau, B. G., Ljungman, M., Fraser, P., … Gilbert, D. M. (2019). Identifying cis Elements for Spatiotemporal Control of Mammalian DNA Replication. *Cell*, *176*(4), 816-830.e18. https://doi.org/10.1016/J.CELL.2018.11.036/ASSET/6464C6FD-4F95-4938-B6FF-A94DF9AF898E/MAIN.ASSETS/GR7.JPG

Singh, N., Asalam, M., Ansari, M. O., Gerasimova, N. S., Studitsky, V. M., & Akhtar, M. S. (2022). Transcription by RNA polymerase II and the CTD-chromatin crosstalk. *Biochemical and Biophysical Research Communications*, *599*, 81–86. https://doi.org/10.1016/J.BBRC.2022.02.039

Skourti-Stathaki, K., Kamieniarz-Gdula, K., & Proudfoot, N. J. (2014). R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature 2014 516:7531*, *516*(7531), 436–439. https://doi.org/10.1038/nature13787

Skourti-Stathaki, K., Proudfoot, N. J., & Gromak, N. (2011). Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Molecular Cell*, *42*(6), 794–805. https://doi.org/10.1016/J.MOLCEL.2011.04.026

So, C. C., Ramachandran, S., & Martin, A. (2019). E3 Ubiquitin Ligases RNF20 and RNF40 Are Required for Double-Stranded Break (DSB) Repair: Evidence for Monoubiquitination of Histone H2B Lysine 120 as a Novel Axis of DSB Signaling and Repair. *Molecular and Cellular Biology*, *39*(8), e00488-18. https://doi.org/10.1128/MCB.00488-18

Sollier, J., Stork, C. T., García-Rubio, M. L., Paulsen, R. D., Aguilera, A., & Cimprich, K. A. (2014). Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability. *Molecular Cell*, *56*(6), 777–785. https://doi.org/10.1016/J.MOLCEL.2014.10.020

Song, C., Hotz-Wagenblatt, A., Voit, R., & Grummt, I. (2017). SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes & Development*, *31*(13), 1370–1381. https://doi.org/10.1101/GAD.300624.117

Sparks, J. L., Chon, H., Cerritelli, S. M., Kunkel, T. A., Johansson, E., Crouch, R. J., & Burgers, P. M. (2012). RNase H2-Initiated Ribonucleotide Excision Repair. *Molecular Cell*, *47*(6), 980. https://doi.org/10.1016/J.MOLCEL.2012.06.035

Sridhara, S. C., Carvalho, S., Grosso, A. R., Gallego-Paez, L. M., Carmo-Fonseca, M., & de Almeida, S. F. (2017). Transcription Dynamics Prevent RNA-Mediated Genomic Instability

through SRPK2-Dependent DDX23 Phosphorylation. *Cell Reports*, *18*(2), 334–343. https://doi.org/10.1016/J.CELREP.2016.12.050

St Germain, C. P., Zhao, H., Sinha, V., Sanz, L. A., Chedin, F., & Barlow, J. H. (2022). Genomic patterns of transcription–replication interactions in mouse primary B cells. *Nucleic Acids Research*, *50*(4), 2051–2073. https://doi.org/10.1093/NAR/GKAC035

Steger, D. J., Lefterova, M. I., Ying, L., Stonestrom, A. J., Schupp, M., Zhuo, D., Vakoc, A. L., Kim, J.-E., Chen, J., Lazar, M. A., Blobel, G. A., & Vakoc, C. R. (2008). DOT1L/KMT4 Recruitment and H3K79 Methylation Are Ubiquitously Coupled with Gene Transcription in Mammalian Cells. *Molecular and Cellular Biology*, *28*(8), 2825–2839. https://doi.org/10.1128/MCB.02076-07

Stork, C. T., Bocek, M., Crossley, M. P., Sollier, J., Sanz, L. A., Chédin, F., Swigut, T., & Cimprich, K. A. (2016). Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. *ELife*, *5*(AUGUST). https://doi.org/10.7554/ELIFE.17548

Stoy, H., Zwicky, K., Kuster, D., Lang, K. S., Krietsch, J., Crossley, M. P., Schmid, J. A., Cimprich, K. A., Merrikh, H., & Lopes, M. (2023). Direct visualization of transcription-replication conflicts reveals post-replicative DNA:RNA hybrids. *Nature Structural & Molecular Biology 2023 30:3*, *30*(3), 348–359. https://doi.org/10.1038/s41594-023-00928-6

