RNA Polymerase II subunit overexpression induces genome instability and deregulate transcription

by

Martina Muste Sadurni

A Thesis submitted to the University of Birmingham for the Degree of DOCTOR OF PHILOSOPHY

Institute of Cancer and Genomic Sciences

College of Medical and Dental Sciences

University of Birmingham

December 2023

Table of contents

Acknowledgements

Most of all, I thank my supervisor Marco Saponaro for the infinite support he gave me, for always being positive when I wasn't and for believing in me more than I did, sometimes. Thank you Marco, your passion, hard work, and all your thousand ideas have made this PhD "enjoyable" even when it got extremely challenging. Thank you for always listening and truly understanding when I was overwhelmed or when I just needed a push! I know that with your work, your research and your passion you'll keep inspiring me in the future, like you did during these years. I rarely say this (we all know) but I feel very lucky I got the chance to meet you and work with you.

A special thanks to all the members of the Saponaro lab, present and past, and to the Gambus and Monteiro labs: all the help, jokes, hugs, tears, support, discussions, suggestions and Friday "I'm done with today" drinks made these years slightly lighter and easier, thank you guys!

Thanks to all my friends, wherever you are in the world, you still manage to be close.

Un grazie speciale alla mia famiglia, che mi ha permesso di arrivare fin qui, insegnandomi quanto sia preziosa l'arte del "lasciar andare".

p.s. Thanks everybody for all the chocolate.

Disclaimer

For plagiarism check purposes, we have to disclose that in the Introduction and in the final Discussion chapters of this thesis the author used text from the review “Deregulations of RNA Pol II Subunits in Cancer, Muste Sadurni and Marco Saponaro, Applied Biosciences 2023, 10.3390/applbiosci2030029”. The entire test of the review and the figure of the review, used in the final Discussion chapter, are an original work of the author of the thesis.

Abstract

Deregulated transcription is a well-established hallmark of cancer. However, studies on deregulations affecting the transcription complexes and specifically on RNA Polymerase II (Pol II) complex, remain elusive. In this work we propose the upregulation (defined as “PolUp” in this thesis) of the subunits POLR2A, POLR2C and POLR2D of Pol could constitute a new tumor biomarker, given the poor survival rates and increased genomic instability found in cancer patients baring at least one of the subunits upregulated. Using the HeLa T-REx inducible cell line model overexpressing the single subunit POLR2A,-2C or -2D, we investigated the consequences of subunits overexpressions on genome stability and transcription. Our finding suggest that each subunit overexpression impacts on DNA damage and transcription uniquely. By immunostaining and γH2AX ChIP-Seq we demonstrated that the overexpressions increase DNA damage formation and we identified the sites of damage. Furthermore, we hypothesised that the increased DNA damage is a result of transcription deregulations arising upon each subunit overexpression. We characterize nascent transcription by TT-chem-Seq and specific Pol II phosphorylation deregulations by ChIP-Seq and our findings suggest that even if there is a general increase in transcription activity, each subunit plays a different role on transcription regulation. Indeed, we identified Ser2 hyper phosphorylation and Ser5 hypo phosphorylation of Pol II CTD upon POLR2C and POLR2D overexpression, linked to a BRD4 and CDK7 defect, respectively. Finally, we were able to specifically target these deregulations with the BRD4 inhibitor JQ1 and the CDK7 inhibitor THZ1, showing how the mechanism identified to be responsible of transcription deregulations could be exploited for therapeutic purposes.

1. Introduction

1.1 RNA Polymerase complexes

Eukaryotic transcription is carried out by three different RNA polymerase machineries: RNA Polymerase I (Pol I), RNA Polymerase II (Pol II) and RNA Polymerase III (Pol III). Pol I is entitled to transcribe ribosomal RNAs (28S rRNA, 18S rRNA, and 5.8S rRNA); Pol II is devoted to the transcription of the protein coding part of the genome, messenger RNA (mRNA) and also for the transcription of long non-coding RNAs (lncRNAs) and small nuclear RNAs (snRNAs); Pol III is responsible for the transcription of transfer RNAs (tRNAs), 5S rRNA and small nuclear RNAs (snRNAs). RNA Polymerases I, II and III are complexes of 14, 12 and 17 subunits, respectively (Cramer et al., 2008; Gerber et al., 2020). Five of the subunits, named DNA-directed RNA Polymerase subunits, RPB5-6-8-10-12 are shared among all three polymerase complexes. These subunits are encoded by the genes POLR2E-F-H-K-L respectively.

- In Pol I, RPB5, RPB6, RPB8, RPB10, and RPB12 together with subunits RPA190, RPA135, RPAC40, RPAC19, and RPA12.2 form the core of the enzyme. Four other subunits, RPA43-RPA14 and the RPA49-RPA34.5 complete the complex (Vannini and Cramer 2012). The subunits specific to RNA Pol I are encoded by the genes *POLR1A-H*.
- In Pol II complex (built up of 12 subunits RPB1-12), RPB1, RPB2, RPB3 and RPB11 share sequence homology with other Pol I and III subunits. RPB4, RPB7 and RPB9 and the disordered C-terminal domain (CTD) of RPB1 subunit are specific to Pol II (Thomas and Chiang 2006). The genes encoding for Pol II 12 subunits are *POLR2A-K*.

- RNA Pol III is constituted of a core of 10 subunits and a peripheral heterodimeric stalk. The TFIIF-like RPC4/5 and the TFIIE-like RPC3/6/7 subcomplexes are Pol III specific. The subunits specific to Pol III are encoded by *POLR3A-K* genes (Ramsay et al. 2020).

Yeast and human Pol II are complexes of 12 different subunits, and the high similarity between the two eukaryotic polymerases has enabled many advances on deciphering their structure and function. Pol II active centre is formed by RPB1 and RPB2 subunits, which sit at two opposite sides of a positively charged "cleft", the entry point of DNA for elongation, together with RPB3, -10, -11, -12 (Hahn, 2004; Patrick Cramer, 2004). RPB1 forms a mobile "clamp" on its side of the cleft, which can be in an open or closed state depending on whether the structure is devoid or not of DNA and RNA, respectively (Hahn, 2004; Cramer et al., 2008). Importantly, a region in RPB1 named trigger loop (TL) is implicated in the stabilization of the incoming nucleotide and therefore in transcription fidelity (He et al., 2016; Gamba & Zenkin, 2018). A small portion of RPB2 and the N-terminal of RPB6 contribute to the structure of the clamp (Woychik and Hampsey 2002). From the base of the clamp protrudes the "linker", connecting RPB1 with its CTD (Cramer, 2004), which coordinates transcription progression and co-transcriptional events. RPB2 side of the cleft constitutes the "lobe" and "protrusion domains", forming a "wall" blocking the end of the cleft (Hahn, 2004; Cramer et al., 2008). Beneath the active centre a "pore" includes a RPB1 loop binding the Mg2+ metal ion (Cramer, 2004). Generally, Pol II is divided into functional domains defined as the clamp (comprehensive of RPB1 mainly and RPB2), the jaw lobe (RPB2, RPB1 and RPB9), the RPB5 jaw and the stalk (RPB4, -7) (Figure 1) (Kostek et al. 2006).

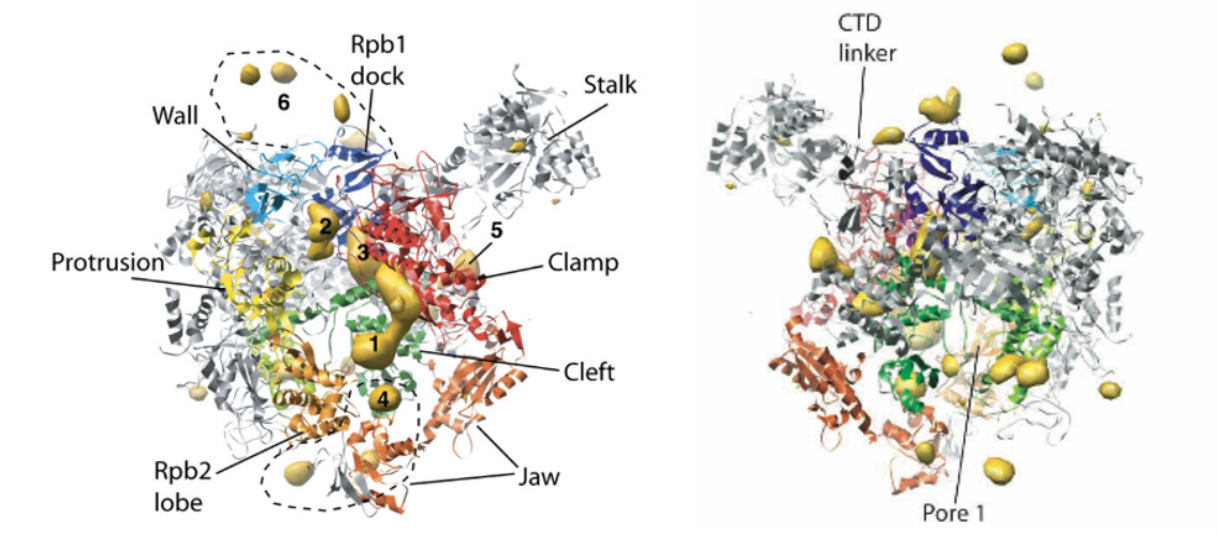


Figure 1 Structure of RNA Pol II functional domains front view (left) and back view (right) of yeast Pol II crystal structure superimposed with electron microscopy (EM) map of human Pol II. From the yeast structure are named the regions of the dock, clamp, pore and jaw (Rpb1), wall, lobe and protrusion (Rpb2), jaw (Rpb5) and stalk (Rpb4-7) and the CTD linker. From human Pol II (yellow) are evident the positions of the pore and clamp head (1), the area where DNA:RNA hybrid binds (2), a region in Rpb2 facing the active site (3), the Rpb1 region contacting Rpb5 and Rpb9 in the jaw (4), Rpb6 N-terminus (5) and unstructured CTD (6). Adapted figure

The latter binds the polymerase core through RPB7 N-terminal domain, called the “tip”. The stalk subunits could function in recruiting factors to the CTD and in mediating the interactions with transcription initiation factors, thus contributing to promoter specificity (Cramer, 2004). Additionally, RPB4 tunes the CTD phosphorylation levels (Allepuz-Fuster et al. 2014) and guides nascent RNA co-transcriptional processing, by directly interacting with the nascent transcripts and recruiting co-transcriptional factors, thus being important in transcription termination (Allepuz-Fuster et al. 2019; Garrido-Godino et al. 2016; Richard et al. 2019, 2021).

1.1.1 RNA Pol II assembly and nuclear import

Pol II assembly occurs in the cytosol and involves multiple subunit-subunit interactions and the formation of three distinct sub-assemblies (Kimura, 2000; Wild & Cramer, 2012).

Moreover, some of the RNA Pol II subunits have a similarity with the bacterial enzyme (Sweetser et al. 1987), which provides basis for modelling the eukaryotic assembly. Indeed, RPB3 subassembly (comprehensive of RPB3, RPB10, RPB11, RPB12 subunits) corresponds to the $\alpha\alpha$ bacterial dimer; RPB2 subassembly (RPB2 and RPB9) to the ß subunit; RPB1 subassembly (RPB1, RPB5, RPB6 and RPB8), to the ß' subunit. In particular, RPB6 as a homolog of ω bacterial subunit is proposed to help in the formation of RPB1 subassembly. Then, the peripheral RPB4/7 stalk should be added as a preassembled complex to Pol II core (Wild & Cramer, 2012). Pol II subunits lack a nuclear localization signal (NLS), therefore additional factors are required for their nuclear import. In yeast, Iwr1 binds the active cleft between the two largest subunits of Pol II and is responsible for Pol II nuclear import (Czeko et al. 2011). Since it binds the active cleft, it could represent a way to ensure that only correctly assembled Pol II is imported in the nucleus (Wild and Cramer 2012). However, yeast strains depleted of Iwr1 are still viable, suggesting that there might be additional import pathways (Czeko et al. 2011). Since RPB1 and RPB2 are not able of passively diffusing, contrary to smaller subunits, including RPB3, it has been proposed that Iwr1-independent pathways may sustain cell viability and that in Iwr1-depleted cells there might be specific pathways for different subunits, implicating that Pol II could enter the nucleus not as a whole (Gómez-Navarro and Estruch 2015). Cytosolic assembly of Pol II depends on a wide range of factors, among which GPN family of small GTPases (which share a unique Gly-Pro-Asn motif) GPN1, GPN2 and GPN3 (Boulon et al. 2010; Carre & Shiekhattar, 2011; Wild & Cramer, 2012; Forget et al., 2010; Forget et al., 2013; Martino et al., 2018; Zeng, F. et al., 2018). In the model proposed by Niesser and colleagues it is suggested that GPN-loop GTPases together with the other factors

function in the cytosolic assembly of Pol II. Once assembled, the enzyme is recognised mainly by Iwr1 and imported into the nucleus (Niesser et al. 2016).

1.2 RNA Pol II transcription

Messanger RNAs (mRNAs) and long non coding RNAs (lncRNAs) are transcribed by RNA Pol II, which works in concert with transcription activators or repressors. Among those, specific transcription factors (TF) bind regulatory elements, promoters and enhancers, and either through direct protein-protein interaction or through a remodelling action on chromatin have an impact on the assembly of the transcription machinery (Soutourina 2018). Transcription is a tightly regulated process, which can be divided in three main phases: initiation, elongation and termination (Figure 2)

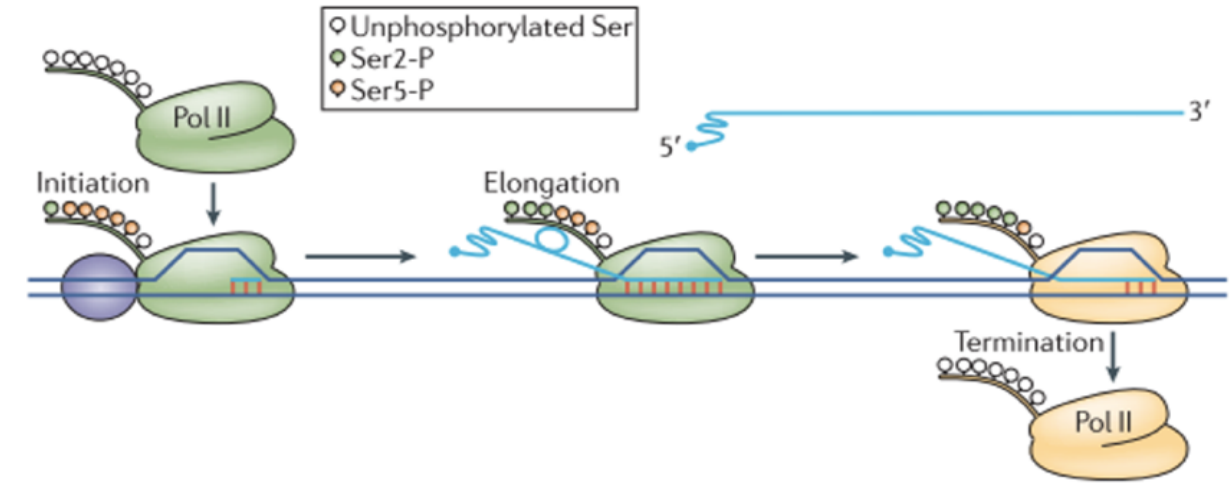


Figure 2 Class II genes transcription. An unphosphorylated Pol II is firstly recruited by initiator factors. Then, CTD becomes hyperphosphorylated at Ser5 (towards the 5'end of the gene) and Ser2 (towards the 3'end) during elongation, until Pol II reaches the termination site, the nascent transcript is cleaved and the polymerase is dephosphorylated and released from DNA. Adapted from Kuehner et al., 2011.

1.2.1 Transcription initiation

The first layer of regulation consists in the formation of the pre-initiation complex (PIC), which is a multiprotein assembly of general transcription factors (GTFs, TFIIA, -B, -D, -E, -F, -H), Pol II, Mediator complex and a large number of co-activators, co-repressor and chromatin remodelers (Figure 3) (Gupta et al. 2016). Transcription initiates at the transcription start site (TSS), located at the 5'-end of a gene. Upstream and downstream the TSS lays the core promoter, which main function is to recruit and bind the PIC elements. Together with the core promoter, in order to increase transcription low basal activity, TFs bind other regulatory elements, enhancers (Haberle and Stark 2018). Although generally all the PIC factors assemble at protein coding genes core promoters, there can be differences in PIC composition at different Pol II transcribed genes or among different cell types, depending on the core promoter elements composition (Schier and Taatjes 2020).

Seven core promoter elements have been identified, usually located around -50bp and +50bp of the TSS. In a simplistic view, eukaryotic promoters can be divided into two different types: the ones containing a functional TATA-box and the so called TATA-less promoters (Tora and Timmers 2010). First, the well-known TATA-box core promoter element, which is recognised by the TATA binding protein (TBP) subunit of TFIID, the first GTF starting the PIC formation (Kornberg, 1999; Thomas & Chiang, 2006). A second core promoter element is the initiator (Inr), which can function either alone or together with TATA-box and/or other core promoter elements. Downstream many TSS in TATA-less promoters in Drosophila and more rarely in humans, is the downstream promoter element (DPE). Both Inr and DPE, which usually are found together, are recognised by the (TBP associated factors) TAFs subunits of TFIID (Burke and Kadonaga 1997; Haberle and Stark 2018). In addition to the DPE, the motif ten element

(MTE) and the downstream core element (DCE) are also found downstream the TSS, and the TAF1 component of TFIID contacts this elements, thus being a crucial player in the establishment of downstream promoter contacts (Thomas & Chiang, 2006; Haberle & Stark, 2018). Notably, TAFs are also found in the Spt–Ada–Gcn5–Acetyl transferase (SAGA) complex, which is a transcriptional co-activator (Gupta et al., 2016, Helmlinger & Tora, 2017). Finally, the last two core promoter elements are constituted by the upstream and downstream TFIIB recognition element (BREu, BREd, respectively), which consensus sequences are indeed recognised by TFIIB (Thomas & Chiang, 2006; Haberle & Stark, 2018). Indeed, the plethora of core promoter elements represents another layer of gene regulation.

According to the sequential assembly pathway, once TFIID recognises via its TBP and TAFs subunits the different core promoter elements, TFIIA and TFIIB join the complex, recruiting next Pol II and TFIIF to the PIC (Thomas and Chiang 2006). TBP bends the double helix to 90° (Liu, Bushnell, and Kornberg 2013), thus wrapping the DNA around the polymerase. In TATA-box promoters, TFIIA binds the upstream region of the core promoter and TBP with a conserved 12-stranded β-barrel domain, therefore stabilizing the complex (Sainsbury, Bernecky, and Cramer 2015). TFIIB maintains different contacts with RPB1 and RPB2 Pol II subunits, DNA and TFIIF as well, thus helping in Pol II recruitment and complex stabilization (Kornberg, 1999; Woychik & Hampsey, 2002; Thomas and Chiang, 2006; Xin Liu et al., 2013). TFIIB is also important for the correct positioning of template DNA in the active site of Pol II, for directing newly synthetized RNA to the exit tunnel and for the TSS selection. The latter is also a function of the heterodimer constituted by RAP74 and RAP30 subunits of TFIIF (Sainsbury et al. 2015). The dimerization module of the two subunits directly contacts Pol II (He et al. 2016), therefore TFIIF has the additional role of stabilizing the complex formed so

far and facilitate the recruitment of TFIIE and TFIIH (Woychik and Hampsey 2002). TFIIH is constituted of ten subunits and seven of them form the core of the GTF. Among those, worth to mention are the xeroderma pigmentosum group B and D complementing protein (XPB, -D) ATPases. The other three subunits constitute the CDK activating module (CAK), which comprises CDK7 (Rimel and Taatjes 2018). TFIIH helicases unwind the transcription start site allowing transcription to proceed (Kornberg 1999): XPB activity is required for promoter DNA unwinding, while XPD functions mainly in Nucleotide excision repair (NER) pathway (Rimel and Taatjes 2018). TFIIE and TFIID promote the integration of the CAK functional domain of TFIIH in the PIC, favouring the phosphorylation of Pol II CTD by said domain, through establishing interaction with Pol II stalk (RPB4/7) and directing the CTD allowing its phosphorylation (Compe et al. 2019; Chen et al. 2021). The core of TFIIH, is crucial to position DNA in the active site of Pol II. These last GTF entering the PIC are released once elongation factors (EFs) are recruited. In fact, same regions of Pol II are engaged in binding both TFIIE and the EF DRB-sensitivity inducing factor (DSIF), suggesting that the first could be substituted by the latter during transcription (Compe et al. 2019).

In summary, promoter is firstly recognised in a closed complex (CC) conformation by TFIID, -A, -B, followed by TFIIF-Pol II and finally TFIIE and TFIIH. Subsequently, in the presence of ATP, DNA is unwound and TFIIF-Pol II and TFIIE, -H stabilize the formation of an open complex (OC), allowing the formation of an initially transcribing complex (ITC) (Figure 3) (Sainsbury et al., 2015; Hantsche & Cramer, 2017). In the OC, DNA is unwound forming the transcription bubble. Here, the ATP-dependent translocase activity of XPB helicase opens between 11 to 15 bp around the TSS, inducing enough torsional strain to position the template DNA in the active site of Pol II, where Pol II clamp slightly closes and RNA synthesis initiation occurs (He et al.

2016). The template DNA is stabilised by interaction with GTFs (Gupta et al., 2016; Rimel & Taatjes, 2018). While the upstream edge of the transcription bubble is fixed by interactions with PIC elements, the downstream edge expands the bubble and first RNA synthesis takes place. Only when nascent RNA reaches a critical length, around 12-13 bp, the bubble collapses, since TFIIB is released, and template strand is exposed and can reanneal (Xin Liu et al., 2013; Sainsbury et al., 2015). In the active site of Pol II a stable 8bp DNA:RNA hybrid is formed (Bernecky et al. 2016). The moment when TFIIB dissociates signs the "promoter escape" phase of transcription initiation, that is going to be followed by promoter proximal pausing (PPP). Importantly, to achieve promoter escape the CTD of RPB1 has to be phosphorylated on Ser5 of the heptad repeats by the CAK module of TFIIH (Fuda et al., 2009; Chen et al., 2018).

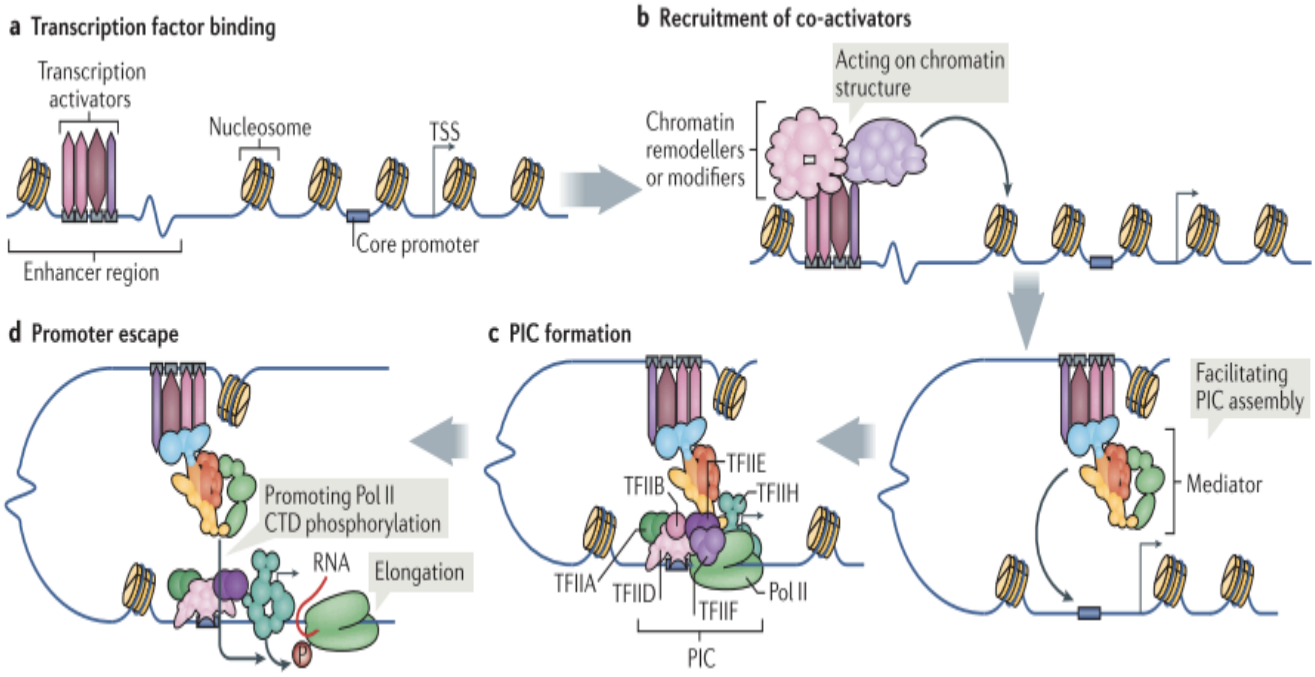


Figure 3 PIC formation and promoter escape Transcription activators bound to enhancer regions and chromatin modifiers (a, b) promote transcription initiation by allowing the interaction between Mediator complex and GTFs, which assemble at promoter regions together with Pol II forming the PIC (c). Subsequent CTD Ser5 phosphorylation marks promoter escape and first RNA synthesis while Pol II proceeds into elongation phase (d). Adapted figure from Soutourina, 2018.

1.2.1.1 Mediator Complex

As previously mentioned, additional factors are required for the correct formation of the PIC and transcription initiation. Among those, the Mediator complex (Kornberg, 1999; Hantsche & Cramer, 2017), highly conserved from yeast to human, with 25 and up to 30 subunits, respectively. It consists of four different modules: the Head, Middle and Tail and CDK8 module, which only associates transiently with the complex (reviewed in Soutourina, 2018). The main role of this complex is to mediate the crosstalk between TFs bound enhancers and promoters and Pol II in the PIC (Ramasamy et al. 2023). Therefore, the Mediator complex is recruited at enhancers by TFs and directly interacts with PIC components at core promoters, stabilizing the complex and positively regulating transcription (Figure 3) (Plaschka et al., 2015; Haberle & Stark, 2018). This is achieved through extensive interactions between the Mediator complex subunits and GTFs and Pol II (Gupta et al., 2016; Schier & Taatjes, 2020). More precisely, Mediator together with TFIID help recruiting Pol II to the PIC, in fact structural studies show how Mediator and Pol II interact through contact sites on RPB1, RPB3-11 and the RPB4/7 stalk (Abdella et al. 2021; Chen et al. 2021; Rengachari et al. 2021; Richter et al. 2022). Mediator and TFIID also help in recruiting and stabilizing TFIIH, thus promoting the phosphorylation of RPB1 CTD on Ser5 by the CAK module (Abdella et al. 2021; Chen et al. 2021; F. X. Chen et al. 2018; Fuda et al. 2009; Haberle and Stark 2018; Rengachari et al. 2021; Robinson et al. 2016; Soutourina 2018)

1.2.2 Elongation

1.2.2.1 Elongation factors and promoter proximal pausing

After promoter escape, the polymerase synthesises RNA for 20 to 60bp, where Pol II pauses, dissociates from PIC elements and finally engages in a functional elongation complex (EC) (Gupta et al. 2016). Pol II pausing is restricted to higher eukaryotes, indeed unicellular organisms and yeast do not show promoter proximal pausing (Booth et al., 2016). Pausing of Pol II is important to ensure the correct recruitment of additional factors involved in co-transcriptional RNA processing (Schier and Taatjes 2020). In fact, CTD phosphorylation also serves as a recruiting signal for capping enzymes, splicing and termination factors (Ghosh et al., 2011; Maita & Nakagawa, 2020; Schier & Taatjes, 2020). Furthermore, it also stimulates the histone-lysine N-methyltransferase (SETD1A), which maintains the trimethylation of histone H3 Lys 4 (H3K4me3) levels, that promotes transcription activation (Ebmeier et al. 2017).

The transition from the initiation complex (IC) to EC is achieved when GTFs are exchanged with EFs (Ehara et al., 2017; Vos et al., 2018; Compe et al., 2019). Conformational changes following DNA unwinding in the transcription bubble and first RNA synthesis enable the EF DSIF to contact the polymerase (He et al. 2016; Bernecky et al. 2016). DSIF is composed of two EFs: SPT4 and SPT5. The latter is formed by an N-terminal acidic domain, known as NusG N-terminal (NGN) domain, a carboxy terminal region and multiple Kyprides, Ouzounis, Woese (KOW) domains, while SPT4 is a small zinc finger protein (Hartzog and Fu 2013). KOW domains and NGN maintain multiple contacts with Pol II subunits as well as RNA, with KOW5 stabilising the strongest interaction with the polymerase (Bernecky et al. 2017). SPT5 C-terminal region, similarly to Pol II CTD, is phosphorylated by either TFIIH and the positive-transcription elongation factor b (P-TEFb), in order to help in capping complex and topoisomerase 1 (TOP1) recruitment (Chen et al., 2018; Haberle and Stark, 2018). DSIF has both a repressive and

activating function on elongation and the latter is favored when its CTR is phosphorylated by P-TEFb (Hartzog and Fu 2013).

Negative elongation factor (NELF) is a complex of four subunits (NELFA, -B, -C, -E) that stabilizes Pol II pausing thus avoiding premature transcription termination (Chen et al. 2018). In particular, NELFA/C dimer binds to RPB1 in the funnel region and stabilizes a DNA:RNA hybrid, non-productive for further elongation, in the active site (Vos et al. 2018). Consistently with the model stating that initiation factors are substituted by elongation factors during IC to EC transition, TFIIS association with Pol II is mutually exclusive with NELF (Vos et al. 2018).

DSIF together with the negative elongation factor (NELF), forms part of the paused RNA Pol II complex, which requires further action of EFs to finally engage in productive elongation (Haberle and Stark, 2018). Recently, also TFIID has been implicated in PPP, indicating a role for the PIC in the establishment of Pol II pausing, whilst DSIF and NELF being fundamental for its maintenance (Fant et al. 2020).

P-TEFb, constituted by the kinase/cyclin pair CDK9 and CCNT1 (Cyclin T1), has the crucial role of smoothing the PPP release into productive elongation (Fuda et al., 2009). P-TEFb is recruited by the transcriptional cofactor BRD4 (Bromodomain containing 4) or the super elongation complex (SEC) (Moon et al. 2005; Yang et al. 2005; Zhou et al. 2022) and through its kinase subunit CDK9 phosphorylates both DSIF and NELF and Pol II CTD at Ser2, releasing Pol II from pausing (Buratowski, 2009; Vos et al., 2018; Haberle and Stark, 2018). The general transcription factor P-TEFb is usually not bound to chromatin and present in an inactive state, bound to 7SK snRNA and HEXIM1, MEPCE and LARP7 (Chen et al. 2004). P-TEF-b components are released as monomers and are activated through the function of either BRD4 or SEC (Chen et al. 2014; Zhou et al. 2022). Normally, BRD4 activity as a positive regulator of transcription

is linked to increased rates of Pol II release in positive elongation and increased Ser2 phosphorylation of the CTD (Altendorfer et al. 2022; Arnold et al. 2021). However, in the lack of BRD4 or inhibition with JQ1 there are evidences suggesting that a compensatory mechanism, possibly relying on HEXIM1 disassociation from the P-TEFb inhibitory complex, may be triggered resulting in increased RNA synthesis and Ser2 phosphorylation (Arnold et al. 2021; Bowry et al. 2018; Song et al. 2020). Worth to mention that BRD4 is an activator of MYC dependent transcription and therefore it has been proposed as a good target for cancer therapy, especially since a specific inhibitor, JQ1, has been found to efficiently target BRD4 (Kotekar et al. 2023; Singh and Alauddin 2023). Upon its phosphorylation by P-TEFb, NELF can dissociate from chromatin and DSIF becomes a positive elongation factor, reinforcing the productive elongating form of Pol II, in which template DNA can easily enter the central cleft (Chen et al. 2018). Following the PPP release, transcription improves and Pol II moves towards the 3'end of the genes (Figure 4) (Jonkers and Lis 2015). Importantly, promoter proximal pausing offers the opportunity for further gene expression regulation, like gene attenuation (Wagner et al. 2023). In fact, not all Pol II successfully escapes promoter to engage in productive elongation, as more evidence shows that it can also undergo premature termination, mediated by the activity of the Integrator complex and specifically by the RNA endonuclease activity of its subunit 11 (INTS11), that cleaves nascent RNA thus promoting Pol II termination (Baillat et al. 2005; Elrod et al. 2019; Wagner et al. 2023).

1.2.2.2 Elongation and splicing

As previously mentioned, RNA Pol II pausing is crucial to ensure the recruitment of factors involved in mRNA processing pathways such as capping, splicing and 3' end processing (Peck et al. 2019). The process of splicing of introns out of mRNA and joining of exons occurs co-

transcriptionally, while Pol II is actively transcribing in the elongation phase, with the spliceosome assembling on the nascent RNA and recognising specific sequences along the introns known as splice sites (ss) (Dujardin et al. 2013). Since elongation and splicing are coupled events, elongations rates can affect splicing and cause alternative splicing events, for example exon skipping or intron retention (Fong et al. 2014; Mimoso and Adelman 2023; Tellier et al. 2020), but also splicing itself can affect transcription, since depletion of components of the spliceosome can cause Pol II accumulation and slow down transcription (Dujardin et al. 2013; Fong and Zhou 2001; Lin et al. 2008), or it can cause premature 3' end processing and transcription termination (Rodríguez-Molina et al. 2023).

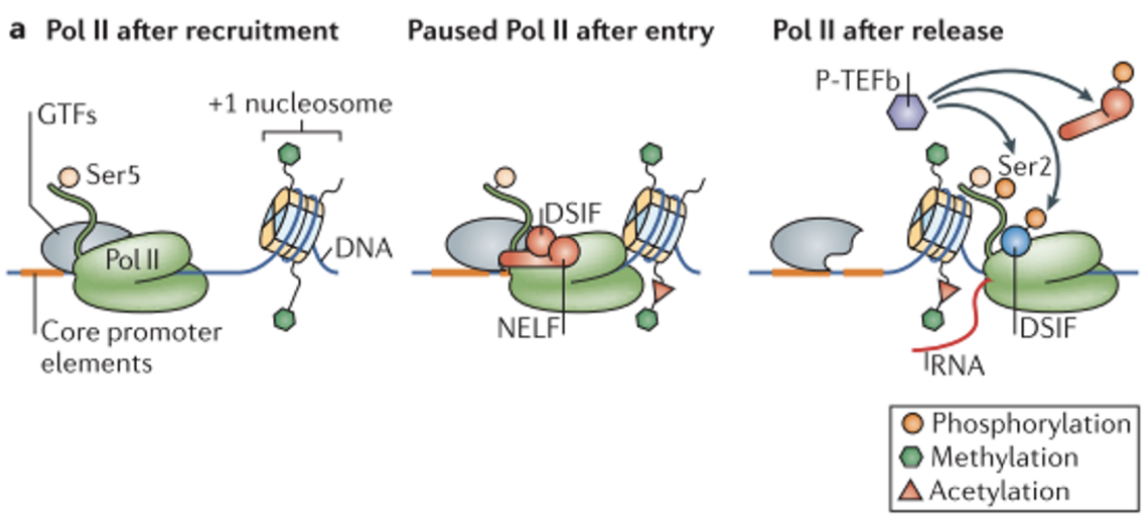


Figure 4 Promoter proximal pausing and release After promoter escape and brief RNA synthesis, Pol II pauses facilitated by the action of negative elongation factors DSIF and NELF. Upon the recruitment of P-TEFb and phosphorylation of DSIF, NELF and Pol II CTD on Ser2, Pol II is released into productive transcription, where phosphorylated DSIF functions as a positive elongation factor. Adapted figure from (Jonkers and Lis 2015)

1.2.3 CTD phosphorylation

CTD is a low complexity domain belonging to the RPB1 subunit. It consists of heptad repeats of YSPTSPS residues, which vary from 26 in yeast to 42 in flies and 52 in humans. In mammals, of the 52 repeats 21 are consensus and 31 non consensus. The consensus repeats are phosphorylated on Tyr1, Ser2, Thr4, Ser5, Ser7, while the two Prolines can undergo cis and trans isomerization (Heidemann et al., 2013; Zaborowska et al., 2016). In the non-consensus repeats, Lys, Arg and Thr can be found at position 7, and when post translationally modified can further influence modifications of neighbouring residues (Zaborowska et al., 2016; Burkholder et al., 2019). Ser5 (pSer5) and Ser2 (pSer2) phosphorylation are the most abundant in almost all the heptads repeats with Tyr1, Thr4 and Ser7 being two times less frequent. It is also unusual to find the same heptad repeat modified at different residues, indicating that multiply phosphorylated repeats represent Pol II intermediates in specific transcription stages (Schüller et al., 2016; Suh et al., 2016).

While Pol II engages in active transcription, towards the 3'-end of the genes, the levels of pSer5 decrease in favour of pSer2 (Buratowski 2009). CTD kinases are often promiscuous: CDK7 can phosphorylate both Ser5 and Ser7 (pSer7) and P-TEFb has been found to phosphorylate also Ser5, especially in the presence of pSer7 (Czudnochowski et al., 2012; Harlen & Churchman, 2017). However, recent studies have shown how Tyr1 phosphorylation (Tyr1P) by ABL1 kinase directs P-TEFb preference for Ser2 upon Ser5 (Mayfield et al. 2019). CTD phosphorylation is fine-tuned during transcription to ensure the occurrence of co-transcriptional processes such as RNA splicing, capping and 3'-end processing (Ghosh et al., 2011; Maita & Nakagawa, 2020). Indeed, not only the positioning of post-translationally modified marks is important, but also their correct removal by phosphatases is pivotal for transcription and co-transcriptional

events (Eick & Geyer, 2013; Heidemann et al., 2013; Genes Robles & Coin, 2019). Interestingly, studies in yeast have highlighted the involvement of RPB4 subunit in recruiting CTD phosphatases Fcp1 and Ssu72, keeping the phosphorylation levels under control (Allepuz-Fuster et al. 2014).

Noteworthily, CTD as a low complexity domain (LCD) can undergo length-dependent liquid-liquid phase separation (LLPS), both in a phosphorylated and unphosphorylated form (Boehning et al. 2018). It has been suggested that Pol II recruitment to promoters may be achieved through the CTD engagement in phase separated compartments together with TFs and transcription activators (Harlen and Churchman 2017). Unphosphorylated CTD is sufficient to undergo phase separation, by CTD-CTD interactions, forming Pol II hubs near gene promoters. Then, Ser5 phosphorylation disrupts the established interactions, enhancing transcription initiation and CTD interaction with other LCD containing factors (Boehning et al. 2018). Indeed, CCNT1 can undergo phase separation with CTD, favouring its phosphorylation by CDK9 (Lu et al. 2018). These features confer CTD the control of Pol II clustering and mobility, directing it to the right droplet (enhancer or splicing droplet) during transcription stages, according to its phosphorylation state (Guo et al. 2019).

1.2.4 Transcription termination

Efficient transcription termination is important to prevent transcribing Pol II from interfering with DNA elements downstream the transcription termination site (TTS) and to ensure that Pol II is recycled for subsequent rounds of transcription (Kuehner et al. 2011). This last step of the process is tightly linked to processing of Pol II transcripts and once again, Pol II CTD proves to be paramount in regulating mRNA processing and transcription termination events

(Buratowski, 2009; Eick & Geyer, 2013; Harlen & Churchman, 2017). Nascent RNA processing occurs via three different pathways, which are gene class specific: the polyadenylation (PolyA) dependent pathway (Figure 5), used for mRNAs; the Nrd1-Sen1-Nab3 (NNS) pathway and the Integrator-dependent pathway, used for small non-coding RNAs (Eick and Geyer 2013).

In the PolyA-dependent pathway, Pol II pauses downstream the polyadenylation site (PAS) AAUAAA to allow cleavage of the nascent transcript (Buratowski 2009). Therefore the product of the cleavage is adenylated upstream the cleavage site, while the downstream product is degraded (Kuehner et al. 2011). In mammals the factors involved are human cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulatory factor (CstF). CPSF recognises the PAS and induces Pol II pausing. Then, CPSF associates with CstF (CPSFCF) stimulating CPSF mediated cleavage and release (Kuehner et al. 2011). In metazoans, CPSF processing endonuclease CPSF73 cleavage at protein coding genes is required for termination (Eaton et al. 2018). To be efficiently released, Pol II requires the action of XRN2 (Rat1 in yeast) 3'-5' exonuclease. Human Senataxin (SETX), the mammalian homolog of Sen1, contributes to PolyA dependent termination by resolving DNA:RNA hybrid structures formed during termination and promoting XRN2 function (Figure 5) (Skourti-Stathaki et al., 2011; Mayer et al., 2018). The collision between the nuclease and Pol II would promote termination (Kuehner et al. 2011). This is also known as the "Torpedo" model (Buratowski 2009). On the contrary, the allosteric model proposes that EC undergoes conformational changes that result in displacing the EFs and decreased Pol II processivity (Mayer et al. 2018). Recently, it has been agreed on a unified model of the allosteric and torpedo models, in which termination is dependent on XRN2 (Eaton et al. 2020; Rodríguez-Molina et al. 2023).

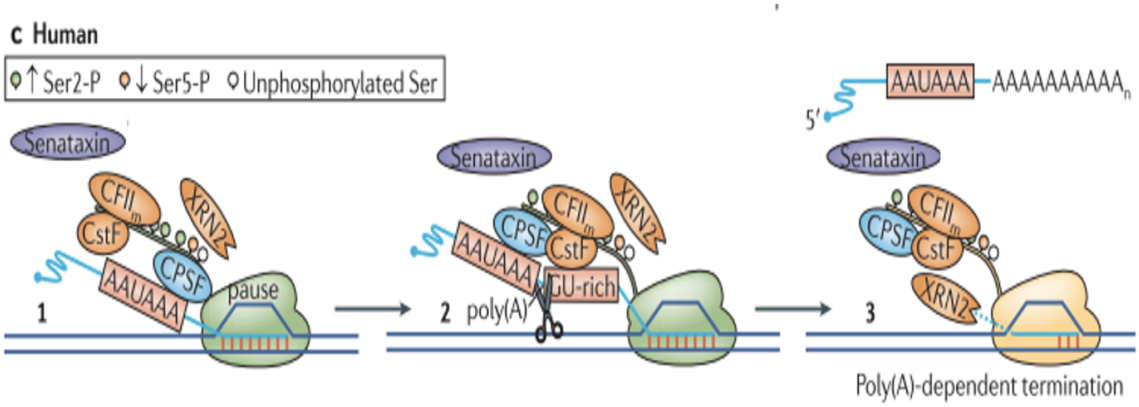


Figure 5 PolyA dependent Transcription termination is dictated by Pol II pausing downstream the PAS and the concerted action of termination factors, which cleave and polyadenylate the nascent transcript and resolve the DNA:RNA hybrid thus displacing Pol II from the double helix. Adapted figure from Kuehner et al., 2011.

1.3 RNA Polymerase complexes and cancer

RNA polymerase complexes work in concert with transcription activators and repressors. Sequence-specific DNA binding transcription factors (TFs) bind regulatory elements and interact via protein-protein interaction with the polymerases, in order to regulate the assembly of the transcription machinery and ensure the smooth course of transcription (Martinez-Rucobo et al., 2015; Plaschka et al., 2015; Ehara et al., 2017; Bernecky et al., 2017; Soutourina, 2018). The activity of TFs is controlled as well by coactivators and corepressors, that are multiprotein complexes with multiple enzymatic functions. Coactivators and corepressors actions can be summarised in two categories: i) bridging sequence-specific DNA binding TFs with the general transcription machinery, like for example the case of the Mediator complex (Allen and Taatjes 2015); ii) altering chromatin accessibility, as for example through nucleosome remodeling by the ATP-dependent complex hSWI/SNF, or post-

translational modifications of histone tails like acetylation by histone acetyl transferases (HATs) associated with transcription activation, deacetylation by histone deacetylases (HDACs) associated with repression, or methylation associated with both activation or repression depending on the site and the complex responsible (Näär et al. 2001; Peterson 2002; Rosenfeld et al. 2006; Spiegelman and Heinrich 2004).

In 2000 a seminal review work by Hanahan and Weinberg defined the six hallmarks of cancer as alterations in the physiology of the cell that represent novel capabilities acquired by cancer cells during tumorigenesis: self-sufficiency in growth signals, insensitivity to antigrowth signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis and tissue invasion and metastasis (Hanahan and Weinberg 2000). In the subsequent years, also genome instability, among others, emerged as new hallmark of cancer together with deregulated transcription (Negrini et al., 2010; Hanahan & Weinberg, 2011). Because of their role as major regulators of gene expression, many transcription factors, activators and repressors, are deregulated (altered gene expression) or mutated in cancers and are well-established oncogenes/tumour suppressor genes (Bywater et al. 2013). Among them, worth to mention the retinoblastoma (RB) tumour suppressor gene and its binding partner E2F1 (and other E2F genes). RB binds E2F in its phosphorylated form (pRB) and represses E2F transcriptional activity. Indeed, RB is phosphorylated by cell cycle dependent kinases (CDKs) and in this way its action and its repression on E2F is controlled through the cell cycle, tightly regulated to control cell proliferation. Cancers commonly bear a disruption of this pRB-E2F regulation (Johnson, 2000; Manickavinayaham et al., 2020); *TP53*, commonly referred to as the guardian of the genome, is a tumour suppressor gene involved in response to stress signals including replication stress, metabolic stress, and DNA damage. It promotes the transcription of genes

involved in cell stress response and it is a central node in cell cycle control and arrest, cell senescence and apoptosis, found to be mutated/deleted across a wide variety of cancer types (Hernández Borrero and El-Deiry 2021). Oncogenes of the *MYC* family (C-L-N- Myc), which regulate about 10% of the expressed genes and also some Pol I and Pol III transcripts, are usually upregulated in a wide variety of tumours (Liu et al., 2016; Tsantoulis & Gorgoulis, 2005; Kent & Leone, 2019; Muller & Vousden, 2013; Strano et al., 2007; Cole & Cowling, 2008; Dang et al., 1999; Bradner et al., 2017). Beyond deregulations of TFs, there are evidences of transcription deregulations also of RNA Pol I and Pol III complexes in cancers (Figure 6) (White, 2004; Bywater et al., 2012; Bywater et al., 2013), while specific Pol II deregulations in relation to cancer remain elusive. Moreover, although the transcription regulation mediated by TFs occurs via their interaction with the subunits composing the RNA polymerase complexes (Martinez-Rucobo et al., 2015; Plaschka et al., 2015; Ehara et al., 2017; Bernecky et al., 2017), studies on the subunits of the RNA Polymerase II machinery itself remain elusive.(Martinez-Rucobo et al., 2015; Plaschka et al., 2015; Ehara et al., 2017; Bernecky et al., 2017), studies on the subunits of the RNA Polymerase II machinery itself remain elusive.

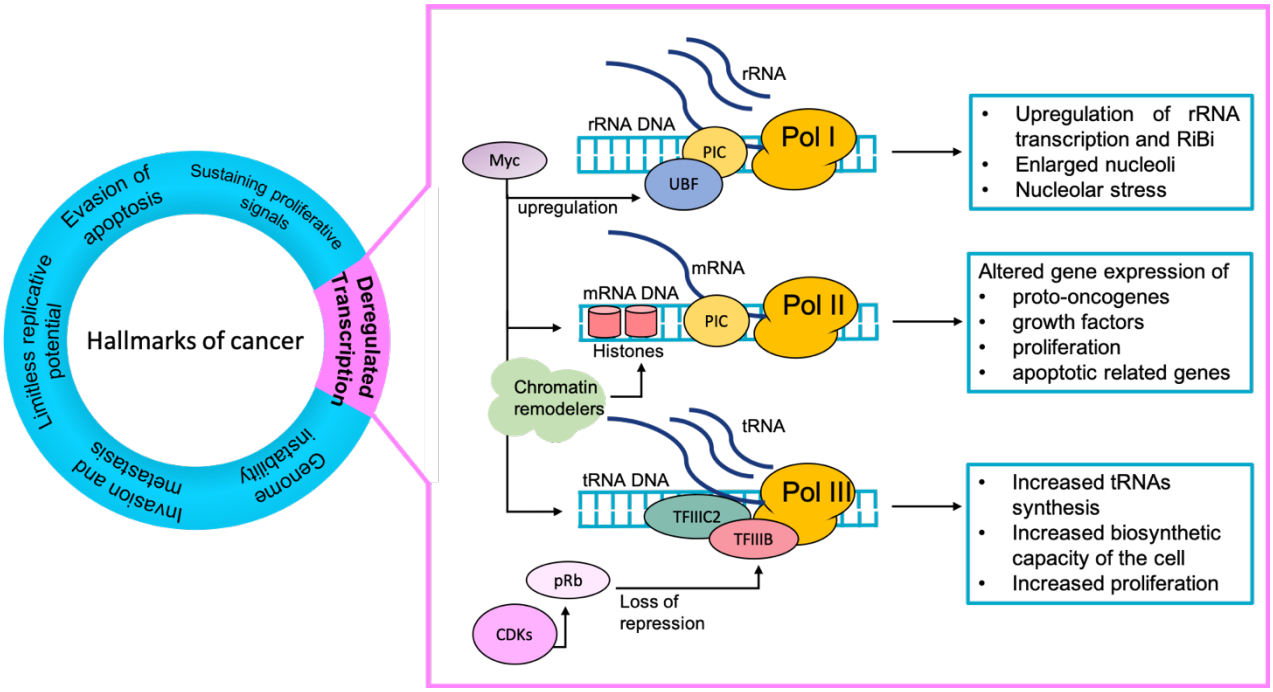


Figure 6 Deregulated transcription is a hallmark of cancer: Transcription factors (TFs), which can be both oncogenes and tumour suppressor genes, and transcription complexes are deregulated in cancer. Pol I rRNA transcription and Ribosomal Biosynthesis (RiBi) can be upregulated in cancer cells. The TF and oncogene Myc upregulates UBF, a component of Pol I transcription machinery, thus increasing Pol I transcription. TFs and chromatin remodellers deregulate the Pol II genes transcription and result in altered gene expression. Cancer cells' increased demand of cellular components to sustain their high proliferation need is satisfied by increased Pol III transcription and increased tRNAs synthesis. TFIIIC2 and TFIIIB upregulations also increase transcription rates. The tumour suppressor gene Rb, when phosphorylated by CDKs (pRb), is no longer able to exercise its repressive function on TFIIIB and results in TFIIIB derepression and increased Pol III transcription. UBF = Upstream Binding Factor; PIC = preinitiation complex; CDKs = cyclin-dependent kinases; Rb = retinoblastoma tumour suppressor gene; pRb = phospho-Rb. Adapted figure from Muste Sadurni et al. 2023

1.3.1 Pol I

Hyperactivation of ribosomal DNA (rDNA) transcription and enlarged nucleoli (nucleolar hypertrophy) are common features of aggressive tumours, together with deregulation of ribosome biosynthesis (RiBi) and overexpression of rRNA (Derenzini et al. 2009). These observations led to suggest that Pol I could be directly targeted in cancer therapy. Indeed, in recent years Pol I inhibitors have been developed and tested in a wide range of tumours. Additionally, many oncogenes and oncoproteins such as *AKT*, PI3K and Ras, play important roles in ribosome biogenesis, making the activation of the ribosomal surveillance pathway and consequent p53 activation a useful therapeutic mean for a wide range of tumours (White, 2004; Bywater et al., 2012; Khot et al., 2019). Furthermore, well-known transcription factors which are tumour suppressor genes or oncogenes, for example *RB* and *MYC,* regulate transcription of Upstream Binding Factor (UBF), an essential component of Pol I pre-initiation complex, resulting in repression or upregulation of Pol I transcription respectively (White, 2004; Drygin et al., 2010; Poortinga et al., 2015). Myc stimulates the expression of UBF upregulating ribosome biosynthesis, thus sensitizing cells to nucleolar stress. Therefore, it has been proposed that Pol I transcription in *MYC* driven tumour cells could be a promising target, as inhibition of rDNA transcription could disrupt the ability of Myc to suppress p53 accumulation in tumour cells more than in normal cells (Poortinga et al. 2015). Among the Pol I inhibitors, CX-5461 prevents the formation of Pol I pre-initiation complex and inhibits transcription initiation, by impeding Pol I release from the promoter, and it's very specific to Pol I and does not affect Pol II or III (Poortinga et al., 2015; Mars et al., 2020). This small-molecule inhibitor has shown promising results in killing Eμ-MYC lymphoma cells, by activating p53 dependent apoptosis (Khot et al. 2019). Interestingly, CX-5461 has been successful also in

the treatment of chemoresistant ovarian cancers, showing promising results in taxane resistant cells, which are instead sensitive to CX-5461 (Cornelison et al., 2017 and citations therein; Sanij et al., 2020). Another inhibitor, CX-3543, binds quadruplex DNA and inhibits Pol I transcription by inhibiting nucleolin-G-quadruplex complex formation, which leads to the translocation of nucleolin into the nucleoplasm and triggering apoptosis (Drygin et al., 2009). BMH-21 is a potent compound capable of inducing p53 activation selectively in different cancer cell lines but not in normal cells, works by binding DNA GC-rich regions. Being rDNA highly GC-rich, it was shown that BMH-21 functions by disassembling Pol I complex at rDNA promoter regions to inhibit rDNA transcription (Ferreira et al. 2020). Recently, it was also found that Pol I activity is upregulated in prostate cancers. CX-5461 blocked proliferation and suppressed prostate cancer tumourigenesis in a p53 dependent manner (Low et al. 2019). Similarly, BMH-21 decreased prostate cancer cell growth *in vitro* and *in vivo* in different cell lines, also in a p53 independent manner (Low et al. 2019).

Deregulations affecting specifically some of the subunits of Pol I complex are also reported to have a role in cancer. POLR1A is overexpressed in tumoral and metastatic cell lines of colorectal cancer (Sanchez-Martin et al. 2021). Similarly, POLR1B is usually upregulated in a broad range of human cancers including in non-small cells lung cancer (NSCLC), where knockdown of POLR1B resulted in reduced cell proliferation (F. Yang et al. 2020). Furthermore, in ER-positive breast cancer it can be considered a biomarker for resistance to fulvestrant (Cheng et al. 2020). The POLR1C subunit is differentially expressed in breast cancer patients compared to normal tissue and either amplified or overexpressed in gastric cancer (Pinto et al. 2016). In colorectal cancer a region in 13q containing *POLR1D* is frequently amplified resulting in POLR1D overexpression (Sheffer et al. 2009) which is associated with higher risk,

poorer overall survival and acquired resistance to bevacizumab, identifying POLR1D as potential therapeutic target (Wang et al., 2019; Zhang et al., 2020; Tian et al., 2020; Zhou et al., 2020). Finally, *POLR1E* expression correlates with increased tumor stage in bladder cancer (Ghafouri-Fard et al. 2014).

1.3.2 Pol III

Similarly to Pol I transcription, also Pol III transcription can contribute to carcinogenesis. Indeed, it was found both in ovarian and breast cancer cells that Pol III transcription can lead to the aberrant production of its transcripts that contribute to cell growth or to the overproduction of tRNAs (Pavon-Eternod et al., 2009; Winter et al., 2000). The accelerated proliferation of cancer cells necessitates an increase in the synthesis of cellular components to match their heightened rate of division. To achieve the necessary boost in biosynthetic capacity for rapid proliferation, the output of Pol III is enhanced (Gjidoda and Henry 2013). Hyper activation of Pol III and components of Pol III transcription such as TFIIIB complex was observed in different cancer types and in mice with myeloma compared to tumour free mice (reviewed in Cabarcas & Schramm, 2011), and TFIIIC2 was found to be overexpressed in ovarian carcinomas (Winter et al., 2000). It was also found that Pol III activity is elevated in primary human fibroblasts carrying loss of function mutations in p53 (Stein et al. 2002) as well as in mice lacking Rb (White et al., 1996; Scott et al., 2001). Indeed, many known mutations in *RB* lay in a region involved in binding TFIIIB and in this way repress Pol III transcription. When this binding is compromised, usually by hyperphosphorylation of RB by CDK-E/D, TFIIIB is de-repressed and further activates Pol III. Hyperactivation of Pol III has also been associated with c-MYC upregulation (White, 2004 and citations therein). This supports the belief that limiting

Pol III activity can impede robust cancer proliferation, as demonstrated by studies in which reduced Pol III activity in mouse xenograft models inhibited tumour formation. Pol III and Pol III-associated transcription factors could potentially serve as valuable biomarkers for cancer diagnosis (Gjidoda & Henry, 2013; Liang et al., 2019). More recently, it has been found that the subunit POLR3G is capable of binding the telomerase reverse transcriptase TERT (Khattar et al. 2016), a central protein in stimulating the proliferative capabilities of many cancer cell types. In support and extension of this finding, it was shown that *POLR3G* inhibition resulted in prostate cancer cell-specific proliferation arrest and cell death (Petrie et al. 2019) and that POLR3G overexpression correlated with poor prognosis of transitional cell carcinoma (Liu et al. 2020). *POLR3G* was also identified in clinical data as one out of five genes in a risk signature predicting the prognosis of patients with hepatocellular carcinoma (Dai et al. 2021). More recently, Lautré and colleagues showed that POLR3G is the only component of the Pol III transcription apparatus significantly overexpressed in triple-negative breast cancer (TNBC) but not in other types of breast cancer. Suppression of POLR3G in the MDA-MB231 TNBC cell line decreases colony formation and invasive growth in vitro. Importantly, *POLR3G* KO impairs tumour growth and metastasis of intraductal xenografts in mice (Lautré et al. 2022). In conclusion, POLR3G expression plays a significant role in regulating tumour formation and metastasis in TNBC, with its deletion leading to altered gene expression patterns and cell fate decisions. Further research will determine whether targeting POLR3G could emerge as a potential therapeutic strategy for TNBC. Other Pol III subunits have been linked to cancer like mutations in *POLR3B* implicated in lung cancer, with the gene found to be differentially methylated, making *POLR3B* a susceptibility and risk gene in lung cancer (Asiedu et al. 2018; Luo, Wang, and Zhang 2018; Musolf et al. 2020; Yunlei et al. 2013). Furthermore, POLR3K

subunit high expression correlated with poorer overall and disease free survival in ovarian cancer (An and Duan 2022).

1.3.3 Pol II

Cancer cells rely on transcription to sustain their growth and survival. Therefore, blocking RNA synthesis via RNA Polymerase inhibitors like DRB or flavopiridol, or nucleotide analogues such as 3'-ethynylcytidine (Ecyd), is an effective strategy both *in vitro* and *in vivo* (Fuda et al., 2009; Turinetto et al., 2009*;* Fukushima et al., 2014*;* Galbraith et al., 2019). Definitely, it is possible to downregulate many oncogenes by reducing overall transcription levels. So far, little is known about Pol II subunits implications in human diseases with a particular focus on cancer. However, recent studies provide evidence of emerging roles of some Pol II subunits. All the deregulations are summarized in Supplementary table 1.

1.3.3.1 POLR2A/RPB1

POLR2A subunit, being at the core of the catalytic Pol II active site, is considered an essential transcription gene and oncogene, being able to promote rapid cell growth and repress apoptosis in tumours (Serra et al., 2019; Jiang et al., 2021). In fact, usually *POLR2A* is highly expressed in tumour tissues compared to neighbouring normal tissues. For example, *POLR2A* overexpression has been found to induce proliferation in gastric cancer cells and inhibit apoptosis both *in vitro* and *in vivo* (Jiang et al. 2021). Furthermore, *POLR2A* expression level was found to be higher in patients derived xenograft models of ovarian cancer (Li et al. 2019). Interestingly, POLR2A high protein levels were found in AML cell lines compared to normal progenitor cells. Accordingly, mRNA expression levels were also higher in primary AML patients, resulted in reduced overall survival and increased tumour growth. Indeed, *POLR2A*

KD reduced cell proliferation and tumour growth *in vitro* and *in vivo* (Yu et al. 2019). Analysing *POLR2A* mRNA expression levels in TCGA dataset, it was also found that *POLR2A* upregulation was occurring not only in AML patients, but also in cholangiocarcinoma and thymoma (Yu et al. 2019). On the contrary, POLR2A was downregulated in adrenocortical carcinoma, testicular germ cell tumours and uterine corpus endometrial carcinoma. Additionally, polymorphisms in *POLR2A* were associated with poor outcomes in non-small cell lung cancer (Yoo et al. 2017). The *POLR2A gene* lays in a neighbouring region of *TP53* on Chr 17 and is often co-deleted together with it as a "passenger gene". In fact, in colorectal cancer *POLR2A* is commonly hemizygously deleted together with *TP53,* rendering cancer cells more sensitive to the transcription elongation inhibitor α-amanitin (Liu et al., 2018) and *POLR2A* depletion with siRNAs, suggesting the possibility of targeting POLR2A in p53-/+ cancer cells or cancers with a more broader 17p loss (Liu et al., 2015; Cance & Bradner, 2015; Li et al., 2021; Van Der Jeught et al., 2018). Similarly, siRNA inhibition of *POLR2A* in hemizygous TP53 TNBCs where *POLR2A* is hemizygously co-deleted, resulted in reduced tumour growth and enhanced tumour suppression (Xu et al. 2019).

1.3.3.2 POLR2B/RPB2

POLR2B is the second largest subunit of Pol II and together with POLR2A forms the catalytic site of the complex. There are not many studies regarding POLR2B deregulations in cancer. However, it has been shown that mutations in *POLR2B* are linked to sensitivity to JQ1, a bromodomain inhibitor targeting BET proteins BRD2-3-4, in breast cancer cell lines. JQ1 sensitivity was assessed in over 1000 cancer cell lines to identify possible markers of sensitivity and *POLR2B* mutations were identified in breast cancer cell lines as predictors of JQ1 sensitivity (Gerlach et al., 2018; Iorio et al., 2016). Furthermore, *POLR2B* downregulation could

play a role in cell resistance to 5-Azacytidine (AZA), a DNA demethylating agents widely used in cancer treatment and more specifically in treatment of patients with AML, with a moderately good response rate. Indeed*, POLR2B* mRNA expression level was found to be significantly decreased in AZA resistant cell lines compared to non-resistant ones (Imanishi et al. 2014). More recent studies analysing TCGA datasets, found *POLR2B* upregulated in hepatocellular carcinoma and high mRNA levels correlated with overall reduced patients' survival rates and were associated with tumour staging. Consistently, high POLR2B protein expression levels were observed in the cancer tissue compared to normal liver tissue. Additionally, POLR2B expression was even higher in virus (HBV or HCV) driven hepatocellular carcinoma. Therefore, *POLR2B* gene has a possible value for further development as diagnostic biomarker in virus-related HCC (Yong-Zheng Zhang 2022).

1.3.3.3 POLR2C/RPB3

As for POLR2B, not much research highlights the role of POLR2C deregulations in general in cancer. However, few studies conducted in hepatocellular carcinoma (HCC) cell lines and tissues found that POLR2C protein expression is increased by immunohistochemical staining in tissue samples of different stages compared to normal tissue, and higher expression (stronger staining) was progressively observed from low to high grade HCC. Furthermore, POLR2C overexpression in patients correlated with poorer overall and disease-free survival. HCC cell lines overexpressing POLR2C had also enhanced proliferation and migration capabilities which translated into enhanced tumour growth in injected mice. POLR2C overexpression upregulated a series of mesenchymal markers, such as N-cadherin, and downregulated epithelial markers, such as E-cadherin, thus promoting epithelial to mesenchymal transition (EMT). Interestingly, POLR2C was found to be a regulator of E-

cadherin expression, through its interaction with Snail, a key regulator of E-cadherin transcription, via its N-terminal domain. Intriguingly, Pol II activity and the expression of the other subunits of the complex remained unchanged. Another mechanism through which POLR2C regulates HCC cell proliferation is due to its regulation of VOPP1 expression, which is upregulated in different cancers and promotes cell proliferation and migration while inhibiting apoptosis (Fang et al., 2014; Hu et al., 2021). Other than in HCC, *POLR2C* expression alterations were found also in osteosarcoma tumour samples, where mRNA expression is significantly lower compared to normal osteoblasts (Chen et al. 2022), and in gastric cancer. In the latter, microarray analysis confirmed from TCGA analysis of differentially expressed genes between gastric cancer patients showing cisplatin resistance and patients showing drug sensitivity, found *POLR2C* to be upregulated together with two other Pol II subunits, *POLR2L* and *POLR2F*, thus indicating that these subunit upregulations could be involved in mechanisms of cisplatin resistance in gastric cancer (Zhou et al. 2018).

1.3.3.4 POLR2D/RPB4

Some research has been conducted on POLR2D functions in tumour development. First evidence comes from colorectal cancer, where it was found that *POLR2D* mRNA expression levels significantly correlated with increased expression of the CA 19-9 tumour marker in patients, indicating that this gene could be used to evaluate disease state (Abba et al. 2012). *POLR2D* was also found to be upregulated in one sample of single cell RNA-Seq analysis performed on malignant ascites cells from ovarian cancer patients (Kim et al. 2021). The upregulation of POLR2D is also important in prostate cancer, where analysing patients on TCGA database it was shown that high *POLR2D* expression associated with significantly lower disease-free survival compared to patients with low expression of *POLR2D* (Yamada et al.

2018). Consistent with these observations, the downregulation of *POLR2D* is considered to be a protective factor in highly grade serous ovarian cancer, being preferentially downregulated in tumour samples compared to normal ones and associated with better overall survival (Dai et al. 2022).

