Employing Polymerase Usage Sequencing to better understand fundamental aspects of eukaryotic replication

Robert Zach

Submitted for the degree of Doctor of Philosophy

University of Sussex

July 2021

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature: Robert Zach

Date: 09/08/2021

Acknowledgements

In few words, I would like to express my gratitude to people, who provided me with guidance and contributed not only to my education, but also personal growth. First of all, I thank Tony Carr who decided to hire me, and thus initiated my journey into the world of replication and genome stability. Throughout the course of my PhD studies, Tony has allowed me to follow my own ideas and make my own mistakes, from which I could learn. Whenever the value of academic line of work became a question, perhaps without even realising, Tony set my mind at ease by providing me with valuable and vibrant conversations on various, mostly scientific, subjects. In a wider perspective, Tony motivated me to develop independent approach and sense of responsibility for undertaken actions. To some, such virtues might seem marginal, but I realise that contemporary world can be a tough place and 'survival' within it often requires stable footing and peace of mind. I do believe that, after spending four years at the Genome Damage and Stability Centre (GDSC), I'm ready for whatever is coming next.

Undertaking a PhD course can be extremely rewarding as well as frustrating. I feel privileged to be able to say that I truly enjoyed the vast majority of my PhD. For that, I'm not only thankful to my boss, but also other colleagues and friends from GDSC. Namely, I thank Karel Naiman for his endless enthusiasm, support and valuable discussions, which often navigated my work the right direction. I thank Adam Watson for teaching me how to clone a plasmid, providing me with technical advice during hard times of troubleshooting and, perhaps most importantly, for being a genuine character. Further, I thank Joanna Fernandez for becoming a new pillar of the lab; Alice Budden for being an excellent colleague, friend and laboratory technician; Owen Wells for his human humour; David Buist for being David Buist; and all other present and past members of the Carr lab for making my PhD smooth, enjoyable and, more or less, stress-free. Thank you.

Thesis summary

DNA replication represents an essential mechanism which ensures short-term survival of all organisms as well as propagation of life in general. In this thesis, we explore fundamental aspects of eukaryotic replication and develop techniques which provide better performance and allow comprehensive analysis of understudied phenomena. In chapter 1, we summarise our attempts to improve standard polymerase usage sequencing protocol. In chapter 2, we test the hypothesis that increased Polδ levels interfere with canonical leading strand replication. In chapter 3, we explore replication dynamics in mutants depleted of senataxin RNA/DNA helicases Sen1^{+} and Dbl8^{+}. In chapter 4, we present a system to study replication of induced heterochromatinised domains. In chapter 5, we characterise basic properties of small Polδ subunit Cdm1^{+}.

Thesis organisation

The Presented work composes of 8 distinct sections including introduction, methods, chapters 1-5 and conclusions. The introduction section introduces general concepts relevant to the main theme of the thesis. The methods section provides comprehensive description of procedures, reagents and biological material used in presented experiments. Chapters 1-5 cover particular research projects developed in the lab. Each chapter provides a specific theoretical background together with description and discussion of acquired results. The conclusion section summarises presented work.

Table of Contents

List of abbreviations

5hmCs	5-hydroxymethyl-cytosines
9-1-1	Rad9–Rad1–Hus1 complex
ahTET	Anhydrotetracycline
ARS	Autonomously replicating sequence
ATM	Ataxia telangiectasia mutated
ATR	Ataxia telangiectasia and Rad3-related
BIR	Break induced replication
BM	Bleomycin
Bp	Base pair
BrdU	Bromodeoxyuridine
CLRC	$Clr4^{+}$ complex
CMG	Cdc45-MCM-GINS complex
Cpc	Copies per cell
CPT	Camptothecin
CTD	C-terminal domain
CUT	Cryptic unstable transcript
DDR	DNA damage repair
DNA-PK	DNA-dependent protein kinase
dNTP	Deoxyribonucleotide triphosphate
DRIPc-Seq	DNA:RNA immunoprecipitation followed by sequencing
dsRNA	Double-stranded RNA
EdU	5-Ethynyl-2'-deoxyuridine
EMM	Edinburgh minimal medium
FGF2	Fibroblast growth factor-2
FPC	Fork protection complex
G4	G-quadruplex
GFP	Green fluorescent protein
GLOE-Seq	Genome-wide ligation of 3’-OH ends followed by sequencing
GO	Gene ontology
HA	Hemagglutinin
HDAC	Histone deacetylase
HMT	Histone methyltransferase
HoRReR	Homologous recombination-restarted replication
HPLC	High performance liquid chromatography
HR	Homologous recombination
HU	Hydroxyurea
HydEn-Seq	Hydrolytic end sequencing
ICL	Inter-strand crosslink
IDT	Integrated DNA Technologies
Ini-Seq	Initiation site sequencing
MCM	Mcm2-7 complex
MMS	Methyl methanesulfonate
MRN	Mre11-Rad50-Nbs1 complex
Net-Seq	Native elongation transcript sequencing
NHEJ	Non-homologous end joining

non-B	Non-canonical DNA structure
NSCLC	Non-small cell lung cancer
NTC	No template control
OD	Optical density
OF	Okazaki fragment
ORC	Origin recognition complex
Ori	Origin of replication
Ori_{Eff}	Origin efficiency
PARP	Poly (ADP-ribose) polymerase
PCNA	Proliferating cell nuclear antigen
PDC	Protein-DNA crosslink
PIP	PCNA interacting protein
Pol	Polymerase
Poly-A	Poly-adenylation
PTM	Post translational modification
Pre-LC	Pre-loading complex
Pre-RC	Pre-replication complex
Pu-Seq	Polymerase usage sequencing
qPCR	Quantitative PCR
QuiLT	Quick Little Tool
RER	Ribonucleotide excision repair
RFB	Replication fork barrier
RITS	RNA-induced initiation of transcriptional gene silencing
RMC	Cre-recombination mediated cassette exchange
RNA-Seq	RNA sequencing
rNMP	Ribonucleotide monophosphate
RPA	Replication protein A
RRM	RNA recognition motif
rSAP	Shrimp alkaline phosphatase
RT	Replication timing; Reverse transcription
RTC	replication transcription collision/conflict
S-CDK	S phase specific CDK kinase
SAM	S-adenyl methionine
SCLC	Small cell lung cancer
siRNA	Short interfering RNA
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNAs
SNS-Seq	Short nascent strand sequencing
ssDNA	Single stranded DNA
StrAv	Streptavidin
TBSt	Tris-buffered saline with tween
tetO	Tetracycline operator
$TetR^{Off}$	"OFF" variant of tetracycline repressor
TLS	Trans-lesion synthesis
Top2	Topoisomerase 2
WT	Wild-type
YES	Yeast extract with supplements

List of figures

Introduction

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Introduction

Genetic information stored in the DNA sequence represents a major determinant of every organism's anatomy and function. Copying and passing on this information constitutes the basis of reproduction, embryonal development and regeneration of damaged body parts, the processes which are indispensable for the survival of individual organisms and, ultimately, propagation of life itself.

The processes involving DNA synthesis can be divided into two major groups. 1) replicative DNA synthesis, also known as DNA replication, facilitates duplication of genomic DNA, two copies of which are subsequently distributed between the two daughter cells. 2) Repair DNA synthesis, a complex system of molecular machines ensuring suppression and repair of diverse DNA lesions, mismanagement of which drives genome instability and consequential cellular and organismal pathologies.

Although DNA damage response (DDR) is not entirely omitted, with respect to the main theme of the thesis, the following text emphasises mechanistic and regulatory details of eukaryotic DNA replication.

Organisation of eukaryotic genomes

In the most simplistic sense, eukaryotic genomes can be regarded as species-specific collections of linear double-stranded DNA molecules, commonly referred to as chromosomes. Across a wide range of eukaryotic organisms, genomic sizes and compositions display considerable variability. For example, while somatic human cells carry 23 pairs of chromosomes (Gardiner, 1995; Lander *et al*., 2001), diploid females of primitive ant species *Myrmecia pilosula* feature only a single chromosome pair (Crosland & Crozier, 1986) and the genome of the blue butterfly species *Polyommatus atlanticus* is composed of as many as 224-226 pairs of chromosomes (Lukhtanov, 2015). Sizes of eukaryotic genomes range from millions of base pairs (bp), as documented in the protozoan parasite *Encephalitozoon intestinalis* (2.3×10^6 bp per genome) (Corradi, *et al.*, 2010), to several billions of bp, as seen in human cells (6×10^9 bp per genome) (Lander *et al.*, 2001).

In eukaryotes, numbers of protein-coding genes do not scale with respective genomic sizes. For instance, while the human genome (6×10^9 bp) encodes ca. 20,000

protein-coding genes, considerably smaller genome of *Arabidopsis thaliana* (1.25×10^8 bp) contains ca. 25,000 protein-coding genes (Arabidopsis Genome Initiative, 2000).

While some species feature highly efficient gene-rich genomes, chromosomes in some organisms contain vast amounts of non-coding DNA. For instance, in canonical cell biology models *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*, protein-coding genes occupy 60.2% and 70.1% of the genome, respectively (Engel *et al*., 2014; Goffeau *et al*., 1996; Wood *et al*., 2002). In stark contrast, protein-coding genes represent only ca. 1.5% of the human genomic DNA (Lander *et al*., 2001).

Naturally, genomic complexity goes far beyond the number of protein-coding loci and is further expanded by employment of alternative splicing and intricate well-tuned regulatory systems unique to each organism.

Despite substantial differences in genome architectures, different eukaryotic species employ highly conserved mechanisms of DNA replication. The upcoming text delineates fundamental aspects of eukaryotic replication and accompanying regulatory systems.

Origins of replication

The process of replication initiates at distinct genomic locations commonly referred to as origins of replication. Unlike replication of characteristically small bacterial genomes, which features activation of a singular origin, duplication of large eukaryotic chromosomes requires numerous spatially and timely separated initiation events.

The vast body of evidence indicates that replication origins in distinct eukaryotic species are determined by different mechanisms. For example, in budding yeast *S. cerevisiae*, most origins of replication are strictly defined by a DNA sequence motif, also known as autonomously replicating sequence (ARS) (Marahrens & Stillman, 1992). For long it had been widely assumed that, in *S. cerevisiae*, each activated origin gives rise to two replication forks, each of which initiates DNA synthesis at a particular ARS site with a single bp accuracy (Garbacz *et al*., 2018). Interestingly, however, it was recently proposed that a subset of non-canonical origins, which represent ca. 5-10% of all initiation sites of a single replication event, operate independently of the ARS motif in *S. cerevisiae* (Müller *et al*., 2019; D. Wang & Gao, 2019).

In fission yeast *S. pombe*, replication initiation sites do not exhibit any apparent consensus sequence, but colocalise with AT-rich intergenic regions (Dai, Chuang, & Kelly, 2005). In *S. pombe*, replication machineries established at slightly loosely defined origins initiate DNA synthesis within ca. 500bp regions (Daigaku *et al*., 2015).

In human cells, genomic positions of replication origins are extremely loosely defined, and are often referred to as replication initiation zones (Petryk *et al*., 2016; Wang *et al*., 2021). Human origins of replication are often situated in close proximity to transcription start sites and tend to display high GC-content (Petryk *et al*., 2016; Wang *et al*., 2021). Several independent studies had claimed that human replication initiation sites are prone to form G-quadruplex (G4) structures, (Cayrou *et al*., 2011; Langley *et al.*, 2016); however, such claims were later challenged by in-depth computational analysis, which indicates that seeming enrichment in G4-forming motifs represents a consequence of underlying bias towards GC-rich sequence composition (Wang *et al*., 2021). In addition to DNA sequence-based characteristics, human origins have been associated with distinct histone modifications such as enhancer-specific H3K4me1 and H3K27ac (Marks *et al.*, 2017; Wang *et al*., 2021).

Origin licencing

Origin licencing constitutes a sequence of events ultimately responsible for the loading of ring-shaped hetero-hexameric helicases Mcm2-7 (MCM), which, upon activation in S phase, facilitate unwinding of the DNA duplex (Abid Ali *et al*., 2017; Bleichert *et al.*, 2017; Douglas *et al.*, 2018).

Origin licencing is conditioned by low CDK activity, and thus is restricted to late mitosis and G1 phase (Nguyen *et al.*, 2001; Remus *et al.*, 2005). Loading of MCM helicases require coordinated action of origin recognition complex (ORC), a hexameric assembly of Orc1-6 subunits, and two co-loading factors, Cdc6 and Cdt1 (Bleichert *et al*., 2017; Guerrero-Puigdevall *et al.*, 2021; Randell *et al.*, 2006; Ticau *et al.*, 2015). In late M and G1 phases, each origin is recognised and bound by the ORC complex associated with Cdc6. Then, DNA-bound ORC-Cdc6 cooperates with Cdt1 to sequentially load two inactive MCM helicases (Bleichert et al., 2017; Coster & Diffley, 2017; Guerrero-Puigdevall et al., 2021; Ticau et al., 2015). Loading of ring-shaped MCM helicases onto DNA exploits a conditional molecular gate formed between

Mcm2 and Mcm5 subunits (Samel *et al*., 2014). While the interaction of ORC with chromatin is stable, Cdc6 and Cdt1 can exhibit dynamic behaviour, operating in cycles of recruitment and subsequent dissociation (Ticau *et al*., 2015). The final product of origin licencing, also known as pre-replication complex (pre-RC), is a stable assembly of two head-to-head oriented MCM helicases, each encircling double-stranded origin DNA.

In order to avoid re-replication, eukaryotes employ a range of mechanisms to supress unscheduled origin licencing. In general, suppression of re-licencing operates via inhibition of crucial origin licencing factors. In *S. cerevisiae*, increased activity of CDK kinases at the end of G1 phase drives proteolytic degradation of Cdc6 and cytoplasmic sequestration of Cdt1 and components of MCM helicases (Nguyen *et al*., 2001; Nguyen *et al.*, 2000; Tanaka & Diffley, 2002). Additionally, in *S. cerevisiae*, it has been shown that Cdc6 and Orc6 functions are inhibited by direct interaction with CDKs (Mimura *et al.*, 2004; Wilmes *et al*., 2004). In *S. pombe* and more complex eukaryotes, Cdt1 and Cdc6 (Cdc18 in *S. pombe*) are phosphorylated by CDKs and targeted for degradation by a Cullin-based E3 ubiquitin ligase system (Arias & Walter, 2005, 2006; Gopalakrishnan *et al*., 2001; Hu & Xiong, 2006; Jallepalli *et al.*, 1997; Walter *et al*., 2016). In human cells, proteolysis of Orc1 also contributes to prevention of re-replication (Méndez *et al*., 2002). In addition to proteolytic degradation, in higher eukaryotes, Cdt1 is inhibited by direct interaction with periodically expressed geminin (Lee *et al*., 2004; Tada *et al.*, 2001; Wohlschlegel *et al*., 2000).

Origin firing

Origin activation, frequently referred to as origin firing, constitutes a sequence of events ultimately leading to activation of replicative MCM helicases deposited on origins of replication. In each S phase, only a subset of licenced origins is activated (DePamphilis, 1993; Sugimoto *et al.*, 2018; Taylor, 1977). Remaining dormant origins function as a backup, being activated to compensate for paused, stalled or damaged replication forks (Blow & Ge, 2009; Brambati *et al*., 2018; Santocanale, 1999; Shima & Pederson, 2017). Underlying steps of origin firing are strictly regulated by the Dbf4-dependent kinase (DDK) and the S phase specific CDK kinase (S-CDK) (Araki, 2010;

Heller *et al*., 2011; Masai *et al*., 2006; Muramatsu *et al.*, 2010; Randell *et al*., 2010; Sheu & Stillman, 2010).

The process of origin activation is initiated by transient recruitment of DDK to an origin of replication, where it facilitates phosphorylation of Mcm4 and Mcm6 subunits of licenced MCM helicases (Masai *et al*., 2006; Randell *et al*., 2010; Sheu & Stillman, 2010). DDK-mediated phosphorylation of the Mcm2-7 complex suppresses the inhibitory function of the N-terminal domain of Mcm4 and promotes downstream recruitment of additional components of replication machinery (Masai *et al*., 2006; Sheu & Stillman, 2010). In *S. cerevisiae*, auxiliary factors Sld2, Sld3, Sld7 and Dpb11 together with the leading strand replicase Polε promote association of each chromatin-bound MCM hexamer with Cdc45 and the GINS heterotetramer composed of Sld5, Psf1, Psf2 and Psf3 subunits (Fragkos *et al.*, 2015; Kamada *et al.*, 2007). Recruitment of Cdc45 and GINS requires S-CDK activity, which modulates complex-forming properties of Sld2 and Sld3 (Araki, 2010; Heller *et al*., 2011; Muramatsu *et al*., 2010). Interestingly, it has been reported that Polε, GINS, Sld2, and Dpb11 are incorporated in a form of four-subunit pre-loading complex (pre-LC) (Muramatsu *et al*., 2010). Each assembly of the MCM hexamer, Cdc45 and GINS represents the *bona fide* 11-subunit replicative helicase, CMG (Cdc45-MCM-GINS), which, upon activation, unwinds DNA duplex and functions as a motor of replication progression (Costa *et al*., 2011; Ilves *et al.*, 2010).

In the final stages of origin activation, each CMG helicase extrudes one of the DNA strands encircled by the MCM hexamer. Extrusion of single-stranded DNA is mediated by the molecular gate formed between Mcm2 and Mcm5 subunits and requires assistance of Mcm10 which stabilises CMG-DNA interaction during the extrusion process (Douglas *et al*., 2018; Kanke *et al.*, 2012; Perez-Arnaiz *et al.*, 2016; Quan *et al*., 2015; van Deursen *et al.*, 2012; Wasserman *et al.*, 2019). Subsequently, at each origin, two head-to-head oriented CMG helicases, each of which encircles the opposite DNA strand, pass each other and form two bi-directional replication forks (Costa *et al*., 2011; Douglas *et al*., 2018). CMG activation and foundation of the canonical replication bubble is accompanied by recruitment of Polα-primase and the main lagging strand replicase Polδ.

Replication timing

Origins of replication across the genome are not activated simultaneously, but fire in a pre-defined temporal order. Consequently, distinct genomic regions are replicated at different times during S phase. Numerous studies employing different experimental systems indicate that replication timing (RT) is, at least in part, regulated by chromatin modifying enzymes. In general, more accessible euchromatin and condensed heterochromatin are replicated in early and late stages of S phase, respectively (Goren *et al.*, 2008; Knott *et al.*, 2009; Mantiero *et al.*, 2011; Shoaib *et al.*, 2018). It has been argued that compact higher-order fold of heterochromatin constrains the access of factors involved in licencing and activation of origins. It has been demonstrated that decompaction of chromatin by siRNA-mediated depletion of SET8 histone methyltransferase results in overloading of Orc1 and components of MCM helicase onto origins of replication in human cells (Shoaib *et al.*, 2018). Accordingly, in *S. cerevisiae*, it has been suggested that early firing origins are licenced with multiple MCM hexamers, and that such increased MCM-occupation drives early origin activation (Das *et al.*, 2015). Several studies further indicate that disruption of histone deacetylase Rpd3 or induced binding of Gcn5 histone acetyltransferase onto chromatin result in untimely activation of dormant and late firing origins, most likely due to unscheduled recruitment of specific origin firing factors in *S. cerevisiae* (Knott *et al.*, 2009; Mantiero *et al.*, 2011; Vogelauer *et al.*, 2002).

Firing of all licenced origins at the same time is also suppressed by the limiting availability of essential firing regulators. In *S. cerevisiae*, limiting intracellular concentrations of origin firing factors Sld2, Sld3, Sld7, Dpb11, Cdc45 and the regulatory subunit of DDK kinase Dbf4 restrict the number of simultaneous replication initiation events (Lynch *et al.*, 2019; Mantiero *et al.*, 2011; Tanaka *et al.*, 2011). Limited availability of origin firing factors has been also utilised as an important parameter in mathematical models of replication (Lygeros *et al.*, 2008).

RT profiles are further refined by locus-specific actions of additional regulators. For instance, it has been postulated that forkhead transcription factors Fkh1 and Fkh2 contribute to establishment of spatially clustered early-firing origin assemblies which efficiently recruit limiting replication initiation regulators in *S. cerevisiae* (Knott *et al.*,

2012). For instance, it was shown that Fkh1/2 recruits DDK kinase to non-centromeric early firing origins (Fang *et al*., 2017). In *S. Pombe*, heterochromatinised pericentromeric regions and mating type locus are replicated early in S phase due to specific Swi6-mediated binding of DDK kinase (Hayashi *et al.*, 2009). Similarly, early replication of centromere-proximal regions is ensured by Ctf19-dependent recruitment of DDK kinase to kinetochores in *S. cerevisiae* (Natsume *et al*., 2013). Interestingly, in both yeast species, late replicating subtelomeric heterochromatin and other genomic regions are believed to be shielded from DDK activity by Rif1-facilitated recruitment of PP1 phosphatase (Davé *et al.*, 2014; Hayano *et al*., 2012). Accordingly, the involvement of Rif1 in the establishment of RT program has been also demonstrated in higher eukaryotes (Cornacchia *et al*., 2012; Yamazaki *et al*., 2012).

Even though the evolutionary advantage of RT maintenance is not completely understood, several lines of evidence indicate that temporal separation of individual origin firing events contributes to genome stability by several mechanisms. For instance, in a divergent set of yeast species, early replication of genes encoding histone proteins ensures adequate expression levels and supply of histone proteins during the course of S phase (Müller & Nieduszynski, 2017). In *S. cerevisiae*, unscheduled activation of additional origins in early S phase results in depletion of deoxyribonucleotides (dNTPs) and replication stress (Mantiero *et al*., 2011). Additionally, impaired RT correlates with cellular pathologies of unknown origin, which are unrelated to depletion of dNTPs (Mantiero *et al*., 2011). Abnormal RT profiles were further correlated with acute lymphoblastic leukaemia (Ryba *et al*., 2012; Sasaki *et al*., 2017) and the Hutchinson-Gilford syndrome (Rivera-Mulia *et al*., 2017). A causal link between disrupted RT and observed pathologies, however, has not been established. Interestingly, ca. 30.5% of the human genome changes RT during differentiation, indicating a role of RT in cell fate specification (Rivera-Mulia *et al*., 2015; Wilson *et al.*, 2016).

General mechanisms of DNA replication

Replication fork progression

Eukaryotic DNA replication is carried out in a semi-conservative manner, using the two strands of the parental DNA molecule as templates, as depicted in Figure I.1

(Burgers & Kunkel, 2017; Leman & Noguchi, 2013; Watson & Crick, 1953). Each template directs the synthesis of its own complement in 5'-3' direction. Upon origin activation, two replication machineries, also known as replication forks or replisomes, are set off (Burgers & Kunkel, 2017). In the most fundamental sense, each replisome incorporates actions of three distinct polymerase (Pol) complexes (Polε, Polδ and Polα-primase) and the CMG helicase (Figure I.1) (Burgers & Kunkel, 2017). While Polα-primase represents an error-prone enzyme, the remaining two replicases, Polε and Polδ, exhibit proofreading 3'-5' exonucleolytic activity, which dramatically increases their fidelity (Bębenek & Ziuzia-Graczyk, 2018; Kunkel *et al.*, 1989). CMG helicases unwind the DNA duplexes in 5'-3' direction and generate transient stretches of ssDNA, which are coated by replication protein A (RPA). Binding of RPA to ssDNA moieties prevents untimely nucleolytic cleavage and supresses formation of potentially deleterious secondary structures (Dueva & Iliakis, 2020). In addition to its protective role, RPA stimulates activities of Polα-primase and Polδ (Braun *et al.*, 1997; Dueva & Iliakis, 2020; Tsurimoto & Stillman, 1989).

Replication forks activated at the same origin of replication are moving in opposite directions and each fork synthesizes DNA strands complementary to both DNA templates. As a consequence of the antiparallel orientation of the two parental strands, one strand, often referred to as the leading strand, is replicated continuously, whereas the other strand, also known as the lagging strand, is copied in a discontinuous manner (Figure I.1) (Burgers & Kunkel, 2017). Continuous leading strand synthesis, which progresses in the same direction as the CMG helicase, is mainly carried out by Polε. The discontinuous lagging strand synthesis, which is carried out in the direction opposite to the CMG movement, is facilitated by Polα-primase and Polδ (Clausen *et al*., 2015; Daigaku *et al*., 2015; Garbacz *et al*., 2018; Miyabe *et al.*, 2011; Zhou *et al.*, 2019).

While highly processive Polε physically interacts with the CMG helicase and specialises in producing long tracks of DNA, Polδ associates with the DNA template only transiently and synthesises short (ca. 200 bp) stretches of DNA, also known as Okazaki fragments (OFs) (Burgers & Kunkel, 2017; Okazaki *et al.*, 1968; Stodola & Burgers, 2017). Polε and Polδ interact with proliferating cell nuclear antigen (PCNA), which function as a DNA clamp and promotes the catalytic potential and processivity

of both DNA replicases (Chilkova *et al*., 2007; Lancey *et al*., 2020; Mondol *et al.*, 2019). While interaction with PCNA is crucial for adequate Polδ function, it only provides minor improvement of Polε activity (Chilkova *et al*., 2007).

Polδ-mediated synthesis of OFs is initiated from short (ca. 20-30 bp) RNA/DNA primers produced by Polα-primase, an inherent component of the replisome, which recognises RPA-coated ssDNA generated by CMG helicases (Figure I.1) (Burgers & Kunkel, 2017; Stodola & Burgers, 2017). In a model situation, Polδ extending the RNA/DNA primer eventually reaches the preceding OF and displaces its 5' end. In most cases, Polδ facilitates nick translation synthesis, only displacing 1 nucleotide at a time. Single nucleotide displacements are also promoted by Polδ idling, which occurs when two or more nucleotides are displaced. Single nucleotide flaps are recognised and digested by the endonuclease Fen1 (Balakrishnan & Bambara, 2013; Lin et al., 2013; Stodola & Burgers, 2016). It has been also suggested that, in addition to Fen1, non-canonical situations involving long 5' end flaps require activity of the nuclease/helicase Dna2 (Rossi *et al.*, 2018). However, this proposition still requires additional experimental validation. It has been estimated that, despite nucleolytic processing of 5' termini of OFs, ca. 1.5% of genomic DNA is produced by low fidelity Polα-primase (Reijns *et al*., 2015). Interestingly, several lines of evidence indicate that potential accumulation of errors generated by Polα-primase is suppressed by proofreading activity of Polδ (Pavlov *et al*., 2006; Perrino & Loeb, 1990). Following the 5' end processing, two consecutive OFs are ligated by the DNA ligase Lig1 (Howes & Tomkinson, 2012).