Straka, J., Khatib, J. B., Pale, L., Nicolae, C. M., & Moldovan, G. L. (2024). CAF-1 promotes efficient PrimPol recruitment to nascent DNA for single-stranded DNA gap formation. *Nucleic Acids Research*, *52*(22), 13865–13880. https://doi.org/10.1093/NAR/GKAE1068

Strobl, L. J., & Eick, D. (1992). Hold back of RNA polymerase II at the transcription start site mediates down-regulation of c-myc in vivo. *The EMBO Journal*, *11*(9), 3307–3314. https://doi.org/10.1002/J.1460-2075.1992.TB05409.X

Sun, H., Ma, L., Tsai, Y. F., Abeywardana, T., Shen, B., & Zheng, L. (2023). Okazaki fragment maturation: DNA flap dynamics for cell proliferation and survival. *Trends in Cell Biology*, *33*(3), 221–234. https://doi.org/10.1016/J.TCB.2022.06.014/ASSET/EA4CADB5-D128-4225-996A-3FA10A226B56/MAIN.ASSETS/GR6.JPG

Sun, Y., Duthaler, S., & Nelson, B. J. (2004). Autofocusing in computer microscopy: selecting the optimal focus algorithm. *Microscopy Research and Technique*, *65*(3), 139–149. https://doi.org/10.1002/JEMT.20118

Šviković, S., Crisp, A., Tan-Wong, S. M., Guilliam, T. A., Doherty, A. J., Proudfoot, N. J., Guilbaud, G., & Sale, J. E. (2019). R-loop formation during S phase is restricted by PrimPol-mediated repriming. *The EMBO Journal*, *38*(3). https://doi.org/10.15252/EMBJ.201899793/SUPPL_FILE/EMBJ201899793-SUP-0002-EVFIGS.PDF

Talwar, T., Vidhyasagar, V., Qing, J., Guo, M., Kariem, A., Lu, Y., Singh, R. S., Lukong, K. E., Wu, Y., & Sung, P. (2017). The DEAD-box protein DDX43 (HAGE) is a dual RNA-DNA helicase and has a K-homology domain required for full nucleic acid unwinding activity. *The Journal of Biological Chemistry*, *292*(25), 10429. https://doi.org/10.1074/JBC.M117.774950

Tan, J., & Lan, L. (2020). The DNA secondary structures at telomeres and genome instability. *Cell and Bioscience*, *10*(1), 1–6. https://doi.org/10.1186/S13578-020-00409-Z/FIGURES/1

Tellier, M., Zaborowska, J., Caizzi, L., Mohammad, E., Velychko, T., Schwalb, B., Ferrer-Vicens, I., Blears, D., Nojima, T., Cramer, P., & Murphy, S. (2020). CDK12 globally stimulates RNA

polymerase II transcription elongation and carboxyl-terminal domain phosphorylation. *Nucleic Acids Research*, *48*(14), 7712–7727. https://doi.org/10.1093/NAR/GKAA514

Thoma, F., Koller, T., & Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *Journal of Cell Biology*, *83*(2), 403–427. https://doi.org/10.1083/JCB.83.2.403

Toledo, L. I., Altmeyer, M., Rask, M. B., Lukas, C., Larsen, D. H., Povlsen, L. K., Bekker-Jensen, S., Mailand, N., Bartek, J., & Lukas, J. (2013). ATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell*, *155*(5), 1088. https://doi.org/10.1016/J.CELL.2013.10.043/ATTACHMENT/E4B5CF34-9000-4356-80A1-141CA824FD68/MMC8.PDF

Tous, C., & Aguilera, A. (2007). Impairment of transcription elongation by R-loops in vitro. *Biochemical and Biophysical Research Communications*, *360*(2), 428–432. https://doi.org/10.1016/J.BBRC.2007.06.098

Tsirkas, I., Dovrat, D., Thangaraj, M., Brouwer, I., Cohen, A., Paleiov, Z., Meijler, M. M., Lenstra, T., & Aharoni, A. (2022). Transcription-replication coordination revealed in single live cells. *Nucleic Acids Research*, *50*(4), 2143–2156. https://doi.org/10.1093/NAR/GKAC069

Utley, R. T., Ikeda, K., Grant, P. A., Côté, J., Steger, D. J., Eberharter, A., John, S., & Workman, J. L. (1998). Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature 1998 394:6692*, *394*(6692), 498–502. https://doi.org/10.1038/28886

Valencia-Sánchez, M. I., De Ioannes, P., Wang, M., Vasilyev, N., Chen, R., Nudler, E., Armache, J. P., & Armache, K. J. (2019). Structural basis of Dot1L stimulation by histone H2B lysine 120 ubiquitination. *Molecular Cell*, *74*(5), 1010. https://doi.org/10.1016/J.MOLCEL.2019.03.029