1.3.3.5 POLR2E/RPB5

So far upregulation and overexpression have been identified as the main deregulations of Pol II subunits in a wide variety of tumour types. However, the alteration linking POLR2E subunit and cancer is not related to a differential expression of the gene, but it is a well characterized polymorphism. It was firstly identified in 2011 by Jin and colleagues that reported in a prostate cancer genome wide association study, the SNP rs3787016 in *POLR2E* associated with increased risk of prostate cancer occurrence (Jin et al. 2011). They proposed that this variant could interfere with correct splicing and therefore produce a dysfunctional POLR2E, thus impacting on Pol II transcription. More recently, these results were further confirmed in a sample of prostate cancer in an Iranian population, and additionally a new variant of *POLR2E*, rs1046040, was found to be a predisposition factor for cancer development (Sattarifard et al. 2019). Interestingly, other studies have shown how this polymorphism located on the fourth intron of *POLR2E* gene, was also significantly associated with oesophageal cancer, breast cancer, papillary thyroid carcinoma and liver cancer (Chen et al. 2018; Chen et al. 2018; Chen et al. 2019; Chu et al. 2017)

1.3.3.6 POLR2F/RPB6

POLR2F encodes the sixth largest subunit of Pol II complex, and it is also one of the shared subunits among the three polymerase complexes (Pol I, Pol II, Pol III). Early studies showed that *POLR2F*, together with two other genes, was significantly overexpressed in colorectal

carcinoma tissues compared to normal tissues, and specifically its overexpression correlated with early disease occurrence and relapse (Antonacopoulou et al. 2008). Previous studies showed *POLR2F* to be upregulated in a colorectal cancer metastatic tumour cell line (Orian-Rousseau et al. 2005). As previously mentioned for the POLR2C subunit, TCGA analysis of differentially expressed genes comparing gastric cancer patients with resistance to cisplatin towards patients showing drug sensitivity, found POLR2F to be upregulated, indicating that these subunit upregulations could be involved in mechanisms of cisplatin resistance in gastric cancer (Zhou et al. 2018). Furthermore, an analysis of microarray datasets TNBC and non-TNBC identified 1075 differentially expressed genes, among which *POLR2F* was found to be upregulated significantly in TNBC and identified as a potential cancer causing gene (Naorem et al. 2019). Consistent with these data, a more recent study focused on identifying genes differentially expressed between malignant epithelial cells and non-malignant epithelial cells in TNBC, identified *POLR2F* as an upregulated gene in the malignant cells population (Yi et al. 2023). Lastly, a retrospective study on prostate cancer integrating transcriptome and clinical and pathological data from different databases identified *POLR2F* upregulation associated with poorer prognosis in patients and predictive of worst outcome in androgen deprivation therapy treatment (Wang et al. 2023). Not only the upregulation/overexpression of POLR2F has a potential prognostic value in tumours, but there is evidence that also downregulation of POLR2F could be relevant. It was found that *POLR2F* expression was significantly lower in glioblastoma tumour tissues compared to normal tissues and that this impacted on overall survival (Yang et al. 2020). However, *in silico* analysis to identify genes important for tumour progression and survival identified *POLR2F* as a survival gene in glioblastoma, and specifically in tumours with IDH1 mutations POLR2F appears to be significantly overexpressed (Masica

and Karchin 2011). Moreover, it was also found that lower levels of the protein are associated with increased risk in colorectal cancer and *POLR2F* lower expression was found to be associated with a poorer prognosis in cervical cancer in patients positive for HPV18 (Sample 2020; Wang et al. 2022)

1.3.3.7 POLR2G/RPB7

Findings on POLR2G subunits are controversial. Li et al. conducted an analysis matching clinical data with mRNA expression in patients with hepatocellular carcinoma, leading to the identification of high-risk factors and protective factors. *POLR2G*, whose expression was upregulated in tumour tissues compared to adjacent normal tissues, was identified as a protective factor in the subset of patients studied, associated therefore with better prognosis and overall survival (Li et al. 2019). However, a more recent study combining single cell RNA-seq and TCGA RNA-seq data identified *POLR2G* (together with another subunit of Pol II, *POLR2L*) as a gene expressed at higher levels in HCC cancer stem cells, which are important for HCC heterogeneity and resistance to treatment. Importantly, the increased expression of mRNA was also associated with a progressive increase in tumour grade (from G1 to G4) and nodal metastasis status, which are associated with tumour progression and aggressiveness (Yao et al. 2022).

1.3.3.8 POLR2H/RPB8

POLR2H is another subunit common to all three RNA polymerase complexes. As for most of the other subunits, its upregulation/overexpression is particularly relevant in different tumours. First evidences show that high mRNA expression of *POLR2H* is found in HPV+ head and neck carcinomas (Slebos et al. 2006). More recent research in rectal cancer analyzing both the TCGA and GEO databases, identified among a total of 18 prognostic associated genes, a

subset of 11 genes including the subunits *POLR2H* and *POLR2J* associated with increased risk of poor overall survival in patients. Importantly, in the high risk group *POLR2H* was overexpressed (Zhang et al. 2020). In agreement with this analysis, both POLR2H and POLR2J are overexpressed in rectal tumour organoids (Normal et al. 2020). *POLR2H* high mRNA levels are also associated with patients overall survival in hepatocellular carcinoma (Ke et al. 2019). Several groups have similarly shown how *POLR2H* is upregulated in prostate cancer and overexpressed in prostate cancer cells compared to the normal tissue, indicating the potential role of POLR2H as a biomarker for prostate cancer prognosis and diagnosis, as well as a potential drug target (Chen and Hu 2019; Fan et al. 2018; Wang et al. 2021). Furthermore, *POLR2H* mRNA levels are significantly higher in lung cancer tissue samples compared to normal samples (Dong et al. 2021) and its high expression significantly correlates with patients survival in lung squamous cell carcinoma (Guttapadu et al. 2023; Jia et al. 2022). Interestingly, more recent data showed that *POLR2H* is upregulated in a subset of breast cancer patients resistant to radiotherapy, implicating its role potentially in the acquisition of radio-resistance (Jia et al. 2015; Miao et al. 2021)

1.3.3.9 POLR2I/RPB9

POLR2I is a subunit specific of Pol II, found to be either amplified or upregulated in a subset of ovarian cancer and head and neck cancer cells. It was speculated that POLR2I could contribute to mechanisms of cancer cells resistance to DNA damaging chemotherapeutics like cisplatin, 5-fluorouracil, PARP inhibitor olaparib and also radiation therapy (Gaponova et al. 2017). Furthermore, in a meta-analysis study of microarray datasets aimed to identify potential prognostic candidates in colorectal cancer cell lines with different levels of aggressiveness, *POLR2I* was found as a candidate gene downregulated in less aggressive cell lines (Long et al.

2016). At the same time, down-regulation of *POLR2I* expression was observed in hepatocellular carcinomas, where it correlated with mutations in the mRNA of other genes (FKH and *FOXP3*) which are not present in the genomic DNA, suggesting that *POLR2I* down-regulation could impair Pol II transcription fidelity and hence contribute to produce aberrant transcripts (Ren et al. 2020; Walmacq et al. 2009).

1.3.3.10 POLR2L/RPB10

POLR2L is one the Pol II subunits shared with the other polymerases, and as for most of the other Pol II subunits, its upregulation has been identified in different cancers. As previously mentioned regarding POLR2C subunit, after screening for thousands of genes and 300 miRNAs associated with cisplatin resistance in gastric cancer, six hub genes were identified and confirmed in TCGA databases. Among these genes, *POLR2L* together with *POLR2C*, *POLR2F* and *POLR2K* resulted significantly upregulated and associated with cisplatin resistance (Zhou et al. 2018). Furthermore, *POLR2L* was among the 7 candidate genes found upregulated to significantly promote cell proliferation compared to control cells, identified via CRISPR/Cas9 activation library screening in two hepatocellular carcinoma cell lines. Zhang and colleagues showed that cell survival is increased by the gain of function of POLR2L in both cell lines. Additionally, the upregulation of these genes promoted liver tumour growth and also its colonization potential with the formation of metastasis to the lung. Indeed, these cell lines showed increased invasion capability compared to the control (Zhang et al. 2023).

1.3.3.11 POLR2J/RPB11

There are different gene copies of *POLR2J (1-4)* on the human Chr 7 and at least 11 alternative mRNAs encoding different POLR2J isoforms. First evidence of a role of POLR2J in cancer came from expression studies in lung tumours where 50 cases were analysed, and nine genes

including POLR2J, were found to show a two-fold increase in expression compared with normal bronchial epithelial cells. POLR2J showed higher expression in 9 of the tumour samples, with 2 matched samples showing more than double expression levels in the tumours (Campbell et al. 2008). TCGA characterized colorectal cancer genome using 224 tumour samples with their corresponding normal tissues, to identify new biomarkers and potential therapeutic targets, found that *POLR2J* expression increased progressively with the aggressiveness of the tumour, and similar results were also found in liver metastatic tumours (Rajput et al. 2013). As mentioned above, *POLR2J* was associated with higher risk and poorer survival in rectal cancers and was found overexpressed in rectal tumour organoids (Normal et al. 2020; Zhang et al. 2020). Furthermore, Li and colleagues observed that POLR2J is upregulated, both at the mRNA and protein levels, in glioblastoma cells compared to normal tissues and resulted in enhanced cell proliferation and metastatic capability (Li 2023). It was already established that upregulation of *POLR2J* correlated with poor prognosis in glioblastoma patients and that its downregulation instead associated with better overall survival. It was shown that knockdown of *POLR2J* reduced cell proliferation and caused cell cycle arrest at G1/G0 phase in glioblastoma cells and additionally *POLR2J* knockdown was able to reduce the migrative and invasive abilities of cancer cells and to suppress epithelial to mesenchymal transition (EMT), indicative of invading cancer cells (Sánchez-Tilló et al. 2012). Consistently, the overexpression of *POLR2J* promoted the migrative and invasive abilities of cells, indicating that it could play a crucial role in EMT and increase metastatic potential in glioblastoma (Li 2023). Finally, POLR2J has also been found associated with patients survival of grade 2 ovarian cancers (Zhao et al. 2021), testicular germ cell tumour prognosis (Yao et al. 2021) and breast cancer susceptibility in a Belgian population (Farahmand et al. 2016).

1.3.3.12 POLR2K/RPB12

POLR2K is the last subunit shared among Pol I, II and III. Initial studies in liver cancer cell lines showed that siRNAs targeting POLR2K significantly reduced the viability of cells , indicating a potential role of the subunit in hepatocellular carcinoma (Woo et al. 2009). Similarly to other Pol II subunits, its increased expression is observed in different cancers. In HER2 breast cancers was observed POLR2K gain of expression by mRNA upregulation or gene amplification (Natrajan et al. 2010); more recent studies showed that low expression of *POLR2K* significantly prolonged patients survival and reduced cancer rates in breast cancer, while the high expression of *POLR2K* was associated with poor patients' survival (Yang and Wang 2021). Recently, along with other 4 genes *POLR2K* higher expression levels was found by machine learning approaches in a high-risk group of patients for breast cancer (Chen et al. 2023). Using mRNA expression profiling, *POLR2K* was firstly identified in 2016 as a differentially expressed gene in prostate cancer tissues compared to normal tissues and later it was evident that *POLR2K* differential expression had a significant prognostic value for patients overall survival (Guo et al. 2021; Kelly et al. 2016). Finally, in bladder cancer patients, high levels of *POLR2K* were associated with poor progression free and overall survival. Furthermore, the knock down of the Pol II subunit reduced the viability of two bladder cancer cell lines, indicating that its decreased expression could slow down cell growth (Yang et al. 2020).

1.4 DNA damage

As previously discussed, transcription is a tightly regulated process, indeed when deregulated transcription can be causative of transcription-associated mutations and transcription-associated recombination events and generally speaking DNA damage formation. Altogether,

these phenomena can be grouped under the general term of transcription-associated genome instability (Gaillard and Aguilera 2016; Kim and Jinks-Robertson 2012), which will be later discussed in more detail.

Each cell present in the human body is threatened by more than 10 thousand DNA lesions per day, which are in majority (75%) represented by single strand DNA breaks (SSBs). Both endogenous and exogenous sources are responsible for modifying DNA and therefore genetic information (Lindahl and Barnes 2000; Tubbs and Nussenzweig 2017).

DNA mismatches and DNA strand breaks, derived from endogenous sources, can be caused for example by errors during DNA replication, replication stress, transcription stress (Branzei and Foiani 2009; Hernández Borrero and El-Deiry 2021; Jackson and Bartek 2009; Saponaro et al. 2014; Shah et al. 2021; Zeman and Cimprich 2014; Zou et al. 2021). Other endogenous DNA lesions causes comprise hydrolytic reactions, non-enzymatic methylation, reactive oxygen species (ROS), that are produced at sites of infection by macrophages and neutrophils, as well as by-products of oxidative respiration (Ciccia and Elledge 2010; Hoeijmakers 2009; Jackson and Bartek 2009). Of these, a particularly dangerous lesion is represented by SSBs, since if not properly repaired they can lead to genome rearrangements, gain or loss of repetitive sequences and mutations, overall increasing genome instability (Caldecott 2024). However, DNA double strand breaks (DSBs) are the most cytotoxic form of DNA lesions as their inaccurate repair influences dramatically the cell survival and fate (Jackson and Bartek, 2009; Tubbs and Nussenzweig 2017). Whenever DNA damage remains unchallenged, essential cell functions are impaired and cause alterations that can lead to tumour development or cell death, in the worst scenario (Tubbs and Nussenzweig 2017). Indeed, failure of DSBs repair leads to substantial genomic rearrangements, altering the balance between proto-oncogene

and tumour-suppressor genes, which can culminate into tumour development and deregulation of cell cycle. One possibility of organism self-protection is to trigger the programmed cell death once detected the presence of unrepaired lesions, in order to avoid spreading the damage during cell division (Aparicio et al. 2014; Frit et al. 2014). To counteract the above mentioned DNA lesions with the final goal of maintaining the intact genetic information to be passed to the next generation, all organisms have evolved a wide repertoire of cellular responses, generally termed as DNA damage response (DDR) (Blackford and Jackson 2017; Ghosal et al. 2013; Mladenov et al. 2016a).

1.4.1 DNA damage sensing

One of the prerequisites for the maintenance of genomic stability and prevention of tumour burden is the fast identification and efficient repair of DNA damage. DSBs are quickly detected by many different proteins, which can be referred to as DSBs "sensor" proteins, which redirect signalling and repair towards either HR or NHEJ. DSBs sensor proteins comprehend: Ku heterodimer (Ku70 and Ku80) that directs repair towards NHEJ; the MRN protein complex (Mre11-Rad50-Nbs1) that favours HR by promoting end resection; polyADPribose polymerases 1 and 2 (PARP1 and PARP2) that recognize both single and double strand breaks and recruit further repair proteins. Additional role of PARP1 may lie in promoting of alternative NHEJ (Brown and Jackson 2015). Once the damage is recognised, signalling pathways aimed to cell cycle regulation are initiated. This is important for slowing down or blocking cell progression in various cell cycle phases, thus allowing DNA repair machinery proteins to properly repair the damage (Liu et al. 2014; Weterings and Chen 2008). If DSBs occur in G1 phase, checkpoint activation stops cell cycle progression into S phase. Likewise,

damage occurring between late S/G2 phase triggers cell cycle checkpoints preventing cells from entering in mitosis (Cannan and Pederson 2016).

Upon DNA damage recognition, three members of phosphatidylinositol 3-kinase related kinases (PIKKs) family are recruited to DNA break sites where they execute the proper DSBR (Figure 7) (Blackford and Jackson 2017). Among six members of the PIKKs, which act in regulating several cellular processes such as cell cycle control, metabolism, transcription, replication, the three main PIKKs acting in DNA damage response are: ATM (ataxia-telangiectasia mutated), ATR (ATM- and Rad3-Related) and DNA-PKcs (DNA dependent protein kinase catalytic subunit) (Parrilla-Castellar et al., 2004). They are recruited by the above mentioned "sensor" proteins Nbs1 (of the MRN complex), ATRIP (ATR interacting protein) and Ku70/80 heterodimer, respectively (Falck, Coates, and Jackson 2005; Gell and Jackson 1999; Singleton et al. 1999; Xu, Bradley, and He 2023; Zou and Elledge 2003)

ATM, together with ATR and DNA-PK (DNA-PKcs and Ku70/80), phosphorylate the histone variant H2AX in a redundant manner, giving rise to gamma-H2AX (γH2AX) (Stiff et al. 2004). In fact, γH2AX is considered a common DNA damage marker. After its phosphorylation, it is then bound by MDC1 (mediator of DNA damage checkpoint1) (Stucki et al. 2005). MDC1 functions in promoting further recruitment of MRN complex and ATM, thereby spreading H2AX phosphorylation on chromatin and amplifying DNA damage response signal and repair (Blackford and Jackson 2017; Brown and Jackson 2015).

On the other hand, ssDNA breaks recruit ATR: Replication protein A (RPA) recognises ssDNA and recruits ATR in association with its obligate partner ATRIP (Zou and Elledge 2003). Finally, ATR activation is dependent on the co-localization of ATR-ATRIP within the 9-1-1 (Rad9-Rad1-Hus1) complex. Since the complex has three different subunits, it can interact with different

proteins at the same time, representing a versatile DNA damage and repair protein. Indeed, the domain in the C-terminal tail of Rad9 protein is responsible for assembling of TOPBP1 (topoisomerase binding partner 1), a crucial ATR activator, to ATR-ATRIP complex (Cimprich and Cortez 2008; Doré et al. 2009). Often the three PIKKs central to DDR work in concert. One example is given by the ATM-dependent ATR activation, since ATM-dependent end resection generates the RPA-ssDNA substrate for ATR-ATRIP recruitment. Furthermore, both ATM and ATR function in phosphorylating DNA-PK, and DNA-PK itself seems to phosphorylate ATM (Blackford and Jackson 2017; Zhou et al. 2017).

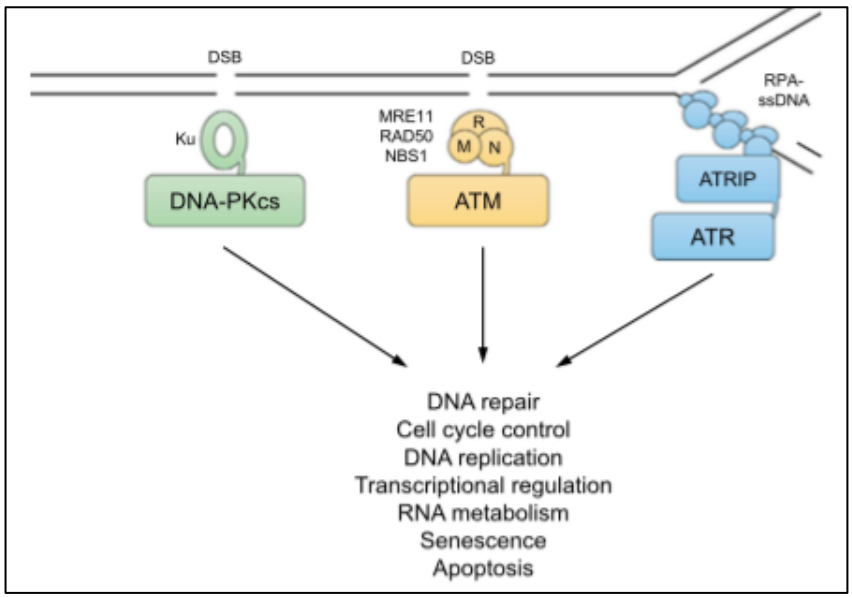


Figure 7 DNA damage sensing The three DDR signalling PIKKs and their recruitment factors at DSBs (Blackford and Jackson, 2017)

Another apical factor in DDR is the tumour suppressor p53 binding protein (53BP1), which is recruited to DSBs and therefore is widely recognized as a useful marker of DNA damage together with γH2AX (Bártová et al. 2019). It is recruited to chromatin through histone

modifications signals, such as methylation and ubiquitylation (Huadong et al.,2011), where it pushes DNA repair towards NHEJ counteracting the action of the HR factor BRCA1, which directs DDR in favour of HR, recruiting downstream proteins(Bouwman et al. 2010; Chapman et al. 2012; Jackson and Durocher 2013; Mirza-Aghazadeh-Attari et al. 2019; Uckelmann and Sixma 2017)

1.4.2 DNA damage repair

Once DNA damage is sensed, it triggers the DNA damage response (DDR). DDR comprehends several pathways: MMR, mismatch repair; BER, base excision repair; NER, nucleotide excision repair; single strand break repair (SSBR) and double strand break repair (DSBR). The last two can be carried out by either homologous recombination (HR) or non-homologous end joining (NHEJ) (Ciccia and Elledge 2010; Kunkel 2015; Ranjha et al. 2018). The first one is considered to be an "error-free" repair mechanism, since it relies on the presence of the intact homologous sister chromatid, while the latter is considered to be an "error-prone" process, since it directly joins broken ends (Liu et al. 2014).

In other words, HR repair re-establishes molecular integrity while saving the exact genetic sequence and information by using the sister chromatid as template to copy. On the contrary, NHEJ repair is not necessarily faithful to the genetic sequence, thus genetic information might be lost or altered, even if molecular integrity is successfully restored (reviewed in Mladenov et al., 2016). In the S phase of cell cycle, the sister chromatid is physically close to the damaged one, thereby HR can occur. On the contrary, outside S/G2 phase, the NHEJ is the preferred option (Lieber 2011). In cells, there is a pathway choice between NHEJ and HR, even if side

pathways such as Alternative End Joining (A-EJ) are also important for damage repair (Figure 8) (Scully et al. 2019).

1.4.2.1 HR repair

HR repair starts with MRN, CtIP and BRCA1 recruitment to the damage site, end resection mediated by MRN complex and by nucleases together with the action of the BLM helicase (Mladenov et al. 2016; Yun and Hiom 2009). The 3' overhang, formed by end resection is then bound by ssDNA binding protein RPA. RPA-coated ssDNA becomes a substrate for the recruitment of the key recombinase in higher eukaryotes RAD51 (Liu et al., 2014; Mladenov et al., 2016). In the pre-synaptic phase of HR, ssDNA bound to RAD51 constitutes the pre-synaptic filament that is stretched and the search for the homologous filament may take place. Then, during the synaptic phase, RAD51 facilitates the establishment of a physical contact between the invading strand and the duplex DNA template, forming a D-loop structure of heteroduplex DNA. In the post-synaptic phase, RAD51 finally dissociates from the DNA filament, thus let the 3'OH overhang free to constitute a primer for DNA synthesis (Krejci et al. 2012). BRCA2 facilitates the loading of RAD51 on ssDNA by competing with RPA (Liu et al., 2014). HR repair can be executed by three different pathways: the first leading to cross-over (in meiotic cells); the second leading to a conservative and non-error prone repair, consisting in strand displacement and strand annealing (SDSA); the third one, occurring when SDSA fails, leads to error-prone repair. The latter, usually occurs after replication fork collapse, and consists in extending the migrating bubble and generating long tracts of ssDNA, which is highly mutagenic (Mladenov et al. 2016; Scully et al. 2019).

1.4.2.2 NHEJ repair

Of the two NHEJ low fidelity pathways, which are responsible for the majority of DSBs repair in human cells, the classical or canonical NHEJ (c-NHEJ) is the prevalent one (Iliakis et al. 2015). In fact, this pathway can operate throughout the whole cell cycle, since it does not require the presence of homologous sequences, unlike the HR. In a simplistic view, this pathway consists in the ligation of two broken ends through an error-prone mechanism, and can thus contribute to the origin of deletions or insertions (Liu et al., 2014). The pathway favours small sequence alterations at the break site and can also be responsible for chromosomal rearrangements (i.e. translocations). Word worthy, errors generated by c-NHEJ are at the basis of CRISPR/Cas9 gene knockout (reviewed in Mladenov et al., 2016). Following the DSB, DNA broken ends are bound by Ku heterodimer (Ku70/Ku80) which subsequently recruits DNA-PKcs to the site of damage (Aparicio et al. 2014; Iliakis et al. 2015; Lieber 2011; Walker et. al 2001). DNA-PK autophosphorylates itself and promotes end tethering and phosphorylation of other NHEJ downstream factors, such as Artemis (Deriano and Roth 2013; Graham et al. 2016; Iliakis et al. 2015). The nuclease Artemis, which is recruited to DSBs by DNA-PK and subsequently phosphorylated, is endowed with a 5'→3' dsDNA nuclease activity and in complex with DNA-PK mediates the processing of a wide variety of DNA ends, thus enabling the repair pathway to proceed towards ligation steps (Goodarzi et al. 2006; Lieber 2011). DNA-PK mediated phosphorylation events also trigger the recruitment of X family DNA polymerases at DSBs that function in further processing of DNA ends, prior to the ligation. Polλ and Polμ are responsible for filling the gaps before break sealing (McElhinny et al. 2005; Mladenov et al. 2016). The final step involves the joining of the processed DNA ends by LIG4 that works together with XRCC4 XLF and PAXX (Ahnesorg et al. 2006; Lieber 2011; Mladenov et al. 2016; Ochi et al. 2015;

Weterings and Chen 2008). In the Alternative pathway, the DNA break is processed by DNA ends resection by MRN complex and CtIP, prior gap filling and ligation involving Polθ, LIG1/3 and XRCC1 (Mladenov et al., 2016).

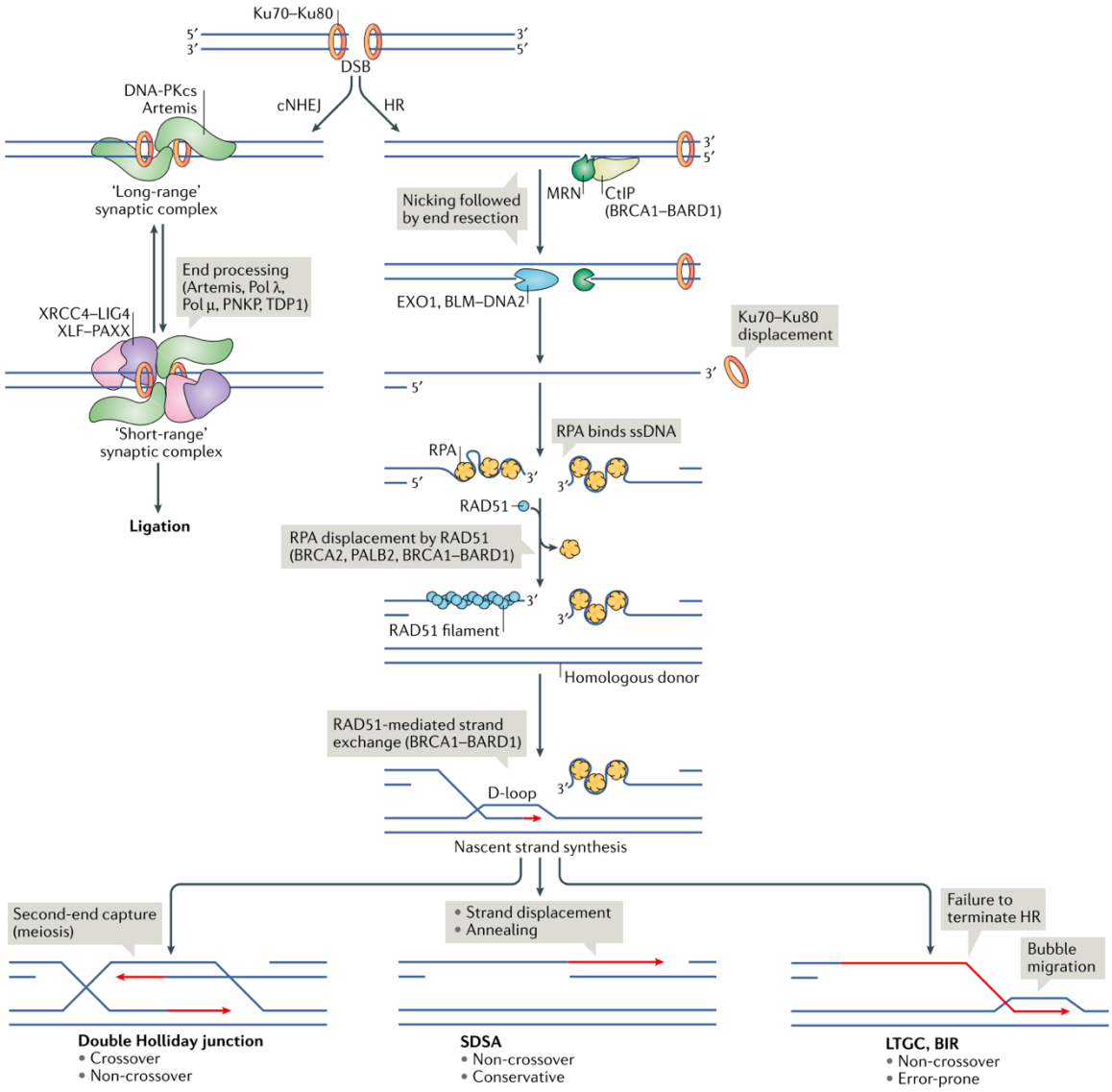


Figure 8 DNA damage response pathways Two main pathways ensure that DNA breaks are repaired, NHEJ and HR. The first is an error prone pathway, involving gap-filling of the break and ligation of DNA ends. The latter is usually an error-free pathway, relying on the sister chromatid to faithfully copy DNA at the site of break in a conservative manner. However, this pathway can result in crossover during meiosis and in error-prone repair when the classical strend displacement strand annealing (SDSA) pathway fails. This mutagenic HR repair is carried out by long tract gene conversion (LTGC) or break induced replication (BIR), both relying on the formation of a migration bubble that moves way past the break side, generating a mutagenic region of ssDNA. Adapted figure from Scully et al., 2019

1.4.3 Replication stress and mitotic aberrations

Insults to DNA integrity or other obstacles such as transcription or R-loops (discussed later) can constitute a road block for the ongoing replication fork, causing it to pause, reverse or collapse in case the DNA damage is not effectively bypassed by repriming or translesion synthesis (Kotsantis et al. 2015; Zeman and Cimprich 2014). Altogether, these problems affecting replication are grouped under the general term of replication stress, which can be itself a direct cause of increased genomic instability, since it can result in DSBs formation or incomplete replication and consequent errors during mitosis (Cortez 2019; Ghosal et al. 2013; Kotsantis et al. 2015). To protect genome integrity, many factors are involved in fork protection and in mitigating replication stress. For example, under replication stress the replication fork structure can be altered and forks are remodelled: the newly synthesised DNA strands anneal together and there is the conversion of a typical replication fork (three-way junction) into a four-way junction, by the action of chromatin remodelers including SMARCAL1 among others (Cortez 2019). This is known as fork reversal, resulting in replication slowing thus providing enough time to the DNA repair machinery to be recruited and repair the damage (Neelsen and Lopes 2015; Qiu et al. 2021). These structures are susceptible to the action of nucleases, consequently there are factors that function in protecting and stabilising the reversed forks (Qui et al., 2021). Worth to mention, BRCA1 and BRCA2, and RAD51 that act separately of their role in HR in regulating fork reversal (Piwko et al. 2016; Quinet et al. 2017; Scully et al. 2019).

A consequence of replication fork slow down or collapse is the formation of replication intermediates that persist past the S phase, representing a challenge for cells entering into mitosis and becoming a primary source of chromosomal instability in daughter cells (Chan and

West 2018; Wilhelm et al. 2020). This replication intermediates together with DNA lesions that are not completely resolved during S phase can give rise to mitotic errors, which can be visualised by the formation of chromatin bridges (anaphase bridges), which are dsDNA and therefore DAPI positive, lagging chromosomes, and ultrafine bridges (UFBs, ssDNA) (Chen et al., 2018). UFBs are resolved, among others, by the action of BLM helicase and the single filament is coated by RPA and failure in solving these structures increases the risk of cytokinesis failure or chromosomal mis-segregation or end-to-end chromosome fusions, which are types of chromosomal instability (Liu et al. 2014; Chan et al. 2018; Kong et al. 2023). Another sign of that cells have not properly repaired any DNA damage before entering mitosis, or because of late replication intermediates that lead to DNA damage, is the formation of 53BP1 bodies, that are large accumulations of 53BP1 usually found in G1 cells (Lukas et al. 2011). On the other hand, mis-segregating chromosomes or persistent bridges can be incorporated in micronuclei, after cell division, giving rise to chromosomal instability (Wilhelm et al. 2020). Micronuclei can also form as consequence of a failure in DSBs repair or nucleus assembly. This structures are dangerous since they are a source of chromothripsis, a massive chromosomal rearrangement, highly frequent in human cancers (Krupina et al. 2021; Petropoulos et al. 2019). Conclusively, replication stress and consequent formation of under-replicated DNA, can give rise to mitotic abberrations such as DNA bridges, chromosomal rearrangements, chromosomal instability, 53BP1 bodies and micronuclei formation (Krupina et al. 2021; Saxena and Zou 2022).

1.4.4 Transcription associated genome instability

Endogenous sources of DNA damage account for DNA replication and transcription defects, among others (Jackson & Bartek, 2009; Gaillard & Aguilera, 2016). Transcription can be causative of transcription-associated mutations and transcription-associated recombination events. Altogether, these phenomena can be grouped under the general term of transcription-associated genome instability (Gaillard and Aguilera 2016; Kim and Jinks-Robertson 2012). Briefly, the main sources of this genomic instability caused by transcription are the topological constraints, supercoiling ahead and behind transcribing Pol II, the formation of DNA-RNA hybrids and non B-DNA structures and stalling and pausing of Pol II, that becomes itself an obstacle (Gómez-González and Aguilera 2019; Huertas and Aguilera 2003; Saponaro et al. 2014) (Figure 9).

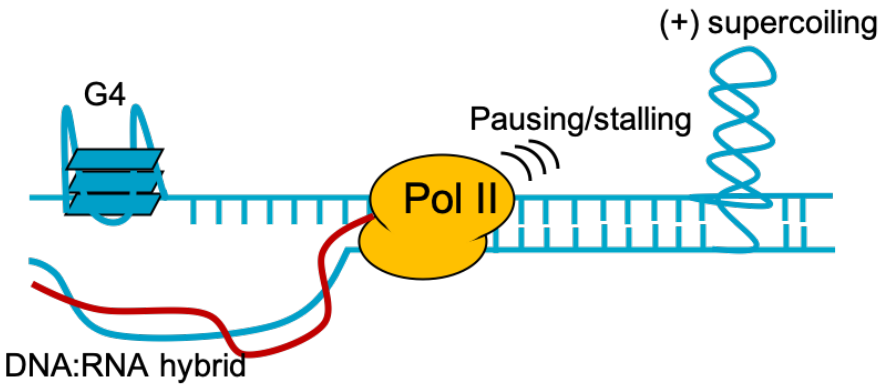


Figure 9 Transcription associated genome instability is caused by topological constraints, non B-DNA structures formation, DNA-RNA hybrid formation and Pol II stalling/pausing.

A primary cause of this increased genomic instability is the complex interplay between transcription and replication, two cellular process that have to coexist in the nucleus and utilize the same DNA as template. For instance, the collision between transcription and

replication machineries, known as transcription-replication collisions, constitutes a main source of transcription-associated DNA damage (Gómez-González and Aguilera 2019). Consistently, it has been found that actively transcribed genes are more prone to become sites of DNA damage, probably due to increased chances of transcription and replication fork collisions (Helmrich et al., 2011; Helmrich et al., 2013). Indeed, many factors among which topoisomerases I and II (TOP I-II), function in preventing the infelicitous encounter (Helmrich et al. 2013). Additionally, since transcription can continue beyond gene boundaries, transcription-replication conflicts are not rare events (Hamperl and Cimprich 2016). Early replicating fragile sites (ERFS), the first sites where replication stalls when cells are treated with hydroxy urea (HU), are also found in actively transcribed genes and are regions particularly subjected to transcription and replication encounters and therefore these regions are highly subjected to amplification or deletions (Barlow et al. 2013; Zeman and Cimprich 2014), indicating the link between this events and genome instability. Additionally, also common fragile sites (CFS) which are late replicated and enriched for long transcribed genes, are hotspots of breakages and rearrangements when replication is perturbed with low doses of aphidicolin (LeTallec et al. 2013; Ma et al. 2012). Replication forks and transcription can encounter in a head to head-on (HO) or co-directional (CD) fashion (Figure 10), with HO collisions a greater source of genomic alterations (Hamperl et al. 2017; Hamperl and Cimprich 2016). Subsequent fork reversal could impair proper chromosomal disjunction resulting in aneuploidy, while fork collapse could result in DNA breaking and formation of recombinogenic 3'-ssDNA (García-Muse and Aguilera 2016; Gómez-González and Aguilera 2019).

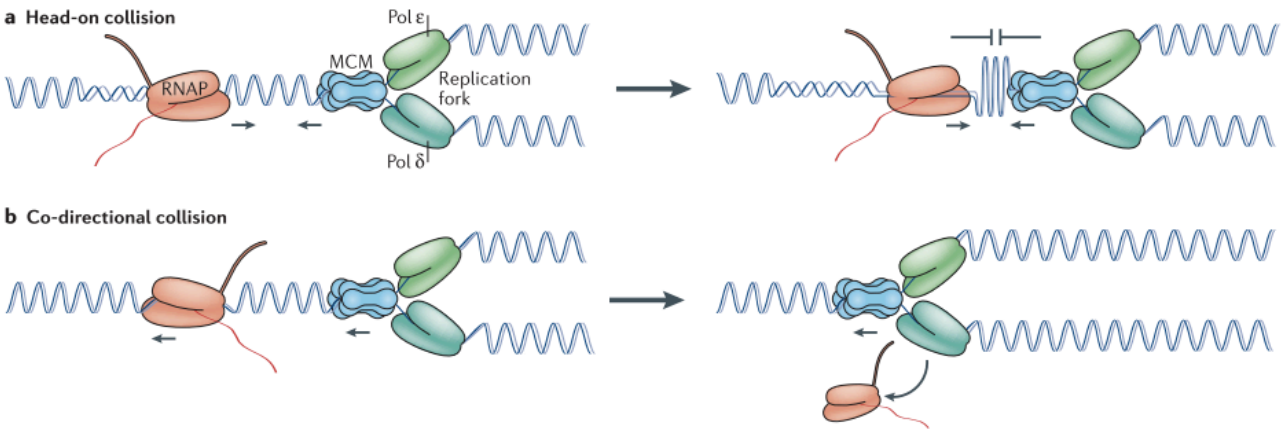


Figure 10 Replication transcription collisions. Head on TRC (a) can be causative of DNA breaks and subsequent transcription-associated recombination. Co-directional TRC (b) are less causative of genomic alterations since they can be resolved with Pol II displacement and replication fork progression. Adapted figure from García-Muse and Aguilera 2016

Another source of transcription associated genome instability is constituted by R-loops, secondary structures formed when the nascent RNA hybridizes with the DNA template strand (Hamperl and Cimprich 2016). R-loops can naturally form during transcription and they can be more or less stable according to the DNA sequence and its GC content (Boulianne and Feldhahn 2018). Although R-loops are normally generated during transcription and serve different biological roles, their accumulation can lead to transcription associated mutagenesis and transcription and replication conflicts, and overall increased genomic instability (Skourti-Stathaki and Proudfoot 2014).

For instance, R-loops associated ssDNA is highly mutagenic and recombinogenic and the resolution of these structures by transcription coupled NER (TC-NER), can result in double strand breaks formation (Hamperl and Cimprich 2016; Sollier et al. 2014; Tubbs and

Nussenzweig 2017). Additionally, R-loops may hinder transcription and increase the chances of transcription replication collisions (Figure 11) (Skourti-Stathaki and Proudfoot 2014).

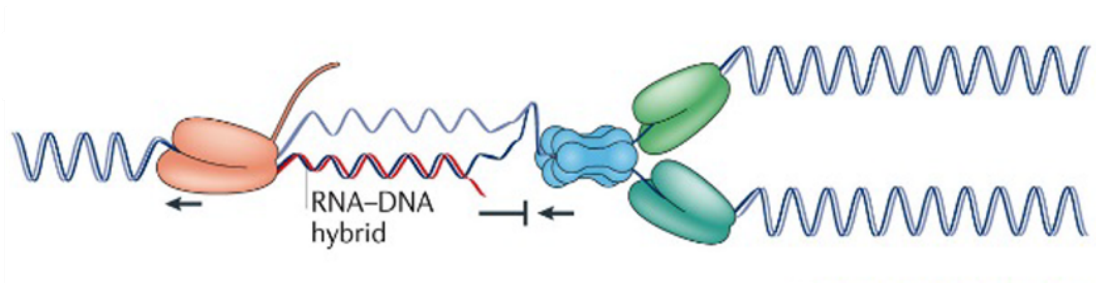


Figure 11 R-loops as a source of transcription associated genome instability. R-loops can result in transcription associated recombination, via the recombinogenic ssDNA, and DNA damage. R-loops are also implicated in promoting transcription collision with the replication fork. Adapted figure from García-Muse and Aguilera 2016

The formation of R-loops is kept under control by different means, including nuclease activity and resolution of topological stress (Allison and Wang 2019). Another mean of resolving R-loops is the Fanconi Anemia (FA) pathway, which has been implicated in active repression of R-loop formation, while helping in stabilizing and resolving replication fork stalling (Schwab et al. 2015). Even if R-loops are considered one of the main sources of transcription associated DNA damage (Brickner, Garzon, and Cimprich 2022), they can form also as a consequence of DNA breaks occurring during transcription. Therefore, it could be speculated that R-loops are both causative and consequence of replication stress (Teloni et al. 2019). Lastly, R-loops have been shown to play an important role in DSBs repair (Marnef and Legube 2021). Indeed, Pol II can be recruited at DSBs sites, together with some general transcription factors, where it generates the formation of DNA:RNA hybrids, which after being processed by the action of

RNaseH enzyme direct the DNA damage repair towards HR (Ohle et al. 2016; Ouyang et al. 2021). Additionally, the presence of R-loops at DNA damage sites prevents the initiation of a next round of transcription, thus ensuring correct DNA repair (Lim et al. 2023; Wang et al. 2023). Similarly to DNA replication machinery, which can be blocked by DNA modifications and lesions encountered on its way, also Pol II can arrest at DNA damage sites. There, it activates TC-NER which rescues transcription (Gómez-González and Aguilera 2019). Importantly, Pol II can also function as a damage sensor via its active site and can trigger different cellular responses including: lesion bypass, transcription slow down, Pol II ubiquitylation and/or removal from the fork, recruitment of chromatin remodelers and DDR proteins (Borsos et al. 2017; Lee et al. 2002; Nakazawa et al. 2020; Shin et al. 2016; Tufegdžić Vidaković et al. 2020; Wang et al., 2023).

1.5 Hypothesis

During transcription, Pol II subunits engage in protein-protein and protein-DNA/RNA interactions with TFs and template DNA and nascent mRNA. Pol II establishes interactions with the general TFs during transcription initiation (He et al. 2016; Kornberg 1999; Liu et al. 2013; Woychik and Hampsey 2002), the Mediator complex (Chen et al. 2018; Gupta et al. 2016; Haberle and Stark 2018; Plaschka et al. 2015; Schier and Taatjes 2020; Soutourina 2018), elongation factors (Bernecky et al. 2017; Compe et al. 2019; Ehara et al. 2017; Vos et al. 2018; Xu et al. 2017) and with components of co-transcriptional processes such as RNA splicing, capping and 3'-end processing (Eick and Geyer 2013; Garg et al. 2023; Genes et al. 2019; Ghosh et al. 2011; Harlen and Churchman 2017; Heidemann et al. 2013; Maita and Nakagawa 2020). Even if Pol II is always studied as a complex, recently it has become of interest how its subunits

could play independent roles in biological processes and can operate outside of the complex (Li et al. 2022, 2023). Given the extensive interactions that Pol II establishes during transcription and that the subunits are the interacting platform with TFs and transcription activators/repressor, we hypothesised that the higher levels of Pol II subunits could interfere with transcription and co-transcriptional processes, with the overexpressed subunits functioning as a sponge scavenging away TFs from transcribing Pol II. It is also conceivable that this is relevant not only for RNA Pol II transcription, but also for RNA Pol I and Pol III transcription. This because some subunits are shared between the three complexes, and because of RNA Pol II contributions to the transcription of the other RNA Polymerases (Abraham et al. 2020; Gerber et al. 2020). We believe therefore, that the study of single subunits should be further explored to determine their specific contribution to the transcription stages and the interchange of TFs, particularly relevant due to their possible valuable role as therapeutic targets or new biomarkers for high-risk patients in a wide variety of tumours.

1.6 Aims

In this work, we aim to describe how the up regulation of the four largest Pol II subunits POLR2A-B-C-D impacts on cancer patients, looking at survival and features of genomic instability (copy number alterations and mutations counts). We then aim to generate cell line models which enable us to characterize Pol II subunits overexpression, each at a time. Firstly, we want to verify if the overexpression of the single subunits induces more DNA damage, a potentially expected outcome of deregulated transcription, and where the DNA damage occurs. We then want to assess the impact of the overexpressions on transcribing RNA Pol II

and how they impact overall on transcription. Finally, once identified the transcription deregulations arising upon subunits overexpression, we want to decipher the mechanisms responsible of these alterations in order to identify specific defects we can target with commercial inhibitors.

2 Materials and Methods

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-Phospho-Histone H2A.X (Ser139)	Abcam	Cat#ab2893
Rabbit monoclonal anti-53BP1	Abcam	Cat#PA1-16565
Mouse monoclonal anti R-loops	The Francis Crick Institute Core Facility	S9.6
Rabbit polyclonal anti-Histone H2A.X	Merck-Millipore	Cat#07-627
Mouse monoclonal anti-Tubulin	The Francis Crick Institute Core Facility	Tat-1
Mouse monoclonal anti-Ser5-RPB1	The Francis Crick Institute Core Facility	4H8
Mouse monoclonal anti-U1 snRNP	Santa Cruz Biotechnoloy	Cat#sc-390988
Mouse monoclonal anti-XRN2	Santa Cruz Biotechnoloy	Cat#sc-365258
Mouse monoclonal anti-SPT5	Santa Cruz Biotechnoloy	Cat#sc-133217
Mouse monoclonal anti-RAD51	Santa Cruz Biotechnoloy	Cat#sc-398587
Rabbit polyclonal anti-Ser2-RPB1	Abcam	Cat#ab26721
Rabbit monoclonal anti-BRD4	Abcam	Cat#ab128874

Goat anti-rabbit IgG (HRP)	Thermofisher Scientific	Cat#31460
Goat anti-mouse IgG (HRP)	Thermofisher Scientific	Cat#31430
Alexa Fluor594-conjugated Goat Anti-Rabbit IgG (H+L)	Thermo Fisher Scientific	Cat#A- 11012
Alexa Fluor488-conjugated Goat Anti-Rabbit IgG (H+L)	Thermo Fisher Scientific	Cat#A-11070
Alexa Fluor488-conjugated Goat Anti-Mouse IgG (H+L)	Thermo Fisher Scientific	Cat#A- 11001
Vybrant DyeCycle Violet stain	Thermo Fisher Scientific	Cat#V35003
Reagents		
Zeocin	Thermo Fisher Scientific	Cat#R25001
Hygromicin	Thermo Fisher Scientific	Cat#J60681.MC
Doxycycline	Sigma-Aldrich	Cat#D9891
Geneticin/G418	Sigma-Aldrich	Cat#G8168
EU	Sigma-Aldrich	Cat#900584
4SU	Glentham LifeSciences	Cat#GN6085
DMF	Sigma-Aldrich	Cat#D4551
MTSEA biotin-XX linker	Biotium	Cat#BT90066
Dynabeads Protein A	Thermo Fisher Scientific	Cat#10002D
FuGENE HD	Promega	Cat#E2311
JQ1	Sigma-Aldrich	Cat#1268524-69-1
THZ1	MedChemExpress	Cat#HY-80013
Commercial assay kits and equipment		

DNA Clean & Concentrator-5	Thermo Fisher Scientific	Cat#10002D
RNeasy Mini Kit	Qiagen	Cat#74104
RNeasy MinElute Cleanup Kit	Qiagen	Cat#74204
Qubit dsDNA HS Assay Kit	Zymo Research	Cat#D4013
Qubit RNA HS Assay Kit	Qiagen	Cat#Q32852
MaXtract high-density phase lock gel tube 1.5mL/2mL	Qiagen	Cat#129046/129056
uMACS streptavidin kit	Miltenyi	Cat#130-133-282
Hybond-N membrane	GE Healthcare	Cat#RPN203N
Ulta II DNA Library prep Kit for Illumina	NEBNext	Cat#E7645S
Ultra RNA Library prep kit for Illumina	NEBNext	Cat#E7770
Multiplex Oligos for Illumina (96 Unique Dual Index Primer Pairs)	NEBNext	Cat#E6440S
Cell lines		
HeLa T-REx	ThermoFisher	Cat#R71407
Oligonucleotides		
See table for Real Time-PCR primer list	IDT	
Software and algorithms		
EaSeq	Lerdrup et al., 2016	https://easeq.net

Bowtie 2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
SAMtools	Li et al., 2009	http://www.htslib.org
STAR	Dobin et al., 2013	https://github.com/alexdobin/STAR
featureCounts	Liao et al., 2014	https://www.rdocumentation.org/packages/Rsubread/versions/1.22.2/topics/featureCounts
Galaxy	Afgan et al., 2018	https://usegalaxy.org
c-BioPortal		http://cbioportal.org
IntOGen		www.intogen.org
Bedtools	Quinlan and Hall, 2010	https://bedtools.readthedocs.io/en/latest/
IGV	Robinson et al., 2011	https://software.broadinstitute.org/software/igv/
CellProfiler	Carpenter et al., 2006	Version 4.0.3
GraphPad Prism	www.graphpad.com	Version 7
Excel	Microsoft	Version 16.35

2.1 Analysis of publicly available datasets

2.1.1 c-BioPortal

For the analysis of POLR2A/D deregulations we used the cBio Cancer Genomics Portal (cBioPortal, http://cbioportal.org).

To obtain the oncoprint of the genes of interest, we queried by selecting "Query by gene" function, for *POLR2A/D* in Breast Invasive Carcinoma (TCGA, Firehose Legacy), Acute Myeloid Leukemia (TCGA, Firehose Legacy), Brain Lower Grade Glioma (TCGA, Firehose Legacy), Head and Neck Squamous Cell Carcinoma (TCGA, Firehose Legacy), Skin Coutaneous Melanoma

(TCGA, PanCancer Atlas) and Liver Hepatocellular Carcinoma (TCGA, Firehose Legacy). Then, we selected for “mutations”, “Putative copy-number alterations from GISTIC”, “mRNA expression z-Scores – relative to all samples (where possible) or relative to diploid samples” with a z-Score threshold of ±2.0 as suggested by default.

To specifically select for patients with High or Low mRNA expression of at least one gene among *POLR2A/D* we used the Onco Query Language (OQL) “datatypes: exp>2, amp” or “datatypes: exp<-2, del”, respectively. From the “download” page we downloaded the list of patients obtained and imported it in an Excel file, where altered patients “altered=1” were selected. We obtained two sub-lists of patients for each of the six tumour types taken in consideration: one for patients that have at least one gene among POLR2A/D with high mRNA expression (referred to as “High” or “PolUP” group of patients) and one for patients with low mRNA expression for the same genes (referred to as “Low”). The “High” and “Low” groups of patients were identified in the complete list of patients downloaded from the “explore selected study” page on c-BioPortal and imported in Excel. Similarly, by selecting the “unaltered” patients, we also obtained the “Non High” sub-list of patients, containing all those patients where neither one of the four subunits was overexpressed. This allowed us to retrieve information for “overall survival in months” and “overall survival status”, “Fraction Genome Altered” and “Mutation Count” in both “altered” and “unaltered” patients. MKI67, PCNA and MCM2-7 gene expression levels comparison between “altered” and “unaltered” groups of patients was performed on cBioPortal, through the command “comparison – mRNA”.

2.1.2 IntOgen

The list of the most deregulated cancer driver genes was obtained from the IntOGen website (www.intogen.org), a platform that identifies cancer driver genes across different tumours,

collecting information available from large data-sets of tumour somatic mutations (Gonzalez-Perez et al. 2013). Similarly, we then restricted this analysis to the six tumours analysed for POLR2A/D upregulation, and retrieved the list of the mutated oncogenes in each cancer study from the TCGA dataset (i.e. “cohort”).

2.2 Cell culture

All cell lines were cultured in T75 Flasks (Corning) in complete DMEM: Dulbecco’s Modified Eagle’s Medium - high glucose (Sigma), supplemented with 0.01 % Pen/Strep (Gibco) 0.01 % Glutamine (Gibco) and 10 % Tet-free Foetal bovine serum (Gibco) (FBS) or Tet-free FBS (TakaraBio), under selection of 5 μg/mL Blasticidin (ThermoFisher) and the appropriate selection marker in a humidified incubator at 37°C and a CO_2 level of 5 %. For maintaining of cells in culture, cells were washed using pre-warmed (37°C) 1X Phosphate buffered saline (PBS, Gibco), trypsinised with 0.25 % Trypsin-EDTA (Gibco) and incubated for 3 minutes. Once detached, cells were diluted in additional volume of warm (37°C) complete DMEM and passaged to a new T75 to a final volume of 15 mL. Cells were usually passaged at around 80 % confluence.

2.2.1 Transfection and generation of stable clones

The host cell line HeLa T-REx (ThermoFisher) was transfected with the Invitrogen Flp-In target site vector, pFRT/lacZeo (ThermoFisher), to introduce an integrated FRT site and generate the host cell line HeLa T-REx Flp-in. Cells were transfected with FuGENE HD Transfection Reagent (Promega) according to manufacturer instructions. Briefly, cells were seeded in 6-well plates (Corning) at 60-70 % confluence and transfected with a mix containing Opti-MEM reduced serum medium (Gibco) media, 2 μg of DNA and FuGENE HD in a 1:3 ratio (DNA:Transfection

reagent). 24 h after transfection, cells were seeded in 10cm dishes (Corning) and culture media was supplemented with 5 µg/mL Blasticidin (ThermoFisher) and 100 µg/mL Zeocin (Invitrogen). Transfected cells were kept under selection until the formation of colonies. Single colonies were picked and moved to a 96-well plate (Corning). Once confluent, cells were expanded and finally moved to a T75 Flask (Corning). Cells were frozen in freezing media (10 % DMSO + 90 % FBS) and stored at -80°C.

HeLa T-REx were transfectred with pT-Rex-DEST31 plasmids (Source Bioscience) containing either Empty vector (EV), POLR2C or POLR2D coding region 6xHis N-terminal tagged. Transfection with Lipofectamine 2000 (Invitrogen) was performed according to the manufacturer instructions for Plasmid DNA transfection for adherent cells cultured in a 6-well plate. Briefly, cells were seeded in 6-well plates (Corning) at 60-70% confluence and transfected with 2 µg of DNA and Lipofectamine 2000 in a 1:3 ratio. 24 h after transfection, cells were seeded in 10 cm dishes (Corning) and culture media was supplemented with 5 µg/mL Blasticidin. In the cells transfected with either *POLR2C* or *POLR2D* expression plasmid, 48 h after transfection culture media was supplemented with 250 µg/mL Geneticin (Thermo Fisher). Transfected cells were kept under selection and once confluent moved in a T75 (Corning) flask. After at least two weeks under selection cells were frozen in freezing media (10 % DMSO + 90 % FBS) and stored at -80°C.

HeLa T-REx Flp-In host cell line was co-transfected with pFRT-TO-RPB1-His WT containing POLR2A coding region 6-His tag at the C-terminus and a Flp recombinase expression plasmid, pOG44 (Thermo Fisher). pFRT-TO-RPB1-His WT CRres si2,4R was a gift from Jesper Svejstrup (Addgene plasmid # 139404 ; http://n2t.net/addgene:139404 ; RRID:Addgene_139404). HeLa T-REx Flp-In cells were also co-transfected with EV (empty vector) and pOGG. Transfection was

performed according to the manufacturer instructions for FuGENE HD transfection agent as previously described. 24 h after transfection culture media was supplemented with 5 µg/mL Blasticidin and 48 h after transfection culture media was supplemented 5ug/mL Hygromycin (Thermofisher) for the cells transfected with POLR2A expression plasmid. Transfected cells were kept under selection until the formation of colonies. Single colonies were picked and moved to a 96-well plate (Corning). Once confluent, cells were expanded and finally moved to a T75 Flask (Corning). Cells were frozen in freezing media (10 % DMSO + 90 % FBS) and stored at -80°C.

2.2.2 Doxycycline Induction

To achieve the overexpression of POLR2A/C/D, 5-10 ng/mL Doxycycline (Sigma-Aldrich) was supplemented to the culture media, for 48 h. Expression level was assessed with western blot and quantitative PCR (qPCR).

2.3 RNA extraction and reverse transcription

Cells were seeded in 6 well plates at 20 % confluence. After 24 h, cells were induced with doxycycline as previously described. 48 h after induction, cells were trypsinised and pellet by centrifugation at 1000rpm. The pellet was processed for RNA extraction with RNeasy kit (Qiagen) according to manufacturer's instructions. Specifically, cells were lysed in 350 µL of RLT Buffer and RNA was eluted in 40 µL of Nuclease free water (Sigma). 1 µg of RNA was reverse transcribed with random hexamers using the SuperScript III reverse transcriptase kit (Invitrogen), according to manufacturers' instructions.

2.4 Primer design

Primers were designed using Primer BLAST (Ye et al. 2012). Primer sequences (**Table 1**) obtained by Primer BLAST were validated using USCS In-Silico PCR (Human Dec. 2013 (GRCh38/hg38) Assembly).

GENE	Forward	Reverse
POLR2A	CCCCAACCTCTCCATTGA	CCCTGCGCACTAGTTCTTGA
POLR2C	GCATGCCAAGTGGAACCCTA	TGTAGTAAAACCTTTCTGGCTTGC
POLR2D	ACAGCTCATCTTTCCTAAAGAGTT	CTGAGAGCTCCTGTTCGTCC
ACTB	CATGTACGTTGCTATCCAGGC	CTCCTTAATGTCACGCACGAT
RAD51	CAAAGGACCCAGAATAGCCGAAATG	AGCCAGGGCAAGACAGCGAGA
intergen C2	GCCTTAAGGTTTATACCAAAATCA	GGAAGGCACTGTTAAAGTTGAG
RPLP0	CAGATTGGCTACCCAACTGTT	GGAAGGTGTAATCCGTCTCCAC

Table 1 Primers list for RT-PCR designed with PrimerBLAST (Ye et al. 2012)

2.5 Quantitative (Real time) PCR

96 well PCR plate (Bio-Rad/Applied Biosystems) was loaded with 7.6 µl of Nuclease free water, 0.4 µL of 10 µM primer mix, 10 µL of SYBR Green super mix (Bioline) and 2 µL of cDNA (1:2.5 previously diluted) per well, in triplicate for each sample. RPLP0 was used as housekeeping reference gene. The fold change (FC) of expression level of POLR2A-C-D induced (dox) vs uninduced (control) was obtained by calculating the ratio of LOG2(dCt (Ct sample induced – Ct housekeeping)) and LOG2(dCt (Ct sample uninduced – Ct housekeeping)).

2.6 Nucleic Acid Gel Electrophoresis

Agarose (Thermofisher) was dissolved to 1 % (W/V) in Ultra Pure Tris borate-EDTA (TBE) Buffer 1X (Thermo Fischer Scientific). The solution was then microwaved until the agarose powder was dissolved and left to cool down for few minutes on the bench before adding SYBR Safe

gel stain 10000X (Thermofisher). The gel was then poured into the tray. Samples for were prepared by adding 6X SDS-Purple dye (NEB) prior loading. The gel was then run at 90V until DNA bands in the DNA ladder (NEB) were separated and imaged at the Chemidoc (Bio-Rad).

2.7 Protein extraction and cell fractionation

Cells were cultured in T75 as previously described and either doxycycline induced or non-induced for 48 h. Cells were then trypsinised and the pellet was washed twice in PBS.

For whole cell extract (WCE) the pellet was lysed in 250 µL of RIPA Buffer (10mM Tris-HCl, pH 8.0, 1mM EDTA, 0.5mM EGTA, 1% Triton X-100, 0.1% Sodium Deoxycholate, 0.1% SDS, 150mM NaCl) supplemented with protease inhibitors (Thermofisher) on ICE for 20'. The WCE was then sonicated using the Bioruptor (Diagenode) on high for 10', 30" on 30" off at 4°C. Sample was then cleared by centrifugation at 4°C at 10000 rpm for 10'. Samples were quantified using the Bradford method. Briefly, 1 µL of sample was diluted in 1 mL of Bradford dye (Bio-Rad) in a cuvette and absorbance was read at nm using a spectrophotometer. The absorbance was then plotted against a standard curve built using increasing concentrations of bovine serum albumin (BSA) to obtain the sample concentration. Between 10 and 30 µg of proteins were used for loading the SDS-PAGE gel. For running the gel samples were prepared by adding 2X SDS Laemmli Buffer (BioRad), supplemented with β-mercaptoethanol (BME), and boiled before loading at 95°C for 5'.

For cell fractionation cells were lysed in 200 µL of Cytosolic Buffer (0.15 % NP40 Alt (V/V), 10 mM TRIS pH 7.0, 150 mM NaCl) for 5' on ice, using a P1000 with cut tip. The lysate was then layered on 500 µL of Sucrose Buffer (10 mM TRIS pH 7.0, 150 mM NaCl, 25 % Sucrose) and centrifuged at 16000xg for 10' at 4°C. The supernatant was collected as fraction 1

(cytoplasmatic fraction). Nuclei were washed twice with 800 µL of Wash Buffer (1 mM EDTA, 0.1 % (V/V) triton in PBS), each time spin down at 1150xg for 1' at 4°C. Then, nuclei were resuspended in 200 µL of Glycerol Buffer (20 mM TRIS pH 8.0, 75 mM NaCl, 0.5 mM EDTA, 50 % glycerol (V/V), 0.85 mM DTT) and then lysed in 200 µL of Nuclear Lysis Buffer (1 % NP40 Alt (V/V), 20 mM HEPES pH 7.5, 300 mM NaCl, 1 M Urea, 0.2 mM EDTA, 1 mM DTT) vortexing on high 5 seconds and then on ice for 5'. Lysate was centrifuged at 18500xg 2' at 4°C and supernatant was collected as fraction 2 (nuclear fraction). Pellet was washed once in PBS and then resuspended in 100 µL PBS supplemented with 1.5 mM $MgCl_2$ and 1 µL of Benzonase (5KU; Sigma) and left 1 h at room temperature (RT) or until digested on ice. Digested material was collected as fraction 3 (Chromatin fraction) after a brief spin. Then fractions were stored at -20°C and analysed by western blot. All buffers are supplemented with 1 % protease and phosphatase inhibitors, freshly added.

2.8 Western blot

SDS-PAGE Gels were prepared at either 8 % or 12 % (Acrylamide : Bis-acrylamide) using 30 % Acrylamide (Bio-Rad) for the separating gel and 4 % for the stacking gel. Gels were run between 90 and 140 V in Running Buffer (Tris Base, Glycine, 20% SDS) until the separation of desired molecular weight was obtained. Gels were then transferred using Transfer Buffer (Running Buffer, 20% Methanol) for 1 h and 30' on Nitrocellulose membrane (BioRad). The membrane was then blocked in 5 % milk ((W/V) in TBS (Tris, NaCl) Tween 0.1 % (V/V) (TBST) for 1 h at RT and incubated with the desired antibodies, diluted 1:1000, in 1 % milk in TBST overnight at 4°C. Membranes were then washed three times for 5' each in TBST and incubated with HRP conjugated secondary antibody in 1 % Milk in TBST for 1 h at RT. Before developing,

membranes where again washed three times in TBST. Membranes were incubated for 2' at RT in the dark with ECL (BioRad) and then visualized using the Chemidoc (Bio-Rad) or by using films (Amersham) in the dark room. Images and bands were then quantified with ImageJ.

2.9 Immunofluorescences and image processing

2.9.1 Immunostaining

Cells were seeded on poly-lysine coated slides in 6-well plates, at 20 % confluence and after 24 h were induced with doxycycline for 48 h. After one wash with PBS, cells were fixed and permeabilized for 20 min in PTEMF buffer (20 mM PIPES pH 6.8, 10 mM EGTA, 0.2 % Triton X-100 (V/V), 1 mM $MgCl_2$, 4 % formaldehyde (V/V)) at RT and then washed again in PBS for three times, 5' each. Fixed samples were stored at 4°C until staining.

Before antibody staining, cells were blocked for 1 h at room temperature with 10 % FBS/PBS (V/V; Sigma Aldrich). Cells were washed in PBS and incubated with 1:500 diluted rabbit anti-53BP1 (Abcam), 1:1000 rabbit anti-Phospho-Histone H2A.X(Ser139) (Abcam), 1:100 mouse anti-R-loops S9.6 and anti-Rad51 (The Francis Crick Institute, SantaCruz) in 1% FBS/PBS for 1 h at RT or at 4°C overnight. Cells were then incubated with 1:1000 diluted Alexa Fluor488-conjugated goat anti-rabbit antibody or Alexa Fluor594-conjugated goat anti-rabbit or Alexa Fluor488-conjugated goat anti-mouse (Thermo Fisher Scientific) in 1 % FBS/PBS for 1 h at RT. After being washed in PBS as previously described, coverslips were mounted on slides with a drop of mounting medium with DAPI (GeneTex).

2.9.2 EU pulse labelling and Click-iT reaction

Cells were grown in the same conditions as for the "immunostaining", but were pulsed with 50 µM ethynyl-uridine (Thermofisher) (EU) for 1 h to label nascent RNA before being fixed and

permeabilized in PTEMF buffer. Fixed cells were then processed according to the Click-iT RNA Alexa Fluor 594 Imaging Kit (Invitrogen) manufacturer's instructions for cells, before proceeding to either DAPI mounting or primary antibody staining.

2.9.3 Image acquisition

All images were acquired with Nikon Eclipse E600 microscope. 53BP1 foci images were acquired with 40X objective, EU click-iT, γH2AX, Rad51 foci and R-loop images with 60X oil objective. For each sample a number > 100 nuclei were counted. For a complete set of samples, the same exposure settings were used.

2.9.4 Image processing

53BP1 and γH2AX foci or γH2AX , EU, S9.6 intensity levels were counted/analysed with CellProfiler 4.0.3 (*Carpenter et al., 2006*). All images were firstly "colored to gray". Nuclei were identified with "identifyAsPrimaryObjects" module (Typical diameter of objects, in pixel units (Min,Max):35,130, 'Discard objects outside the diameter range?:Yes', 'Discard objects touching the border of the image?:Yes', 'Method to distinguish clumped objects:Shape', 'Method to draw dividing lines between clumped objects:Shape', 'Size of smoothing filter:10', 'Thresholding method:Otsu'). Nuclei where then used to mask the foci image with "MaskImage" module, and then foci where identified with "IdentifyPrimaryObjects" module (Typical diameter of objects, in pixel units (Min,Max):3,20 'Discard objects outside the diameter range?:Yes', 'Discard objects touching the border of the image?:No', 'Method to distinguish clumped objects:Intensity', 'Method to draw dividing lines between clumped objects:Intensity', 'Size of smoothing filter:3', 'Thresholding method:Robust Background'). Intensity was measured on all gray images and on masked images with "MeasureObjectIntensity" module. Finally, each "children" foci was associated to a "parent"

nulceus with “RelateObjects” module and results were exported on an Excel sheet with “ExportToSpreadSheet” module.

53BP1 bodies and mitotic errors (anaphase bridges and lagging chromosomes) were counted by eye.

2.10 Flourescence activated cell sorting (FACS) – cell cycle analysis

Cells were seeded in a 60mm dish (corning), induced or non-induced for 48h, trypsinised and fixed in 70% Ethanol (V/V) and stored at 4°C. Cells were washed twice with PBS at RT, resuspended in 500 µL of PBS with vybrant dye cycle (Thermofisher) according to manufacturers’ instructions and incubated O/N in dark. Cells were analysed with Beckman Coulter CytoFLEX S. Cells profiles were obtained with CytExptert Software for CytoFLEX.

2.11 Crystal violet staining assay

Cells were induced or non-induced for 48 h, counted and 1000 cells were seeded in 96 well plate in 50 µL of DMEM. Each condition was seeded in quadruplicate. After 24 h, cells were treated with either DMSO, JQ1 500nM, THZ1 1uM, with a top up of media of 50 µL. After 24 h the first plate was washed twice with PBS and stained with 4 % crystal violet (V/V; Sigma Aldrich) in 20 % Methanol (V/V) for at least 30’. Then, plate was washed until all the excess staining was removed and left to air dry. A plate was stained every day for a total of up to 7 days (up to when the untreated uninduced control was confluent). Every 2 days media was replaced to the remaining plates with DMEM +/- Doxycycline supplemented with either DMSO or inhibitors. 50 µL of 10 % (v/v) acetic acid glacial were used to dissolve crystal violet and

absorbance was read at 595nm with a plate reader (EnSpire multimode plate reader, PerkinElmer).