Even though the precise mechanism of replisome coordination has not been elucidated, studies using *S. cerevisiae* as a model system indicate that the main replisome components and their auxiliary factors are, at least in part, organised around the Ctf4 trimer, which functions as a multifaceted interaction hub (Simon *et al*., 2014; Villa *et al*., 2016). Accordingly, the Ctf4 orthologues AND-1 and Mcl1^{+} contribute to replisome assembly and organisation in human and *S. pombe*, respectively (Figure I.1) (Im *et al*., 2009; Tsutsui *et al*., 2005).

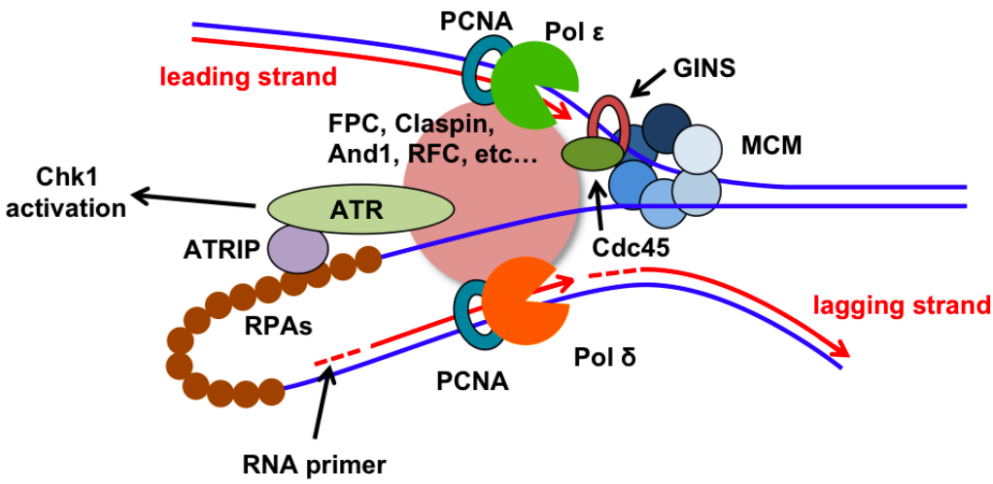

Figure I.1 – Schematic representation of the replisome complex. Depiction of the canonical replisome organisation including the core factors involved in the leading strand and the lagging strand synthesis, as well as replisome coordination and regulation. Figure was adopted from the review article by A. Leman and E. Noguchi (Leman & Noguchi, 2013).

Replication initiation

Establishment of functional replication machinery takes place shortly after origin activation. Originally, it was believed that origin firing is accompanied by deposition of four RNA/DNA primers by Polα-primase. According to this model, each of the two established replication forks utilise two individual RNA/DNA primers to initiate leading and lagging strand synthesis. Recent *in vitro* as well as *in vivo* studies utilising *S. cerevisiae* as a model system, however, challenged such notion and established an alternative mechanism featuring only two priming events (Aria & Yeeles, 2018; Garbacz *et al.*, 2018; Zhou *et al.*, 2019). According to this model, two initiator primers are deposited on lagging strand templates of the two divergent replication forks. Each initiator primer is recognised by Polδ which facilitates primer extension until it reaches the co-directionally moving CMG helicase with associated Polε. Highly processive Polε then takes over the extension of the 3' end and establishes continuous leading strand replication. As replisomes progress, lagging strand replication operates in a dynamic manner, as described in the previous text. As an inevitable consequence of this mechanism, short stretches of both DNA strands

constituting origins of replication are synthesised by Polδ (Aria & Yeeles, 2018; Garbacz *et al*., 2018; Zhou *et al*., 2019).

Replication termination

Replication termination occurs when the two replication forks moving in the opposite direction converge. It is widely accepted that most replication termination sites are sequence independent, and their genomic positions purely depend on activation timing and firing probabilities of adjacent origins of replication (Clausen *et al*., 2015; Daigaku *et al*., 2015; Dewar & Walter, 2017; Petryk *et al*., 2016). In contradiction, it has been also suggested that, at least in *S. cerevisiae*, replication termination sites form around natural impediments of replisome progression, such as centromeric sequences and actively transcribed genes (Fachinetti *et al*., 2010).

The exact mechanism of replication termination in eukaryotes has not yet been collectively agreed on; however, a prominent model emerged from a study employing cell-free *Xenopus laevis* egg extracts and plasmid-based system allowing transient block and synchronous release of two converging replication forks (Dewar *et al.*, 2015). Data presented by J. Dewar and colleagues indicate that converging CMG helicases, each of which encircles the leading strand template (Costa *et al*., 2011; Douglas *et al*., 2018), pass each other and promote Polε-mediated elongation of the 3’ end of the leading strand until they reach the 5’ end of the converging lagging strand. Newly synthesised leading and lagging strands of the same orientation are subsequently ligated (Dewar *et al*., 2015). It has been proposed that, after reaching the 5’ end of the converging lagging strand, CMG helicases keep moving along the dsDNA, which represents a substrate that can be accommodated by the MCM helicase (Costa *et al*., 2011; Dewar *et al*., 2015; Kaplan *et al.*, 2003). It is hypothesised that such molecular dynamics triggers K48-linked polyubiquitination of Mcm7, which function as a signal for p97/Cdc48-mediated dissociation of the CMG helicase in *X. laevis* and *S. cerevisiae* (Dewar *et al*., 2015; Maric *et al.*, 2014; Moreno *et al.*, 2014).

Convergence of two replisomes is accompanied by accumulation of torsional stress which prevents dissolution of the last stretch of unreplicated DNA. It is believed that obstructing supercoiling is relieved by rotation of converging replication machineries. Rotational movement introduces intertwined (catenated) DNA

structures behind both forks (Dewar *et al*., 2015). The significant body of evidence indicates that decatenation is carried out by topoisomerase 2 (Top2) enzymes, whose replication termination-specific function is indispensable for unperturbed segregation of chromosomes during mitosis (Baxter & Diffley, 2008; DiNardo *et al.*, 1984; Fachinetti *et al*., 2010).

Versatile and modular functions of Polδ

Apart from being the main lagging strand replicase, Polδ has been implicated in DNA repair processes such as homologous recombination-restarted replication (HoRReR), break induced replication (BIR), post-repair gap filling and bypass of oxidative DNA lesions (Donnianni *et al*., 2019; Guilliam & Yeeles, 2021, 2020b; Miyabe *et al*., 2015; Naiman *et al*., 2021). Additionally, experimental evidence indicates various situations, where Polδ mediates leading strand synthesis in unperturbed *in vivo* and *in vitro* systems (Guilliam & Yeeles, 2020a). For instance, in *S. cerevisiae*, Polδ establishes leading strand synthesis at origins of replication and, based on preliminary evidence, Polδ has been suggested to take over the leading strand replication during final stages of replisome progression (Aria & Yeeles, 2018; Garbacz *et al*., 2018; Zhou *et al*., 2019). Additionally, according to *in vitro* experiments utilising the reconstituted *S. cerevisiae* replisome, the main leading strand replicase Polε is stochastically replaced by Polδ, and the occurrence of such polymerase switch events depends on Polδ concentration (Yeeles *et al.*, 2017). In chapter 3, we utilise *S. pombe* as a model system to investigate whether the hypothesised stochastic replicase switching manifests in living cells.

In most eukaryotic organisms including human and *S. pombe*, Polδ complex is composed of four distinct components including the catalytic subunit p125/Cdc6 and three regulatory subunits p50/Cdc1, p68/Cdc27 and p12/Cdm1 (Hughes *et al.*, 1999; Iino & Yamamoto, 1997; Lee *et al.*, 1991; Liu *et al.*, 2000; MacNeill *et al.*, 1996; Reynolds *et al.*, 1998; Zhang *et al.*, 1995). Four-subunit Polδ composition is not accommodated by *S. cerevisiae*, which features a trimeric Polδ assembly of Pol3, Pol31 and Pol32, the respective orthologues of p125/Cdc6, p50/Cdc1 and p68/Cdc27 (Jain *et al*., 2019).

In human and *S. pombe*, p12/Cdm1, the smallest component of Polδ complex, represents the only non-essential Polδ subunit. Several studies have indicated that p12 is targeted for degradation in S phase and as a consequence of chronic DNA damage (Darzynkiewicz *et al*., 2015; Terai *et al.*, 2013; Zhang *et al*., 2013, 2007; Zhao *et al*., 2014). According to biochemical analyses, the accessory p12 subunit modulates enzymatic properties of human Polδ. While S phase specific three-subunit Polδ complex ($Pol\delta_3$) lacking p12 is characterised by lower replication rate and increased fidelity, four-subunit Polδ complex ($Pol\delta_4$) exhibits increased polymerisation rate and is considered error-prone (Huang *et al*., 2010; Meng *et al.*, 2010; Meng *et al*., 2009; Podust *et al.*, 2002). Additionally, $Pol\delta_3$ only displaces short 5'-end flaps, whereas $Pol\delta_4$ facilitates long track strand displacement synthesis (Lin *et al*., 2013). Based on available evidence, it is currently assumed that S-phase specific $Pol\delta_3$ represents the main Polδ form involved in replication, whereas $Pol\delta_4$ facilitates homologous recombination and DNA repair (Lee *et al*., 2019). In chapter 5, we set out to test whether functions of p12 and Cdm1 are evolutionary conserved, primarily aiming to establish *S. pombe* as a simple and versatile system to study Cdm1-dependent modulation of Polδ function.

Replication stress

General definition of replication stress

Replication fork progression is frequently challenged by endogenous impediments arising from inherent chemical properties of DNA molecules and physiological processes of DNA metabolism. Additionally, replication machineries face obstacles imposed by exogenous agents which compromise the integrity of DNA duplexes and alter metabolic and regulatory pathways critical for successful duplication of the genome (Mirkin & Mirkin, 2007; Vesela *et al.*, 2017; Zeman & Cimprich, 2014). In a broad sense, cellular conditions and molecular contexts which impair the progression of replication forks are commonly referred to as replication stress.

Impediments of replication progression can manifest on both DNA strands. Obstacles situated on the lagging strand template can be bypassed by the lagging

strand synthesis machinery. Bypass of lagging strand replication blocks generates a short gap in the newly synthesised strand, which is subsequently filled in by designated trans-lesion polymerases or a not yet fully characterised template switching mechanism, which allows utilisation of the newly synthesised chromatid as an alternative template (Taylor & Yeeles, 2018). Lesions situated on the leading strand template inhibit continuous leading strand polymerisation, stall replisome progression and, in general, are considered more toxic (Taylor & Yeeles, 2018).

Most replication blocks inhibit canonical DNA polymerisation but allow unhindered progression of CMG helicases. Uncoupling of helicase and polymerase activities is accompanied by accumulation of RPA-bound ssDNA, which is recognised as a signature of compromised replication fork function and triggers replication checkpoint signalling cascade (Byun *et al.*, 2005; Pacek & Walter, 2004). A subset of DNA lesions, such as inter-strand crosslinks (ICLs) and some protein-DNA crosslinks (PDCs), do not permit free bypass of CMG helicases, and thus function as potent blocks of whole replisome complexes (Deans & West, 2011; Hizume & Araki, 2019).

Mechanisms compensating for compromised replication fork function

Stalled or otherwise immobilised replication forks represent fragile structures which compromise timely completion of replication and drive genomic instability (Zeman & Cimprich, 2014). Stalled replication forks can be rescued by converging replisomes (Al Mamun *et al*., 2016; Brambati *et al*., 2018; Jieqiong Zhang *et al*., 2015) or salvaged by several distinct mechanisms including lesion bypass by trans-lesion synthesis (TLS) (Guilliam & Yeeles, 2020b; Yang & Gao, 2018), PrimPol-mediated re-priming and re-initiation (Guilliam & Doherty, 2017), template switching (Carr & Lambert, 2013), HoRReR (Miyabe *et al*., 2015; Naiman *et al*., 2021), and lesion traverse facilitated by FANCM or DONSON (Zhang *et al*., 2020). In *S. cerevisiae*, stability of stalled replication forks is at least partially dependent on fork protection complex (FPC) composed of accessory replisome components Csm3/Tof1 (Timeless/Tipin in human) and Mrc1 (claspin in human) (Baretić *et al*., 2020). In several experimental systems, replication fork stalling is accompanied by replisome backtracking and formation of a four-way chicken foot structure. It has been proposed that accommodation of a chicken foot structure also contributes to replisome stabilisation

and repair (Quinet *et al.*, 2017). If all safety mechanisms fail, stalled replication forks collapse and generate one-ended double-strand break, which can be repaired by homology-directed BIR (Donnianni *et al*., 2019; Kramara *et al.*, 2018).

A subset of molecular mechanisms mediating replication fork restart are error-prone and can drive genomic instability. TLS polymerases, which facilitate replication block bypass and post-replicative gap filling following PrimPol-mediated re-priming, lack proof-reading activity, and thus introduce mutations at higher frequency (Yang & Gao, 2018). Homology directed mechanisms including template switching, HoRReR and BIR can utilise non-allelic repair templates and introduce deletions, insertions and gross chromosomal rearrangements (Hu *et al*., 2013; Miyabe *et al*., 2015; Mizuno *et al.*, 2009). Additionally, HoRReR and BIR feature non-canonical Polδ-mediated polymerisation, which is considered error-prone (Donnianni *et al*., 2019; Naiman *et al*., 2021).

Replication checkpoint

The most fundamental function of replication checkpoint signalling is to prevent cells with under-replicated genomes from entering mitosis. Consequences of checkpoint activation include suppression of late origin firing, inhibition of further cell cycle progression, deployment of factors facilitating adequate DNA repair, stabilisation of compromised replication forks and promotion of replication fork repair and/or restart (Molinari *et al.*, 2000; Santocanale & Diffley, 1998; Shechter *et al.*, 2004; Shell *et al*., 2009; Trenz *et al.*, 2006; Zeng *et al*., 1998; Zhao *et al.*, 2002).

The following text uses human nomenclature. The Respective *S. cerevisiae* orthologues are indicated as superscripts. The replication checkpoint is a complex multilayer signalling cascade involving numerous modulatory and functional elements. Replication checkpoint is mainly governed by the ataxia telangiectasia and Rad3-related (ATR^{Mec1}) kinase and its effector, the $CHK1^{Chk1}$ kinase (Recolin *et al.*, 2014). According to the current consensus, ATR is recruited to stalled replication forks by its co-factor $ATRIP^{Ddc2}$, which directly interacts with ssDNA-bound RPA (Ball *et al.*, 2005; Recolin *et al*., 2014; Zou & Elledge, 2003). Activation of ATR is mediated by autophosphorylation and requires action of additional factors including the checkpoint clamp 9-1-1 ($RAD9^{Ddc1}$-$RAD1^{Rad17}$-$HUS1^{Mec3}$) complex and its loader

RAD17^{Rad24}-RFC (Bermudez *et al.*, 2003; Zou, 2002), the multidomain modulator TopBP1^{Dbp11} (Hashimoto *et al.*, 2006; Parrilla-Castellar & Karnitz, 2003), the MRN (MRE11^{Mre11}-RAD50^{Rad50}-NBS1^{Nbs1}) complex (Carson *et al.*, 2009; Duursma *et al.*, 2013) and, in human cells, the checkpoint mediator MDC1 (Wang *et al.*, 2011). Additionally, in human cells, several lines of evidence indicate that ATR activation also requires activities of DNA helicase BACH1/FANCJ and Polα-primase. While BACH1/FANCJ helicase facilitates further propagation of ssDNA tracks, Polα-primase, by synthesising short RNA/DNA primers, generates ssDNA/dsDNA junctions which contribute to lesion recognition (Gong *et al.*, 2010; Michael *et al.*, 2000; Van *et al.*, 2010). In human cells, ATR can be also activated via a distinct pathway involving ETAA1 (Achuthankutty *et al.*, 2019; Thada & Cortez, 2019).

Active ATR^{Mec1} kinases phosphorylate numerous substrates including histone variant H2AX, commonly referred to as γH2AX and CHK1 (Liu *et al.*, 2000; Lopez-Girona *et al.*, 2001; Ward & Chen, 2001; Ward *et al.*, 2004; Zhao & Piwnica-Worms, 2001). In broad molecular contexts, γH2AX phosphorylation signifies compromised DNA integrity and functions as a docking site for MDC1 and other factors (Ibuki & Toyooka, 2015; Leimbacher *et al.*, 2019). MDC1 functions as a molecular scaffold which facilitates recruitment of additional ATR molecules which further propagate γH2AX phosphorylation and amplify the checkpoint signal. Other direct targets of ATR are RPA (Liu *et al.*, 2012) and various replication and DNA repair factors (Matsuoka *et al.*, 2007). When activated, the checkpoint effector kinase CHK1^{Chk1} inactivates CDC25 phosphatase (Furnari *et al.*, 1997; Sanchez *et al.*, 1997) and DDK kinase (Heffernan *et al.*, 2007; Shechter *et al.*, 2004). Inhibition of CDC25 prevents activation of CDK kinases and effectively stalls further cell cycle progression (Molinari *et al.*, 2000; Zeng *et al.*, 1998; Zhao *et al.*, 2002). Additionally, combined inactivation of CDK and DDK kinases and their substrates is believed to play a role in global suppression of origin firing and consequently slower S phase progression (Petermann *et al.*, 2010; Shechter *et al.*, 2004; Zegerman & Diffley, 2010).

Depending on the type of DNA lesion and the cellular context, ATR^{Mec1}-CHK1^{Chk1} signalling can functionally overlap with other DDR responders, such as ataxia telangiectasia mutated (ATM) together with its main effector kinase CHK2 and DNA-dependent protein kinase (DNA-PK) (Blackford & Jackson, 2017).

Endogenous sources of replication stress

The most prevalent natural causes of replication stress include repetitive DNA sequences, which accommodate non-canonical (non-B) DNA conformations, and collisions between replication and transcription machineries (Mirkin & Mirkin, 2007; Zeman & Cimprich, 2014).

Repetitive DNA elements and non-canonical DNA structures

Repetitive DNA elements and motifs prone to form non-B DNA conformations, such as G-quadruplexes, left-handed Z-DNA helices, triple-stranded H-DNA structures and cruciforms, have been associated with signs of genome instability including DNA double-strand breaks, increased frequency of base substitutions, deletions, duplications and broad range of chromosomal rearrangements (Bacolla & Wells, 2004; Guiblet *et al*., 2021; Samadashwily *et al.*, 1997; Tognetti & Speck, 2016; Wang & Vasquez, 2014).

Unperturbed replisome progression is challenged by sequences composed of trinucleotide tandem repeats of various compositions and lengths (Samadashwily *et al*., 1997; Tognetti & Speck, 2016). In yeast models as well as mice and human cells, trinucleotide tandem repeats have been correlated with replisome stalling, increased occurrence of chromosomal breaks and contraction or expansion of respective trinucleotide elements (Bontekoe *et al*., 2001; De Temmerman *et al*., 2008; Gellon *et al*., 2019; Razidlo & Lahue, 2008; Voineagu *et al.*, 2009). In human cells, unrestrained expansion of trinucleotide repeats has been linked to several clinical disorders including Huntington disease, fragile X disorder, myotonic dystrophy type 1 and others (Paulson, 2018). Similarly to trinucleotide repeats, low complexity AT-rich sequences also represent well characterised fragile sites in yeast as well as human cells (Sinai *et al*., 2019; Zhang & Freudenreich, 2007).

In mammalian organisms, the most prevalent species of repetitive DNA are the endogenous signatures of previous retroviral infections also known as retrotransposons. Stable genomic remnants of retroviral genomes collectively constitute at least one third of human DNA and represent hotspots of chromosomal rearrangements (Ade *et al.*, 2013; Burwinkel & Kilimann, 1998; Lander *et al*., 2001; Lee *et al*., 2012). Another form of repetitive DNA, collectively referred to as satellite DNA,

composes of tandem arrays of repeated DNA elements of various lengths and base compositions (Garrido-Ramos, 2017). In most organisms, satellite DNA forms specific genomic domains such as telomeres, centromeres and pericentromeric regions, all of which are classified as "difficult to replicate" domains (Forsburg, 2013; Forsburg & Shen, 2017; Hartley & O'Neill, 2019; Stroik & Hendrickson, 2020).

According to the current consensus, genomic instability associated with repetitive DNA mainly originates from frequent replication stalling events and illegitimate fork repair or restart (Zeman & Cimprich, 2014). Intriguingly, many forms of repetitive DNA associate with heterochromatin, the condensed conformation of which has been also discussed as a potential impediment of replication fork progression (Beeharry *et al.*, 2013; Forsburg, 2013; Garrido-Ramos, 2017; Stroik & Hendrickson, 2020). Publicly available data and molecular tools addressing the impact of heterochromatinisation on replication fork progression, however, are relatively scarce. In chapter 4, we establish a robust system to analyse replication of artificial heterochromatin domains.

Collisions between replication and transcription

Replication and transcription operate on the same DNA template, which inevitably leads to physical and/or functional interference between replisomes and RNA polymerases. So-called replication transcription collisions (RTCs) have been associated with signatures of genome instability (Gottipati *et al.*, 2008; Hamperl *et al.*, 2017; Helmrich *et al.*, 2011; Prado & Aguilera, 2005) and contribute to propagation of clinical disorders such oculomotor apraxia type 2 (Anheim *et al*., 2009; Moreira *et al*., 2004), amyotrophic lateral sclerosis type 4 (Chen *et al*., 2004) and cancer (Hatchi *et al*., 2015; Zhao *et al*., 2010).

Replisomes and RNA polymerases thread along the DNA template in 5'-3' direction and can encounter each other in co-directional or head-on orientation. Head-on RTCs are accompanied by replication fork pausing and significant increase in occurrence of homologous recombination events, whereas co-directional clashes only cause minor disturbance to replisome progression (Brüning & Marians, 2020; Prado & Aguilera, 2005). Thus, it is generally assumed that head-on RTCs are more deleterious

then co-directional encounters (Brüning & Marians, 2020; Hamperl *et al*., 2017; Hamperl & Cimprich, 2016; Prado & Aguilera, 2005).

According to the general consensus, transcription-dependent obstruction of replisome progression arises from physical interreference between replisomes and RNA polymerases (Hamperl & Cimprich, 2016). Alternatively, it has been also proposed that accumulation of supercoiled structures ahead of two converging machineries can result in topological constrains and consequential inhibition of DNA unwinding, which obstructs progression of replisomes and RNA polymerases (Bermejo *et al.*, 2012).

A specific form of RTCs involve triple-stranded RNA/DNA structures, also known as R-loops, which form when the 5' end of the nascent RNA re-anneals with the template DNA (Crossley *et al.*, 2019). Formation of R-loops is believed to be consequential to RNA polymerase stalling or defective termination of transcription (Crossley *et al*., 2019). Many studies correlate R-loop accumulation with signs of genomic instability, and it is widely accepted that R-loop structures represent potent replication blocks (Alzu *et al*., 2012; Brambati *et al*., 2018; Crossley *et al*., 2019; Mischo *et al*., 2011; Stirling *et al*., 2012). Interestingly; however, it has been also proposed that progressing replisomes are only stalled by protein-bound R-loops (Brüning & Marians, 2020).

Cellular mechanisms responsible for prevention and resolution of RTCs are complex and, up to this day, only partially understood. Early studies employing human and mouse cell cultures indicate that, in S phase, activities of replication and transcription machineries are spatially separated (Wansink *et al*., 1994; Wei *et al*., 1998). It is believed that strict demarcation of replication and transcription zones represents a basic mechanism to minimise occurrence of RTCs. Additionally, in human cells, actively transcribed loci seem to be preferentially oriented in a direction which favours co-directional, rather than more toxic head-on RTCs (Chen *et al*., 2019; Petryk *et al*., 2016).

Nonetheless, from accumulating evidence it is clear that RTCs represent a potent source of genomic instability. Factors and mechanisms involved in supressing pathological RTCs, however, are poorly understood. Several studies indicate that senataxin RNA/DNA helicases could represent factors involved in suppression or immediate resolution of RTCs (Alzu *et al*., 2012; Appanah *et al.*, 2020; Brambati *et al*.,

2018; Mischo *et al*., 2011; Stirling *et al*., 2012). In chapter 3, we employ polymerase usage sequencing to investigate the role of senataxin helicases Sen1^{+} and Dbl8^{+} in the maintenance of unperturbed replication in *S. pombe*.

Means to study genome-wide replication dynamics

Accurate assessment of genome-wide replication progression had been a long-lasting experimental challenge. In the past decade, numerous methods addressing genome-wide positions of replication origins, replication fork directionality and replication timing have been developed. In mammalian cells, the most prominent approaches include short nascent strand sequencing (SNS-Seq) (Besnard *et al*., 2012), Bubble-Seq (Mesner *et al*., 2013), chromatin immunoprecipitation of ORC components followed by deep sequencing (Kirstein *et al*., 2021), initiation site sequencing (Ini-Seq) (Langley *et al*., 2016), Okazaki fragment sequencing (OK-Seq) (Petryk *et al*., 2016) and methods detecting base analogues, such as 5-Ethynyl-2'-deoxyuridine (EdU), incorporated after transient hydroxyurea-mediated G1/S block (EdUseq-HU) (Macheret & Halazonetis, 2018). The most recent method developed in human cells is the genome-wide ligation of 3'-OH ends followed by sequencing (GLOE-Seq) which allows detection of single-strand DNA breaks and intermediates of lagging strand replication (Sriramachandran *et al*., 2020).

Additional approaches, which were developed in yeast models, utilise analysis of DNA copy number in synchronised cell populations or asynchronous cells sorted by their DNA content (Müller *et al*., 2014), detection of incorporated bromodeoxyuridine (BrdU) in HU-treated cells (Gan *et al*., 2017; Peace *et al.*, 2014) and detection of incorporated BrdU in unperturbed cells by nanopore sequencing (Müller *et al*., 2019).