Veloso, A., Kirkconnell, K. S., Magnuson, B., Biewen, B., Paulsen, M. T., Wilson, T. E., & Ljungman, M. (2014). Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Research*, *24*(6), 896–905. https://doi.org/10.1101/GR.171405.113

Velychko, T., Mohammad, E., Ferrer-Vicens, I., Parfentev, I., Werner, M., Studniarek, C., Schwalb, B., Urlaub, H., Murphy, S., Cramer, P., & Lidschreiber, M. (2024). CDK7 kinase activity promotes RNA polymerase II promoter escape by facilitating initiation factor release. *Molecular Cell*, *84*(12), 2287-2303.e10. https://doi.org/10.1016/J.MOLCEL.2024.05.007

Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W. M. P., van Schaik, F. M. A., Varier, R. A., Baltissen, M. P. A., Stunnenberg, H. G., Mann, M., & Timmers, H. T. M. (2007). Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4. *Cell*, *131*(1), 58–69. https://doi.org/10.1016/J.CELL.2007.08.016/ATTACHMENT/5841369F-07BD-4B5F-B930-174F625D70ED/MMC2.XLS

Vos, S. M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., & Cramer, P. (2018). Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature 2018 560:7720*, *560*(7720), 607–612. https://doi.org/10.1038/s41586-018-0440-4

Vos, S. M., Farnung, L., Urlaub, H., & Cramer, P. (2018). Structure of paused transcription complex Pol II–DSIF–NELF. *Nature 2018 560:7720*, *560*(7720), 601–606. https://doi.org/10.1038/s41586-018-0442-2

Wang, E., Kawaoka, S., Yu, M., Shi, J., Ni, T., Yang, W., Zhu, J., Roeder, R. G., & Vakoc, C. R. (2013). Histone H2B ubiquitin ligase RNF20 is required for MLL-rearranged leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(10), 3901–3906. https://doi.org/10.1073/PNAS.1301045110/-/DCSUPPLEMENTAL

Wang, J. D., Berkmen, M. B., & Grossman, A. D. (2007). Genome-wide coorientation of replication and transcription reduces adverse effects on replication in Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(13), 5608–5613. https://doi.org/10.1073/PNAS.0608999104/ASSET/64E243FE-5373-44C2-8CEE-BBA773C7852B/ASSETS/GRAPHIC/ZPQ0100756050005.JPEG

Wang, J., Rojas, P., Mao, J., Mustè Sadurnì, M., Garnier, O., Xiao, S., Higgs, M. R., Garcia, P., & Saponaro, M. (2021). Persistence of RNA transcription during DNA replication delays duplication of transcription start sites until G2/M. *Cell Reports*, *34*(7), 108759. https://doi.org/10.1016/J.CELREP.2021.108759

Wei, P. C., Chang, A. N., Kao, J., Du, Z., Meyers, R. M., Alt, F. W., & Schwer, B. (2016). Long Neural Genes Harbor Recurrent DNA Break Clusters in Neural Stem/Progenitor Cells. *Cell*, *164*(4), 644–655. https://doi.org/10.1016/J.CELL.2015.12.039

Weigert, M., Schmidt, U., Haase, R., Sugawara, K., & Myers, G. (2020). Star-convex polyhedra for 3D object detection and segmentation in microscopy. *Proceedings - 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020*, 3655–3662. https://doi.org/10.1109/WACV45572.2020.9093435

Weissmann, F., Greiwe, J. F., Pühringer, T., Eastwood, E. L., Couves, E. C., Miller, T. C. R., Diffley, J. F. X., & Costa, A. (2024). MCM double hexamer loading visualized with human proteins. *Nature 2024 636:8042*, *636*(8042), 499–508. https://doi.org/10.1038/s41586-024-08263-6

Werner, M., Trauner, M., Schauer, T., Ummethum, H., Márquez-Gómez, E., Lalonde, M., Lee, C. S., Tsirkas, I., Sajid, A., Mur iello, A. C., not Längst, G., & Hamperl, S. (2025). Transcription-replication conflicts drive R-loop-dependent nucleosome eviction and require DOT1L activity for transcription recovery. *Nucleic Acids Research*, *53*(4), 13–14. https://doi.org/10.1093/NAR/GKAF109

Wilson, M. D., Harreman, M., & Svejstrup, J. Q. (2013). Ubiquitylation and degradation of elongating RNA polymerase II: the last resort. *Biochimica et Biophysica Acta*, *1829*(1), 151–157. https://doi.org/10.1016/J.BBAGRM.2012.08.002