2.12 Chromatin immuno-precipitation

For chromatin immune-precipitation (ChIP) cells were seeded in 150 mm dishes (Corning) at 20 % confluence and induced with Doxycycline. For 6 ChIP reactions, two 150 mm dishes per conditions were used. After 48 h, cells were trypsinised and collected in culture media in a 50 mL tube. Then, cells were crosslinked with 1% formaldehyde final concentration for 10' on rotation. Crosslinking was quenched by adding glycine to 125 mM final concentration for 5' on rotation. Cells where then pelleted by centrifugation at 1000 rpm for 5' with cell pellet washed twice in cold PBS and transferred to a 2 mL centrifuge tube. The cell pellet was lysed in 1.5 mL of Cell lysis Buffer (5 mM HEPES pH 8, 85 mM KCl, 0.5% NP40 Alt (V/V)) on ice for 5'. Nuclei were pelleted by centrifugation at 3900g for 5' at 4 °C. Nuclei were lysed in 1.5 mL Nuclear Lysis Buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 1 % SDS (V/V)) for 5' on ice. Prior to sonication, samples were divided into 300 µL aliquots and nuclear lysates were sheared by sonication using a tip sonicator set at 30 % Amp, 15 s ON, 25 s OFF, for 8 cycles. Samples were kept on ice between cycles of 2, or alternatively sonicated on a cooling rack. DNA size was measured on a 1% Agarose gel and samples were deemed sonicated appropriately when there was a mean smear visible between 200 and 600 bp. Sonicated chromatin was cleared by centrifugation at 20,000 g for 10' at 4 °C. Cleared chromatin was stored at -80 °C until immunoprecipitation.

For the immune-precipitation (IP) reaction, 15 µL of Protein A Dynabeads (Invitrogen) were washed twice with 5% BSA in PBS (W/V, Thermofisher) and incubated for 1 h at RT with 1ug

of the corresponding antibody: rabbit anti-Phospho-Histone H2A.X(Ser139) (Abcam); rabbit anti-Phospho-Histone H2A.X (Merck-Millipore); mouse anti Ser5-RPB1 (4H8) (Cell Signalling); mouse anti NTD-RPB1 (Cell Signalling); rabbit anti Ser2-RPB1 (Abcam). Before the IP reaction, the beads were washed in BSA/PBS and for each reaction 150 µL of sheared chromatin were diluted 1:5 with dilution buffer (16.7 mM Tris-HCl pH 8.0, 1.2 mM EDTA pH 8.0, 0.01 % SDS, 1.1 % triton (V/V), 167 mM NaCl) and incubated overnight at 4 °C with the prepared beads. 10 % chromatin was kept aside as the 10 % input. Up to this point, all buffers were supplemented with 1 % protease inhibitors.

After overnight incubations beads were washed twice subsequentially in ChIP low salt buffer (0.1 % SDS, 1 % Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl), ChIP high salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 500 mM NaCl), ChIP LiCl buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 % NP40, 1 % (W/V) deoxycholic acid and 1 mM EDTA) and once in TE buffer (Tris 10mM, EDTA 1mM). Beads were centrifuged at 1000g for 3' to remove the last wash and centrifuged again at 16000g for 1' to carefully remove all the remaining supernatant. DNA was eluted in 40 µL of Elution Buffer (50mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 1 % SDS (V/V)) by carefully resuspending the beads by pipetting at least 10 times, at 65°C for 10'.

Eluted DNA was incubated at 65°C overnight for de-crosslinking together with the 10 % input DNA. Followingly, samples were diluted in TE buffer to dilute the SDS concentration in the elution buffer to 100 µL final volume and treated with 1 µL of 20 mg/mL RNAse A (Applichem) for 1 h and 3 µL of Proteinase K (Qiagen) for 2 h at 40°C. Finally, DNA was purified by PCR Mini Elute purification Kit (Qiagen) according to manufacturer's instructions and eluted in 10 µL Nuclease Free water. 2 µL of the purified DNA were used for quantification with Qubit dsDNA

HS Assay Kit (Thermo Fisher Scientific) using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific) according to manufacturer's instructions and 2 µL were used to test ChIP efficiency by ChIP-qPCR, with the remaining material used for the library preparation

2.12.1 ChIP-qPCR

2 µL of DNA were diluted 1:25 in nuclease free water for single gene analysis by qPCR. The input was serially diluted 1:3, 1:9, 1:27 and was used to create a calibration curve. The ChIP efficiency was then calculated as % over the input. Primers were designed at the transcription start site (TSS) or in the gene body (GB) of housekeeping genes.

2.12.2 Library Preparation and Sequencing

Library preparation was done with the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (NEBNext) according to the manufacturer's instructions. The libraries were indexed with the 96 Unique dual index primer pairs (NEBNext). Indexed libraries were quantified, normalized and pooled, and sequenced on Illumina NextSeq system in four lanes on a single end run.

2.12.3 Bioinformatics analysis

The library quality was assessed using the FastQC. Reads were aligned to the hg38 genome using Bowtie2 v.2.3.4.2 (Langmead and Salzberg 2012) on the online platform Galaxy (https://usegalaxy.org; Afgan et al. 2018). BAM files from single lanes were merged using SAMtools Merge (Li et al. 2009). Read coverage profiles were generated using the EaSeq v 1.101, with the function "Average" around 120% relative to the middle of the gene (Lerdrup et al. 2016). ChIP-Seq signal levels were obtained using the function "Quantify" in EaSeq, from the start to the end of the gene of all annotated transcribed genes, or else as indicated in figure legends. The relative level of γH2AX/H2AX was calculated over the gene or the genomic region. Single genes snapshots were captured on IGV (Robinson et al. 2011).

2.13 TT-chem Seq

2.13.1 RNA labelling and extraction

Cells were seeded in 10 cm dishes (Corning) at 20 % confluence and induced with Doxycycline after 24 h. 48h after induction cells were processed as described in Gregersen et al. 2020.

Cells were labeled with 1 mM 4-thiouridine (4SU) (Glentham Life Sciences) for 15’ at 37°C. The labelling was stopped by removing the culture media and adding 1 mL of TRIzol (Thermo Fisher). Cells were collected into a microcentrifuge tube by scraping the plate with a cell scraper. The culture media was removed in the tissue culture while the TRIzol was added under the fume hood.

RNA was then extracted by adding 200 µL of chloroform (Thermo Fisher Scientific) to the TRIzol-cell mixture. After adding chloroform tubes were vigorously shaken for 30’’ and centrifuged at 12000g for 15’ at 4°C. The aqueous phase was then loaded into a MaXtract high-density phase lock gel tube (Qiagen) and chloroform/isoamyl alcohol (24:1; Sigma-Aldrich) was added. The tubes were shaken and centrifuged at 12000g for 5’ at 4°C to separate the organic phase (that will be found below the resin) and the aqueous phase (found above the resin). The recovered aqueous phase was transferred to a new centrifuge tube and isopropanol was added in a 1:1.1 ratio and the mix was incubated for 20’ at RT. After incubation the tube was centrifuged at 12000g for 20’ at 4°C. The so obtained RNA pellet was washed in 750 µL of 85 % (V/V) ethanol without resuspending the pellet and centrifuged at 7500g for 5’ at 4°C. Once all the ethanol was removed the pellet was left to air dry and then it was resuspended in 50 µL of nuclease-free water. The RNA concentration was measured at the Qubit according to manufacturer’s instructions.

2.13.2 Dot blot to assess 4SU incorporation

2 µg of RNA were diluted in nuclease-free water in a final volume of 247 µL. 3 µL of biotin buffer (833 mM Tris-HCl, pH 7.4, and 83.3 mM EDTA) and 50 µL of 0.1 mg/ml MTSEA biotin-XX linker (MTSEA biotincapcap (2-((6-((6-((biotinoyl)amino)hexanoyl)amino)hex-anoyl)amino)ethylmethanethiosulfonate); Biotium) dissolved in DMF (dimethylformamide; Sigma-Aldrich) were added and the mix was incubated for 30' at RT in the dark. Once RNA is biotinylated it is purified using phase-lock-gel tubes. Prior to adding the RNA to the tube it was mixed with 250 µL of phenol/chloroform/isoamyl alcohol (25:24:1 (vol/vol/vol); Thermo Fisher). Then the tube is mixed and centrifuged at 12000g for 5' at 4°C and the aqueous phase containing RNA is transferred into a new centrifuge tube. To precipitate the RNA, 5 M NaCl in a 10:1 ratio and isopropanol in a 1:1.1 ratio are added to the aqueous phase and mixing by inverting the tubes 3 to 5 times and incubating for 10' at RT. The RNA is then pelleted by centrifugation at 20000g for 20' at 4°C and washed in 500 µL of 85 % (V/V) ethanol and centrifuged again at 20000g for 5' at 4°C. After the accurate removal of all ethanol the pellet is left to air dry and is finally resuspended in 10 µL of nuclease free water. To facilitate the loading of the RNA, 0.001% (W/V) bromophenol blue (Sigma-Aldrich) was added to the eluted RNA, which was then dropped on a Hybond-N membrane (GE Healthcare) using a dot blot apparatus. The membrane was UV-crosslinked at 0.2 J/cm2 (254 nm) in a Stratalinker (Stratagene). The membrane was blocked in blocking buffer (10 % (W/V) SDS, 1 mM EDTA in 1X PBS) for 20' at RT. After blocking the membrane was incubated in blocking buffer with 1 mg/mL HRP-conjugated streptavidin diluted 1:50000, for 15' at RT. After probing, the membrane was washed in blocking buffer and twice in washing buffer I (1 % (V/V) SDS in PBS) and twice in washing buffer II (0.1 % (V/V) SDS in PBS). Each wash was performed twice for 10'

at RT. The membrane was then incubated with ECL (diluted 1:5 in water) for 2' in dark and the signal was visualized at the Chemidoc (Bio-Rad). To serve as loading control, the membrane was then stained in staining buffer (0.5 M sodium acetate and 0.5 % (W/V) methylene blue) for 10' at RT and de-stained with washes in water (repeated as necessary until convenient).

2.13.3 RNA fragmentation

100 µg or 50 µg of 4SU labeled RNA were mixed with 1 µg or 0.5 µg of *S. cerevisiae* BY4741 RNA labeled with 4TU to serve as spike in, in a total volume of 100 µL. 20 µL of 1 M NaOH were added to the RNA for 20' on ice to chemically fragment the RNA. The RNA fragmentation was then stopped by adding 80 µL of 1 M Tris pH 6.8 and quickly proceeding to the clean-up reaction using the Micro Bio-Gel spin columns, following manufacturer's instructions. The clean-up with the columns was repeated twice per sample to ensure that the RNA solution reached a neutral pH and avoid further RNA fragmentation. The RNA was then run on an Agilent 2200 Tapestation system as for manufacturer's instructions for High sensitivity RNA screen tape, in order to confirm that the fragmentation was successful (Figure 12)

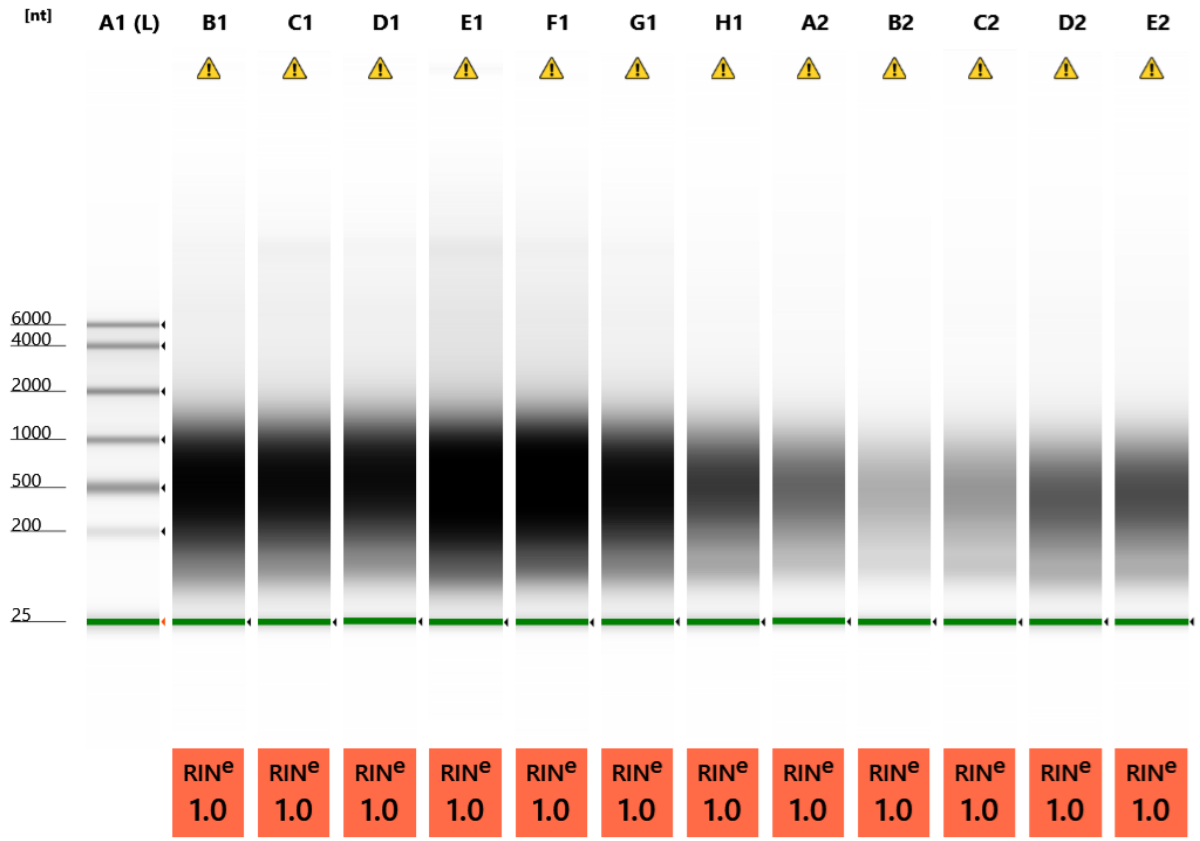


Figure 12 Fragmentation control by Tapestation analysis. Fragmented RNA samples were run on a Tapestation system to assess the correct fragmentation. Loading from lane A1 to lane G1 = ladder, first replicate samples WT ctl, WT dox, 2C ctl, 2C dox, 2D ctl, 2D ctl, lanes H1-E2 = second replicate samples WT ctl, WT dox, 2C ctl, 2C dox, 2D ctl, 2D ctl.

2.13.4 Biotinylation and streptavidin pull-down

3 µL of biotin buffer and 50 µL of 0.1 mg/mL MTSEA biotin-XX linker (in DMF) were added to the fragmented RNA and the reaction was incubated for 30' at RT in the dark. Then, the biotinylated RNA was purified as described in "2.11.3". Once biotinylated, the RNA was denatured at 65°C for 10' quickly followed by incubation on ice for 5'. 200 µL of uMACS streptavidin MicroBeads (Miltenyi) were added to the RNA and incubated on rotation for 15' at RT. Then, the streptavidin MicroBeads were applied to a uColumn placed on the uMACS

magnetic separator and prepared according to the kit's instructions (Miltenyi). The column was then washed with warmed at 55°C pull-down wash buffer (100 mM Tris-HCl, pH 7.4, 10 mM EDTA, 1 M NaCl and 0.1 % (V/V) Tween 20). The biotinylated RNA was eluted in 100 µL of elution buffer (100 mM DTT freshly made in nuclease-free water) at RT. A second elution was repeated after 5' and the two eluted samples were pooled together in a total volume of 200 µL. The eluted RNA was concentrated using the RNeasy MinElute Cleanup Kit (Qiagen), following manufacturer's instruction but adding 700 µL of RLT buffer and 1050 µL of 100% ethanol (for a 200 µL sample). This is important to retain fragments smaller than 200 nucleotides that otherwise would be discarded in the flow through. Finally, the RNA was eluted in 15 µL of nuclease-free water and quantified at the Qubit prior to library preparation.

2.13.5 Library Preparation and Sequencing

Library preparation was done with the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina (NEBNext) according to the manufacturer's instructions. Given the size distribution of the RNA, the protocol followed was the one for "degraded samples". The libraries were indexed with the 96 Unique dual index primer pairs (NEBNext). Indexed libraries were quantified, normalized and pooled, and sequenced on Illumina NextSeq system in four lanes on a single end run.

2.13.6 Bioinformatics analysis

The library quality was assessed using the FastQC, on the online platform Galaxy (https://usegalaxy.org; Afgan et al., 2018), reads were aligned using STARRNA (Dobin et al. 2013) to the hg38 genome and to the sacCer3 genome, using existing gene annotation information. The Gene annotation files (GTFs) were downloaded from Ensemble release 110 with reads aligned against each reference genome. The total number of unique mapped reads

of the yeast spike in, in the BAM file, was used to calculate a scale factor among different samples. For each replicate the WT empty vector control (uninduced) sample was used as reference normalizing factor. Strand specific bedgraph files were then created using the SAM tools (Li et al. 2009) to split the BAM file in + and – strand and using the advance settings to include the scale factor previously calculated. Read coverage profiles were generated using the EaSeq v 1.101, with the function “Average” around transcription start sites (TSS) or transcription termination sites (TTS) as indicated in the figure legends (Lerdrup et al. 2016). TT-Seq levels were obtained using the function “Quantify” in EaSeq, from the start to the end of the gene of all annotated transcribed genes, or else as indicated in figure legends.

2.14 Statistics

Kaplan-Meier (KM) survival curves and box plots were generated using software Prism (GraphPad 7). The statistical equivalence between the curves was evaluated with the Log-rank test and with the Wilcoxon test while Student t-test was used to compare the two groups. Student t test was carried out with Prism (GraphPad 7) or Excel (16.35). P value below 0.05 was considered statistically significant.

3. Results

3.1 PolUP is associated with patients' poor survival and increased genome instability in cancer patients

3.1.1 Cancer genomic data analysis

Genome instability and deregulated transcription are established hallmarks of cancer (Negrini et al., 2010; Hanahan & Weinberg, 2011). Indeed, the acquired capabilities of cancer cells during tumorigenesis lead eventually to an alteration in gene expression, and many well-known oncogenes and tumour suppressor genes are transcription factors (Hanahan & Weinberg, 2011; Bywater et al., 2013). For this reason RNA Polymerase complexes are considered to be valid direct or indirect targets for anticancer therapies (Bywater et al., 2013; Ferreira et al., 2020; Pan, 2010; Drygin et al., 2010; Petrie et al., 2019; Bywater et al., 2012; Bradner et al., 2017). Interestingly, little is known about specific functions of the RNA Polymerase complexes' subunits and in particular of Pol II subunits, both in transcription and in other cellular functions. However, recently deregulations and mutations in RNA Pol II subunits have been identified in cancer patients as mentioned previously in the introduction (Muste Sadurni and Saponaro 2023). Therefore, we wanted to better investigate whether any deregulation in Pol II subunits could have a potential role in tumourigenesis and whether cancers with deregulation in Pol II subunits shared common features. We focused on the four largest subunits of the complex, RPB1-4, encoded by *POLR2A/D* genes, since they mediate the majority of interactions occurring between Pol II and transcription associated complexes (Martinez-Rucobo et al., 2015; Plaschka et al., 2015; Ehara et al., 2017; Bernecky et al., 2017).

We extended our analysis to the other eight Pol II subunits, but when included to our search the phenotypes described in this chapter were lost (see below).

We exploited the The Cancer Genome Atlas (TCGA) data available in cBioPortal. The portal allows the research community to easily access data generated by cancer genomic projects such as TCGA, consenting the exploration of cancer genomic datasets and analysis across genes. cBioPortal genomic data type include: somatic mutations, DNA copy-number alterations (CNAs), mRNA expression level (RNA-Seq data or microarray), microRNA expression level, DNA methylation, protein and phosphoprotein abundance (Cerami et al., 2012; Gao et al., 2014). We firstly assessed what deregulation and mutations affected the four major subunits of Pol II in the Breast Invasive Carcinoma (BIC, TCGA Firehose Legacy), being this the largest study in number of patients present on the web portal. When exploring this study we selected for “mutations”, “putative copy number alterations from GISTIC”, “mRNA expression – z-scores relative to diploid samples (log RNA Seq V2 RSEM)”, with a z-score of 2 as default (Figure 13).

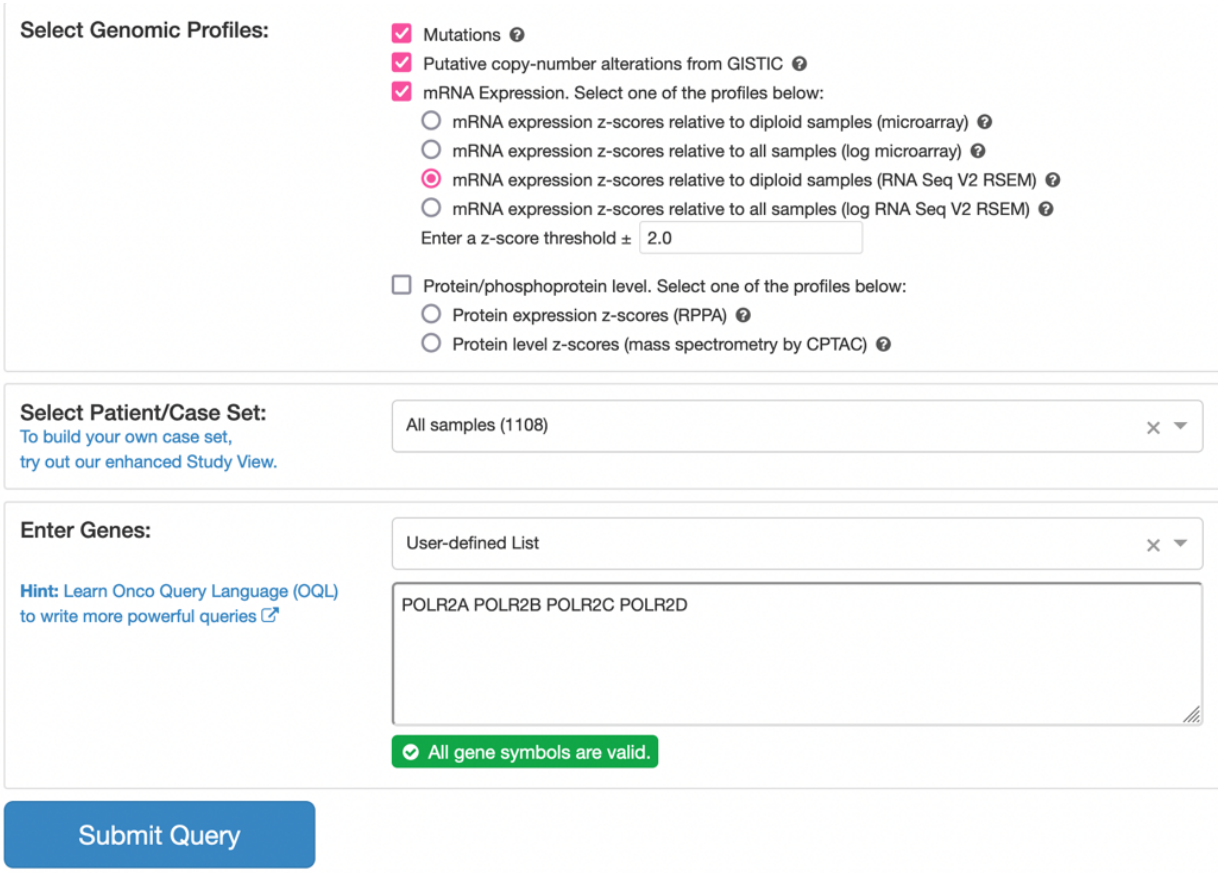


Figure 13 Query in Breast Invasive Carcinoma Screenshot from cBio Portal of the query on the selected study of Breast Invasive Carcinoma (TCGA, Firehose)

For mRNA expression, the z-score is calculated only using patients' data. The z-score is a statistical measurement indicating how a value is related to the mean of a group of values. Hence, z-score is measured in terms of standard deviations from the mean. This means that a z-score of 0 indicates that the data point's score is identical to the mean. A z-score of 2 indicates that a value is two standard deviations from the mean. Z-scores may be positive or negative, with a positive value indicating the score is above the mean and a negative score indicating it is below the mean. In cBio-portal, the z-score can be calculated according to the distribution on either all samples or only diploid samples. According to the web site, in the

distribution based on all sample, the expression distribution of the gene is estimated by calculating the mean and variance of all samples with expression values; if the gene has samples whose expression values are all zeros or non-numeric, then its normalized expression is reported as NA. In the distribution based on diploid samples, the expression distribution for unaltered copies of the gene is estimated by calculating the mean and variance of the expression values for samples in which the gene is diploid; if the gene has no diploid samples, then its normalized expression is reported as NA.

Therefore, we obtained the oncoprint for BIC, showing all the deregulations affecting *POLR2A/D* in all the patients (Figure 14A). We observed that *POLR2A-D* genes are upregulated or downregulated in a great number of patients (16%). We then narrowed our search to only down (low mRNA expression) or up (high mRNA expression) regulation affecting the genes of interest. We observed that for down regulation the altered patients are the 7.3%, while for up regulation the 8.6% (Figure 14B). We therefore investigated whether the down or up-regulation of these subunits could be associated to a specific outcome in patients, in particular patients' overall survival status. For our analysis we compared the altered group of patients (i.e. patients with at least one among the four subunits up or down-regulated) with the unaltered population (i.e. all other patients). Interestingly, an up or down regulation in at least one among the four largest Pol II subunits pooled together resulted in an opposite phenotype. Altered patients for higher mRNA expression showed a significant drop in their overall survival, whereas patients with a lower mRNA expression showed no significant difference compared to the unaltered population (Figure 15)

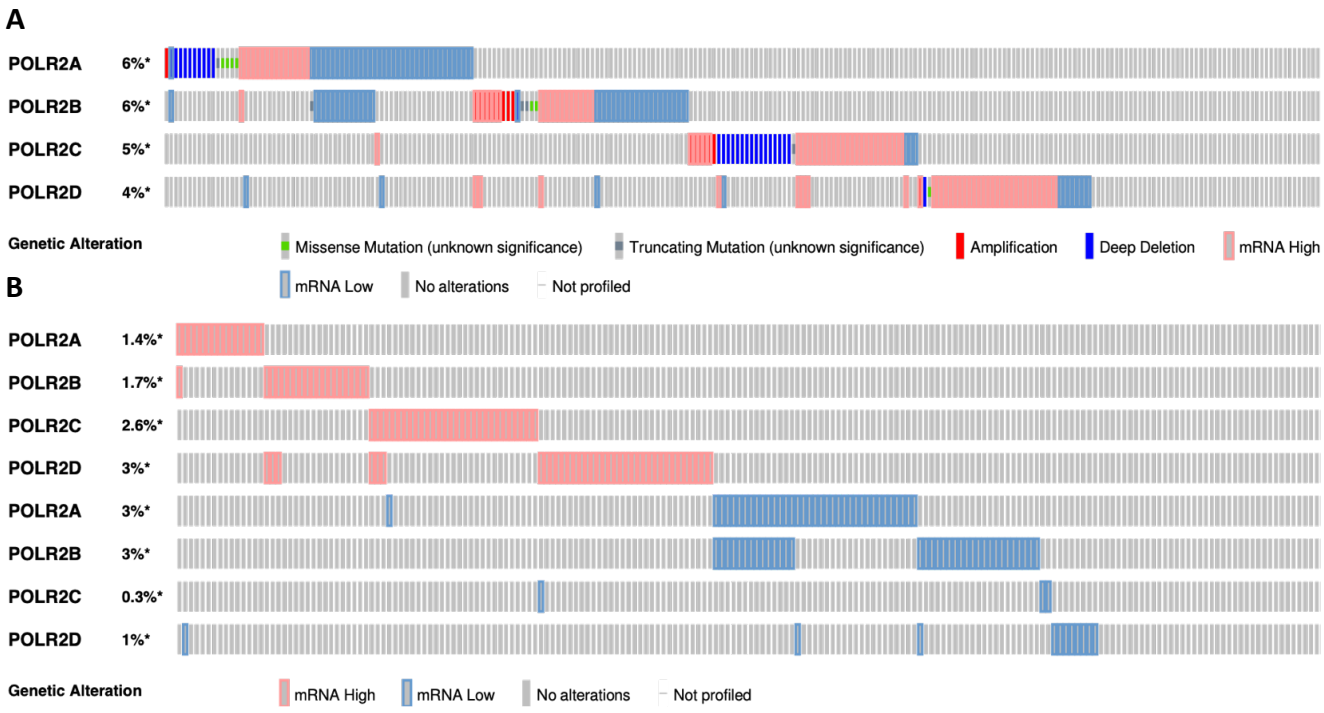


Figure 14 Deregulation of Pol II subunits in BIC A) Oncoprint of BIC showing in the panel the oncoprint of all deregulations affecting POLR2A-D in Breast Invasive Carcinoma patients (TCGA, Firehose Legacy). B) the oncoprint separating patients with upregulation from those with downregulation in POLR2A/D in all patients. The settings used are: mRNA expression – Z-score relative to all samples (log RNA Seq V2 RSEM), Z-score +/- 2 as default. Source: www.cbioportal.org

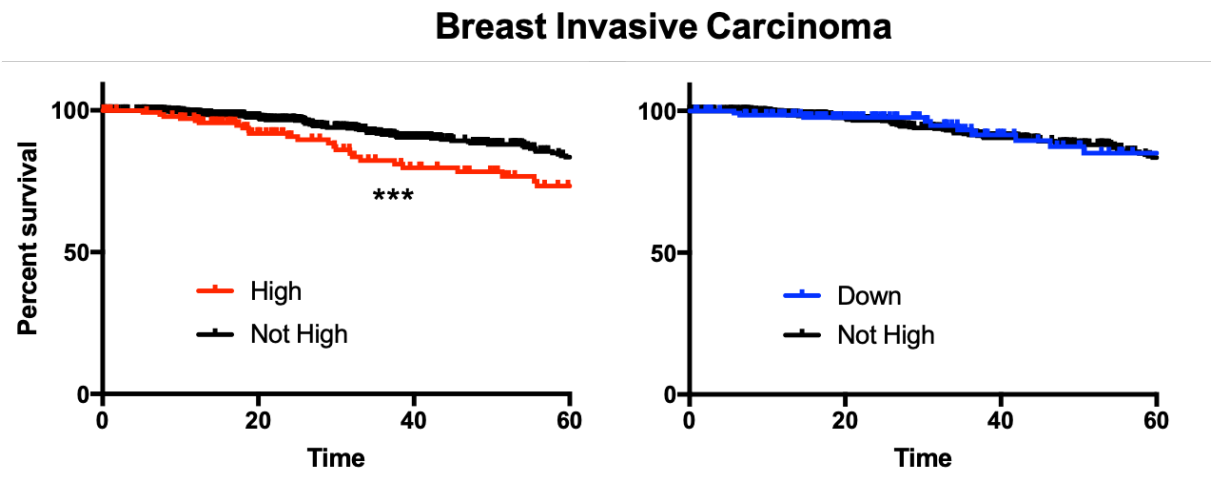


Figure 15 Patients overall survival in Breast Invasive Carcinoma Kaplan-Meier survival curves showing: Left, overall survival (in months) of Breast Invasive Carcinoma patients with high mRNA expression of either POLR2A-D genes (altered group, red) compared to all other patients, all non-upregulated patients (unaltered group, black). Right, overall survival (in months) of Breast Invasive Carcinoma patients with low mRNA expression of either POLR2A/D genes (altered group, blue) compared to all non-downregulated patients (unaltered group, blue). Source: www.cbioportal.org

We then expanded our study to the other cancer types that included mRNA expression data from RNA Seq, to assess whether we could observe the same phenotype as in BIC patients. We firstly queried for any deregulation affecting each of the four Pol II subunits. Among the many datasets analysed, at the time the data were analysed specifically we observed that upregulation of the four subunits were showing a prominent correlation with poor cancer patients' outcomes also in other tumour types, among which: Acute Myeloid Leukemia (AML, TCGA, Firehose Legacy), Liver Hepatocellular Carcinoma (LHC, TCGA, Firehose Legacy), Head and Neck Squamous Cell Carcinoma (HNSCC, TCGA, Firehose Legacy), Brain Lower Grade Glioma (BLGG, TCGA, Firehose Legacy) and Skin Cutaneous Melanoma (TCGA, Firehose Legacy)

(Figure 16). When we analysed the oncoprint of all the deregulations in these tumour types, we found that overall more than 15% of patients for all tumour types were presenting an upregulation in either one of the Pol II four subunits (Figure 17).

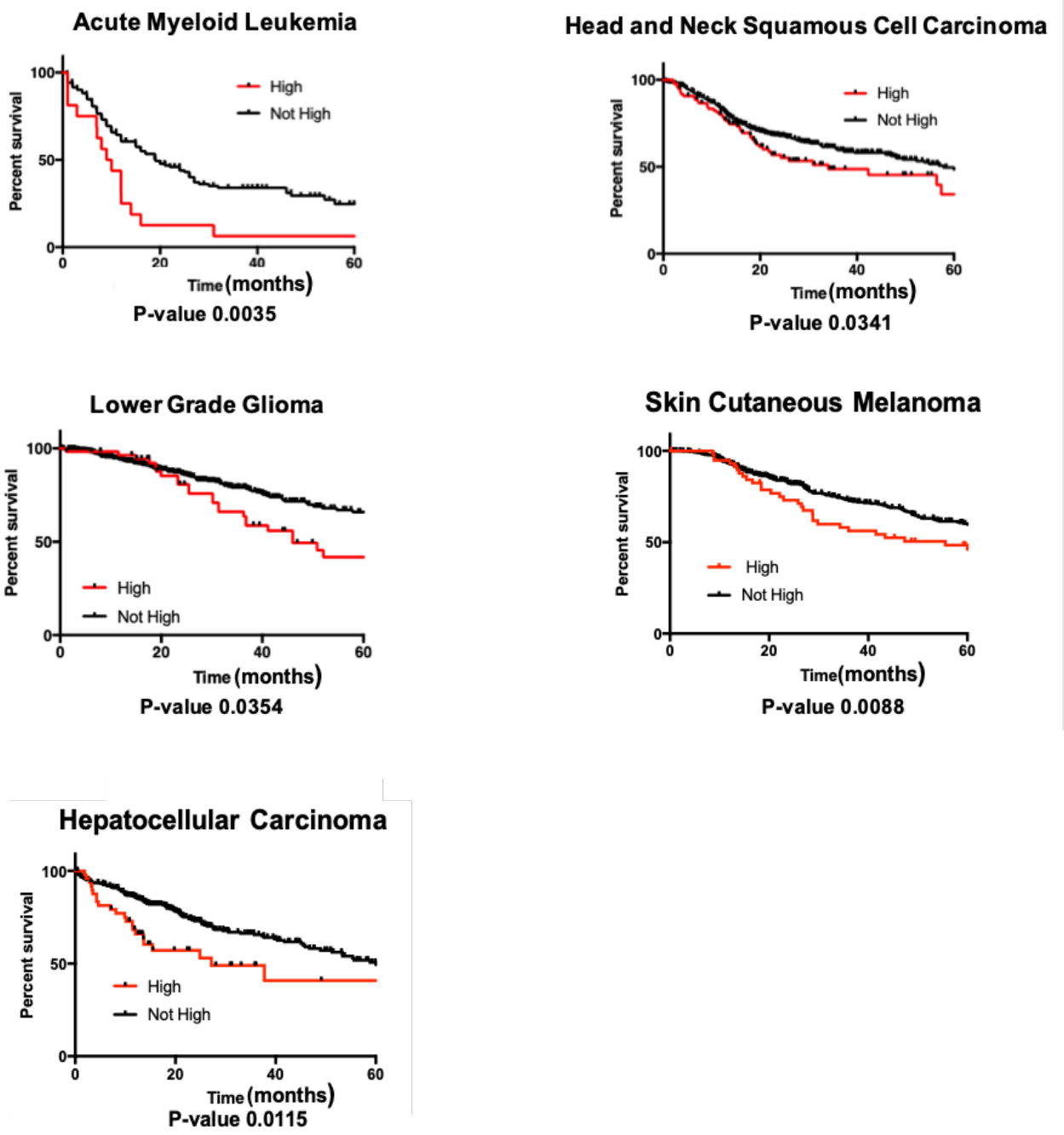


Figure 16 Kaplan-Meier survival curves in patients that overexpress at least one among POLR2A-D subunits (altered, red) and control patients (unaltered, black) in the listed tumour types. Log-Rank test p values shown in figure.

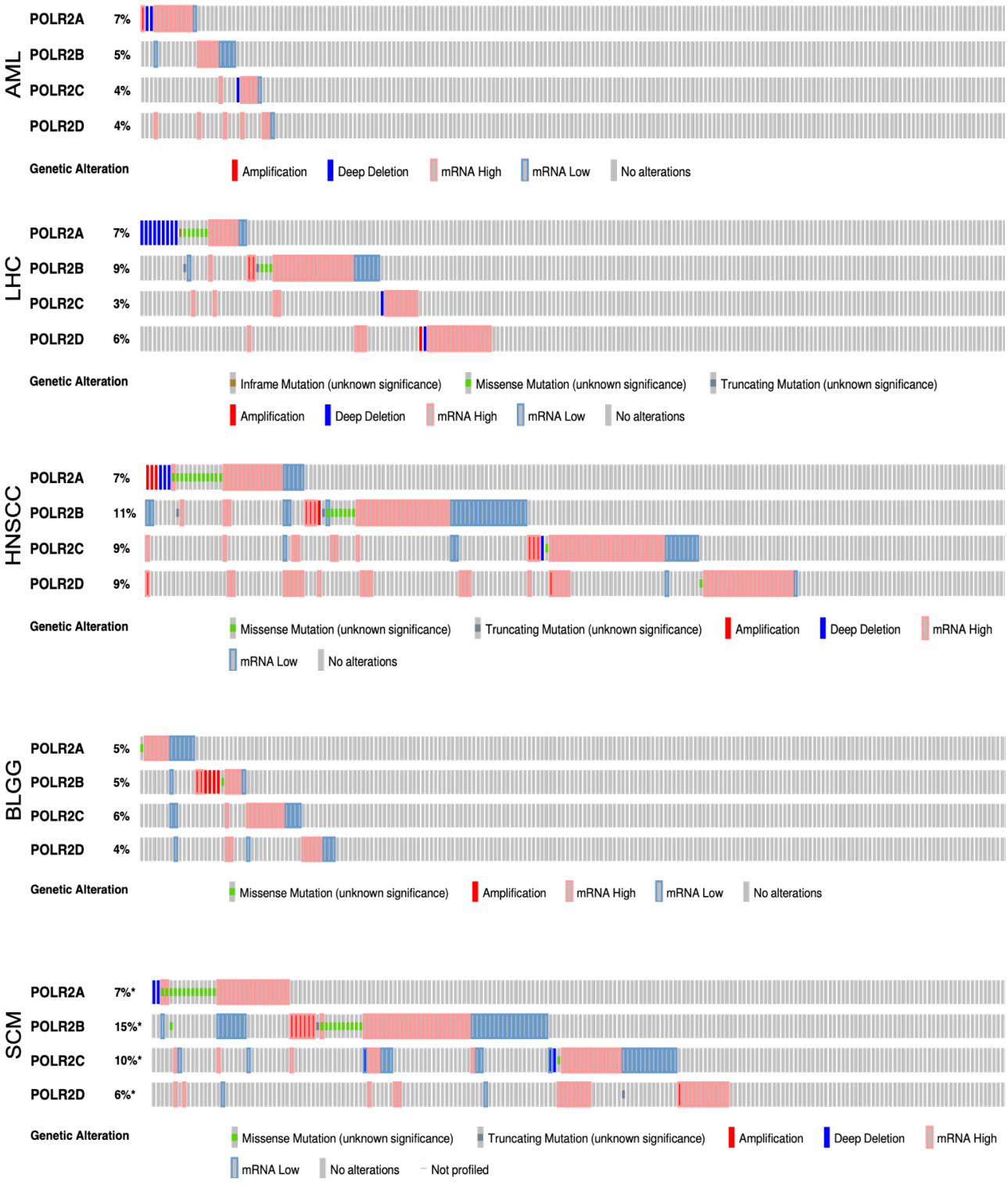

AML
POLR2A 7%
POLR2B 5%
POLR2C 4%
POLR2D 4%
Genetic Alteration
Amplification
Deep Deletion
mRNA High
mRNA Low
No alterations
LHC
POLR2A 7%
POLR2B 9%
POLR2C 3%
POLR2D 6%
Genetic Alteration
Inframe Mutation (unknown significance)
Missense Mutation (unknown significance)
Truncating Mutation (unknown significance)
Amplification
Deep Deletion
mRNA High
mRNA Low
No alterations
HNSCC
POLR2A 7%
POLR2B 11%
POLR2C 9%
POLR2D 9%
Genetic Alteration
Missense Mutation (unknown significance)
Truncating Mutation (unknown significance)
Amplification
Deep Deletion
mRNA High
mRNA Low
No alterations
BLGG
POLR2A 5%
POLR2B 5%
POLR2C 6%
POLR2D 4%
Genetic Alteration
Missense Mutation (unknown significance)
Amplification
mRNA High
mRNA Low
No alterations
SCM
POLR2A 7%*
POLR2B 15%*
POLR2C 10%*
POLR2D 6%*
Genetic Alteration
Missense Mutation (unknown significance)
Truncating Mutation (unknown significance)
Amplification
Deep Deletion
mRNA High
mRNA Low
No alterations
Not profiled

Figure 17 Oncoprint of POLR2A-D deregulations in Acute Myeloid Leukemia (AML), Liver hepatocellular carcinoma (HCC), Head and Neck Squamous Cell Carcinoma (HNSCC), Brain Lower Grade Glioma (BLGG) and Skin Cutaneous Melanoma (SCM), using for all of them data from TCGA Firehose Legacy studies. The settings used for mRNA are: mRNA expression – Z-score relative to diploid samples (log RNA Seq V2 RSEM), Z-score +/- 2 as default. Source: www.cbioportal.org

The oncoprints in the different tumour types were showing that upregulations of two or more subunits in the same patient were a rare event. Therefore, we analysed into more details the relationship between the expression levels of POLR2A-D across the six tumour types where we found a correlation with patients' survival. To do this we performed a Spearman correlation analysis of the transcription levels of each subunit against the other 3 in all the cancer patients of the six tumours analysed in this chapter. Interestingly, this analysis showed that overall the expression levels of POLR2A-D were not consistently linked to each other, and the landscape of the different deregulations suggests that the expression of the four genes might be under the control of different factors and independent from each other (Figure 18)

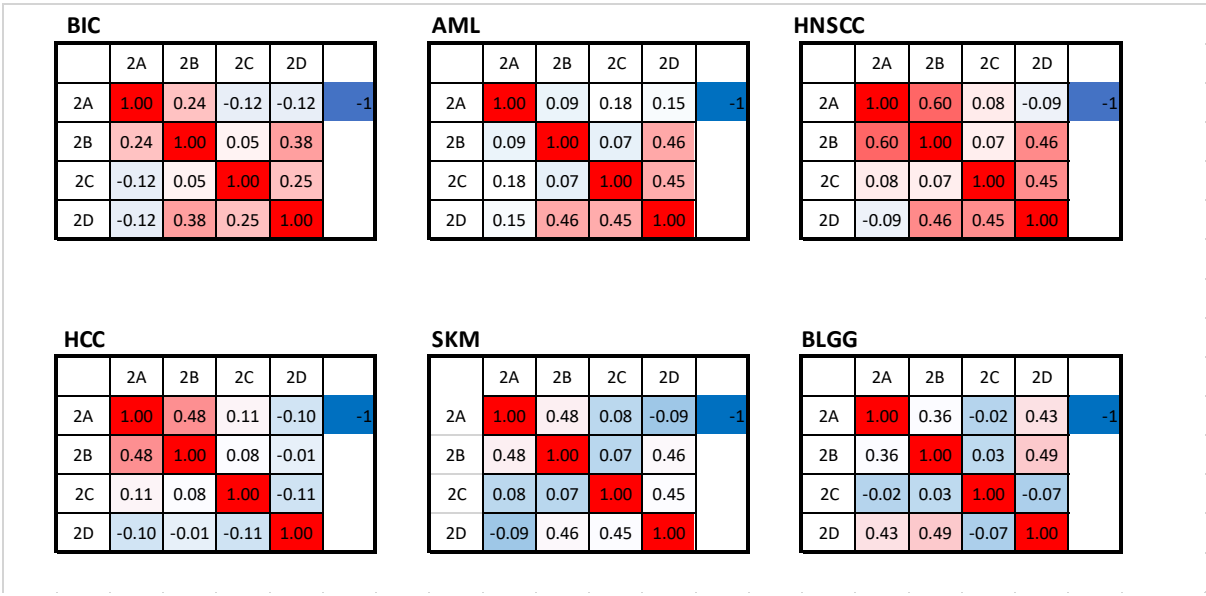


Figure 18 POLR2A-D subunits expression correlation analysis. Spearman correlation analysis of the transcription levels of each of the four largest subunit of Pol II against the other three in BIC, AML, HNSCC, HCC, SKM and BLGG, using for all of them RNA-Seq data from TCGA Firehose Legacy studies, showing no correlation between the expression levels of the single

3.1.1.1 Copy Number Alteration and Mutation Count analysis

We hypothesised that the upregulation of the four largest subunits of Pol II could undermine the smooth course of transcription, directly interfering with transcription regulation and resulting in defective transcription. For example, impaired transcription could result in increased stalling of Pol II or defective termination, increasing chances of transcription replication collision, R-loop formation and consequent replication stress (Hamperl & Cimprich, 2016; García-Muse & Aguilera, 2016; Gómez-González & Aguilera, 2019). Indeed, one well known characteristic of defective transcription is also an increase in genome instability. To assess whether patients with high mRNA levels of at least one among the four major subunits of Pol II (altered group) showed also increased genome instability, we checked for Copy

Number Alterations (CNAs) levels, referred to as Fraction genome altered on cBioPortal, and mutation counts, in all six tumour types.

Interestingly, CNA levels were always significantly higher in the altered group of patients with upregulations in any of the Pol II subunits compared to the unaltered control group (Figure 19). However, when we analysed the mutation count, although apparently following the same trend of the CNA levels, only the BIC, AML and BLGG cancer studies showed a significant difference between the two patients groups (Figure 20).

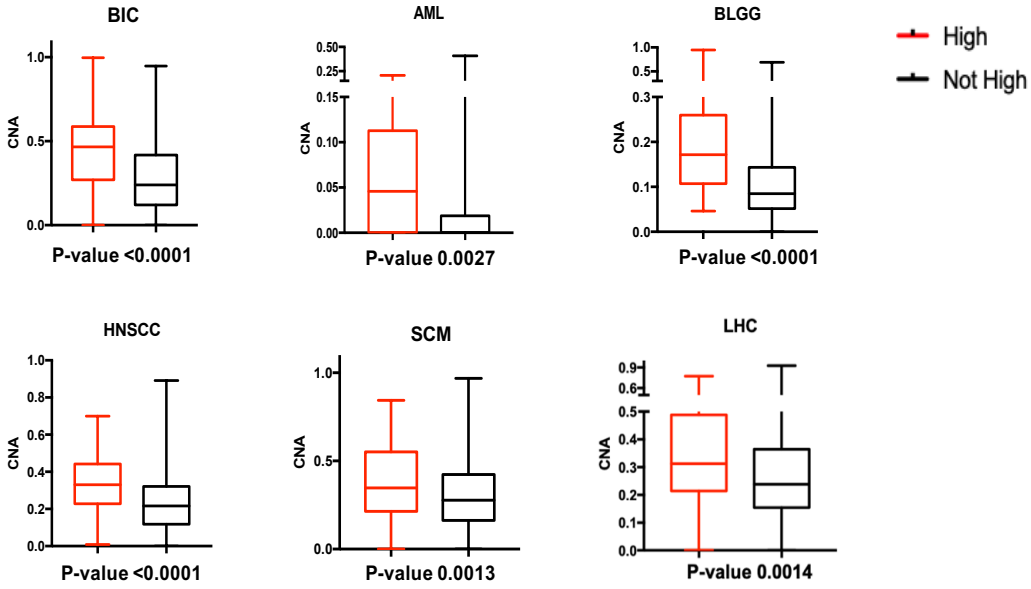


Figure 19 Copy Number Alterations levels comparison between patients with upregulation of at least one among POLR2A-D subunits (High) and control patients (Not High). Statistic is non-parametric t-test, p-value shown in figure.

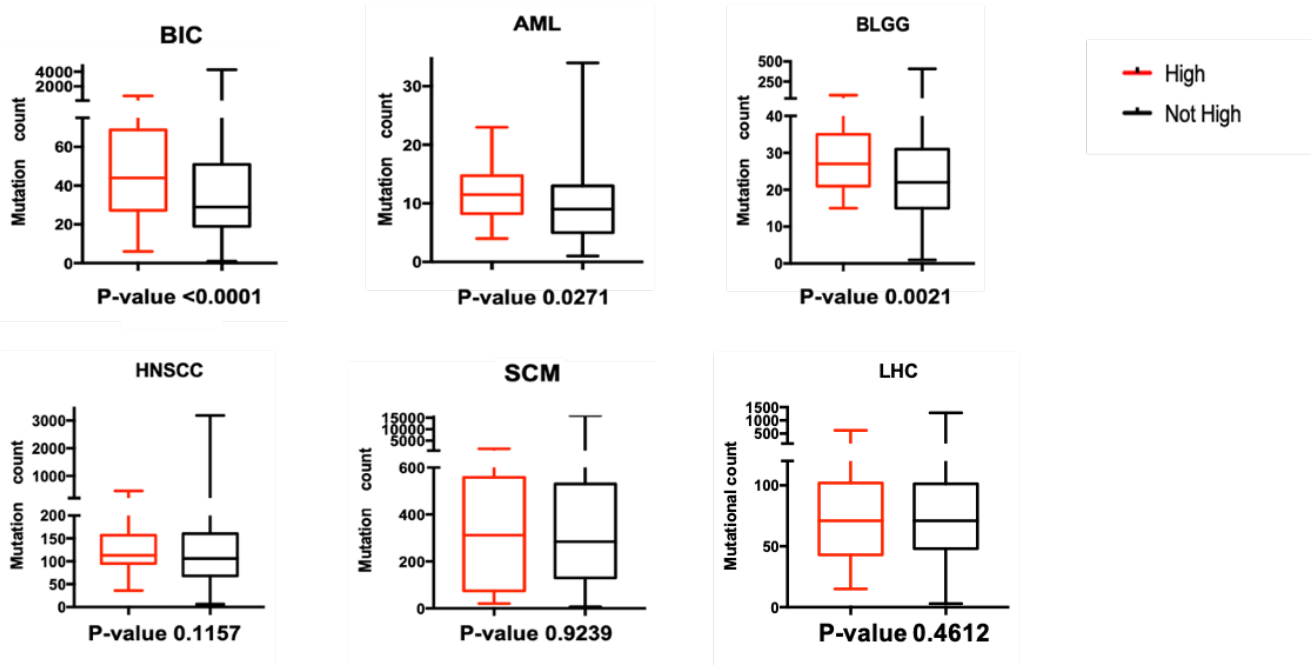


Figure 20 Mutation count comparison between patients with over-expression of at least one among POLR2A-D subunits (High) and control patients (Not High). Statistic is non-parametric t-test, p-value shown in figure.

3.1.1.2 Co-occurrence with the most mutated oncogenes from IntOgen

Given the results showing how PolUp is associated to poorer survival in patients and to increased genome instability, we extended our investigation performing further analysis. In particular, we wanted to better characterize whether the phenotypes were indeed primarily correlated with the high expression of the major Pol II subunits, and not related to a collateral oncogenic event. To do this, first of all we obtained a list of the top ten most mutated cancer driver genes from the IntOgen web portal, a platform that collects information from different cancer projects (referred to as cohorts) and identifies all the cancer driver genes across different tumours (Gonzalez-Perez et al. 2013) (Table 2).

Symbol	Mutations	Samples	Samples (%)
TP53	8,504	7,695	27.41
KRAS	2,378	2,234	7.96
PIK3CA	2,453	2,157	7.68
BRAF	1,39	1,222	4.35
LRP1B	3,871	1,019	3.63
KMT2D	2,02	1,01	3.6
APC	1,761	994	3.54
ARID1A	1,282	954	3.4
KMT2C	1,887	924	3.29
PTEN	1,143	904	3.22

Table 2 Most recurrently mutated cancer driver genes The top 10 most mutated cancer driver genes list downloaded from IntOgen, ranked according to Samples. Samples (%) = aggregated % of samples in cohorts were the gene was dected as driver. Source: www.intogen.org

Therefore, on cBioPortal we firstly identified patients bearing a deregulation in any of the ten oncogenes, selecting specifically where possible for activating mutations or loss of function ones, according to the oncogenes' mode of action. This procedure was done specifically in the same TCGA cancer studies analysed before: BIC, AML, BLGG, HNSCC, SCM and LHC. First of all, we evaluated whether patients with PolUP had an enrichment for any of the ten oncogenes. This result showed that PolUp is rarely associated with significantly increased mutations in any of the top 10 most mutated oncogenes according to IntOgen, except of TP53 found significantly more mutated in all tumours but BLGG (Table 3).

	BIC	BLGG	AML	HNSCC	SCM	LHC
TP53	✓	✗	✓	✓	✓	✓
KRAS	✓	✗	✗	✗	✗	✗
PIK3CA	✗	✗	✗	✗	✗	✗
BRAF	✓	✗	✗	✗	✗	✗
LRP1B	✗	✗	✗	✗	✗	✗
KMT2D	✓	✗	✗	✗	✓	✗
APC	✓	✓	✗	✗	✗	✗
ARID1A	✗	✗	✗	✗	✗	✗
KMT2C	✗	✗	✗	✗	✗	✗
PTEN	✓	✗	✗	✗	✗	✗

Table 3 Mutations affecting the top 10 most mutated cancer driver genes according to IntOgen (www.intogen.org) in BIC, BLGG, AML, HNSCC, SCM and LHC. **X**= not significantly enriched in PolUp, ✓= significantly enriched in PolUp, pvalue < 0.05.

Next, we aimed to determine whether PolUP was having an impact on the survival of patient on top of the impact of the oncogene deregulation alone. For most of the oncogenes taken under consideration, there was no significant difference in the survival curves of the two distinct groups of patients, indicating that while deregulations in the oncogenes were potentially important for the onset of the diseases, they rarely associated with more aggressive subtypes (Table 4).

	BIC	BLGG	AML	HNSCC	SCM	LHC
TP53	✗	✗	✗	✗	✗	✓
KRAS	✗	✗	✗	✗	✓	✗
PIK3CA	✗	✗	✗	✗	✗	✗
BRAF	✗	✗	✗	✗	✗	✗
LRP1B	✗	✗	✗	✗	✗	✓
KMT2D	✗	✗	✗	✗	✗	✗
APC	✗	✗	✗	✓	✗	✗
ARID1A	✗	✗	✗	✓	✗	✗
KMT2C	✗	✗	✗	✗	✗	✗
PTEN	✗	✓	✗	✗	✗	✗

Table 4 The top 10 most mutated cancer driver genes according to IntOgen (www.intogen.org) impact on patients survival in BIC, BLGG, AML, HNSCC, SCM and LHC. **X**= no significant difference from unaltered patients, ✓= significant difference from unaltered patients. Log Rank t-test, p-value <0.05.

We extended this analysis selecting all those patients with high mRNA expression of at least one among POLR2A-D, obtaining a list of patients with both a deregulation in the oncogenes and subunits high expression. As before, we compared the overall survival of patients carrying a deregulation in the oncogene alone with the overall survival of those patients carrying both a deregulation in the oncogene and Pol II subunits (Table 5). In agreement with what was previously observed, in many cases no significant difference was found, although we had instances where there were no Pol II subunits upregulated in the sub-lists of patients (for example in BLGG patients carrying a deregulation in *PTEN*).

	BIC	BLGG	AML	HNSCC	SCM	LHC
TP53	✓	✓	✗	✗	✗	✓
KRAS	✗	✓	✗	✗	O	✓
PIK3CA	✗	✗	✗	✗	✗	✗
BRAF	✗	✗	✗	✗	✗	✓
LRP1B	✗	✗	✗	✗	✗	O
KMT2D	✗	✗	✗	✗	✗	✓
APC	✗	✗	✗	O	✗	✗
ARID1A	✗	✗	✗	O	✓	✗
KMT2C	✓	✗	✗	✗	✗	✗
PTEN	✓	O	✗	✗	✗	✓

Table 5 Comparison between survival rates of patients carrying a deregulation in each of the top 10 most mutated cancer driver genes according to IntOgen (www.intogen.org) and patients carrying both a deregulation in the oncogene and a PolUp phenotype in BIC, BLGG, AML, HNSCC, SCM, LHC. **X**= no significant difference ✓ = significant difference, O= no patients found. patients Log Rank t-test, p-value <0.05.

Importantly however, where we could assess the impact of Pol II subunits upregulation in patients in the context of deregulated oncogenes, like in the specific case of deregulations of TP53 in BLGG, we observed not only a significant decrease in overall survival (Figure 21A), but also an increase in CNA and mutation counts levels (Figure 21B). Strikingly, whenever we observed a reduction in survival for the one of the oncogenes identified on IntOgen alone, there was either a further significant drop in survival when patients had also an upregulated Pol II subunits, or no Pol II subunits upregulation were found in patients (Table 5).

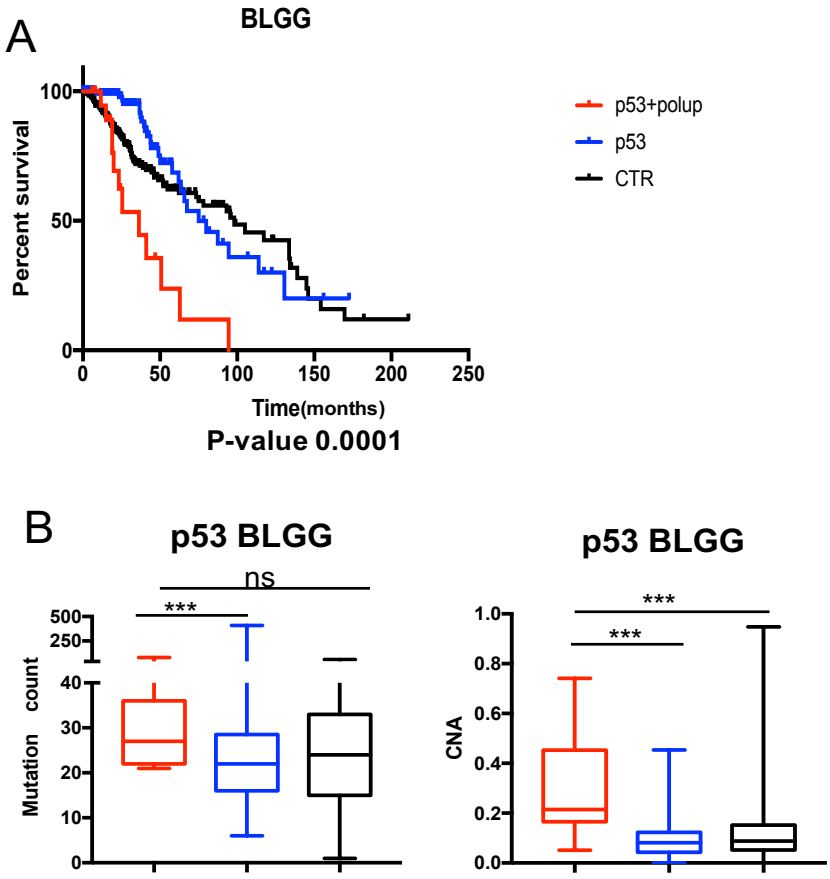


Figure 21 A) Kaplan-Meier survival curves in p53 deficient BLGG (p53) alone or combined with PolUp (p53+polup) vs control patients (CTR). Log-Rank test p values shown in figure. B) Mutation count and copy number alteration level in "p53", "p53+polup" and "CTR" Statistical analysis is non-parametric t-test, p-value < 0.001 = ***, p-value > 0.05 = ns

To exclude that results were affected by the low number of patients in the selected cancer studies, and the risk that we were not selecting the most appropriate oncogenic drivers for our specific cancer sets, we decided to repeat this analysis with the top 5 most mutated oncogenes specific for each tumour type, identified as before using IntOgen website (Table 6). As before, we evaluated whether patients with PolUP had an enrichment for any of the selected oncogenes. This result showed that PolUp is rarely associated with significantly

increased mutations. However, TP53 is found to be significantly more mutated in all tumours but BLGG (Table 7).

	Symbol	Mutations	Samples	Samples (%)
BIC	PIK3CA	359	327	33.61
	TP53	294	287	29.5
	CDH1	72	94	9.66
	GATA3	22	94	9.66
	KMT2C	86	79	8.12
AML	DNMT3A	13	10	7.14
	NPM1	1	9	6.43
	RUNX1	7	9	6.43
	TP53	11	8	5.71
	IDH2	8	8	5.71
BLGG	IDH1	395	395	77.6
	TP53	292	236	46.37
	ATRX	133	137	26.92
	CIC	82	88	17.29
	PIK3CA	35	35	6.88
HNSCC	TP53	380	313	62.35
	FAT1	115	110	21.91
	LRP1B	123	94	18.73
	PIK3CA	88	86	17.13
	NOTCH1	93	83	16.53
SCM	BRAF	269	220	47.21
	FAT3	191	129	27.68
	NRAS	124	117	25.11
	PTPRB	118	91	19.53
	TRRAP	79	61	13.09
LHC	TP53	98	93	25.83
	CTNNB1	97	92	25.56
	ARID1A	23	28	7.78
	KMT2D	23	20	5.56
	AXIN1	21	19	5.28

Table 6 The top 5 most mutated cancer driver genes list downloaded from IntOgen, ranked according to Samples (%), after selecting for the specific TCGA study of the cancer study of interest (BIC, AML, BLGG; HNSCC, SCM, LHC)

BIC		**AML**		**SCM**	
PIK3CA	✗	DNMT3A	✗	BRAF	✗
TP53	✓	NPM1	✗	FAT3	✗
CDH1	✓	RUNX1	✗	NRAS	✗
GATA3	✗	TP53	✓	PTPRB	✗
KMT2C	✗	IDH2	✗	TRRAP	✗
BLGG		**HNSCC**		**LHC**	
IDH1	✗	TP53	✓	TP53	✓
TP53	✗	FAT1	✗	CTNNB1	✗
ATRX	✗	LRP1B	✗	ARID1A	✗
CIC	✗	PIK3CA	✗	KMT2D	✗
PIK3CA	✗	NOTCH1	✗	AXIN1	✗

Table 7 Mutations affecting the top 10 most mutated cancer driver genes according to IntOgen (www.intogen.org) in BIC, BLGG, AML, HNSCC, SCM and LHC. X= not significantly enriched in PolUp, ✓= significantly enriched in PolUp, pvalue < 0.05.

We Again, we firstly evaluated the impact on patients' survival rate when a deregulation in each of the oncogenes was present (Table 8). Secondly, we checked for differences in patients' survival status between the group of patients bearing the deregulated oncogene alone, and the group of patients bearing both the deregulated oncogene and Pol II subunits up-regulation (Table 9).

BIC		**AML**		**SCM**	
PIK3CA	✗	DNMT3A	✓	BRAF	✗
TP53	✗	NPM1	✗	FAT3	✗
CDH1	✗	RUNX1	✓	NRAS	✗
GATA3	✗	TP53	✗	PTPRB	✓
KMT2C	✗	IDH2	✗	TRRAP	✗
BLGG		**HNSCC**		**LHC**	
IDH1	✗	TP53	✗	TP53	✓
TP53	✗	FAT1	✗	CTNNB1	✗
ATRX	✗	LRP1B	✗	ARID1A	✗
CIC	✗	PIK3CA	✗	KMT2D	✗
PIK3CA	✗	NOTCH1	✗	AXIN1	✗

Table 8 Comparison between survival rates of patients carrying a deregulation in each of the top 5 most mutated cancer driver genes according to IntOgen (www.intogen.org) and control patients in BIC, BLGG, AML, HNSCC, SCM, LHC. **X**= no significant difference ✓= significant difference,. patients Log Rank t-test, p-value <0.05.

BIC		**AML**		**SCM**	
PIK3CA	✗	DNMT3A	✓	BRAF	✗
TP53	✓	NPM1	✓	FAT3	✗
CDH1	✗	RUNX1	O	NRAS	✗
GATA3	✗	TP53	✗	PTPRB	✓
KMT2C	✓	IDH2	✓	TRRAP	✗
BLGG		**HNSCC**		**LHC**	
IDH1	✓	TP53	✗	TP53	✓
TP53	✓	FAT1	✗	CTNNB1	✗
ATRX	✓	LRP1B	✗	ARID1A	✗
CIC	✗	PIK3CA	✗	KMT2D	O
PIK3CA	✗	NOTCH1	✗	AXIN1	✗

Table 9 Comparison between survival rates of patients carrying a deregulation in each of the top 5 most mutated cancer driver genes according to IntOgen (www.intogen.org) and patients carrying both a deregulation in the oncogene and a Pol II subunits upregulation phenotype in BIC, BLGG, AML, HNSCC, SCM, LHC. **X**= no significant difference ✓= significant difference, O= no patients found. patients Log Rank t-test, p-value <0.05.

In agreement with the previous analysis, it was evident that when there was a significant difference between patients' survival it was associated with the concomitant deregulation of the oncogene and the presence of Pol II subunits up-regulation (Figure 22A). We also investigated CNAs and mutation counts, which resulted to be higher in PolUp compared to patients carrying a mutation in the cancer driver gene alone (Figure 22B).