Relatively recent methodological advancement in studies of replication progression in fission yeast *S. pombe* and budding yeast *S. cerevisiae* capitalise on mutated versions of replicative polymerase complexes which are characterised by reduced base selectivity and incorporate ribonucleotides (rNMPs) at markedly increased frequencies (Clausen *et al*., 2015; Daigaku *et al*., 2015; Koh *et al.*, 2015; Reijns *et al*., 2015). Misincorporated rNMPs are stabilised by disruption of ribonucleotide excision repair (RER). Positions of stabilised genomic rNMPs are determined by rNMP-specific DNA nicking by alkali or recombinant RNase H2 followed

by denaturation and deep sequencing of generated ssDNA fragments. The most successful methods utilising rNMP traces as a proxy for determination of respective polymerase activities are hydrolytic end sequencing (HydEn-Seq), developed in *S. cerevisiae*, and polymerase usage sequencing (Pu-Seq), developed in *S. pombe* (Clausen *et al*., 2015; Daigaku *et al*., 2015). Studies employing HydEn-Seq and Pu-Seq methodologies provided additional insights into genome-wide activities of replicative polymerases Polα, Polδ and Polε during unperturbed replication as well as DNA repair processes such as BIR and HoRReR (Clausen *et al*., 2015; Daigaku *et al*., 2015; Donnianni *et al*., 2019; Garbacz *et al*., 2018; Naiman *et al*., 2021; Zhou *et al*., 2019). In chapter 1, we revisit the original Pu-Seq protocol and propose an updated version, which aims to provide lower cost, better processivity and increased accuracy.

PhD aims and brief descriptions of presented projects

The PhD work presented in this thesis aims to introduce new methodological and theoretical concepts, which will be beneficial for researchers interested in molecular mechanisms of eukaryotic DNA replication and genome stability maintenance. In chapter 1, we introduce an updated protocol for preparation of Pu-Seq libraries, which, in comparison with the original version, minimise potential experimental bias and, at the same time, improves processivity and cost-efficiency. We believe that the newly developed Pu-Seq procedure can become a standardly used protocol, and thus represents a valuable outcome of our work. In chapter 2, we cover our efforts to investigate the effects of moderately increased Polδ levels on replication dynamics and cell physiology. We argue that presented data provide interesting insights into mechanics of DNA replication, and we emphasise the importance of experimental validation (Zach & Carr, 2021). In chapter 3, we utilise Pu-Seq technology and investigate roles of senataxin paralogs Sen1^{+} and Dbl8^{+} in the maintenance of unperturbed replication fork progression. We provide the first evidence of the role of Sen1^{+} and Dbl8^{+} senataxins in prevention and/or resolution of collisions between replication transcription machineries. Moreover, we postulate a novel layer of functional divergence between Sen1^{+} and Dbl8^{+} paralogs. In chapter 4, we create an elaborate system, which allows assessment of replication progression across induced heterochromatin domains. We argue that utilisation of our system can provide new

mechanistic insights into replication dynamics of heterochromatinised domains. In chapter 5, we initiate a project which aims to provide a better understanding of Cdm1^{+} (p12 in human), the small subunit of Polδ, which functions as a modulator of Polδ enzymatic properties. We present the preliminary data, which indicate that the function of the small Polδ subunit is conserved between human and *S. pombe*. We believe that the discussed work represents a solid foundation for further project development.

Methods

Cultivation and cell biology techniques

Yeast culture and transformation

S. pombe cells were grown according to standard procedures (Petersen & Russell, 2016) in yeast extract with supplements (YES; Formedium PCM0155, PSU0110) or Edinburgh minimal medium (EMM; Formedium, PMD0210) supplemented with adenine (225 mg/L), uracil (225 mg/L) and leucine (225 mg/L). In experiments involving anhydrotetracycline (ahTET), the ahTET stock solution (2.5 mg/mL) was prepared by dissolving ahTET (SIGMA, 37919) in DMSO. ahTET was added to the cultivation medium to the final concentration 2.5 µg/mL. In experiment involving thiamine, the thiamine stock solution (30 mM) was prepared by dissolving thiamine (SIGMA, T4625) in distilled H_2O. Thiamine was added to the cultivation medium to the final concentration 15 µM. For spot-tests, 10-fold serial dilutions of exponentially growing cells were spotted onto a freshly prepared solid medium plate of interest. Optical densities (OD) were measured by WPA CO8000 Cell Density Meter (Biochrom). Doubling times were calculated using the formula: DT = 1/k, where DT stands for doubling time and k represents the slope of linear regression computed from a time-series of $\log_2$-transformed OD measurements. Cells were transformed by the lithium-acetate based method (Bähler *et al.*, 1998). Optical densities were measured by WPA CO8000 Cell Density Meter (Biochrom). A list of strains used in this study is provided in Table 1.

Microscopy

1 mL of exponentially growing fission yeast cells (OD_{600} = 0.5; 5 × 10^6 cells/mL) was centrifuged (1000 × g, 5 min, 25°C) and cell pellet resuspended in 1 mL of 70% ethanol. 500 µL of fixed cells were collected by centrifugation (1000 × g, 5 min, 25°C) and re-suspended in 50 µL of H_2O containing 1 µM 4′,6-diamidino-2-phenylindole (DAPI). Cells were incubated at room temperature in the dark for at least 15 min, and then analysed by microscopy using a Nikon E400 system. Cell lengths were determined from DIC images by measuring the distance between the opposite poles of the cell using ImageJ software (version 1.51m9) (Schneider, Rasband, & Eliceiri, 2012). At least 200 cells per sample were scored.

Cell synchronisation and DNA content analysis

Exponentially growing *cdc2*^{asM17} cells (OD_{600} = 0.1-0.2; 1-2 × 10^6 cells/mL) were treated with 2 µM 3-Br-PP1 (abcam, ab143756) for 3 h. A 50 mL fraction 3-Br-PP1-treated culture was centrifuged (1000 × g, 5 min, 25°C) and the cell pellet washed with 50 mL of fresh YES medium pre-heated to 30°C. Washed cells were re-suspended in

50 mL of fresh pre-heated YES and incubated at 30°C. In 15-min intervals, 1 mL aliquots of synchronous cell culture were centrifuged (1000 × g, 3 min, 25°C) and collected cells fixed in 1 mL of 70% ethanol. For each time point, 500 µL of fixed cells were centrifuged (1000 × g, 3 min, 25°C), the supernatant was discarded, and the cell pellet re-suspended in 500 µL of sodium citrate (50 mM, pH = 7) containing 1 mg/mL RNase A (NEB, T3018L). The resulting cell suspension was incubated for 3 h at 37°C, and then mixed with 500 µL of sodium citrate (50 mM, pH = 7) containing 2 µM SYTOX Green (Invitrogen, S7020). Samples were analysed using an Accuri C6 flow cytometry system (Beckman Coulter). Data were analysed by BD CSampler software (version 1.0.264.21) and R (version 4.0.0) (https://www.R-project.org; R Core Team, 2020).

Cloning

Cre-recombination mediated cassette exchange

Cre-recombination mediated cassette exchange (RMC) was performed as described previously (Watson, Garcia, Bone, Carr, & Armstrong, 2008). Briefly, leucine-auxotrophic cells carrying a LoxP-LoxM3 integration site were transformed with a pAW8-derived Cre-Lox integration vector carrying an insert of interest. Successful transformants (leucine-prototrophic) carrying a pAW8-derived Cre-Lox integration vector were selected on EMM plates lacking leucine. Single clones were isolated and grown over-night in liquid YES medium. 200-500 cells were plated onto a plate (containing leucine) selecting for a given integration. Leucine-auxotrophic colonies, which lacked transformed Cre-Lox vector, carrying a selection marker associated with a given integration construct were selected (Watson *et al*., 2008).

Construction of heterochromatin-induction system

Strains carrying *I-3325162:LoxP-kanR-tetO7-LoxM3* construct were generated by Cre-Lox mediated cassette exchange (Watson *et al*., 2008). Briefly, *rts-ura4*$^{+}$ construct in previously constructed *I-3325162:LoxP-rts-ura4+-LoxM3* cells was replaced with *kanR-tetO7* utilising pAW8 (Addgene, 110222)-derived Cre-Lox integration vector pAW8_kanR-tetO7 (pRZ10). Strains carrying *II-1389186:LoxM3-tetO7-natR-LoxP* and *III-1609353:LoxP-ura4-tetO7-LoxM3* constructs were generated de-novo by PCR-based gene targeting (Bähler *et al*., 1998) using *LoxM3-tetO7-natR-LoxP* and *LoxP-ura4-tetO7-LoxM3* PCR fragments carrying 5' and 3' 100 bp extensions homologous to the respective target loci. Transformed DNA fragments were generated by PCR templated by pAW8-derived plasmids; either pAW8_natR-tetO7 (pRZ08) or pAW8_ura4+-tetO7 (pRZ09). pRZ08, pRZ09 and pRZ10 were generated by standard restriction insertion coning using *Sph*I and *Sac*I restriction sites and T4 DNA ligase (NEB, M0202S). *Sph*I (NEB, R3182S) and *Sac*I-HF (NEB, R3156S) restriction enzymes were used. Strains carrying *I-5230932:LoxM3-tetR-2xFLAG-clr4cdd-LoxP* and *I-5230932:LoxM3-tetR-2xFLAG-clr4H410Kcdd-LoxP* constructs were generated by Cre-

Lox mediated cassette exchange (Watson *et al*., 2008). Briefly, KanR cassette in cells carrying *I-5230932:LoxM3-kanR-LoxP* integration locus was replaced with either *tetR-2xFLAG-clr4cdd* or *tetR-2xFLAGclr4H410Kcdd* utilising pAW8 (Addgene, 110222)-derived Cre-Lox integration vectors pAW8_nmt1-tetR(off)-2xFLAG-clr4cdd (pRZ19) or pAW8_nmt1-tetR(off)-2xFLAG-clr4H410Kcdd (pRZ18), respectively. pRZ18 and pRZ19 were generated by standard restriction insertion cloning using *Sac*I and *Sal*I restriction sites and T4 DNA ligase (NEB, M0202S). *Sac*I-HF (NEB, R3156S) and *Sal*I-HF (NEB, R3138S) restriction enzymes were used. *nmt1-tetR(off)-2xFLAG-clr4cdd* construct was obtained from R. Allshire lab (Audergon *et al*., 2015). H410K mutation in *nmt1-tetR(off)-2xFLAG-clr4H410Kcdd* construct was introduced by overlap extension PCR. Strains carrying *tetR(off)-2xFLAG-dfp1-bleoR* were generated by altering the endogenous *dfp1*$^{+}$ locus by PCR-based gene targeting (Bähler *et al*., 1998). Transforming *tetR(off)-2xFLAG-dfp1-bleoR* fragment was produced by overlap extension PCR. PCR amplifications were done with high-fidelity KOD Hot Start DNA Polymerase (Merck Millipore, 71085–3) or Q5 high-fidelity polymerase (NEB, M0492S). PCR fragments were purified with QIAquick PCR Purification Kit (QIAGEN, 28104) or QIAquick Gel Extraction Kit (QIAGEN, 28704). Plasmids were isolated with QIAprep Spin Miniprep Kit (QIAGEN, 27104).

Construction of Polδ over-expression system

Cre-Lox integration plasmids carrying 4 Polδ genes (*cdc1*$^{+}$, *cdc27*$^{+}$, *cdm1*$^{+}$, *cdc6*$^{+}$ or *cdc6*^{L591G}) and one of the three selection markers (NatR, KanR, ura4^{+}) were derived from the previously characterised pAW8 vector (Addgene, 110222) (Watson *et al.*, 2008) by standard restriction insertion cloning. Briefly, insert DNA fragments (*Sph*I-*cdc1*$^{+}$-*Apa*I, *Apa*I-*cdc27*$^{+}$-*Xho*I, *Xho*I-*ura4*$^{+}$-*Sac*I, *Sac*I-*cdc6*$^{+}$-*Sbf*I, *Sac*I-*cdc6*^{L591G}-*Sbf*I, *Sbf*I-*cdm1*$^{+}$-*Spe*I, *Apa*I-*cdc27*$^{+}$-NatMX6-*Sac*I, *Apa*I-*cdc27*$^{+}$-KanMX6-*Sac*I) were generated by high-fidelity PCR with KOD Hot Start DNA Polymerase (Merck Millipore, 71085-3) and purified with QIAquick PCR Purification Kit (QIAGEN, 28104) or QIAquick Gel Extraction Kit (QIAGEN, 28704). *Sac*I-*cdc6*^{L591G}-*Sbf*I, *Apa*I-*cdc27*$^{+}$-NatMX6-*Sac*I and *Apa*I-*cdc27*$^{+}$-KanMX6-*Sac*I were produced by overlap extension PCR. Generated Polδ gene fragments contained intact 5'UTR and 3'UTR sequences, as well as upstream and downstream regions of 663-980 bp. pAW8 vector and insert DNA fragments were digested by respective restriction enzymes and ligated over-night at 18°C using T4 DNA ligase (NEB, M0202S). Each ligation reactions contained 50 ng of pAW8 vector and 3-fold molar excess of respective DNA fragments. Ligation reactions were incubated in T3 Thermocycler (Biometra). Restriction digestion reactions were performed according to manufacturer's instructions with the following restriction enzymes: *Sph*I-HF (NEB, R3182S); *Apa*I (NEB, R0114S); *Xho*I (NEB, R0146S); *Sac*I-HF (NEB, R3156S); *Spe*I-HF (NEB, R3133S); *Sbf*I-HF (NEB, R3642S); *Sal*I-HF (NEB, R3138S); *Bam*HI-HF (NEB, R3136S). Ligation products were transformed into DH5-Alpha *E. coli*

competent cells. Plasmids were purified using a QIAprep Spin Miniprep Kit (QIAGEN, 27104). A list of generated plasmids is provided in Table 2.

Protein manipulation techniques

Protein extraction

10-20 mL of exponentially growing fission yeast cells (OD_{600} = 0.5, 5×10^6 cells/mL) were centrifuged (1000 × g, 3 min, 4°C), supernatant removed and cell pellet snap-frozen in liquid nitrogen. At this point, cell pellet could be stored at -80°C. Cell pellet was resuspended in 300 µL of NP-40 buffer (6 mM Na_2HPO_4, 4 mM $NaH_2PO_4{\cdot}H_2O$, 1% NONIDET P-40/IGEPAL CA-630, 150 mM NaCl, 2 mM EDTA, 50 mM NaF) supplemented with protease (Roche, Complete™, 4693159001) and phosphatase (Roche, PhosSTOP™, 4906837001) inhibitors. Cell suspension was transferred into a clean screw-cap tube and mixed with equal volume of ice-cold glass beads (SIGMA, G8772). Cells were lysed using Fastprep-24 instrument (6.5 m/s, 60 s, 2 cycles). The tube containing lysed cells was pierced at the bottom with a hot needle and placed on the top of a clean cap-less 1.5 mL centrifuge tube. Both tubes were put into a 15 mL centrifuge tube and centrifuged (4,000 × g, 2 min, 4°C). Supernatant from the bottom tube was transferred into a clean 1.5 mL centrifuge tube and centrifuged (16 000 × g, 10 min, 4°C). Supernatant was transferred into a clean 1.5 mL centrifuge tube and stored at -80°C.

SDS-PAGE

10% resolving gel (for two gels: 3.2 mL of deionised H_2O, 2.6 mL of 1.5 M Tris-HCl (pH 8.8), 2 mL of 40% acrylamide, 200 µL of 20% SDS, 20 µL of 30% ammonium persulfate and 8 µl of TEMED) was introduced into the vertical gel assembly apparatus (Bio-Rad), leaving approximately 1 cm of free space on the top. Resolving gel was covered by deionised H_2O and left to solidify. Excess H_2O was removed and solidified resolving gel was covered by 4% stacking gel (for two gels: 3.9 mL of deionised H_2O, 500 µL of 1.0 M Tris-HCl (pH 6.8), 500 µL of 40% acrylamide, 100 µL of 20% SDS, 16 µL of 30% ammonium persulfate and 8 µL of TEMED). At this point, a comb was introduced into the stacking gel and the assembly was left at room temperature until the stacking gel solidified. Comb was carefully removed, and gel assembly introduced into the vertical electrophoresis apparatus (Bio-Rad). The apparatus was placed in a vertical electrophoretic tank (Bio-Rad). The electrophoretic tank was subsequently filled with 1× running buffer with SDS (for 2 L of 10× running buffer: 288 g of glycine, 60.4 g of Tris base, 20 g of SDS and deionised H_2O up to 2 L). Protein samples dissolved in 1× Laemmli buffer (2× Laemmli buffer: 4% SDS, 20% glycerol, 120mM Tris-HCl, 0.02% bromophenol blue) were incubated at 95° for 10 min, centrifuged (16 000 × g, 10 min, 25°C), and 10 µL were loaded onto the gel alongside the protein size marker

Page Ruler (Thermo Fisher Scientific, 26616). Proteins were separated at constant 100 V for ca. 15 min, and then at constant 200 V until the head dye reached the bottom of the gel. Gels containing resolved proteins were subsequently used for western blotting.

Western blot

PVDF membrane (Amersham, 10600023) was washed with absolute ethanol and, subsequently, with deionised H_2O. Before putting together the western blot assembly, washed PVDF membrane, the acrylamide gel containing resolved proteins and two thick blot filter papers (Bio-Rad, 1703932) were equilibrated in 1× transfer buffer (for 1 L of 1× transfer buffer: 2.9 g of Tris base, 10 mL of 10% SDS, 14.4 g of glycine, 200 mL of absolute ethanol and deionised H_2O up to 1 L) for 15 min at room temperature. The western blot assembly was put together by placing thick blot filter paper, PVDF membrane, gel containing resolved proteins and another thick blot filter paper on top of each other in order listed from the bottom to the top. The whole assembly was soaked in 1× transfer buffer and introduced into the Trans-Blot SD Semi-Dry Transfer Cell system (Bio-Rad). Protein transfer was carried out at constant 15 V for 30 min. Following the protein transfer, PVDF membrane was washed with deionised H_2O, cut to size and placed in 50mL centrifugation tube(s) with protein-containing sides facing inwards. Each PVDF membrane was blocked in 5 mL of 5% low fat milk solution for 1 h on a tilt/roller mixer. Next, each PVDF membrane was washed three times with 5 mL of 1× TBS (for 1 L of 10× TBS: 80 g of NaCl, 200 mL of 1M Tris-HCl (pH 7.5) and deionised H_2O up to 1 L) containing 0.1% Tween 20 (TBSt). Following the three washes, each PVDF membrane was incubated with 5 mL of TBSt containing primary antibody on a tilt/roller at 4°C. The duration of incubation depended on the primary antibody used. α-tubulin was probed for 1 h by Mouse Monoclonal Anti-α-Tubulin, clone B-5-1-2 (SIGMA, T5168). FLAG-tagged proteins were probed for 1 h by ANTI-FLAG(R) M2 antibody (SIGMA, F3165). GFP was probed over night by Anti-GFP (clones 7.1 and 13.1) (SIGMA, 11814460001). Next, each PVDF membrane was washed three times with 5 mL of TBSt. Following the three washes, each membrane was incubated with 5 mL of TBSt containing the secondary antibody Anti-Mouse IgG (whole molecule) F(ab′)2 fragment–Cy3 antibody produced in sheep (SIGMA, C2181) on a tilt/roller for 2 h at 4°C. Each membrane was washed three times with TBSt and subsequently treated with Western Lightning Plus-ECL Enhanced Chemiluminescence Substrate (PerkinElmer, NEL104001EA). Membranes were imaged by ImageQuant LAS 4000 system.

DNA and RNA manipulation techniques

Chromosomal DNA extraction – Phenol/Chloroform method

200 mL of exponentially growing fission yeast cells (OD_{600} = 0.5, 5×10^6 cells/mL) were collected by centrifugation, resuspended in 1 mL of CSE buffer (50 mM Na_2HPO_4, 1.2 M sorbitol, 40 mM EDTA) containing 2.5 mg/mL zymolyase 100T (amsbio, 120493-1) and incubated at 37°C for 30 min. Cells were centrifuged (16,000 × g, 1 min, 4°C), washed twice with 1 mL of nuclease-free water and re-suspended in 450 µL of 5×TE. Resulting cell suspension was mixed with 50 µL of 10% SDS by inversion and incubated at 65°C for 10 min. Next, 150 µL of 5M potassium acetate were added. The cell suspension was mixed by inversion and incubated on ice for 10 min. Sample was centrifuged (16,000 × g, 10 min, 4°C) and the supernatant transferred into a clean 1.5mL centrifuge tube. Collected supernatant was mixed with 600 µL of isopropanol by inversion, incubated on ice for 10 min and centrifuged (16,000 × g, 10 min, 4°C). Supernatant was removed and the pellet washed with 1 mL of 70% ethanol. The pellet was resuspended in 190 µL of 5×TE containing 5 µL of RNase cocktail (Invitrogen, AM2286) and incubated at 37° for 2-3 h. During the course of incubation, samples were mixed by pipetting every ca. 30 min. 10 µL of proteinase K (20 mg/mL; NEB, P8107S) were added and samples were incubated at 55˚C for 2 h. Next, samples were mixed with 200 µL of phenol/chloroform/isoamyl alcohol (50:49:1; SIGMA, 77617) by vortexing and centrifuged (16,000 × g, 10 min, 4°C). 150 µL of the aqueous fraction were collected and DNA recovered either by isopropanol precipitation or with NucleoSpin gDNA Clean-up kit (MACHEREY-NAGEL, 740230.50). Concentration and quality of extracted DNA was assessed with NanoDrop spectrophotometer ND1000 (Thermo Scientific).

Chromosomal DNA extraction – QIAGEN Genomic-tip method

800 mL of exponentially growing fission yeast cells (OD_{600} = 0.5, 5×10^6 cells/mL) were collected by centrifugation (6,000 × g, 10 min, 4°C). Supernatant was removed, cells resuspended in 40 mL of sterile H_2O and the resulting suspension transferred into 50 mL centrifugation tube. Sample was centrifuged (4,000 × g, 5 min, 4°C) and supernatant removed. Cells were resuspended in 2 mL of NIB buffer (17% glycerol; 50mM MOPS; 150mM potassium acetate; 40 mM EDTA; pH 7.2 adjusted with KOH) containing 5 mg of zymolyase 100T (amsbio, 120493-1) and incubated at 37°C for 30 min. The suspension was mixed with 20 mL of sterile H_2O by vortexing and centrifuged (4,000 × g, 10 min, 4°C). Supernatant was removed and the cell pellet resuspended in 2 mL of G2 buffer (QIAGEN, 1014636). The suspension was mixed with 100 µL of RNase A (10 mg/mL) and incubated at 37°C for 30 min. Next, 100 µL of 30% N-Lauroyl Sarcosine (SIGMA, 61747) and 100 µL of freshly prepared proteinase K (20 mg/mL) were added and the resulting suspension was incubated at 55°C for 1 h. Sample was centrifuged (4,000 × g, 15 min, 4°C) and the supernatant transferred into a clean 15

mL centrifugation tube. Remaining cell pellet was resuspended in 1 mL of G2 buffer with 50 µL of 30% N-Lauroyl Sarcosine (SIGMA, 61747) and 50 µL of freshly prepared proteinase K (20 mg/mL). Resulting suspension was incubated at 55°C for 1 h, centrifuged (4,000 × g, 15 min, 4°C) and supernatant was pooled with the liquid fraction collected previously. DNA was purified with Genomic-tip 100/G (QIAGEN, 10243) according to manufacturer's instructions. Concentration and quality of extracted DNA was assessed with NanoDrop spectrophotometer ND1000 (Thermo Scientific).

Isolation of total RNA

Total RNA was isolated using MasterPure Yeast RNA purification kit (Cambio Ltd, MPY03100). The following protocol describes processing of 1 sample. No more than 12 samples should be processed at a time. The 1-2 mL aliquot of an exponentially growing culture (OD_{600} = 0.5; 5 × 10^6 cells/mL) was collected by centrifugation (1,000 × g, 3 min, 25°C). Supernatant was removed and the cell pellet snap-frozen in liquid nitrogen. At this stage, the sample could be stored at -80°C. The cell pellet was thawed on ice and resuspended in 300 µL of Extraction Reagent for RNA containing 50 µg/µL proteinase K. The suspension was mixed by vortexing and incubated at 70°C for 15 min. During incubation, the suspension was mixed by vortexing every 5 min. The sample was put on ice for 5 min and subsequently mixed with 175 µL of MPC Protein Precipitation Reagent by vortexing for 10 s. The sample was centrifuged (16,000 × g, 10 min, 4°C) and the supernatant transferred into a clean 1.5 mL centrifugation tube. The supernatant was mixed with 500 µL of ice-cold isopropanol by inverting the tube 40 times. Precipitated nucleic acids were sedimented by centrifugation (16 000 × g, 10 min, 4°C). The supernatant was discarded and the pellet resuspended in 200 µL of DNase-containing solution (175 µL H_2O, 20 µL 10x DNase Buffer and 5 µL RNase-Free DNase I (1 U/µL)). DNase reaction was carried out at 37°C for 30 min. Following the DNase treatment, 200 µL of 'T & C Lysis Solution' were added. The resulting solution was vortexed for 5 s and subsequently mixed with 200 µL of MPC Protein Precipitation Reagent by vortexing for 10 s. The resulting mixture was incubated on ice for 5 min and centrifuged (16,000 × g, 10 min, 4°C). The supernatant was transferred into a clean 1.5 mL centrifugation tube and mixed with 500 µL of ice-cold isopropanol by inverting the tube 40 times. Sample was centrifuged (16,000 × g, 10 min, 4°C), the supernatant removed, and the pellet washed twice with ice-cold 70% ethanol. Ethanol was removed and the pellet resuspended in 35 µL of H_2O containing 1 µL of RiboGuard RNAse inhibitor. RNA concentration was determined by NanoDrop spectrophotometer ND1000 (Thermo Scientific). RNA was stored at -80°C.

Reverse transcription

Reverse transcription (RT) was performed using RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, K1621). The following protocol describes processing of 1 sample. RT reaction components were mixed on ice. Purified RNA (0.1 - 2 µg) was mixed with 1 µL of random hexamer primers (0.2 µg/µL) and RNase-free H_2O, so the final volume was 12 µL. The mixture was incubated at 65°C for 5 min and subsequently put on ice. Additional RT reaction components including 4 µL of 5x Reaction Buffer, 1 µL of RNase inhibitor RiboLock™ (20 U/µL), 2 µL of dNTPs (10 mM) and 1 µL of RevertAid™ M-MuLV Reverse Transcriptase (200 U/µL) were added. The RT reaction was mixed by vortexing, briefly centrifuged and incubated at 25°C for 5 min. The RT reaction was carried out at 42°C for 60 min and subsequently terminated at 70°C for 5 min. cDNA was stored at -80°C.