Wilson, M. D., Harreman, M., Taschner, M., Reid, J., Walker, J., Erdjument-Bromage, H., Tempst, P., & Svejstrup, J. Q. (2013). XProteasome-mediated processing of Def1, a critical step in the cellular response to transcription stress. *Cell*, *154*(5), 983–995. https://doi.org/10.1016/J.CELL.2013.07.028/ATTACHMENT/60D7E559-16FA-4145-95CA-39D8749DE600/MMC2.PDF

Wong, K. H., Jin, Y., & Struhl, K. (2014). TFIIH Phosphorylation of the Pol II CTD Stimulates Mediator Dissociation from the Preinitiation Complex and Promoter Escape. *Molecular Cell*, *54*(4), 601–612. https://doi.org/10.1016/J.MOLCEL.2014.03.024/ATTACHMENT/01E98ED5-A19C-4E2A-AB77-AE5696CCB0EC/MMC3.PDF

Wu, A., Zhi, J., Tian, T., Cihan, A., Cevher, M. A., Liu, Z., David, Y., Muir, T. W., Roeder, R. G., & Yu, M. (2021). DOT1L complex regulates transcriptional initiation in human

erythroleukemic cells. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(27), e2106148118. https://doi.org/10.1073/PNAS.2106148118/SUPPL_FILE/PNAS.2106148118.SAPP.PDF

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., & Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Molecular Cell*, *21*(2), 227–237. https://doi.org/10.1016/J.MOLCEL.2005.11.024/ATTACHMENT/CEECC2CD-EA2F-45A1-A313-2E8A10AB0CD7/MMC1.PDF

Yamaguchi, Y., Shibata, H., & Handa, H. (2013). Transcription elongation factors DSIF and NELF: Promoter-proximal pausing and beyond. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1829*(1), 98–104. https://doi.org/10.1016/J.BBAGRM.2012.11.007

Yang, S., Winstone, L., Mondal, S., & Wu, Y. (2023). Helicases in R-loop Formation and Resolution. *The Journal of Biological Chemistry*, *299*(11), 105307. https://doi.org/10.1016/J.JBC.2023.105307

Yates, J. L., Camiolo, S. M., & Bashaw, J. M. (2000). The Minimal Replicator of Epstein-Barr Virus oriP. *Journal of Virology*, *74*(10), 4512. https://doi.org/10.1128/JVI.74.10.4512-4522.2000

Yu, K., Chedin, F., Hsieh, C. L., Wilson, T. E., & Lieber, M. R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature Immunology 2003 4:5*, *4*(5), 442–451. https://doi.org/10.1038/ni919

Yuan, Z., Georgescu, R., Li, H., & O'Donnell, M. E. (2023). Molecular choreography of primer synthesis by the eukaryotic Pol α-primase. *Nature Communications 2023 14:1*, *14*(1), 1–14. https://doi.org/10.1038/s41467-023-39441-1

Zack, T. I., Schumacher, S. E., Carter, S. L., Cherniack, A. D., Saksena, G., Tabak, B., Lawrence, M. S., Zhang, C. Z., Wala, J., Mermel, C. H., Sougnez, C., Gabriel, S. B., Hernandez, B., Shen, H., Laird, P. W., Getz, G., Meyerson, M., & Beroukhim, R. (2013). Pan-cancer patterns of somatic copy number alteration. *Nature Genetics 2013 45:10*, *45*(10), 1134–1140. https://doi.org/10.1038/ng.2760

Zampetidis, C. P., Galanos, P., Angelopoulou, A., Zhu, Y., Polyzou, A., Karamitros, T., Kotsinas, A., Lagopati, N., Mourkioti, I., Mirzazadeh, R., Polyzos, A., Garnerone, S., Mizi, A., Gusmao, E. G., Sofiadis, K., Gál, Z., Larsen, D. H., Pefani, D. E., Demaria, M., … Gorgoulis, V. G. (2021). A recurrent chromosomal inversion suffices for driving escape from oncogene-induced senescence via subTAD reorganization. *Molecular Cell*, *81*(23), 4907-4923.e8. https://doi.org/10.1016/J.MOLCEL.2021.10.017/ASSET/E1CB12E4-49EC-4253-9225-CDB9AC849F8F/MAIN.ASSETS/GR7.JPG

Zeman, M. K., & Cimprich, K. A. (2013). Causes and consequences of replication stress. *Nature Cell Biology 2014 16:1*, *16*(1), 2–9. https://doi.org/10.1038/ncb2897

Zhang, Hongshan, Shi, Z., Banigan, E. J., Kim, Y., Yu, H., Bai, X. chen, & Finkelstein, I. J. (2023). CTCF and R-loops are boundaries of cohesin-mediated DNA looping. *Molecular Cell*, *83*(16), 2856-2871.e8. https://doi.org/10.1016/J.MOLCEL.2023.07.006