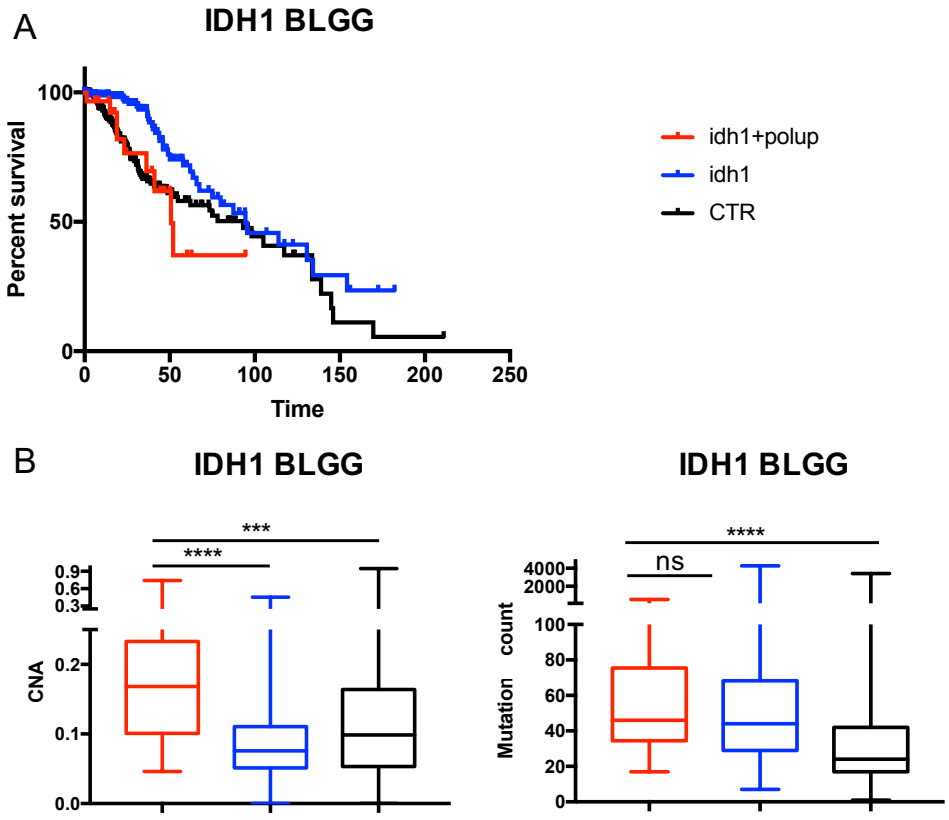


Figure 22 Survival, CAN and Mutation counts in BLGG A) Kaplan-Meier survival curves in IDH1 deficient BLGG patients (idh1) alone or combined with PolUp (idh1+polup) vs control patients (CTR). Log-Rank test p values shown in figure. B) Mutation count and copy number alteration level in "p53", "p53+polup" and "CTR" Statistical analysis is non-parametric t-test, p-value <

3.2 Discussion

While characterising the role of transcription factors, and their deregulation in diseases, has become central to understand transcription regulation, the study of each component of the RNA Pol II complex yet remains elusive. Our cancer genomic data is the first step in evaluating the impact of the four largest subunits of RNA Pol II, identifying for the first time that they share a common negative correlation with cancer patients' survival and increased genome instability across a range of different tumours.

First of all, we decided to focus on the 4 largest subunits because being the largest, we hypothesised that they could be a platform for interactions with transcription factors, perhaps also outside and independently of when included in the Pol II complex. Indeed, structural studies have shown that TF-Pol II interactions can be mapped to specific regions of Pol II subunits (Bernecky et al. 2017; Compe et al. 2019; Ghosh et al. 2011; Hantsche and Cramer 2017; He et al. 2016; Liu et al. 2018; Nojima et al. 2018). .However, it has to be commented on the fact that deregulations of other subunits form the ones investigated in this analysis and in this thesis (i.e., POLR2A-D), may have similar phenotypes to the ones we observed. For example, this is the case of POLR2E and POLR2K (reviewed in Muste Sadurni and Saponaro 2023), which show poorer survival in specific tumours (data not shown). In our analysis however, we were interested in looking at the impact of Pol II subunits deregulations as an ensemble, as this allowed us to have enough patients to expand our analysis to multiple tumour types, so that our conclusion would be more robust. Therefore, we focused our analysis only on the four largest subunits of the complex because when we expanded this analysis including more subunits, the correlation with cancer patients' survival was lost (data not shown). The correlation with patients' survival (Figure 15; Figure 16) in the BIC study and

subsequently confirmed in AML, BLGG, LHC, HNSCC, SCM, suggested that the correlation of Pol II subunits upregulation might not be specifically associated with a single tumour type.

High transcription activity is a common feature of more aggressive tumours (Galbraith et al. 2019; Zatzman et al. 2022). Therefore, it could have been possible that what we observed was just an indirect consequence of the fact that we were selecting patients with high transcription activity, rather than because of the specific Pol II single subunit upregulation. We hypothesised that if that was the case, we would have observed multiple if not all Pol II subunits expression levels correlating together, as an indication that the whole complex might be potentially overexpressed. In agreement with our starting hypothesis, the subunits expression levels are not significantly correlating, nor consistently upregulated together in all tumour types studied (Figure 18). This analysis strengthened our believe that the upregulation of the single subunit was on its own correlated with the phenotype observed. Indeed, the fact that their up regulation is not concurrent and that there are many cases in which their upregulations are mutually exclusive, could indicate that the mechanism(s) driving the upregulation events are different among the four subunits, and perhaps specific for each one of them for each tumour type.

We hypothesised therefore that the upregulation of the single subunits could translate in defective transcription and consequentially affecting genome integrity, as mentioned before. We hypothesised that the overexpressed subunits could interfere with the correct interaction and recruitment of TFs and Pol II interactors at transcribing Pol II, thus impeding the smooth course of transcription. Indeed, an altered recruitment of specific TFs to transcribing Pol II could lead to increased stalling and pausing of Pol II at the transcription start site or defective termination (Cui and Denis 2003; Jonkers and Lis 2015; Core et al. 2008; Mayer et al. 2018;

Morales et al. 2016). As transcription and DNA replication have been shown to occur together on the same DNA template, any transcription defect could increase chances of transcription and replication collisions, resulting in replication stress and DNA damage (Hamperl et al. 2017; Kotsantis et al. 2016). However, it has also been shown that Pol II is a DNA damage sensor and therefore it could also be speculated that defective transcription or impaired Pol II, could fail in DNA damage sensing role for subsequent repair (Gómez-González and Aguilera 2019; Shin et al. 2016; Wang et al. 2023). Altogether, we thought that if upregulation of Pol II subunits could impact on genome instability, this could be potentially observed in patients as well. Strikingly, PolUp patients showed significantly increased genome instability, in the form of "fraction genome altered" and "mutations counts". It appears that the phenotypes observed in patients were therefore dependent on PolUp. At the same time, we wanted to exclude the possibility that PolUp was linked to another oncogenic mutation and/or deregulation, that could have been then responsible for the phenotypes observed in patients. Consistent with our hypothesis we found no recurrent mutation in the top 10 oncogenes according to intOgen across the tumour types analysed, apart from TP53. In the case of BIC we observed 7 oncogenes to be significantly more mutated in PolUp patients compared to unaltered patients (Table 3; Table 7). However, we can't exclude that this high level of overlap of PolUP with other oncogenic deregulations might have been found only in BIC because it has the largest cohort of patients of all TCGA studies and/or because of the high recurrence of those oncogenic deregulations specifically in BIC. Importantly, however, the phenotypes observed in PolUp patients (i.e., reduced survival and increased genomic instability) were confirmed even in patients baring a deregulation in drivers' oncogenes, showing that even in patients with mutated oncogenes the addition of Pol II subunits upregulation further decreased

patients' survival and increased genome instability. For example this is the case of TP53 mutated altered patients in BLGG (Figure 21).

Altogether, this cancer genomic data analysis shows that poor patients' survival and increased genome instability link strongly to a potential direct consequence of the overexpressed subunits as they do not seem to depend on co-occurring deregulations in cancer driver genes, indicating that PolUp could constitute a new Bio Marker for high-risk patients.

The limitation of these analysis lays in the fact that Pol II subunits upregulation was analysed as an ensemble, to increase the number of patients and therefore the robustness of the analysis. Indeed, while patients survival and genome alteration status were consistent among different tumour types, other clinical features such as tumour stage were not included in the analysis. Being the expression of the four subunits independent one from another (Figure 18), we anticipate that there could be differences between each subunit upregulation. Therefore, it would be interesting to confirm the phenotypes observed for single subunits and to expand the single subunit analysis to other tumour types as well. However, this analysis would need a larger number of patients, and at the time would perhaps only possible in BIC as it has the highest numbers of patients overall, however with the limitations and the caveats listed earlier.

3.2 Overexpression of single subunits induces genome instability

3.2.1 Establishment of doxycycline inducible cell lines

In order to investigate the impact and the consequences of the over expression of the single RNA Pol II major subunits, we established stable clones expressing either POLR2A, POLR2C or POLR2D in a doxycycline inducible system, using HeLa T-REx cells as host (Figure 23).

This system relies on two fundamental components: the inducible expression plasmid and the regulatory vector. The first, contains the genes of interest under the control of CMV promoter into which have been inserted two copies of the tet Operator 2 (TetO2) sequence. Each copy of TetO2 is bound by two Tet repressor (TetR) molecules (Figure 23B). pcDNA6/TR is instead the regulatory vector which expresses the TetR (Figure 23A). The transcription of the genes under the control of the inducible promoter is allowed only when tetracycline/doxycycline are

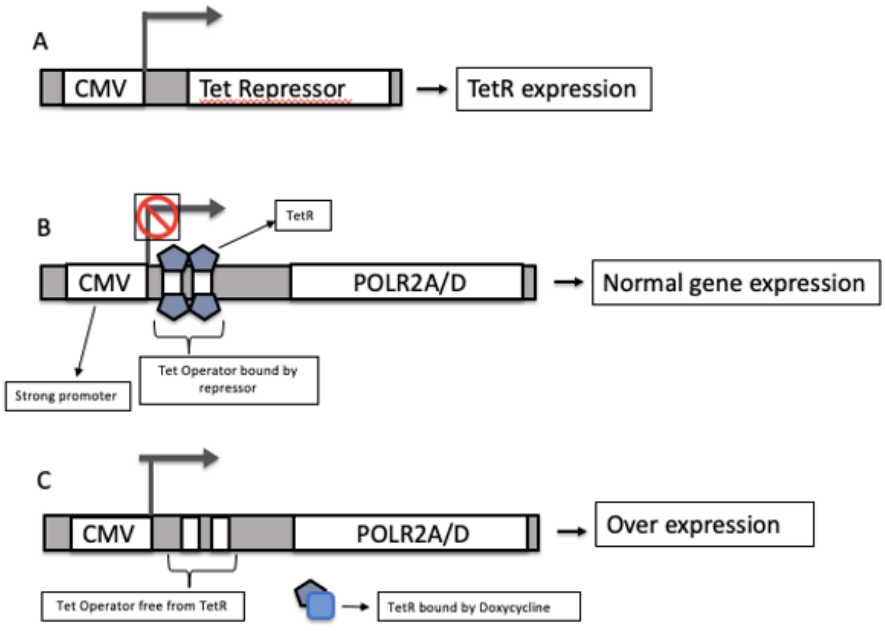


Figure 23 Doxycycline inducible expression system in HeLa T-REx: A) the first component of the system is the Tet repressor (TetR) B) the second component of the system is the operator sequence flanking CMV promoter. When bound by TetR, the operator blocks transcription. C) Upon doxycycline addiction, TetR is removed by preferentially binding Doxycycline and transcription can initiate.

provided to cells in culture, since they bind the TetR and abolish its binding to the TetO2 (Figure 23C) (Kallunki et al. 2019). Doxycycline and tetracycline are interchangeable, however we used doxycycline since it is more stable with a half-life double that of tetracycline.

The first cell line generated was for the overexpression of the POLR2C subunit (referred to as POLR2C cell line). We built a dose response curve testing different doxycycline concentrations and checking the over expression of the subunit by real time PCR after 48h induction (Figure 24A). We used decreasing concentrations of doxycycline, in order to identify the lowest working concentration to achieve efficient and consistent overexpression. We found that 5ng/mL of doxycycline were enough to induce the over expression of the POLR2C subunit. Therefore, we tested whether the same levels of expression could be achieved inducing with doxycycline for only 24h, but it was not the case (Figure 24A). Finally, we confirmed our qPCR data by western blot with antibodies against RPB3 (Figure 24B).

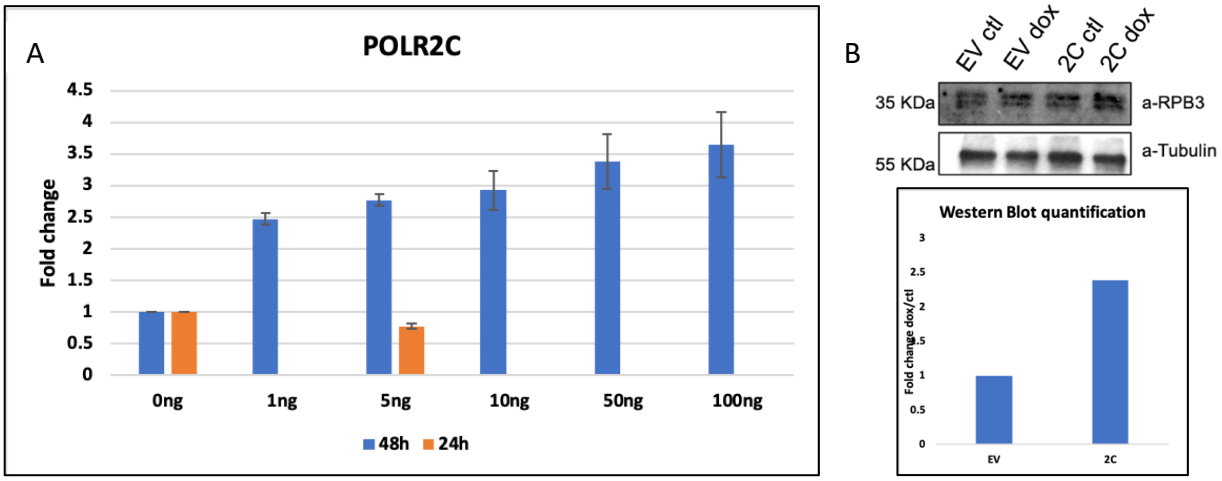
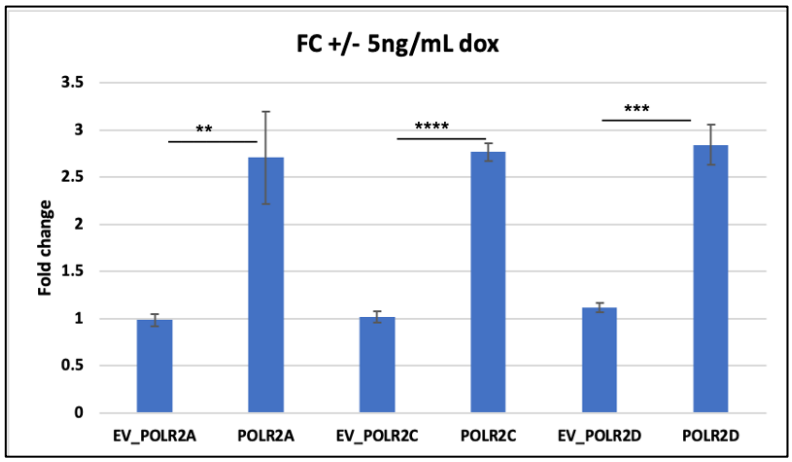


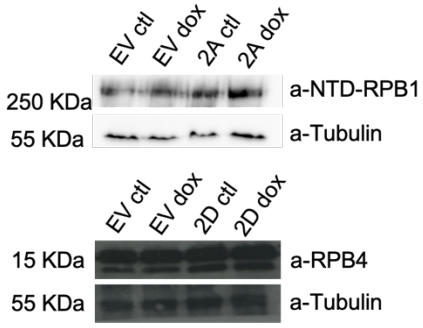
Figure 24 Fold Change increase after Doxycycline induction: A: qPCR analysis of plasmids overexpression in POLR2C cell line. Inducible cell line was treated for 48h with increasing doses of doxycycline and for 24h with 5ng/mL doxycycline, N=3. B: Western blot showing the overexpression of POLR2C with a-RPB3 antibody and western blot quantification; tubulin serves as loading control, N=2. Error bars show standard error mean.

Once we established the conditions for the overexpression of POLR2C, we applied these same conditions also to the other cell lines generated (POLR2A and POLR2D cell lines, each for the overexpression of the corresponding subunit). As anticipated, we used the same host cell line to generate all of our cell lines, therefore it was not surprising that the same concentration of 5ng/mL doxycycline worked for all of them. We also made sure that the overexpression was only observed in our inducible cell lines and that doxycycline on its own was not altering the expression levels of any of the three subunits tested in the control empty vector (EV) cell line, confirming the correct functioning of the inducible system and excluding any indirect effect of doxycycline on the endogenous expression of the subunits (Figure 25A). As before, we confirmed the qPCR by western blot (Figure 25B)

A



B



C

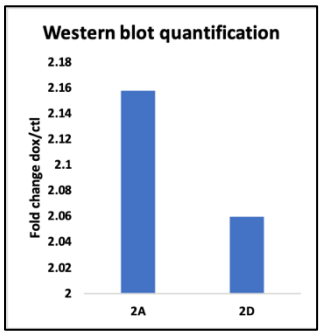


Figure 25 Fold Change increase after Doxycycline induction A) qPCR analysis of overexpression of POLR2A,2C and 2D in the corresponding cell lines (POLR2A, POLR2C and POLR2D) and in the EV cell line after 48h induction with 5ng/mL doxycycline N=3. B) western blot in EV, POLR2A and POLR2D cell line after 48h induction with 5ng/mL doxycycline; Tubulin serves as loading control. C) Western blot quantification normalized to EV, N=2. Error bars show standard error mean. **= p value <0.01, ***= p value <0.001, ****= p value <0.0001.

To understand whether our system could be used to recapitulate the phenotypes observed in the Pol Up patients, we calculated how much subunits are upregulated in patients compared to the average expression level. To do so, we retrieved the mRNA-Seq data from TCGA of BIC, AML, BLGG and LHC studies and we calculated the average expression level (average RPKM) of *POLR2A-D* in all patients. We then selected the patients with a z-score > 2 (i.e., patients with expression levels of at least 2 standard deviations higher than the average), therefore identifying the Pol Up patients. Finally, we calculated the fold change of expression between Pol Up patients and the average expression in all other patients. Importantly, we were able to show that the overexpression levels achieved in our cell line models, between 2 and 3 FC, were comparable to the upregulation in patients (Figure 26).

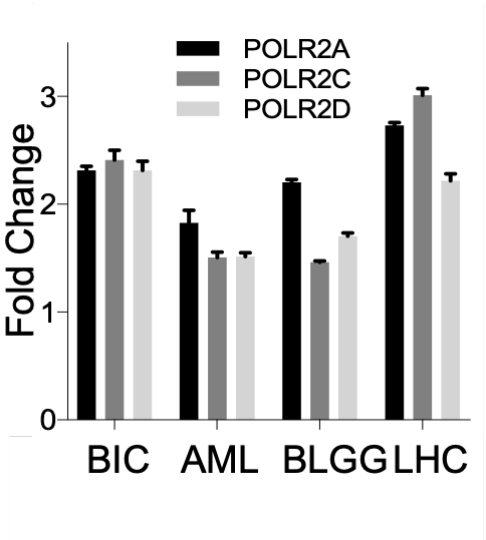
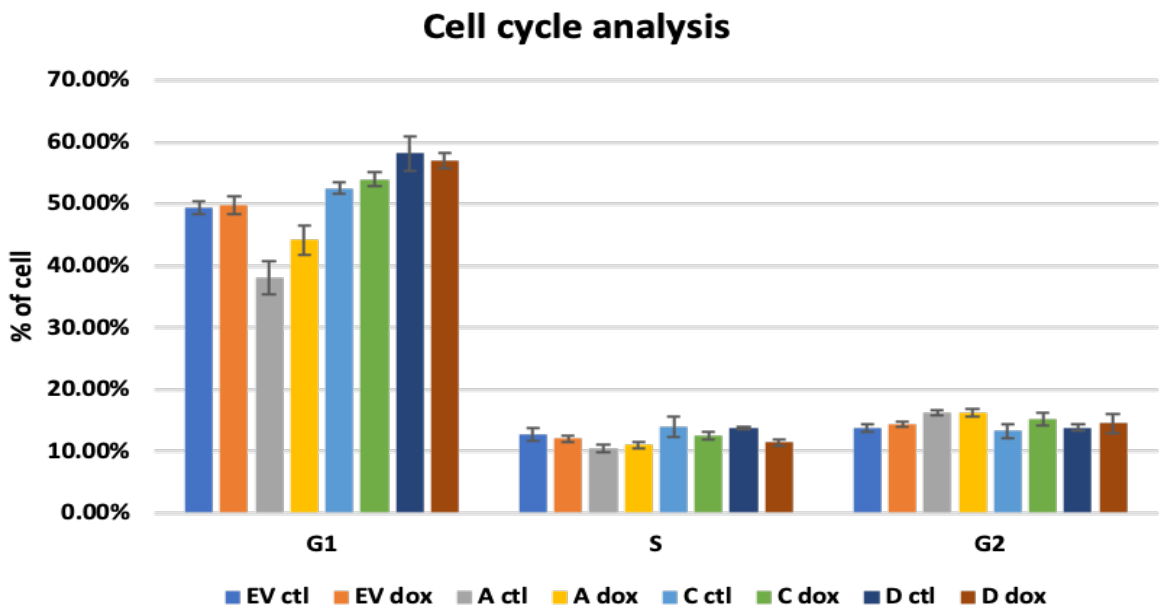


Figure 26 Fold Change of upregulation of POLR2A, POLR2C and POLR2D in BIC, AML, BLGG, LHC between Pol Up patients and patients without upregulation of the subunit. Error bars show standard deviation.

Once we assessed that our cell line models were correctly overexpressing the single subunits, we firstly assessed whether the overexpression of the single subunit impacted on cell cycle progression, or whether this could be affected by the addiction of doxycycline, even at the low concentrations we used for the overexpression. After inducing cells for 48h with doxycycline, we fixed them and prepared the samples for FACS analysis, using DNA violet to stain for the DNA content. We could observe that neither doxycycline on its own, nor the single subunits overexpression impacted on cell cycle progression (Figure 27), given that we found the same % of cells in the three main phases of the cell cycle.



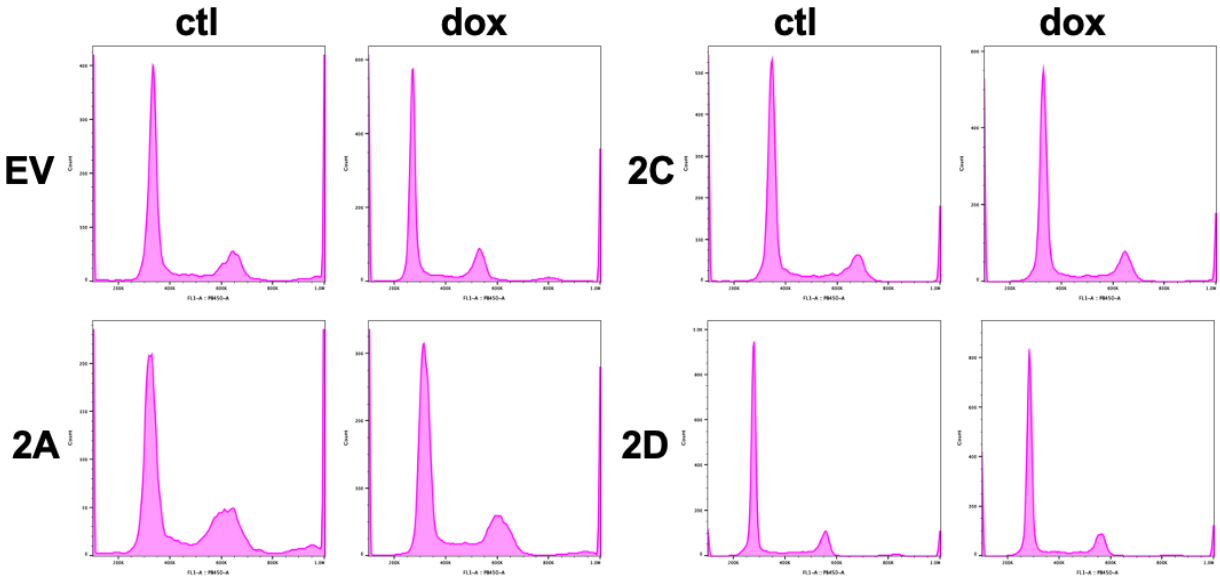


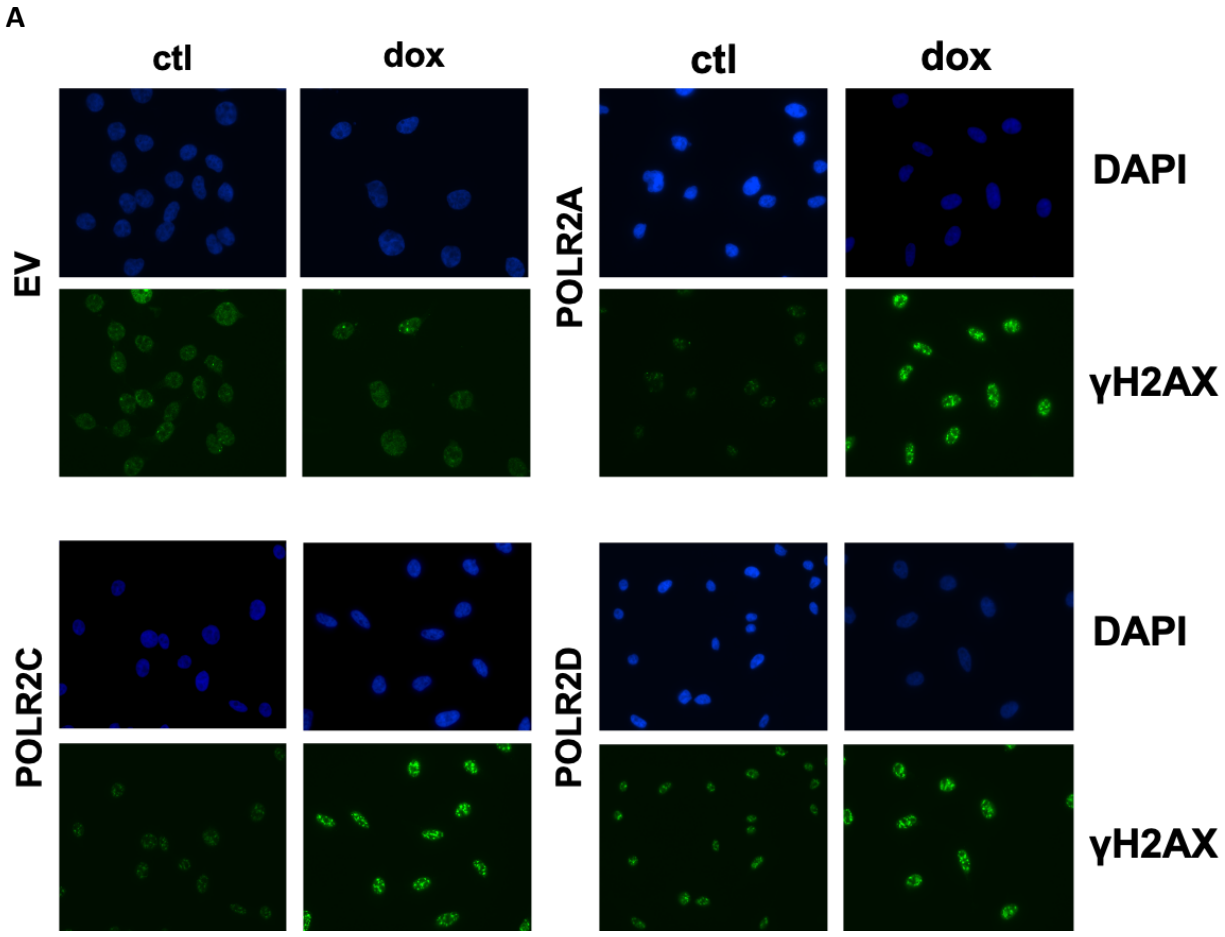
Figure 27 POLR2A-D subunits overexpression does not impact on cell cycle progression Up, distribution of cells (% of cells) in the three phases of cell cycle, G1, S and G2. Bottom, FACS profile of the cell cycle after sorting cells according to DNA content. Samples are empty vector (EV) uninduced (ctl) and induced (dox), POLR2A ctl and dox, POLR2C ctl and dox, POLR2D ctl and dox. Error bars show standard error mean, N=2

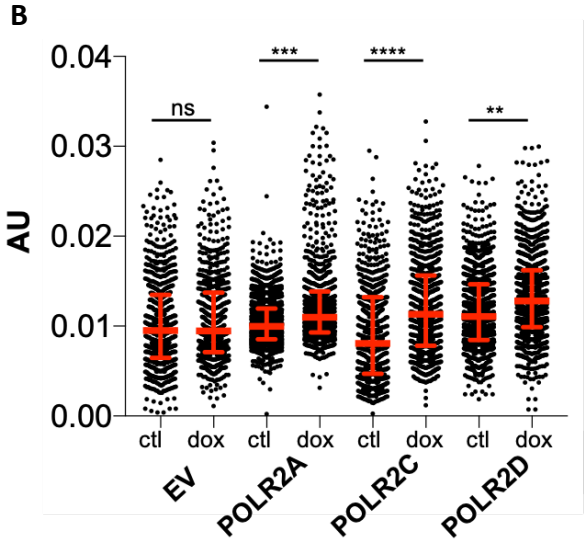
3.2.2 Immunofluorescence of DNA damage markers

3.2.2.1 Analysis of γH2AX foci and intensity

Firstly, we wanted to assess whether the phenotype observed in patients could be observed also in our cell line models and therefore could be considered a direct consequence of the overexpression of the single subunits. Our analysis in cBio Portal showed that in the subset of patients with high expression of POLR2A-D subunits there is also a significant increase in copy number alteration levels, which as previously mentioned is indicative of increased genomic instability. Therefore, we decided to check for the presence of increased DNA damage markers by immunofluorescence. One of the most common DNA damage marker is γH2AX,

phosphorylated histone 2A family member X, whose phosphorylation is carried out in a redundant manner by different DNA damage sensor kinases, such as ATM, ATR and DNA-PK (Blackford and Jackson 2017). As for the conditions established in our models, we treated the cell lines overexpressing POLR2A-D and the EV cell line with 5ng/mL doxycycline for 48h and then we proceeded with the staining. We observed that in all three cell lines when cells where induced with doxycycline, the global nuclear levels of γH2AX increased, indicating an overall increase in DNA damage (Figure 28A). This is easly observed calculating the mean intensity of γH2AX signal in each single nucleus (Figure 28B), and this is overall significantly increased upon subunit overexpresions. Importantly, it is not increased in the empty vector cell line, indicating that the increased DNA damage levels observed are dependent on the single subunits overexpression and not because of the doxycycline treatment. Finally, we counted the foci present in the nuclei, and calculated the fold change (FC) of foci per cell between doxycycline induced and uninduced samples, in order to take into account the possible doxyxycicline effect on samples. Consistently with the γH2AX intensity data, the γH2AX foci number fold change is significantly increased in the subunits overexpressing cell lines, compared to the empty vector (Figure 28C).


A
ctl
dox
ctl
dox
EV
POLR2A
DAPI
γH2AX
POLR2C
POLR2D
DAPI
γH2AX


B
AU
0.04
0.03
0.02
0.01
0.00
ns
***
****
**
ctl
dox
ctl
dox
ctl
dox
ctl
dox
EV
POLR2A
POLR2C
POLR2D

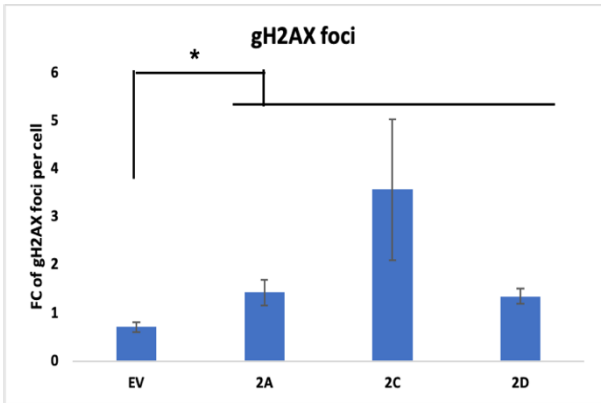
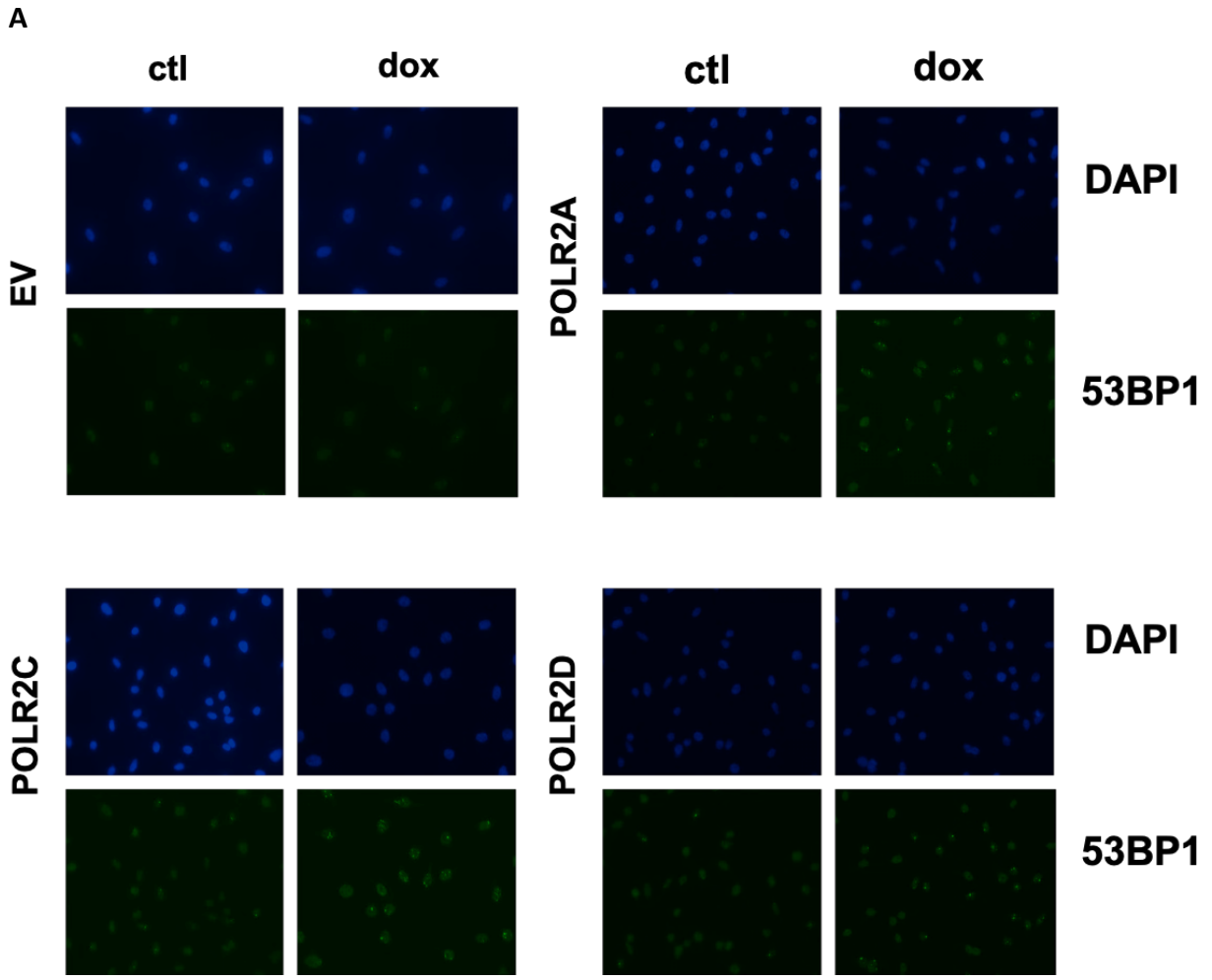

C
gH2AX foci
*
FC of gH2AX foci per cell
6
5
4
3
2
1
0
EV
2A
2C
2D

Figure 28 γH2AX increase upon subunits overexpression A) Representative microscope image of γH2AX intensity in induced (ctl) and uninduced (dox) cells in EV (empty vector), POLR2A, POLR2C and POLR2D cell lines. B) Quantification of γH2AX mean intensity per cell in EV, POLR2A, POLR2C and POLR2D cell lines in ctl and dox samples. C) Fold change (FC) of γH2AX foci number per cell in induced vs uninduced samples in EV, POLR2A, POLR2C and POLR2D cell lines. N=3, ns= not significant, *= p value <0.05, **= p value <0.01, ***= p value <0.001, ****= p value <0.0001.

3.2.2.2 Analysis of 53BP1 foci

53BP1 nuclear accumulations, together with γH2AX, are the most commonly used markers of DNA damage. Specifically, 53BP1 interacts with p53, a key protein in orchestrating DDR, and is usually dispersed in the nuclei of cells, but specifically after DNA damage double strand breaks (DSBs) it localizes in relatively small foci (Rappold et al., 2001; Markova et al., 2015). That is why both 53BP1 and γH2AX are widely recognized as markers of DNA damage and easily visualized by fluorescent microscopy. To confirm and support our γH2AX results, we also performed staining for 53BP1. Similarly with what observed with γH2AX intensity increase, the number of 53BP1 foci per cell also increased upon the overexpression of the single subunits, while it did not change in our control empty vector cell line (Figure 29A). However, differently from γH2AX, the increase in foci was not significant upon the overexpression of the single subunits compared to their uninduced control (Figure 29B). We observed that there was a significant decrease in the number of foci formation upon doxycycline treatment in the empty vector cell line, therefore we decided to take into account the possible doxycycline effect which was not previously observed in the γH2AX intensity measure. Interestingly, when we calculated the fold change number of foci of induced (dox) over uninduced (ctl) samples, we found that the fold change in POLR2A, POLR2C and POLR2D cell lines is significantly higher

compared to the fold change in the EV control (Figure 29C), indicating that taking in account any effect doxycycline could have on cells, there is a significant increase in the average number of 53BP1 foci per cell, upon single subunits overexpression. The slightly different behaviour between the cell lines and the difference in γH2AX levels compared to 53BP1 foci increase, indicates that even though increasing the global levels of DNA damage, each subunit could have a different impact on genome stability, perhaps through different mechanisms inducing damage.



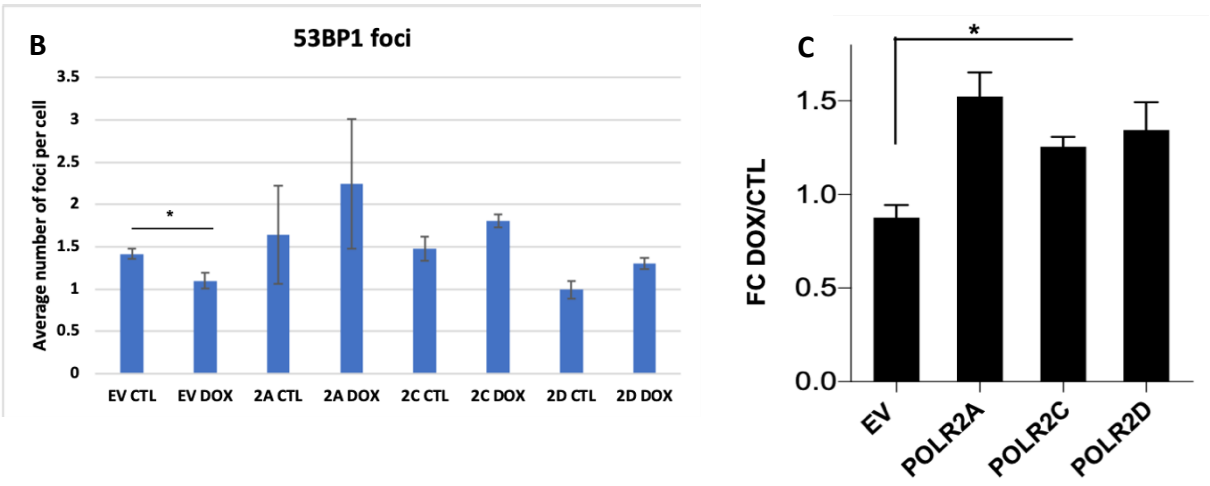
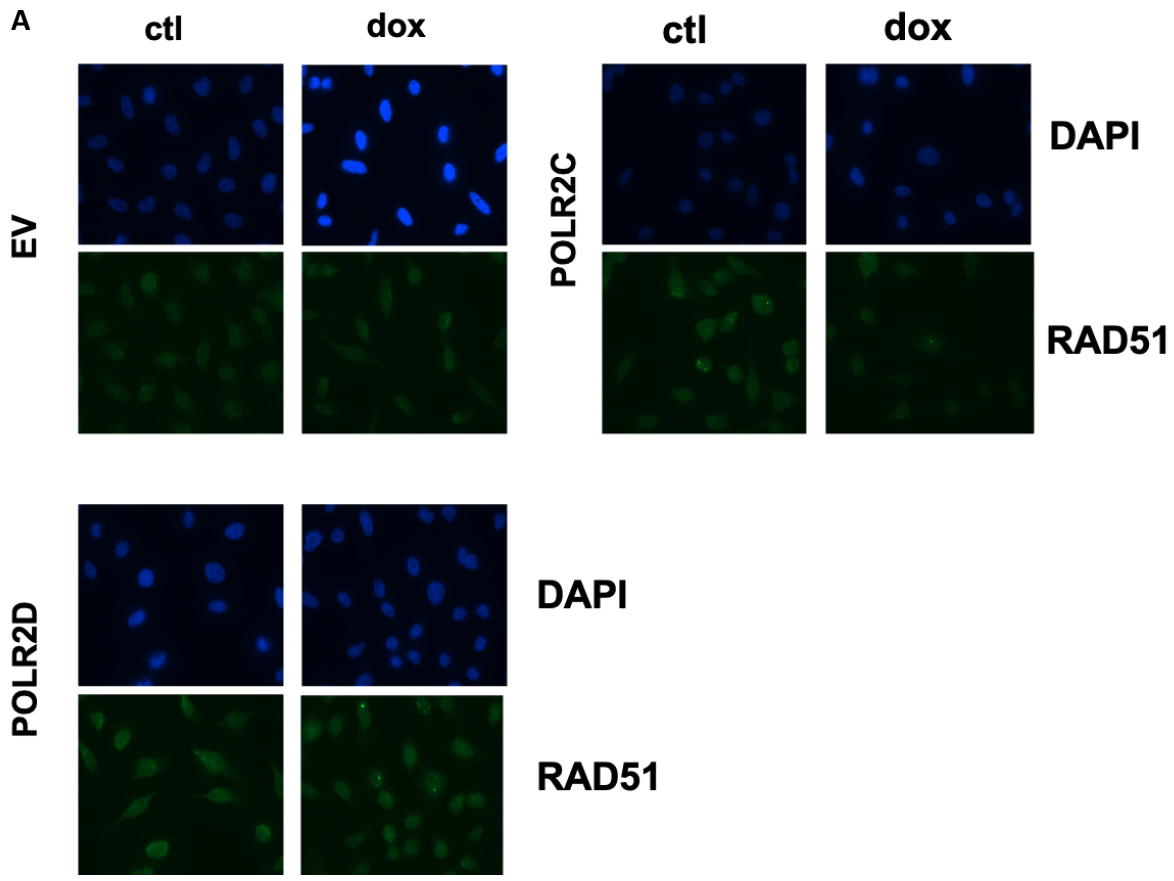


Figure 29 53BP1 foci formation increases upon subunits overexpression A) Representative microscope image of 53BP1 staining in induced (ctl) and uninduced (dox) empty vector (EV), POLR2A, POLR2C and POLR2D cell lines. B) Quantification of 53BP1 average number of foci per cell in EV, POLR2A, POLR2C and POLR2D cell lines in ctl and dox samples. C) Fold Change induced over uninduced (dox/ctl) of 53BP1 average foci per cell in EV, POLR2A, POLR2C and POLR2D cell lines. Error bars show standard error mean, N=3, *= p value <0.05,

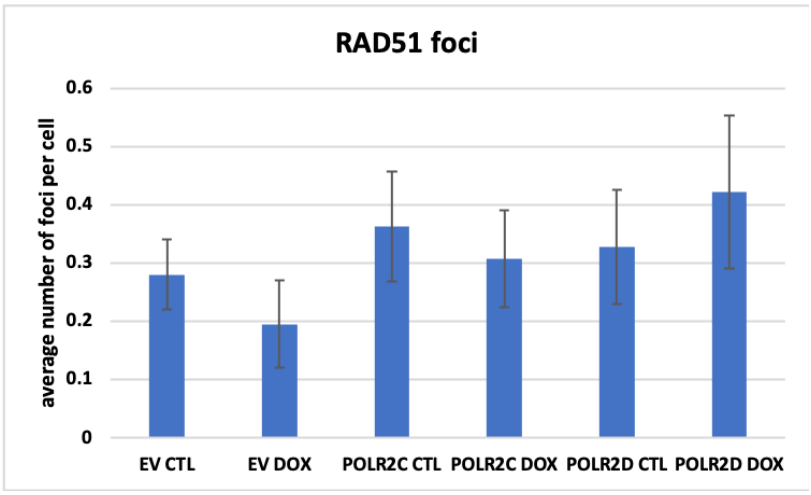
3.2.2.3 Analysis of RAD51 foci

Together with γH2AX and 53BP1, we also decided to look at RAD51 foci formation as another marker of DNA damage. Indeed, RAD51 is one of the key players of homologous recombination (HR) repair and therefore can be used as double strand break marker during S-G2 phases (Baumann and West 1998; Huang and Zhou 2021). Unfortunately, due to time constrains, these experiments were not performed in the POLR2A cell line. Interestingly, even if we were able to observe an increase in both γH2AX intensity and 53BP1 foci formation in both cell lines, when we counted RAD51 foci we found that only upon the overexpression of POLR2D subunit there was an increase in RAD51 foci number. On the contrary, the overexpression of POLR2C subunit did not affect the RAD51 foci formation, similarly to what

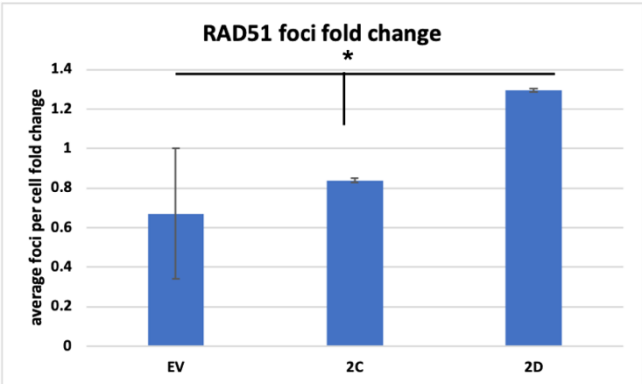
observed in the empty vector (Figure 30A-B). More precisely, both the empty vector and the POLR2C cell line showed a slight decrease in the formation of RAD51 foci. Therefore, as previously done for 53BP1 foci, we calculated the fold change of induced vs uninduced samples to consider the doxycycline effect. By doing so, consistently with the other data, POLR2D overexpression significantly increases the RAD51 foci formation, while POLR2C overexpression does not (Figure 30C). To further elucidate the lack of RAD51 foci formation upon POLR2C overexpression, we decided to quantify the mRNA level of RAD51 by quantitative PCR. By doing so and calculating the fold change compared to the induced samples in the EV and POLR2C cell lines, we found a significant decrease in RAD51 expression (Figure 30D), which could be responsible for the phenotype observed with our staining.



B



C



D

Figure 30 RAD51 foci formation upon subunits overexpression A) Representative microscope image of RAD51 staining in induced (ctl) and uninduced (dox) empty vector (EV), POLR2C and POLR2D cell lines. B) Quantification of RAD51 average number of foci per cell in EV, POLR2C and POLR2D cell lines in ctl and dox samples. C) Fold Change induced over uninduced (dox/ctl) of RAD51 average foci per cell in EV, POLR2C and POLR2D cell lines. D) Fold change of RAD51 mRNA quantification by real time PCR in empty vector dox/ctl and POLR2C dox/ctl. Error bars show standard error mean, N=3, *= p value <0.05, **= p value <0.01.

3.2.3 Analysis of 53BP1 bodies, micronuclei and mitotic bridges

Using the same set of images, we integrated our analysis by quantifying the number of 53BP1 bodies and micronuclei. 53BP1 bodies are larger accumulations of 53BP1 and are usually found in G1 cells, as they are the evidence of unresolved DNA lesions caused by replication stress which are not correctly repaired before cell division (Lukas et al, 2011). Indeed, when the replication fork stalls, the under-replicated DNA can form DNA bridges which give rise to chromosomal rearrangements, chromosomal instability and 53BP1 bodies and micronuclei formation (Saxena and Zou 2022; Krupina et al., 2021). Indeed, 53BP1 bodies and micronuclei are to be considered markers of replication stress, mitotic problems and associated DNA damage.

We observed a significant increase of 53BP1 bodies and micronuclei formation in POLR2C and POLR2D cell lines upon the overxpression of the single subunits, compared to uninduced (ctl) and induced (dox) empty vector. Interestingly, the overexpression of POLR2A did not significantly increased neither the number of bodies nor the number of micronuclei, suggesting that the mechanism by which its overexpression induces increased DNA damage might not be due to increased replication stress (Figure 31A-B).

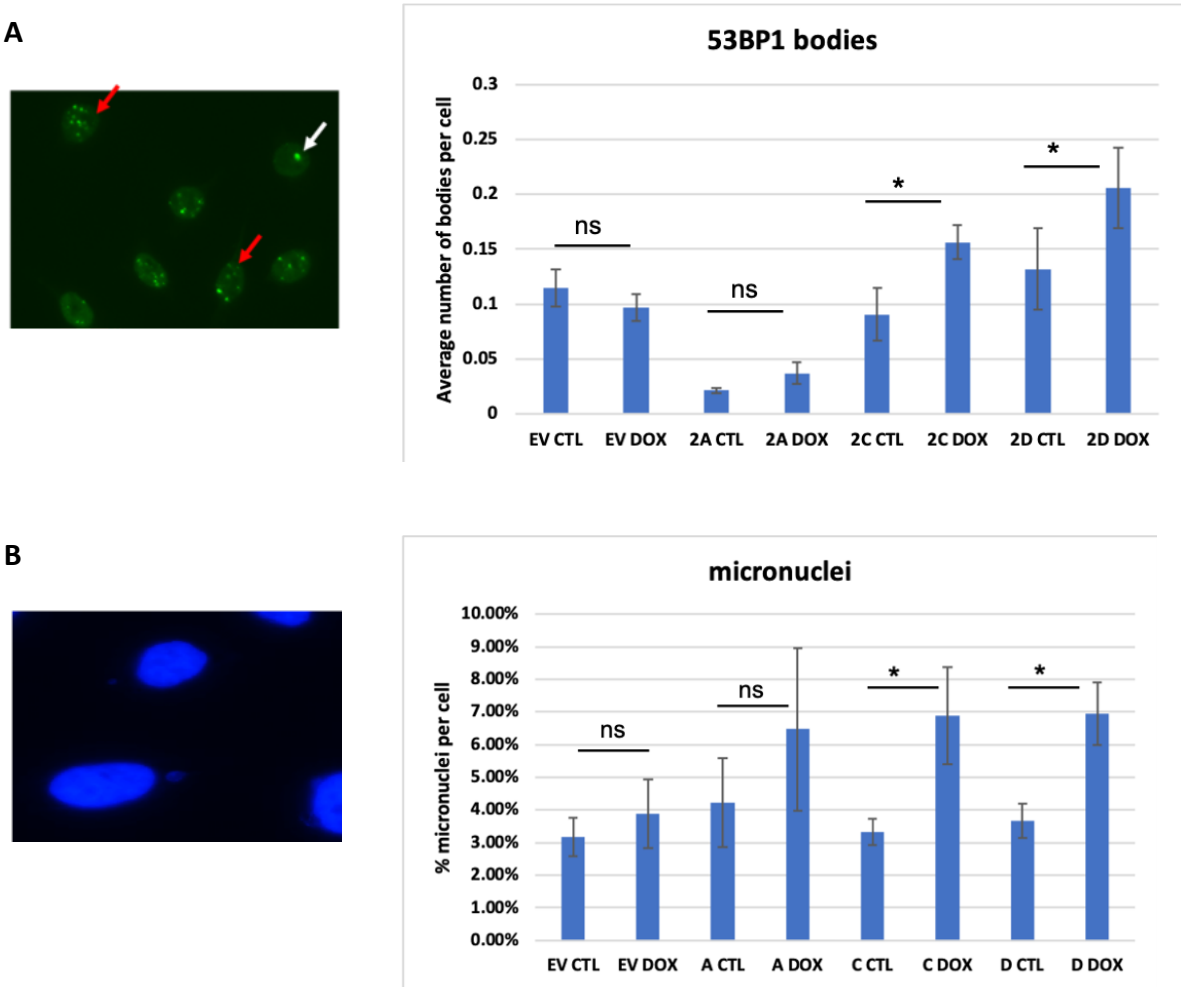


Figure 31 53BP1 bodies and micronuclei formation increase upon subunits overexpression

A) Representative microscope image of 53BP1 bodies (white arrow) compared to foci (red arrow) and quantification of 53BP1 average number of bodies per cell in EV, POLR2A, POLR2C and POLR2D cell lines uninduced (ctl) and induced (dox). B) Representative microscope image of micronuclei and quantification of micronuclei % average number per cell in EV, POLR2A, POLR2C and POLR2D cell lines uninduced (ctl) and induced (dox). Error bars show standard error mean, N=3, ns = not significant, *= p value <0.05,

As previously mentioned, replication intermediates or DNA lesions that are not repaired or resolved during S phase can give rise to mitotic errors, which can be visualised by the formation of chromatin bridges (anaphase bridges), which are dsDNA and therefore DAPI

positive, lagging chromosomes or ultrafine bridges (UFBs, ssDNA) (Chen et al., 2018). We observed the formation of anaphase DAPI positive bridges in mitotic cells, increased upon POLR2C and POLR2D subunits overexpression (Figure 32).

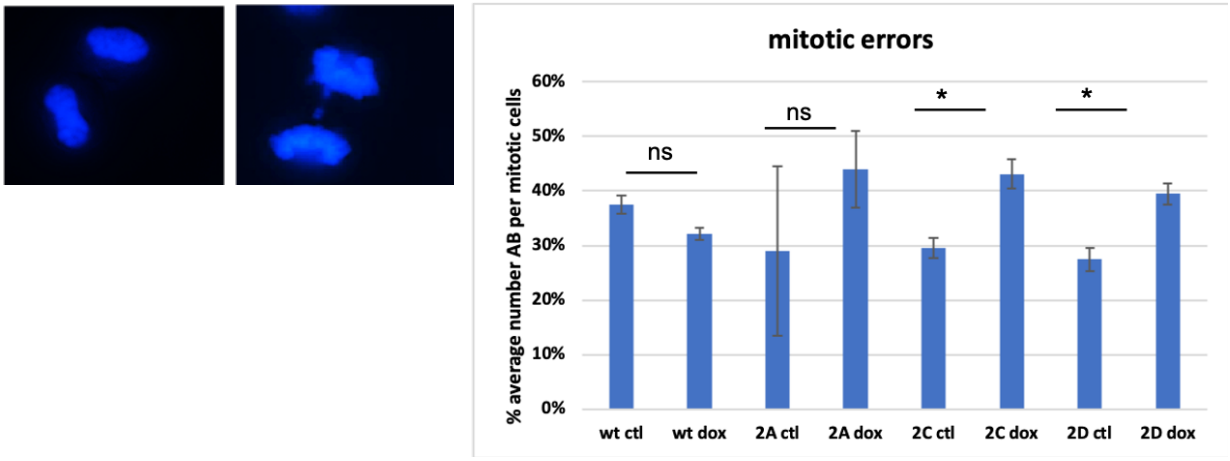
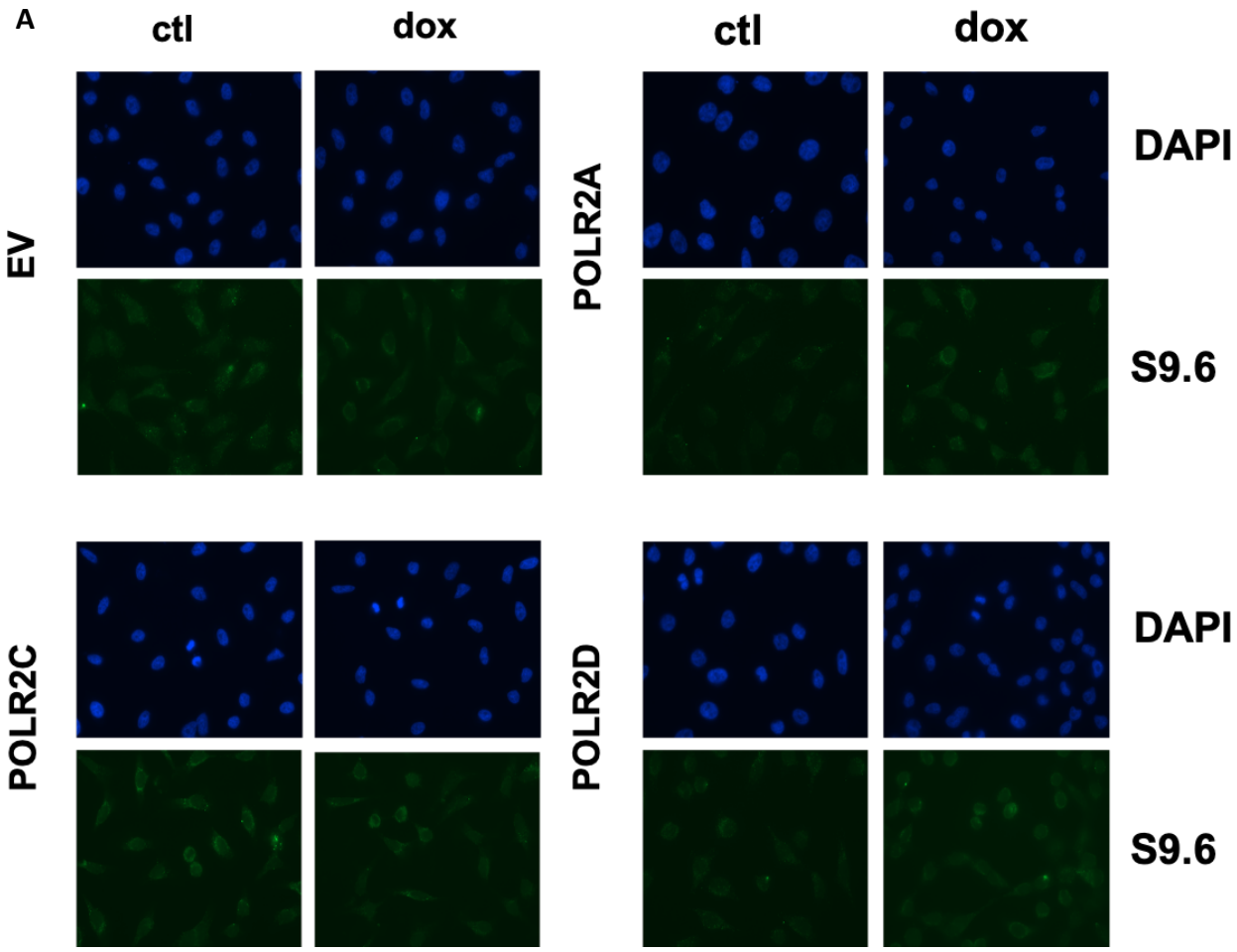


Figure 32 Representative figure of a normal anaphase cell on the left and an anaphase cell with an anaphase bridge on the right, and % average number of anaphase bridges (AB) per mitotic cells in EV, POLR2A, POLR2C and POLR2D uninduced (ctl) and induced (dox). Error bars show standard error mean, N=3, ns = not significant, *= p value <0.05

3.2.4 Analysis of R-loops

Another source of transcription associated genome instability is constituted by the accumulation or persistence of R-loops, a three-stranded secondary structures formed by the nascent RNA hybridized with the DNA template strand (Hamperl and Cimprich 2016), which can lead to transcription associated mutagenesis and recombination events (Skourti-Stathaki and Proudfoot 2014). Indeed, when not resolved, the formation of this structures can also evolve in DSBs formation (Sollier et al., 2014; Hamperl & Cimprich, 2016; Tubbs & Nussenzweig, 2017). Given the general increased DNA damage levels observed, we checked whether subunits upregulation could lead to an accumulation in the formation of R-loops.

After inducing the cells for 48h with doxycycline, we performed an immunofluorescence staining using the S9.6 antibody, which recognises R-loop structures as it binds DNA:RNA hybrids (Figure 33A). Surprisingly, we did not observe increases in R-loop intensity upon doxycycline induction, independently of the subunit overexpressed (Figure 33B).



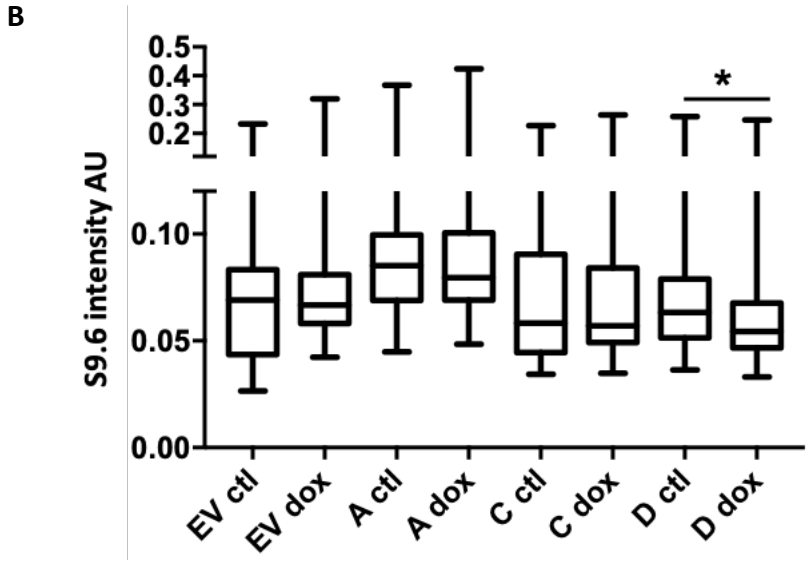


Figure 33 A) Representative image of S9.6 staining (R-loops) in EV, POLR2A, POLR2C and POLR2D uninduced (ctl) and induced (dox). B) Quantification of the S.96 intensity in EV, POLR2A, POLR2C and POLR2D uninduced (ctl) and induced (dox). N=3, *= p value <0.05, non-significant p values not included in the figure.

3.2.5 γH2AX ChIP-Seq

To map the regions where DNA damage observed by IF is occurring, we decided to perform a chromatin immuno-precipitation (ChIP) followed by sequencing (ChIP-Seq) for γH2AX, upon subunits overexpression and in the empty vector cell line, with doxycycline induction for 48h. In parallel, to normalize for γH2AX ChIP levels, we also performed a ChIP-Seq of the unmodified H2AX in the same samples. Firstly we performed the analysis on all the genes extracted from the GRCh38 (human reference genome) and the average metagene profiles showed almost no difference between induced and uninduced samples (data not shown). Therefore we performed our analysis specifically across transcribed regions, on the list of genes transcribed in HeLa cells. To determine these, we identified these genes calculating the

reads per kilobase per million (RPKM) on an RNA-Seq dataset of HeLa cells from Encode, and selecting genes with RPKM value greater than 1.

First, we looked at the average metagene profile over all transcribed genes from the transcription start site (TSS) to the transcription termination site (TTS) in the empty vector cell line and in POLR2A, POLR2C and POLR2D cell lines upon the overexpression of the corresponding subunits (Figure 34). Interestingly, we noticed that the three cell lines behaved differently, with each of the overexpression inducing an increase in γH2AX levels across genes in specific positions. For example, the overexpression of the POLR2A subunit appears to increase the accumulation of damage near the transcription start site and the beginning of the gene, as shown by the POLR2A dox profile in black, compared to its uninduced control in orange (Figure 34B). However, the more we move into the gene body and towards the transcription termination site, the less pronounced the difference between the two profiles becomes until they overlap, indicating that DNA damage is accumulating at the beginning and in the first half of the gene. Differently, when we observe the average profile over transcribed genes in POLR2C cell line we can appreciate how in the POLR2C dox (in black) the γH2AX levels are always higher than its uninduced control (in orange) (Figure 34C). Finally, the average γH2AX profile upon POLR2D overexpression (in black) shows how the accumulation of γH2AX starts more into the gene body and increases towards the TTS (Figure 34D). These data appear indicating how the overexpression of each subunit affects genome instability levels in a specific way, in agreement and explaining some of the differences previously identified when we analysed DNA damage levels in cells by immunofluorescence. Importantly, similarly to what observed by immunofluorescence, we noticed that the doxycycline treatment was impacting on the γH2AX levels in the empty vector cell line. Indeed, the average profile shows

a slight decrease in γH2AX levels upon doxycycline induction (in black) compared to the uninduced sample (in orange) (Figure 34A).

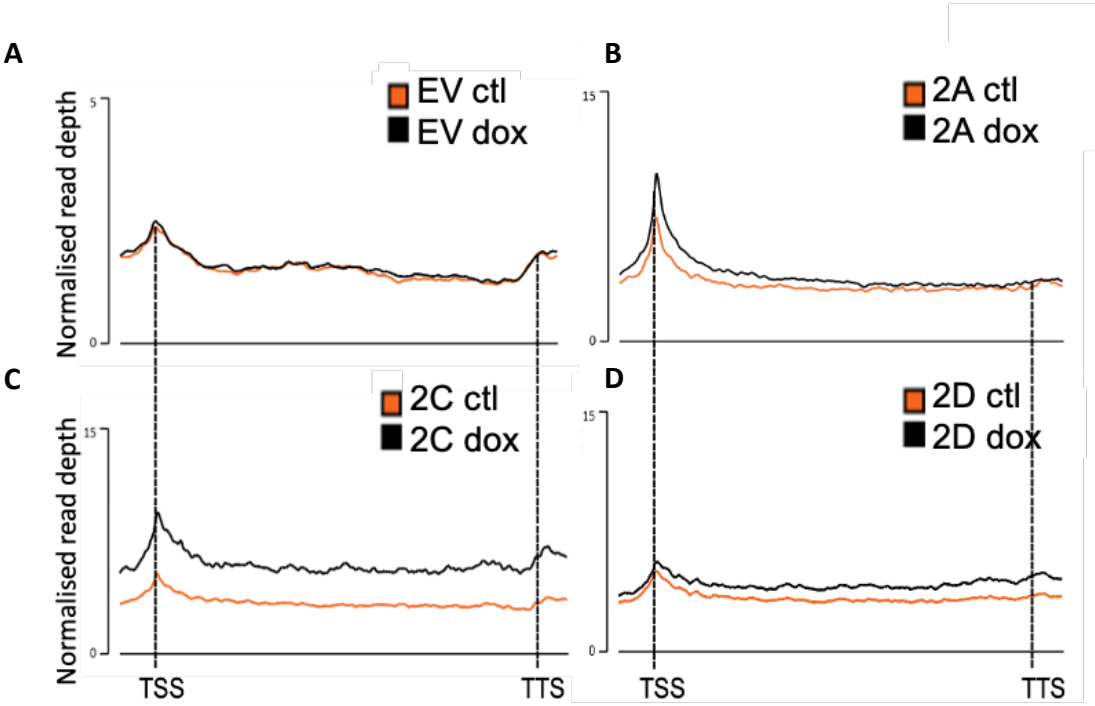


Figure 34 Metagene profile of γH2AX ChIP Seq from the transcription start site (TTS) to the transcription termination site (TTS) on transcribed genes in HeLas in: A) Empty vector (EV) uninduced (ctl) in orange and induced (dox) in black, B) POLR2A ctl and dox, C) POLR2C ctl and dox and D) POLR2D ctl and dox.