Quantitative PCR

Quantitative PCR (qPCR) was performed with Luna Universal qPCR Master Mix (NEB, M3003E) and AriaMx Real-time PCR System (Agilent Technologies). The following protocol describes processing of 1 sample. A 20 µL qPCR reaction was prepared by mixing Luna Universal qPCR Mix (10 µL), forward primer [10 µM] (0.5 µL), reverse primer [10 µM] (0.5 µL), template DNA (2 µL of 500-fold diluted cDNA; 5 µL of immunoprecipitated DNA) and corresponding volume of nuclease-free H_2O. Template DNA in no template control (NTC) and no reverse transcriptase (No-RT) reactions was replaced by nuclease-free H_2O or 500-fold diluted No-RT sample, respectively. For every unique qPCR reaction, 3 technical replicates were analysed and averaged using AriaMX Software (Agilent, version 1.71). In RT-qPCR experiments, relative transcript levels were calculated using the formula $RNA_{REL} = 2^{-Cq(T)} / 2^{-Cq(R)}$, where RNA_{REL} represents relative transcript level of a gene of interest, and Cq(T) and Cq(R) stand for PCR quantification values of target and reference genes, respectively. *act1* was used as the reference. In ChIP-qPCR experiments, the percent of input was calculated using $I_{Perc} = 100 \times 2^{Cq(adjusted\ input) - Cq(IP)}$, where I_{Perc} stands for percent of input, Cq(adjusted input) for adjusted PCR quantification value of input, and Cq(IP) for PCR quantification value of immunoprecipitated DNA. Primer efficiency was calculated from a linear regression model of log_2-transformed PCR quantification values of 5-fold serial dilutions of DNA using $E = 10^{-1/k} \times 100$, where E stands for primer efficiency and k represents the slope of linear regression model. Only primers with an estimated efficiency > 90% were used. A list of qPCR primers is provided in Table 2.

Polymerase usage sequencing (original version)

20 µg of genomic DNA suspended in 70 µL of nuclease-free H_2O were mixed with 30 µL of freshly prepared 1M NaOH and incubated at 55°C for 2 h. DNA fragments were separated by agarose gel electrophoresis. The agarose gel was stained with

acridine orange (5 µg/mL) for 2 h at room temperature and de-stained overnight in 500 mL of distilled H_2O. 300-500 bp fragments were excised and purified with NucleoSpin Gel and PCR Clean-up (MACHEREY-NAGEL, 740609.50). Concentration of ssDNA was measured by NanoDrop spectrophotometer ND1000 (Thermo Scientific). 100 ng of ssDNA were transferred into a clean PCR tube. Volume of the sample was adjusted to 20 µL with nuclease-free H_2O. ssDNA was mixed with 5 µL of random 8-mers (3 mg/mL) and 5 µL of 10× NEB2.1 buffer. The sample was incubated at 95°C for 5 min, and then put on ice for 5 min. Next, the sample was mixed with 4 µL of nuclease-free H_2O, 5 µL of dNTPs (2mM each) containing deoxy uridine instead of deoxy thymidine and 1 µL T4 DNA polymerase (NEB, M0203S/L). The resulting mixture was incubated at 37°C for 20 min and subsequently mixed with 5 µL of 0.5M EDTA (pH 8.0). The sample was transferred into a clean 1.5 mL tube and mixed with 99 µL of AMPure XP beads (Beckman Coulter, A63881). The mixture was incubated at room temperature for 5 min. Beads were separated from the solution using a magnetic tube holder. The supernatant was removed and beads washed two times with 80% ethanol. Beads were air-dried and DNA eluted in 60 µL of nuclease-free H_2O. At this point, the quality of dsDNA was assessed by Bioanalyzer (Agilent) using High Sensitivity DNA Chip. 50 µL of purified dsDNA were mixed with components of NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, E7645S/L), 7 µL of NEBNext Ultra II End Prep Reaction Buffer, and 3 µL of NEBNext Ultra II End Prep Enzyme Mix. The mixture was incubated at 20°C for 20 min and, subsequently, at 65°C for 30 min. Next, the following components were added: 30 µL NEBNext Ultra II Ligation Master Mix, 1 µL of NEBNext Adaptor for Illumina (1.5 µM), 2.5 µL of nuclease-free H_2O. Reaction was incubated at 20°C for 15 min. Afterwards, the sample volume was adjusted to 100 µL with nuclease-free H_2O, transferred to a clean 1.5 mL tube and mixed with 30 µL of AMPure XP beads. Beads were separated from the solution using a magnetic tube holder and the supernatant was transferred into a clean 1.5 mL tube. The collected fraction was mixed with fresh 40 µL of AMPure XP beads. Beads were separated from the solution using a magnetic tube holder and washed three times with 200 µL of 80% ethanol. Beads were air-dried and DNA eluted in 25µL of 0.1× TE (pH 8.0). 20 µL of eluted DNA was transferred into a clean PCR tube and mixed with 3 µL of USER enzyme (NEB, M5505S/L), 25 µL of NEBNext Ultra II Q5 Master Mix (NEB, M0544S/L), 2.5 µL of universal PCR primer (10 µM) and 2.5 µL of index primer (10 µM). Resulting reaction mix was incubated at 37°C for 15, and then PCR-amplified with 75 s of annealing/extension at 65°C using 9-13 cycles. The enriched DNA library was purified twice with equal volume of AMPure XP beads and eluted in 23 µL of nuclease-free H_2O. The quality of prepared DNA library was assessed by Bioanalyzer (Agilent) using High Sensitivity DNA Chip.

Next generation sequencing data analysis

Calculation of polymerase track values

Used in chapters 1, 2 and 3. Polymerase tracks at any given 300 bp bin were calculated using the equation $PT = (R_T - R_B) / (R_T + R_B)$, where PT represents polymerase track, and R_T and R_B stand for ribonucleotides (rNMPs) mapped to the top and the bottom DNA strands, respectively. Polymerase tracks were determined for each biological repeat separately then averages of the individual repeats were used for subsequent analysis.

Estimation of origin efficiencies from polymerase track values

Used in chapter 2. Positions and efficiencies of origins of replication were determined from differential values of polymerase tracks, similarly to (Daigaku *et al.* (2015). Briefly, for all three datasets (Polδ, Polε, Polα), the difference of each neighbouring datapoint of polymerase track values (smoothed by simple moving average of 3) was calculated as $Diff_i = PT_i - PT_{i-1}$, where $Diff_i$ represents the differential value at position i, and $PT_i - PT_{i-1}$ stand for smoothed polymerase track values at positions i and i-1, respectively. The differential value of the first bin on a given chromosome was assigned 0. Polε differentials and the opposites of Polδ and Polα differentials were averaged and smoothed by simple moving average of 3. Then, positive peaks (indicating sharp inclinations in the data) were selected. Differential peaks containing two or more distinct maxima separated by at least 4 bins were treated as unresolved independent peaks. Peaks with maxima bellow 30th quantile were disregarded. Each independent differential peak represented an origin of replication, the efficiency of which was estimated as 50% of the sum of its values. 259 replication initiation regions and 147 termination zones were selected using wild-type (WT) origin efficiency data. For comparison purposes, origin efficiencies were normalised assuming that the efficiency of the most efficient origin was 100%. Data were analysed in R (https://www.R-project.org; R Core Team, 2020) using a custom script (see *Software availability*). Source code available from: https://github.com/R-Zach/Pu-Seq_ polymerase_delta_over-expression. Archived source code at time of publication: https://doi.org/ 10.5281/zenodo.4516546 (Zach, 2021). License: Apache License 2.0.

Estimation of relative replication fork stalling

Used in chapter 3. Differentials of averaged polymerase tracks were calculated using $Diff_i = PT_i - PT_{i+1}$, where $Diff_i$ represents differential value at position i, and $PT_i - PT_{i+1}$ stand for smoothed polymerase track values at positions i and i+1, respectively. Differential values were smoothed by simple moving average (window of 3 bins) and negative values indicating sites of replication initiation were assigned 0. Next,

differences in non-negative differentials between senataxin mutants (*sen1Δ*, *dbl8Δ*, *sen1Δdbl8Δ*) and WT were calculated. Subsequently, maximum values and areas of distinct positive peaks were determined, and peaks characterised by maxima and areas above respective 0.98 quantiles were selected. Selected peaks represented hypothetical replication fork stalling sites. Repetitive genomic regions were omitted from the analysis.

Calculation of polymerase usage values

Used in chapter 4 and 5. Polymerase usage on the top and the bottom strands were determined as described previously (Daigaku *et al*., 2015). For each strand, Polδ usage was determined using $Usage\delta_i = \delta_i / (\delta_i + \varepsilon_i)$, where $Usage\delta_i$ represents Polδ usage at position i, and δ_i and ε_i stand for normalised count of mapped rNMPs incorporated by Polδ and Polε at position i, respectively. Similarly, for each strand, Polε usage was determined using $Usage\varepsilon_i = \varepsilon_i / (\delta_i + \varepsilon_i)$, where $Usage\varepsilon_i$ represents Polε usage at position i, and δ_i and ε_i stand for normalised count of mapped ribonucleotides incorporated by Polδ and Polε at position i, respectively.

Estimation of origin efficiencies from polymerase usage values

Used in chapter 4 and 5. Origin efficiencies were estimated as described previously (Daigaku *et al*., 2015). Briefly, differentials of polymerase usage values were calculated using $Diff_i = PT_i – PT_{i+1}$, where $Diff_i$ represents differential value at position i, and $PT_i – PT_{i+1}$ stand for smoothed polymerase track values at positions i and i+1, respectively. Differential values were smoothed by simple moving average (window of 3 bins). Differential values of Polδ on the top strand and Polε on the bottom strand were inverted. Negative and positive differential values indicated sites of replication initiation and termination, respectively. Differentials were averaged and smoothed by simple moving average of 3. Then, positive peaks (indicating sharp inclinations in the data) were selected. Differential peaks containing two or more distinct maxima separated by at least 4 bins were treated as unresolved independent peaks. Peaks with maxima bellow 30th quantile were disregarded. Each independent differential peak represented an origin of replication, the efficiency of which was estimated as the sum of its values.

Other next generation sequencing data

RNA-Seq data were retrieved from Marguerat *et al*., 2012. Start and end positions of listed genes were corrected to include 5’ UTR and 3’ UTR regions using current genomic annotation (Lock *et al*., 2019). Raw native elongation transcript sequencing (Net-Seq) data (Wery *et al*., 2018) were retrieved from the NCBI Gene Expression Omnibus repository (accession number GEO: GSE72382) and mapped onto the *S. pombe* reference genome using Bowtie2 (Langmead & Salzberg, 2012). Net-Seq

density values represent log_2-transformed raw counts. Net-Seq densities were binned into 100 bp-bins. Strand specific DNA:RNA immunoprecipitation followed by sequencing (DRIPc-Seq) data (Hartono *et al*., 2018) were retrieved from the NCBI Gene Expression Omnibus repository (accession number GEO: GSE101086). DRIPc-Seq score values on the top and the bottom DNA strands were binned into 300bp bins. Binned data were smoothed by simple moving average with 3-bin window.

Tables

Table 1 – List of strains

ID	Genotype	Origin	chapter
RZ07	*h- ade6-704 leu1-32 ura4-D18*	Stock	5
RZ08	*h+ ade6-704 leu1-32 ura4-D18*	Stock	5
RZ42	*h- ade6-704 leu1-32 ura4-D18 I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6-cdm1-LoxM3]*	This study	2
RZ47	*h- ade6-704 leu1-32*	Stock	2
RZ93	*h- ade6-704 leu1-32 ura4-D18 I-3325162:[LoxP-cdc1-cdc27-kanR-cdc6-cdm1-LoxM3] I-4734015:[LoxP-cdc1-cdc27-natR-cdc6-cdm1-LoxM3] I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6-cdm1-LoxM3]*	This study	2
655	*h- ade6-? leu1-32 ura4-D18 rnh201::kanR cdc20M630F*	Stock	2
856	*h- ade6-704 leu1-32 ura4-D18 rnh201::kanR cdc6L591G*	Stock	2
1141	*h- ade6-704 leu1-32 ura4-D18 rnh201::kanR pol1L850F*	Stock	2
RZ57	*h- ade6-704 leu1-32 ura4-D18 I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6L591G-cdm1-LoxM3] rnh201::hygR cdc6L591G*	This study	2
RZ62	*h- ade6-704 leu1-32 ura4-D18 I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6-cdm1-LoxM3] rnh201::hygR cdc20M630F*	This study	2
RZ68	*h- ade6-704 leu1-32 ura4-D18 I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6-cdm1-LoxM3] rnh201::hygR pol1L850F*	This study	2
RZ112	*h- ade6-704 leu1-32 ura4-D18 I-3325162:[LoxP-cdc1-cdc27-kanR-cdc6L591G-cdm1-LoxM3] I-4734015:[LoxP-cdc1-cdc27-natR-cdc6L591G-cdm1-LoxM3] I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6L591G-cdm1-LoxM3] rnh201::hygR cdc6L591G*	This study	2
RZ116	*h- ade6-704 leu1-32 ura4-D18 I-3325162:[LoxP-cdc1-cdc27-kanR-cdc6-cdm1-LoxM3] I-4734015:[LoxP-cdc1-cdc27-natR-cdc6-cdm1-LoxM3] I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6-cdm1-LoxM3] rnh201::hygR pol1L850F*	This study	2
RZ118	*h- ade6-704 leu1-32 ura4-D18 I-3325162:[LoxP-cdc1-cdc27-kanR-cdc6-cdm1-LoxM3] I-4734015:[LoxP-cdc1-cdc27-natR-cdc6-cdm1-LoxM3] I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6-cdm1-LoxM3] rnh201::hygR cdc20M630F*	This study	2
RZ121	*h- ade6-704 leu1-32 ura4-D18 cdm1::natR*	This study	5
RZ122	*h+ ade6-704 leu1-32 ura4-D18 cdm1::natR*	This study	5
RZ123	*h- ade6-704 leu1-32 ura4-D18 rad3::natR*	Stock	5
RZ147	*h- ade6-704 leu1-32 ura4-D18 cdc2asM17*	Stock	5
RZ150	*h- ade6-704 leu1-32 ura4-D18 loxP-rnh201-RED:hphMX6-loxM3 cdc6L591G*	This study	1
RZ151		This study	1
RZ152		This study	1
RZ153	*h- ade6-704 leu1-32 ura4-D18 loxP-rnh201-RED:hphMX6-loxM3 cdc20M630F*	This study	1
RZ154		This study	1
RZ155		This study	1

RZ159	*h- ade6-704 leu1-32 ura4-D18 cdm1::natR loxP-rnh201-RED:hphMX6-loxM3 cdc6L591G*	This study	5
RZ160	*h- ade6-704 leu1-32 ura4-D18 cdm1::natR loxP-rnh201-RED:hphMX6-loxM3 cdc20M630F*	This study	5
RZ201	*h+ ade6-704 leu1-32 ura4-D18 kanR:cdm1:nGFP*	This study	5
RZ219	*h+ ade6-704 leu1-32 ura4-D18 kanR:cdm1:nGFP spd1::hph*	This study	5
RZ220	*h+ ade6-704 leu1-32 ura4-D18 kanR:cdm1:nGFP spd1::hph rad3::natR*	This study	5
RZ221	*h+ ade6-704 leu1-32 ura4-D18 kanR:cdm1:nGFP spd1::hph tel1::ura4*	This study	5
RZ222	*h+ ade6-704 leu1-32 kanR:cdm1:nGFP spd1::hph cdt2::natR*	This study	5
RZ259	*h- ade6-704 ura4-D18 leu1-32 LoxP-rnh201-RED-hphMX6-LoxM3 ARS:nmt1:rnhA:LEU2 cdc6L591M*	This study	1, 3
RZ260		This study	1, 3
RZ261	*h- ade6-704 ura4-D18 leu1-32 LoxP-rnh201-RED-hphMX6-LoxM3 ARS:nmt1:rnhA:LEU2 sen1::NatR cdc6L591M*	This study	1, 3
RZ262		This study	1, 3
RZ263	*h- ade6-704 ura4-D18 leu1-32 LoxP-rnh201-RED-hphMX6-LoxM3 ARS:nmt1:rnhA:LEU2 dbl8::KanR cdc6L591M*	This study	1, 3
RZ264		This study	1, 3
RZ265	*h- ade6-704 ura4-D18 leu1-32 LoxP-rnh201-RED-hphMX6-LoxM3 ARS:nmt1:rnhA:LEU2 sen1::NatR dbl8::KanR cdc6L591M*	This study	1, 3
RZ266		This study	1, 3
RZ267	*h- ade6-704 ura4-D18 leu1-32 LoxP-rnh201-RED-hphMX6-LoxM3 ARS:nmt1:rnhA:LEU2 cdc20M630F*	This study	1, 3
RZ268		This study	1, 3
RZ269	*h- ade6-704 ura4-D18 leu1-32 LoxP-rnh201-RED-hphMX6-LoxM3 ARS:nmt1:rnhA:LEU2 sen1::NatR cdc20M630F*	This study	1, 3
RZ270		This study	1, 3
RZ271	*h- ade6-704 ura4-D18 leu1-32 LoxP-rnh201-RED-hphMX6-LoxM3 ARS:nmt1:rnhA:LEU2 dbl8::KanR cdc20M630F*	This study	1, 3
RZ272		This study	1, 3
RZ273	*h- ade6-704 ura4-D18 leu1-32 LoxP-rnh201-RED-hphMX6-LoxM3 ARS:nmt1:rnhA:LEU2 sen1::NatR dbl8::KanR cdc20M630F*	This study	1, 3
RZ274		This study	1, 3
RZ280	*h+ ade6-704 leu1-32 ura4-D18 kanR:cdm1R26A:nGFP*	This study	5
RZ282	*h- leu1-32 cdc24-m38 pREP1*	This study	5
RZ283	*h- leu1-32 cdc24-m38 pREP1_cdc24*	This study	5
RZ284	*h- leu1-32 cdc24-m38 pREP1_cdm1*	This study	5
RZ285	*h- leu1-32 cdc24-m38 pREP1_cdm1-nGFP*	This study	5
RZ286	*h- leu1-32 cdc24-m38 pREP1_cdm1-nmEos3.2*	This study	5
RZ287	*h- leu1-32 cdc27-p11 pREP1*	This study	5
RZ288	*h- leu1-32 cdc27-p11 pREP1_cdc27*	This study	5
RZ289	*h- leu1-32 cdc27-p11 pREP1_cdm1*	This study	5
RZ290	*h- leu1-32 cdc27-p11 pREP1_cdm1-nGFP*	This study	5
RZ291	*h- leu1-32 cdc27-p11 pREP1_cdm1-nmEos3.2*	This study	5
RZ310	*h- ade6-704 loxP-rnh201-RED-hphMX6-loxM3 leu1+:nmt81xTetR"off"-2xFLAG-clr4-cdd I-3325162::LoxP-kanR-tetO7-LoxM3 ura4::[4xTetO-ade6+] cdc6L591G*	This study	4
RZ313	*h- ade6-704 loxP-rnh201-RED-hphMX6-loxM3 leu1+:nmt81xTetR"off"-2xFLAG-clr4-cdd I-3325162::LoxP-kanR-tetO7-LoxM3 ura4::[4xTetO-ade6+] cdc20M630F*	This study	4
RZ316	*h- ade6-704 leu1-32 ura4-D18 ARS:nmt1:RnhA:LEU2+*	This study	3
RZ317	*h- ade6-704 leu1-32 ura4-D18 sen1::natR ARS:nmt1:RnhA:LEU2+*	This study	3
RZ318	*h- ade6-704 leu1-32 ura4-D18 dbl8::kanR ARS:nmt1:RnhA:LEU2+*	This study	3
RZ319	*h- ade6-704 leu1-32 ura4-D18 dbl8::kanR sen1::natR ARS:nmt1:RnhA:LEU2+*	This study	3
RZ331	*h- ade6-704 leu1-32 cdc2asM17*	This study	2
RZ332	*h- ade6-704 leu1-32 cdc2asM17 I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6-cdm1-LoxM3]*	This study	2
RZ333	*h- ade6-704 leu1-32 cdc2asM17 I-3325162:[LoxP-cdc1-cdc27-kanR-cdc6-cdm1-LoxM3] I-4734015:[LoxP-cdc1-cdc27-natR-cdc6-cdm1-LoxM3] I-5230932:[LoxP-cdc1-cdc27-ura4-cdc6-cdm1-LoxM3]*	This study	2

RZ340	*h+ ade6-704 leu1-32 ura4-D18 I-3325162:LoxP-kanR-tetO7-LoxM3 II-1389186:LoxM3-tetO7-natR-LoxP III-1609353:LoxP-ura4-tetO7-LoxM3 I-5230932:LoxM3-tetR-clr4cdd-LoxP loxP-rnh201-RED-hphMX6-loxM3 cdc6L591G*	This study	4
RZ341		This study	4
RZ342	*h+ ade6-704 leu1-32 ura4-D18 I-3325162:LoxP-kanR-tetO7-LoxM3 II-1389186:LoxM3-tetO7-natR-LoxP III-1609353:LoxP-ura4-tetO7-LoxM3 I-5230932:LoxM3-tetR-clr4cdd-LoxP loxP-rnh201-RED-hphMX6-loxM3 cdc20M630F*	This study	4
RZ343		This study	4
RZ344	*h+ ade6-704 leu1-32 ura4-D18 I-3325162:LoxP-kanR-tetO7-LoxM3 II-1389186:LoxM3-tetO7-natR-LoxP III-1609353:LoxP-ura4-tetO7-LoxM3 I-5230932:LoxM3-tetR-clr4H410Kcdd-LoxP loxP-rnh201-RED-hphMX6-loxM3 cdc6L591G*	This study	4
RZ345		This study	4
RZ346	*h+ ade6-704 leu1-32 ura4-D18 I-3325162:LoxP-kanR-tetO7-LoxM3 II-1389186:LoxM3-tetO7-natR-LoxP III-1609353:LoxP-ura4-tetO7-LoxM3 I-5230932:LoxM3-tetR-clr4H410Kcdd-LoxP loxP-rnh201-RED-hphMX6-loxM3 cdc20M630F*	This study	4
RZ347		This study	4
RZ348	*h+ ade6-704 leu1-32 ura4-D18 I-3325162:LoxP-kanR-tetO7-LoxM3 II-1389186:LoxM3-tetO7-natR-LoxP III-1609353:LoxP-ura4-tetO7-LoxM3 dfp1:tetR:bleoR loxP-rnh201-RED-hphMX6-loxM3 cdc6L591G*	This study	4
RZ349		This study	4
RZ350	*h+ ade6-704 leu1-32 ura4-D18 I-3325162:LoxP-kanR-tetO7-LoxM3 II-1389186:LoxM3-tetO7-natR-LoxP III-1609353:LoxP-ura4-tetO7-LoxM3 dfp1:tetR:bleoR loxP-rnh201-RED-hphMX6-loxM3 cdc20M630F*	This study	4
RZ351		This study	4

Table 2 – List of plasmids

ID	Annotation	Origin	chapter
pRZ01	pAW8_cdc1:cdc27:ura4	This study	2
pRZ02	pAW8_cdc1:cdc27:ura4:cdc6:cdm1	This study	2
pRZ03	pAW8_cdc1:cdc27:ura4:cdc6L591G:cdm1	This study	2
pRZ04	pAW8_cdc1:cdc27:natR:cdc6:cdm1	This study	2
pRZ05	pAW8_cdc1:cdc27:natR:cdc6L591G:cdm1	This study	2
pRZ06	pAW8_cdc1:cdc27:kanMX6:cdc6:cdm1	This study	2
pRZ07	pAW8_cdc1:cdc27:kanMX6:cdc6L591G:cdm1	This study	2
pRZ08	pAW8_natR:tetO7	This study	4
pRZ09	pAW8_ura4:tetO7	This study	4
pRZ10	pAW8_kanR:tetO7	This study	4
pRZ11	pAW8_kanR-cdm1(5'UTR)-nGFP	This study	5
pRZ12	pAW8_kanR:cdm1:nmEos3.2	This study	5
pRZ13	pREP1_cdm1	This study	5
pRZ14	pREP1_cdm1-nGFP	This study	5
pRZ15	pREP1_cdm1-nmEos3.2	This study	5
pRZ16	pREP1_cdc24	This study	5
pRZ17	pREP1_cdc27	This study	5
pRZ18	pAW8_nmt1-tetR(off)-2xFLAG-clr4H410Kcdd	This study	4
pRZ19	pAW8_nmt1-tetR(off)-2xFLAG-clr4	This study	4

Table 3 – List of qPCR primers

Primer ID	Sequence	Target gene	Origin	Chapter(s)
RZ67	CAACTATCCTTCCTCAACAG	*cdc1⁺*	This study	2
RZ68	GCTAGTAGCCAACACAAAATG			
RZ69	CGTTCACGATTCTGAAGATG	*cdc27⁺*	This study	2
RZ70	ATAATTTCCTGAGGTTCGT			
RZ75	CCTGCAATAAATCCTGAGAAG	*cdc6⁺*	This study	2
RZ76	CATTGTCAGTAACACCAAAC			
RZ81	TTCATTCTAGTACCGCAGTG	*cdm1⁺*	This study	2

RZ82	TGTGGGATTGACTTGAATTAC			
RZ87	TCCTCATGCTATCATGCGTCTT	*act1⁺*	(Převorovský et al., 2016)	2
RZ88	CCACGCTCCATGAGAATCTTC			
RZ185	CTGTTGAACAAGTCTGGAAAG	kanR	This study	4
RZ186	ACTCGTCCAACATCAATACAAC			
RZ283	GGCAGGGCATACTCATGTAG	natR	This study	4
RZ284	GGAGGTCACCAACGTCAAC			
RZ279	TGGCTACTGGTTCCTACAC	*ura4⁺*	This study	4
RZ280	CTTTAACATCCAAGCCGATAC			