Zhang, Huimin, Rigo, F., & Martinson, H. G. (2015). Poly(A) Signal-Dependent Transcription Termination Occurs through a Conformational Change Mechanism that Does Not Require Cleavage at the Poly(A) Site. *Molecular Cell*, *59*(3), 437–448.

https://doi.org/10.1016/J.MOLCEL.2015.06.008

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nussbaum, C., Myers, R. M., Brown, M., Li, W., & Shirley, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biology*, *9*(9). https://doi.org/10.1186/GB-2008-9-9-R137

Zhao, H., Zhu, M., Limbo, O., & Russell, P. (2018). RNase H eliminates R-loops that disrupt DNA replication but is nonessential for efficient DSB repair. *EMBO Reports*, *19*(5). https://doi.org/10.15252/EMBR.201745335/SUPPL_FILE/EMBR201745335-SUP-0003-TABLEEV2.DOCX

Zhao, P. A., Sasaki, T., & Gilbert, D. M. (2020). High-resolution Repli-Seq defines the temporal choreography of initiation, elongation and termination of replication in mammalian cells. *Genome Biology*, *21*(1), 1–20. https://doi.org/10.1186/S13059-020-01983-8/FIGURES/7

Zhou, J., Snyder, A. R., & Lieberman, P. M. (2009). Epstein-Barr virus episome stability is coupled to a delay in replication timing. *Journal of Virology*, *83*(5), 2154–2162. https://doi.org/10.1128/JVI.02115-08

Zhou, Q., Li, T., & Price, D. H. (2012). RNA polymerase II elongation control. *Annual Review of Biochemistry*, *81*(Volume 81, 2012), 119–143. https://doi.org/10.1146/ANNUREV-BIOCHEM-052610-095910/CITE/REFWORKS

Zhu, F., Farnung, L., Kaasinen, E., Sahu, B., Yin, Y., Wei, B., Dodonova, S. O., Nitta, K. R., Morgunova, E., Taipale, M., Cramer, P., & Taipale, J. (2018). The interaction landscape between transcription factors and the nucleosome. *Nature 2018 562:7725*, *562*(7725), 76–81. https://doi.org/10.1038/s41586-018-0549-5

Zou, L., & Elledge, S. J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science*, *300*(5625), 1542–1548. https://doi.org/10.1126/SCIENCE.1083430/SUPPL_FILE/ZOU.SOM.PDF

8. Abbreviations

Abbreviation	Meaning
53BP1	TP53-binding protein 1
ANOVA	analysis of variance
APH	Aphidicolin
AQR	Aquarius
ATAD5	ATPase family AAA domain-containing protein 5
ATM	Ataxia Telangiectasia Mutated
ATR	Ataxia Telangiectasia and Rad3 related
ATRIP	ATR-Interacting Protein
A.U.	Absolute Units
AUC	Area under the curve
BLM	Bloom
bp	base pairs
BPE	Bovine Pituitary Extract
BRCA2	Breast cancer type 2 susceptibility protein
BrdU	Bromodeoxyuridine
BRG1	transcription activator BRG1
CD	co-directional
CDC	Cell Division Cycle (number)
CDK	cyclin-dependent kinase
CDK9	cyclin-dependent kinase 9
CFS	common fragile site
ChIP	chromatin immunoprecipitation
CHK1	checkpoint kinase 1
CHK2	checkpoint kinase 2
CMG	CDC45-MCM2-7-GINS
CTCF	CCCTC-binding factor
CTD	C-terminal domain
CTD1	Chromatin licensing and DNA replication factor 1
DAPI	4′,6-diamidino-2-phenylindole
DDX	DEAD-box helicase
DHX	DEAH-box helicase
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulfoxide
DOT1L	disruptor of telomeric silencing-1-like
DOX	Doxycycline
DRB	5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole
DRIP	DNA-RNA immunoprecipitation
DSB	double strand breaks