Following the average metagene analysis, we decided to perform a gene-to-gene analysis. As the average metagene assesses only the average γH2AX profile across all transcribed genes, it does not allow for example determining whether there are only a few genes that are highly affected by DNA damage, or many genes on which there is a slight increase in DNA damage. Given the metagene profile observed in the Empty vector (Figure 34A), we decided to take into account the doxycycline effect when performing our gene-to-gene analysis. To do so, the gene-to-gene fold change (FC) that we measured upon each subunit overexpression was

normalized to the gene-to-gene FC in the Empty vector cell line. Therefore, following normalization to the EV, we observed that over all transcribed genes the differences in the increase in γH2AX from the TSS to the TTS depended on which subunit was overexpressed (Figure 35A). For example, a high level of increase in γH2AX can be appreciated mainly in the POLR2C overexpressing cell line upon the specific subunit overexpression, with the majority of genes having a FC above 1, while POLR2D overexpression shows a slight increase in γH2AX, POLR2A overexpression seems to have no effect, with most of the genes with a FC below 1, as the metagene analysis was indicating an increase in γH2AX exclusively around the TSS region. Then, we selected the genes with a fold change increase in γH2AX levels >1.5 in all three cell lines, and we ranked the affected genes by their gene length, observing that genes with higher γH2AX FC tend to be shorter genes, especially in the POLR2D cell line (Figure 35B).

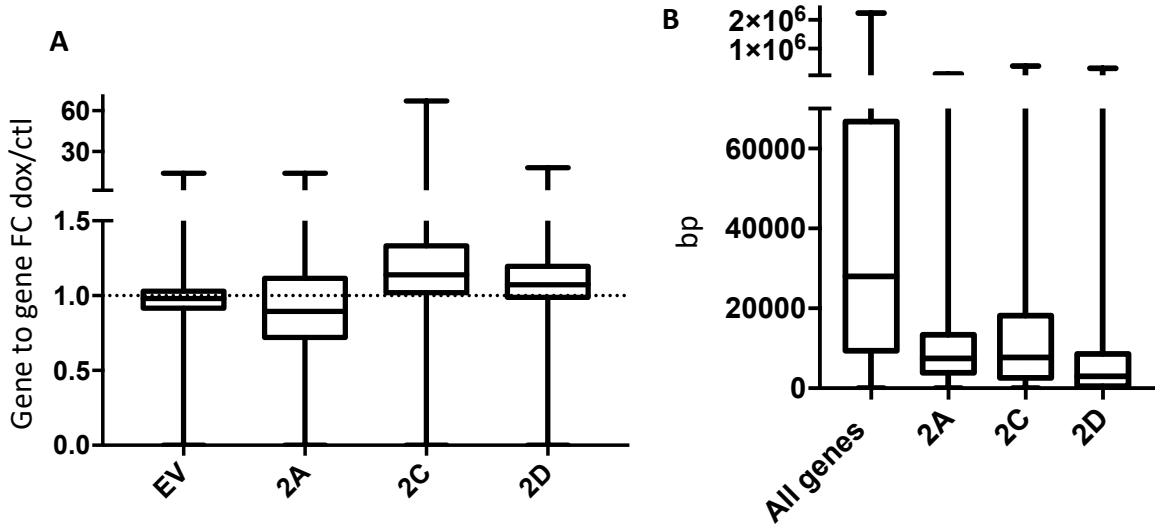


Figure 35 γH2AX gene to gene quantification and gene length distribution A) gene-to-gene fold change FC induced/uninduced in empty vector (EV), POLR2A (2A), POLR2C (2C) POLR2D (2D) cell lines of γH2AX level quantified in HeLa transcribed genes. B) Gene length distribution of all genes transcribed in HeLa and of genes with a γH2AX induced/uninduced gene-to-gene FC above 1.5 in POLR2A, POLR2C and POLR2D cell lines, p value not included in the figure <0.001

Next, we quantified the γH2AX/H2AX at specific genomic regions only, in particular at the TSS (TSS->TSS+1kb), gene body (GB) (from TSS+1kb to TTS-1kb) and TTS (TTS+1kb->TTS) of transcribed genes, to better characterize the DNA damage phenotype (Figure 36 left). Consistently with what previously described, we observe that the higher fold change in POLR2A induced vs uninduced is found only at the TSS of genes and lost when proceeding into the GB. On the opposite, when looking at the effect of POLR2D overexpression we observe an increase in the gene-to-gene fold change towards the TTS. POLR2C overexpression, as before, seems to induce more γH2AX independently of the position in the gene, with slightly more increase at the TSS. We hypothesized that this could be related to the different impact of the overexpressed subunits on transcription regulation. However, as we observed that short genes where the most affected by the γH2AX increase we wanted to see whether the same trend was conserved even on longer genes. Therefore, we quantified γH2AX levels at the TSS, GB and TTS of genes longer than 30kb and interestingly we observed the same trend as for all genes (Figure 36 right).

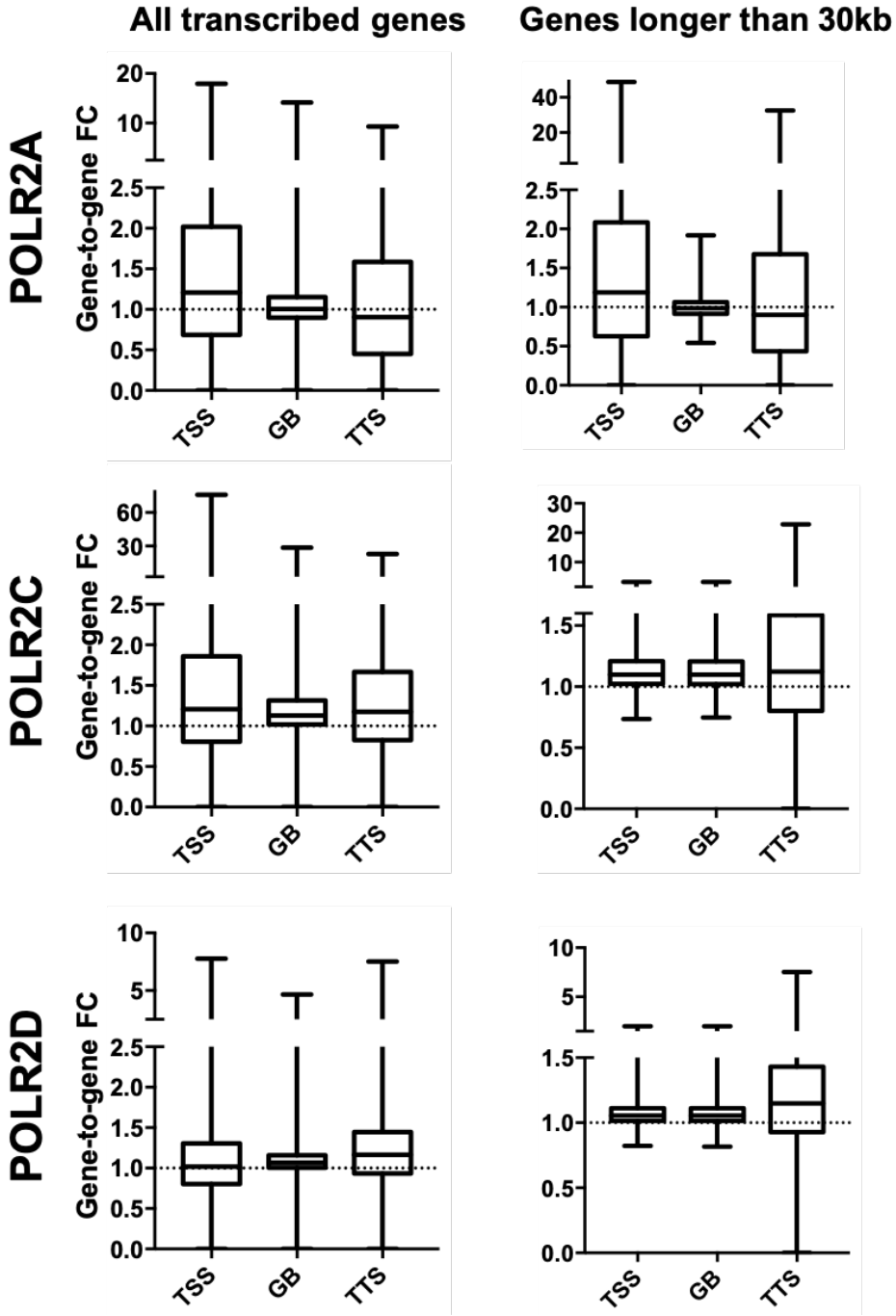


Figure 36 γH2AX quantifications at the TSS, GB, TTS of all and long genes gene-to-gene fold change of γH2AX in POLR2A, POLR2C and POLR2D cell lines induced vs uninduced on all transcribed genes in HeLa (left) and on gene longer than 30Kb (right) measured at the TSS->TSS+1kb, the GB (TSS+1Kb to TTS-1kb) and at the TTS+1kb->TTS.

3.2.7 Discussion

Our findings on cBioPortal show that Pol II major subunits overexpression is associated with patients poor survival. In order to understand what phenotypes are associated with the overexpression of the single subunits in cells, we established a series of cell line models to characterize the overexpression of the single subunits. We generated four cell lines using HeLa T-REx inducible system, one empty vector (EV) and one each for POLR2A, -2C and -2D overexpressing cell lines. We decided to use this model instead of transient transfection since it allows for fast and consistent overexpression, without relying on the transfection efficiency. Each cell line has its own internal control, the uninduced sample, plus we used the EV cell line as a control for any undesired secondary effect of doxycycline treatment alone. The system we established for overexpressing the single subunits shows fold changes (FC) of expression in induced (dox) over uninduced (ctl) samples to be between 2 and 3. Although these are not very high FC, we observed that the overexpression in our system well recapitulates the upregulation levels in the patients (for BIC, AML, BLGG and LHC)(Figure 24; Figure 25; Figure 26) . Additionally, the phenotypes observed in patients were a result of the upregulation of the four subunits (POLR2A-D) as an ensemble, since the low number of patients affected by a single subunit overexpression would have compromised the robustness of our analysis. Therefore, the generation of cell lines overexpressing one subunit at time allowed us to characterize the impact of each single subunit deregulation, highlighting differences and similarities arising upon each subunit overexpression.

Firstly, we have shown how the overexpression of the single subunits POLR2A, POLR2C and POLR2D alone can cause an increase in DNA damage, resulting in the accumulation of common DNA damage markers (Figure 28; Figure 29; Figure 30). However, even if the general levels of

damage increase, the mechanism responsible for its increase may differ according to which subunit is overexpressed. γH2AX foci formation is indicative of a general increase in DNA damage, while 53BP1 foci form at DNA DSBs (Williamson et al. 2020). We observed a pan-nuclear staining of γH2AX, which indicates a general increase in DNA damage but also could indicate the presence of increased replication stress (Moeglin et al. 2019; Parsels et al. 2018). The increased formation of 53BP1 foci upon each single subunit overexpression (Figure 29), however, indicates that there is also increased formation of DNA DSBs. These results indicate that the overexpression of the single subunits is sufficient to increase DNA damage formation in cells, without requiring additional treatment with common exogenous DNA damaging agents. The increased presence of DNA damage markers could be indicative of either an increased challenge to DNA stability by subunit overexpression or a defective/ slower DNA damage repair. However, the RAD51 staining results in particular upon POLR2C overexpression, may suggest at least in part a defective repair capability, which would result in an accumulation of endogenous DNA damage. Further experiments would be needed to confirm this. For example, we could perform an HR reporter assay to better understand whether the lack of RAD51 foci formation upon POLR2C overexpression is a phenotype linked to defective HR, or we could do a kinetics of RAD51 foci formation after inducing DNA damage in the presence or absence of subunits overexpression. Moreover, the qPCR analysis of RAD51 mRNA levels indicate that the gene is significantly less expressed (Figure 30D). This could be a direct consequence of a transcription defect arising due to the subunit upregulation, and that's reason why we observe the foci formation defect. Notably, the RAD51 foci formation upon POLR2D overexpression was significant increased compared to the EV, in line with our other DNA damage markers results (Figure 28, Figure 29). Given that RAD51 is involved in HR

(Baumann and West 1998; Mladenov et al. 2016; Wassing and Esashi 2021), is important to mention that the RAD51 formation defects observed in POLR2C, if compared to the POLR2D cell line, is not dependent on changes in the cell cycle progression between the two lines, as our FACS data shows that all four cell lines progress through the cell cycle similarly (Figure 27). Finally, to better understand whether the overexpression of the single subunits is associate to a damage repair defect, we could treat cells with exogenous sources of DNA damage following doxycycline induction. We could then perform sequential staining with DNA damage markers at different time points, to assess if the overexpression of the subunits impacts on how fast cell recover following DNA damage induction. As previously mentioned, some of the phenotypes observed upon the overexpression of POLR2C and POLR2D, more specifically the observation of increased 53BP1 bodies, micronuclei and mitotic errors formation (Figure 31; Figure 32), suggest that the induced DNA damage may be consequence of an increase in replication stress (Wilhelm et al. 2020). Interestingly, the overexpression of POLR2A seems to have no significant effect on these genomic instability markers, further highlighting how the single subunits upregulations drive DNA damage formation by different mechanisms. These results suggest that even if it can cause DNA damage, the overexpression of the largest Pol II subunits could do this without specifically inducing replication stress. To validate this point, it would be useful to perform a characterization of DNA replication following subunits overexpression, for example performing fiber assays in order to assess replication rates (Quinet, Carvajal-Maldonado, et al. 2017).

Also the DNA:RNA hybrids R-loops are implicated in the induction of replication stress, since their formation and persistence can increase, among others, the chances of transcription-replication collisions, and consequently DNA damage (Allison and Wang 2019; Hamperl and

Cimprich 2016; Skourti-Stathaki and Proudfoot 2014). However, there is conversive evidence of R-loops also being implicated in the correct recruitment of DNA repair proteins, and therefore R-loops playing a protective role in genome stability maintenance (Crossley et al. 2019). Given this dual role of R-loops, we could not speculate on the impact of the overexpression of the subunits on their formation and accumulation. According to our immunostaining, the overexpression of POLR2A and POLR2C subunits did not impact on the R-loops formation, while POLR2D overexpression if anything significantly decreased their average nuclear intensity (Figure 33). However, even though the S9.6 antibody is widely used to analyse R-loops levels, it is not absolutely specific for R-loops, as it recognises DNA:RNA hybrids in general but also RNA:RNA hybrids (Smolka et al. 2021). Indeed, we acknowledge that our experiments lack of a positive control, which would have clarified if our results were biased by the quality of the staining/ use of antibody. R-loops are specifically resolved by RNAseH (Amon and Koshland 2016; Lockhart et al. 2019), therefore further experiments using the catalytic dead form of RNAseH, which binds these specific structures, could better clarify the role of the subunits overexpression in affecting R-loops levels. For instance, a ChIP-Seq using the catalytic dead version of RNAseH, which binds but does not process R-loops, to immune precipitate R-loop, could identify the regions more affected by their formation when subunits are overexpressed (Chen et al. 2017). In addition, it could be studied whether the genomic regions more affected by potential R-loops accumulation upon subunits overexpression do correlate with regions of increased γH2AX ChIP-Seq levels.

The differences in the three subunits overexpression on DNA damage formation become more evident when analysing the γH2AX ChIP-Seq data. Indeed, even if there is a general increase in DNA damage upon each subunit overexpression, the average metagene profiles of γH2AX

on transcribed genes are specific for each subunit (Figure 34). What observed in the metagene profiles is also confirmed by the gene-to-gene fold change analysis especially when we quantified γH2AX in the different areas of the gene, TSS, GB and TTS (Figure 36). The fact that γH2AX accumulates at the TSS (in POLR2A and POLR2C), or towards the TTS (in POLR2D), and in the GB of POLR2C as well, suggests that the mechanisms leading to increased DNA damage depend on a specific transcription defect, which differs according to which subunit is overexpressed. When we calculated the gene-to-gene fold change from the TSS to the TTS we observed that the genes with higher γH2AX (FC>1.5) were relatively short genes in all three cell lines (Figure 35), particularly in POLR2D cell line. However, since different regions of the gene (TSS, GB, TTS) were affected according to which subunit was overexpressed, we showed that even longer genes (>30Kb) were affected following the same trend observed on all genes at TSS, GB and TTS, although to a lower extent compared to shorter genes (Figure 36 right). Pol II transcription levels and regulation can vary accordingly to gene length and intron and exons composition and splicing regulation (Chiaromonte et al. 2003; Grishkevich and Yanai 2014). Short genes are usually more expressed than longer genes, in fact many house-keepijg genes are usually short, with both short introns and exons (Brown 2021; Eisenberg et al. 2003; Lopes et al. 2021). Additionally, transcription is a tightly regulated processes and long genes present more sites of Pol II pausing, increased splicing regulation and increased termination regulation (Chathoth et al. 2014; Gilchrist et al. 2010). Finally, transcription levels can also depend on cell cycle stage, which impacts on gene expression and regulation (Bertoli et al. 2013; Bregman et al. 2000; Cho et al. 2001; Cosma 2002; Yonaha et al. 1995). In this light, our results point towards the hypothesis of transcription defects arising with Pol II subunits overexpression that may result in DNA damage accumulation. In order to confirm our

hypothesis, assays aimed to transcription characterisation in the presence of subunits overexpression are needed.

3.3 Single subunit overexpression deregulates transcription

3.3.1 Transient transcriptome sequencing (TT-Seq) shows that subunit overexpression increases transcription activity

Following the characterization of the DNA damage phenotype and identifying genomics regions more affected by DNA damage (by γH2AX ChIP-Seq, 3.2.), we hypothesized that the mechanisms leading to DNA damage accumulation differ from each subunit and could be linked to a possible transcription defect arising upon subunit overexpression. To characterize transcription and identify any possible deregulations, we decided to perform transient transcription sequencing (TT-Seq) upon each subunit induction for 48h with doxycycline. TT-Seq would be able to identify any differences between our samples on actively transcribing RNA Pol II. Following the TT-Seq protocol as described in Gregersen et al., induced and uninduced cells were pulsed with 4-thiouridine (4SU) for 15 minutes, to label nascent RNA with the nucleotide analogue. The incorporation of the 4SU was assessed by dot blot (Figure 37). Once nascent RNA was labelled, total RNA was extracted and fragmented chemically (treating the samples with NaOH). Since we used chemical fragmentation, we refer to this technique as TT-Chem-Seq. The nascent labelled RNA was then biotinylated and pulled down using streptavidin conjugated beads. Once 4SU-labelled RNA was isolated, we proceeded to library preparation and next generation sequencing (NGS). For a relative quantification of nascent RNA abundance we included also a spike-in of *S. cerevisiae* RNA labeled with 4TU (4-thiouracil), which can be imported in yeast and then converted into 4SU, before RNA fragmentation in each sample. The number of mapped reads of the yeast spike-in in each sample was indeed used to calculate a scaling factor according to which each sample is normalized to after sequencing. The scaling factors calculated were then used to produce an

accurate bedgraph file for each sample. Using the same scaling factor, we generated two strand specific bedgraph files (forward and reverse), per sample. The procedure is summarized in (Figure 38).

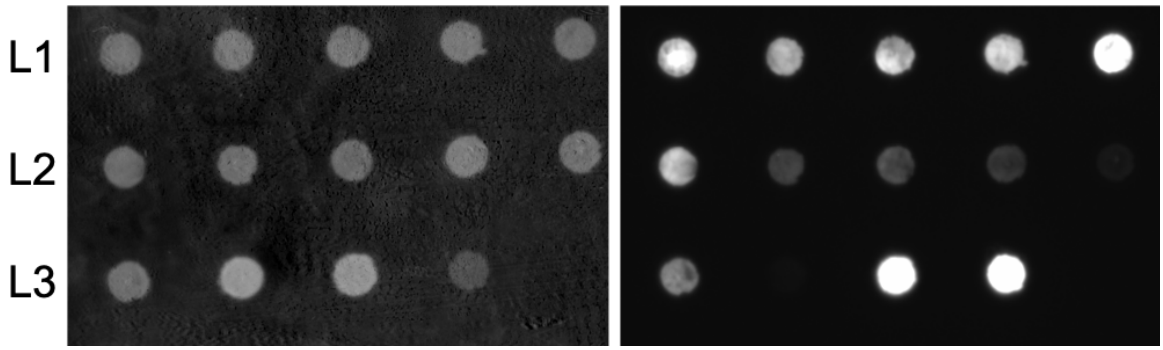


Figure 37 4SU incorporation assessment. Dot blot to assess the incorporation of 4SU. Loading control (left) and dot blot (right) of one representative replicate. Loading from left to right L1: WT ctl, WT dox, 2C ctl, 2C dox, 2D ctl; L2: 2D dox, WT ctl, WT dox, 2A ctl, 2A dox; L3: positive control, negative control, yeast spike in, yeast positive control

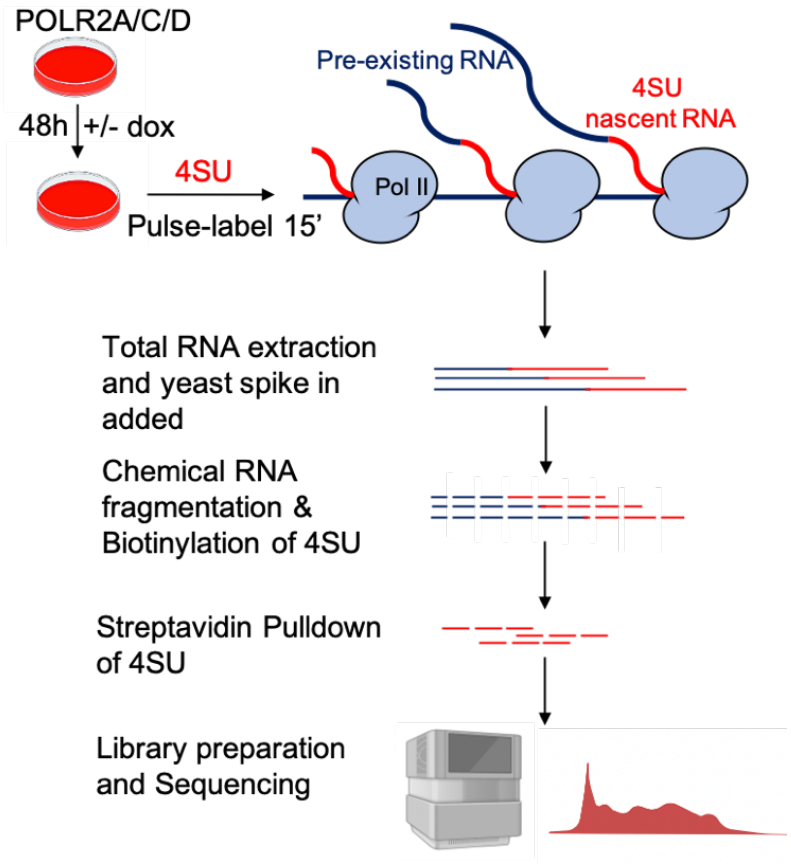


Figure 38 TT-Chem-Seq procedure Representative schematic figure of transient transcription sequencing with chemical fragmentation of RNA

Since our sequencing was strand specific we obtained strand specific results, with reads mapping either on the forward or the reverse strand according to the gene direction. In order to generate accurate average metagene profiles we generated two gene lists, one including the genes on the forward strand and one for the genes on the reverse. First of all, we observed that doxycycline had little effect on its own, as it can be observed from the average metagene profile from the TSS to TTS all over transcribed genes, and if anything decreased RNA Pol II transcription activity on genes on the reverse strand (Figure 39). Differently from what

observed in the empty vector cell line, the overexpression of the single subunits had however a clear impact on transcribing Pol II (Figure 39).

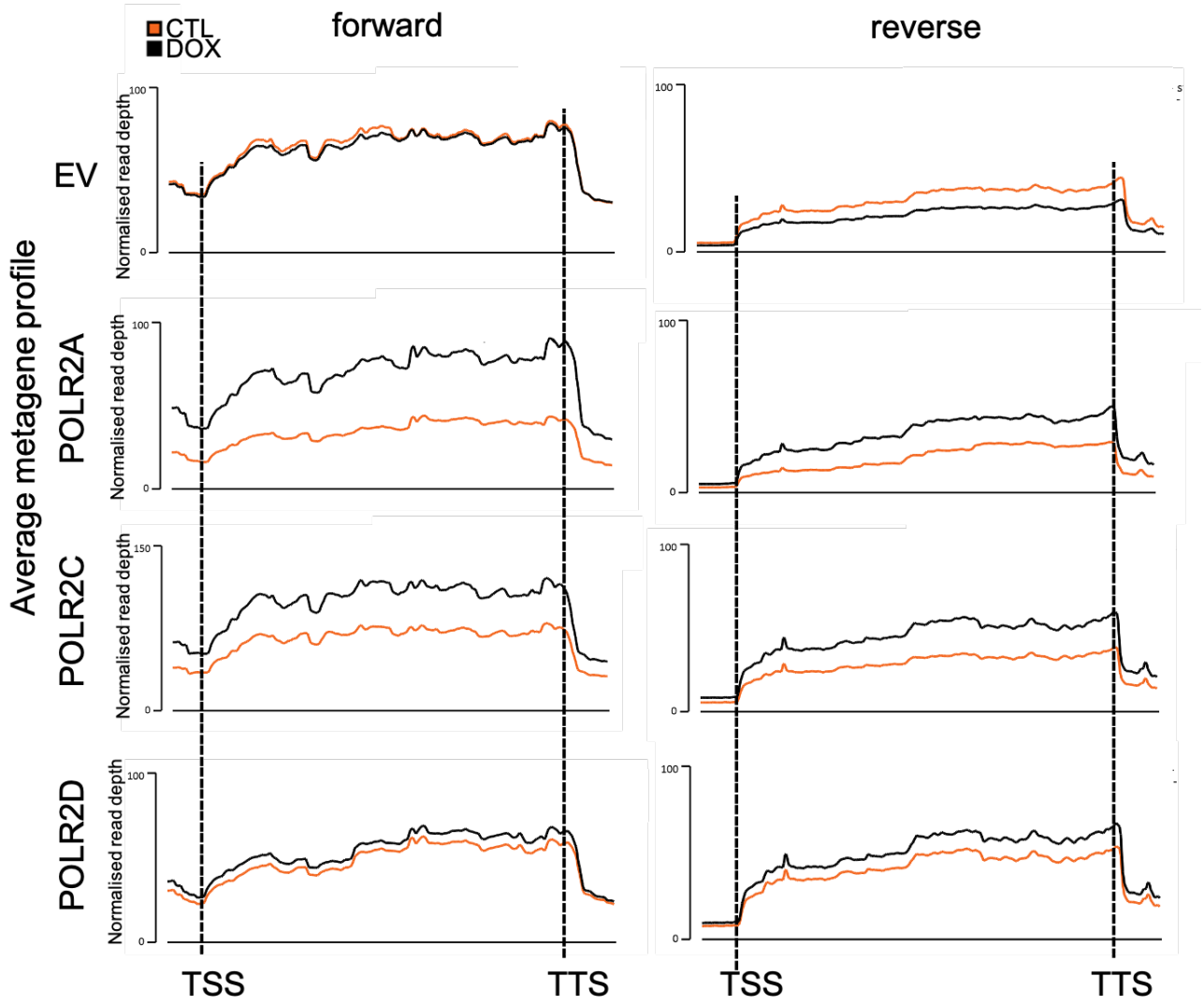


Figure 39 Representative average metagene profiles from transcription start site (TTS) to transcription termination site (TTS) on transcribed genes in HeLa in: Empty vector (EV) uninduced (ctl) in orange and induced (dox) in black, POLR2A ctl and dox, POLR2C ctl and dox and POLR2D ctl and dox, according to forward (right) and reverse (left) strands.

The metagene profiles from the TSS to the TTS of transcribed genes, showed that the overexpression of all three subunits, to a different extent, lead to higher transcription levels (Figure 39). The doxycycline treated samples (dox, in black) show increased transcription levels compared to the corresponding uninduced control samples (in orange), independently of the gene region (TSS, GB or TTS) (Figure 39).

As metagene profiles present only an average picture of the impact of the overexpressions of the subunits on transcription activity, to determine how many genes were specifically affected we calculated the gene-to-gene fold change of induced (dox) vs uninduced (ctl) samples on transcribed genes. For this, we quantified TT-Chem-Seq transcription levels from TSS to the TTS of all genes, in a strand-specific manner specifically for the sense transcription. We also removed all genes that had no TT-Seq quantification (value = 0), in order to facilitate our analysis. We observed that the effect of doxycycline was slightly decreasing the FCs in the EV cell line, and this was observed in both repeats (Figure 40). Therefore, to take into account the doxycycline effect, when calculating the gene-to-gene fold change in the other three cell lines, we normalized the results to the corresponding EV FCs. What we observed in the average metagene profiles was confirmed when calculating gene-to-gene fold changes. In fact, all three cell lines showed average FCs values above one for both replicates (Figure 40). POLR2A cell line appeared to have the strongest phenotype among all three cell lines. The same trend observed in the two replicates of each cell line was efficiently recapitulated also when averaging the two replicates together (Figure 40E). Therefore, the next analysis were done on the average of the replicates.

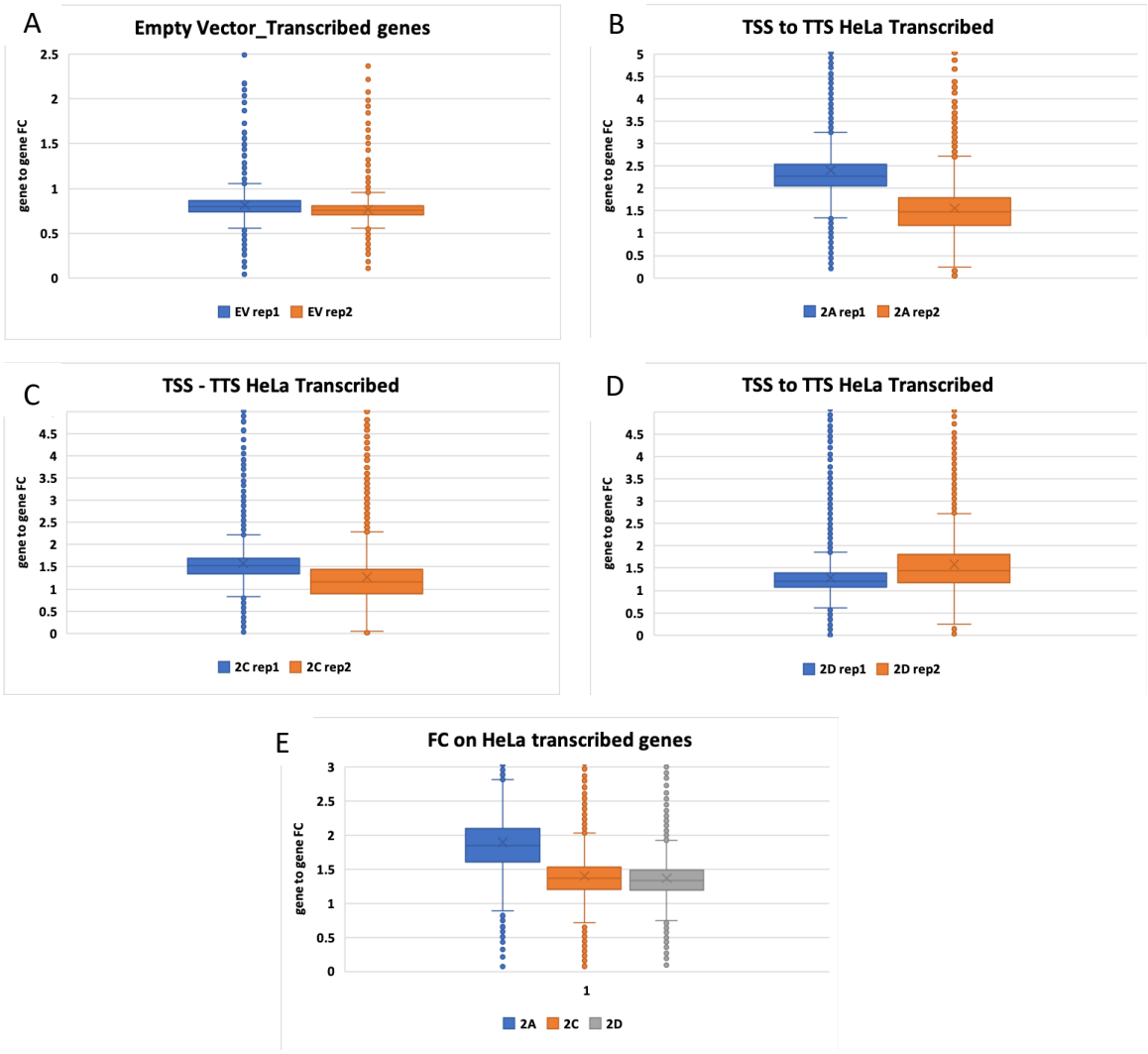


Figure 40 Gene-to-gene fold change quantification in induced vs uninduced calculated from TSS to TTS of transcribed genes in A) Empty vector uninduced, B) POLR2A, C) POLR2C D) POLR2D cell lines and E) gene-to-gene fold change induced vs uninduced from TSS to TTS of transcribed genes of averaged replicates.

To analyse which genes are more affected by transcription deregulation, we selected genes with a FC (dox/ctl) above or equal to 1.5 in the three cell lines and analysed the gene length distribution. Overall, with some variability, the average gene length of genes with highest transcription is between 10-30Kb, with a distribution from 10 to 70Kb, generally shorter than the average gene length (Figure 41) .

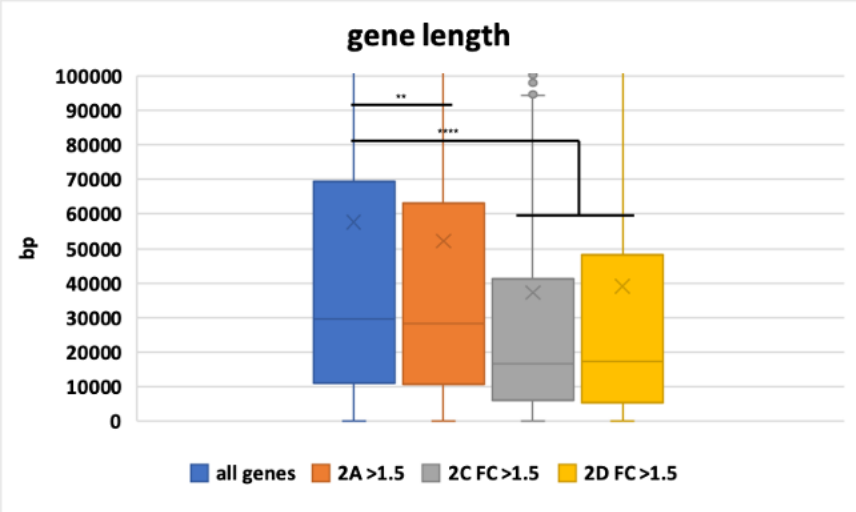


Figure 41 Average gene length of all genes and of genes with dox/ctl FC above or equal to 1.5 in POLR2A, POLR2D and POLR2D cell lines.

3.3.2 Characterization of transient transcription at the TSS, GB and TTS

To further characterize the transcriptional phenotype and better understand where in the gene the effect of the subunits overexpression was mainly impacting, whether more at the beginning, the middle or towards the end of the gene, we quantified TT-Seq levels at different locations. This would be important to identify potentially the specific phase(s) of the transcription process that were more affected, initiation, elongation or termination. For this, we analysed TT-Seq levels specifically around the transcription start site, (TSS + 1Kb), in the

gene body (GB, comprised between TSS +1Kb and TTS), and at the transcription termination site (TTS + 1Kb). Next, we calculated the gene-to-gene fold changes dox/ctl and normalized them to the EV dox/ctl (Figure 42A-C). As before, we looked at the average between the two replicates (Figure 42 left). Additionally, given that the average gene length distribution of genes most affected was comprised between 10 and 70Kb, we decided to specifically assess what happened to the longer genes (100Kb and more), to assess whether they would maintain the same general trend or they would behave differently from shorter genes (Figure 42 right). In the POLR2C, more than in POLR2A and POLR2D, cell line the higher FCs are at the TSS of the gene, with no difference between long and all genes (Figure 42B-E). In the POLR2D and POLR2A cell lines, the overexpression of the corresponding subunit is causing an increase in transcription activity at the TSS, less pronounced than in POLR2C (Figure 42A-C), even on long genes (Figure 42D-F).

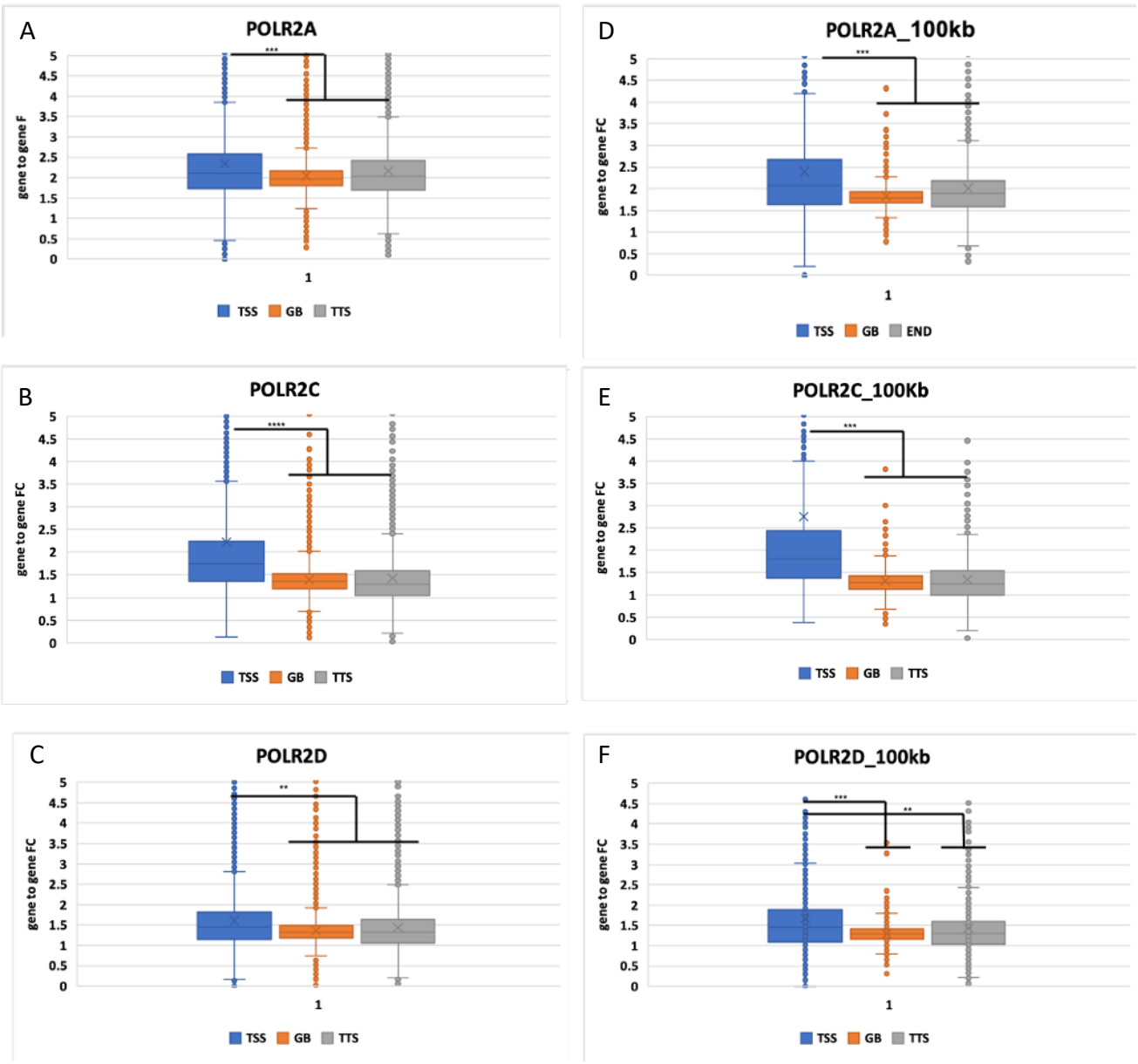


Figure 42 Gene to gene fold change quantification of transcription Left, Gene to gene fold change at the TSS + 1Kb, GB (TSS+1Kb to TTS-1Kb), and TTS-1Kb of induced vs uninduced on HeLa transcribed genes, in A) POLR2A, B) POLR2C, C) POLR2D cell lines. Right, Gene to gene fold change at the TSS + 1Kb, GB (TSS+1Kb to TTS-1Kb), and TTS-1Kb of induced vs uninduced on HeLa transcribed genes longer or equal to 100Kb, in D) POLR2A, E) POLR2C, F) POLR2D cell lines. All values were normalised to the Empty Vector.

3.3.3 Travel ratio

In order to better characterize if there is any possible defect arising at the beginning of the genes, in particular a difference between the amount of RNA Pol II that is loaded at the TSS that then progresses into active elongation, we decided to carry a further analysis and to calculate the travel ratio (TR) upon each subunit overexpression. To do so, we quantified the transcription levels on the gene list of all transcribed genes longer than 1Kb. The quantifications were done in two different regions of the gene, at the TSS (-50bp, +300bp) and at the TSS+300bp to TSS+1KB, in a strand-specific manner. For convenience, we named these regions TSS and GB respectively. To calculate the TR then the quantifications done at the GB region are divided by the quantification at the TSS, thus obtaining a ratio value. If this value is greater than one it indicates that more Pol II is accumulating at the TSS and not proceeding in the GB. Plotting TR values from the lowest to the highest for the different cell lines showed that on average overall, none of the subunit overexpression was changing the initiation-elongation transcription dynamics of cells (Figure 43).

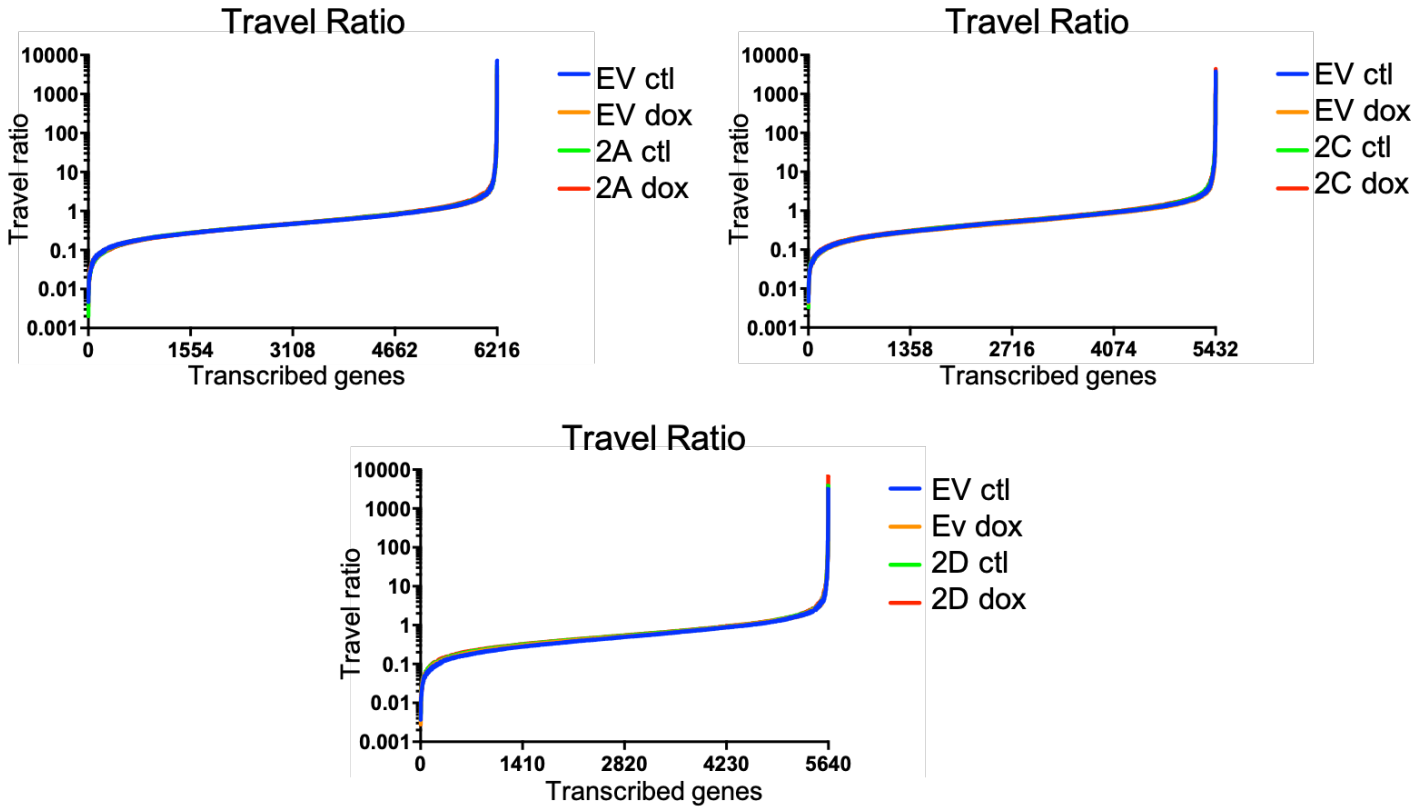


Figure 43 Travel Ratio in POLR2A (left), POLR2C (right) and POLR2D (bottom) cell lines and corresponding empty vector, ranked from the lowest to the highest.

To validate this further, we calculated the gene-to-gene FC of induced vs uninduced samples in both gene regions to see if we could observe a change specifically in either the TSS or GB region. This because a decrease in the GB region in respect to the TSS region could indicate that Pol II is stalling/pausing more at the TSS (and therefore there is an accumulation of reads) and not moving further into the gene, underlying a possible problem in promoter proximal pausing regulation; at the same time an increase in the GB region respect to the TSS region could indicate that the Pol II is instead stalling/pausing less at the TSS and moving more frequently into the gene, again indicating a deregulation of the promoter proximal regulation of transcription. As before, we did our analysis on both all transcribed genes and specifically long genes and there appears to be no impact of the overexpression of POLR2A subunit in Pol II escaping the promoter proximal pausing and moving into the gene body (Figure 44). This is line with our previous data showing increase transcription levels throughout the gene length. However, there seems to be increased accumulation of signal at the TSS of both POLR2C and even more in POLR2D cell lines upon subunits overexpression, indicating a possible increased stalling and release defect, both on short and long genes (Figure 44 right).

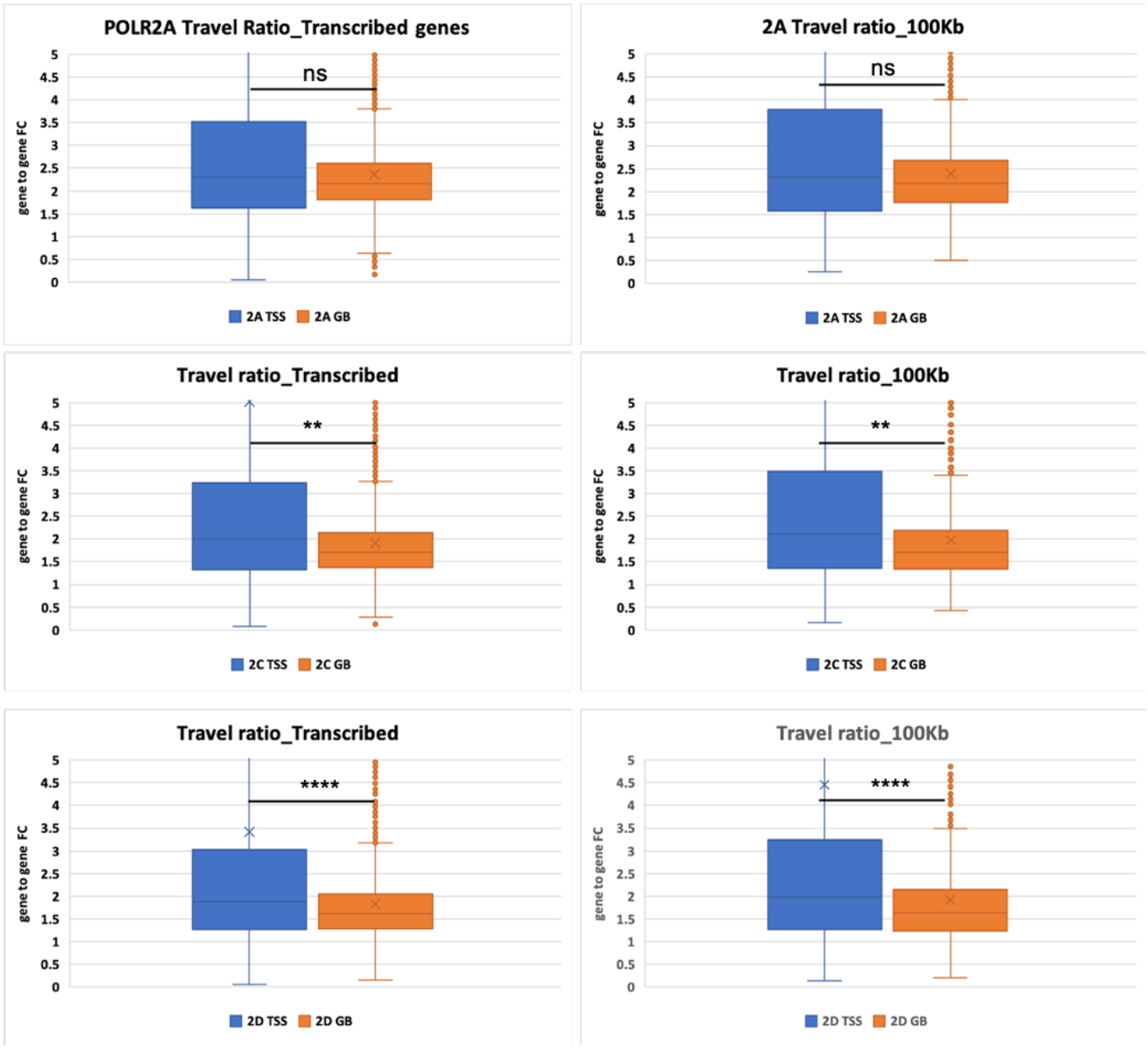


Figure 44 Gene to gene fold change quantification of travel ratio Left, gene to gene fold change at the TSS -50bp/+300bp (TSS), and TSS +300bp-1Kb (GB) of induced vs uninduced on transcribed genes, in A) POLR2A, B) POLR2C, C) POLR2D cell lines. Right, gene to gene fold change at the TSS -50bp/+300bp (TSS), and TSS +300bp-1Kb (GB) of induced vs uninduced on transcribed genes longer or equal to 100Kb, in D) POLR2A, E) POLR2C, F) POLR2D cell lines. All values were normalised to the Empty Vector.

3.3.4 Elongation and termination analysis

Our results show that the Pol II subunits overexpression had effect on changing the transition between initiation and elongation in the POLR2C and POLR2D cell lines, but not in the POLR2A cell line. Since these results are restricted to the very beginning of the gene and are indicative of the promoter proximal pausing stage, we expanded our analysis further into the gene, at a greater distance from the TSS, therefore looking at the elongation stage. To do so, we selected all genes longer than 5Kb and we calculated the FC dox/ctl at the TSS+2Kb, TTS, and the region in between (referred to as GB in the figures). We also repeated the same quantifications specifically on long genes (>100Kb) (Figure 45 right). Also in this case, the POLR2C and POLR2D overexpression affect how Pol II moves further in the gene (Figure 45). POLR2A overexpression affect how Pol II moves into the gene body, far from the TSS. The fact that this phenotype can be observed when looking far from the TSS (2Kb), and not with travel ratio in the case of POLR2A cell line, indicates that possibly POLR2A overexpression is interfering with transcription in productive elongation (Figure 45). The trend is also conserved on long genes (Figure 45 right).

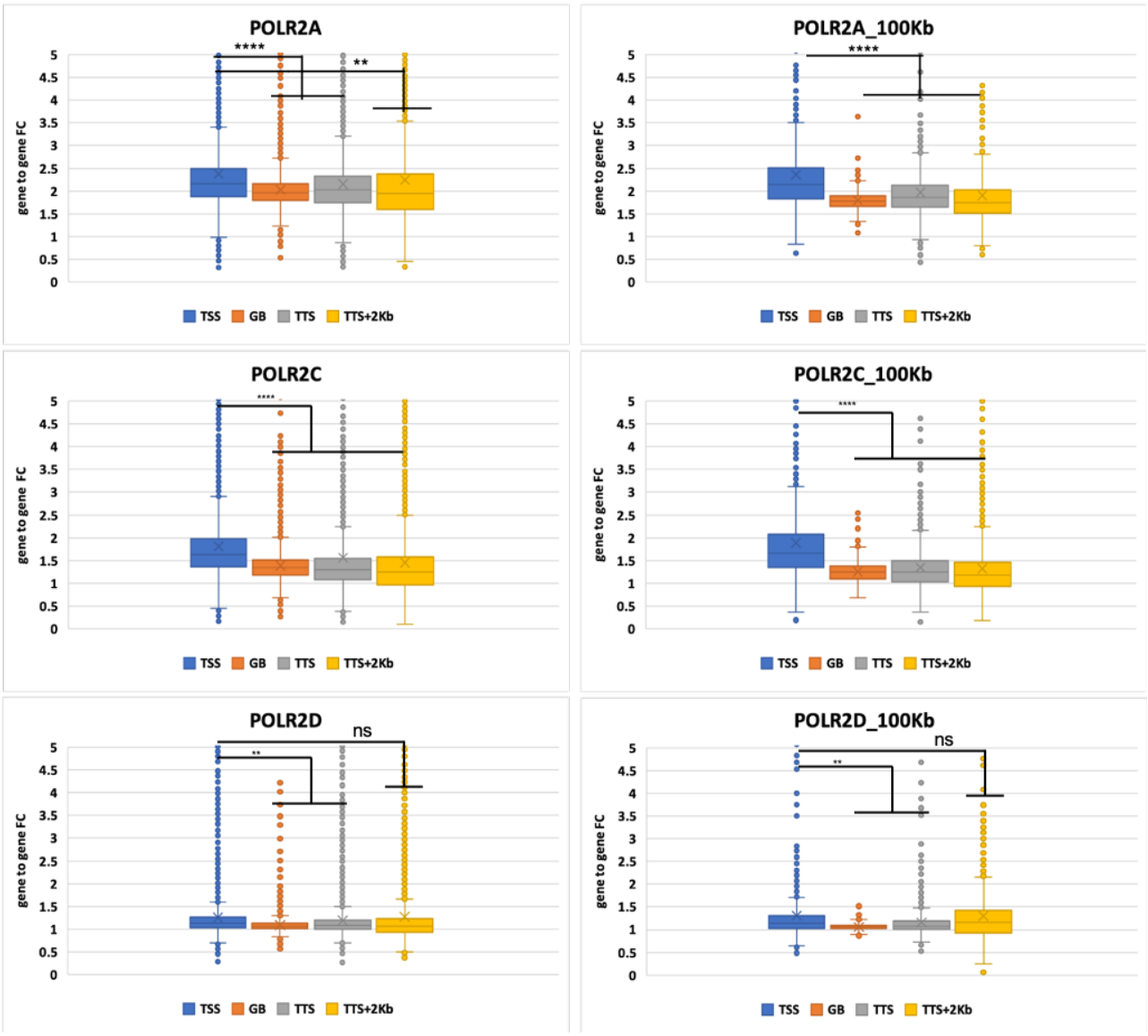


Figure 45 Gene to gene fold change quantification of transcription levels gene to gene fold change at the TSS+2Kb (TSS), TSS+2Kb to TTS (GB), TTS and TTS+2Kb of induced vs uninduced on transcribed genes longer than 5Kb, in POLR2A, POLR2C, POLR2D cell lines in all transcribed genes (left) and genes longer than 100Kb (right). All values were normalised to the Empty Vector.

Finally, we wanted to assess whether the increase of transcription observed upon subunits overexpression could be seen also past the TTS, maybe indicating increased Pol II readthrough past the termination sites. Therefore, we quantified in the region from the end of the gene to 2Kb (TTS+2Kb), on all transcribed genes and only on long genes (Figure 45). The increased fold change above 1 is observed also past the TTS, on all genes and long genes, with the same drop in TT-chem-Seq levels in respect to the TSS observed at GB and TTS. Surprisingly, POLR2D overexpression seems to cause some readthrough activity, especially on long genes (Figure 45 right), since there is no difference from the TSS quantifications, with still high levels of transcription activity.

3.3.5 γH2AX genes have more transcription

Given the DNA damage phenotype previously observed (3.1), we wanted to understand whether there was a link between the increase damage and deregulated transcription. To do so, we assessed how much TT-Seq signal is deregulated in the list of genes with increased γH2AX fold change (>1.5) identified in each cell line, identifying therefore the most affected genes upon each subunit overexpression. Once we had the gene lists we compared the transcription levels of the γH2AX affected genes and all the other genes. By doing so, we observed that independently of the cell line, the genes with more γH2AX showed also a slightly more increased in transcription levels compared to other genes, especially in the POLR2D cell line, indicating a possible link between transcription deregulation and DNA damage increase (Figure 46).

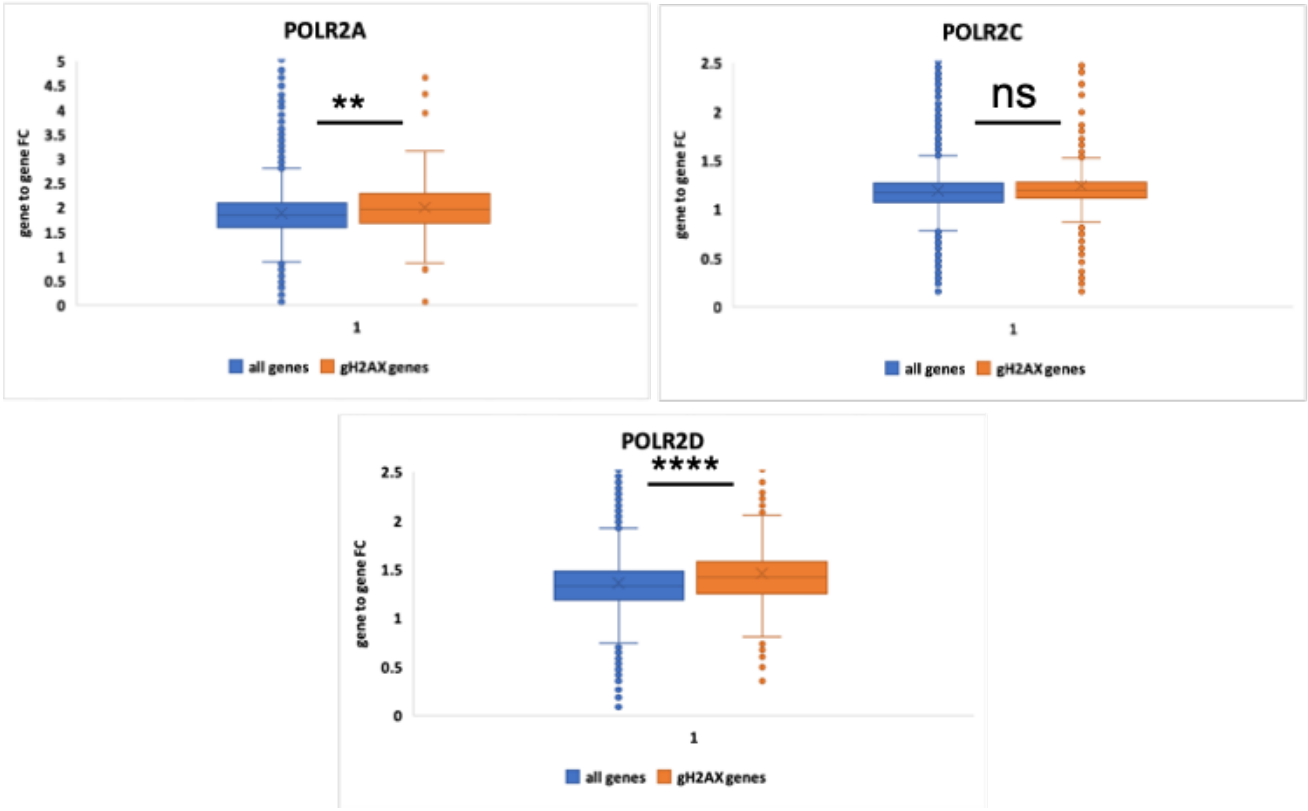


Figure 46 Transcription levels at genes affected by more DNA damage Gene to gene fold change from TSS to TTS of HeLa transcribed genes of all genes and of genes with γH2AX FC >1.5 in A) POLR2A, B) POLR2C and C) POLR2D cell lines

3.3.6 RNA Pol II ChIP-Seq identifies specific CTD phosphorylation defects

Transcription is a tightly regulated process and part of its regulation relies in the correct phosphorylation of its C-terminal domain (CTD). Therefore, we thought that the mechanism lying behind the transcriptional defects observed could be identified by studying the CTD phosphorylation upon subunits overexpression. To do so, we performed a Chromatin Immuno Precipitation followed by sequencing (ChIP-Seq) of the phospho-Ser5 and phosphor-Ser2 Pol II, using specific antibodies according to the phosphorylated CTD form. We chose these phosphorylated forms of the CTD being the most abundant and because they mark the initiation phase and elongation phase (Schüller et al. 2016). We also performed a ChIP-Seq of the total Pol II, using an antibody for the N-terminal domain (NTD), in order to verify Pol II occupancy on genes, unbiased from its CTD phosphorylations.

After the chromatin IP reaction and DNA purification, also 10% of the input DNA was purified. Firstly, we verified by qPCR the ChIP efficiency and specificity, calculating the percentage over the input on a house keeping gene and in a gene desert region. We used primers for the gene body of ACTB gene and primers amplifying a region on Chr2 where no genes are annotated (N2 region). We evaluated the percentage over the input for each IP reaction in both regions (Figure 47) to assess that the IP reaction was comparable between the induced and uninduced samples.

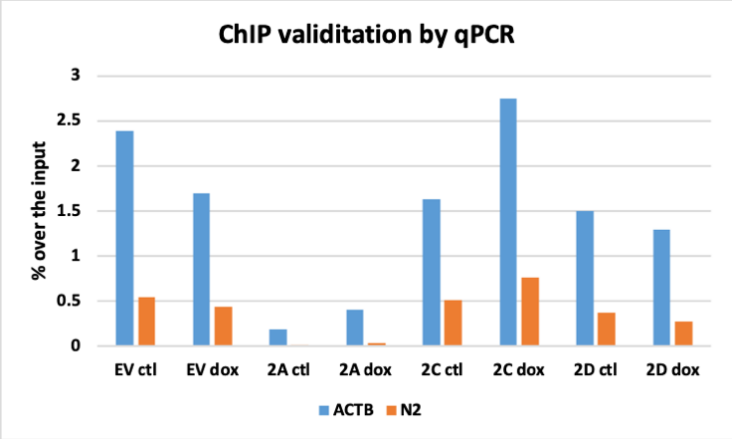
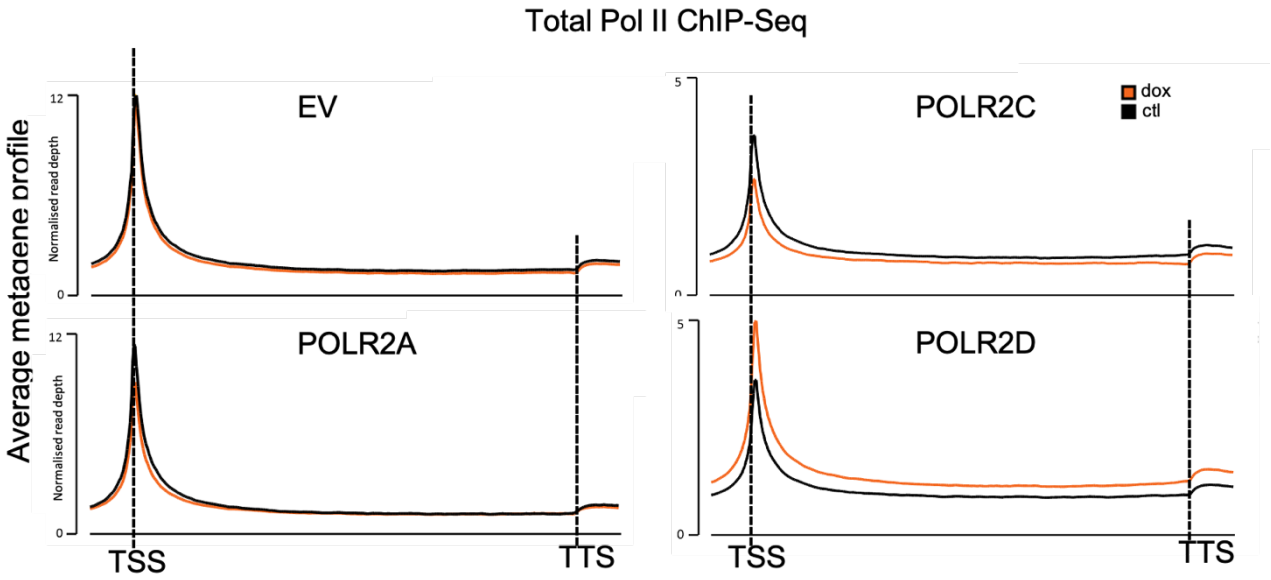


Figure 47 ChIP validation by qPCR Representative figure of ChIP qPCR validation. qPCR on ACTB gene body and gene desert region on Ch2 (N2) on phospho-Ser2 ChIP reactions in EV, POLR2A, POLR2C and POLR2D cell lines induced and uninduced.

After validating our ChIP reactions by qPCR, we proceeded to library preparation and sequencing. As previously done for our sequencing results, we generated the average metagene profile over HeLa transcribed genes, from the TSS to the TTS, and we compared the profiles in uninduced (in orange) and induced (in black) samples in the four cell lines (EV, 2A, 2C and 2D) (Figure 48). To our surprise, our ChIP-Sequencing results revealed important differences among our cell lines, with the overexpression of each subunit affecting Pol II in a specific and unique manner. For instance, while no difference can be appreciated between induced and uninduced samples upon POLR2C overexpression on the Ser5-CTD phosphorylated Pol II, it is striking the hyperphosphorylation on its CTD Ser2 residue upon the subunit overexpression, from the beginning to the end of the gene (Figure 48, black line). On the contrary, POLR2D overexpression appears not to affect the Ser2-CTD phosphorylation, however it seems to profoundly impair Ser5 phosphorylation from the beginning of the gene (Figure 48). Finally, the overexpression of POLR2A subunit has no effect, or only a very mild

effect, on either Ser5 or Ser2 phosphorylation. Importantly, these CTD deregulations observed, which vary according to which subunit is overexpressed, are not observed in the empty vector cell line (Figure 48). Furthermore, our ChIP-Seq metagene profile on total Pol II shows no particular difference between induced and uninduced samples, indicating that we were able to identify specific CTD defects and that the phenotypes observed are not biased by Pol II being more or less abundant in the induced samples. Importantly, a slight increase of total Pol II can be observed at the TSS of transcribed genes upon POLR2A overexpression (Figure 48).