Table 4 – Number of uniquely mapped reads in presented Pu-Seq experiments

#	Strain ID	Chapter(s)	Cultivation Conditions	Pu-Seq protocol	Uniquely Mapped Reads	
					Top strand	Bottom strand
1	RZ150	1	YES, 30°C	New	11706065	11463269
2	RZ151	1	YES, 30°C	New	14265002	13968934
3	RZ152	1	YES, 30°C	New	10603251	10338344
4	RZ153	1	YES, 30°C	New	10976477	11017589
5	RZ154	1	YES, 30°C	New	11995941	12039317
6	RZ155	1	YES, 30°C	New	11847866	11897286
7	RZ259	1, 3	EMM /wo thiamine, 30°C	New	12137811	12057953
8	RZ259	1, 3	EMM /w thiamine, 30°C	New	12236321	12172004
9	RZ260	1, 3	EMM /wo thiamine, 30°C	New	6886877	6804289
10	RZ260	1, 3	EMM /w thiamine, 30°C	New	5132405	5106112
11	RZ261	1, 3	EMM /wo thiamine, 30°C	New	13138124	12997397
12	RZ261	1, 3	EMM /w thiamine, 30°C	New	12329119	12258944
13	RZ262	1, 3	EMM /wo thiamine, 30°C	New	8120680	8086333
14	RZ262	1, 3	EMM /w thiamine, 30°C	New	6660877	6633520
15	RZ263	1, 3	EMM /wo thiamine, 30°C	New	15176411	14813733
16	RZ263	1, 3	EMM /w thiamine, 30°C	New	9843129	9706805
17	RZ264	1, 3	EMM /wo thiamine, 30°C	New	11134750	10982996
18	RZ264	1, 3	EMM /w thiamine, 30°C	New	6781886	6724002
19	RZ265	1, 3	EMM /wo thiamine, 30°C	New	17194954	16858078
20	RZ265	1, 3	EMM /w thiamine, 30°C	New	14053759	13841294
21	RZ266	1, 3	EMM /wo thiamine, 30°C	New	10046578	9939481
22	RZ266	1, 3	EMM /w thiamine, 30°C	New	8329676	8273365
23	RZ267	1, 3	EMM /wo thiamine, 30°C	New	10509362	10506913
24	RZ267	1, 3	EMM /w thiamine, 30°C	New	13242034	13264423
25	RZ268	1, 3	EMM /wo thiamine, 30°C	New	7296141	7271488
26	RZ268	1, 3	EMM /w thiamine, 30°C	New	8576583	8519311
27	RZ269	1, 3	EMM /wo thiamine, 30°C	New	11798696	11796380
28	RZ269	1, 3	EMM /w thiamine, 30°C	New	15986616	15913589
29	RZ270	1, 3	EMM /wo thiamine, 30°C	New	8515177	8480432
30	RZ270	1, 3	EMM /w thiamine, 30°C	New	5335476	5306692
31	RZ271	1, 3	EMM /wo thiamine, 30°C	New	10751608	10923119
32	RZ271	1, 3	EMM /w thiamine, 30°C	New	11523790	11646472
33	RZ272	1, 3	EMM /wo thiamine, 30°C	New	7476165	7548913
34	RZ272	1, 3	EMM /w thiamine, 30°C	New	6081183	6113185
35	RZ273	1, 3	EMM /wo thiamine, 30°C	New	9550640	9544479
36	RZ273	1, 3	EMM /w thiamine, 30°C	New	14497203	14404915
37	RZ274	1, 3	EMM /wo thiamine, 30°C	New	6998435	6978342
38	RZ274	1, 3	EMM /w thiamine, 30°C	New	7116437	7079629
39	655	2	YES, 30°C	Original, tech1	7104619	7096152
40	655	2	YES, 30°C	Original, tech2	10144678	10129860
41	655	2	YES, 30°C	New	11507861	11439288

42	856	2	YES, 30°C	Original, tech1	6025330	5926320
43	856	2	YES, 30°C	Original, tech2	13541100	13316893
44	856	2	YES, 30°C	New	9030680	8879936
45	1141	2	YES, 30°C	Original, tech1	7363234	7318476
46	1141	2	YES, 30°C	Original, tech2	12005387	11933832
47	1141	2	YES, 30°C	New	8055797	8006914
48	RZ57	2	YES, 30°C	Original, tech1	7339726	7196858
49	RZ57	2	YES, 30°C	Original, tech2	12257107	12029000
50	RZ57	2	YES, 30°C	New	9979134	9820222
51	RZ62	2	YES, 30°C	Original, tech1	7505695	7468044
52	RZ62	2	YES, 30°C	Original, tech2	10737538	10679413
53	RZ62	2	YES, 30°C	New	8479955	8441774
54	RZ68	2	YES, 30°C	Original, tech1	8438787	8375584
55	RZ68	2	YES, 30°C	Original, tech2	13438109	13337833
56	RZ68	2	YES, 30°C	New	11392885	11295575
57	RZ112	2	YES, 30°C	Original, tech1	7072644	6989621
58	RZ112	2	YES, 30°C	Original, tech2	13519861	13355765
59	RZ112	2	YES, 30°C	New	12378277	12212269
60	RZ116	2	YES, 30°C	Original, tech1	7927477	7864342
61	RZ116	2	YES, 30°C	Original, tech2	14121460	13992024
62	RZ116	2	YES, 30°C	New	8597613	8520711
63	RZ118	2	YES, 30°C	Original, tech1	7594174	7621479
64	RZ118	2	YES, 30°C	Original, tech2	11621178	11660226
65	RZ118	2	YES, 30°C	New	8728257	8769787
66	RZ310	4	EMM /wo ahTET, 30°C	New	13204901	12935505
67	RZ310	4	EMM /w ahTET, 30°C	New	14588257	14340504
68	RZ313	4	EMM /wo ahTET, 30°C	New	11933655	11975488
69	RZ313	4	EMM /w ahTET, 30°C	New	12098718	12143117
70	RZ150	5	YES, 30°C	Original	6154485	6061563
71	RZ153	5	YES, 30°C	Original	8169112	8223990
72	RZ159	5	YES, 30°C	Original	2529781	2473708
73	RZ160	5	YES, 30°C	Original	7861260	7909076

Table 5 – List of genes causing replication fork stalling in sen1Δ mutants

	Systematic ID	Gene	Product
1	SPAC9.09	*met26⁺*	Homocysteine methyltransferase Met26
2	SPAC1002.13c	*psu1⁺*	Cell wall beta-glucosidase Psu1 (predicted)
3	SPAC1071.10c	*pma1⁺*	Plasma membrane P-type proton exporting ATPase, P3-type Pma1
4	SPAPB15E9.01c	*pfl2⁺*	Cell surface glycoprotein, flocculin Pfl2
5	SPAC27E2.11c	*NA*	*Schizosaccharomyces* specific protein
6	SPBC839.15c	*tef103⁺*	Translation elongation factor EF-1 alpha Ef1a-c
7	SPBC3D6.02	*but2⁺*	But2 family protein But2, similar to cell surface molecules
8	SPBC32F12.11	*tdh1⁺*	Glyceraldehyde-3-phosphate dehydrogenase Tdh1
9	SPCC622.09	*htb1⁺*	Histone H2B Htb1

Table 6 - List of genes causing replication fork stalling in dbl8Δ mutants

	Systematic ID	Gene	Product
1	SPAC12G12.04	*mcp60⁺*	Mitochondrial heat shock protein Hsp60/Mcp60
2	SPAC24H6.07	*rps901⁺*	40S ribosomal protein S9

3	SPAC22G7.06c	*ura1⁺*	Carbamoyl-phosphate synthase (glutamine hydrolyzing), aspartate carbamoyltransferase Ura1
4	SPAC18G6.09c	*edc1⁺*	Dcp2-Dcp1 mRNA-decapping complex subunit Edc1
5	SPAC4H3.10c	*pyk1⁺*	Pyruvate kinase
6	SPAC926.04c	*hsp90⁺*	Hsp90 chaperone
7	SPNCRNA.53	*prl53⁺*	Non-coding RNA, poly(A)-bearing (predicted)
8	SPAC26H5.10c	*tif51⁺*	Translation elongation and termination factor eIF5A (predicted)
9	SPAC9E9.09c	*atd1⁺*	Aldehyde dehydrogenase (predicted)
10	SPAC17C9.03	*tif471⁺*	Translation initiation factor eIF4G
11	SPAPB8E5.06c	*rpl302⁺*	60S ribosomal protein L3
12	SPAC664.11	*ssc1⁺*	Mitochondrial (2Fe-2S) cluster assembly chaperone Ssc1
13	SPAC6B12.15	*cpc2⁺*	Ribosome-associated signalling scaffold, receptor of activated C kinase (RACK1) ortholog Cpc2
14	SPAC26A3.04	*rpl2002⁺*	60S ribosomal protein L20 (predicted)
15	SPBC359.03c	*aat1⁺*	Plasma membrane amino acid transmembrane transporter Aat1
16	SPBP35G2.16c	*ecl2⁺*	Extender of chronological lifespan protein Ecl2
17	SPBP8B7.16c	*dpb2⁺*	ATP-dependent RNA helicase Dbp2
18	SPBC1685.10	*rps27⁺*	40S ribosomal protein S27 (predicted)
19	SPBC119.10	*asn1⁺*	Asparagine synthetase
20	SPBC1734.11	*mas5⁺*	Hsp40 family DNAJ domain protein Mas5 (predicted)
21	SPSNORNA.21	*snoU14⁺*	Small nucleolar RNA U14
22	SPBC19C2.07	*fba1⁺*	Fructose-bisphosphate aldolase Fba1
23	SPBC29A3.04	*rpl8⁺*	60S ribosomal protein L7a/L8 (predicted)
24	SPBC2G5.05	*N/A*	Transketolase (predicted)
25	SPCC794.09c	*tef101⁺*	Translation elongation factor EF-1 alpha Ef1a-a
26	SPCC1795.11	*sum3⁺*	Translation initiation RNA helicase Sum3
27	SPCC622.08c	*hta1⁺*	Histone H2A alpha
28	SPCC13B11.01	*adh1⁺*	Alcohol dehydrogenase Adh1
29	SPCC1739.13	*ssa2⁺*	Hsp70 family heat shock protein Ssa2
30	SPCC576.07	*ret3⁺*	Coatomer zeta subunit (predicted)

Table 7 - List of genes causing replication fork stalling in sen1Δdbl8Δ mutants only

	Systematic ID	Gene	Product
1	SPAC13G6.10c	*asl1⁺*	Cell wall protein Asl1, predicted O-glucosyl hydrolase
2	SPAC12G12.13c	*cid14⁺*	TRAMP complex poly(A) polymerase subunit Cid14
3	SPAC24H6.08	*NA*	*Schizosaccharomyces* specific protein
4	SPAC4G8.06c	*trm12⁺*	tRNA 4-demethylwyosine alpha-amino-alpha-carboxypropyltransferase Trm12 (predicted)
5	SPAC222.12c	*atp2⁺*	F1-FO ATP synthase beta subunit Atp2
6	SPAC56F8.16	*esc1⁺*	DNA-binding transcription factor Esc1 (predicted)
7	SPAC10F6.01c	*sir1⁺*	Sulfite reductase beta subunit Sir1
8	SPAC1420.01c	*NA*	GATA-like domain protein (predicted)
9	SPAC1565.08	*cdc48⁺*	AAA family ATPase involved in ubiquitin-mediated protein degradation Cdc48

10	SPAC6F12.10c	*ade3*$^{+}$	Phosphoribosylformylglycinamidine synthase Ade3
11	SPAC19E9.03	*pas1*$^{+}$	Cyclin Pas1
12	SPAC20G8.06	*not1*$^{+}$	CCR4-Not complex scaffold subunit 1
13	SPAC9.03c	*brr2*$^{+}$	U5 snRNP complex subunit Brr2
14	SPAC110.01	*ppk1*$^{+}$	Serine/threonine protein kinase Ppk1 (predicted)
15	SPAC1783.08c	*rpl1502*$^{+}$	60S ribosomal protein L15b (predicted)
16	SPAC4G9.03	*adk1*$^{+}$	Adenylate kinase Adk1
17	SPAC4G9.08c	*rpc2*$^{+}$	DNA-directed RNA polymerase III complex subunit Rpc2
18	SPAC6B12.12	*tom70*$^{+}$	Mitochondrial TOM complex subunit Tom70 (predicted)
19	SPAC19A8.15	*trp2*$^{+}$	Tryptophan synthase (predicted)
20	SPAC4A8.16c	*tif303*$^{+}$	Translation initiation factor eIF3c
21	SPAC1805.11c	*rps2602*$^{+}$	40S ribosomal protein S26 (predicted)
22	SPAC8F11.10c	*pvg1*$^{+}$	Golgi pyruvyltransferase Pvg1
23	SPACUNK4.16c	*tps3*$^{+}$	Alpha,alpha-trehalose-phosphate synthase (predicted)
24	SPAC513.07	*NA*$^{+}$	Flavonol reductase/cinnamoyl-CoA reductase family
25	SPAPB24D3.09c	*pdr1*$^{+}$	ABC transmembrane transporter Pdr1
26	SPAC31G5.11	*pac2*$^{+}$	cAMP-independent regulatory protein Pac2
27	SPAC24C9.06c	*aco1*$^{+}$	Aconitate hydratase Aco1 (predicted)
28	SPAC24C9.12c	*shm1*$^{+}$	Serine hydroxymethyltransferase Shm1 (predicted)
29	SPAC3G9.12	*peg1*$^{+}$	CLASP family microtubule-associated protein
30	SPAPB1E7.07	*glt1*$^{+}$	Glutamate synthase Glt1
31	SPAC3A11.07	*nde1*$^{+}$	External mitochondrial NADH dehydrogenase (ubiquinone) Nde1/Nde2 (predicted)
32	SPAP7G5.04c	*lys1*$^{+}$	Aminoadipate-semialdehyde dehydrogenase
33	SPAC926.09c	*fas1*$^{+}$	Fatty acid synthase beta subunit Fas1
34	SPAC27E2.03c	*NA*	Obg-like ATPase, human OLA1 ortholog (predicted)
35	SPAC25B8.12c	*NA*	HAD superfamily hydrolase, unknown role
36	SPAC25B8.13c	*isp7*$^{+}$	2-OG-Fe(II) oxygenase superfamily protein
37	SPAC25G10.08	*tif302*$^{+}$	Translation initiation factor eIF3b (p84)
38	SPAC23D3.12	*NA*	Plasma membrane inorganic phosphate transmembrane transporter (predicted)
39	SPAC9E9.04	*NA*	Bcap family homolog, implicated in vesicle-mediated transport (predicted)
40	SPAC9E9.13	*wos2*$^{+}$	p23 homolog, predicted co-chaperone Wos2
41	SPAC17C9.07	*alg8*$^{+}$	Dolichyl pyrophosphate Glc1Man9GlcNAc2 alpha-1,3-glucosyltransferase Alg8 (predicted)
42	SPAC1250.01	*snf21*$^{+}$	RSC-type complex ATP-dependent DNA helicase Snf21
43	SPAC26F1.13c	*lrs1*$^{+}$	Cytoplasmic leucine-tRNA ligase Lrs1 (predicted)
44	SPAC19D5.04	*ptr1*$^{+}$	HECT-type ubiquitin-protein ligase E3 Ptr1
45	SPAC14C4.11	*vtc2*$^{+}$	Vacuolar transporter chaperone (VTC) complex polyphosphate synthetase subunit Vtc2/3 (predicted)
46	SPAC29B12.08	*clr5*$^{+}$	Clr5 protein
47	SPBC839.04	*rpl803*$^{+}$	60S ribosomal protein L8/L2 (predicted)
48	SPBC947.03c	*naa38*$^{+}$	NatC N-acetyltransferase non-catalytic Sm-like domain subunit Naa38 (predicted)
49	SPBPJ4664.06	*gpt1*$^{+}$	UDP-glucose-glycoprotein glucosyltransferase Gpt1

50	SPBC530.10c	*anc1⁺*	Mitochondrial carrier, ATP:ADP antiporter Anc1
51	SPBC530.15c	*NA*	Spermidine family transmembrane transporter (predicted)
52	SPBC1709.06	*dus2⁺*	tRNA dihydrouridine synthase Dus2 (predicted)
53	SPBC1709.07	*erg27⁺*	3-keto sterol reductase Erg27 (predicted)
54	SPBC1709.08	*cft1⁺*	mRNA cleavage and polyadenylation specificity factor complex, WD repeat protein Cft1
55	SPBC725.02	*mpr1⁺*	Histidine-containing response regulator phosphotransferase Mpr1
56	SPBC8D2.20c	*sec31⁺*	COPII-coated vesicle component Sec31 (predicted)
57	SPBC32H8.12c	*act1⁺*	Actin Act1
58	SPBC29B5.01	*atf1⁺*	DNA-binding transcription factor, Atf-CREB family Atf1
59	SPBC28F2.12	*rpb1⁺*	RNA polymerase II large subunit Rpb1
60	SPBC2F12.03c	*ebs1⁺*	EST1 family nonsense-mediated mRNA decay (NMD) pathway protein Ebs1
61	SPBC18H10.04c	*sce3⁺*	Translation initiation factor (predicted)
62	SPBC3H7.02	*NA*	Sulfate transmembrane transporter (predicted)
63	SPBC16E9.13	*ksp1⁺*	Serine/threonine protein kinase Ksp1 (predicted)
64	SPBC1E8.05	*NA*	Conserved fungal cell surface protein, Kre9/Knh1 family
65	SPBP23A10.04	*apc2⁺*	Anaphase-promoting complex cullin family subunit Apc2
66	SPBC17G9.11c	*pyr1⁺*	Pyruvate carboxylase Pyr1
67	SPBC14C8.10	*mrpl24⁺*	Mitochondrial ribosomal protein subunit L28 (predicted)
68	SPBC16H5.02	*pfk1⁺*	6-phosphofructokinase pfk1
69	SPBC19G7.05c	*bgs1⁺*	Primary septum and spore wall linear 1,3-beta-glucan synthase catalytic subunit Bgs1
70	SPBC12C2.10c	*pst1⁺*	Clr6 histone deacetylase complex subunit Pst1
71	SPBC29A10.07	*pom152⁺*	Nucleoporin Pom152
72	SPBC1826.01c	*mot1⁺*	TATA-binding protein-associated transcription initiation factor Mot1 (predicted)
73	SPBC3E7.16c	*leu3⁺*	2-isopropylmalate synthase Leu3
74	SPBC1703.07	*acl1⁺*	ATP citrate synthase subunit 1 (predicted)
75	SPBC1703.13c	*NA*	Mitochondrial carrier, inorganic phosphate/copper (predicted)
76	SPBC2A9.04c	*san1⁺*	Sir antagonist, ubiquitin-protein ligase E3
77	SPBC609.01	*NA*	Cytoplasmic P body 3'-5'-exoribonuclease, Dis3L2-related (predicted)
78	SPBC776.09	*ste13⁺*	ATP-dependent RNA helicase Ste13/Dhh1
79	SPBC17D11.05	*tif301⁺*	Translation initiation factor eIF3a
80	SPBC1718.07c	*zfs1⁺*	Zf-CCCH tandem zinc finger protein, human Tristetraprolin homolog Zfs1, involved in mRNA catabolism
81	SPBP8B7.05c	*nce103⁺*	Carbonic anhydrase (predicted)
82	SPBC21C3.13	*rps1901⁺*	40S ribosomal protein S19 (predicted)
83	SPBC1604.05	*pgi1⁺*	Glucose-6-phosphate isomerase (predicted)
84	SPBC14F5.04c	*pgk1⁺*	Phosphoglycerate kinase Pgk1 (predicted)
85	SPBC1289.03c	*spi1⁺*	Ran GTPase Spi1
86	SPCC794.12c	*mae2⁺*	Malic enzyme, malate dehydrogenase (oxaloacetate decarboxylating), Mae2
87	SPCC553.10	*NA*	Conserved fungal cell surface protein, Kre9/Knh1 family (predicted)
88	SPCC736.15	*pil1⁺*	Eisosome BAR domain protein Pil1
89	SPCC970.04c	*mob2⁺*	Protein kinase activator Mob2

90	SPCC962.04	*rps1201⁺*	40S ribosomal protein S12 (predicted)
91	SPCC1672.02c	*sap1⁺*	Switch-activating protein Sap1
92	SPCC1672.03c	*gud1⁺*	Guanine deaminase Gud1 (predicted)
93	SPCC1672.11c	*NA*	P-type ATPase P5 type (predicted)
94	SPCC1183.11	*msy1⁺*	MS calcium ion channel protein Msy1
95	SPCC16C4.02c	*NA*	Armadillo-type fold protein, DUF1941 family protein, human neurochondrin ortholog, implicated in signal transduction
96	SPCC16C4.09	*sts5⁺*	Cytoplasmic P body 3'-5'-exoribonuclease, Dis3L2-related (predicted)
97	SPCC14G10.04	*NA*	*Schizosaccharomyces* specific protein
98	SPCC1393.08	*fil1⁺*	DNA-binding transcription factor, zf-GATA type
99	SPCPB16A4.03c	*ade10⁺*	Bifunctional IMP cyclohydrolase/phosphoribosylaminoimidazolecarboxamideformyltransferase
100	SPCC162.08c	*nup211⁺*	Nucleoporin nup211
101	SPCC417.08	*tef3⁺*	Translation elongation factor eEF3
102	SPCC191.02c	*acs1⁺*	Acetyl-CoA ligase (predicted)
103	SPCC737.08	*mdn1⁺*	Midasin, Mdn1
104	SPCC18.03	*NA*	Shuttle craft like transcriptional repressor/ubiquitin-protein ligase E3 (predicted)
105	SPCC1739.01	*NA*	Zf-CCCH type zinc finger protein
106	SPCC1620.14c	*snf22⁺*	SWI/SNF ATP-dependent DNA helicase subunit Snf22
107	SPCC1840.02c	*bgs4⁺*	Cell wall and secondary septum 1,6 branched 1,3-beta-glucan synthase catalytic subunit Bgs4
108	SPCC965.04c	*yme1⁺*	Mitochondrial inner membrane i-AAA protease complex subunit Yme1 (predicted)

Chapter 1 – Establishment of a new protocol for preparation of polymerase usage sequencing libraries

Background 1

Polymerase usage sequencing (Pu-Seq) is a procedure that allows assessment of genome-wide activities of replicative polymerases Polε, Polδ and Polα (Daigaku *et al*., 2017; Keszthelyi *et al.*, 2015; Naiman *et al*., 2021; Zach & Carr, 2021). Pu-Seq methodology is based on stabilisation and subsequent detection of genomic ribonucleotides (rNMPs) misincorporated by mutated variants of replicative polymerases which exhibit reduced base-selectivity. Stabilisation of genomic rNMPs is ensured by disruption of RNase H2, an essential constituent of RER pathway (Kellner & Luke, 2020). Inactivation of RNase H2 is achieved by deletion of *rnh201* or, alternatively, by introduction of the separation of function allele *rnh201-RED*, which abrogates RER, but preserves RNase H2 function in resolution of longer R-loops (Daigaku *et al*., 2017; Naiman *et al*., 2021). Reduced base-selectivity of replicative polymerases is achieved by introduction of mutated alleles of respective polymerase catalytic subunits (Polε – *cdc20*^{M630F}, Polδ – *cdc6*^{L591G} or *cdc6*^{L591M}, and Polα – *pol1*^{L850F}) (Daigaku *et al*., 2017; Naiman *et al*., 2021).

The original protocol for preparation of Pu-Seq libraries comprises five main steps: 1) fragmentation and denaturation of rNMP-containing DNA by alkali treatment at 65°C; 2) size selection of 300-600bp single stranded DNA (ssDNA) fragments; 3) synthesis of the complementary DNA strands; 4) ligation of the Illumina adapters and 5) library enrichment and indexing (Daigaku *et al*., 2015; Keszthelyi *et al*., 2015).

Although the original procedure performs well, as demonstrated by numerous successfully prepared and analysed Pu-Seq libraries (Daigaku *et al*., 2015; Naiman *et al*., 2021; Zach & Carr, 2021), we reasoned that certain adjustments could improve Pu-Seq accuracy and, at the same time, reduce the cost and the time required for completion.

Design of the new Pu-Seq procedure was inspired by GLOE-Seq, a recently elaborated method designed to detect single-stranded DNA breaks (Sriramachandran *et al*., 2020), and the modified version of hydrolytic end sequencing (RHII-HydEn-seq),

an experimental procedure based on principles analogous to Pu-Seq (Donnianni *et al*., 2019; Zhou *et al*., 2019).

Unlike the original Pu-Seq procedure, the newly developed protocol does not involve size-selection, which, in theory, could be generating a yet uncharacterised bias. Additionally, the newly developed protocol introduces rNMP-specific ssDNA nicks by employing recombinant RNase H2 instead of a less specific alkali treatment. We argue that introduction of purified RNase H2 increases specificity of genomic rNMP detection, as was previously discussed (Zhou *et al*., 2019).

The following text summarises key concepts of the new Pu-Seq procedure, discusses the preliminary analysis of newly generated data and provides a detailed protocol, which provides sufficient guidance on the preparation of Pu-Seq libraries.

Results 1

Overview of the proposed Pu-Seq protocol

Genomic DNA is first treated with shrimp alkaline phosphatase (rSAP) which removes 5'- and 3'-phosphate groups from DNA ends. Implementation of this step was inspired by a recent study introducing RHII-HydEn-seq (Zhou *et al*., 2019) and ensures that random DNA breaks do not represent a substrate for subsequent DNA ligation reactions (Figure 1.1A; steps 1-3).

Next, DNA is treated with recombinant RNase H2, which introduces single-stranded DNA (ssDNA) breaks at positions of misincorporated ribonucleotides (rNMPs). Resulting DNA nicks carry 3'-OH and 5'-phosphate groups. Arguably, implementation of RNase H2 improves the specificity of rNMP detection. DNA is subsequently denatured at 95°C and ssDNA fragments incubated on ice to prevent re-annealing. Recovery of digested ssDNA was verified by agarose gel electrophoresis followed by acridine orange staining (Figure 1.1A and 1.1B; steps 4-5).

Recovered ssDNA fragments are subsequently ligated to Illumina adapters-1 using a technique elaborated in studies analysing low-quality ancient DNA and GLOE-Seq (Gansauge *et al*., 2017; Sriramachandran *et al*., 2020). Briefly, biotinylated adapter-1/splinter oligonucleotides carrying splinter 5'-overhangs of six random nucleotides (NNNNNN) are annealed to ssDNA. Random annealing of NNNNNN-overhangs to 5'-phosphorylated ends of ssDNA fragments create substrates

recognised by T4 DNA ligase. Subsequently, T4 DNA ligase seals nicks between phosphorylated 5’ ends of ssDNA fragments and adjacent 3’-OH ends of adapter-1 molecules. Ligation of biotinylated adapters-1 to 5’ ends of ssDNA fragments marks genomic positions of misincorporated rNMPs (Figure 1.1A; step 6).

DNA is then purified with magnetic AMPure XP beads under custom DNA binding conditions (1.25M NaCl, 7.5% PEG8000) previously used in the GLOE-Seq protocol (Sriramachandran *et al*., 2020). Purified DNA is then fragmented by sonication. The sonication step reduces the average size of ssDNA fragments, and thus allows next generation sequencing analysis of originally longer DNA species using conventional Illumina sequencing technology (Figure 1.1A and 1.1C; steps 7-11).