dsDNA	double-stranded DNA
DSIF	DRB sensitivity inducing factor
E. coli	*Escherichia coli*
EBNA1	Epstein-Barr nuclear antigen 1
ECFP	enhanced cyan fluorescent protein
EdU	5-Ethynyl-2′-deoxyuridine
EME1	crossover junction endonucleases EME1
ERFS	early replicating fragile sites
ETAA1	Ewing's tumor-associated antigen 1
FA	Fanconi Anemia
FANCD2	Fanconi anemia group D2 protein
FANCM	Fanconi anemia complementation group M
FBS	fetal bovine serum
G4	G-quadruplex
G-MiDS	G2/M DNA synthesis
GRO-seq	Global run-on sequencing
H2Ak119cr	histone H2A lysine 119 ubiquitination
H2AK119ub	histone H2A lysine 119 crontonylation
H2BK120ub	histone H2B at lysine 120 ubiquitination
H3K27me3	H3 lysine 27 trimethylation
H3K36me3	H3 lysine 36 trimethylation
H3K4me	H3 lysine 4 trimethylation
H3K79me2	H3 lysine 79 dimethylation
H3K79me3	H3 lysine 79 trimethylation
H3S10P	H3 serine 10 phosphorylation
HBEC	Human Bronchial Epithelial Cells
HO	head-on
HR	homologous recombination
hTERT	Telomerase reverse transcriptase
IF	Immunofluorescence
INO80	INO80 Complex ATPase Subunit
mAIRN	murine Antisense Of IGF2R Non-Protein Coding RNA
Mb	Megabase
MCM	minichromosome maintenance protein complex
MNase	Micrococcal nuclease
MOPS	3-(*N*-morpholino)propanesulfonic acid
mRNA	messenger RNA
MTA2	Metastasis-associated protein MTA2
MUS81	crossover junction endonucleases MUS81
NELF	negative elongation factor
NHEJ	non-homologous end joining
ORC	origin recognition complex
ori	origin of replication

PAGE	polyacrylamide gel electrophoresis
PAS	polyadenylation signal
PBS	phosphate-buffered saline
PCNA	proliferation cell nuclear antigen
PIC	pre-initiation complex
PLA	Proximity Ligation Assay
PNUTS	PP1 nuclear targeting subunit
POLD	DNA polymerase δ
POLE	DNA polymerase ε
PP1	protein phosphatase 1
P-TEFb	Positive elongation factor b
qDRIP	Quantitative DRIP
qPCR	quantitative PCR
RAD52	DNA repair protein RAD52 homolog
RDC	recurrent DNA break cluster
RECQ1	ATP-dependent DNA helicase Q1
RECQ5	ATP-dependent DNA helicase Q5
rEGF	recombinant Epidermal Growth Factor
RNAPI	RNA polymerase I
RNAPII	RNA polymerase II
RNAPIII	RNA polymerase III
RNase	Ribonuclease
RNF20/40	RNF20/40 E3 ubiquitin-protein ligase complex
RPA	Replication Protein A
RPA32pS33	Serine 33 phosphorylated replication protein A 32 kDa subunit
RT-qPCR	reverse transcriptase-qPCR
SD	standard deviation
SDS	sodium dodecyl sulfate
Ser2P	Serine 2 phosphorylation
Ser5P	Serine 5 phosphorylation
Ser7P	Serine 7 phosphorylation
SETX	Senataxin
SLX4	structure-specific endonuclease subunit SLX4
SMARCA5	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5
ssDNA	single stranded DNA
TAD	topologically associated domain
TC-NER	transcription-coupled nucleotide excision repair
Thr4P	Threonine 4 phosphorylation
TRC	transcription-replication conflict
tRNA	transfer RNA
TSS	transcription start site

TSS	transcription start site
Tyr1P	Tyrosine 1 phosphorylation
WDR82	WD repeat-containing protein 82
γH2AX	histone H2AX phosphorylation

9. Appendix

9.1 Table S1: Integration site positions

CHROMOSOME	START LOCATION	END LOCATION	NAME	PHASE	STRAND
2	104941714	104941978	Chr 2 site 1	0	.
2	168374974	168375175	Chr 2 site 2	0	.
5	134521192	134521451	Chr 5 site	0	.
9	137626501	137626702	Chr 9 site	0	.
10	87030690	87030891	Chr 10 site	0	.