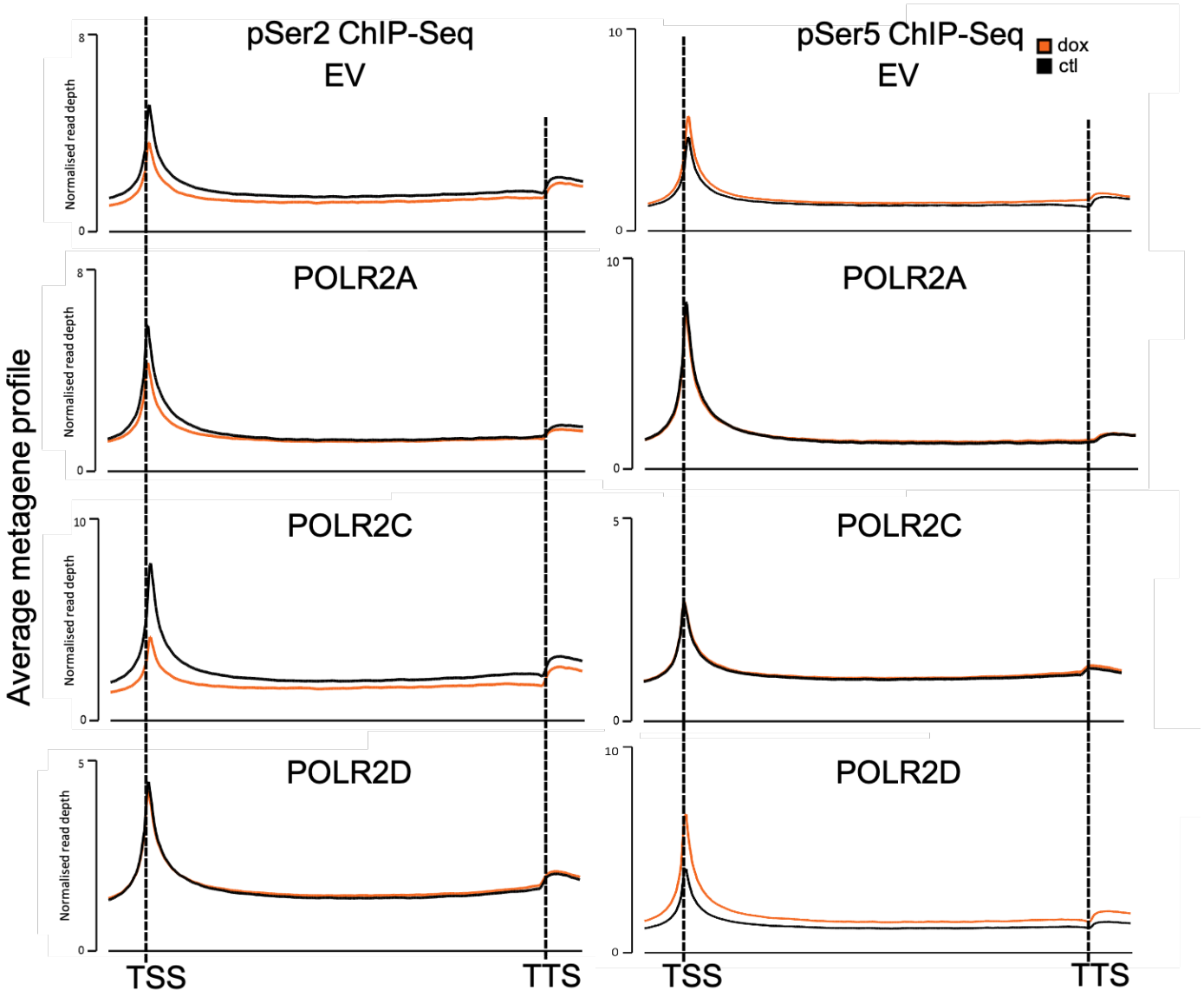


Figure 48 Metagene profile of ChIP-Seq for Total Pol II, phosphor-Ser2 (pSer2) and phosphor-Ser5 (pSer5) from the transcription start site (TTS) to the transcription termination site (TTS) on transcribed genes in HeLas in: Empty vector (EV) uninduced (ctl) in orange and induced (dox) in black, POLR2A ctl and dox, POLR2C ctl and dox and POLR2D ctl and dox.

As we have previously done for our other sequencing data, to be more precise with our analysis, we calculated the gene-to-gene fold change of the ChIP-Seq levels quantified on all transcribed genes, from the TSS to the TSS of the gene. As before, we also normalized our quantifications to the empty vector cell line, in order to consider the possible doxycycline effect. Strikingly, the gene-to-gene fold changes clearly show that each subunit overexpression induces a different Pol II deregulation: it can be appreciated that only in POLR2A cell line the fold change induced vs uninduced of total Pol II is above one, while there is no effect observed in the other two cell lines (Figure 49); POLR2C cell line shows increased pSer2 (fold change >1.5) upon induction, even if also in the POLR2D cell line there appears to be an increase in pSer2 FCs, which was not observed in the metagene profile (Figure 49); The overexpression of POLR2D clearly causes a drop in Ser5 phosphorylation (FC <1), while it can't be observed in the other two cell lines (Figure 49). These phenotypes are also conserved, even if slightly less prominent, on long genes (Figure 49 right). POLR2C cell lines and also POLR2C cell lines show an increase in the pSer2 Pol II, more pronounced in POLR2C cell line, upon subunits overexpression, which indicates increased transcription in the gene body, consistently with our TT-Seq results. Similarly, POLR2A overexpression increases the levels of total Pol II also in line with the TT-Seq results. However, it has to be considered that the TT-Seq analysis were done normalising all the samples to the yeast spike-in, which may have influenced the phenotypes observed. Further analysis without normalising to the spike-in are needed in order to better compare the TT-Seq and the ChIP-Seq results.

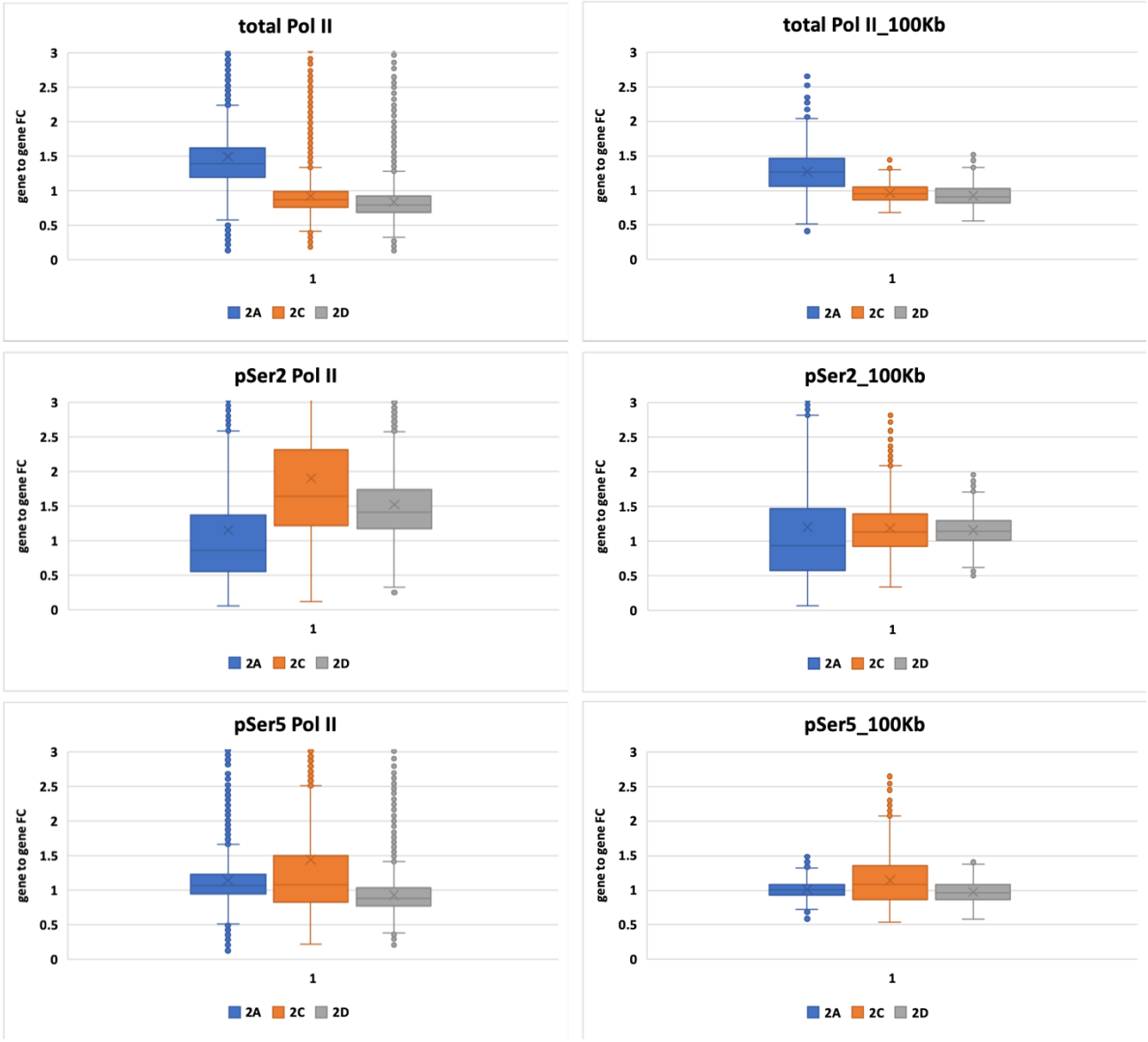


Figure 49 ChIP gene to gene fold change quantification Left, gene-to-gene fold change FC induced/uninduced in POLR2A (2A), POLR2C (2C) POLR2D (2D) cell lines of total Pol II, pSer2 and pSer5 ChIP-Seq level quantified in all transcribed genes from the TSS to the TTS. Right, as in left but on genes longer than 100Kb. All the FC are normalized to the EV and are the average of the two replicates.

3.3.7 Pol II deregulations and DNA damage

We identified each RNA Pol II deregulation happening upon specific subunit overexpression and we have showed that each subunits impacts differently on how Pol II and/or its CTD are deregulated. We wanted to link the Pol II phenotypes we identified with the increased DNA damage. For POLR2A cell line, we compared the gene-to-gene fold change of total Pol II and all transcribed genes with the fold change on genes with high γH2AX levels (FC >1.5). We found that the increased FC upon POLR2A overexpression on total Pol II levels is even greater in the genes affected by higher DNA damage (Figure 50). With the same approach, we compared the Ser2 and Ser5 phosphorylation levels upon the overexpression of POLR2C and POLR2D respectively, among all transcribed genes and γH2AX increased genes. Strikingly, the Pol II CTD deregulation is exacerbated on those genes that have higher γH2AX FC, in both cell lines (Figure 50).

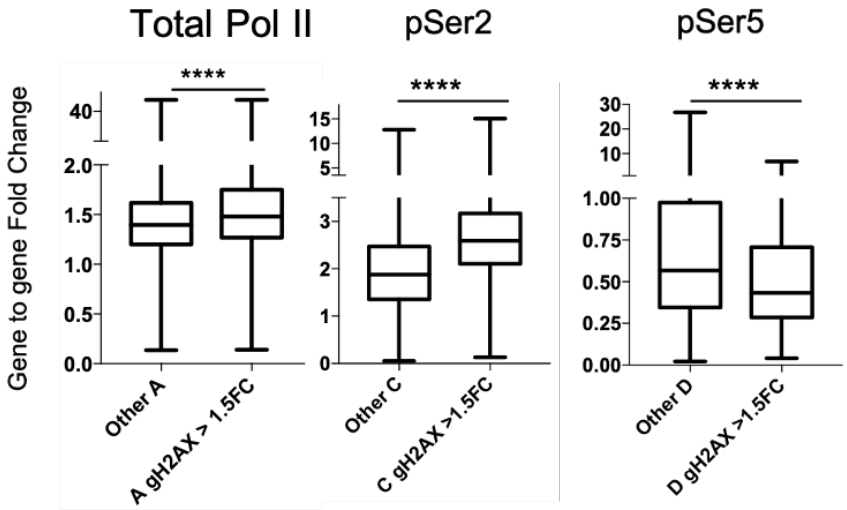
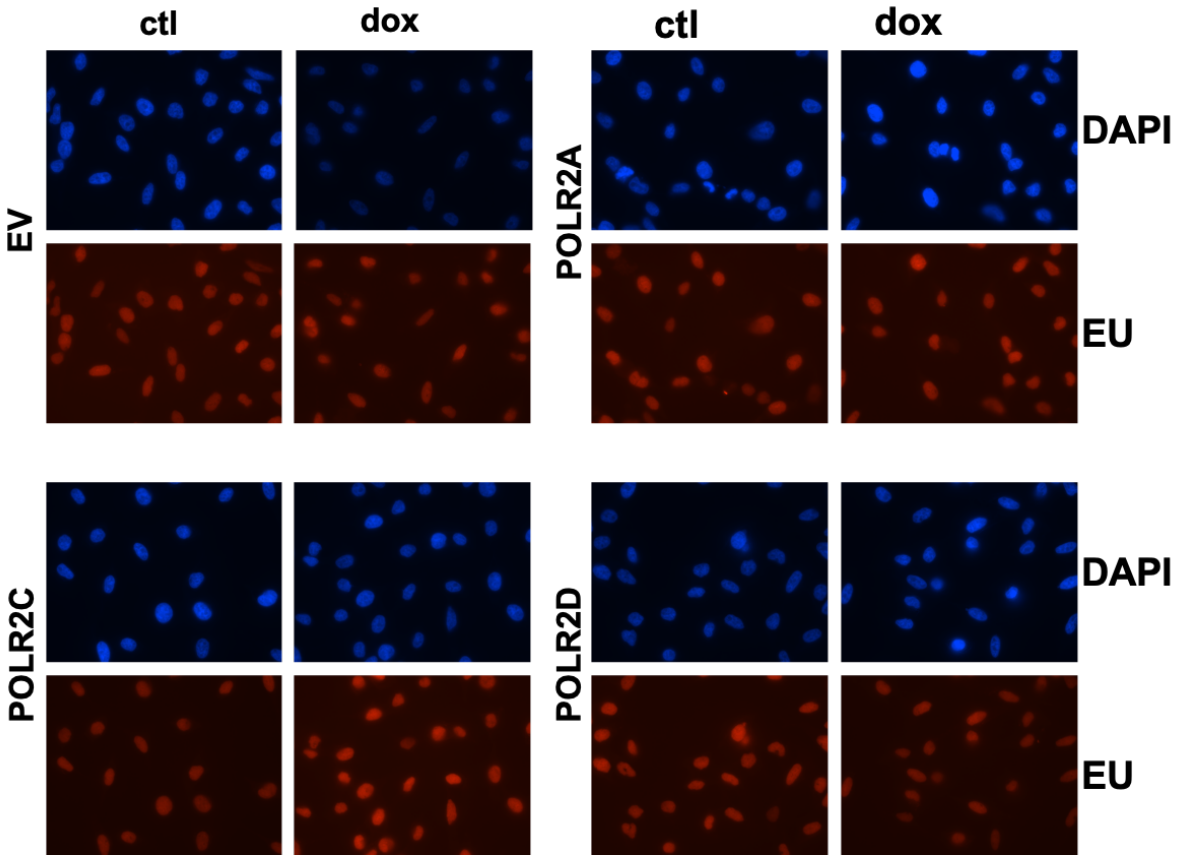


Figure 50 ChIP quantification at genes affected by DNA damage gene-to-gene fold change FC induced/uninduced over all transcribed genes and over genes with high (FC>1.5) γH2AX levels of A) total Pol II in POLR2A (2A); B) pSer2 ChIP-Seq in POLR2C (2C) and C) pSer5 ChIP-Seq in POLR2D (2D). The FCs are normalized to the empty vector and are the average of two

3.3.8 Transcription levels by EU labeling

An alternative approach to asses a transcription phenotype is by EU (ethyl uridine) RNA labelling followed by Click-it reaction. After 48h of induction, cells were labeled with 1mM EU for 1h at 37 degrees, before proceeding with staining and imaging.

Our results show specifically increase in EU nuclear intensity in the POLR2C cell line in the doxycycline induced sample, consistently with our results. However, in the two other cell lines there is a slight decrease upon POLR2A overexpression and a significant decrease upon POLR2D overexpression (Figure 51).



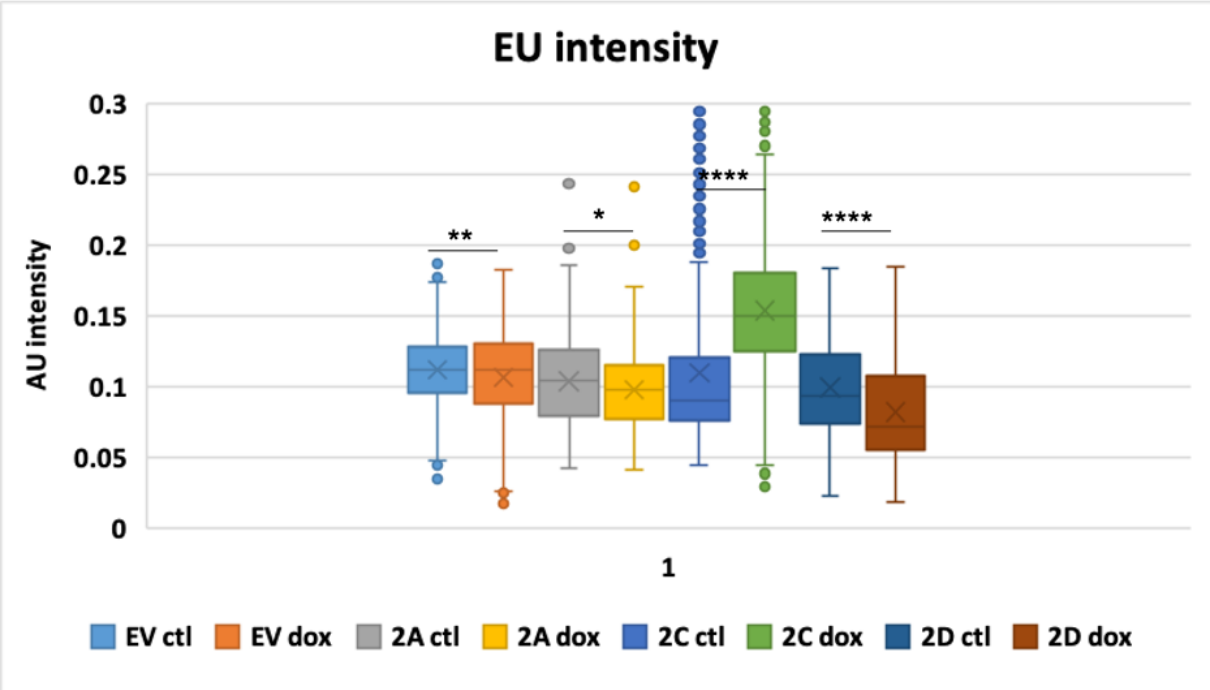


Figure 51 EU intensity analysis Above, representative figure of EU-Click-it microscope images. Bottom, EU intensity measured in the nuclei of EV, POLR2A, POLR2C and POLR2D cell lines uninduced (ctl) and induced (dox). N=3.

3.3.9 TT-Chem-Seq and ChIP-Seq single gene examples

Apart of presenting ensemble analyses of the different genome-wide experiments, below are snapshots of two housekeeping gene examples, ACTB and GAPDH, respectively reverse and forward strand, presenting the normalized TT-Chem-Seq profiles (Figure 52) and of the ChIP-Seqs (total Pol II, pSer-2 and p-Ser5-Pol II) (Figure 53), for all the four cell lines with and without subunits overexpression.

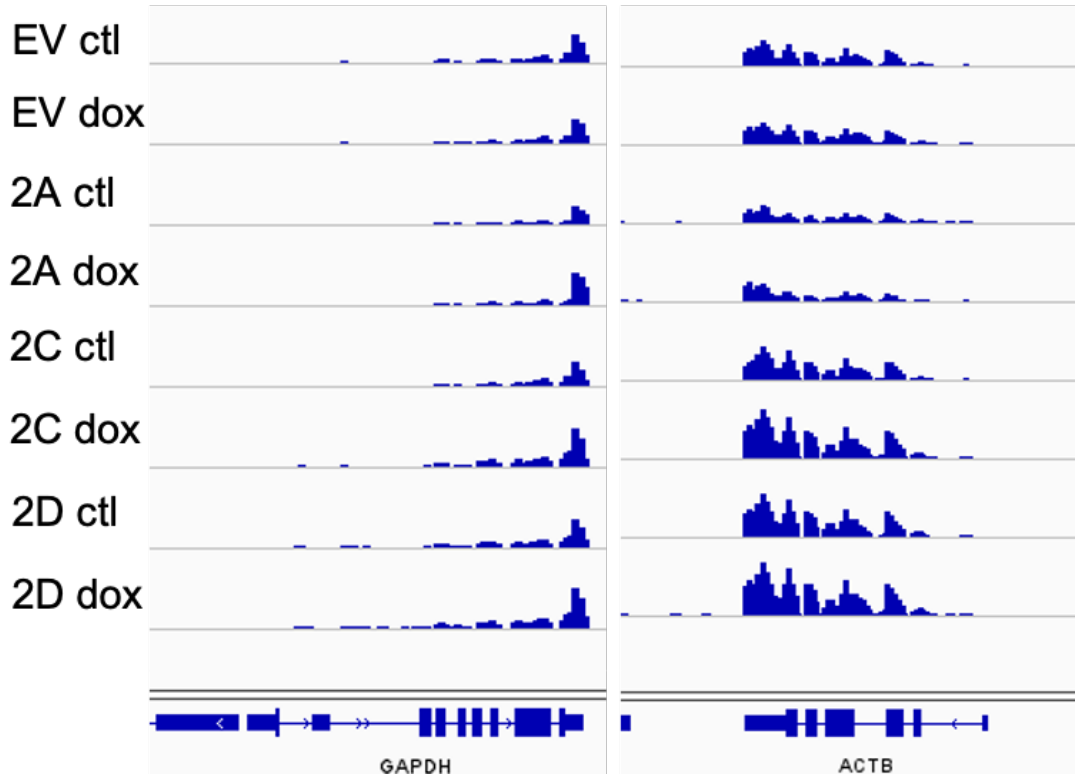


Figure 52 Single gene snapshot of TT-Chem-Seq forward (left) and reverse (right), in all samples uninduced (ctl) and induced (dox). A single replicate is shown as an example

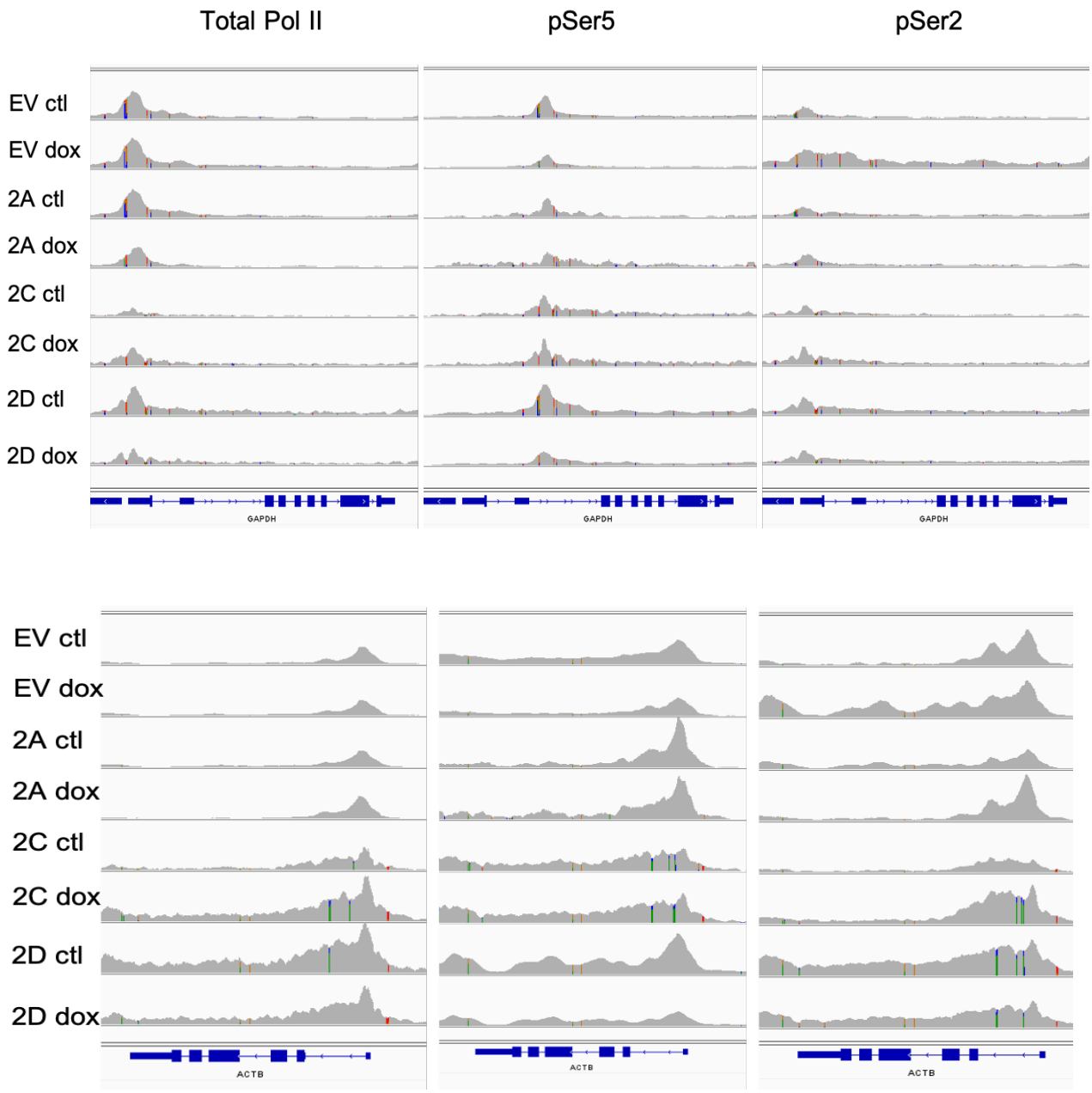


Figure 53 Single gene snapshots of ChIP-Seqs on GAPDH (above) and ACTB (bottom) in all samples uninduced (ctl) and induced (dox). A single replicate is shown as an example

3.3.10 Subunit overexpression interferes with correct TFs recruitment to chromatin

Giving the phenotypes observed with the TT-Chem-Seq and ChIP-Seq analysis, we decided to analyse recruitment to the chromatin of specific TFs, known to play a role in either transcription termination (XRN2), promoter proximal pausing and release into elongation (SPT5), or elongation (BRD4, phospho-SPT5) (Altendorfer et al. 2022; Hartzog and Fu 2013; Skourti-Stathaki et al. 2011). We found that the overexpression of POLR2C appears to impair the correct recruitment of BRD4 to chromatin. On the other hand, POLR2D overexpression affects SPT5 recruitment to chromatin while POLR2A overexpression results in XRN2 accumulation (Figure 54B). We also wanted to assess whether the transcription deregulations observed upon the overexpression of POLR2C and POLR2D could be corroborated by western blotting for the Ser2 and Ser5 phosphorylated forms of Pol II, respectively in POLR2C and POLR2D. Consistently with our sequencing data, we can see increased Ser2 phosphorylation in POLR2C dox treated sample compared to its control and similarly we detect less Ser5 phosphorylated Pol II in POLR2D dox sample, on the chromatin fraction (Figure 54C). Due to time constrain, the analysis of the CTD phosphorylation patterns and TFs recruitment to chromatin is only preliminary .

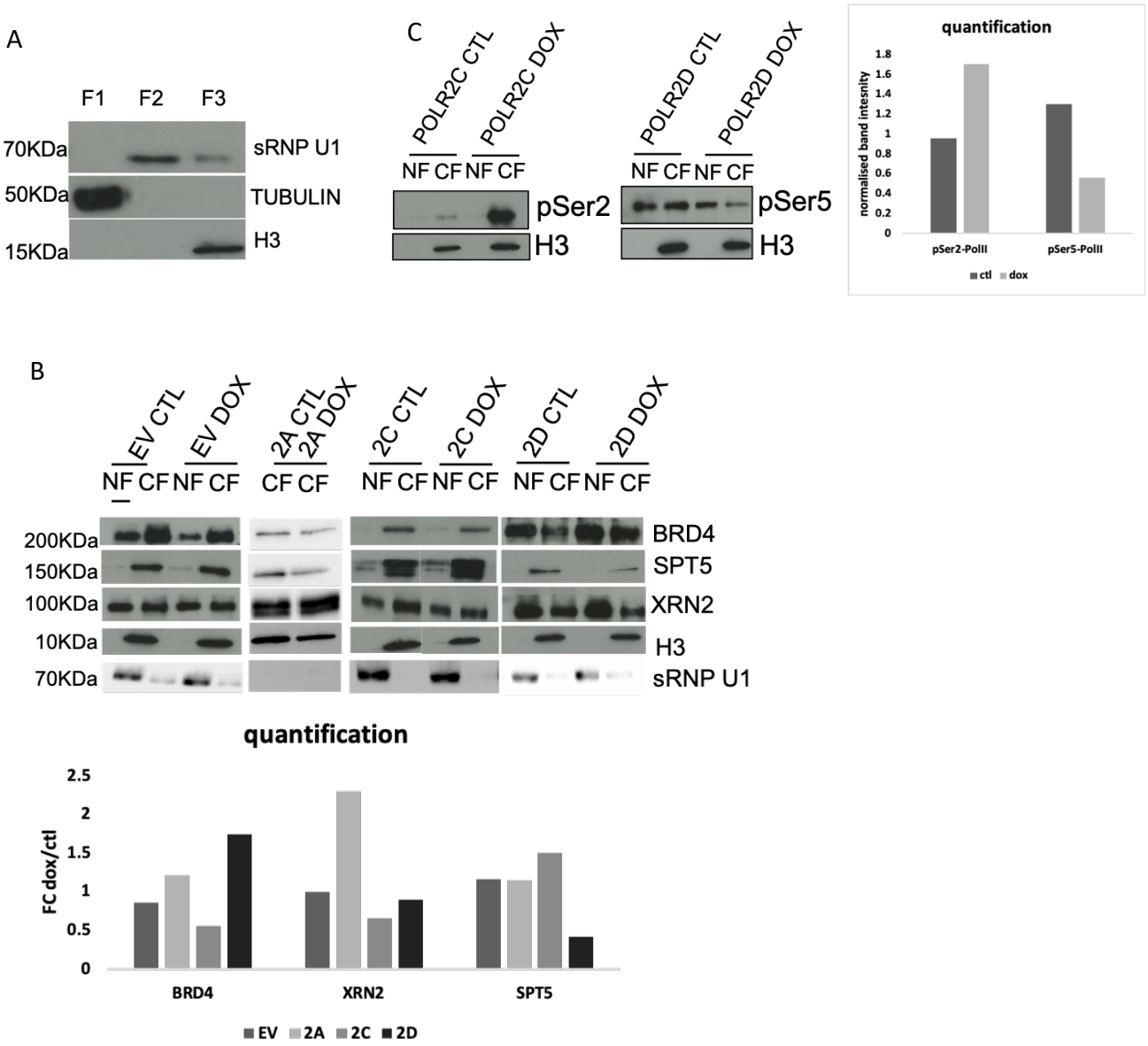


Figure 54 Recruitment of transcription factors to chromatin upon single subunits overexpression A) representative western blot to characterise the three cellular fractions, cytoplasmatic, nuclear and chromatin (F1, F2 and F3 respectively), blotting for TUBULIN in F1, sRNP U1 in F2 and histone H3 in F3. B)Representative western blot in the nuclear and chromatin fractions (NF and CF respectively) in EV, POLR2C, POLR2D cell lines and only CF in POLR2A cell line, induced and uninduced (ctl and dox), of BRD4, SPT5 and XRN2. Histone H3 and sRNP70 serve as controls of CF and NF respectively. C) Western blot on NF and CF in POLR2C and POLR2D cell lines induced and uninduced (ctl and dox) of pSer2 in POLR2C and pSer5 in POLR2D. H3 serves as loading control for CF. N=2

3.3.11 BRD4 overexpression rescues POLR2C overexpression phenotypes and can be specifically targeted

As shown by chromatin fractionation (Figure 54), the overexpression of POLR2C impairs the correct recruitment of BRD4 to the chromatin. BRD4 is a transcription factor belonging to the BET bromodomain containing protein family, and is involved in pTEFb activation and consequent Ser2 phosphorylation by CDK9 (Altendorfer et al. 2022) . Given the hyper Ser2 phosphorylation arising upon POLR2C overexpression, we hypothesized that the deregulation in BRD4 could be at least in part responsible for the phenotype observed, as there are evidences of treatment with the BET bromodomain inhibitor JQ1 increasing Ser2 phosphorylation (Bowry et al. 2018). To characterise this into more details, we performed an immunoprecipitation of total Pol II from the chromatin to analyse whether apart of the reduced BRD4 recruitment to chromatin, there is also impaired interaction of BRD4 with Pol II. Indeed, our cell fractionation (Figure 54) and the subsequent immuno-precipitation with total Pol II (Figure 55) show decreased BRD4 on chromatin and decreased BRD4 interaction with Pol II, respectively, upon POLR2C overexpression. Our initial hypothesis was that the overexpression of single subunits could affect the recruitments of specific TFs to transcribing Pol II. Hence, to understand whether the defective recruitment of BRD4 could be responsible, at least in part, for the transcription deregulation and the DNA damage phenotypes observed following POLR2C overexpression, we decided to check whether overexpression of BRD4 could rescue some of the phenotypes observed following the overexpression of POLR2C. Interestingly, the overexpression of BRD4 was sufficient to significantly reduced the EU intensity increase and to reduce the γH2AX intensity upon POLR2C overexpression (Figure

56A-B). Due to time constraints, we were not able to characterize the TF altered recruitment further in POLR2A and POLR2D cell line.

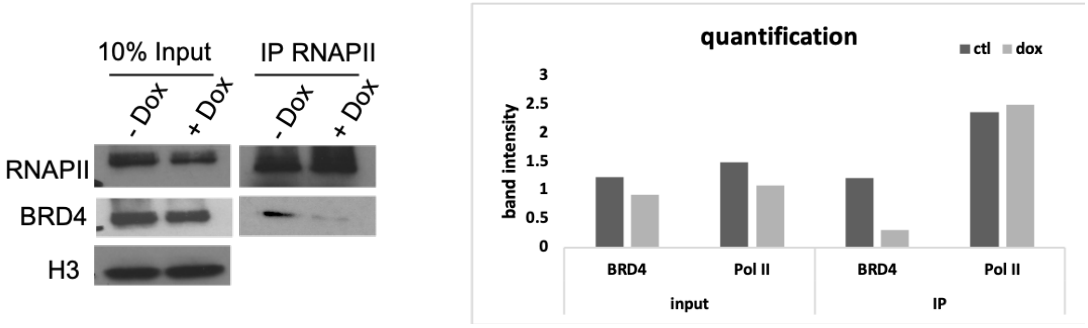


Figure 55 BRD4 Co-immunoprecipitation with Pol II western blot of BRD4 and Total Pol II (RNAPII) in Immuno precipitation of Total Pol II in POLR2C cell line uninduced (-dox) and induced (+dox).

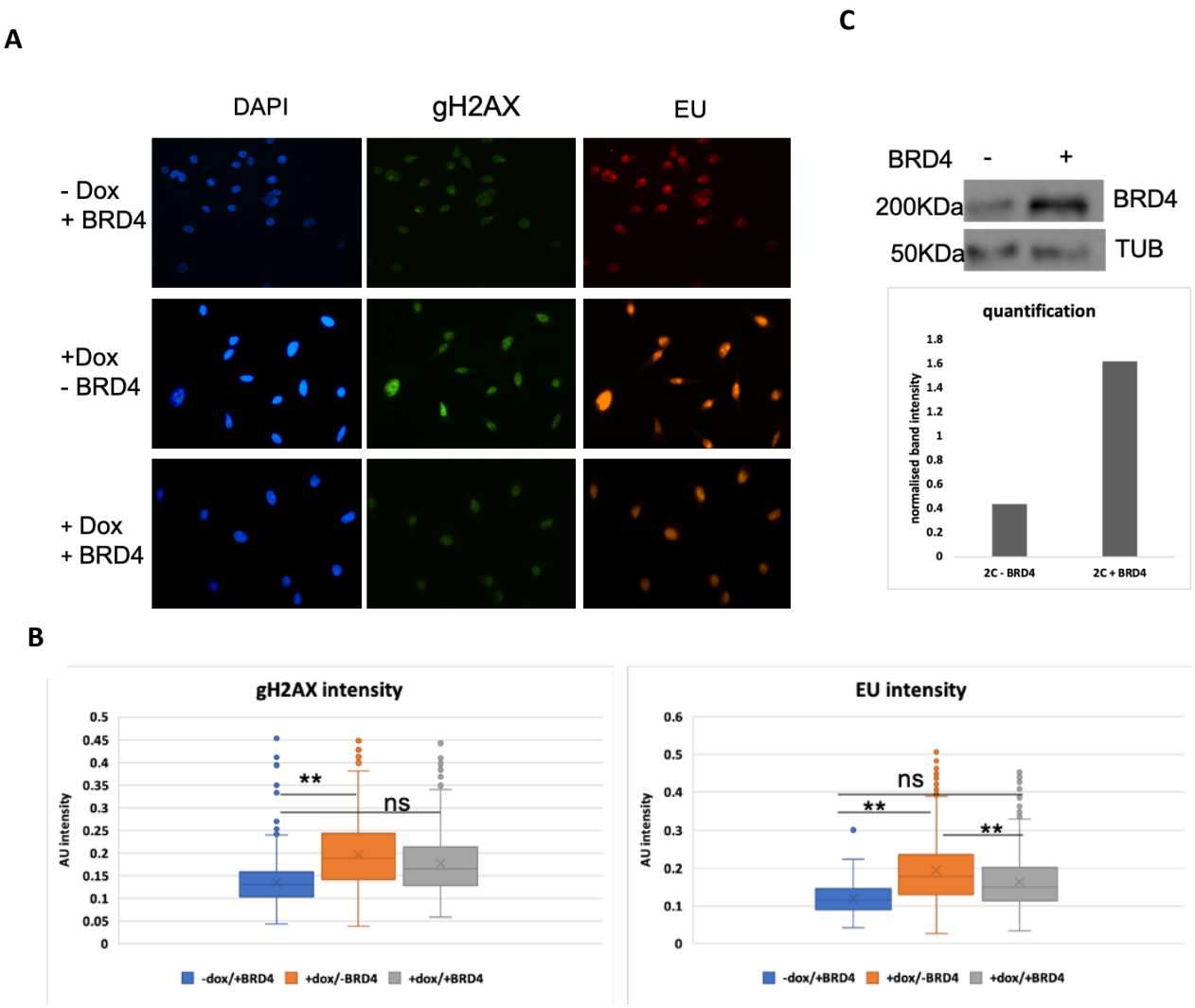


Figure 56 BRD4 overexpression rescues POLR2C hyper-transcription A) representative figure of microscope images staining for γH2AX and EU-clickit in POLR2C cell line untreated + BRD4, treated (dox) – BRD4 and treated (dox) + BRD4. B) Corresponding γH2AX nuclear intensity (left) and EU nuclear intensity (right). C) representative WB showing BRD4 overexpression, tubulin serves as loading control.

3.3.12 Transcription deregulations can be targeted

Our analysis showed that we were able to characterize the defective recruitment of BRD4 upon POLR2C overexpression, and the defective recruitment of SPT5 upon POLR2D overexpression. Moreover, we found an important Ser5 hypo phosphorylation upon POLR2D overexpression, altogether indicating perhaps a defect in CDK7 kinase activity and/or regulation. Therefore, we decided to assess in this study the sensitivity towards the inhibitors JQ1 and THZ1, because of their specific role. The first one, JQ1, is an inhibitor of BET bromodomain containing proteins, among which four is also BRD4 (Jiang et al. 2020). The latter, THZ1, is a specific inhibitor of CDK7, the kinase devoted to CTD phosphorylation on Ser5 (Kwiatkowski et al. 2014). We tested therefore cells' sensitivity to these inhibitors on all the cell lines after doxycycline treatment. We measured cell survival for 5 days after inhibitor treatment and we normalized our results to the DMSO control and to day 1. As cells were stained with crystal violet, cell survival was measured by measuring the optical density with a plate reader of the crystal violet stain. We found that the inhibitors had the strongest effect on day 2 and that then this effect was lost unless we treated cells again on day 3 (data not shown). Therefore, we calculated the survival on day 2. Strikingly, the effect of JQ1 is strongest on cells that overexpress POLR2C, while it has no effect on POLR2D and POLR2A overexpressing samples. On the other hand, THZ1 significantly sensitizes POLR2D overexpressing cells, while it has no significant effect on POLR2D and POLR2A cell lines with or without induction (Figure 57). POLR2A induced samples show a slight sensitivity to both drugs, but it is not significant.

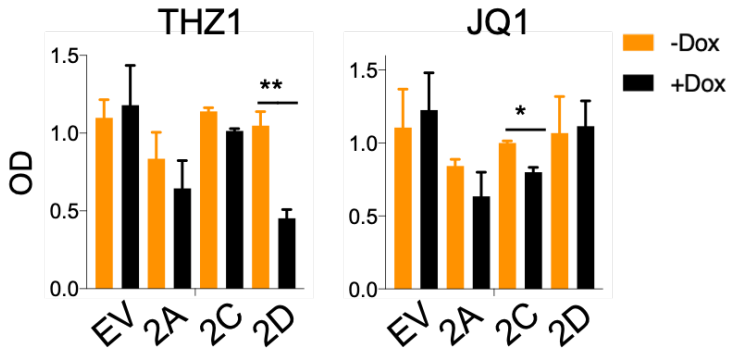
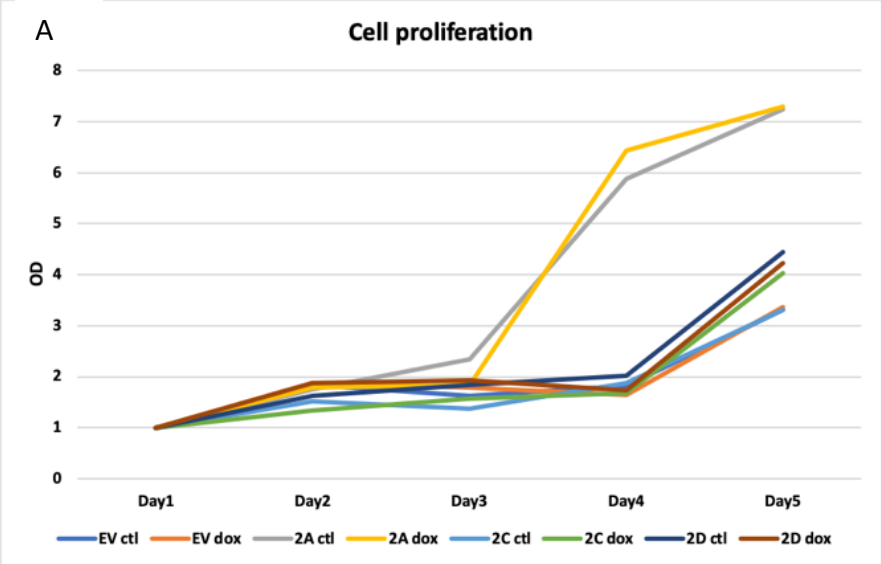


Figure 57 Drug resistance analysis upon subunits overexpression THZ1 and JQ1 impact on cell survival in empty vector (EV) POLR2A-2C-2D cell lines uninduced (-dox) and induced (+dox) at day2 post treatment. OD= optical density

Next, we used our data from the DMSO treated samples in order to assess whether the overexpression of the single subunits itself could impact on cell proliferation. We did not find any difference between induced and uninduced samples (Figure 58A). Given the variability between experiments, we calculated the fold change (dox/ctl) on each day for all four cell lines. Only POLR2D overexpression appears to slightly reduce cell proliferation, though not significantly (Figure 58B). These data are in line with our FACS data, which showed no cell cycle defect arising upon subunits overexpression.



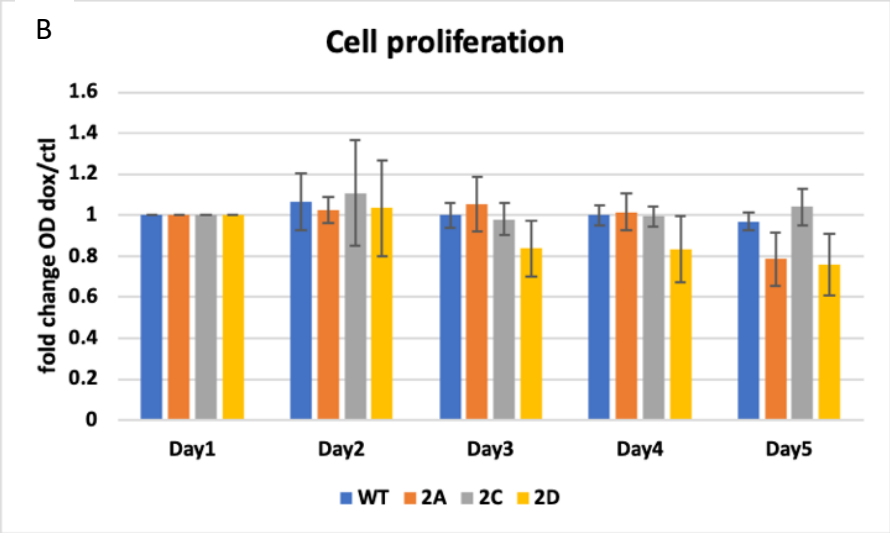


Figure 58 Cell proliferation A) representative image of cell proliferation curve on 5 days measurements in all four cell lines treated (dox) and non-treated (ctl). B) Fold change of cell growth (measured as OD) between treated and non-treated samples at each day over a period of 5 days.

3.3.13 Discussion

In this chapter we have described how the upregulation of the three Pol II subunits (POLR2A-2C-2D) affects transcription regulation. Moreover, we began to characterise the mechanism lying behind transcription deregulations to potentially exploit it therapeutically by targeting it with commercial inhibitors. The first approach we decided to pursue for transcription characterization is the TT-chem-Seq, which is the sequencing of transient transcription. There are different approaches to track nascent RNA synthesis, as reviewed by Wissink and colleagues (Wissink et al. 2019). The chromatin associated RNA Sequencing (ca-RNA-Seq) (Bhatt et al. 2012), which uses strong salt washes to separate mature nuclear and cytoplasmatic RNA from the RNA still bound to chromatin (nascent). It is indeed a powerful tool when combined with other methods investigating co-transcriptional processes, however, it can also capture some mature RNAs (such as lncRNAs) that are strongly bound to chromatin. For example, this method can be combined with the enzymatic digestion of RNAs not protected by 5' cap to provide high resolution information on TSS and initiation. Similarly, NET-Seq (native elongating transcript sequencing) (Nojima et al. 2015) captures the RNA bound to Pol II, thus enriching specifically for Pol II nascent RNA, excluding Pol I and Pol III transcription; however, it is a method biased by the efficiency of Pol II immuno-precipitation. Indeed, the choice of antibody (against specific CTD phosphorylated forms for instance), can select for specific phases of transcription. Other methods for nascent transcription sequencing rely on isolating RNAs that are undergoing synthesis, i.e. RNAs from transcriptionally competent Pol II. Among them, worth of mentioning is GRO-Seq (Global run-on Sequencing) (Leighton et al. 2008), which relies on the labelling of nascent RNA with 5-bromouridine 5'-triphosphate (BrUTP) and its isolation by BrUTP pull down, and the PRO-Seq (Precision run-on Sequencing)

(Kwak et al. 2013), which instead incorporates biotin-labelled terminating NTPs, generating nucleotide resolution maps and reducing background signal. While in the run-on RNA labeling occurs in vitro, the TT-Seq (Gregersen et al. 2020)(method used in this work) relies on the metabolic reaction occurring in cells that incorporate a modified nucleoside, such as 4SU, by active transcription. This technique does not allow single nucleotide resolution, but it allows to measure the RNA synthesis in living cells. While other methods can be Pol II specific or have nucleotide resolution, TT-Seq is the only method that allows capturing RNAs from actively transcribing Pol II without having to extract the nuclei like in GRO/PRO-Seq, a process that can cause stress and interfere with transcription regulation by removing regulatory factors from Pol II interaction. The main limitation of this method is the 4SU incorporation itself, which can't be shorter than 5 minutes, and the longer it is done for the higher the chance of capturing co-transcriptionally processed RNAs (Wissink et al. 2019 and citations there in). We thought it would have been the best method to identify transcription alterations arising upon subunits overexpression. This method, like the others mentioned, primarily looks at Pol II produced nascent RNAs, hence we did not assess the impact of subunits overexpression on Pol I or Pol III transcription. Given that Pol II can play a role in regulating Pol I and Pol III transcription, and that some transcription factors and Pol II subunits itself are shared between polymerases (Arimbasseri et al. 2013; Cramer et al. 2008; Gerber et al. 2020; Michels and Hernandez 2006; Sanij et al. 2015), we do not exclude that the overexpression of the single subunits could also impact on Pol I and Pol III transcription.

Another method widely used to characterize transcription is ChIP-Seq for RNA Pol II. With this method we identify the regions on the DNA that are bound by Pol II. ChIP-Seq allowed us to characterize Pol II CTD modifications. These are very important throughout the different

phases of transcription, and specific CTD modifications identify the different transcription stages (Heidemann et al. 2013). Combining ChIP-Seq with TT-chem-Seq allows us to identify deregulations in the transcription process itself, looking at actively transcribing Pol II and any specific defect affecting the polymerase in different parts of the gene.

From our nascent RNA sequencing results, we observed that the overexpression of Pol II single subunits increased the transcription activity, especially upon the overexpression of POLR2A and POLR2C overexpression (Figure 39). However, the metagene profiles represent an average coverage of quantitative data from the TSS to the TTS of the genes, so as an average some phenotypes could be over or under-estimated if only a group of genes is particularly affected. With the quantification of the sequencing signal in different part of the gene, from the TSS to the TTS, we could look at how transcription is generally affected and what happens when moving from the start of the gene towards the end of the gene. With the gene-to-gene fold change between induced and uninduced samples, from the TSS to the TTS of transcribed genes, we confirmed the increased transcription levels observed in the metagene plots and then we assessed which genes, shorts or longs, were more affected (Figure 40; Figure 41). First of all, we selected all the actively transcribed genes and calculated the median gene length, which we compared to the gene length of genes with increased transcription levels in the POLR2A-2C-2D cell lines. While the genes with increased fold change (dox/ctl) in POLR2A cell lines show only slight decrease in regard to gene lengths to all genes transcribed, those affected in POLR2C and POLR2D cell lines are significantly shorter (Figure 41). Long and short genes generally undergo different transcription regulation, therefore they can be impacted differently. For instance, many house-keeping genes are short genes and are ubiquitously highly transcribed. Non house-keeping genes, on the contrary, are generally longer than

house-keeping, and are regulated with a wide plethora of TFs according to environmental stimuli, cell cycle stimuli, in response to DNA damage, are subjected to increased stall/pausing (Bertoli et al. 2013; Bregman et al. 2000; Brown 2021; Chiaromonte et al. 2003; Cho et al. 2001; Cosma 2002; Eisenberg et al. 2003; Grishkevich and Yanai 2014; Lopes et al. 2021; Yonaha et al. 1995). For this reason, we decided in our analysis to characterize transcription on all transcribed genes but also to specifically analyze longer genes, which may show phenotypes not observed on shorter genes, or sustain the same phenotypes. When quantifying TT-chem-Seq fold change levels in different regions of the gene (TSS, GB and TTS), we observed that transcription levels seem to drop in respect when moving in to the gene body compared to how much there is at the beginning of the gene, upon the overexpression of POLR2A-2C and to less extent POLR2D subunit, especially on long genes (Figure 42). A change between how much Pol II is found at the start of the gene compared to the gene body can be indicative of changes in the regulation of the transition between the initiation and elongation phases of Pol II transcription, or highlight an inability to bring to completion all transcription events, because of an increase in stalling/pausing of Pol II along the way, a defect basically in the processivity of Pol II (Boettiger et al. 2011; Core and Adelman 2019). However, calculating the travel ratio (Reppas et al. 2006; Saponaro et al. 2014) in each cell line showed no differences between induced and uninduced samples, indicating that the drop in the gene body may not be due to an increased promoter proximal pausing and/or overall changes in the levels of Pol II loaded at the TSS that will enter productive elongation (Figure 43). However, the quantifications at the TSS and in the region between TSS+300bp and TSS+1Kb (GB region), there is a significant drop in TT-Seq levels in the GB region in POLR2C-2D cell lines, but not in POLR2A that shows no defect in escaping the promoter region (Figure 44). From the

quantification on start, middle, and end of the gene, we understand that subunits overexpression increase overall transcription activity, but there seems to be a drop after the TSS+1Kb in the transcription activity, more pronounced in long genes and upon POLR2A and POLR2C overexpression. These observations seem to indicate that Pol II may be encountering issues at some point after the very beginning of the gene (1kb), but not be affected by PPP specifically. Pol II is paused at the PPP also to give time for correct recruitment of TFs acting co-transcriptionally in the following stages of transcription (Arnold et al. 2021; Fong et al. 2017; Kuehner et al. 2011; Nojima et al. 2018). Thus, Pol II may be similarly released from its paused state when subunits are overexpressed, but its interaction with the splicing, elongation or termination components may be affected, thus interfering with its correct movement into the gene. Indeed, POLR2D quantifications at PPP sites (Figure 44), show that the overexpression of this subunits is causing an accumulation of signal at the TSS specifically. The fourth largest subunit of the Pol II complex, as being part of its stalk domain, interacts and regulates co-transcriptional processes, such as splicing and 3'end processing (Allepuz-Fuster et al. 2014, 2019; Compe et al. 2019; Garrido-Godino et al. 2016; Rengachari et al. 2021; Richard et al. 2019, 2021). For this reason, we also decided to look at the gene body even further away from the TSS, 2Kb past the TSS, when Pol II is productively elongating. We also quantified the TT-chem-Seq levels 2Kb past the TTS, in order to evaluate if Pol II encounters any termination defect upon subunits overexpressions (Figure 45). Especially on longer genes, the nascent transcription levels appear to be higher in the first part of the gene, while there is a drop further away from the start site, indicating that transcription levels decrease far from the TSS, as mentioned above, indicative of a potential processivity problem. In POLR2D cell line this elongation drop is not as pronounced as in the other cell lines (Figure 45), but it is the

only cell line showing increased FCs after the termination site, possibly indicating a defect in the termination of Pol II. This is in line with the afore mentioned role of POLR2D in regulating co-transcriptional processes.

Overall, we observed increased nascent transcription levels upon the single subunits overexpression, with each subunit showing a slightly different phenotype from the other. POLR2A and POLR2C highest transcription levels appear to be restricted at the TSS of transcribed genes and more on long genes, with POLR2D showing a possible termination defect, also on long genes. POLR2A and POLR2C hyper transcription is found mainly at the TSS of the gene and also at 2Kb after the TSS we can observe a drop in transcription levels, indicating a possible elongation defect. The phenotypes observed are in line with the γH2AX ChIP-Seq profiles, showing increased levels at the TSS and more towards the TTS in POLR2A-POLR2C and POLR2D, respectively.

While the TT-chem-Seq identified the impact of the subunits overexpressions on actively transcribing Pol II, the ChIP-Seq identified specific alterations on Pol II itself. We know that the CTD can be phosphorylated all over its heptad repeat, with each phosphorylation being fundamental for the correct transcription process (Schüller et al., 2016; Suh et al., 2016). We performed ChIP-Seq for the main two phosphorylated forms of Pol II, Ser5 and Ser2, that decorate the CTD when Pol II moves from the transcription initiation to productive elongation (Suh et al. 2016).

POLR2A overexpression causes an increase in total Pol II presence on transcribed genes (Figure 48; Figure 49), which is consistent with what observed in the TT-Chem-Seq. When analyzing which transcription factors may be differentially recruited on chromatin upon POLR2A overexpression we found that XRN2 is more present on chromatin upon doxycycline induction

(Figure 54B). This data is preliminary, but it could point towards a premature transcription termination. Indeed, we observe increased Pol II only at the TSS of the genes and a drop in nascent transcription levels past the beginning of the gene (Figure 48). Despite being able to characterize the transcription defects arising upon POLR2A overexpression we were not able to define the specific mechanism behind the phenotypes observed, therefore we failed in identifying an inhibitor that could have targeted the subunit overexpression in a specific manner. Further studies to better address the possible termination defect in POLR2A dox samples are needed to understand how to target this Pol II deregulation for therapeutic purposes.

POLR2C overexpression results in hyper-phosphorylation of the Ser2 residue, which is reduced on genes higher than 100Kb. Indeed, we noticed a decrease of transcription activity in the gene body (Figure 45 left) and especially on long genes by TT-Chem-Seq (Figure 45 right). Ser2 phosphorylation is highly dependent on the activity of CDK9, the kinase module of the positive elongation factor pTEF-b (Buratowski 2009; Fuda et al., 2009; Vos et al. 2018; Yang et al. 2005). Ser2 phosphorylation is also dependent on the activity of the transcription regulator BRD4, that positively regulates pTEF-b and stimulates transcription (Chen et al. 2014; Zhou et al. 2022). Albeit BRD4 being considered a positive regulator of transcription, and evidence in support of this show how knock down or inhibition of BRD4 reduces transcription levels and pSer2-Pol II (Devaiah et al. 2012; Khoueiry et al. 2019; Yang et al. 2005), there is also evidence showing that reduced BRD4 levels/activity associated with hyper transcription and readthrough (Arnold et al. 2021; Bowry et al. 2018). It is suggested that BRD4 KD could impact differently on different classes of genes and that its loss could trigger a compensatory mechanism relying on HEXIM1 mediated pTEF-b regulation (Bowry et al. 2018), resulting in

hyper Ser2 phosphorylation and transcription. By cell fractionation and immune precipitation our preliminary data indicate that BRD4 is defectively recruited to chromatin (Figure 54B) and to Pol II (Figure 55) upon POLR2C overexpression. We hypothesized that this defect could be responsible for the phenotypes observed by ChIP-Seq and TT-chem-Seq, and by EU staining as well (Figure 51). Given the multiple roles of BRD4 in transcription elongation regulation (Altendorfer et al. 2022), its reduced interaction with Pol II could be also responsible further into the gene for the phenotype observed in TT-chem-Seq levels drop past the TSS. Our hypothesis of BRD4 impaired Pol II interaction being crucial for the phenotypes observed was confirmed with our rescue experiments with γH2AX staining and EU staining (Figure 56). Further experiments are needed to prove how BRD4 is scavenged away from Pol II on chromatin by the overexpressed POLR2C. It would be important to show that BRD4 is retained in the nuclear fraction together with POLR2C and to rescue the hyper phosphorylation of Ser2 by ChIP and western blot. By ChIP for BRD4 upon POLR2C overexpression we could identify which genes promoters/enhancers are less bound by BRD4 and assess whether those genes are the same in which we observe more Ser2 phosphorylation in POLR2C dox sample. We then tested if the BRD4 defect could be targeted by the commercial inhibitor JQ1, which is widely used (Donati et al. 2018; Singh and Alauddin 2023). Strikingly cells that overexpress POLR2C are also more sensitive to the BRD4 inhibitor (Figure 57), further confirming that its impaired recruitment/interaction is the mechanism happening when POLR2C is overexpressed.

Lastly, POLR2D overexpression causes a drop in Ser5 phosphorylation (Figure 48; Figure 49), while it only slightly affects Ser2 phosphorylation (which is increased) and does not affect Total Pol II levels on transcribed genes. However, the TT-chem-Seq shows a slight drop in transcription in the elongation stage and it even shows increased transcription past the

termination site (Figure 45). We could speculate that the defective Ser5 phosphorylation does not impact in the processive transcription, but it may affect the correct timing of Pol II pausing in the TSS region and therefore interfere with the correct recruitment of co-transcriptional factors, such as 3'end processing factors, capping factors, splicing factors (Ghosh et al., 2011; Maita & Nakagawa, 2020). As a result, Pol II which is correctly phosphorylated on its Ser2 residue could proceed into the gene with regulatory elements that could be lost. Indeed, we found that POLR2D overexpression interferes with SPT5 and XRN2 recruitment on chromatin (Figure 54B), which are recruited at the TSS. We would need to confirm by Pol II IP that these factors present reduced interaction with the polymerase in the dox treated samples. The CTD of Pol II is phosphorylated on Ser5 by CDK7 kinase module of TFIIH , among other kinases (Abdella et al. 2021; Chen et al. 2021; Chen et al. 2018; Fuda et al. 2009; Haberle and Stark 2018; Rengachari et al. 2021; Robinson et al. 2016; Soutourina 2018). Noticeably, CDK7 inhibitor THZ1 (Fisher 2019; Kwiatkowski et al. 2014) is effective on POLR2D overexpressing cells, confirming the important role of Ser5 hypo phosphorylation indicative of a defect in CDK7 activity and regulation. Also, SPT5 can be phosphorylated by CDK7 (Larochelle et al. 2006; Patel and Simon 2010). It should be assessed whether it is specifically the phosphorylated form of SPT that is less present on chromatin upon POLR2D overexpression. However, SPT5 is mainly phosphorylated by CDK9 in order to become a positive elongation factor (Hartzog and Fu 2013), therefore the mechanism leading to its impaired recruitment on chromatin needs to be further investigated.

Ultimately, we wanted to link the transcription defects with the DNA damage phenotypes. For POLR2A and POLR2D cell lines, but not POLR2C, the genes with increased DNA damage levels are also the genes with slightly increased nascent transcription levels (Figure 46). This is the

first step for linking transcription deregulation and genome instability. Remarkably, when looking at the specific Pol II deregulation (identified by ChIP-Seq) we found that each deregulation (increased Pol II, hyper-pSer2 and hypo-pSer5) are even more exacerbated in those genes with higher γH2AX ChIP-Seq levels (Figure 50), strongly suggesting that it is indeed the transcriptional alteration arising upon single subunits overexpressions, more than just increased transcription levels, the cause of the increased genomic instability.

4. Discussion and conclusions

In this work we propose Pol II single subunits POLR2A, POLR2B, POLR2C and POLR2D could constitute a new biomarker for high-risk patients. We found that patients baring upregulated at least one of the four largest subunits of Pol II, showed poorer survival and increased genomic instability (in the form of copy number alteration levels and mutation counts), compared to all other patients. We were also able to show that the phenotypes observed in patients appear to be independent from other oncogenes/tumour suppressor genes deregulation that may be concurrent, finding that the upregulation of Pol II subunit is generally not significantly associated with other drivers being mutated. Additionally, these phenotypes are not tumour specific, as we observed them in six different cancer studies. However, as our analyses assessed the impact of all subuntis together, we can't exclude that for a specific tumour only one among the subunits could be leading in causing the phenotypes observed; unfortunately, the low number of patients obtained when only selecting for one subunit at the time compromises the possibility of performing a robust analysis. It is important to consider that new studies and new patients' cohorts are continuously added to cBio-Portal, therefore it could be worth interrogating again the future the same and other cancer studies, thus building more robust and up to date data.

The analysis carried out on cBio-Portal paved the way for this research, that characterized the role of Pol II subunits upregulations in a model cell line, in order to establish whether some of the phenotypes observed in the patients were directly dependent on Pol Up and to characterize the cause and consequences of the subunits overexpression. The cell line model we used to generate was based on the doxycycline inducible HeLa T-REx cell line, which is a modified version of HeLa cells that allows the conditional overexpression of a desired gene.

We chose this model because it allowed a consistent overexpression level that could be tuned with appropriate doxycycline concentrations, which we found being comparable with the expression levels in the cancer patients. We thought that this model was easy to use, very consistent and reproducible in expression levels in each experimental replicate, and it also allowed us to maintain the same levels of expression across the three subunits, without relying on high transfection efficiency for achieving the overexpression. Finally, as the cancer patients phenotypes were not specific to a single organ or tissue, we sought to use a potentially more generic cell line for our analysis. However, HeLa cells are known for their great genomic instability and high division rates (Landry et al. 2013), so we would like to confirm some of our results in less challenged cell lines and ideally in primary cell lines, to test whether the overexpression of the single subunits is sufficient to give to the cells features of cancer cells.

Once our cell models were established, we were able to show that the single subunit overexpression increased DNA damage levels assessed by the formation and accumulation of a series of common DNA damage markers, and genomic instability, in the form of micronuclei and mitotic aberrations. Despite all subunits overexpression inducing increased DNA damage, there were some differences in the induction of specific markers, such as micronuclei formation or DNA bridges, indicating that overexpressed subunits may be impacting on genomic instability differently. This is also evident when looking at the γH2AX ChIP-Seq metagene profiles and gene-to-gene quantifications, which differ from each subunit. The γH2AX ChIP-Seq allowed us to identify the genes with higher γH2AX levels, and to link the increased DNA damage to specific transcription defects. Indeed, we found that genes with increased damage are also the genes in which the transcription defects arising upon the subunits overexpression are even more pronounced, indicating the strong link between

transcription deregulation and genomic instability. Strikingly, each transcription deregulation found in POLR2A, POLR2C and POLR2D, whether more transcription at the TSS or more towards the termination of the gene and past respectively, agrees with where γH2AX accumulated in the ChIP-Seq metagene profiles. It is important to mention that we quantified the γH2AX levels as well as ChIP-Seq and TT-chem-Seq exclusively of Pol II on transcribed genes, therefore excluding from our analysis other genomic regions, such as enhancers, which could also be affected by Pol II subunits overexpressions, since Pol II pausing is involved in stabilizing the interactions between enhancers and promoters (Barshad et al. 2023), or Pol I and Pol III transcribed genes, given the impact of Pol II transcription also on Pol I and Pol III (Arimbasseri et al. 2013; Cramer et al. 2008; Gerber et al. 2020; Michels and Hernandez 2006; Oler et al. 2010; Sanij et al. 2015). Furthermore, the DNA damage phenotypes could be explored more, characterizing for example the replication stress phenotype that could be arising upon subunits overexpression. First, we could look at how replication is affected by DNA fiber assay and by BrdU-Seq, which uses bromodehoxyuridine (BrdU) to label DNA synthesis in cells in vivo in order to look at replication dynamics (Hulke et al. 2020; Peace et al. 2016). Secondly, we could characterize if the increased DNA damage observed is consequence of more damage formation, perhaps caused by defective transcription affecting DNA replication and resulting in replication stress, or whether it is a consequence of impaired DNA damage repair abilities of spontaneous endogenous lesions. The lack in RAD51 foci formation, and its significantly reduced mRNA levels upon POLR2C overexpression, suggest that the HR pathway could be to some extent compromised, which could be confirmed with an HR-reporter assay. Moreover, it would be necessary to check the kinetics of RAD51 foci formation at DNA damage sites after inducing DNA damage with ionizing radiaton (IR).

Additionally, the defective recruitment to chromatin of BRD4 upon POLR2C overexpression could increase replication fork reversal, as previously shown with inhibition of BRD4. In this scenario, RAD51 activity would be crucial to prevent/ reduce fork reversal and avoid DNA damage formation (Bowry et al. 2018).