To remove non-specific ligatable DNA ends originating from sonication, sheared ssDNA is dephosphorylated by rSAP. Subsequently, ssDNA fragments successfully ligated to biotinylated Illumina adapters-1 are recovered by streptavidin (StrAv) coated magnetic beads and, while bound to the beads, ligated to Illumina adapter-2 using 5’-phosphorylated adapter-2/splinter oligonucleotides carrying splinter 3’-NH_2-NNNNNN overhangs. The splinter 3’ amino modification NH_2 prevents the formation of adapter-2/splinter homodimers. Unligated adapter-2/splinter oligonucleotides are removed by multiple washes with low-salt SSC buffer. Potential splinter oligonucleotides annealed to the ligated adapters are removed by mild alkaline wash (Figure 1.1A; steps 12-18).

ssDNA libraries are subsequently eluted from the StrAv-coated magnetic beads and PCR-amplified using NextSeq index primers for Illumina sequencing platforms. Following PCR amplification, libraries are purified with AMPure XP magnetic beads and analysed by Bioanalyzer (Agilent) using High Sensitivity DNA Chip (Figure 1.1D; steps 19-22).

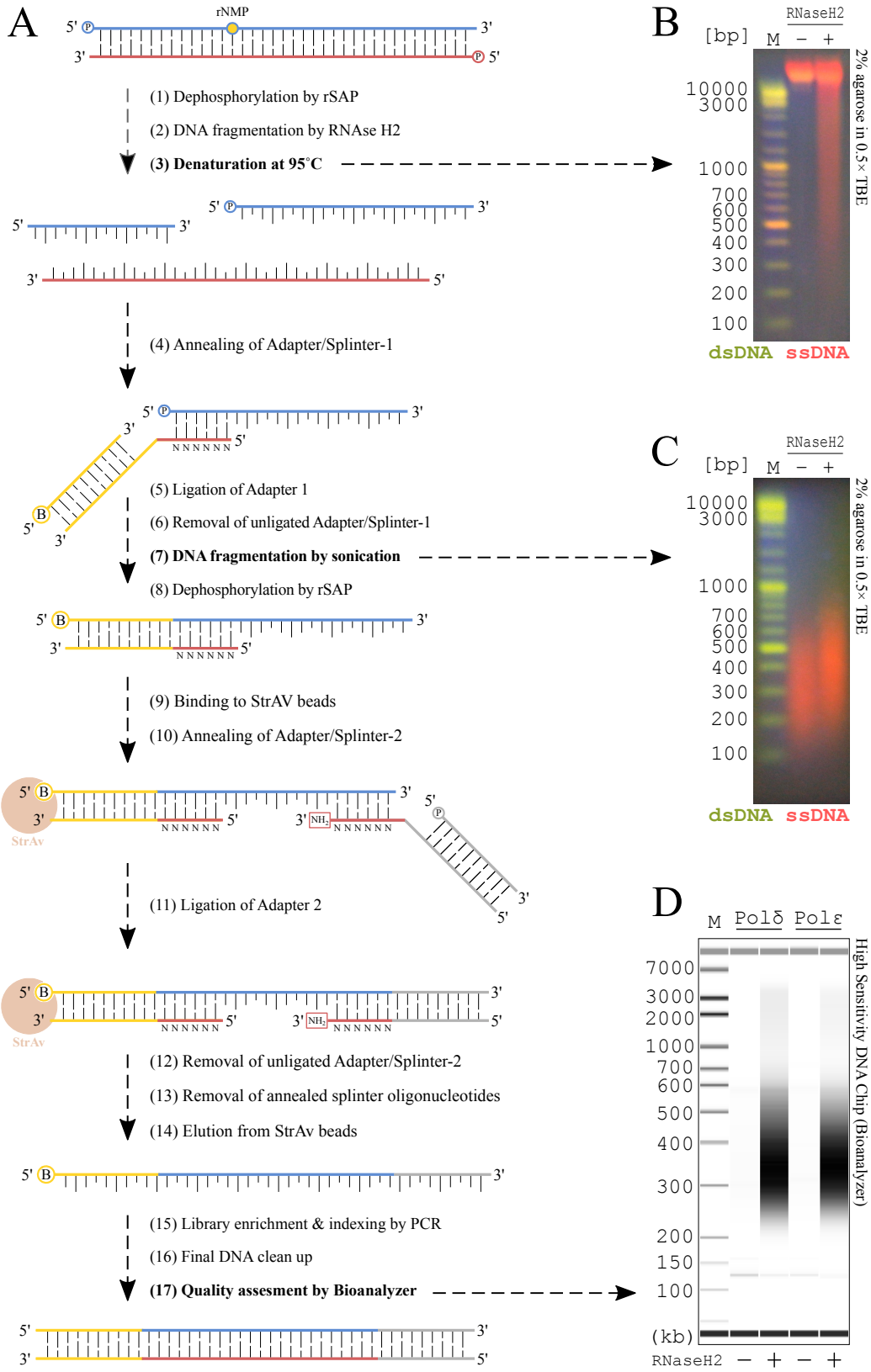

Figure 1.1 – Overview of proposed Pu-Seq protocol. (A) Simplified representation of key steps in newly the developed Pu-Seq protocol. **(B, C)** Control and RNase H2-treated DNA containing

high levels of rNMPs. Denatured DNA fragments were separated by agarose gel electrophoresis and stained with acridine orange. Distributions of DNA fragments prior to sonication (B) and after sonication (C) are shown. **(D)** Quality control of Illumina sequencing libraries prepared using DNA extracted from *cdc6*^{L591G} *rnh201Δ* (Polδ) and *cdc20*^{M630F} *rnh201Δ* (Polε) cells. Libraries were analysed by Bioanalyzer (Agilent) using High Sensitivity DNA Chip. P – phosphate group PO_4^-, B – biotin, N – random nucleotide, StrAv – streptavidin, NH_2 – DNA end amino-modification.

Preliminary results

As indicated in (Figure 1.1D), the newly established Pu-Seq procedure generated DNA libraries composed of DNA fragments, 5' ends of which represented RNase H2-sensitive positions of rNMPs misincorporated by mutated versions of Polδ and Polε.

Next, to test the performance of the newly developed Pu-Seq procedure, we prepared and sequenced three Pu-Seq libraries for both *cdc6*^{L591G} and *cdc20*^{M630F} mutants, each of which expressed the separation of function allele Rnh201-RED. For each independent dataset, we calculated polymerase track values. Briefly, polymerase track values are calculated using equation PT = (RT – RB) / (RT + RB), where PT represents polymerase track, and RT and RB stand for rNMPs mapped to the top and the bottom DNA strands, respectively. In a simplistic sense, positive and negative polymerase track values indicate predominant polymerase activity on the top and the bottom DNA strand, respectively. Calculation of polymerase track values was inspired by a previously published Okazaki fragment sequencing analysis (Petryk *et al*., 2016).

Independent pilot Polδ and Polε tracks generated by the newly developed Pu-Seq procedure were highly consistent and comparable to randomly selected Pu-Seq data produced by the original protocol (Figure 1.2A).

To address the newly generated data in a more comprehensive manner, we plotted histograms of genome wide Polδ and Polε track values and compared them with equivalent distributions deduced from data generated by the original protocol. We found that histograms of three newly generated and two control polymerase track values were highly similar (Figure 1.2B). Distributions of Polδ and Polε track values in control datasets were noticeably different (Figure 1.2B). We argue that this difference is attributable to inter-experimental variability which is commonly observed between

independent Pu-Seq datasets produced by the original protocol. We speculate that observed inconsistencies arise from sub-optimal input DNA which is already fragmented or introduction of non-specific DNA breaks during alkali treatment. Experimental evidence backing such claims, however, is currently lacking.

Since one of the most important features of Pu-Seq is determination of positions and efficiencies of replication origins, we compared distributions of origin efficiencies derived from data generated by the new and the old Pu-Seq procedures. Distributions of origin efficiencies estimated by the new protocol showed very little variability (Figure 1.2C). Histograms of origin efficiencies determined by the original protocol followed the same trend, however, results from three individual experiments displayed apparent variability (Figure 1.2C), which, as stated previously, we have not comprehensively addressed.

Collectively, preliminary experiments indicate that the newly developed Pu-Seq procedure can generate high quality data, which are consistent with datasets produced by the original version of Pu-Seq protocol. It is tempting to speculate that, when fully optimised, the new Pu-Seq procedure could eliminate or reduce inter-experimental variability, however, this notion will have to be validated by additional experiments.

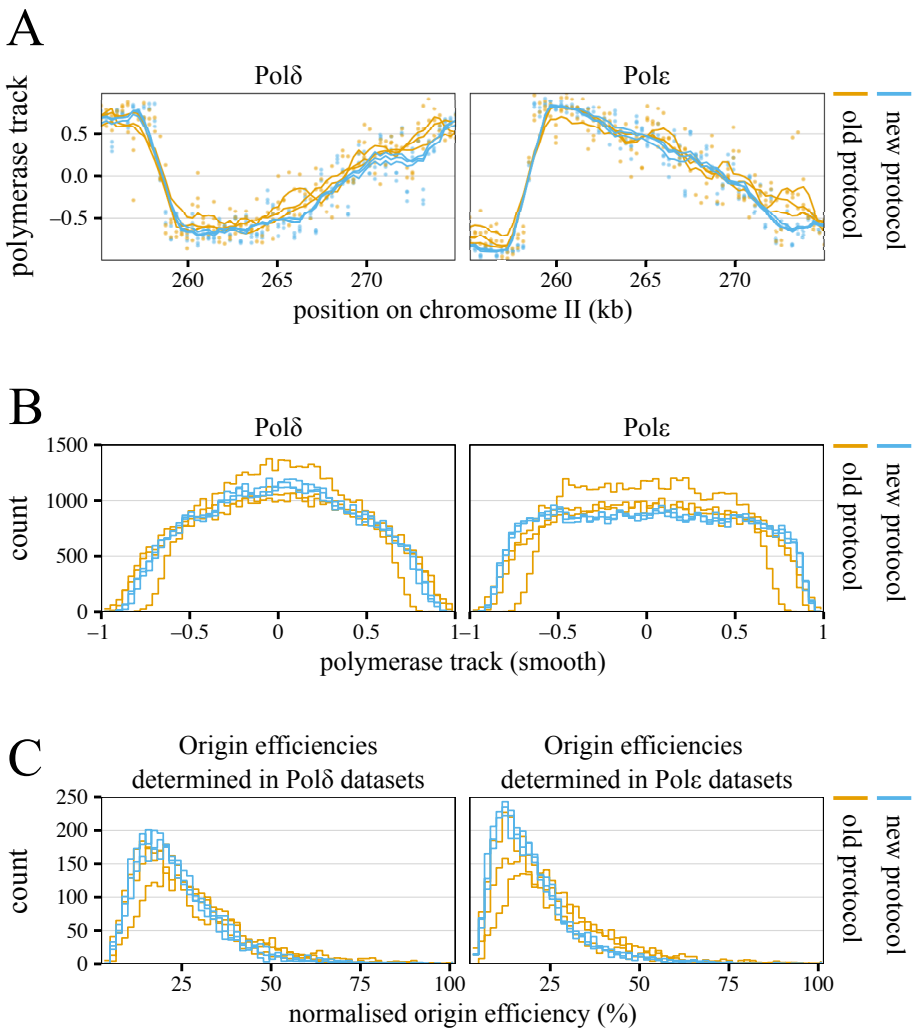

Figure 1.2 – Comparison of data generated by the original and the new Pu-Seq protocols. (A) Polδ and Polε track values across representative region on chromosome II (255,000-275000 bp). Efficient origin of replication and adjacent termination zone are indicated. Points and solid lines represent raw and smooth polymerase track values, respectively. For both polymerases, Polδ and Polε, three independent datasets generated by the original or the new protocol are presented. **(B)** Distributions of Polδ and Polε track values in datasets generated by the original or the new protocol. Distinct distributions correspond to independent experiments. **(C)** Histograms of origin efficiencies determined from Polε and Polδ track datasets generated by the original or the new Pu-Seq protocol. Distinct distributions correspond to independent experiments. Origin efficiencies were normalised, so the values fall between 0% and 100%.

Sub-optimal results of subsequent experiments

The newly developed Pu-Seq procedure was successfully used to generate high quality data presented in chapter 2 and chapter 3. Disappointingly, in significant portion of subsequent experiments, the new Pu-Seq protocol performed sub-

optimally and produced data which could not be analysed due to high levels of non-specific signal. Results of one such unsuccessful experiment (also discussed in chapter 4) are presented in Figure 1.3, where we address a representative genomic section as well as genome-wide distributions of Polδ and Polε tracks (Figure 1.3A and 1.3B). Interestingly, in this particular experiment, only data covering Polδ activity showed noticeably reduced polymerase track amplitude (Figure 1.3A and 1.3B). However, since other colleagues encountered the same issue in their Polε Pu-Seq datasets, it is unlikely that this artefact is somehow linked to the employed polymerase mutation. Also, since the presented analysis involved two independent isolates of four different genetic backgrounds, each of which was grown under two different conditions (details discussed in chapter 4), it is unlikely that the sub-optimal performance originated form human error, such as inoculation of an incorrect strain.

Complete distortion of characteristic oscillatory profiles of polymerase tracks, which we believe is attributable to increased levels of non-specific signal, could be caused by detection of non-specific DNA breaks and/or unscheduled ligation of Illumina adapters. This technical issue represents a significant caveat and is currently under investigation.

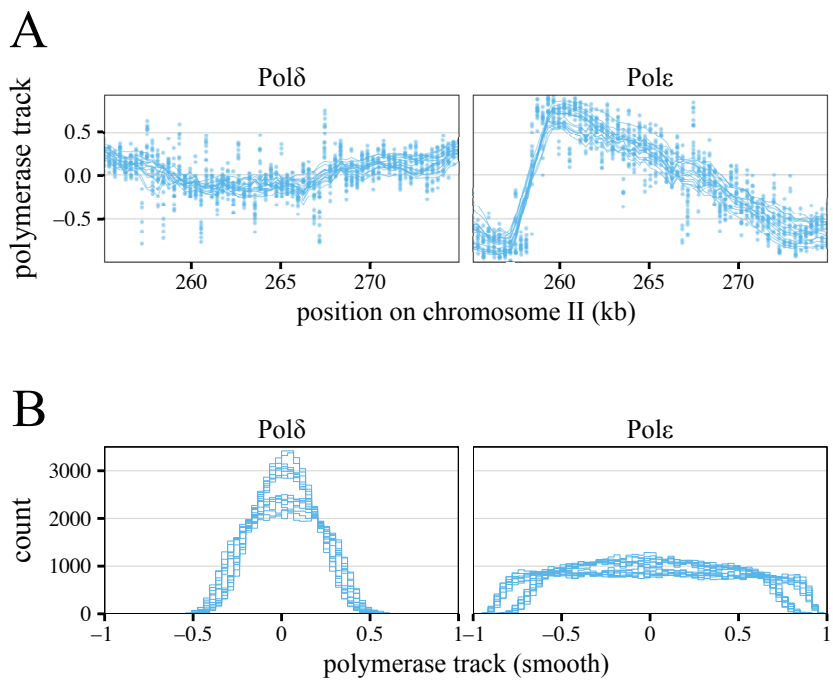

Figure 1.3 – Representative sub-optimal Pu-Seq data. (A) Polδ and Polε track values across representative region on chromosome II (255,000-275000 bp). Points and solid lines represent raw and smooth polymerase track values, respectively. **(B)** Distributions of Polδ and Polε track values in datasets generated by the new protocol. Distinct distributions correspond to independent experiments. (A, B) For each polymerase, 16 datasets were analysed.

Discussion 1

In this chapter, we summarised our efforts to establish a new version of Pu-Seq protocol, which is based on principles accommodated by recently developed GLOE-Seq and RHII-HydEn-seq (Sriramachandran *et al*., 2020; Zhou *et al*., 2019). We provide evidence that an updated Pu-Seq procedure can generate high quality data, which are comparable to those produced by the original Pu-Seq experimental pipeline (Daigaku *et al*., 2015; Keszthelyi *et al*., 2015). At the moment, however, the new Pu-Seq procedure stochastically generates sub-optimal results, which carry significant amount of non-specific signal and are not suitable for further analysis.

We reason that high levels of noise in sub-optimal Pu-Seq datasets originates from non-specific DNA shearing and/or illegitimate ligation of Illumina adapters. To address the issue, we propose the following optimisation experiments. 1) Verify that input genomic DNA samples and other reagents do not contain contaminant DNA nucleases, and test different genomic DNA extraction procedures. 2) Test whether rSAP phosphatase removes all ligatable DNA ends after sonication step. 3) Test whether additional purification of commercial adapter and splinter oligonucleotides by high performance liquid chromatography (HPLC) or other standard methods reduces non-specific signal. 4) Determine the optimal concentration of adapter/splinter oligonucleotides in both ligation reactions.

Assuming successful optimisation, from the practical standpoint, the newly established approach to construct Pu-Seq libraries carries several benefits including higher processivity (at least 8 samples can be processed at the time), lower cost (estimated to be 30% lower) and less time required for completion (two days in comparison with original three to four days).

From methodological and analytical perspectives, the new Pu-Seq protocol bypasses potentially problematic steps such alkali-based DNA fragmentation and size selection, which are suspected to generate yet uncharacterised forms of bias.

Step-by-step protocol of the newly designed Pu-Seq procedure

Premixes and Solutions

Phosphatase reaction premix (30 µL)

Genomic DNA	1 µg
10× CutSmart buffer	3 µL
rSAP	1 µL
mpH_2O	up to 30 µL

RNase H2 reaction premix (10 µL)

10× CutSmart buffer	1 µL
RNase H2	1 µL
mpH_2O	8 µL

Adapter 1 ligation reaction premix (30 µL)

10 mM ATP	7 µL
10× CutSmart buffer	3 µL
50 µM adapter/splinter 1	3 µL
PEG8000	14 µL
H_2O	2 µL
T4 DNA ligase (2,000,000 U/mL)	1µL

Dilution mix (70 µL)

5M NaCl	35 µL
Nuclease-free H_2O	28 µL
50% PEG8000	7 µL

Adapter 2 ligation reaction premix (50 µL)

10× T4 DNA Ligase Reaction Buffer	5 µL
50 µM adapter/splinter 2	2 µL
50% PEG8000	10 µL
mpH_2O	32 µL
T4 DNA ligase (2,000,000 U/mL)	1 µL

2× Binding and Washing (B & W) buffer

Tris-HCl (pH 7.5)	10 mM
EDTA	1 mM
NaCl	2M

20× SSC buffer (pH 7.0, adjusted with HCl)

Na-citrate	0.3 M
NaCl	3 M

Oligonucleotides

RZ199 (adapter 1)
/5BiodT/ACACTCTTTCCCTACACGACGCTCTTCCGATCT

RZ201 (splinter 1)
NNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

RZ202 (adapter 2)
/5Phos/AGATCGGAAGAGCACACGTCTGAACTCCAGT*C

RZ203 (splinter 2)
GACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNN*N/3AmMO/

Universal primer
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

Index primer
CAAGCAGAAGACGGCATACGAGATXXXXXXGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

Listed oligonucleotides are obtained from Integrated DNA Technologies (IDT). 100 µM stock solutions of adapter and splinter oligonucleotides are prepared by dissolving lyophilised oligonucleotides in nuclease-free H_2O. To prepare adapter/splinter duplexes, complementary 100µM adapters and splinters are mixed in 1:1 ratio and incubated according to *Adapter/splinter annealing thermal profile*.

Following oligo-modifications were used: 5BiodT – 5' Biotin-dT; 5Phos – 5' Phosphorylation; 3AmMO – 3' Amino Modifier; * – Phosphorothioate. Sequences of respective oligonucleotides are listed below. “XXXXXX” represents an index sequence.

Thermal profiles

Adapter/splinter annealing thermal profile.

1.	95°C	2 min
2.	75°C	1 s
3.	50°C	1 s
4.	25°C	∞

Adapter 1 ligation thermal profile.

1.	25°C	60 min
2.	22°C	60 min
3.	20°C	60 min
4.	16°C	∞

Note: over-nigh incubation
is recommended

Index PCR thermal profile

1.	98°C	30 s	
2.	98°C	10 s	
3.	72°C	40 s	Go to 2. [10×]
4.	72°C	2 min	
5.	4°C	∞	

Consumables

Product	Manufacturer	Catalog #
rSAP (1,000 U/mL)	NEB	M0371S/L
RNase H2 (5,000 U/mL)	NEB	M0288S/L
T4 DNA ligase (2,000,000 U/mL)	NEB	M0202T/M
AMPure XP beads	Beckman Coulter	A63881
Dynabeads MyOne Streptavidin C1	Invitrogen	65001
Index Primers Set 1	NEB	E7335S/L
Index Primers Set 2	NEB	E7500S/L
Index Primers Set 3	NEB	E7710S/L
Index Primers Set 4	NEB	E7730S/L
NEBNext Ultra II Q5 Master Mix	NEB	M0544S/L

Procedure

1. Prepare the *Phosphatase reaction premix* (30 µL).

2. Incubate the mixture at 37°C for 30 min and, subsequently, at 65°C for 10 min.

3. Cool the mixture down at room temperature for 5 min.

4. Add 10 µL of *RNase H2 reaction premix* into the dephosphorylated DNA (30 µL) and mix by pipetting.

5. Incubate at 37°C for 1 h and, subsequently, at 95°C for 10 min. Immediately after, put the mixture on slushy ice (ice bath) and incubate for at least 5 min.

6. Prepare the *Adapter 1 ligation reaction premix* (30 µL) and mix with denatured DNA (40 µL) by pipetting. Incubate according to the *Adapter 1 ligation thermal profile*

7. Mix the ligation reaction (70 µL) with the *Dilution mix* (70 µL) by pipetting.

8. Transfer 100 µL of AMPure XP beads into 1.5 mL centrifuge tube. Place the tube on a magnetic stand and, when solution is clear (ca. 5 min), remove the original storage buffer. Resuspend the beads in the mixture from step 7 (140 µL) by pipetting and incubate at room temperature for 5 min.

9. Place the tube on a magnetic stand, wait until solution is clear (ca. 5 min), and then remove supernatant. Rinse the beads twice with 200 µL of 80% ethanol and air-dry for ca. 5-10 min. Elute the DNA in 72 µL of nuclease-free H_2O.

10. Collect 70 µL of the eluate and mix with 70 µL of fresh AMPure XP beads. Then, proceed as in step 20. Elute the DNA in 105 µL of nuclease-free H_2O. Transfer 100 µL of the eluate into a clean 0.5 mL tube.

11. Using the sonicator Q800 (Qsonica), sonicate purified DNA (100 µL). Settings: 48 cycles, 20 s ON/40 s OFF, amplitude 70%.

12. Mix the sonicated DNA (100 µL) with 12 µL of 10× CutSmart buffer and 6 µL of rSAP phosphatase. Incubate at 37°C for 30 min and, subsequently, at 95°C for 10 min.

13. Resuspend Streptavidin (StrAv) magnetic beads by vortexing for 30 s. Add 20 µL of StrAv beads into a clean 1.5 mL centrifuge tube and mix with 100 µL of *B & W buffer*. Place the tube on a magnetic stand and, when solution is clear, remove and discard supernatant. Wash the StrAv beads with 100 µL of B & W buffer two more times.

14. Resuspend the StrAv beads in the solution containing sonicated DNA (118 µL), place the tube on a rotating wheel and incubate for 15-30 min at room temperature.

15. Place the tube on a magnetic stand and, when solution is clear, remove supernatant. Wash the StrAv beads three times with 100 µL of SSC buffer. In between washes, place the tube on a rotating wheel and incubate at room temperature for 5 min.

16. Resuspend the StrAv beads in the *Adapter 2 ligation reaction premix* (50 µL) and incubate at room temperature for 60 min on a rotating wheel.

17. Place the tube on a magnetic stand, wait until solution is clear (ca. 5min), and then remove supernatant. Wash the beads twice with 100 µL of *SSC buffer* and once with 100 µL of 20 mM NaOH. In between washes, place the tube on a rotating wheel and incubate at room temperature for 5 min. After the NaOH-wash, quickly wash the beads with 100 µL of nuclease-free H_2O.

18. To elute the DNA, resuspend the StrAv beads in 24 µL of nuclease-free H_2O and incubate at 95°C for 10 min. Don't spin and place the tube on a magnetic stand, wait for ca. 1 min and collect 18 µL of the supernatant.

19. Mix 18 µL of eluted DNA with 1 µL of universal primer, 1 µL of index primer and 20 µL of 2× Q5 Master Mix. Incubate according to the *Index PCR thermal profile.*

20. Mix the PCR reaction (40 µL) with 40 µL of AMPure XP beads. Place the tube on a magnetic stand, wait until solution is clear (ca. 5 min), and then remove supernatant. Rinse the beads twice with 200 µL of 80% ethanol and air-dry for ca. 5-10 min. Elute the DNA in 42 µL of nuclease-free H_2O. Transfer 40 µL of eluate into a clean 1.5 mL tube.

21. Mix the eluate (40 µL) with fresh 40 µL of AMPure XP beads. Place the tube on a magnetic stand, wait until solution is clear (ca. 5 min), and then remove supernatant. Rinse the beads twice with 200 µL of 80% ethanol and air-dry for ca. 5-10 min. Elute the DNA in 22 µL of nuclease-free H_2O. Transfer 20 µL of the eluate into a clean 1.5 mL tube.

22. Mix 1 µL of the library prep with 4 µL of nuclease-free H_2O. Analyse 1 µL of 1:5 diluted library by Bioanalyzer (Agilent) using high sensitivity DNA chip.

Chapter 2 - Increased expression of Polδ does not alter the canonical replication program *in vivo*

This chapter has been constructed based on recent publication (Zach & Carr, 2021). The original draft of the submitted manuscript was prepared by the author of this thesis. Relevant protocols are included in the *Methods* sections.

Background 2

Unchallenged duplication of the eukaryotic genome requires the coordinated action of three replicative polymerase complexes: Polα-primase (hereafter referred to as Polα), Polδ and Polε (Burgers & Kunkel, 2017). According to the canonical model of eukaryotic replication, Polα and Polδ cooperate to discontinuously synthesise the lagging strand via the iterative production of short Okazaki fragments (OF), ca. 150 bp, whereas Polε caries out continuous leading strand replication (Clausen *et al*., 2015; Keszthelyi *et al*., 2015; Miyabe *et al*., 2011) Interestingly, such strict division of labour does not always apply, and deviations have been documented (Guilliam & Yeeles, 2021).

While polymerase activities of Polα and Polδ are indispensable for cell survival, the polymerase domain of Polε is not required for completion of replication in either *S. cerevisiae* or *S. pombe* (Feng & D'Urso, 2001; Kesti *et al.*, 1999). In both yeast experimental models it has been demonstrated that Polδ facilitates the leading strand synthesis when catalytically-inactive Polε is expressed (Garbacz *et al*., 2018; Miyabe *et al*., 2015). Such findings have found support in *in vitro* experiments utilising reconstituted replisome system (Yeeles *et al*., 2017), confirming that, under certain circumstances, Polδ is competent in the leading strand replication.