9.2 Publications

- Werner, M., Trauner, M., Schauer, T., Ummethum, H., Márquez-Gómez, E., Lalonde, M., Lee, C.S., Tsirkas, I., Sajid, A., Mur iello, A.C., et al. (2025). Transcription-replication conflicts drive R-loop-dependent nucleosome eviction and require DOT1L activity for transcription recovery. Nucleic Acids Res. *53*, 13–14. 10.1093/NAR/GKAF109.
- Marmolejo, C.O., Sanchez, C., Lee, J., Werner, M., Roberts, P., Hamperl, S., and Saldivar, J.C. (2024). A phosphorylation code coordinating transcription condensate dynamics with DNA replication. bioRxiv, 2024.05.10.593572. 10.1101/2024.05.10.593572 (in revision at Cell).
- Velychko, T., Mohammad, E., Ferrer-Vicens, I., Parfentev, I., Werner, M., Studniarek, C., Schwalb, B., Urlaub, H., Murphy, S., Cramer, P., et al. (2024). CDK7 kinase activity promotes RNA polymerase II promoter escape by facilitating initiation factor release. Mol. Cell *0*. 10.1016/J.MOLCEL.2024.05.007.
- Werner, M., and Hamperl, S. (2024). A quick restart: RNA polymerase jumping onto post-replicative chromatin. Mol. Cell 84, 186–188. 10.1016/j.molcel.2023.12.029.
- Chanou, A., Weiβ, M., Holler, K., Sajid, A., Straub, T., Krietsch, J., Sanchi, A., Ummethum, H., Lee, C.S.K., Kruse, E., et al. (2023). Single molecule MATAC-seq reveals key determinants of DNA replication origin efficiency. Nucleic Acids Res. 51, 12303–12324. 10.1093/NAR/GKAD1022.
- Ummethum, H., Lalonde, M., Werner, M., Trauner, M., Chanou, A., Weiβ, M., Lee, C.S.K., Kruse, E., Ettinger, A., and Hamperl, S. (2023). The CGG triplet repeat binding protein 1 counteracts DNA secondary structure-induced transcription-replication conflicts. bioRxiv, 2023.03.09.531843. 10.1101/2023.03.09.531843. (in revision at EMBO Reports)
- Weiβ, M., Chanou, A., Schauer, T., Tvardovskiy, A., Meiser, S., König, A.C., Schmidt, T., Kruse, E., Ummethum, H., Trauner, M., et al. (2023). Single-copy locus proteomics of early- and late-

firing DNA replication origins identifies a role of Ask1/DASH complex in replication timing control. Cell Rep. 42. 10.1016/j.celrep.2023.112045.

- Werner, M.*, Dyas, A.*, Parfentev, I.*, Schmidt, G.E., Mieczkowska, I.K., Müller-Kirschbaum, L.C., Müller, C., Kalkhof, S., Reinhardt, O., Urlaub, H., et al. (2022). ROBO3s: a novel ROBO3 short isoform promoting breast cancer aggressiveness. Cell Death Dis. 2022 139 13, 1–11. 10.1038/s41419-022-05197-7.
- Mieczkowska, I.K., Pantelaiou-Prokaki, G., Prokakis, E., Schmidt, G.E., Müller-Kirschbaum, L.C., Werner, M., Sen, M., Velychko, T., Jannasch, K., Dullin, C., et al. (2021). Decreased PRC2 activity supports the survival of basal-like breast cancer cells to cytotoxic treatments. Cell Death Dis. 12, 1118. 10.1038/s41419-021-04407-y.
- Lalonde, M.*, Trauner, M.*, Werner, M.*, and Hamperl, S. (2021). Consequences and Resolution of Transcription–Replication Conflicts. Life 2021, Vol. 11, Page 637 11, 637. 10.3390/LIFE11070637.

* = equal contribution

10. Acknowledgements

I started this incredible journey 4.5 years ago and while I always wanted to do a PhD, I had limited knowledge about what that would actually mean. Equipped with a lot of motivation and curiosity I was brave and joined a relatively young lab, hoping to do some cool science. Reflecting at the end of this journey, I must say that it was an absolutely amazing experience and a time of my life that I will always remember very fondly. Despite some hardships and struggles in the beginning, both professionally and personally, I never questioned my choice and pushed through. This would not have been possible without the continuous support of many people who were there for me all the time and also the ones, who I met along the way. Thank you so much everyone for being there for me! This work is not only my achievement but also yours.

First and foremost, I would like to thank Stephan for accepting me as a PhD student and giving me the opportunity to be part of your lab. I could not have chosen a better place to do my PhD and I am incredibly thankful for all your support throughout the years. You put together a wonderful group of people, who are not only a team but a bunch of friends doing awesome science together. I truly loved being part of the Hamperl lab! Thank you also for being the best supervisor and mentor one could ask for. You are incredibly smart, supportive, and caring. You taught me a lot about how things work in the lab but also helped me to develop my own scientific thinking and approach. These things will forever stay with me and will be extremely helpful in the future! You

are a role model boss, and I deeply admire your drive and dedication. I am eternally grateful for everything you have done in all these years, and I am sure we will always stay in touch!

Thank you, Maria-Elena, for building such an amazing institute in which science can flourish to its maximum extent. Being part of and experiencing this great international environment with its fun atmosphere is something I grew to love and hope to experience again. Moreover, I highly appreciated your scientific advice during all the TAC meetings, talks, and seminars. Challenging my viewpoints has not only helped me to improve my presentations but most importantly sharpened my critical thinking about science. I am particularly grateful for this and am convinced that this will help me even more moving forward. Thank you!