Although we observed increased DNA damage, when we tested the effect of DDR protein inhibitors in the cell proliferation assays in the different cell lines following subunits overexpression, they did not show differences in survival of cells , but our screening could be explored further, for example it could be worth to test inhibitors targeting the HR pathway in POLR2C cell line, given its defective RAD51 foci formation phenotype and decreased RAD51 expression (Berte et al. 2016; Chernikova et al. 2012; Lord and Ashworth 2008; Tseng et al. 2021). The description of specific transcription defects by TT-chem-Seq and ChIP-Seq pointed us towards the understanding of the mechanisms and the consequent identification of inhibitors to which cells overexpressing the subunits were more sensitive. We hypothesized that the overexpressed subunits could be interfering with transcribing Pol II, by interfering with the correct CTD phosphorylation and the correct interaction of Pol II with transcription factors. It has been shown recently that Pol II subunits can be found not in the complex (Li et al. 2023) and that each Pol II subunit can have different roles aside of the structural/enzymatic role in the transcribing complex (Li et al. 2022). The ChIP-Seq data is limited to the CTD modifications we investigated. It is not excluded that subunits overexpression may interfere with other CTD modifications, as important as pSer2 and pSer5. Further Pol II ChIP-Seq experiments with all the CTD phosphorylations are crucial to completely characterize the impact of Pol II subunits overexpression on transcription regulation. Furthermore, in the transcription factor screening by cell fractionation and immune-precipitation, we only

investigated a limited number of transcription factors, which we know interact with Pol II to regulate its phosphorylated forms or are efficiently recruited during transcription. However, we did not consider in our analysis other factors that may be directly involved in the CTD phosphorylation phenotype observed, such as phosphatases, which together with kinases are also crucial regulators of transcription (reviewd in Cossa et al. 2021).

However, our results with rescue experiments of POLR2C cell line phenotypes with BRD4 overexpression and the effectiveness of JQ1 and THZ1 in POLR2C and POLR2D cell lines respectively, point towards the fact that the altered CTD phosphorylation is dependent on BRD4 in POLR2C and CDK7 in POLR2D. It is important to consider that JQ1 is an inhibitor targeting all Bromodomain containing proteins, so not only BRD4 but also BRD3 and BRD2, which have not been tested in our screening (Coudé et al. 2015; Singh and Alauddin 2023). Therefore it will be essential performing in future experiments with BRD4-specific inhibitor to assess that it is indeed the partial loss of BRD4 responsible of the phenotypes observed and sensitizing cells to JQ1 and not a secondary effect of targeting BRD3 and BRD2.

In conclusion, we propose that overexpressed/upregulated subunits (PolUp) could constitute a new biomarker for high-risk patients across multiple cancer types, and with the understanding of the transcription deregulations and increased genomic instability arising upon single subunits overexpression we are one step closer to identifying a potentially more suitable targeted therapy. Additionally, we firstly describe new roles of the single subunits in the regulation of the CTD phosphorylation and efficient transcription beyond the boundaries of the transcribing Pol II complex, which are unique to each subunit. In our view, the overexpressed subunits actively interfere with transcription, by deregulating Pol II CTD

phosphorylation and interaction with transcription factors on chromatin (Figure 59), even if further studies are needed in order to define all the CTD modifications and identify the factors defectively recruited to transcribing Pol II, potentially expanding the pool of suitable targets for therapy. Indeed, our working model proposes that when subunits are overexpressed they interfere with TFs recruitment to transcribing Pol II, serving as a sponge scavenging away TFs. At the same time, when interfering with Pol II CTD phosphorylation events other transcription associated processes or the recruitment of TFs dependent on the CTD phosphorylation pattern may be altered. Altogether these deregulations would result in transcription defects and increased DNA damage (Figure 59). Ultimately, to give this research and our findings a more translational breath, it would be worth it to expand our work on primary cells and organoids, recapitulating some of the phenotypes observed in these models. It would be interesting, for example, to assess the drug sensitivity in organoids, which are complex systems and better represent the genomic complexity in which the deregulations of Pol II subunits may occur, giving us a clearer idea on the possibility of exploiting the Pol II subunit upregulation as a new biomarker and on how to precisely target each upregulation for therapeutic purposes.

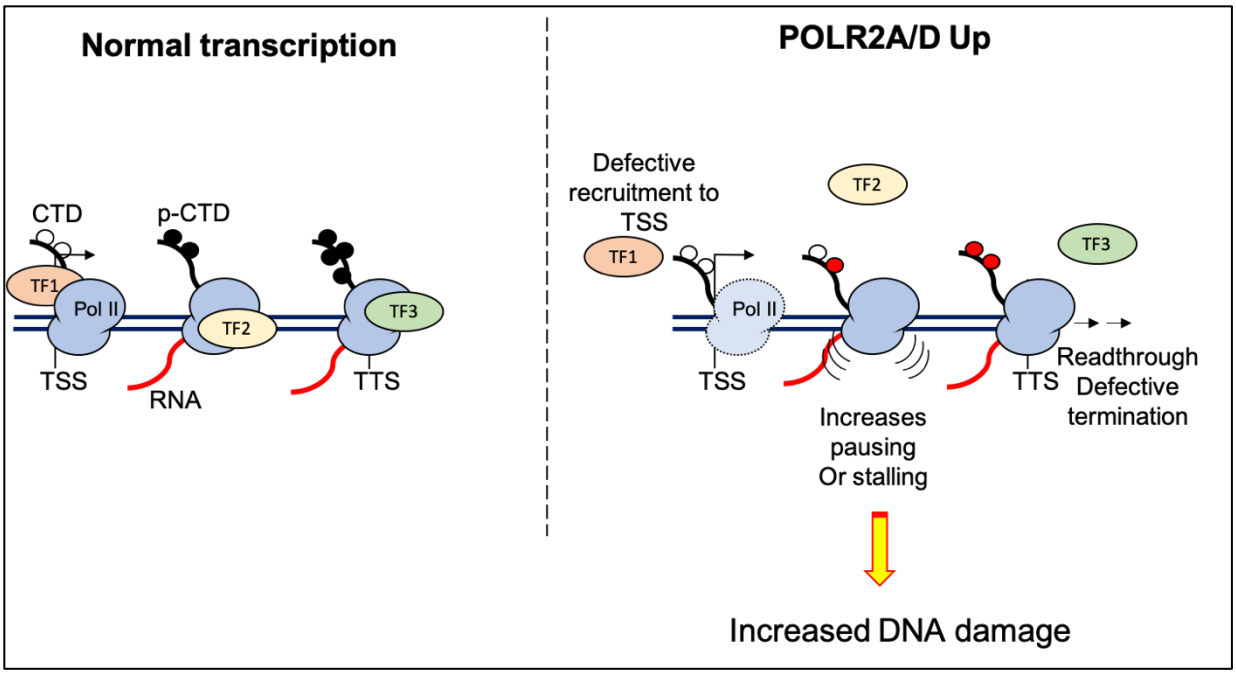


Figure 59 Working model During unperturbed transcription, Pol II CTD is decorated with phosphorylation at the different stages of transcription, from the TSS to the TTS of the gene and transcription factors are correctly co-transcriptionally recruited to transcribing Pol II. When single subunits are overexpressed, they can interfere with transcription factors (TF1,2,3) recruitment and/or with correct CTD phosphorylation, resulting in hyper transcription, elongation defects, termination defects which impair the correct course of transcription and increase the chances of challenging DNA integrity, resulting in increased genomic instability.

Supplementary

Gene	Deregulation	Associated Cancer	References
POLR2A	Upregulated/Overexpressed	Gastric cancer Ovarian cancer Acute Myeloid Leukemia Cholangiocarcinoma Thymoma	Jiang et al., 2021 Li et al., 2019 Yu et al., 2019
	Downregulated	Adrenocortical carcinoma Testicular germ cell carcinoma Endometrial carcinoma	Yu et al., 2019
	Polymorphisms	Non-small cell lung cancer	Yoo et al., 2017
	Co-deletion with TP53	Colorectal cancer Triple-negative breast cancer	Liu et al., 2018 Liu et al., 2015 Cance & Bradner, 2015 Van Der Jeught et al., 2018 Yujing Li et al., 2021 Xu et al., 2019
POLR2B	Mutated	Breast cancer	Luo et al., 2009, Huang & Vakoc, 2016, Gerlach et al., 2018, Iorio et al., 2016
	Upregulated/Overexpressed	Hepatocellular carcinoma	Zhang, 2022
	Downregulated	Acute myeloid leukemia	Imanishi et al., 2014
POLR2C	Upregulated/Overexpressed	Hepatocellular carcinoma Gastric cancer	Fang et al., 2014 Hu et al., 2021 Zhou et al., 2018

	Downregulated	Osteosarcoma	Chen et al., 2022
POLR2D	Upregulated/Overexpressed	Colorectal cancer Ovarian cancer Prostate cancer	Abba et al., 2012 Kim et al., 2021 Yamada et al., 2018
POLR2E	Polymorphisms	Prostate cancer Esophageal cancer Breast cancer Papillary thyroid carcinoma Liver cancer	Jin et al., 2011 Sattarifard et al., 2019 Chu et al., 2017 Chen et al., 2019; Wang, et al., 2018 Li, et al., 2018 Chen et al., 2018
POLR2F	Upregulated/Overexpressed	Colorectal cancer Gastric cancer Triple negative breast cancer Prostate cancer Glioblastoma	Antonacopoulou et al., 2008 Orian-Rousseau et al., 2005 Zhou et al., 2018 Naorem et al., 2019 Yi et al., 2023 Wang et al., 2023 Masica & Karchin, 2011
	Downregulated	Cervical cancer (HPV18+)	Sample, 2020 Wang et al., 2022
POLR2G	Upregulated/Overexpressed	Hepatocellular carcinoma	Yao et al., 2022 Li et al., 2019
POLR2H	Upregulated/Overexpressed	Head and neck carcinomas Colorectal cancer Hepatocellular carcinoma Prostate cancer Lung cancer	Slebos et al., 2006 Zhang et al., 2020 Normal et al., 2020 Ke et al., 2019 Fan et al., 2018

		Breast cancer	Chen & Hu, 2019 Wang et al., 2021 Dong et al., 2021 Jia et al., 2022 Guttapadu et al., 2023 Jia et al., 2015 Miao et al., 2021
POLR2I	Upregulated/Overexpressed	Ovarian cancer Head and neck cancer	Gaponova et al., 2017 Long et al., 2016
	Downregulated	Hepatocellular carcinoma	Walmacq et al., 2009 Ren et al., 2020
POLR2L	Upregulated/Overexpressed	Gastric cancer Hepatocellular carcinoma	Zhou et al., 2018 Zhang et al., 2023
POLR2J	Upregulated/Overexpressed	Lung cancer Colorectal cancer Glioblastoma Ovarian cancer Testicular germ cell carcinoma Breast cancer	Campbell et al., 2008 Rajput et al., 2013 Zhang et al., 2020 Normal et al., 2020 Li, 2023 Sánchez-Tilló et al., 2012 Zhao et al., 2021 Yao et al., 2021 Farahmand et al., 2016
POLR2K	Upregulated/Overexpressed	Hepatocellular carcinoma Breast cancer Prostate cancer Bladder cancer	Woo et al., 2009 Natrajan et al., 2010 Yang & Wang, 2021 Chen et al., 2023 Kelly et al., 2016

			Guo et al., 2021 Yang et al., 2020

Supplementary table 1 The main deregulations affecting each Pol II subunit in cancer

References

Abba, Mohammed, Stephanie Laufs, Monireh Aghajany, Bernhard Korn, Axel Benner, and Heike Allgayer. 2012. “Look Who’s Talking: Deregulated Signaling in Colorectal Cancer.” *Cancer Genomics and Proteomics* 9(1):15–25.

Abdella, R., A. Talyzina, S. Chen, C. J. Inouye, R. Tjian, and Y. He. 2021. “Structure of the Human Mediator-Bound Transcription Preinitiation Complex.” *Science* 372(6537):52–56.

Abraham, Karan J., Negin Khosraviani, Janet N. Y. Chan, Aparna Gorthi, Anas Samman, Dorothy Y. Zhao, Miling Wang, Michael Bokros, Elva Vidya, Lauren A. Ostrowski, Roxanne Oshidari, Violena Pietrobon, Parasvi S. Patel, Arash Algouneh, Rajat Singhania, Yupeng Liu, V. Talya Yerlici, Daniel D. De Carvalho, Michael Ohh, Brendan C. Dickson, Razq Hakem, Jack F. Greenblatt, Stephen Lee, Alexander J. R. Bishop, and Karim Mekhail. 2020. “Nucleolar RNA Polymerase II Drives Ribosome Biogenesis.” *Nature* 585(7824):298–302.

Afgan, Enis, Dannon Baker, Bérénice Batut, Marius Van Den Beek, Dave Bouvier, Martin Ech, John Chilton, Dave Clements, Nate Coraor, Björn A. Grüning, Aysam Guerler, Jennifer Hillman-Jackson, Saskia Hiltemann, Vahid Jalili, Helena Rasche, Nicola Soranzo, Jeremy Goecks, James Taylor, Anton Nekrutenko, and Daniel Blankenberg. 2018. “The Galaxy Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2018 Update.” *Nucleic Acids Research* 46(W1):W537–44.

Ahnesorg, Peter, Philippa Smith, and Stephen P. Jackson. 2006. “XLF Interacts with the XRCC4-DNA Ligase IV Complex to Promote DNA Nonhomologous End-Joining.” *Cell* 124(2):301–13.

Allen, Benjamin L., and Dylan J. Taatjes. 2015. “The Mediator Complex: A Central Integrator of Transcription.” *Nature Reviews Molecular Cell Biology* 16(3):155–66.

Allepuz-Fuster, Paula, Verónica Martínez-Fernández, Ana I. Garrido-Godino, Sergio Alonso-Aguado, Steven D. Hanes, Francisco Navarro, and Olga Calvo. 2014. “Rpb4/7 Facilitates RNA Polymerase II CTD Dephosphorylation.” *Nucleic Acids Research* 42(22):13674–88.

Allepuz-Fuster, Paula, Michael J. O’Brien, Noelia González-Polo, Bianca Pereira, Zuzer Dhoondia, Athar Ansari, and Olga Calvo. 2019. “RNA Polymerase II Plays an Active Role in the Formation of Gene Loops through the Rpb4 Subunit.” *Nucleic Acids Research* 47(17):8975–87.

Allison, David F., and Gang Greg Wang. 2019. “R-Loops: Formation, Function, and Relevance to Cell Stress.” *Cell Stress* 3(2):38–46.

Altendorfer, Elisabeth, Yelizaveta Mochalova, and Andreas Mayer. 2022. “BRD4: A General Regulator of Transcription Elongation.” *Transcription* 13(1–3):70–81.

Amon, Jeremy D., and Douglas Koshland. 2016. “RNase H Enables Efficient Repair of R-Loop Induced DNA Damage.” *ELife* 5:1–20.

An, Yuanyuan, and Hua Duan. 2022. “The Comprehensive Analysis of Interferon-Related Prognostic Signature with Regard to Immune Features in Ovarian Cancer.” *Disease*

Markers 2022.

Antonacopoulou, Anna G., Petros D. Grivas, Lambros Skarlas, Melpomeni Kalofonos, Chrisoula D. Scopa, and Haralabos P. Kalofonos. 2008. “POLR2F, ATP6V0A1 and PRNP Expression in Colorectal Cancer: New Molecules with Prognostic Significance?” *Anticancer Research* 28(2 B):1221–27.

Aparicio, Tomas, Richard Baer, and Jean Gautier. 2014. “DNA Double-Strand Break Repair Pathway Choice and Cancer.” *DNA Repair* 19:169–75.

Arimbasseri, Aneeshkumar G., Keshab Rijal, and Richard J. Maraia. 2013. “Comparative Overview of RNA Polymerase II and III Transcription Cycles, with Focus on RNA Polymerase III Termination and Reinitiation.” *Transcription* 5(DEC):1–13.

Arnold, Mirjam, Annkatrin Bressin, Olga Jasnovidova, David Meierhofer, and Andreas Mayer. 2021. “A BRD4-Mediated Elongation Control Point Primes Transcribing RNA Polymerase II for 3′-Processing and Termination.” *Molecular Cell* 81(17):3589-3603.e13.

Asiedu, Michael K., Charles F. Thomas, Jie Dong, Sandra C. Schulte, Prasidda Khadka, Zhifu Sun, Farhad Kosari, Jin Jen, Julian Molina, George Vasmatzis, Ray Kuang, Marie Christine Aubry, Ping Yang, and Dennis A. Wigle. 2018. “Pathways Impacted by Genomic Alterations in Pulmonary Carcinoid Tumors.” *Clinical Cancer Research* 24(7):1691–1704.

Baillat, David, Mohamed Ali Hakimi, Anders M. Näär, Ali Shilatifard, Neil Cooch, and Ramin Shiekhattar. 2005. “Integrator, a Multiprotein Mediator of Small Nuclear RNA Processing, Associates with the C-Terminal Repeat of RNA Polymerase II.” *Cell* 123(2):265–76.

Barlow, Jacqueline H., Robert B. Faryabi, Elsa Callén, Nancy Wong, Amy Malhowski, Hua Tang Chen, Gustavo Gutierrez-Cruz, Hong Wei Sun, Peter McKinnon, George Wright, Rafael Casellas, Davide F. Robbiani, Louis Staudt, Oscar Fernandez-Capetillo, and André Nussenzweig. 2013. “Identification of Early Replicating Fragile Sites That Contribute to Genome Instability.” *Cell* 152(3):620–32.

Barshad, Gilad, James J. Lewis, Alexandra G. Chivu, Abderhman Abuhashem, Nils Krietenstein, Edward J. Rice, Yitian Ma, Zhong Wang, Oliver J. Rando, Anna Katerina Hadjantonakis, and Charles G. Danko. 2023. “RNA Polymerase II Dynamics Shape Enhancer–Promoter Interactions.” *Nature Genetics* 55(8):1370–80.

Bártová, Eva, Sońa Legartová, Miroslav Dundr, and Jana Suchánková. 2019. “A Role of the 53BP1 Protein in Genome Protection: Structural and Functional Characteristics of 53BP1-Dependent DNA Repair.” *Aging* 11(8):2488–2511.

Baumann, Peter, and Stephen C. West. 1998. “Role of the Human RAD51 Protein in Homologous Recombination and Double- Stranded-Break Repair.” *Trends in Biochemical Sciences* 23(7):247–51.

Bernecky, Carrie, Franz Herzog, Wolfgang Baumeister, Jürgen M. Plitzko, and Patrick Cramer. 2016. “Structure of Transcribing Mammalian RNA Polymerase II.” *Nature* 529(7587):551–54.

Bernecky, Carrie, Jürgen M. Plitzko, and Patrick Cramer. 2017. “Structure of a Transcribing RNA

Polymerase II-DSIF Complex Reveals a Multidentate DNA-RNA Clamp." *Nature Structural and Molecular Biology* 24(10):809–15.

Berte, Nancy, Andrea Piée-Staffa, Nadine Piecha, Mengwan Wang, Kerstin Borgmann, Bernd Kaina, and Teodora Nikolova. 2016. "Targeting Homologous Recombination by Pharmacological Inhibitors Enhances the Killing Response of Glioblastoma Cells Treated with Alkylating Drugs." *Molecular Cancer Therapeutics* 15(11):2665–78.

Bertoli, Cosetta, Jan M. Skotheim, and Robertus A. M. De Bruin. 2013. "Control of Cell Cycle Transcription during G1 and S Phases." *Nature Reviews Molecular Cell Biology* 14(8):518–28.

Bhatt, Dev M., Amy Pandya-Jones, Ann Jay Tong, Iros Barozzi, Michelle M. Lissner, Gioacchino Natoli, Douglas L. Black, and Stephen T. Smale. 2012. "Transcript Dynamics of Proinflammatory Genes Revealed by Sequence Analysis of Subcellular RNA Fractions." *Cell* 150(2):279–90.

Blackford, Andrew N., and Stephen P. Jackson. 2017. "ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response." *Molecular Cell* 66(6):801–17.

Boehning, Marc, Claire Dugast-Darzacq, Marija Rankovic, Anders S. Hansen, Taekyung Yu, Herve Marie-Nelly, David T. McSwiggen, Goran Kokic, Gina M. Dailey, Patrick Cramer, Xavier Darzacq, and Markus Zweckstetter. 2018. "RNA Polymerase II Clustering through Carboxy-Terminal Domain Phase Separation." *Nature Structural and Molecular Biology* 25(9):833–40.

Boettiger, Alistair N., Peter L. Ralph, and Steven N. Evans. 2011. "Transcriptional Regulation: Effects of Promoter Proximal Pausing on Speed, Synchrony and Reliability." *PLoS Computational Biology* 7(5).

Borsos, Barbara N., Ildikó Huliák, Hajnalka Majoros, Zsuzsanna Ujfaludi, Ákos Gyenis, Peter Pukler, Imre M. Boros, and Tibor Pankotai. 2017. "Human P53 Interacts with the Elongating RNAPII Complex and Is Required for the Release of Actinomycin D Induced Transcription Blockage." *Scientific Reports* 7(December 2016):1–11.

Boulianne, B., and N. Feldhahn. 2018. "Transcribing Malignancy: Transcription-Associated Genomic Instability in Cancer." *Oncogene* 37(8):971–81.

Boulon, Séverine, Bérengère Pradet-Balade, Céline Verheggen, Dorothée Molle, Stéphanie Boireau, Marya Georgieva, Karim Azzag, Marie Cécile Robert, Yasmeen Ahmad, Henry Neel, Angus I. Lamond, and Edouard Bertrand. 2010. "HSP90 and Its R2TP/Prefoldin-like Cochaperone Are Involved in the Cytoplasmic Assembly of RNA Polymerase II." *Molecular Cell* 39(6):912–24.

Bouwman, Peter, Amal Aly, Jose M. Escandell, Mark Pieterse, Jirina Bartkova, Hanneke Van Der Gulden, Sanne Hiddingh, Maria Thanasoula, Atul Kulkarni, Qifeng Yang, Bruce G. Haffty, Johanna Tommiska, Carl Blomqvist, Ronny Drapkin, David J. Adams, Heli Nevanlinna, Jiri Bartek, Madalena Tarsounas, Shridar Ganesan, and Jos Jonkers. 2010. "53BP1 Loss Rescues BRCA1 Deficiency and Is Associated with Triple-Negative and BRCA-Mutated Breast Cancers." *Nature Structural and Molecular Biology* 17(6):688–95.

Bowry, Akhil, Ann Liza Piberger, Patricia Rojas, Marco Saponaro, and Eva Petermann. 2018. "BET Inhibition Induces HEXIM1- and RAD51-Dependent Conflicts between Transcription and Replication." *Cell Reports* 25(8):2061-2069.e4.

Bradner, James E., Denes Hnisz, and Richard A. Young. 2017. "Transcriptional Addiction in Cancer." *Cell* 168(4):629–43.

Branzei, Dana, and Marco Foiani. 2009. "The Checkpoint Response to Replication Stress." *DNA Repair* 8(9):1038–46.

Bregman, D., Pestell, R., Kidd, V. 2000. "CELL CYCLE REGULATION AND RNA POLYMERASE II." *Frontiers in Bioscience* 5(3):244–57.

Brickner, Joshua R., Jada L. Garzon, and Karlene A. Cimprich. 2022. "Walking a Tightrope: The Complex Balancing Act of R-Loops in Genome Stability." *Molecular Cell* 82(12):2267–97.

Brown, Jay C. 2021. "Role of Gene Length in Control of Human Gene Expression: Chromosome-Specific and Tissue-Specific Effects." *International Journal of Genomics* 2021(1).

Brown, Jessica S., and Stephen P. Jackson. 2015. "Ubiquitylation, Neddylation and the DNA Damage Response." *Open Biology* 5(4).

Buratowski, Stephen. 2009. "Progression through the RNA Polymerase II CTD Cycle." *Molecular Cell* 36(4):541–46.

Burke, Thomas W., and James T. Kadonaga. 1997. "The Downstream Core Promoter Element, DPE, Is Conserved from Drosophila to Humans and Is Recognized by TAF(II)60 of Drosophila." *Genes and Development* 11(22):3020–31.

Burkholder, Nathaniel T., Sarah N. Sipe, Edwin E. Escobar, Mukeshkumar Venkatramani, Seema Irani, Wanjie Yang, Haoyi Wu, Wendy M. Matthews, Jennifer S. Brodbelt, and Yan Zhang. 2019. "Mapping RNAPII CTD Phosphorylation Reveals That the Identity and Modification of Seventh Heptad Residues Direct Tyr1 Phosphorylation." *ACS Chemical Biology* 14(10):2264–75.

Bywater, Megan J., Richard B. Pearson, Grant A. McArthur, and Ross D. Hannan. 2013. "Dysregulation of the Basal RNA Polymerase Transcription Apparatus in Cancer." *Nature Reviews Cancer* 13(5):299–314.

Bywater, Megan J., Gretchen Poortinga, Elaine Sanij, Nadine Hein, Abigail Peck, Carleen Cullinane, Meaghan Wall, Leonie Cluse, Denis Drygin, Kenna Anderes, Nanni Huser, Chris Proffitt, Joshua Bliesath, Mustapha Haddach, Michael K. Schwaebe, David M. Ryckman, William G. Rice, Clemens Schmitt, Scott W. Lowe, Ricky W. Johnstone, Richard B. Pearson, Grant A. McArthur, and Ross D. Hannan. 2012. "Inhibition of RNA Polymerase I as a Therapeutic Strategy to Promote Cancer-Specific Activation of P53." *Cancer Cell* 22(1):51–65.

Cabarcas, Stephanie, and Laura Schramm. 2011. "RNA Polymerase III Transcription in Cancer : The BRF2 Connection." 1–10.

Caldecott, Keith W. 2024. "Causes and Consequences of DNA Single-Strand Breaks." *Trends in*

Biochemical Sciences 49(1):68–78.

Campbell, Jennifer M., William W. Lockwood, Timon P. H. Buys, Raj Chari, Bradley P. Coe, Stephen Lam, and Wan L. Lam. 2008. "Integrative Genomic and Gene Expression Analysis of Chromosome 7 Identified Novel Oncogene Loci in Non-Small Cell Lung Cancer." *Genome* 51(12):1032–39.

Cance, R., and James E. Bradner. 2015. "An Essential Passenger with P53 Stillbirth Prevented by Signal Blockade." *Nature* 520:626–27.

Cannan, Wendy J., and David S. Pederson. 2016. "Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin." *Journal of Cellular Physiology* 231(1):3–14.

Carre, C., and R. Shiekhattar. 2011. "Human GTPases Associate with RNA Polymerase II To Mediate Its Nuclear Import." *Molecular and Cellular Biology* 31(19):3953–62.

Cerami et al. 2012. "The CBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data." *Cancer Discov.* 32(7):736–40.

Chan, Ying Wai, Kasper Fugger, and Stephen C. West. 2018. "Unresolved Recombination Intermediates Lead to Ultra-Fine Anaphase Bridges, Chromosome Breaks and Aberrations." *Nature Cell Biology* 20(1):92–103.

Chan, Ying Wai, and Stephen C. West. 2018. "A New Class of Ultrafine Anaphase Bridges Generated by Homologous Recombination." *Cell Cycle* 17(17):2101–9.

Chapman, J. Ross, Alex J. Sossick, Simon J. Boulton, and Stephen P. Jackson. 2012. "BRCA1-Associated Exclusion of 53BP1 from DNA: Damage Sites Underlies Temporal Control of DNA Repair." *Journal of Cell Science* 125(15):3529–34.

Chathoth, Keerthi T., J. David Barrass, Shaun Webb, and Jean D. Beggs. 2014. "A Splicing-Dependent Transcriptional Checkpoint Associated with Prespliceosome Formation." *Molecular Cell* 53(5):779–90.

Chen, Bifeng, Yuwei Jiao, Fukang Yaolong, Tiansong Li, Yuxiao Liu, Meiqin Wang, Gu Xiuli, and Xianhong Feng. 2019. "The POLR2E Rs3787016 Polymorphism Is Strongly Associated with the Risk of Female Breast and Cervical Cancer." *Pathology Research and Practice* 215(5):1061–65.

Chen, Bifeng, Jinfeng Li, Can Yi, Yuwei Jiao, Xiuli Gu, and Xianhong Feng. 2018. "Long Non-Coding RNA POLR2E Rs3787016 Is Associated with the Risk of Papillary Thyroid Carcinoma in Chinese Population." *Pathology Research and Practice* 214(7):1040–44.

Chen, Bifeng, Shang Wang, Guangxin Ma, Jin Han, Jingli Zhang, Xiuli Gu, and Xianhong Feng. 2018. "The Association of POLR2E Rs3787016 Polymorphism and Cancer Risk: A Chinese Case–Control Study and Meta-Analysis." *Bioscience Reports* 38(6):1–10.

Chen, Chen, Junxiao Wang, Chao Dong, David Lim, and Zhihui Feng. 2023. "Development of a Risk Model to Predict Prognosis in Breast Cancer Based on CGAS-STING-Related Genes." (March):1–17.

Chen, Fei Xavier, Edwin R. Smith, and Ali Shilatifard. 2018. "Born to Run: Control of

Transcription Elongation by RNA Polymerase II." *Nature Reviews Molecular Cell Biology* 19(7):464–78.

Chen, Liang, Jia Yu Chen, Xuan Zhang, Ying Gu, Rui Xiao, Changwei Shao, Peng Tang, Hao Qian, Daji Luo, Hairi Li, Yu Zhou, Dong Er Zhang, and Xiang Dong Fu. 2017. "R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-Loops with Transcriptional Pausing at Gene Promoters." *Molecular Cell* 68(4):745-757.e5.

Chen, Ruichuan, Zhiyuan Yang, and Qiang Zhou. 2004. "Phosphorylated Positive Transcription Elongation Factor b (P-TEFb) Is Tagged for Inhibition through Association with 7SK SnRNA." *Journal of Biological Chemistry* 279(6):4153–60.

Chen, Ruichuan, Jasper H. N. Yik, Qiao Jing Lew, and Sheng Hao Chao. 2014. "Brd4 and HEXIM1: Multiple Roles in P-TEFb Regulation and Cancer." *BioMed Research International* 2014.

Chen, Xiujin, Nan Zhang, Yuanyuan Zheng, Zhichao Tong, Tuanmin Yang, Xin Kang, Yan He, and Liang Dong. 2022. "Identification of Key Genes and Pathways in Osteosarcoma by Bioinformatics Analysis." *Computational and Mathematical Methods in Medicine* 2022.

Chen, Xizi, Xiaotong Yin, Jiabei Li, Zihan Wu, Yilun Qi, Xinxin Wang, Weida Liu, and Yanhui Xu. 2021. "Structures of the Human Mediator and Mediator-Bound Preinitiation Complex." *Science* 372(6546).

Chen, Zhiqiang, and Haiyi Hu. 2019. "Identification of Prognosis Biomarkers of Prostatic Cancer in a Cohort of 498 Patients from TCGA." *Current Problems in Cancer* 43(6):100503.

Cheng, Ran, Liqiang Qi, Xiangyi Kong, Zhongzhao Wang, Yi Fang, and Jing Wang. 2020. "Identification of the Significant Genes Regulated by Estrogen Receptor in Estrogen Receptor-Positive Breast Cancer and Their Expression Pattern Changes When Tamoxifen or Fulvestrant Resistance Occurs." *Frontiers in Genetics* 11(September):1–13.

Chernikova, Sophia B., John C. Game, and J. Martin Brown. 2012. "Inhibiting Homologous Recombination for Cancer Therapy." *Cancer Biology and Therapy* 13(2):61–68.

Chiaromonte, Francesca, Webb Miller, and Eric E. Bouhassira. 2003. "Gene Length and Proximity to Neighbors Affect Genome-Wide Expression Levels." *Genome Research* 13(12):2602–8.

Cho, Raymond J., Mingxia Huang, Michael J. Campbell, Helin Dong, Lars Steinmetz, Lisa Sapinoso, Garret Hampton, Stephen J. Elledge, Ronald W. Davis, and David J. Lockhart. 2001. "Transcriptional Regulation and Function during the Human Cell Cycle." *Nature Genetics* 27(1):48–54.

Chu, Haiyan, Yaoyao Chen, Qinbo Yuan, Qiuhan Hua, Xu Zhang, Meilin Wang, Na Tong, Wei Zhang, Jinfei Chen, and Zhengdong Zhang. 2017. "The HOTAIR, PRNCR1 and POLR2E Polymorphisms Are Associated with Cancer Risk: A Meta-Analysis." *Oncotarget* 8(26):43271–83.

Ciccia, Alberto, and Stephen J. Elledge. 2010. "The DNA Damage Response: Making It Safe to Play with Knives." *Molecular Cell* 40(2):179–204.

Cimprich, Karlene A., and David Cortez. 2008. “ATR: An Essential Regulator of Genome Integrity.1. Cimprich KA, Cortez D: ATR: An Essential Regulator of Genome Integrity. Nat Rev Mol Cell Biol 2008, 9:616–27.” *Nature Reviews. Molecular Cell Biology* 9(8):616–27.

Cole, Michael D., and Victoria H. Cowling. 2008. “Transcription-Independent Functions of MYC: Regulation of Translation and DNA Replication.” *Nature Reviews Molecular Cell Biology* 9(10):810–15.

Compe, Emmanuel, Carlos M. Genes, Cathy Braun, Frederic Coin, and Jean Marc Egly. 2019. “TFIIE Orchestrates the Recruitment of the TFIIH Kinase Module at Promoter before Release during Transcription.” *Nature Communications* 10(1).

Core, Leighton, and Karen Adelman. 2019. “Promoter-Proximal Pausing of RNA Polymerase II: A Nexus of Gene Regulation.” *Genes and Development* 33(15–16):960–82.

Cornelison, Robert, Zachary C. Dobbin, Ashwini A. Katre, Dae Hoon Jeong, Yinfeng Zhang, Dongquan Chen, Yuliya Petrova, Danielle C. Llaneza, Adam D. Steg, Laura Parsons, David A. Schneider, and Charles N. Landen. 2017. “Targeting RNA-Polymerase I in Both Chemosensitive and Chemoresistant Populations in Epithelial Ovarian Cancer.” *Clinical Cancer Research* 23(21):6529–40.

Cortez, David. 2019. “Replication-Coupled DNA Repair.” *Molecular Cel* 74(5):866–76.

Cosma, Maria Pia. 2002. “Ordered Recruitment: Gene-Specific Mechanism of Transcription Activation.” *Molecular Cell* 10:227–36.

Cossa, Giacomo, Pabitra K. Parua, Martin Eilers, and Robert P. Fisher. 2021. “Protein Phosphatases in the RNAPII Transcription Cycle: Erasers, Sculptors, Gatekeepers, and Potential Drug Targets.” *Genes and Development* 35(9–10):658–76.

Coudé, Marie Magdelaine, Thorsten Braun, Jeannig Berrou, Mélanie Dupont, Sibyl Bertrand, Aline Masse, Emmanuel Raffoux, Raphaël Itzykson, Marc Delord, Maria E. Riveiro, Patrice Herait, André Baruchel, Hervé Dombret, and Claude Gardin. 2015. “BET Inhibitor OTX015 Targets BRD2 and BRD4 and Decreases C-MYC in Acute Leukemia Cells.” *Oncotarget* 6(19):17698–712.

Cramer, P., K. J. Armache, S. Baumli, S. Benkert, F. Brueckner, C. Buchen, G. E. Damsma, S. Dengl, S. R. Geiger, A. J. Jasiak, A. Jawhari, S. Jennebach, T. Kamenski, H. Kettenberger, C. D. Kuhn, E. Lehmann, K. Leike, J. F. Sydow, and A. Vannini. 2008. “Structure of Eukaryotic RNA Polymerases.” *Annual Review of Biophysics* 37(1):337–52.

Cramer, Patrick. 2004. “RNA Polymerase II Structure: From Core to Functional Complexes.” *Current Opinion in Genetics and Development* 14(2):218–26.

Crossley, Madzia P., Michael Bocek, and Karlene A. Cimprich. 2019. “R-Loops as Cellular Regulators and Genomic Threats.” *Molecular Cell* 73(3):398–411.

Cui, Yajun, and Clyde L. Denis. 2003. “In Vivo Evidence That Defects in the Transcriptional Elongation Factors RPB2, TFIIS, and SPT5 Enhance Upstream Poly(A) Site Utilization.” *Molecular and Cellular Biology* 23(21):7887–7901.

Czeko, Elmar, Martin Seizl, Christian Augsberger, Thorsten Mielke, and Patrick Cramer. 2011. "Iwr1 Directs RNA Polymerase II Nuclear Import." *Molecular Cell* 42(2):261–66.

Czudnochowski, Nadine, Christian A. Bösken, and Matthias Geyer. 2012. "Serine-7 but Not Serine-5 Phosphorylation Primes RNA Polymerase II CTD for P-TEFb Recognition." *Nature Communications* 3(May).

Dai, Danian, Qiang Li, Pengfei Zhou, Jianjiang Huang, Hongkai Zhuang, Hongmei Wu, and Bo Chen. 2022. "Analysis of Omics Data Reveals Nucleotide Excision Repair-Related Genes Signature in Highly-Grade Serous Ovarian Cancer to Predict Prognosis." *Frontiers in Cell and Developmental Biology* 10(June):1–15.

Dai, Xiaoyan, Wei Jiang, Liang Ma, Jie Sun, Xiaodi Yan, Jing Qian, Yan Wang, Yu Shi, Shujie Ni, and Ninghua Yao. 2021. "A Metabolism-Related Gene Signature for Predicting the Prognosis and Therapeutic Responses in Patients with Hepatocellular Carcinoma." *Annals of Translational Medicine* 9(6):500–500.

Dang, Chi V., Linda M. S. Resar, Eileen Emison, Sunkyu Kim, Qing Li, Julia E. Prescott, Diane Wonsey, and Karen Zeller. 1999. "Function of the C-Myc Oncogenic Transcription Factor." *Experimental Cell Research* 253(1):63–77.

Derenzini, Massimo, Lorenzo Montanaro, and Davide Treré. 2009. "What the Nucleolus Says to a Tumour Pathologist." *Histopathology* 54(6):753–62.

Deriano, Ludovic, and David B. Roth. 2013. "Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage." *Annual Review of Genetics* 47(September):433–55.

Devaiah, Ballachanda N., Brian A. Lewis, Natasha Cherman, Michael C. Hewitt, Brian K. Albrecht, Pamela G. Robey, Keiko Ozato, Robert J. Sims, and Dinah S. Singer. 2012. "BRD4 Is an Atypical Kinase That Phosphorylates Serine2 of the RNA Polymerase II Carboxy-Terminal Domain." *Proceedings of the National Academy of Sciences of the United States of America* 109(18):6927–32.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29(1):15–21.

Donati, Benedetta, Eugenia Lorenzini, and Alessia Ciarrocchi. 2018. "BRD4 and Cancer: Going beyond Transcriptional Regulation." *Molecular Cancer* 17(1):1–13.

Dong, Ao, Zi Wen Wang, Na Ni, Lu Li, and Xiang Yang Kong. 2021. "Similarity and Difference of Pathogenesis among Lung Cancer Subtypes Suggested by Expression Profile Data." *Pathology Research and Practice* 220(November 2020).

Doré, Andrew S., Mairi L. Kilkenny, Neil J. Rzechorzek, and Laurence H. Pearl. 2009. "Crystal Structure of the Rad9-Rad1-Hus1 DNA Damage Checkpoint Complex-Implications for Clamp Loading and Regulation." *Molecular Cell* 34(6):735–45.

Drygin, Denis, William G. Rice, and Ingrid Grummt. 2010. "The RNA Polymerase I Transcription Machinery: An Emerging Target for the Treatment of Cancer." *Annual Review of*

Pharmacology and Toxicology 50(1):131–56.

Dujardin, Gwendal, Celina Lafaille, Ezequiel Petrillo, Valeria Buggiano, Luciana I. Gómez Acuña, Ana Fiszbein, Micaela A. Godoy Herz, Nicolás Nieto Moreno, Manuel J. Muñoz, Mariano Alló, Ignacio E. Schor, and Alberto R. Kornblihtt. 2013. "Transcriptional Elongation and Alternative Splicing." *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1829(1):134–40.

Eaton, Joshua D., Lee Davidson, David L. V. Bauer, Toyoaki Natsume, Masato T. Kanemaki, and Steven West. 2018. "Xrn2 Accelerates Termination by RNA Polymerase II, Which Is Underpinned by CPSF73 Activity." *Genes and Development* 32(2):127–39.

Eaton, Joshua D., Laura Francis, Lee Davidson, and Steven West. 2020. "A Unified Allosteric/Torpedo Mechanism for Transcriptional Termination on Human Protein-Coding Genes." *Genes & Development* 34(1–2):132–45.

Ebmeier, Christopher C., Benjamin Erickson, Benjamin L. Allen, Mary A. Allen, Hyunmin Kim, Nova Fong, Jeremy R. Jacobsen, Kaiwei Liang, Ali Shilatifard, Robin D. Dowell, William M. Old, David L. Bentley, and Dylan J. Taatjes. 2017. "Human TFIIH Kinase CDK7 Regulates Transcription-Associated Chromatin Modifications." *Cell Reports* 20(5):1173–86.

Ehara, Haruhiko, Takeshi Yokoyama, Hideki Shigematsu, Shigeyuki Yokoyama, Mikako Shirouzu, and Shun Ichi Sekine. 2017. "Structure of the Complete Elongation Complex of RNA Polymerase II with Basal Factors." *Science* 357(6354):921–24.

Eick, Dirk, and Matthias Geyer. 2013. "The RNA Polymerase II Carboxy-Terminal Domain (CTD) Code." *Chemical Reviews* 113(11):8456–90.

Eli Eisenberg and Erez Y. Levanon. 2003. "Human Housekeeping Genes Are Compact." *Trends in Genetics* 19(7):356–62.

Elrod, Nathan D., Telmo Henriques, Kai-Lieh Huang, Deirdre C. Tatomer, Jeremy E. Wilusz, Eric J. Wagner, and Karen Adelman. 2019. "The Integrator Complex Attenuates Promoter-Proximal Transcription at Protein-Coding Genes." *Molecular Cell* 76(5):738-752.e7.

Falck, Jacob, Julia Coates, and Stephen P. Jackson. 2005. "Conserved Modes of Recruitment of ATM, ATR and DNA-PKcs to Sites of DNA Damage." *Nature* 434(7033):605–11.

Fan, Shutong, Zumu Liang, Zhiqin Gao, Zhiwei Pan, Shaojie Han, Xiaoying Liu, Chunling Zhao, Weiwei Yang, Zhifang Pan, and Weiguo Feng. 2018. "Identification of the Key Genes and Pathways in Prostate Cancer." *Oncology Letters* 16(5):6663–69.

Fang, Zhe Ping, Bei Ge Jiang, Fa Biao Zhang, Ai Dong Wang, Yi Ming Ji, Yong Fu Xu, Ji Cheng Li, Wei Ping Zhou, Wei Jie Zhou, and Hai Xiong Han. 2014. "Rpb3 Promotes Hepatocellular Carcinoma through Its N-Terminus." *Oncotarget* 5(19):9256–68.

Fant, Charli B., Cecilia B. Levandowski, Kapil Gupta, Zachary L. Maas, John Moir, Jonathan D. Rubin, Andrew Sawyer, Meagan N. Esbin, Jenna K. Rimel, Olivia Luyties, Michael T. Marr, Imre Berger, Robin D. Dowell, and Dylan J. Taatjes. 2020. "TFIID Enables RNA Polymerase II Promoter-Proximal Pausing." *Molecular Cell* 78(4):785-793.e8.

Farahmand, S., S. Goliaei, N. Ansari-Pour, and Z. Razaghi-Moghadam. 2016. “GTA: A Game Theoretic Approach to Identifying Cancer Subnetwork Markers.” *Molecular BioSystems* 12(3):818–25.

Ferreira, Rita, John S. Schneekloth, Konstantin I. Panov, Katherine M. Hannan, and Ross D. Hannan. 2020. “Targeting the RNA Polymerase I Transcription for Cancer Therapy Comes of Age.” *Cells* 9(2):266.

Fisher, Robert P. 2019. “Cdk7: A Kinase at the Core of Transcription and in the Crosshairs of Cancer Drug Discovery.” *Transcription* 10(2):47–56.

Fong, Nova, Hyunmin Kim, Yu Zhou, Xiong Ji, Jinsong Qiu, Tassa Saldi, Katrina Diener, Ken Jones, Xiang Dong Fu, and David L. Bentley. 2014. “Pre-MRNA Splicing Is Facilitated by an Optimal RNA Polymerase II Elongation Rate.” *Genes and Development* 28(23):2663–76.

Fong, Nova, Tassa Saldi, Ryan M. Sheridan, Michael A. Cortazar, and David L. Bentley. 2017. “RNA Pol II Dynamics Modulate Co-Transcriptional Chromatin Modification, CTD Phosphorylation, and Transcriptional Direction.” *Molecular Cell* 66(4):546-557.e3.

Fong, Y. W., and Q. Zhou. 2001. “Stimulatory Effect of Splicing Factors on Transcriptional Elongation.” *Nature* 414(6866):929–33.

Forget, Diane, Andrée Anne Lacombe, Philippe Cloutier, Racha Al-Khoury, Annie Bouchard, Mathieu Lavallée-Adam, Denis Faubert, Célia Jeronimo, Mathieu Blanchette, and Benoit Coulombe. 2010. “The Protein Interaction Network of the Human Transcription Machinery Reveals a Role for the Conserved GTPase RPAP4/GPN1 and Microtubule Assembly in Nuclear Import and Biogenesis of RNA Polymerase II.” *Molecular and Cellular Proteomics* 9(12):2827–39.

Forget, Diane, Andrée Anne Lacombe, Philippe Cloutier, Mathieu Lavallée-Adam, Mathieu Blanchette, and Benoit Coulombe. 2013. “Nuclear Import of RNA Polymerase II Is Coupled with Nucleocytoplasmic Shuttling of the RNA Polymerase II-Associated Protein 2.” *Nucleic Acids Research* 41(14):6881–91.

Frit, Philippe, Nadia Barboule, Ying Yuan, Dennis Gomez, and Patrick Calsou. 2014. “Alternative End-Joining Pathway(s): Bricolage at DNA Breaks.” *DNA Repair* 17:81–97.

Fuda, Nicholas J., M. Behfar Ardehali, and John T. Lis. 2009. “Defining Mechanisms That Regulate RNA Polymerase II Transcription in Vivo.” *Nature* 461(7261):186–92.

Fukushima, Hiroto, Tetsuya Abe, Kazuki Sakamoto, Hiroaki Tsujimoto, Shinji Mizuarai, and Shinji Oie. 2014. “3’-Ethynylcytidine, an RNA Polymerase Inhibitor, Combined with Cisplatin Exhibits a Potent Synergistic Growth-Inhibitory Effect via Vaults Dysfunction.” *BMC Cancer* 14(1):1–12.

Gaillard, Hélène, and Andrés Aguilera. 2016. “Transcription as a Threat to Genome Integrity.” *Annual Review of Biochemistry* 85(1):291–317.

Galbraith, Matthew D., Heather Bender, and Joaquín M. Espinosa. 2019. “Therapeutic Targeting of Transcriptional Cyclin-Dependent Kinases.” *Transcription* 10(2):118–36.

Gamba, Pamela, and Nikolay Zenkin. 2018. “Transcription Fidelity and Its Roles in the Cell.” *Current Opinion in Microbiology* 42:13–18.

Gao, Jianjiong. 2014. “Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the CBioPortal.” *Sci Signal* 6(269):1–34.

Gaponova, Anna V., Alexander Y. Deneka, Tim N. Beck, Hanqing Liu, Gregory Andrianov, Anna S. Nikonova, Emmanuelle Nicolas, Margret B. Einarson, Erica A. Golemis, and Ilya G. Serebriiskii. 2017. “Identification of Evolutionarily Conserved DNA Damage Response Genes That Alter Sensitivity to Cisplatin.” *Oncotarget* 8(12):19156–71.

García-Muse, Tatiana, and Andrés Aguilera. 2016. “Transcription-Replication Conflicts: How They Occur and How They Are Resolved.” *Nature Reviews Molecular Cell Biology* 17(9):553–63.

Garg, Gaurika, Christian Dienemann, Lucas Farnung, Juliane Schwarz, Andreas Linden, Henning Urlaub, and Patrick Cramer. 2023. “Structural Insights into Human Co-Transcriptional Capping.” *Molecular Cell* 83(14):2464-2477.e5.

Gargi Ghosal and Junjie Chen. 2013. “DNA Damage Tolerance: A Double-Edged Sword Guarding the Genome.” *Transl Caner Res.* 2(3):107–29.

Garrido-Godino, A. I., M. C. García-López, J. García-Martínez, V. Pelechano, D. A. Medina, J. E. Pérez-Ortín, and F. Navarro. 2016. “Rpb1 Foot Mutations Demonstrate a Major Role of Rpb4 in MRNA Stability during Stress Situations in Yeast.” *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1859(5):731–43.

Gell, David, and Stephen P. Jackson. 1999. “Mapping of Protein-Protein Interactions within the DNA-Dependent Protein Kinase Complex.” *Nucleic Acids Research* 27(17):3494–3502.

Genes Robles, Carlos Mario, and Frédéric Coin. 2019. “Conducting the CTD Orchestra.” *Nature Chemical Biology* 15(2):97–98.

Gerber, Alan, Keiichi Ito, Chi-Shuen Chu, and Robert G. Roeder. 2020. “Gene-Specific Control of TRNA Expression by RNA Polymerase II.” *Molecular Cell* 1–14.

Gerlach, Daniel, Ulrike Tontsch-Grunt, Anke Baum, Johannes Popow, Dirk Scharn, Marco H. Hofmann, Harald Engelhardt, Onur Kaya, Janina Beck, Norbert Schweifer, Thomas Gerstberger, Johannes Zuber, Fabio Savarese, and Norbert Kraut. 2018. “The Novel BET Bromodomain Inhibitor BI 894999 Represses Super-Enhancer-Associated Transcription and Synergizes with CDK9 Inhibition in AML.” *Oncogene* 37(20):2687–2701.

Ghafouri-Fard, Soudeh, Leili Nekoohesh, and Elahe Motevaseli. 2014. “Bladder Cancer Biomarkers: Review and Update.” *Asian Pacific Journal of Cancer Prevention* 15(6):2395–2403.

Ghosh, Agnidipta, Stewart Shuman, and Christopher D. Lima. 2011. “Structural Insights to How Mammalian Capping Enzyme Reads the CTD Code.” *Molecular Cell* 43(2):299–310.

Gilchrist, Daniel A., Gilberto Dos Santos, David C. Fargo, Bin Xie, Yuan Gao, Leping Li, and Karen Adelman. 2010. “Pausing of RNA Polymerase II Disrupts DNA-Specified Nucleosome

Organization to Enable Precise Gene Regulation." *Cell* 143(4):540–51.

Gjidoda, Alison, and R. William Henry. 2013. "RNA Polymerase III Repression by the Retinoblastoma Tumor Suppressor Protein." *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1829(3–4):385–92.

Gómez-González, Belén, and Andrés Aguilera. 2019. "Transcription-Mediated Replication Hindrance: A Major Driver of Genome Instability." *Genes and Development* 33(15–16):1008–26.

Gómez-Navarro, Natalia, and Francisco Estruch. 2015. "Different Pathways for the Nuclear Import of Yeast RNA Polymerase II." *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1849(11):1354–62.

Gonzalez-Perez, Abel, Christian Perez-Llamas, Jordi Deu-Pons, David Tamborero, Michael P. Schroeder, Alba Jene-Sanz, Alberto Santos, and Nuria Lopez-Bigas. 2013. "IntOGen-Mutations Identifies Cancer Drivers across Tumor Types." *Nature Methods* 10(11):1081–82.

Goodarzi, Aaron A., Yaping Yu, Enriqueta Riballo, Pauline Douglas, Sarah A. Walker, Ruiqiong Ye, Christine Härer, Caterina Marchetti, Nick Morrice, Penny A. Jeggo, and Susan P. Lees-Miller. 2006. "DNA-PK Autophosphorylation Facilitates Artemis Endonuclease Activity." *EMBO Journal* 25(16):3880–89.

Graham, Thomas G. W., Johannes C. Walter, and Joseph J. Loparo. 2016. "Two-Stage Synapsis of DNA Ends during Non-Homologous End Joining." *Molecular Cell* 61(6):850–58.

Gregersen, Lea H., Richard Mitter, and Jesper Q. Svejstrup. 2020. "Using TTchem-Seq for Profiling Nascent Transcription and Measuring Transcript Elongation." *Nature Protocols* 15(2):604–27.

Grishkevich, Vladislav, and Itai Yanai. 2014. "Gene Length and Expression Level Shape Genomic Novelties." *Genome Research* 24(9):1497–1503.

Guo, Hanxu, Zhichao Zhang, Yuhang Wang, and Sheng Xue. 2021. "Identification of Crucial Genes and Pathways Associated with Prostate Cancer in Multiple Databases." *Journal of International Medical Research* 49(6).

Guo, Yang Eric, John C. Manteiga, Jonathan E. Henninger, Benjamin R. Sabari, Alessandra Dall Agnese, Krishna Shrinivas, Brian J. Abraham, Nancy M. Hannett, Jan-hendrik Spille, Lena K. Afeyan, V. Alicia, Ann Boija, Tim-michael Decker, Jenna K. Rimel, Charli B. Fant, Tong Ihn Lee, Ibrahim I. Cisse, Phillip A. Sharp, Dylan J. Taatjes, and Richard A. Young. 2019. "Transcriptional and Splicing Condensates." *Nature*.

Gupta, Kapil, Duygu Sari-Ak, Matthias Haffke, Simon Trowitzsch, and Imre Berger. 2016. "Zooming in on Transcription Preinitiation." *Journal of Molecular Biology* 428(12):2581–91.

Guttapadu, Ranjitha, Teesta Katte, Deepak Sayeeram, Saloni Bhatia, Anika Rachel Abraham, Kiran Rajeev, Anish Raju R. Amara, Sharadhi Siri, Kavitha Bommana, Avinash Arvind Rasalkar, Rajyalakshmi Malempati, Mohammed S. Mustak, Prathibha Narayanan, and S.

Divijendra Natha Reddy. 2023. “Identification of Novel Biomarkers for Lung Squamous Cell Carcinoma.” *3 Biotech* 13(2):1–9.

Haberle, Vanja, and Alexander Stark. 2018. “Eukaryotic Core Promoters and the Functional Basis of Transcription Initiation.” *Nature Reviews Molecular Cell Biology* 19(10):621–37.

Hahn, Steven. 2004. “Structure and Mechanism of the RNA Polymerase II Transcription Machinery.” *Nature Structural and Molecular Biology* 11(5):394–403.

Hamperl, Stephan, Michael J. Bocek, Joshua C. Saldivar, Tomek Swigut, and Karlene A. Cimprich. 2017. “Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses.” *Cell* 170(4):774-786.e19.

Hamperl, Stephan, and Karlene A. Cimprich. 2016. “Conflict Resolution in the Genome: How Transcription and Replication Make It Work.” *Cell* 167(6):1455–67.

Hanahan, Douglas, and Robert A. Weinberg. 2000. “The Hallmarks of Cancer.” *Cell* 100:57–70.

Hanahan, Douglas, and Robert A. Weinberg. 2011. “Hallmarks of Cancer: The next Generation.” *Cell* 144(5):646–74.

Hantsche, Merle, and Patrick Cramer. 2017. “Conserved RNA Polymerase II Initiation Complex Structure.” *Current Opinion in Structural Biology* 47:17–22.

Harlen, Kevin M., and L. Stirling Churchman. 2017. “The Code and beyond: Transcription Regulation by the RNA Polymerase II Carboxy-Terminal Domain.” *Nature Reviews Molecular Cell Biology* 18(4):263–73.

Hartzog, Grant A., and Jianhua Fu. 2013. “The Spt4-Spt5 Complex: A Multi-Faceted Regulator of Transcription Elongation.” *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1829(1):105–15.

He, Yuan, Chunli Yan, Jie Fang, Carla Inouye, Robert Tjian, Ivaylo Ivanov, and Eva Nogales. 2016. “Near-Atomic Resolution Visualization of Human Transcription Promoter Opening.” *Nature* 533(7603):359–65.

Heidemann, Martin, Corinna Hintermair, Kirsten Voß, and Dirk Eick. 2013. “Dynamic Phosphorylation Patterns of RNA Polymerase II CTD during Transcription.” *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1829(1):55–62.

Helmlinger, Dominique, and László Tora. 2017. “Sharing the SAGA.” *Trends in Biochemical Sciences* 42(11):850–61.

Helmrich, Anne, Monica Ballarino, Evgeny Nudler, and Laszlo Tora. 2013. “Transcription-Replication Encounters, Consequences and Genomic Instability.” *Nature Structural and Molecular Biology* 20(4):412–18.

Helmrich, Anne, Monica Ballarino, and Laszlo Tora. 2011. “Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes.” *Molecular Cell* 44(6):966–77.

Hernández Borrero, Liz J., and Wafik S. El-Deiry. 2021. “Tumor Suppressor P53: Biology,

Signaling Pathways, and Therapeutic Targeting." *Biochimica et Biophysica Acta - Reviews on Cancer* 1876(1).

Hoeijmakers, Jan H. J. 2009. "DNA Damage, Aging, and Cancer." *NEJM* 15(361):1475–85.

Hu, Peng, Binfeng Wang, Ting Chen, Yongfu Xu, Guoqun Zheng, Yu Zhu, and Xuefeng Du. 2021. "RNA Polymerase II Subunit 3 Regulates Vesicular, Overexpressed in Cancer, Prosurvival Protein 1 Expression to Promote Hepatocellular Carcinoma." *Journal of International Medical Research* 49(4).

Huadong Pei1, Lindsey Zhang*,4, Kuntian Luo*,1, Yuxin Qin2, Marta Chesi3, Frances Fei4, P. Leif Bergsagel3, Liewei Wang2, Zhongsheng You4, and Zhenkun Lou1. 2011. "MMSET Regulates Histone H4K20 Methylation and 53BP1 Accumulation at DNA Damage Sites Huadong." *Nature* 470(7332):124–28.

Huang, Ruixue, and Ping Kun Zhou. 2021. *DNA Damage Repair: Historical Perspectives, Mechanistic Pathways and Clinical Translation for Targeted Cancer Therapy*. Vol. 6. Springer US.

Huertas, Pablo, and Andrés Aguilera. 2003. "Cotranscriptionally Formed DNA:RNA Hybrids Mediate Transcription Elongation Impairment and Transcription-Associated Recombination." *Molecular Cell* 12(3):711–21.

Hulke, Michelle L., Dashiell J. Massey, and Amnon Koren. 2020. "Genomic Methods for Measuring DNA Replication Dynamics." *Chromosome Research* 28(1):49–67.

Iliakis, George, Tamara Murmann, and Aashish Soni. 2015. "Alternative End-Joining Repair Pathways Are the Ultimate Backup for Abrogated Classical Non-Homologous End-Joining and Homologous Recombination Repair: Implications for the Formation of Chromosome Translocations." *Mutation Research - Genetic Toxicology and Environmental Mutagenesis* 793:166–75.

Imanishi, Satoshi, Tomohiro Umezu, Kazushige Ohtsuki, Chiaki Kobayashi, Kazuma Ohyashiki, and Junko H. Ohyashiki. 2014. "Constitutive Activation of the ATM/BRCA1 Pathway Prevents DNA Damage-Induced Apoptosis in 5-Azacytidine-Resistant Cell Lines." *Biochemical Pharmacology* 89(3):361–69.

Iorio, Francesco, Theo A. Knijnenburg, Daniel J. Vis, Graham R. Bignell, Michael P. Menden, Michael Schubert, Nanne Aben, Emanuel Gonçalves, Syd Barthorpe, Howard Lightfoot, Thomas Cokelaer, Patricia Greninger, Ewald van Dyk, Han Chang, Heshani de Silva, Holger Heyn, Xianming Deng, Regina K. Egan, Qingsong Liu, Tatiana Mironenko, Xeni Mitropoulos, Laura Richardson, Jinhua Wang, Tinghu Zhang, Sebastian Moran, Sergi Sayols, Maryam Soleimani, David Tamborero, Nuria Lopez-Bigas, Petra Ross-Macdonald, Manel Esteller, Nathanael S. Gray, Daniel A. Haber, Michael R. Stratton, Cyril H. Benes, Lodewyk F. A. Wessels, Julio Saez-Rodriguez, Ultan McDermott, and Mathew J. Garnett. 2016. "A Landscape of Pharmacogenomic Interactions in Cancer." *Cell* 166(3):740–54.

Jackson, Stephen P., and Jiri Bartek. 2009. "The DNA-Damage Response in Human Biology and Disease." *Nature* 461(7267):1071–78.

Jackson, Stephen P., and Daniel Durocher. 2013. “Regulation of DNA Damage Responses by Ubiquitin and SUMO.” *Molecular Cell* 49(5):795–807.

Van Der Jeught, Kevin, Han Chen Xu, Yu Jing Li, Xiong Bin Lu, and Guang Ji. 2018. “Drug Resistance and New Therapies in Colorectal Cancer.” *World Journal of Gastroenterology* 24(34):3834–48.

Jia, Bin, Ting Gong, Bingsheng Sun, Zhenfa Zhang, Diansheng Zhong, and Changli Wang. 2022. “Identification of a DNA Damage Repair Gene-Related Signature for Lung Squamous Cell Carcinoma Prognosis.” *Thoracic Cancer* 13(8):1143–52.

Jia, Zhuomin, Xing Ai, Fengling Sun, Tong Zang, Yawei Guan, and Feng Gao. 2015. “Identification of New Hub Genes Associated with Bladder Carcinoma via Bioinformatics Analysis.” *Tumori* 101(1):117–22.

Jiang, Guojuan, Wanglong Deng, Yang Liu, and Chengde Wang. 2020. “General Mechanism of JQ1 in Inhibiting Various Types of Cancer.” *Molecular Medicine Reports* 21(3):1021–34.

Jiang, Qiuyu, Jinyuan Zhang, Fang Li, Xiaoping Ma, Fei Wu, Jiyu Miao, Qian Li, Xiaofei Wang, Ruifang Sun, Yang Yang, Lingyu Zhao, and Chen Huang. 2021. “POLR2A Promotes the Proliferation of Gastric Cancer Cells by Advancing the Overall Cell Cycle Progression.” *Frontiers in Genetics* 12(November):1–12.

Jin, Guangfu, Jielin Sun, Sarah D. Isaacs, Kathleen E. Wiley, Seong Tae Kim, Lisa W. Chu, Zheng Zhang, Hui Zhao, Siqun Lilly Zheng, William B. Isaacs, and Jianfeng Xu. 2011. “Human Polymorphisms at Long Non-Coding RNAs (LncRNAs) and Association with Prostate Cancer Risk.” *Carcinogenesis* 32(11):1655–59.

Johnson, David G. 2000. “The Paradox of E2F1: Oncogene and Tumor Suppressor Gene.” *Molecular Carcinogenesis* 27(3):151–57.

Jonkers, Iris, and John T. Lis. 2015. “Getting up to Speed with Transcription Elongation by RNA Polymerase II.” *Nature Reviews Molecular Cell Biology* 16(3):167–77.

Kallunki, Tuula, Marin Barisic, Marja Jäättelä, and Bin Liu. 2019. “How to Choose the Right Inducible Gene Expression System for Mammalian Studies?” *Cells* 8(8):1–16.

Ke, Rui sheng, Kun Zhang, Li zhi Lv, Ya ping Dong, Fan Pan, Fang Yang, Qiu cheng Cai, and Yi Jiang. 2019. “Prognostic Value and Oncogene Function of Heterogeneous Nuclear Ribonucleoprotein A1 Overexpression in HBV-Related Hepatocellular Carcinoma.” *International Journal of Biological Macromolecules* 129:140–51.

Kelly, Rachel S., Jennifer A. Sinnott, Jennifer R. Rider, Ericka M. Ebot, Travis Gerke, Michaela Bowden, Andreas Pettersson, Massimo Loda, Howard D. Sesso, Philip W. Kantoff, Neil E. Martin, Edward L. Giovannucci, Svitlana Tyekucheva, Matthew Vander Heiden, and Lorelei A. Mucci. 2016. “The Role of Tumor Metabolism as a Driver of Prostate Cancer Progression and Lethal Disease: Results from a Nested Case-Control Study.” *Cancer & Metabolism* 4(1):1–9.

Kent, Lindsey N., and Gustavo Leone. 2019. “The Broken Cycle: E2F Dysfunction in Cancer.” *Nature Reviews Cancer* 19(6):326–38.

Khattar, Ekta, Pavanish Kumar, Chia Yi Liu, Semih Can Akincilar, Anandhkumar Raju, Manikandan Lakshmanan, Julien Jean Pierre Maury, Yu Qiang, Shang Li, Ern Yu Tan, Kam M. Hui, Ming Shi, Yuin Han Loh, and Vinay Tergaonkar. 2016. "Telomerase Reverse Transcriptase Promotes Cancer Cell Proliferation by Augmenting TRNA Expression." *Journal of Clinical Investigation* 126(10):4045–60.

Khot, Amit, Natalie Brajanovski, Donald P. Cameron, Nadine Hein, Kylee H. Maclachlan, Elaine Sanij, John Lim, John Soong, Emma Link, Piers Blombery, Ella R. Thompson, Andrew Fellowes, Karen E. Sheppard, Grant A. McArthur, Richard B. Pearson, Ross D. Hannan, Gretchen Poortinga, and Simon J. Harrison. 2019. "First-in-Human RNA Polymerase I Transcription Inhibitor CX-5461 in Patients with Advanced Hematologic Cancers: Results of a Phase I Dose-Escalation Study." *Cancer Discovery* 9(8):1036–49.

Khoueiry, P., A. Ward Gahlawat, M. Petretich, A. M. Michon, D. Simola, E. Lam, E. E. Furlong, V. Benes, M. A. Dawson, R. K. Prinjha, G. Drewes, and P. Grandi. 2019. "BRD4 Bimodal Binding at Promoters and Drug-Induced Displacement at Pol II Pause Sites Associates with I-BET Sensitivity." *Epigenetics and Chromatin* 12(1):1–18.

Kim, Nayun, and Sue Jinks-Robertson. 2012. "Transcription as a Source of Genome Instability." *Nature Reviews Genetics* 13(3):204–14.

Kim, Soochi, Youngjin Han, Se Ik Kim, Juwon Lee, Hyun A. Jo, Wenyu Wang, Untack Cho, Woong Yang Park, Thomas A. Rando, Danny N. Dhanasekaran, and Yong Sang Song. 2021. "Computational Modeling of Malignant Ascites Reveals CCL5–SDC4 Interaction in the Immune Microenvironment of Ovarian Cancer." *Molecular Carcinogenesis* 60(5):297–312.

Kimura, M. 2000. "Involvement of Multiple Subunit-Subunit Contacts in the Assembly of RNA Polymerase II." *Nucleic Acids Research* 28(4):952–59.

Kong, Nannan, Zeyuan Liu, and Ying Wai Chan. 2023. "RIF1 Suppresses the Formation of Single-Stranded Ultrafine Anaphase Bridges via Protein Phosphatase 1." *Cell Reports* 42(2):112032.

Kornberg, Roger D. 1999. "Eukaryotic Transcriptional Control." *Trends in Biochemical Sciences* 24(12):46–49.

Kostek, Seth A., Patricia Grob, Sacha De Carlo, J. Slaton Lipscomb, Florian Garczarek, and Eva Nogales. 2006. "Molecular Architecture and Conformational Flexibility of Human RNA Polymerase II." *Structure* 14(11):1691–1700.