Indeed, it has been reported that Polδ replicates both DNA strands during HoRReR in *S. pombe* (Miyabe *et al*., 2015) and BIR in *S. cerevisiae* (Donnianni *et al*., 2019). Additionally, genomic analysis by Pu-Seq or HydEn-seq revealed that Polδ is involved in the initiation of the leading strand replication in unperturbed *S. cerevisiae* and *S. pombe* cells, respectively (Daigaku *et al*., 2015; Garbacz *et al*., 2018; Zhou *et al*., 2019). In agreement with such findings, PCNA-associated Polδ has been shown to play

an important role in early stages of leading strand replication *in vitro* (Yeeles *et al*., 2017; Zhou *et al*., 2019). Moreover, it has recently been proposed that Polδ takes over leading strand synthesis prior to replication fork termination (Zhou *et al*., 2019). The exact role of Polδ during the final stages of replisome progression is, however, yet to be clarified.

Apart from homologous recombination dependent DNA synthesis and replication initiation, Polδ-mediated leading strand synthesis has been shown to occur in the context of polymerase uncoupling. It has been reported that cyclobutane pyrimidine dimer driven disengagement of CMG-associated Polε from the leading strand 3'OH generates a gap, the efficient filling of which requires the translesion synthesis machinery, as well as the action of Polδ (Guilliam & Yeeles, 2020b). Additionally, it has been demonstrated that Polδ takes over the leading strand synthesis and performs an error-free bypass of oxidative DNA adducts thymine glycol and 8-oxoguanine (Guilliam & Yeeles, 2021). In further support of a more generic function of Polδ in leading strand synthesis, Polδ has been shown to proofread errors introduced by Polε in hyper mutator *pol2-M644G* mutants (Guilliam & Yeeles, 2020b). In line with all aforementioned observations, it has been shown that CMG-associated Polε exists in two mutually-exclusive conformations, of which only one facilitates DNA synthesis (Zhou *et al*., 2017).

Intriguingly, according to *in vitro* studies of eukaryotic replication, two-fold and four-fold increase in Polδ concentration reduces the rate of the leading strand synthesis (Yeeles *et al*., 2017). It has been suggested that the observed retardation of leading strand replication represents a consequence of stochastic polymerase switching, during which Polδ outcompetes Polε and temporarily facilitates inefficient extension of the leading 3' end. Since the effect of Polδ concentration on replisome progression and the hypothetical phenomenon of leading strand polymerase switching has not been investigated *in vivo*, we aimed to test whether a similar phenomenon manifests in living cells, potentially shedding light on a yet uncomprehended promiscuity of replicative polymerases.

Results 2

Brief overview of Pu-Seq

Pu-Seq methodology determines the genome-wide polymerase activities by detecting the traces of rNMPs misincorporated by mutated Polδ (*cdc6*^{L591G}), Polε (*cdc20*^{M630F}) or Polα (*pol1*^{L850F}) (Daigaku *et al*., 2015; Keszthelyi *et al*., 2015). In Pu-Seq, respective polymerase mutant strains also carry a deletion of *rnh201*, the catalytic subunit of RNase H2 complex, disruption of which abrogates RER and thus stabilises misincorporated rNMPs (Daigaku *et al*., 2015). To assess activities of individual replicative polymerases, we employed a strategy previously used to analyse Okazaki fragment sequencing data (Petryk *et al*., 2016). Briefly, activities of Polδ, Polε and Polα at any given locus are expressed as polymerase tracks, which are proportional differences of rNMPs misincorporated in the top and the bottom DNA strands (Figure 2.1).

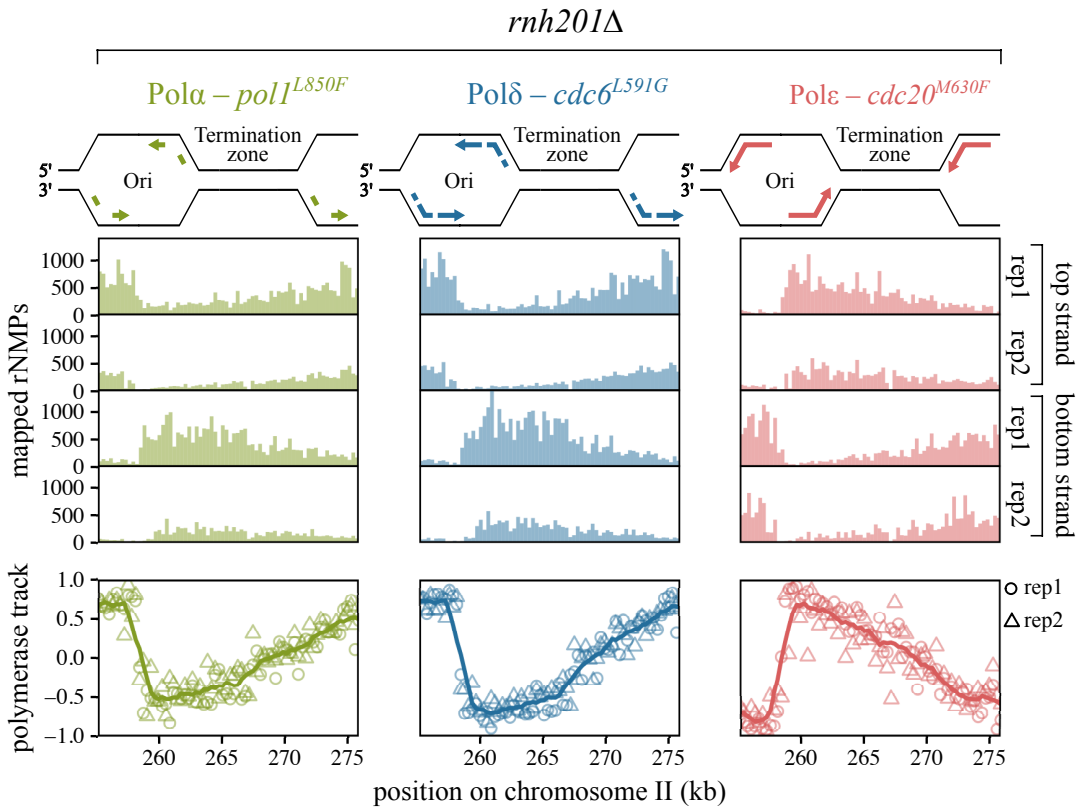

Figure 2.1 – Representative example of a basic Pu-Seq analysis. Top panel – cartoon representation of Polα, Polδ and Polε activities around an origin of replication (Ori) and across

adjacent termination zone. Respective polymerase mutations employed in Pu-Seq are indicated. Middle panel – Example of genomic ribonucleotides (rNMPs; presented as 300 bp bins) detected by Pu-Seq in *rnh201Δ* cells expressing Cdc6^{L591G}, Cdc20^{M630F}, and Pol1^{L850F}. A representative locus adjacent to an origin of replication is shown. Bottom panel – Polymerase tracks calculated for Polα, Polδ and Polε at the representative locus. For each polymerase, polymerase tracks are calculated from rNMPs mapped to the top and the bottom DNA strands as: PT = (RT – RB) / (RT + RB), where PT represents polymerase track, and R_T and R_B stand for rNMPs mapped to the top and the bottom DNA strands, respectively. Positive and negative values indicate predominant polymerase activity on the top and the bottom DNA strands, respectively. Data from 2 independent experiments are shown.

Construction and characterisation of Polδ-overexpressing strains

To achieve an approximate two-fold and four-fold upregulation of the whole Polδ complex, we aimed to increase the genomic copy number of all four Polδ genes. We constructed a set of Cre-Lox integration vectors, each of which carried a distinct selection marker (NatR, KanR, *ura4*$^+$) and all four genes constituting either WT (*cdc6*$^+$, *cdc1*$^+$, *cdc27*$^+$, *cdm1*$^+$) or L591G-mutated (*cdc6*^{L591G}, *cdc1*$^+$, *cdc27*$^+$, *cdm1*$^+$) Polδ (Figure 2.2A). Employing Cre-Lox recombination mediated cassette exchange (Watson *et al.*, 2008), we generated three distinct Polδ genomic integrations and created strains carrying either one (2×Polδ) or three (4×Polδ) extra copies of either WT or L591G-mutated Polδ holoenzyme (Figure 2.2B). Using WT Polδ integrations, we constructed 2×Polδ and 4×Polδ strains expressing Cdc2^{asM17} (Cdk1 variant inhibitable by the ATP analogue 3-Br-PP1), which allowed us to synchronise cells in G2 and assess their progression through the S-phase (Aoi *et al.*, 2014). Additionally, we constructed 2×Polδ *rnh201Δ* and 4×Polδ *rnh201Δ* mutants expressing either Cdc20^{M630F} or Pol1^{L850F}, which allowed us to determine whether the activities of Polε and Polα were altered in cells over-expressing Polδ. In a similar manner, utilising L591G-mutated Polδ integrations, we produced 2×Polδ *rnh201Δ* and 4×Polδ *rnh201Δ* mutants exclusively expressing Cdc6^{L591G}, which allowed us to asses activity of Polδ at different expression levels.

To validate that 2×Polδ and 4×Polδ mutants displayed increased expression of Polδ genes, we measured relative transcript levels of *cdc1*$^+$, *cdc27*$^+$, *cdc6*$^+$/*cdc6*^{L591G}, and *cdm1*$^+$ by RT-qPCR. In all genetic backgrounds tested, 2×Polδ and 4×Polδ mutants

displayed a significant increase in relative transcript levels of all four Polδ genes (Figure 2.2C). Unfortunately, due to the unavailability of commercial antibodies recognising Polδ subunits in *S. pombe*, we were unable to confirm that protein levels of the Polδ subunits were also increased. It has been previously reported, however, that plasmid-based over-expression of each of the four Polδ subunits is achievable in *S. pombe* (Kang *et al*., 2000; MacNeill *et al*., 1996; Reynolds *et al*., 1998). Consequently, we reasoned that elevation of Polδ transcript levels represented sufficient proof of bona-fide upregulation.

To determine the fundamental cellular consequences of Polδ-overexpression, we assessed growth rate and cellular morphology of WT, 2×Polδ and 4×Polδ cells. Polδ-overexpressing mutants displayed WT-like growth parameters and did not develop any cellular or nuclear defects (Figure 2.2D and 2.2E). Accordingly, increased Polδ expression did not alter the distribution of cell sizes (Figure 2.2F). To assess whether increased Polδ expression influenced progression through S-phase specifically, we synchronised WT, 2×Polδ and 4×Polδ cells with a Cdc2^{asM17} background in G2 by the addition of 3-Br-PP1 and analysed changes in DNA content in 15-min intervals after release. Progression through S-phase in 2×Polδ and 4×Polδ mutants was undistinguishable from WT cells (Figure 2.2F), suggesting that the over-production of Polδ did not change S-phase progression. Taken together, a moderate increase in Polδ expression did not have a notable impact on cell cycle or replication progression.

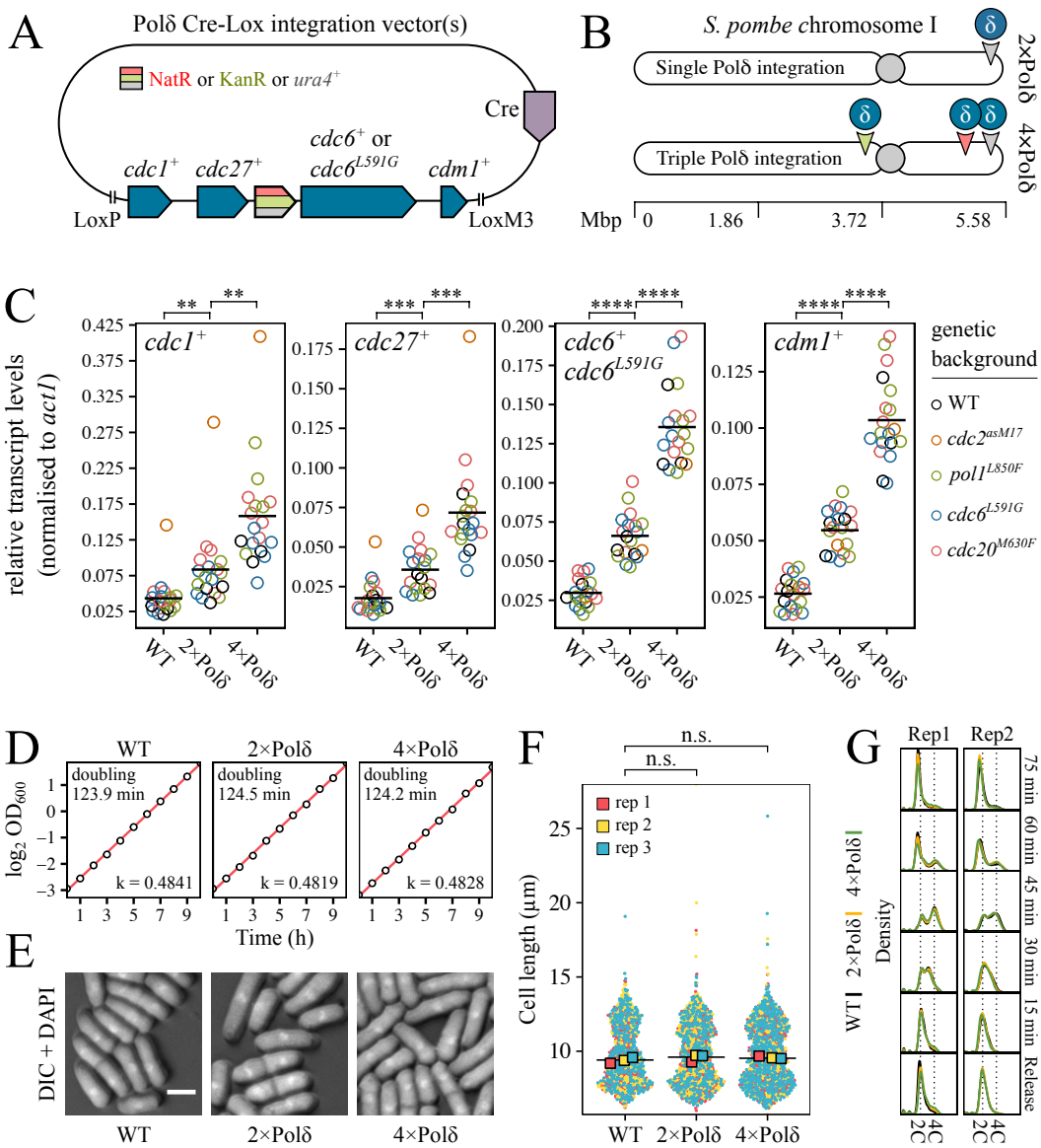

Figure 2.2 – Construction and characterisation of mutants over-expressing Polδ. (A) Simplified map of Cre-Lox vector(s) that were used to integrate extra copies of Polδ genes. Each vector carries genes constituting wild-type (WT) or L591G-mutated Polδ holoenzyme and one of three selection markers: NatR or KanR or *ura4*$^{+}$. Cre – Cre recombinase **(B)** Graphical representation of genomic Polδ integration site(s) in 2×Polδ and 4×Polδ cells. **(C)** Relative transcript levels of Polδ genes (*cdc1*$^{+}$, *cdc27*$^{+}$, *cdc6*$^{+}$/*cdc6*^{L591G} and *cdm1*$^{+}$) in the indicated mutants measured by RT-qPCR. Mutants designated as *pol1*^{L850F}, *cdc6*^{L591G} and *cdc20*^{M630F} also carried *rnh201Δ*. Individual points represent data from independent experiments. For WT, 2×Polδ and 4×Polδ cells, 19 independent measurements were taken (all genetic backgrounds combined). Horizontal lines represent means. Statistical significance was determined by the unpaired two-sample t-test. ** $p \leq 0.01$; *** $p \leq 0.001$; **** $p \leq 0.0001$ **(D)** Representative growth curves of WT, 2×Polδ and 4×Polδ cells. Optical density (OD) was measured in 1h intervals for total 10 h. Time-series of $\log_2$-transformed OD measurements are presented. Red lines represent linear regression models. Slopes of linear regression models (k) and calculated

doubling times are indicated. **(E)** Representative images of WT, 2×Polδ and 4×Polδ cells stained with DAPI. Composite images of DIC and DAPI channels are shown. Scale bar represents 5 µm. **(F)** Distributions of cell lengths of WT, 2×Polδ and 4×Polδ cells. Data from three independent experiments are shown. Squares represent medians of individual experiments. Horizontal line represents the median of merged data. Statistical significance was determined by the unpaired two-sample Wilcoxon test. n.s. = not significant. **(G)** DNA profiles of WT, 2×Polδ and 4×Polδ cells synchronised in G2. Results from two independent experiments are shown.

Replication dynamics

To investigate the potential influence of Polδ-overexpression on replication dynamics in greater detail, we performed two independent Pu-Seq experiments, each of which addressed activities of Polδ, Polε and Polα, in WT, 2×Polδ and 4×Polδ cells. While the first experiment utilised the canonical Pu-Seq procedure (Daigaku *et al.*, 2015; Keszthelyi *et al.*, 2015), the second repeat was carried out according to the newly developed protocol introduced in Chapter 1. Overall, in all genetic backgrounds tested, Polδ, Polε and Polα tracks displayed very little variation (Figure 2.3), suggesting that increased Polδ levels did not dramatically alter the properties of replication.

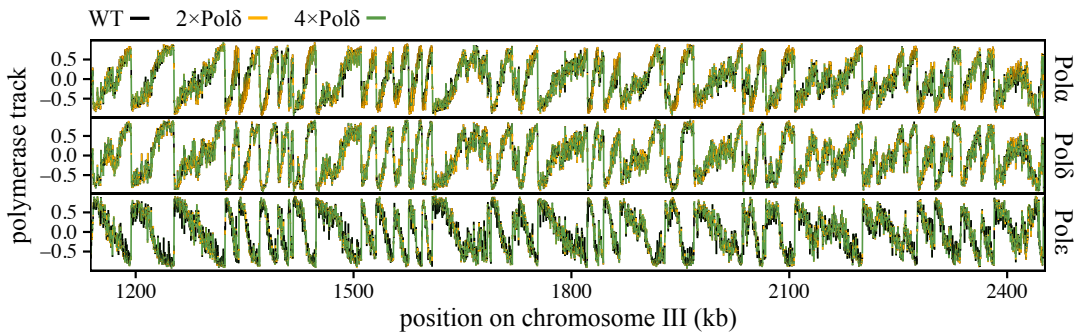

Figure 2.3 – Representative polymerase tracks in cells over-expressing Polδ. Polδ, Polε and Polα tracks across the right arm of chromosome III in WT, 2×Polδ and 4×Polδ cells. Means of two independent experiments are shown.

To capture a genome-wide view of replication, we examined regions around efficient origins of replication [characterised by estimated firing efficiency (Ori_{Eff}) of at least 40%] and regions constituting replication termination zones, which were defined by two efficient origins (Ori_{Eff} > 40%) and did not contain any intermediary efficiency origins (20% < Ori_{Eff} < 40%). Analysis of Polδ and Polε tracks associated with 259

efficient origins and 147 termination zones did not reveal any notable differences (Figure 2.4A and 2.4B). We observed that Polα tracks in 2×Polδ cells displayed marginal deviation from the WT profile (Figure 2.4A and 2.4B); however, considering that the observed difference was not reflected in 4×Polδ cells, we concluded this observation represented a technical, rather than biological phenomenon. We reasoned that if increased Polδ levels negatively affected replisome progression, 2×Polδ and 4×Polδ mutants would be expected to display increased activity of low and intermediary efficiency origins. Polδ-overexpressing cells, however, retained a WT-like distribution of genome-wide origin efficiencies, which further indicated normal replication progression (Figure 2.4C and 2.4D). Taken together, we concluded that, in our experimental system, a moderate increase in Polδ levels did not result in any observable changes in replication dynamics.

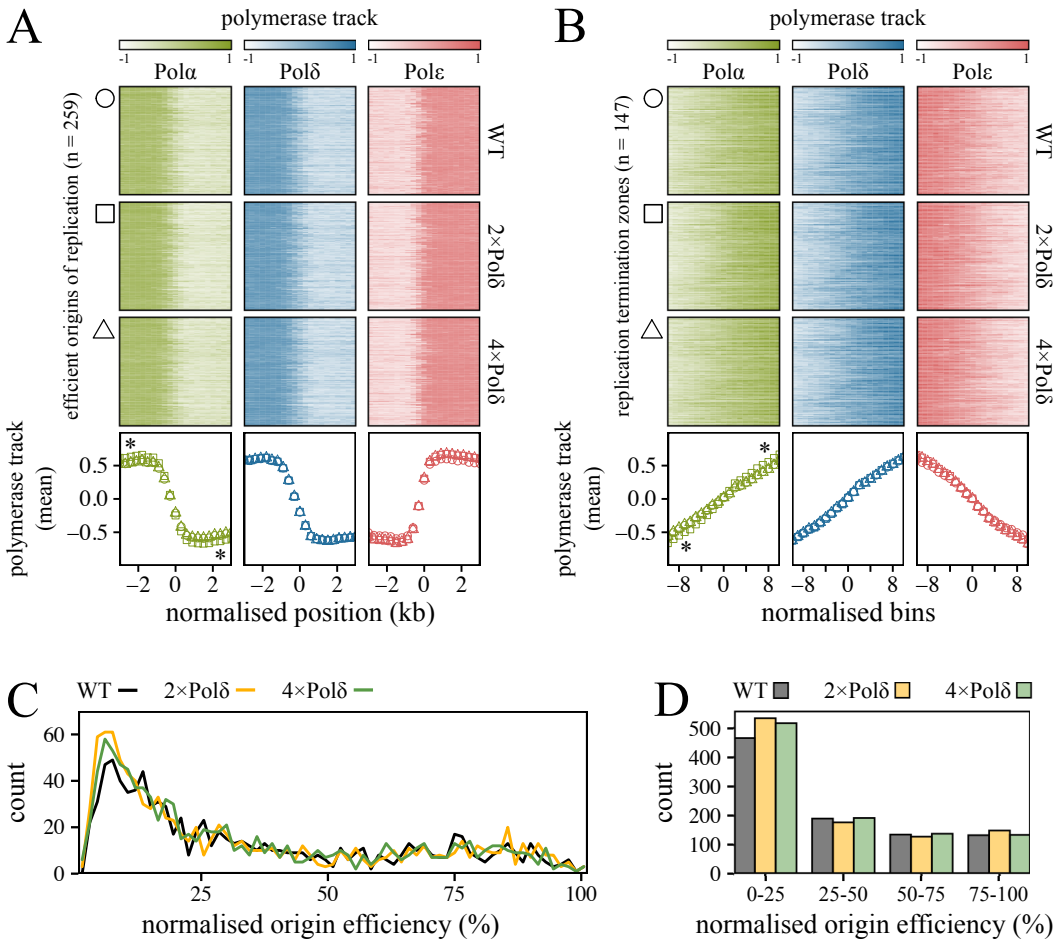

Figure 2.4 – Pu-Seq analysis of mutants over-expressing Polδ. (A, B) Polδ, Polε and Polα tracks around 259 efficient origins of replication (A) and across 147 termination zones (B). Individual regions and means are shown. Polδ expression levels are indicated: circles – wild-type (WT);

squares – 2×Polδ; triangles – 4×Polδ. Chromosomal coordinates around efficient origins were centred relative to the position of an origin. Data constituting termination zones were equally binned, and bins were centred relative to the midpoint of a termination zone. * Minor deviations in Polα tracks **(C, D)** Distribution of normalised origin efficiencies in WT, 2×Polδ and 4×Polδ cells. (A–D) Means of two independent experiments were analysed.

Discussion 2

In this study, we tested whether a moderate (2–4-fold) increase in Polδ expression impairs, or in any way alters, replication dynamics under normal conditions in *S. pombe*. The presented experiments were inspired by report that a two-fold and four-fold increase in Polδ concentration reduces the rate of the leading strand synthesis *in vitro*, hypothesised to be due to stochastic polymerase switching, during which Polδ outcompetes Polε and temporarily facilitates the extension of the leading strand (Yeeles *et al*., 2017).

We constructed a set of strains carrying either one or three extra copies of all Polδ genes and validated that these Polδ integrations resulted in increased transcription of the respective Polδ components: *cdc1*$^{+}$, *cdc27*$^{+}$, *cdc6*$^{+}$ and *cdm1*$^{+}$. We were unable to explore if the Polδ subunits were upregulated at the protein level. However, considering that successful ectopic over-production of Polδ subunits has been reported in the seminal literature (Kang *et al*., 2000; MacNeill *et al*., 1996; Reynolds *et al*., 1998), we argue that our experimental design conveyed a genuine Polδ over-production.

We determined that cells characterised by up to four-fold increased Polδ expression do not exhibit defects in growth and cell cycle progression. Furthermore, utilising Pu-Seq methodology, we demonstrated that genome-wide replication dynamics in 2×Polδ and 4×Polδ mutants is virtually indistinguishable from WT, arguing against the notion of stochastic polymerase switching or any other impairment of DNA replication induced by over-production of Polδ.

Naturally, it is still possible that we simply did not reach the threshold of Polδ expression that is required for the polymerase-switch to occur at frequencies detectable by Pu-Seq. Higher cellular levels of Polδ could be achieved by ectopic or strong promoter-driven expression of Polδ genes; however, we argue that such an

extensive Polδ over-production would constitute a non-physiological system, which would no longer be biologically relevant in relation to the reported *in vitro* data (Yeeles *et al*., 2017). Moreover, it has been shown that gross over-expression of *cdc6*$^{+}$ is detrimental to overall cell physiology (MacNeill *et al*., 1996), which would likely make Pu-Seq experiments difficult to interpret or impossible to carry out. We also argue that promoter manipulation or plasmid-based over-expression would disrupt the stoichiometry of Polδ subunits, which could be detrimental to Polδ folding and function.

While we established that moderate over-expression of Polδ does not noticeably affect canonical replication, we acknowledge that presented data do not sufficiently disprove the natural occurrence of the stochastic switch from Polε to Polδ. Nevertheless, our data do imply that, if such events occur *in vivo*, they manifest at low frequencies and likely represent only a marginal disturbance to an overwhelmingly robust replication program.