I would also like to thank all group leaders at the IES, Antiono, Eva, and Nico, who make this institute such an enjoyable working environment. Your comments during countless seminars and lab meetings were a great source of inspiration for me and helped to shape this project in many ways.

I am also tremendously thankful for the support from my TAC members Christof and Josh, who helped me to establish my project and always provided great suggestions. Moreover, I would also like to thank Josh for bringing me in for our collaboration and many of our discussions. It was truly a pleasure to work together, and I am sure we will stay in touch.

Where to start when talking about the Hamperl lab – I am so thankful for all the support everyone gave me over the years. My success is also all your success, I could not have done it without you. Thank you, Manuel, for sharing this project with me it was a pleasure to work together with you and I think our project is a great example of two people with different approaches and backgrounds coming together and making it work. I am already missing our talks about science, music, stocks, life, and all the other stuff. Thank you, Henning, for a lot of help with experiments, protocols, and getting started in the beginning. This really made my journey much smoother. Thank you also for the many hilarious lunch discussions. I am also missing our regular bouldering sessions. Thank you, Elisabeth, for your incredible support over the years. Your passion and determination made this lab truly exceptional, and I really miss my bay buddy and our small and sometimes not-so-small chats during the day. Thank you, Maxime, for the great scientific advice, the master macro, endless jokes and humor, and our weekly One Piece discussions. You will be a great PI one day! Thank you also to Anna and Matthias, who helped me a lot at the beginning of my project. You were always there when I needed something, and I could always rely on you! Thank you also to Clare, Atiqa, and Ioannis. It has been a pleasure to work with you and the

countless discussions about scientific and fun stuff have not only helped to shape my project but also made lab life incredibly fun! I will miss the cooking and baking skills of all of you! Thank you, Elizabeth, for being so brave to join as your bioinformatician and helping to push my project over the finish line. Also, thank you for your relentless energy and great vibes that are a true embodiment of our lab's spirit! Thank you to Augusto and Xanthoula for keeping the lab running. Without you, we would have been doomed long ago. Your passion and resilience were always admirable. Also thank you, Xanthoula, for being my occasional private driver. Finally, thank you to all the students and interns over the years – Ekin, Melina, Zhi-Hoon, Sara, August, Simon, and Zhao - who not only made significant scientific contributions to this work but also were crucial for many great events over the years. Thank you all!

I am tremendously grateful to everyone at the IES. It is an amazing scientific environment with even more amazing people. Thank you for the great atmosphere and countless events such as our hikes, the international lunch, the movie nights, and the sports day. Also thank you for countless discussions and feedback during my many talks and seminars. Your feedback helped to improve my work a lot. While I cannot mention everyone, I would like to particularly thank Tamas for his incredible bioinformatic support. Without you, many things we have done would have been possible. Moreover, I would like to give a special thanks to the afternoon snack kitchen crew – especially Ana, Meghana, and Mrinmoy. I will miss these discussions a lot! I would also like to thank Mrinmoy again for his help with countless sequencing experiments. You were always there if I had a question. Thanks to our secretaries Marie-Sophie, Laura, and Thomas. Your work and dedication made my life much easier countless times. I am also grateful for the microscopy support from Andreas. Thank you everyone!

Apart from the lab and institute I would like to thank all my friends for their continuous support throughout the years. You all always had an open ear for me and were there for me when I needed it. Special thanks to Nathalie for always being there and sharing this journey with me for almost 10 years. Even though our professional paths split many years ago, I could not ask for a better best friend, who was and will always be there for me. Thank you, Yuliia, for your unending support and for being there whenever I needed you. Thank you to the Göttingen gang for the fun times throughout the years. You kept the spirits high in good and bad times. Thank you, Maxi, Hannes, Steven, and Christian. You guys always believed in me right from the start, you were always there, and we will always be in touch!

Mein größter Dank gilt selbstverständlich meiner Familie und meiner Freundin, die mich über all die Jahre immer bedingungslos unterstützt haben. Ihr wart immer da sowohl in guten als auch in

schlechten Zeiten. Ohne euch, eure Hilfe und eure Liebe wäre all dies nicht möglich gewesen. Danke Melanie für deine nie endende Unterstützung, Inspiration und Liebe. Ich könnte nicht glücklicher sein als mit dir an meiner Seite. Mein unendlicher Dank gilt meinen Eltern Ingrid und Frank, die immer an mich geglaubt und mir all dies ermöglicht haben. Ich kann mich nicht glücklicher schätzen, euch zu haben. Danke für alles!