Kotekar, Aparna, Amit Kumar Singh, and Ballachanda N. Devaiah. 2023. "BRD4 and MYC: Power Couple in Transcription and Disease." *FEBS Journal* 290(20):4820–42.

Kotsantis, Panagiotis, Rebecca M. Jones, Martin R. Higgs, and Eva Petermann. 2015. *Cancer Therapy and Replication Stress: Forks on the Road to Perdition*. Vol. 69. 1st ed. Elsevier Inc.

Kotsantis, Panagiotis, Lara Marques Silva, Sarah Irmscher, Rebecca M. Jones, Lisa Folkes, Natalia Gromak, and Eva Petermann. 2016. "Increased Global Transcription Activity as a

Mechanism of Replication Stress in Cancer." *Nature Communications* 7:1–13.

Krejci, Lumir, Veronika Altmannova, Mario Spirek, and Xiaolan Zhao. 2012. "Homologous Recombination and Its Regulation." *Nucleic Acids Research* 40(13):5795–5818.

Krupina, Ksenia, Alexander Goginashvili, and Don W. Cleveland. 2021. "Causes and Consequences of Micronuclei." *Current Opinion in Cell Biology* 70:91–99.

Kuehner, Jason N., Erika L. Pearson, and Claire Moore. 2011. "Unravelling the Means to an End: RNA Polymerase II Transcription Termination." *Nature Reviews Molecular Cell Biology* 12(5):283–94.

Kunkel, Thomas A. 2015. "Celebrating DNA's Repair Crew." *Cell* 163(6):1301–3.

Landry, Jonathan J. M., Paul Theodor Pyl, Tobias Rausch, Thomas Zichner, Manu M. Tekkedil, Adrian M. Stütz, Anna Jauch, Raeka S. Aiyar, Gregoire Pau, Nicolas Delhomme, Julien Gagneur, Jan O. Korbel, Wolfgang Huber, and Lars M. Steinmetz. 2013. "The Genomic and Transcriptomic Landscape of a Hela Cell Line." *G3: Genes, Genomes, Genetics* 3(8):1213–24.

Langmead, Ben, and Steven L. Salzberg. 2012. "Fast Gapped-Read Alignment with Bowtie 2." *Nature Methods* 9(4):357–59.

Larochelle, Stéphane, Jasmin Batliner, Matthew J. Gamble, Nora M. Barboza, Brian C. Kraybill, Justin D. Blethrow, Kevan M. Shokat, and Robert P. Fisher. 2006. "Dichotomous but Stringent Substrate Selection by the Dual-Function Cdk7 Complex Revealed by Chemical Genetics." *Nature Structural and Molecular Biology* 13(1):55–62.

Lautré, Wiebke, Elodie Richard, Jean Paul Feugeas, Hélène Dumay-Odelot, and Martin Teichmann. 2022. "The POLR3G Subunit of Human RNA Polymerase III Regulates Tumorigenesis and Metastasis in Triple-Negative Breast Cancer." *Cancers* 14(23).

Lee, Keng Boon, Dong Wang, Stephen J. Lippard, and Phillip A. Sharp. 2002. "Transcription-Coupled and DNA Damage-Dependent Ubiquitination of RNA Polymerase II in Vitro." *Proceedings of the National Academy of Sciences of the United States of America* 99(7):4239–44.

Leighton J. Core*, Joshua J. Waterfall*, and John T. Lis†. 2008. "Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters." 322(5909):1845–48.

Leighton J. Core and John T. Lis. 2008. "Transcription Regulation Through Promoter-Proximal Pausing of RNA Polymerase II." *Science* 319(5871):1791–92.

Lerdrup, Mads, Jens Vilstrup Johansen, Shuchi Agrawal-Singh, and Klaus Hansen. 2016. "An Interactive Environment for Agile Analysis and Visualization of ChIP-Sequencing Data." *Nature Structural and Molecular Biology* 23(4):349–57.

LeTallec, Benoît, Gaël Armel Millot, Marion Esther Blin, Olivier Brison, Bernard Dutrillaux, and Michelle Debatisse. 2013. "Common Fragile Site Profiling in Epithelial and Erythroid Cells Reveals That Most Recurrent Cancer Deletions Lie in Fragile Sites Hosting Large Genes."

Cell Reports 4(3):420–28.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, and Richard Durbin. 2009. “The Sequence Alignment/Map Format and SAMtools.” *Bioinformatics* 25(16):2078–79.

Li, Lan Ying, Hee Jung Kim, Sun Ae Park, So Hyun Lee, Lee Kyung Kim, Jung Yun Lee, Sunghoon Kim, Young Tae Kim, Sang Wun Kim, and Eun Ji Nam. 2019. “Genetic Profiles Associated with Chemoresistance in Patient-Derived Xenograft Models of Ovarian Cancer.” *Cancer Research and Treatment* 51(3):1117–27.

Li, Na, Lan Zhao, Chunyan Guo, Chang Liu, and Yongyu Liu. 2019. “Identification of a Novel DNA Repair-Related Prognostic Signature Predicting Survival of Patients with Hepatocellular Carcinoma.” *Cancer Management and Research* 11:7473–84.

Li, Yuanjun, Jie Huang, Lijun Bao, Junyi Zhu, Wenjia Duan, Haonan Zheng, Hui Wang, Yongpeng Jiang, Weiwei Liu, Meiling Zhang, Yang Yu, Chengqi Yi, and Xiong Ji. 2023. “RNA Pol II Preferentially Regulates Ribosomal Protein Expression by Trapping Disassociated Subunits.” *Molecular Cell* 83(8):1280-1297.e11.

Li, Yuanjun, Jie Huang, Junyi Zhu, Lijun Bao, Hui Wang, Yongpeng Jiang, Kai Tian, Rui Wang, Haonan Zheng, Wen Jia Duan, Weifeng Lai, Xiao Yi, Yi Zhu, Tiannan Guo, and Xiong Ji. 2022. “Targeted Protein Degradation Reveals RNA Pol II Heterogeneity and Functional Diversity.” *Molecular Cell* 82(20):3943-3959.e11.

Li, Yujing, Yifan Sun, Michael Kulke, Torsten Hechler, Kevin Van Der Jeught, Tianhan Dong, Bin He, Kathy D. Miller, Milan Radovich, Bryan P. Schneider, Andreas Pahl, Xinna Zhang, and Xiongbin Lu. 2021. “Targeted Immunotherapy ForHER2-Low Breast Cancer with 17p Loss.” *Science Translational Medicine* 13(580).

Li, Zhi-di. 2023. “POLR2J Is a Potential Biomarker for Abnormal Tumor Progression , Vorinostat Sensitization , Immune Infiltration , and Prognosis of Glioblastoma Multiform.” 1–29.

Liang, Xia, Renxiang Xie, Jinfeng Su, Bingqi Ye, Saisai Wei, Zhibing Liang, Rongpan Bai, Zhanghui Chen, Zhongxiang Li, and Xiangwei Gao. 2019. “Inhibition of RNA Polymerase III Transcription by Triptolide Attenuates Colorectal Tumorigenesis.” *Journal of Experimental and Clinical Cancer Research* 38(1):1–13.

Lieber, Michael R. 2011. “The Mechanism of DSB Repair by the NHEJ.” *Annual Review of Biochemistry* 79(3):181–211.

Lim, Gunhyoung, Seungha Hwang, Kilwon Yu, Jin Young Kang, Changwon Kang, and Sungchul Hohng. 2023. “Translocating RNA Polymerase Generates R-Loops at DNA Double-Strand Breaks without Any Additional Factors.” *Nucleic Acids Research* 51(18):9838–48.

Lin, Shengrong, Gabriela Coutinho-Mansfield, Dong Wang, Shatakshi Pandit, and Xiang-Dong Fu. 2008. “The Splicing Factor SC35 Has an Active Role in Transcriptional Elongation.” *Nature Structural & Molecular Biology* 15(8):819–26.

Lindahl, T., and D. E. Barnes. 2000. “Repair of Endogenous DNA Damage.” *Cold Spring Harbor Symposia on Quantitative Biology* 65:127–33.

Liu, Chao, Sriganesh Srihari, Kim Anh Lê Cao, Georgia Chenevix-Trench, Peter T. Simpson, Mark A. Ragan, and Kum Kum Khanna. 2014. “A Fine-Scale Dissection of the DNA Double-Strand Break Repair Machinery and Its Implications for Breast Cancer Therapy.” *Nucleic Acids Research* 42(10):6106–27.

Liu, Fan, Lan Wang, Fabiana Perna, and Stephen D. Nimer. 2016. “Beyond Transcription Factors: How Oncogenic Signalling Reshapes the Epigenetic Landscape.” *Nature Reviews Cancer* 16(6):359–72.

Liu, Xiangyang, Lucas Farnung, Christoph Wigge, and Patrick Cramer. 2018. “Cryo-EM Structure of a Mammalian RNA Polymerase II Elongation Complex Inhibited by α-Amanitin.” *Journal of Biological Chemistry* 293(19):7189–94.

Liu, Xin, David A. Bushnell, and Roger D. Kornberg. 2013. “RNA Polymerase II Transcription: Structure and Mechanism.” *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1829(1):2–8.

Liu, Yang, Wenxiao Jia, Ji Li, Hui Zhu, and Jinming Yu. 2020. “Identification of Survival-Associated Alternative Splicing Signatures in Lung Squamous Cell Carcinoma.” *Frontiers in Oncology* 10(September):1–14.

Liu, Ying, Christian F. Nielsen, Qi Yao, and Ian D. Hickson. 2014. “The Origins and Processing of Ultra Fine Anaphase DNA Bridges.” *Current Opinion in Genetics and Development* 26:1–5.

Liu, Yunhua, Xinna Zhang, Cecil Han, Guohui Wan, Xingxu Huang, Cristina Ivan, Dahai Jiang, Cristian Rodriguez-Aguayo, Gabriel Lopez-Berestein, Pulivarthi H. Rao, Dipen M. Maru, Andreas Pahl, Xiaoming He, Anil K. Sood, Lee M. Ellis, Jan Anderl, and Xiongbin Lu. 2015. “TP53 Loss Creates Therapeutic Vulnerability in Colorectal Cancer.” *Nature* 520(7549):697–701.

Lockhart, Arianna, Vanessa Borges Pires, Fabio Bento, Vanessa Kellner, Sarah Luke-Glaser, George Yakoub, Helle D. Ulrich, and Brian Luke. 2019. “RNase H1 and H2 Are Differentially Regulated to Process RNA-DNA Hybrids.” *Cell Reports* 29(9):2890-2900.e5.

Long, Nguyen Phuoc, Wun Jun Lee, Nguyen Truong Huy, Seul Ji Lee, Jeong Hill Park, and Sung Won Kwon. 2016. “Novel Biomarker Candidates for Colorectal Cancer Metastasis: A Meta-Analysis of in Vitro Studies.” *Cancer Informatics* 15:11–17.

Lopes, Inês, Gulam Altab, Priyanka Raina, and João Pedro de Magalhães. 2021. “Gene Size Matters: An Analysis of Gene Length in the Human Genome.” *Frontiers in Genetics* 12(February).

Lord, Christopher J., and Alan Ashworth. 2008. “Targeted Therapy for Cancer Using PARP Inhibitors.” *Current Opinion in Pharmacology* 8(4):363–69.

Low, Jin Yih, Paul Sirajuddin, Michael Moubarek, Shreya Agarwal, Apurv Rege, Gunes Guner, Hester Liu, Zhiming Yang, Angelo M. De Marzo, Charles Bieberich, and Marikki Laiho. 2019. “Effective Targeting of RNA Polymerase I in Treatment-Resistant Prostate Cancer.” *Prostate* 79(16):1837–51.

Lu, Huasong, Dan Yu, Anders S. Hansen, Sourav Ganguly, Rongdiao Liu, Alec Heckert, Xavier Darzacq, and Qiang Zhou. 2018. “Phase-Separation Mechanism for C-Terminal Hyperphosphorylation of RNA Polymerase II.” *Nature* 558(7709):318–23.

Lukas, Claudia, Velibor Savic, Simon Bekker-Jensen, Carsten Doil, Beate Neumann, Ronni Sølvhøj Pedersen, Merete Grøhfte, Kok Lung Chan, Ian David Hickson, Jiri Bartek, and Jiri Lukas. 2011. “53BP1 Nuclear Bodies Form around DNA Lesions Generated by Mitotic Transmission of Chromosomes under Replication Stress.” *Nature Cell Biology* 13(3):243–53.

Luo, Wei Ming, Zheng Yu Wang, and Xin Zhang. 2018. “Identification of Four Differentially Methylated Genes as Prognostic Signatures for Stage I Lung Adenocarcinoma.” *Cancer Cell International* 18(1):1–8.

Ma, Ke, Li Qiu, Kristin Mrasek, Jun Zhang, Thomas Liehr, Luciana Gonçalves Quintana, and Zheng Li. 2012. “Common Fragile Sites: Genomic Hotspots of DNA Damage and Carcinogenesis.” *International Journal of Molecular Sciences* 13(9):11974–99.

Maita, Hiroshi, and Shinichi Nakagawa. 2020. “What Is the Switch for Coupling Transcription and Splicing? RNA Polymerase II C-Terminal Domain Phosphorylation, Phase Separation and Beyond.” *Wiley Interdisciplinary Reviews: RNA* 11(1):1–15.

Manickavinayaham, Swarnalatha, Renier Velez-Cruz, Anup K. Biswas, Jie Chen, Ruifeng Guo, and David G. Johnson. 2020. “The E2F1 Transcription Factor and RB Tumor Suppressor Moonlight as DNA Repair Factors.” *Cell Cycle* 19(18):2260–69.

Marnef, Aline, and Gaëlle Legube. 2021. “R-Loops as Janus-Faced Modulators of DNA Repair.” *Nature Cell Biology* 23(4):305–13.

Mars, Jean Clément, Michel G. Tremblay, Mélissa Valere, Dany S. Sibai, Marianne Sabourin-Felix, Frédéric Lessard, and Tom Moss. 2020. “The Chemotherapeutic Agent CX-5461 Irreversibly Blocks RNA Polymerase I Initiation and Promoter Release to Cause Nucleolar Disruption, DNA Damage and Cell Inviability.” *NAR Cancer* 2(4):1–15.

Martinez-Rucobo, Fuensanta W., Rebecca Kohler, Michiel van de Waterbeemd, Albert J. R. Heck, Matthias Hemann, Franz Herzog, Holger Stark, and Patrick Cramer. 2015. “Molecular Basis of Transcription-Coupled Pre-MRNA Capping.” *Molecular Cell* 58(6):1079–89.

Martino, Fabrizio, Mohinder Pal, Hugo Muñoz-Hernández, Carlos F. Rodríguez, Rafael Núñez-Ramírez, David Gil-Carton, Gianluca Degliesposti, J. Mark Skehel, S. Mark Roe, Chrisostomos Prodromou, Laurence H. Pearl, and Oscar Llorca. 2018. “RPAP3 Provides a Flexible Scaffold for Coupling HSP90 to the Human R2TP Co-Chaperone Complex.” *Nature Communications* 9(1):1–13.

Masica, David L., and Rachel Karchin. 2011. “Correlation of Somatic Mutation and Expression Identifies Genes Important in Human Glioblastoma Progression and Survival.” *Cancer Research* 71(13):4550–61.

Mayer, Andreas, Heather M. Landry, and L. Stirling Churchman. 2018. “Pause & Go: From the

Discovery of RNA Polymerase Pausing to Its Functional Implications.” 72–80.

Mayfield, Joshua E., Seema Irani, Edwin E. Escobar, Zhao Zhang, Nathaniel T. Burkholder, Michelle R. Robinson, M. Rachel Mehaffey, Sarah N. Sipe, Wanjie Yang, Nicholas A. Prescott, Karan R. Kathuria, Zhijie Liu, Jennifer S. Brodbelt, and Yan Zhang. 2019. “TYR1 Phosphorylation Promotes Phosphorylation of Ser2 on the C-Terminal Domain of Eukaryotic RNA Polymerase Ii by p-TEFB.” *ELife* 8:1–27.

McElhinny, Stephanie A. Nic., Jody M. Havener, Miguel Garcia-Diaz, Raquel Juárez, Katarzyna Bebenek, Barbara L. Kee, Luis Blanco, Thomas A. Kunkel, and Dale A. Ramsden. 2005. “A Gradient of Template Dependence Defines Distinct Biological Roles for Family X Polymerases in Nonhomologous End Joining.” *Molecular Cell* 19(3):357–66.

Miao, Weili, David Bade, and Yinsheng Wang. 2021. “Targeted Proteomic Analysis Revealed Kinome Reprogramming during Acquisition of Radioresistance in Breast Cancer Cells.” *Journal of Proteome Research* 20(5):2830–38.

Michels, Annemieke A., and Nouria Hernandez. 2006. “Does Pol I Talk to Pol II? Coordination of RNA Polymerases in Ribosome Biogenesis.” *Genes and Development* 20(15):1982–85.

Mimoso, Claudia A., and Karen Adelman. 2023. “U1 SnRNP Increases RNA Pol II Elongation Rate to Enable Synthesis of Long Genes.” *Molecular Cell* 83(8):1264-1279.e10.

Mirza-Aghazadeh-Attari, Mohammad, Amir Mohammadzadeh, Bahman Yousefi, Ainaz Mihanfar, Ansar Karimian, and Maryam Majidinia. 2019. “53BP1: A Key Player of DNA Damage Response with Critical Functions in Cancer.” *DNA Repair* 73(November 2018):110–19.

Mladenov, Emil, Simon Magin, Aashish Soni, and George Iliakis. 2016a. “DNA Double-Strand-Break Repair in Higher Eukaryotes and Its Role in Genomic Instability and Cancer: Cell Cycle and Proliferation-Dependent Regulation.” *Seminars in Cancer Biology* 37–38:51–64.

Mladenov, Emil, Simon Magin, Aashish Soni, and George Iliakis. 2016b. “DNA Double-Strand-Break Repair in Higher Eukaryotes and Its Role in Genomic Instability and Cancer: Cell Cycle and Proliferation-Dependent Regulation.” *Seminars in Cancer Biology* 37–38:51–64.

Moeglin, Eric, Dominique Desplancq, Sascha Conic, Mustapha Oulad-Abdelghani, Audrey Stoessel, Manuela Chiper, Marc Vigneron, Pascal Didier, Laszlo Tora, and Etienne Weiss. 2019. “Uniform Widespread Nuclear Phosphorylation of Histone H2AX Is an Indicator of Lethal DNA Replication Stress.” *Cancers* 11(3):1–22.

Moon, Kyoo Jang, Kazuki Mochizuki, Meisheng Zhou, Ho Sang Jeong, John N. Brady, and Keiko Ozato. 2005. “The Bromodomain Protein Brd4 Is a Positive Regulatory Component of P-TEFb and Stimulates RNA Polymerase II-Dependent Transcription.” *Molecular Cell* 19(4):523–34.

Morales, Julio C., Patricia Richard, Praveen L. Patidar, Edward A. Motea, Tuyen T. Dang, James L. Manley, and David A. Boothman. 2016. “XRN2 Links Transcription Termination to DNA

Damage and Replication Stress." *PLoS Genetics* 12(7):1–22.

Muller, Patricia A. J., and Karen H. Vousden. 2013. "P53 Mutations in Cancer." *Nature Cell Biology* 15(1):2–8.

Musolf, Anthony M., Bilal A. Moiz, Haiming Sun, Claudio W. Pikielny, Yohan Bosse, Diptasri Mandal, Mariza De Andrade, Colette Gaba, Ping Yang, Yafang Li, Ming You, Ramaswamy Govindan, Richard K. Wilson, Elena Y. Kupert, Marshall W. Anderson, Ann G. Schwartz, Susan M. Pinney, Christopher I. Amos, and Joan E. Bailey-Wilson. 2020. "Whole Exome Sequencing of Highly Aggregated Lung Cancer Families Reveals Linked Loci for Increased Cancer Risk on Chromosomes 12q, 7p, and 4q." *Cancer Epidemiology Biomarkers and Prevention* 29(2):434–42.

Muste Sadurni, Martina, and Marco Saponaro. 2023. "Deregulations of RNA Pol II Subunits in Cancer." *Applied Biosciences* 2(3):459–76.

Näär, A. M., B. D. Lemon, and R. Tjian. 2001. "Transcriptional Coactivator Complexes." *Annual Review of Biochemistry* 70(May):475–501.

Nakazawa, Yuka, Yuichiro Hara, Yasuyoshi Oka, Okiru Komine, Diana van den Heuvel, Chaowan Guo, Yasukazu Daigaku, Mayu Isono, Yuxi He, Mayuko Shimada, Kana Kato, Nan Jia, Satoru Hashimoto, Yuko Kotani, Yuka Miyoshi, Miyako Tanaka, Akira Sobue, Norisato Mitsutake, Takayoshi Suganami, Akio Masuda, Kinji Ohno, Shinichiro Nakada, Tomoji Mashimo, Koji Yamanaka, Martijn S. Luijsterburg, and Tomoo Ogi. 2020. "Ubiquitination of DNA Damage-Stalled RNAPII Promotes Transcription-Coupled Repair." *Cell* 180(6):1228-1244.e24.

Naorem, Leimarembi Devi, Mathavan Muthaiyan, and Amouda Venkatesan. 2019. "Integrated Network Analysis and Machine Learning Approach for the Identification of Key Genes of Triple-Negative Breast Cancer." *Journal of Cellular Biochemistry* 120(4):6154–67.

Natrajan, Rachael, Æ. Britta Weigelt, Æ. Alan Mackay, Æ. Felipe C. Geyer, Anita Grigoriadis, Æ. David S. P. Tan, Æ. Chris Jones, Æ. Christopher J. Lord, Radost Vatcheva, Æ. Socorro M. Rodriguez-pinilla Æ. Jose, and Alan Ashworth Æ. Jorge S. Reis-filho. 2010. "An Integrative Genomic and Transcriptomic Analysis Reveals Molecular Pathways and Networks Regulated by Copy Number Aberrations in Basal-like , HER2 and Luminal Cancers." 575–89.

Neelsen, Kai J., and Massimo Lopes. 2015. "Replication Fork Reversal in Eukaryotes: From Dead End to Dynamic Response." *Nature Reviews Molecular Cell Biology* 16(4):207–20.

Negrini, Simona, Vassilis G. Gorgoulis, and Thanos D. Halazonetis. 2010. "Genomic Instability an Evolving Hallmark of Cancer." *Nature Reviews Molecular Cell Biology* 11(3):220–28.

Nicholas Kwiatkowski1, 2, 3,*, Tinghu Zhang1, 2,*, Peter B Rahl3, Brian J Abraham3, Jessica Reddy3, 4, Scott B Ficarro1, 2, 5, Anahita Dastur6, Arnaud Amzallag6, 7, Sridhar Ramaswamy6, 7, Bethany Tesar8, 9, Christopher R Jenkins10, Nancy M Hannett3, Dougl, 2 1Department. 2014. "Targeting Transcription Regulation in Cancer with a Covalent CDK7 Inhibitor." *Nature* 511(7511):616–20.

Niesser, Jürgen, Felix R. Wagner, Dirk Kostrewa, Wolfgang Mühlbacher, and Patrick Cramer. 2016. “Structure of GPN-Loop GTPase Npa3 and Implications for RNA Polymerase II Assembly.” *Molecular and Cellular Biology* 36(5):820–31.

Nojima, Takayuki, Tomás Gomes, Ana Rita Fialho Grosso, Hiroshi Kimura, Michael J. Dye, Somdutta Dhir, Maria Carmo-Fonseca, and Nicholas J. Proudfoot. 2015. “Mammalian NET-Seq Reveals Genome-Wide Nascent Transcription Coupled to RNA Processing.” *Cell* 161(3):526–40.

Nojima, Takayuki, Kenny Rebelo, Tomás Gomes, Ana Rita Grosso, Nicholas J. Proudfoot, and Maria Carmo-Fonseca. 2018. “RNA Polymerase II Phosphorylated on CTD Serine 5 Interacts with the Spliceosome during Co-Transcriptional Splicing.” *Molecular Cell* 72(2):369-379.e4.

Normal, Patient-derived, Tumor Colon, Alba Costales-carrera, Pilar Bustamante-madrid, Aurora Burgos, Antonio Barb, and Alberto Muñoz. 2020. “Comparative Study of Organoids From.”

Ochi, Takashi, Andrew N. Blackford, Julia Coates, Satpal Jhujh, Shahid Mehmood, Naoka Tamura, Jon Travers, Qian Wu, Viji M. Draviam, Carol V Robinson, L. Tom, and Stephen P. Jackson. 2015. “Europe PMC Funders Group Europe PMC Funders Author Manuscripts PAXX , a Paralog of XRCC4 and XLF , Interacts with Ku to Promote DNA Double-Strand Break Repair **.” 347(6218):185–88.

Ohle, Corina, Rafael Tesorero, Géza Schermann, Nikolay Dobrev, Irmgard Sinning, and Tamás Fischer. 2016. “Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair.” *Cell* 167(4):1001-1013.e7.

ojoong Kwak, Nicholas J. Fuda, Leighton J. Core, and John T. Lis. 12435. “Precise Maps of RNA Polymerase Reveal How Promoters Direct Initiation and Pausing.” *Science* 339(6122):950–53.

Oler, Andrew J., Ravi K. Alla, Douglas N. Roberts, Alexander Wong, Peter C. Hollenhorst, Katherine J. Chandler, Patrick A. Cassiday, Cassie A. Nelson, Curt H. Hagedorn, Barbara J. Graves, and Bradley R. Cairns. 2010. “Human RNA Polymerase III Transcriptomes and Relationships to Pol II Promoter Chromatin and Enhancer-Binding Factors.” *Nature Structural and Molecular Biology* 17(5):620–28.

Orian-Rousseau, Véronique, Sigrun Mink, Jörg Mengwasser, Harm HogenEsch, Feng Guo, Wolf Gerolf Thies, Martin Hofmann, Peter Herrlich, and Helmut Ponta. 2005. “Genes Upregulated in a Metastasizing Human Colon Carcinoma Cell Line.” *International Journal of Cancer* 113(5):699–705.

Ouyang, Jian, Tribhuwan Yadav, Jia Min Zhang, Haibo Yang, Esther Rheinbay, Hongshan Guo, Daniel A. Haber, Li Lan, and Lee Zou. 2021. “RNA Transcripts Stimulate Homologous Recombination by Forming DR-Loops.” *Nature* 594(7862):283–88.

Pan, Jingxuan. 2010. “RNA Polymerase - An Important Molecular Target of Triptolide in Cancer Cells.” *Cancer Letters* 292(2):149–52.

Parsels, Leslie A., Joshua D. Parsels, Daria M. Tanska, Jonathan Maybaum, Theodore S. Lawrence, and Meredith A. Morgan. 2018. “The Contribution of DNA Replication Stress Marked by High-Intensity, Pan-Nuclear ΓH2AX Staining to Chemosensitization by CHK1 and WEE1 Inhibitors.” *Cell Cycle* 17(9):1076–86.

Patel, Shetal A., and M. Celeste Simon. 2010. “Functional Analysis of the Cdk7·cyclin H·Mat1 Complex in Mouse Embryonic Stem Cells and Embryos.” *Journal of Biological Chemistry* 285(20):15587–98.

Pavon-Eternod, Mariana, Suzanna Gomes, Renaud Geslain, Qing Dai, Marsha Rich Rosner, and Tao Pan. 2009. “TRNA Over-Expression in Breast Cancer and Functional Consequences.” *Nucleic Acids Research* 37(21):7268–80.

Peace, Jared M., Sandra K. Villwock, John L. Zeytounian, Yan Gan, and Oscar M. Aparicio. 2016. “Quantitative BrdU Immunoprecipitation Method Demonstrates That Fkh1 and Fkh2 Are Rate-Limiting Activators of Replication Origins That Reprogram Replication Timing in G1 Phase.” *Genome Research* 26(3):365–75.

Peck, Sarah A., Katlyn D. Hughes, Jose F. Victorino, and Amber L. Mosley. 2019. “Writing a Wrong: Coupled RNA Polymerase II Transcription and RNA Quality Control.” *Wiley Interdisciplinary Reviews: RNA* 10(4):1–22.

Peterson, Craig L. 2002. “Chromatin Remodeling: Nucleosomes Bulging at the Seams.” *Current Biology* 12(7):245–47.

Petrie, John L., Caroline Swan, Richard M. Ingram, Fiona M. Frame, Anne T. Collins, Hélène Dumay-Odelot, Martin Teichmann, Norman J. Maitland, and Robert J. White. 2019. “Effects on Prostate Cancer Cells of Targeting RNA Polymerase III.” *Nucleic Acids Research* 47(8):3937–56.

Petropoulos, Michalis, Spyridon Champeris Tsaniras, Stavros Taraviras, and Zoi Lygerou. 2019. “Replication Licensing Aberrations, Replication Stress, and Genomic Instability.” *Trends in Biochemical Sciences* 44(9):752–64.

Pinto, Joseph A., Jhajaira Araujo, Nadezhda K. Cardenas, Zaida Morante, Franco Doimi, Tatiana Vidaurre, Justin M. Balko, and Henry L. Gomez. 2016. “A Prognostic Signature Based on Three-Genes Expression in Triple-Negative Breast Tumours with Residual Disease.” *Npj Genomic Medicine* 1(May 2015).

Piwko, Wojciech, Lucie J. Mlejnkova, Karun Mutreja, Lepakshi Ranjha, Diana Stafa, Alexander Smirnov, Mia ML Brodersen, Ralph Zellweger, Andreas Sturzenegger, Pavel Janscak, Massimo Lopes, Matthias Peter, and Petr Cejka. 2016. “The MMS22L–TONSL Heterodimer Directly Promotes RAD51-dependent Recombination upon Replication Stress.” *The EMBO Journal* 35(23):2584–2601.

Plaschka, C., L. Larivière, L. Wenzeck, M. Seizl, M. Hemann, D. Tegunov, E. V. Petrotchenko, C. H. Borchers, W. Baumeister, F. Herzog, E. Villa, and P. Cramer. 2015. “Architecture of the RNA Polymerase II-Mediator Core Initiation Complex.” *Nature* 518(7539):376–80.

Poortinga, G., L. M. Quinn, and R. D. Hannan. 2015. “Targeting RNA Polymerase i to Treat MYC-

Driven Cancer." *Oncogene* 34(4):403–12.

Qiu, Shan, Guixing Jiang, Liping Cao, and Jun Huang. 2021. "Replication Fork Reversal and Protection." *Frontiers in Cell and Developmental Biology* 9(May):1–8.

Quinet, Annabel, Denisse Carvajal-Maldonado, Delphine Lemacon, and Alessandro Vindigni. 2017. *DNA Fiber Analysis: Mind the Gap!* Vol. 591. 1st ed. Elsevier Inc.

Quinet, Annabel, Delphine Lemaçon, and Alessandro Vindigni. 2017. "Replication Fork Reversal: Players and Guardians." *Molecular Cell* 68(5):830–33.

Rajput, Ashwani, Ekta Agarwal, Premila Leiphrakpam, Michael G. Brattain, and Sanjib Chowdhury. 2013. "Establishment and Validation of an Orthotopic Metastatic Mouse Model of Colorectal Cancer." *ISRN Hepatology* 2013:1–9.

Ramasamy, Shyam, Abrar Aljahani, Magdalena A. Karpinska, T. B. Ngo. Cao, Taras Velychko, J. Neos Cruz, Michael Lidschreiber, and A. Marieke Oudelaar. 2023. *The Mediator Complex Regulates Enhancer-Promoter Interactions*. Vol. 30. Springer US.

Ramsay, Ewan Phillip, Guillermo Abascal-Palacios, Julia L. Daiß, Helen King, Jerome Gouge, Michael Pilsl, Fabienne Beuron, Edward Morris, Philip Gunkel, Christoph Engel, and Alessandro Vannini. 2020. "Structure of Human RNA Polymerase III." *Nature Communications* 11(1).

Ranjha, Lepakshi, Sean M. Howard, and Petr Cejka. 2018. "Main Steps in DNA Double-Strand Break Repair: An Introduction to Homologous Recombination and Related Processes." *Chromosoma* 127(2):187–214.

Ren, Jianwei, Yi Liu, Shanshan Wang, Yu Wang, Wende Li, Siyu Chen, Dexuan Cui, Shengli Yang, Ming Yue Li, Bo Feng, Paul B. S. Lai, and George G. Chen. 2020. "The FKH Domain in FOXP3 MRNA Frequently Contains Mutations in Hepatocellular Carcinoma That Influence the Subcellular Localization and Functions of FOXP3." *Journal of Biological Chemistry* 295(16):5484–95.

Rengachari, Srinivasan, Sandra Schilbach, Shintaro Aibara, Christian Dienemann, and Patrick Cramer. 2021. "Structure of the Human Mediator–RNA Polymerase II Pre-Initiation Complex." *Nature* 594(7861):129–33.

Reppas, Nikos B., Joseph T. Wade, George M. M. Church, and Kevin Struhl. 2006. "The Transition between Transcriptional Initiation and Elongation in E. Coli Is Highly Variable and Often Rate Limiting." *Molecular Cell* 24(5):747–57.

Richard, Stephen, Lital Gross, Jonathan Fischer, Keren Bendalak, Tamar Ziv, Shira Urim, and Mordechai Choder. 2021. "Numerous Post-Translational Modifications of RNA Polymerase II Subunit Rpb4/7 Link Transcription to Post-Transcriptional Mechanisms." *Cell Reports* 34(2):108578.

Richard, Stephen, Lital Gross Gross, Jonathan Fischer, Keren Bendalak, Tamar Ziv, Shira Urim, and Mordechai Choder. 2019. "Numerous Post-Translational Modifications of RNA Polymerase II Subunit Rpb4 Link Transcription to Post-Transcriptional Mechanisms." *SSRN Electronic Journal*.

Richter, William F., Shraddha Nayak, Janet Iwasa, and Dylan J. Taatjes. 2022. “The Mediator Complex as a Master Regulator of Transcription by RNA Polymerase II.” *Nature Reviews Molecular Cell Biology* 23(11):732–49.

Rimel, Jenna K., and Dylan J. Taatjes. 2018. “The Essential and Multifunctional TFIIH Complex.” *Protein Science* 27(6):1018–37.

Robinson, James T., Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, and Jill P. Mesirov. 2011. “Integrative Genomics Viewer.” *Nature Biotechnology* 29(1):24–26.

Robinson, Philip J., Michael J. Trnka, David A. Bushnell, Ralph E. Davis, Pierre Jean Mattei, Alma L. Burlingame, and Roger D. Kornberg. 2016. “Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex.” *Cell* 166(6):1411-1422.e16.

Rodríguez-Molina, Juan B., Steven West, and Lori A. Passmore. 2023. “Knowing When to Stop: Transcription Termination on Protein-Coding Genes by Eukaryotic RNAPII.” *Molecular Cell* 83(3):404–15.

Rosenfeld, Michael G., Victoria V. Lunyak, and Christopher K. Glass. 2006. “Sensors and Signals: A Coactivator/Corepressor/Epigenetic Code for Integrating Signal-Dependent Programs of Transcriptional Response.” *Genes and Development* 20(11):1405–28.

Sainsbury, Sarah, Carrie Bernecky, and Patrick Cramer. 2015. “Structural Basis of Transcription Initiation by RNA Polymerase II.” *Nature Reviews Molecular Cell Biology* 16(3):129–43.

Sample, Klarke M. 2020. “DNA Repair Gene Expression Is Associated with Differential Prognosis between HPV16 and HPV18 Positive Cervical Cancer Patients Following Radiation Therapy.” *Scientific Reports* 10(1):1–9.

Sanchez-Martin, Victoria, David A. Schneider, Matilde Ortiz-Gonzalez, Ana Soriano-Lerma, Angel Linde-Rodriguez, Virginia Perez-Carrasco, Jose Gutierrez-Fernandez, Marta Cuadros, Carlos González, Miguel Soriano, and Jose A. Garcia-Salcedo. 2021. “Targeting Ribosomal G-Quadruplexes with Naphthalene-Diimides as RNA Polymerase I Inhibitors for Colorectal Cancer Treatment.” *Cell Chemical Biology* 28(11):1590-1601.e4.

Sánchez-Tilló, Ester, Yongqing Liu, Oriol De Barrios, Laura Siles, Lucia Fanlo, Miriam Cuatrecasas, Douglas S. Darling, Douglas C. Dean, Antoni Castells, and Antonio Postigo. 2012. “EMT-Activating Transcription Factors in Cancer: Beyond EMT and Tumor Invasiveness.” *Cellular and Molecular Life Sciences* 69(20):3429–56.

Sanij, Elaine, Jeannine Diesch, Analia Lesmana, Gretchen Poortinga, Nadine Hein, Grace Lidgerwood, Donald P. Cameron, Jason Ellul, Gregory J. Goodall, Lee H. Wong, Amardeep S. Dhillon, Nourdine Hamdane, Lawrence I. Rothblum, Richard B. Pearson, Izhak Haviv, and Ross D. Hannan. 2015. “A Novel Role for the Pol I Transcription Factor Ubtf in Maintaining Genome Stability through the Regulation of Highly Transcribed Pol II Genes.” *Genome Research* 25(2):201–12.

Sanij, Elaine, Katherine M. Hannan, Jiachen Xuan, Shunfei Yan, Jessica E. Ahern, Anna S. Trigos, Natalie Brajanovski, Jinbae Son, Keefe T. Chan, Olga Kondrashova, Elizabeth Lieschke,

Matthew J. Wakefield, Daniel Frank, Sarah Ellis, Carleen Cullinane, Jian Kang, Gretchen Poortinga, Purba Nag, Andrew J. Deans, Kum Kum Khanna, Linda Mileshkin, Grant A. McArthur, John Soong, Els M. J. J. Berns, Ross D. Hannan, Clare L. Scott, Karen E. Sheppard, and Richard B. Pearson. 2020. "CX-5461 Activates the DNA Damage Response and Demonstrates Therapeutic Efficacy in High-Grade Serous Ovarian Cancer." *Nature Communications* 11(1).

Saponaro, Marco, Theodoros Kantidakis, Richard Mitter, Gavin P. Kelly, Mark Heron, Hannah Williams, Johannes Söding, Aengus Stewart, and Jesper Q. Svejstrup. 2014. "RECQL5 Controls Transcript Elongation and Suppresses Genome Instability Associated with Transcription Stress." *Cell* 157(5):1037–49.

Sattarifard, Hedieh, Mohammad Hashemi, Shekoufeh Hassanzarei, Abbas Basiri, Behzad Narouie, and Saeid Ghavami. 2019. "Long Non-Coding RNA POLR2E Gene Polymorphisms Increased the Risk of Prostate Cancer in a Sample of the Iranian Population." *Nucleosides, Nucleotides and Nucleic Acids* 38(1):1–11.

Saxena, Sneha, and Lee Zou. 2022. "Hallmarks of DNA Replication Stress." *Molecular Cell* 82(12):2298–2314.

Schier, Allison C., and Dylan J. Taatjes. 2020. "Structure and Mechanism of the RNA Polymerase II Transcription Machinery." *Genes & Development* 34(7–8):465–88.

Schüller, Roland, Ignasi Forné, Tobias Straub, Amelie Schreieck, Yves Texier, Nilay Shah, Tim Michael Decker, Patrick Cramer, Axel Imhof, and Dirk Eick. 2016. "Heptad-Specific Phosphorylation of RNA Polymerase II CTD." *Molecular Cell* 61(2):305–14.

Schwab, Rebekka A., Jadwiga Nieminuszczy, Fenil Shah, Jamie Langton, David Lopez Martinez, Chih Chao Liang, Martin A. Cohn, Richard J. Gibbons, Andrew J. Deans, and Wojciech Niedzwiedz. 2015. "The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription." *Molecular Cell* 60(3):351–61.

Scott, P. H., C. A. Cairns, J. E. Sutcliffe, H. M. Alzuherri, A. McLees, A. G. Winter, and R. J. White. 2001. "Regulation of RNA Polymerase III Transcription during Cell Cycle Entry." *Journal of Biological Chemistry* 276(2):1005–14.

Scully, Ralph, Arvind Panday, Rajula Elango, and Nicholas A. Willis. 2019. "DNA Double-Strand Break Repair-Pathway Choice in Somatic Mammalian Cells." *Nature Reviews Molecular Cell Biology* 20(11):698–714.

Serra, O., M. Galán, M. M. Ginesta, M. Calvo, N. Sala, and R. Salazar. 2019. "Comparison and Applicability of Molecular Classifications for Gastric Cancer." *Cancer Treatment Reviews* 77(March):29–34.

Shah, Pragya, Chad M. Hobson, Svea Cheng, Marshall J. Colville, Matthew J. Paszek, Richard Superfine, and Jan Lammerding. 2021. "Nuclear Deformation Causes DNA Damage by Increasing Replication Stress." *Current Biology* 31(4):753-765.e6.

Sheffer, Michal, Manny D. Bacolod, Or Zuk, Sarah F. Giardina, Hanna Pincas, Francis Barany, Philip B. Paty, William L. Gerald, Daniel A. Notterman, and Eytan Domany. 2009.

"Association of Survival and Disease Progression with Chromosomal Instability: A Genomic Exploration of Colorectal Cancer." *Proceedings of the National Academy of Sciences of the United States of America* 106(17):7131–36.

Shin, Ji Hyun, Liang Xu, and Dong Wang. 2016. "RNA Polymerase II Acts as a Selective Sensor for DNA Lesions and Endogenous DNA Modifications." *Transcription* 7(3):57–62.

Singh, Vindhya Vikram, and Shafqat Alauddin. 2023. "Review on: BRD4 Inhibitors for Anticancer Research." *Human Gene* 37(May):201196.

Singleton, B. K., M. I. Torres-Arzayus, S. T. Rottinghaus, G. E. Taccioli, and P. A. Jeggo. 1999. "The C Terminus of Ku80 Activates the DNA-Dependent Protein Kinase Catalytic Subunit." *Molecular and Cellular Biology* 19(5):3267–77.

Skourti-Stathaki, Konstantina, and Nicholas J. Proudfoot. 2014. "A Double-Edged Sword: R Loops as Threats to Genome Integrity and Powerful Regulators of Gene Expression." *Genes and Development* 28(13):1384–96.

Skourti-Stathaki, Konstantina, Nicholas J. Proudfoot, and Natalia Gromak. 2011. "Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination." *Molecular Cell* 42(6):794–805.

Slebos, Robbert J. C., Yajun Yi, Kim Ely, Jesse Carter, Amy Evjen, Xueqiong Zhang, Yu Shyr, Barbara M. Murphy, Anthony J. Cmelak, Brian B. Burkey, James L. Netterville, Shawn Levy, Wendell G. Yarbrough, and Christine H. Chung. 2006. "Human Cancer Biology Gene Expression Differences Associated with Human Papillomavirus Status in Head and Neck Squamous Cell Carcinoma." 12(3):701–9.

Smolka, John A., Lionel A. Sanz, Stella R. Hartono, and Frédéric Chédin. 2021. "Recognition of Rna by the S9.6 Antibody Creates Pervasive Artifacts When Imaging Rna:Dna Hybrids." *Journal of Cell Biology* 220(6).

Sollier, Julie, Caroline Townsend Stork, María L. García-Rubio, Renee D. Paulsen, Andrés Aguilera, and Karlene A. Cimprich. 2014. "Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability." *Molecular Cell* 56(6):777–85.

Song, Yawei, Gongcheng Hu, Jinping Jia, Mingze Yao, Xiaoshan Wang, Wenliang Lu, Andrew P. Hutchins, Jiekai Chen, Keiko Ozato, and Hongjie Yao. 2020. "DNA Damage Induces Dynamic Associations of BRD4/P-TEFb With Chromatin and Modulates Gene Transcription in a BRD4-Dependent and -Independent Manner." *Frontiers in Molecular Biosciences* 7(December):1–12.

Soutourina, Julie. 2018. "Transcription Regulation by the Mediator Complex." *Nature Reviews Molecular Cell Biology* 19(4):262–74.

Spiegelman, Bruce M., and Reinhart Heinrich. 2004. "Biological Control through Regulated Transcriptional Coactivators." *Cell* 119(2):157–67.

Stein, Torsten, Diane Crighton, John M. Boyle, Jennifer M. Varley, and Robert J. White. 2002. "RNA Polymerase III Transcription Can Be Derepressed by Oncogenes or Mutations That

Compromise P53 Function in Tumours and Li-Fraumeni Syndrome." *Oncogene* 21(19):2961–70.

Stiff, Tom, Mark O'Driscoll, Nicole Rief, Kuniyoshi Iwabuchi, Markus Löbrich, and Penny A. Jeggo. 2004. "ATM and DNA-PK Function Redundantly to Phosphorylate H2AX after Exposure to Ionizing Radiation." *Cancer Research* 64(7):2390–96.

Strano, S., S. Dell'Orso, S. Di Agostino, G. Fontemaggi, A. Sacchi, and G. Blandino. 2007. "Mutant P53: An Oncogenic Transcription Factor." *Oncogene* 26(15):2212–19.

Stucki, Manuel, Julie A. Clapperton, Duaa Mohammad, Michael B. Yaffe, Stephen J. Smerdon, and Stephen P. Jackson. 2005. "MDC1 Directly Binds Phosphorylated Histone H2AX to Regulate Cellular Responses to DNA Double-Strand Breaks." *Cell* 123(7):1213–26.

Suh, Hyunsuk, Scott B. Ficarro, Un Beom Kang, Yujin Chun, Jarrod A. Marto, and Stephen Buratowski. 2016a. "Direct Analysis of Phosphorylation Sites on the Rpb1 C-Terminal Domain of RNA Polymerase II." *Molecular Cell* 61(2):297–304.

Suh, Hyunsuk, Scott B. Ficarro, Un Beom Kang, Yujin Chun, Jarrod A. Marto, and Stephen Buratowski. 2016b. "Direct Analysis of Phosphorylation Sites on the Rpb1 C-Terminal Domain of RNA Polymerase II." *Molecular Cell* 61(2):297–304.

Sweetser, D., M. Nonet, and R. A. Young. 1987. "Prokaryotic and Eukaryotic RNA Polymerases Have Homologous Core Subunits." *Proceedings of the National Academy of Sciences of the United States of America* 84(5):1192–96.

Tellier, Michael, Isabella Maudlin, and Shona Murphy. 2020. "Transcription and Splicing: A Two-Way Street." *Wiley Interdisciplinary Reviews: RNA* 11(5):1–25.

Teloni, Federico, Jone Michelena, Aleksandra Lezaja, Sinan Kilic, Christina Ambrosi, Shruti Menon, Jana Dobrovolna, Ralph Imhof, Pavel Janscak, Tuncay Baubec, and Matthias Altmeyer. 2019. "Efficient Pre-MRNA Cleavage Prevents Replication-Stress-Associated Genome Instability." *Molecular Cell* 73(4):670-683.e12.

Thomas, Mary C., and Cheng Ming Chiang. 2006. "The General Transcription Machinery and General Cofactors." *Critical Reviews in Biochemistry and Molecular Biology* 41(3):105–78.

Tian, Yong, Fan Sun, Yong Zhong, Wenhua Huang, Guowu Wang, Chaoqing Liu, Yujie Xiao, Jiangdong Wu, and Lati Mu. 2020. "Expression and Clinical Significance of POLR1D in Colorectal Cancer." *Oncology (Switzerland)* 98(3):138–45.

Tora, Laszlo, and H. Th Marc Timmers. 2010. "The TATA Box Regulates TATA-Binding Protein (TBP) Dynamics in Vivo." *Trends in Biochemical Sciences* 35(6):309–14.

Tsantoulis, P. K., and V. G. Gorgoulis. 2005. "Involvement of E2F Transcription Factor Family in Cancer." *European Journal of Cancer* 41(16):2403–14.

Tseng, Wei Che, Chi Yuan Chen, Ching Yuh Chern, Chu An Wang, Wen Chih Lee, Ying Chih Chi, Shu Fang Cheng, Yi Tsen Kuo, Ya Chen Chiu, Shih Ting Tseng, Pei Ya Lin, Shou Jhen Liou, Yi Chen Li, and Chin Chuan Chen. 2021. "Targeting Hr Repair as a Synthetic Lethal Approach to Increase Dna Damage Sensitivity by a Rad52 Inhibitor in Brca2-Deficient

Cancer Cells.” *International Journal of Molecular Sciences* 22(9).

Tubbs, Anthony, and André Nussenzweig. 2017. “Endogenous DNA Damage as a Source of Genomic Instability in Cancer.” *Cell* 168(4):644–56.

Tufegdžić Vidaković, Ana, Richard Mitter, Gavin P. Kelly, Michelle Neumann, Michelle Harreman, Marta Rodríguez-Martínez, Anna Herlihy, Juston C. Weems, Stefan Boeing, Vesela Encheva, Liam Gaul, Laura Milligan, David Tollervey, Ronald C. Conaway, Joan W. Conaway, Ambrosius P. Snijders, Aengus Stewart, and Jesper Q. Svejstrup. 2020. “Regulation of the RNAPII Pool Is Integral to the DNA Damage Response.” *Cell* 180(6):1245-1261.e21.

Turinetto, Valentina, Paola Porcedda, Luca Orlando, Mario De Marchi, Antonio Amoroso, and Claudia Giachino. 2009. “The Cyclin-Dependent Kinase Inhibitor 5, 6-Dichloro-1-Beta-D-Ribofuranosylbenzimidazole Induces Nongenotoxic, DNA Replication-Independent Apoptosis of Normal and Leukemic Cells, Regardless of Their P53 Status.” *BMC Cancer* 9:281.

Uckelmann, Michael, and Titia K. Sixma. 2017. “Histone Ubiquitination in the DNA Damage Response.” *DNA Repair* 56(June):92–101.

Vannini, Alessandro, and Patrick Cramer. 2012. “Conservation between the RNA Polymerase I, II, and III Transcription Initiation Machineries.” *Molecular Cell* 45(4):439–46.

Vos, Seychelle M., Lucas Farnung, Henning Urlaub, and Patrick Cramer. 2018. “Structure of Paused Transcription Complex Pol II–DSIF–NELF.” *Nature* 560(7720):601–6.

Wagner, Eric J., Liang Tong, and Karen Adelman. 2023. “Integrator Is a Global Promoter-Proximal Termination Complex.” *Molecular Cell* 83(3):416–27.

Walker, John R., Richard A. Corpina, and Jonathan Goldberg. 2001. “Structure of the Ku Heterodimer Bound to Dna and Its Implications for Double-Strand Break Repair.” *Nature* 412(6847):607–14.

Walmacq, Celine, Maria L. Kireeva, Jordan Irvin, Yuri Nedialkov, Lucyna Lubkowska, Francisco Malagon, Jeffrey N. Strathern, and Mikhail Kashlev. 2009. “Rpb9 Subunit Controls Transcription Fidelity by Delaying NTP Sequestration in RNA Polymerase II.” *Journal of Biological Chemistry* 284(29):19601–12.

Wang, Jianming, Martina Muste Sadurni, and Marco Saponaro. 2023. “RNAPII Response to Transcription-Blocking DNA Lesions in Mammalian Cells.” *FEBS Journal* 290(18):4382–94.

Wang, Lijuan, Jie Chen, Susanna Larsson, Malcolm Dunlop, and Susan Farrington. 2022. “Identification of Novel Protein Biomarkers and Drug Targets for Colorectal Cancer by Integrating Human Plasma Proteome with Genome.” 1–17.

Wang, Mingqing, Wenbo Niu, Rong Hu, Yanjing Wang, Yangyang Liu, Lingyu Liu, Juan Zhong, Cha Zhang, Haiyan You, Jiaxing Zhang, Lu Lu, Lianbo Wei, and Wei Xiao. 2019. “POLR1D Promotes Colorectal Cancer Progression and Predicts Poor Prognosis of Patients.” (December 2018):735–48.

Wang, Tzu Hao, Cheng Yang Lee, Tzong Yi Lee, Hsien Da Huang, Justin Bo Kai Hsu, and Tzu Hao Chang. 2021. “Biomarker Identification through Multiomics Data Analysis of Prostate Cancer Prognostication Using a Deep Learning Model and Similarity Network Fusion.” *Cancers* 13(11).

Wang, Xuan, Zhengtong Lv, Haoran Xia, Xiaoxiao Guo, Jianye Wang, Jianlong Wang, and Ming Liu. 2023. “Biochemical Recurrence Related Metabolic Novel Signature Associates with Immunity and ADT Treatment Responses in Prostate Cancer.” *Cancer Medicine* 12(1):862–78.

Wassing, Isabel E., and Fumiko Esashi. 2021. “RAD51: Beyond the Break.” *Seminars in Cell and Developmental Biology* 113(June 2020):38–46.

Weterings, Eric, and David J. Chen. 2008. “The Endless Tale of Non-Homologous End-Joining.” *Cell Research* 18(1):114–24.

White, Robert J. 2004. “RNA Polymerase III Transcription and Cancer.” *Oncogene* 23(18):3208–16.

White, Robert J., Didier Trouchet, Klaus Martint, Stephen P. Jackson, Tony Kouzaridest, and D. T. France. 1996. “Retinoblastoma Protein.” 382(July):88–90.

Wild, Thomas, and Patrick Cramer. 2012. “Biogenesis of Multisubunit RNA Polymerases.” *Trends in Biochemical Sciences* 37(3):99–105.

Wilhelm, Therese, Maha Said, and Valeria Naim. 2020. “Dna Replication Stress and Chromosomal Instability: Dangerous Liaisons.” *Genes* 11(6):1–35.

Williamson, Josh, Ciara M. Hughes, George Burke, and Gareth W. Davison. 2020. “A Combined γ-H2AX and 53BP1 Approach to Determine the DNA Damage-Repair Response to Exercise in Hypoxia.” *Free Radical Biology and Medicine* 154(April):9–17.

Winter, Andrew G., George Sourvinos, Simon J. Allison, Kerrie Tosh, Pamela H. Scott, Demetrios A. Spandidos, and Robert J. White. 2000. “RNA Polymerase III Transcription Factor TFIIIC2 Is Overexpressed in Ovarian Tumors.” *Proceedings of the National Academy of Sciences of the United States of America* 97(23):12619–24.

Wissink, Erin M., Anniina Vihervaara, Nathaniel D. Tippens, and John T. Lis. 2019. “Nascent RNA Analyses: Tracking Transcription and Its Regulation.” *Nature Reviews Genetics* 20(12):705–23.

Woo, Hyun Goo, Eun Sung Park, Ju-seog Lee, Yun-han Lee, Tsuyoshi Ishikawa, Yoon Jun Kim, and Snorri S. Thorgeirsson. 2009. “Identification of Potential Driver Genes in Human Liver Carcinoma by Genomewide Screening.” (9):4059–66.

Woychik, Nancy A., and Michael Hampsey. 2002. “The RNA Polymerase II Machinery: Structure Illuminates Function.” *Cell* 108(4):453–63.

Xu, Jiangsheng, Yunhua Liu, Yujing Li, Hai Wang, Samantha Stewart, Kevin Van der Jeught, Pranay Agarwal, Yuntian Zhang, Sheng Liu, Gang Zhao, Jun Wan, Xiongbin Lu, and Xiaoming He. 2019. “Precise Targeting of POLR2A as a Therapeutic Strategy for Human

Triple Negative Breast Cancer." *Nature Nanotechnology* 14(4):388–97.

Xu, Jingfei, Noah Bradley, and Yuan He. 2023. "Structure and Function of the Apical PIKKs in Double-Strand Break Repair." *Current Opinion in Structural Biology* 82:102651.

Xu, Youwei, Carrie Bernecky, Chung Tien Lee, Kerstin C. Maier, Björn Schwalb, Dimitry Tegunov, Jürgen M. Plitzko, Henning Urlaub, and Patrick Cramer. 2017. "Architecture of the RNA Polymerase II-Paf1C-TFIIS Transcription Elongation Complex." *Nature Communications* 8.

Yamada, Yasutaka, Rika Nishikawa, Mayuko Kato, Atsushi Okato, Takayuki Arai, Satoko Kojima, Kazuto Yamazaki, Yukio Naya, Tomohiko Ichikawa, and Naohiko Seki. 2018. "Regulation of HMGB3 by Antitumor MiR-205-5p Inhibits Cancer Cell Aggressiveness and Is Involved in Prostate Cancer Pathogenesis." *Journal of Human Genetics* 63(2):195–205.

Yang, Fan, Haitao Liu, Junjie Zhao, Xingjie Ma, and Weibo Qi. 2020. "POLR1B Is Upregulated and Promotes Cell Proliferation in Non-Small Cell Lung Cancer." *Oncology Letters* 19(1):671–80.

Yang, Liliang, Kaizhen Wang, and Xian Chen. 2020. "Gene Expression and Regulatory Webwork of POLR2K in Bladder Carcinogenesis by Integrated Bioinformatics Approaches." 1–28.

Yang, Tianshu, and Xin Wang. 2021. "Prognostic Analysis of Patients with Breast Cancer Based on Tumor Mutational Burden , DNA Damage Repair Genes , and Immune Infiltration." 1–26.

Yang, Ying, Ranran Yan, Liwen Zhang, Xiangli Meng, and Wen Sun. 2020. "Primary Glioblastoma Transcriptome Data Analysis for Screening Survival-Related Genes." *Journal of Cellular Biochemistry* 121(2):1901–10.

Yang, Zhiyuan, Jasper H. N. Yik, Ruichuan Chen, Nanhai He, Kyoo Jang Moon, Keiko Ozato, and Qiang Zhou. 2005. "Recruitment of P-TEFb for Stimulation of Transcriptional Elongation by the Bromodomain Protein Brd4." *Molecular Cell* 19(4):535–45.

Yao, Fuwen, Yongqiang Zhan, Changzheng Li, Ying Lu, Jiao Chen, Jing Deng, Zijing Wu, Qi Li, Yi'an Song, Binhua Chen, Jinjun Chen, Kuifeng Tian, Zuhui Pu, Yong Ni, and Lisha Mou. 2022. "Single-Cell RNA Sequencing Reveals the Role of Phosphorylation-Related Genes in Hepatocellular Carcinoma Stem Cells." *Frontiers in Cell and Developmental Biology* 9(January):1–18.

Yao, Liangyu, Rong Cong, Chengjian Ji, Xiang Zhou, Jiaochen Luan, Xianghu Meng, and Ninghong Song. 2021. "RNA-Binding Proteins Play an Important Role in the Prognosis of Patients With Testicular Germ Cell Tumor." *Frontiers in Genetics* 12(March):1–21.

Ye, Jian, George Coulouris, Irena Zaretskaya, Ioana Cutcutache, Steve Rozen, and Thomas L. Madden. 2012. "Primer-BLAST: A Tool to Design Target-Specific Primers for Polymerase Chain Reaction." *BMC Bioinformatics* 13:134.

Yi, Li, Third Xiangya Hospital, Hubei Provincial Women, Children Hospital, and Read Full License. 2023. "Malignant Cell Related Gene Signature for Patients with Triple-Negative Breast Cancer Based on Integrated Analysis of Single Cell and Bulk RNA Sequencing."

Yonaha, Masatomo, Taku Chibazakura, Shigetaka Kitajima, and Yukio Yasukochi. 1995. “Cell Cycle-Dependent Regulation of RNA Polymerase II Basal Transcription Activity.” *Nucleic Acids Research* 23(20):4050–54.

Yong-Zheng Zhang, Amir Zeb and Lu-Feng Cheng* Department. 2022. “Exploring the Molecular Mechanism of Hepatitis Virus Inducing Hepatocellular Carcinoma by Microarray Data and Immune Infiltrates Analysis Yong-Zheng.” *Frontiers in Immunology*.

Yoo, Seung Soo, Mi Jeong Hong, Jang Hyuck Lee, Jin Eun Choi, Shin Yup Lee, Jaehee Lee, Seung Ick Cha, Chang Ho Kim, Yangki Seok, Eungbae Lee, Sukki Cho, Sanghoon Jheon, and Jae Yong Park. 2017. “Association between Polymorphisms in MicroRNA Target Sites and Survival in Early-Stage Non-Small Cell Lung Cancer.” *Thoracic Cancer* 8(6):682–86.

Yu, Qingfeng, Ying Xu, Haifeng Zhuang, Zhaoxing Wu, Lei Zhang, Jinfan Li, Linlin Yang, Bowen Wu, Ping Wang, Xuzhao Zhang, Xiaoxian Gan, Yun Liang, Shu Zheng, Xiaofang Yu, Ying Gu, and Rongzhen Xu. 2019. “Aberrant Activation of RPB1 Is Critical for Cell Overgrowth in Acute Myeloid Leukemia.” *Experimental Cell Research* 384(2):111653.

Yun, Maximina H., and Kevin Hiom. 2009. “CtIP-BRCA1 Modulates the Choice of DNA Double-Strand-Break Repair Pathway throughout the Cell Cycle.” *Nature* 459(7245):460–63.

Yunlei, Zhou, Chen Zhe, Lei Yan, Wang Pengcheng, Zheng Yanbo, Sun Le, and Liang Qianjin. 2013. “INMAP, a Novel Truncated Version of POLR3B, Represses AP-1 and P53 Transcriptional Activity.” *Molecular and Cellular Biochemistry* 374(1–2):81–89.

Zaborowska, Justyna, Sylvain Egloff, and Shona Murphy. 2016. “The Pol II CTD: New Twists in the Tail.” *Nature Structural and Molecular Biology* 23(9):771–77.

Zatzman, Matthew, Fabio Fuligni, Ryan Ripsman, Tannu Suwal, Federico Comitani, Lisa Monique Edward, Rob Denroche, Gun Ho Jang, Faiyaz Notta, Steven Gallinger, Saravana P. Selvanathan, Jeffrey A. Toretsky, Matthew D. Hellmann, Uri Tabori, Annie Huang, and Adam Shlien. 2022. “Widespread Hypertranscription in Aggressive Human Cancers.” *Science Advances* 8(47).

Zeman, Michelle K., and Karlene A. Cimprich. 2014. “Causes and Consequences of Replication Stress.” *Nature Cell Biology* 16(1):2–9.

Zeng, Fanli, Yu Hua, Xiaoqin Liu, Sijie Liu, Kejing Lao, a Ze Zhang, Daochun Kong. 2018. “Gpn2 and Rba50 Directly Participate in the Assembly of the Rpb3 Subcomplex in the Biogenesis of RNA Polymerase II.” 1–11.

Zhang, Bei, Zhiyao Ren, Hongmei Zheng, Meilan Lin, Guobing Chen, Oscar Junhong Luo, and Guodong Zhu. 2023. “CRISPR Activation Screening in a Mouse Model for Drivers of Hepatocellular Carcinoma Growth and Metastasis.” *IScience* 26(3):106099.

Zhang, Zhong Yi, Qing Zhi Yao, Hui Yong Liu, Qiao Nan Guo, Peng Jun Qiu, Jian Peng Chen, and Jian Qing Lin. 2020. “Metabolic Reprogramming-Associated Genes Predict Overall Survival for Rectal Cancer.” *Journal of Cellular and Molecular Medicine* 24(10):5842–49.

Zhao, Jian, Xiaofeng Song, Tianyi Xu, Qichang Yang, Jingjing Liu, Bin Jiang, and Jing Wu. 2021. “Identification of Potential Prognostic Competing Triplets in High-Grade Serous Ovarian

Cancer." *Frontiers in Genetics* 11(January):1–11.

Zhou, Danyang, Xing Li, Hengyu Zhao, Banghao Sun, Anqi Liu, Xue Han, Zhongqi Cui, and Lijie Yuan. 2018. "Combining Multi-Dimensional Data to Identify a Key Signature (Gene and MiRNA) of Cisplatin-Resistant Gastric Cancer." *Journal of Cellular Biochemistry* 119(8):6997–7008.

Zhou, Kai, Songkuan Zhuang, Fulong Liu, Yanheng Chen, You Li, Shihui Wang, Yuxuan Li, Huixin Wen, Xiaohua Lin, Jie Wang, Yue Huang, Cailing He, Nan Xu, Zongshu Li, Lang Xu, Zixuan Zhang, Lin Feng Chen, Ruichuan Chen, and Min Liu. 2022. "Disrupting the Cdk9/Cyclin T1 Heterodimer of 7SK SnRNP for the Brd4 and AFF1/4 Guided Reconstitution of Active P-TEFb." *Nucleic Acids Research* 50(2):750–62.

Zhou, Yi, Ji Hoon Lee, Wenxia Jiang, Jennie L. Crowe, Shan Zha, and Tanya T. Paull. 2017. "Regulation of the DNA Damage Response by DNA-PKcs Inhibitory Phosphorylation of ATM." *Molecular Cell* 65(1):91–104.

Zou, Lee, and Stephen J. Elledge. 2003. "Sensing DNA Damage through ATRIP Recognition of RPA-SsDNA Complexes." *Science* 300(5625):1542–48.

Zou, Xueqing, Gene Ching Chiek Koh, Arjun Scott Nanda, Andrea Degasperi, Katie Urgo, Theodoros I. Roumeliotis, Chukwuma A. Agu, Cherif Badja, Sophie Momen, Jamie Young, Tauanne Dias Amarante, Lucy Side, Glen Brice, Vanesa Perez-Alonso, Daniel Rueda, Celine Gomez, Wendy Bushell, Rebecca Harris, Jyoti S. Choudhary, John C. Ambrose, Prabhu Arumugam, Emma L. Baple, Marta Bleda, Freya Boardman-Pretty, Jeanne M. Boissiere, Christopher R. Boustred, Helen Brittain, Mark J. Caulfield, Georgia C. Chan, Clare E. H. Craig, Louise C. Daugherty, Anna de Burca, Andrew Devereau, Greg Elgar, Rebecca E. Foulger, Tom Fowler, Pedro Furió-Tarí, Adam Giess, Joanne M. Hackett, Dina Halai, Angela Hamblin, Shirley Henderson, James E. Holman, Tim J. P. Hubbard, Kristina ibáñez, Rob Jackson, Louise J. Jones, Dalia Kasperaviciute, Melis Kayikci, Athanasios Kousathanas, Lea Lahnstein, Kay Lawson, Sarah E. A. Leigh, Ivonne U. S. Leong, Javier F. Lopez, Fiona Maleady-Crowe, Joanne Mason, Ellen M. McDonagh, Loukas Moutsianas, Michael Mueller, Nirupa Murugaesu, Anna C. Need, Pter O'Donovan, Chris A. Odhams, Andrea Orioli, Christine Patch, Mariana Buongermino Pereira, Daniel Perez-Gil, Dimitris Polychronopoulos, John Pullinger, Tahrima Rahim, Augusto Rendon, Pablo Riesgo-Ferreiro, Tim Rogers, Mina Ryten, Kevin Savage, Kushmita Sawant, Richard H. Scott, Afshan Siddiq, Alexander Sieghart, Damian Smedley, Katherine R. Smith, Samuel C. Smith, Alona Sosinsky, William Spooner, Helen E. Stevens, Alexander Stuckey, Razvan Sultana, Mélanie Tanguy, Ellen R. A. Thomas, Simon R. Thompson, Carolyn Tregidgo, Arianna Tucci, Emma Walsh, Sarah A. Watters, Matthew J. Welland, Eleanor Williams, Katarzyna Witkowska, Suzanne M. Wood, Magdalena Zarowiecki, Josef Jiricny, William C. Skarnes, and Serena Nik-Zainal. 2021. "A Systematic CRISPR Screen Defines Mutational Mechanisms Underpinning Signatures Caused by Replication Errors and Endogenous DNA Damage." *Nature Cancer* 2(6):643–57.