Chapter 3 – The role of senataxin helicases in the maintenance of the canonical DNA replication program

Background 3

Senataxin has been characterised as a widely conserved member of the Upf1-like family of helicases (Fairman-Williams *et al.*, 2010; Jankowsky, 2011; Leonaitė *et al*., 2017). *In vitro* studies established that purified *S. cerevisiae* and *S. pombe* senataxin helicases (Sen1p, Sen1^{+}) translocate in 5'-3' orientation and, in an ATP-dependent manner, unwind DNA/DNA, RNA/RNA as well as RNA/DNA duplexes (Kim *et al.*, 1999; Leonaitė *et al*., 2017; Martin-Tumasz & Brow, 2015).

In mammalian systems, mutations in senataxin, SETX, have been associated with neurological disorders ataxia with oculomotor apraxia type 2 (Anheim *et al*., 2009; Moreira *et al*., 2004) and amyotrophic lateral sclerosis type 4 (Chen *et al*., 2004). Additionally, SETX has been identified as a putative tumour-suppressor gene (Zhao *et al*., 2010) and is required for meiotic recombination and inactivation of sex chromosomes (Becherel *et al*., 2013; Yeo *et al*., 2015). Disruptions of the SETX gene have been further associated with signatures of genomic instability, including shortening of telomeres (De Amicis *et al*., 2011), illegitimate DSB repair (Brustel *et al.*, 2018; Skourti-Stathaki, Proudfoot, & Gromak, 2011), increased incidence of chromosomal aberrations (Kazadi *et al*., 2020) and increased sensitivity to H_2O_2, mitomycin C and camptothecin (Airoldi *et al*., 2010; De Amicis *et al*., 2011; Suraweera *et al*., 2007). Interestingly, patient-derived lymphoblastoid cells carrying the nonsense allele SETX-C1669T do not develop sensitivity to any of the above stated DNA-damaging stressors, suggesting a possibility that DNA damage response defect associated with at least certain SETX variants represents a consequence of altered or newly acquired SETX function rather than complete loss of activity (Airoldi *et al*., 2010).

In human cells, SETX predominantly localises to the nuclear compartment and coimmunoprecipitates with various RNA processing factors and RNAP2 (Suraweera *et al*., 2007, 2009). Disruption of SETX function causes defects in RNA metabolism of a subset of protein-coding genes manifested by lower RNAP2 occupancy, decreased transcript levels, aberrant splicing patterns and transcriptional readthrough (Nahas *et al.*, 2007; Suraweera *et al*., 2009). Analysis of RNAP2-transcribed artificial gene

constructs as well as endogenous loci indicates that SETX is required for the resolution of R-loops formed downstream of the poly-adenylation (poly-A) signal and facilitates recruitment of the exoribonuclease Xrn2/Rat1, which promotes termination of transcription by digesting the 3' end of the nascent RNA (Skourti-Stathaki *et al*., 2011).

Moreover, it has been elaborated that SETX is recruited at induced double-strand breaks colocalising with actively transcribed gene bodies, where it resolves break-induced RNA/DNA hybrids and promotes formation of Rad51 filaments and accurate DNA repair via either homologous recombination (HR) or non-homologous end joining (NHEJ) pathways (Cohen *et al*., 2018). Perhaps counter-intuitively, SETX mutants do not exhibit increased sensitivity to ionising radiation, a potent source of DSBs, an observation that is not yet fully understood (Nahas *et al*., 2007).

In *S. cerevisiae*, senataxin Sen1p has been mainly implicated in regulation of transcription termination of RNAP2-transcribed small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), cryptic unstable transcripts (CUTs) and a subset of protein coding genes (Arigo *et al.*, 2008; Steinmetz *et al.*, 2001; Steinmetz *et al*., 2006; Vasiljeva *et al.*, 2008). Additionally, some experimental evidence indicates a role of Sen1p in termination and/or processing or RNAP1-transcribed rRNAs and RNAP3-transcribed tRNAs (DeMarini *et al.*, 1992; Jamonnak *et al*., 2011; Kawauchi *et al*., 2008; Wlotzka *et al.*, 2011).

Disruption of Sen1p is accompanied by accumulation of RNA/DNA hybrid structures (R-loops), which are believed to be a direct consequence of impaired termination of transcription and have been extensively studied as potent drivers of genomic instability (Alzu *et al*., 2012; Appanah *et al*., 2020; Crossley *et al*., 2019; García-Muse & Aguilera, 2019; Helmrich *et al*., 2011; Mischo *et al*., 2011; Stirling *et al*., 2012). Accumulation of R-loops in Sen1p-deficient mutants correlates with increased frequency of HR, checkpoint activation and prevalent replication fork pausing at head-on, but not codirectionally oriented RNAP2-transcribed genes (Alzu *et al*., 2012; Brambati *et al*., 2018; Mischo *et al*., 2011; Stirling *et al*., 2012). Interestingly, disruption of Sen1p does not impede replication fork progression at RNAP3-transcribed loci, suggesting a specific activity of Sen1p in regulation of RNAP2-mediated transcription (Alzu *et al*., 2012). In accordance with the hyper-recombinogenic phenotype associated with Sen1p-deficiency, combining the

temperature sensitive *sen1-1* allele with mutations in genes involved in HR (Rad51p, Rad52p, Srs2p, Mre11p, and Sgs1p) results in synthetic lethality or severe cellular defects (Alzu *et al*., 2012; Mischo *et al*., 2011).

Most studies addressing Sen1p activity *in vivo* implicate that Sen1p functions as a constituent of the NNS (Nrd1p-Nab3p-Sen1p) complex, which physically interacts with the nuclear exosome and has been firmly established as a factor involved in termination of short, predominantly untranslated RNAP2-generated transcripts (Arndt & Reines, 2015; Carroll *et al.*, 2004; Chinchilla *et al*., 2012; Porrua & Libri, 2015; Steinmetz *et al*., 2001; Vasiljeva & Buratowski, 2006). NNS complex has been shown to associate with the C-terminal domain (CTD) of RNAP2 via a Nrd1-facilitated interaction with phosphorylated serine 5 (CTD-Ser5P). Additionally, the NNS complex binds specific sequence motifs on nascent RNA transcripts, which are recognised by RNA recognition motif (RRM) domains of Nrd1p and Nab3p (Vasiljeva *et al*., 2008). Interestingly, *in vitro* experiments suggest that Sen1p can displace transcribing RNAP2 without an assistance of additional auxiliary factors, indicating a possibility of NNS-independent Sen1p function (Leonaitė *et al*., 2017; Porrua & Libri, 2013; Wang *et al.*, 2019). In support of NSS-independent Sen1p activity, S-phase specific accumulation of R-loops at RNAP2-transcribed loci and accompanying replication fork stalling observed in *sen1-1* cells does not manifest in the temperature-sensitive *nrd1-102* mutants (Alzu *et al*., 2012). Additionally, it has been reported that Sen1p can interact with RNAP2 independently of Nrd1p and Nab3p via direct interaction with CTD phosphorylated on serine 2 (CTD-Ser2P) (Chinchilla *et al*., 2012; Ursic *et al.*, 2004). Since RNAP2 CTD-Ser5P and CTD-Ser2P modifications manifest in early and later stages of transcription cycle, respectively, it has been hypothesised that NNS-complexed Sen1p plays a role in termination of short transcripts, while free Sen1p manages longer transcription units (Komarnitsky *et al.*, 2000; Ursic *et al*., 2004).

It has been proposed that Sen1p translocating along the nascent RNA in 5'-3' direction eventually reaches progressing RNAP2 and facilitates its displacement from the transcribed DNA strand. In accordance with this model, *in vivo* and *in vitro* data indicate that Sen1p is more likely to displace slow, stalled or immobilised RNAP2, suggesting a dynamic relationship between ongoing transcription and termination (Hazelbaker *et al.*, 2013; Porrua & Libri, 2013; Wang *et al*., 2019). Some experimental

evidence also implies that, in unperturbed yeast cells, Sen1p cooperates with ribonucleases Rat1p and/or Rnt1p to promote termination of transcription by a 'torpedo-like' mechanism. According to this model, Sen1p resolves obstructing R-loops, promotes efficient degradation and/or processing of nascent RNAs, and ultimately contributes to displacement of RNAP1 and RNAP2 (Kawauchi *et al*., 2008; Mischo *et al*., 2011; Rondón *et al.*, 2009; Tollervey, 2004; Ursic *et al*., 2004).

Interestingly, it has been also reported that Sen1p associates with the replisome via interaction with Ctf4p and Mrc1p subunits (Alzu *et al*., 2012; Appanah *et al*., 2020). The separation of function mutant *sen1-3*, which no longer binds the replisome but preserves transcription termination activity, only develops minor genomic instability defects (Appanah *et al*., 2020). However, in genetic backgrounds characterised by significant accumulation of R-loops, *sen1-3* mutants are inviable or exhibit severe cellular defects (Appanah *et al*., 2020). In conclusion, even though the exact mechanisms of Sen1p activities are not completely clear, it is likely that Sen1p switches between multiple modes of action depending on the underlying regulation, actual interaction partners and/or biological context.

Unlike mammalian cells and *S. cerevisiae*, fission yeast expresses two senataxin paralogs, Sen1^{+} and Dbl8^{+}, which, when compared to *S. cerevisiae* senataxin orthologue Sen1p, exhibit 26% and 27% identity, respectively (Lemay *et al*., 2016). Interestingly, *sen1^{+}* as well as *dbl8^{+}* are not required for viability and disruption of either or both senataxin genes has little effect on cell physiology and morphology (Lemay *et al*., 2016). It has been demonstrated that Sen1^{+} and Dbl8^{+} interact with RNAP3 and RNAP1, respectively, but stable association with RNAP2 has not been documented for either of the two paralogs (Legros *et al.*, 2014; Rivosecchi *et al*., 2019). Accordingly, Sen1^{+} has been shown to play a role in termination of RNAP3-trancribed genes (tRNAs, 5S rRNA, srp7 and U6 snRNA) and Dbl8^{+} localises to RNAP1-transcribed rDNA (18S, 28S) loci and some RNAP2-transcribed protein-coding genes (Legros *et al*., 2014; Rivosecchi *et al*., 2019). Similarly to human senataxin, analysis of the snR3 locus suggests that Sen1^{+} and Dbl8^{+} do not play a role in termination of RNAP2-transcribed non-coding genes in *S. pombe* (Lemay *et al*., 2016).

While at least certain functions of Sen1p are carried out in the NNS (Nrd1p-Nab3p-Sen1p) complex in *S. cerevisiae*, neither Sen1^{+} nor Dbl8^{+} interacts with Nab3^{+}

and Seb1^{+}, the *S. pombe* orthologues of Nab3p and Nrd1p, respectively. In fact, it has been determined that Nab3^{+} and Seb1^{+} have functionally diverted from their *S. cerevisiae* counterparts (Lemay *et al*., 2016). Additionally, an equivalent of the *S. cerevisiae* NNS complex is likely absent in mammalian cells (Larochelle *et al*., 2018). Collectively, such evidence further implicates fundamental differences in molecular contexts of senataxin functions in *S. cerevisiae*, *S. pombe* and other eukaryotic species.

Even though experimental evidence proving the interaction between Sen1^{+} and/or Dbl8^{+} helicases with the replisome has not been publicly communicated, protein sequence alignments indicate that the amino acid residues required for Sen1p interaction with the replisome in *S. cerevisiae* are conserved in both senataxin helicases in *S. pombe* (Appanah *et al*., 2020). Notably, according to the protein sequence alignment of Sen1p and human SETX, human senataxin helicase likely does not feature a homologous replisome interaction motif. It cannot yet be ruled out, however, that the human senataxin SETX employs functionally analogous but structurally different means to interact with the replication machinery.

Across a wide range of experimental models, senataxin RNA/DNA helicases have been implicated in termination of transcription, suppression of pathological R-loops and maintenance of genome integrity. It has been inferred that at least some defects associated with senataxin-deficiency represent a consequence of increased occurrence of replication/transcription conflicts (RTCs), which, especially in head-on orientation, cause replication fork stalling and impose a threat to genome integrity (Hamperl *et al*., 2017; Hamperl & Cimprich, 2016; Pomerantz & O'Donnell, 2010; Rudolph *et al.*, 2007). Despite significant efforts put into the studies addressing RTCs, experimental evidence establishing concrete factors involved in resolution of transcription-induced replication fork stalling is lacking. In theory, senataxin helicases might represent such molecular suppressors of RTCs. In this chapter, we set out to investigate how senataxin-deficiency affects genome-wide replication dynamics in *S. pombe*.

Results 3

System to analyse the role of senataxin helicases in canonical DNA replication

Considering the well-established function of senataxin helicases in termination of transcription, we hypothesised that Sen1^{+} and Dbl8^{+}, the two senataxin paralogs in *S. pombe*, could be involved in suppression of naturally occurring RTCs and consequential replication fork stalling. To address the impact of senataxin-deficiency on replication progression, we generated a set of senataxin mutants including *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ*. Since some experimental evidence indicates that accumulation of R-loops plays a role in pathology of senataxin-impaired *S. cerevisiae* and human cells (Appanah *et al*., 2020; Helmrich *et al*., 2011), we also introduced a thiamine-regulatable construct encoding bacterial RNase A, the over-expression of which dramatically decreases R-loop levels (Hartono *et al*., 2018). If the analysis of constructed senataxin-deficient strains revealed any relevant phenotype(s), implementation of the regulatable RNase A would allow us to test a potential link between observed defects and R-loops levels. While *dbl8Δ* mutant was constructed de-novo by PCR-based gene targeting (Bähler *et al*., 1998), parent strains carrying thiamine-regulatable RNase A and *sen1Δ* constructs were kindly provided by V. Vanoosthuyse.

In accordance with the literature (Lemay *et al*., 2016), unperturbed *sen1Δ* and *senΔ1dbl8Δ* mutants developed minor growth defects, whereas *dbl8Δ* cells retained WT-like growth (Figure 3.1A). Interestingly, upon induction of RNase A expression, all senataxin-deficient mutants and WT cells displayed marginally reduced growth (Figure 3.1A). According to a genome-wide screen of factors involved in DNA damage response in *S. pombe*, *sen1Δ* cells exhibit sensitivity to DNA-damaging drugs including hydroxyurea (ribonucleotide reductase inhibitor; HU), methyl methanesulfonate (alkylating agent; MMS) and UV (Pan *et al*., 2012). To verify whether senataxin-deficiency was accompanied by sensitivity to DNA-damaging agents, we treated WT and senataxin-deficient mutants with the topoisomerase inhibitor camptothecin (CPT), UV and HU. According to our analysis, CPT and UV did not cause apparent impairment of growth in any of the three senataxin-deficient mutants tested (Figure 3.1A), whereas treatment with HU resulted in subtle growth retardation in

senΔ1dbl8Δ, but not *sen1Δ* and *dbl8Δ* mutants (Figure 3.1A). Interestingly, induction of RNase A expression had a slight effect on growth of senataxin-deficient as well as WT cells exposed to CPT, UV and HU (Figure 3.1A). Based on this analysis, we inferred that Sen1^{+} and Dbl8^{+} senataxins do not play a major role in DNA damage response; however, combined depletion of both senataxin genes might be associated with physiological consequences that negatively synergize with HU-induced replication stress. Since gross ectopic over-expression systems are often accompanied by detrimental effects, it is perhaps not that surprising that expression of RNase A driven by strong promoter had a negative impact on general cell physiology.

To be able to address replication dynamics by Pu-Seq, we introduced *rnh201-RED* allele, which abrogates RER but preserves RNase H2 function in resolution of longer R-loops (Naiman *et al*., 2021), and either *cdc6*^{L591M} or *cdc20*^{M630F}, which increase the frequency of genomic ribonucleotides misincorporated by Polδ and Polε, respectively (Figure 3.1B). Initially, we aspired to utilise the *cdc6*^{L591G} allele standardly used in Pu-Seq experiments (Daigaku *et al*., 2015); however, we did not manage to recover the combination of the *cdc6*^{L591G} allele with both senataxin mutations, indicating synthetic lethality. The historically-determined tendency to preferentially employ the *cdc6*^{L591G} allele is currently being discussed due to concerns about defective properties of Polδ complexes accommodating Cdc6^{L591G}.

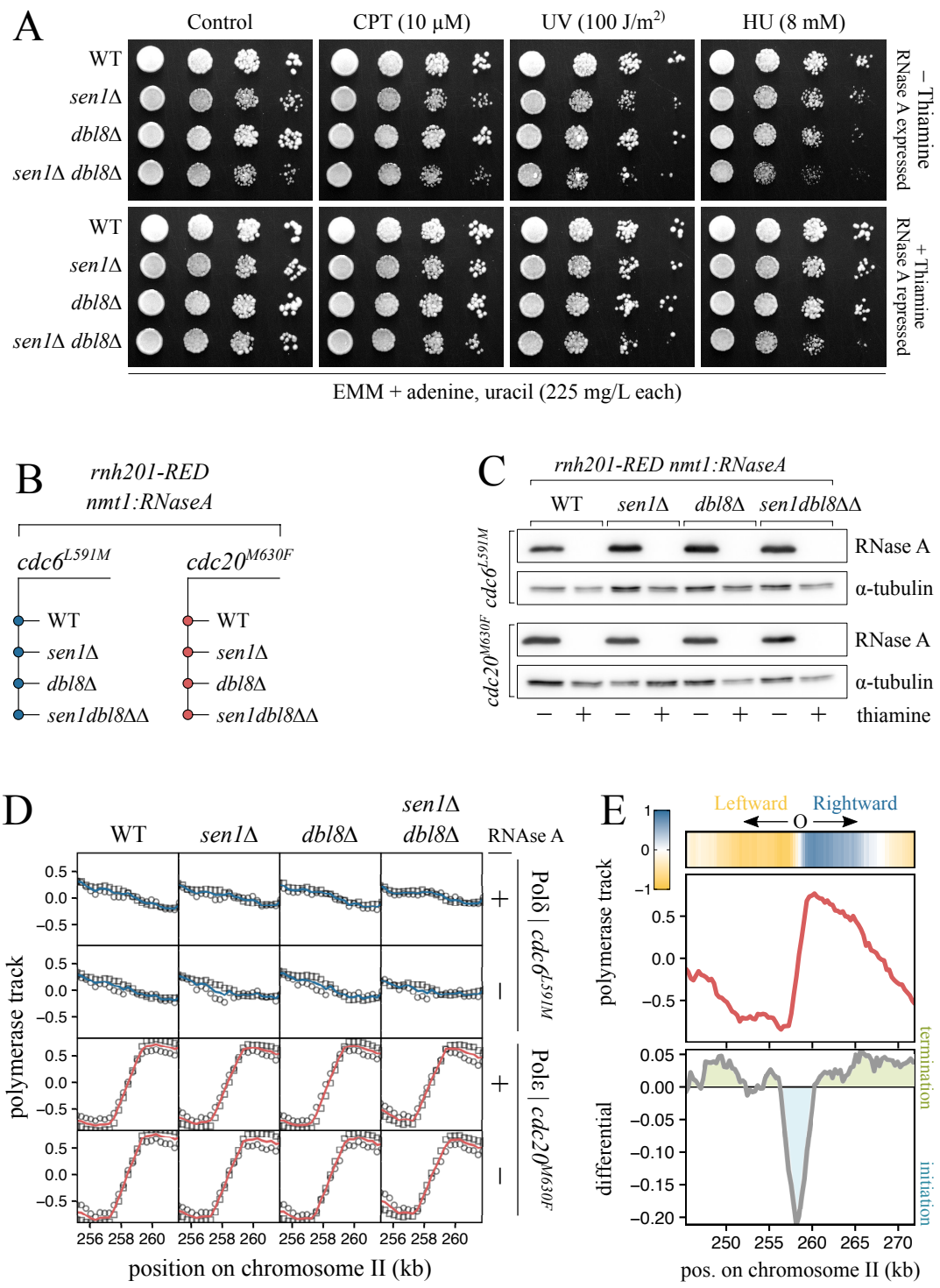

Figure 3.1 – Basic characterisation of senataxin-deficient mutants and the core Pu-Seq analysis **(A)** Growth parameters of WT and senataxin-deficient cells expressing or repressing bacterial RNase A under normal and genotoxic conditions. All tested strains carried a thiamine-regulatable RNase A construct. Exponentially growing cells were 10-fold serially diluted and spotted onto EMM plates containing listed genotoxic drugs at indicated doses. Cells subjected to UV-irradiation were first plated on EMM plate and then UV-irradiated. Images were taken 3 days after spotting. Experiment was performed once. **(B)** Schematic representation of

constructed senataxin-deficient strains used in Pu-Seq experiments. **(C)** Western-blot analysis showing bacterial RNase A levels in indicated mutants grown without (–) or with (+) thiamine. The expressed RNase A construct carried Flag-tag, which was used for detection. Exposure times: RnhA-FLAG – 8 s, alpha-tubulin – 60 s. **(D)** Polδ and Polε tracks at a representative origin of replication in WT and indicated mutant strains grown under conditions permitting (– thiamine) or repressing (+ thiamine) bacterial RNAse A expression. Positive and negative polymerase track values correspond to predominant polymerase activity on top and bottom DNA strands, respectively. Individual points (squares and circles) represent smoothed data from 2 independent experiments. Solid lines represent means of two experiments. While Polε tracks display the expected profiles, Polδ tracks exhibit a significantly reduced amplitude which is consequential of high levels of non-specific signal present in the raw sequencing data. **(E)** Upper panel – Polε track represented as a line or a heatmap. Negative and positive Polε track values indicate replication facilitated by leftward and rightward moving forks, respectively (as indicated). O – origin of replication. Lower panel – differential Polε track values. Negative and positive differential values indicate replication initiation and termination, respectively (as indicated).

Pu-Seq analysis

We performed two independent Pu-Seq experiments, each of which addressed replication dynamics in WT, *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* cells grown under conditions permitting (- thiamine) or repressing (+ thiamine) expression of bacterial RNase A. We determined that thiamine sufficiently repressed RNase A expression (Figure 3.1C), but we did not experimentally cover R-loop levels in any of the above stated genetic backgrounds and growth conditions. In future, such analysis would have to be undertaken to demonstrate functionality of the exogenous RNase A construct. Pu-Seq libraries were prepared according to the newly developed Pu-Seq protocol (discussed in chapter 1) and analysed as described in the *Methods* section. Briefly, polymerase tracks were calculated as $PT = (R_T - R_B) / (R_T + R_B)$, where PT represents polymerase track, and R_T and R_B stand for rNMPs mapped to the top and the bottom DNA strands, respectively. Positive and negative PT values indicate predominant polymerase activity on the top and the bottom DNA strands, respectively. While all generated Polε tracks displayed characteristic oscillatory patterns, Polδ tracks showed high levels of non-specific signal and were not suitable for further analysis (Figure 3.1D). We reasoned that the high degree of noise in Polδ datasets represented a technical artefact attributed to the sub-optimally performing Pu-Seq library preparation protocol, which, at the time, was not fully optimised. The

concrete reason behind such faulty performance, however, remained unclear. Nevertheless, Polε tracks on their own provided valuable information on replication dynamics, including replication initiation, termination and fork progression (Figure 3.1E).

Preliminary analysis of replication defects in senataxin-deficient mutants

First, we addressed replication dynamics in WT and senataxin-deficient cells grown under conditions inhibiting the expression of bacterial RNase A. Visual examination of replication profiles revealed that the loss of Sen1$^+$ and Dbl8$^+$ helicases caused local disturbances in replication profiles (Figure 3.2A). The observed replication defects manifested as sharp transitions between predominant Polε activity on the top and the bottom DNA strands with positive "right to left" slope (Figure 3.2A). As indicated in (Figure 3.1E), such profiles mark increased propensity for replication fork termination. Consequently, we hypothesised that these loci represented sites characterised by increased probability of replication fork stalling.

Since senataxin helicases play a role in regulation of transcription and accompanying RNA/DNA by-products, we aspired to correlate Pu-Seq data with genome-wide profiles of transcription levels. In order to construct such correlations, we retrieved two publicly available datasets addressing genome wide expression levels by either RNA sequencing (RNA-Seq) or native elongation transcript sequencing (Net-Seq) (Marguerat *et al*., 2012; Wery *et al*., 2018). While RNA-Seq data provided the information on steady state numbers of transcript copies per cell (cpc), Net-Seq data indicated the activity of RNAP2 on both DNA strands. In accordance with the function of senataxin helicases in regulation of transcription, identified replication fork stalling sites exhibited high transcriptional activity (Figure 3.2 A and 3.2B).

Interestingly, replication defects observed in *sen1Δ* and *dbl8Δ* mutants were mutually exclusive and correlated with co-directional and head-on oriented transcription, respectively (Figure 3.2A and 3.2B). Such observation indicated a possibility that Sen1$^+$ and Dbl8$^+$ helicases play distinct roles in preventing and/or resolving co-directional and head-on RTCs.

Notably, replication defects observed in *sen1Δ* and *dbl8Δ* mutants also manifested in *sen1Δdbl8Δ* double mutant. Interestingly, replication perturbations in

sen1Δdbl8Δ double mutant were sometimes, but not always, more pronounced then their parallels identified in *sen1Δ* and *dbl8Δ* cells (Figure 3.2A). Such observation indicated a negative genetic interaction between the two senataxin mutations. Even though not entirely comprehensive, these primary observations indicated a separation of function between Sen1$^+$ and Dbl8$^+$ senataxin helicases in resolution and/or prevention of transcription-linked replication fork perturbations and prompted us to analyse the data in more detail.

Identification of replication fork stalling sites in senataxin-deficient mutants

To address the impact of Sen1$^+$ and Dbl8$^+$ depletions on replication progression in a more comprehensive manner, we devised an analytical pipeline that determined the most dramatic replication impediments accompanying senataxin-deficiency as described in the *Methods* section. Our approach was based on a premise that replication fork stalling sites are characterised by increased probability of replication termination. In line with this postulate, we aimed to identify genomic sites which, in *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* mutants, exhibited increased occurrence of replication termination events with respect to the WT control.

First, similarly to (Daigaku *et al*., 2015), we determined replication termination and initiation sites by calculating differentials (slopes) of polymerase track profiles. As indicated in (Figure 3.1D), initiation and termination sites were characterised by negative and positive differential values, respectively (Figure 3.2C). Calculation of differential values and all subsequent analyses used mean polymerase track values derived from two independent experiments. Next, we calculated the difference in positive differentials between each senataxin mutant (*sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ*) and the WT control. Positive change at any given site indicated relative increase in probability of termination, hereafter referred to as relative replication fork stalling (Figure 3.2D). To avoid unspecific and low-penetrance hits, we only considered positive peaks characterised by the area and the maximum values above 0.98 quantile (Figure 3.1E). Lastly, using polymerase track values, we determined directionality of stalled forks at all identified stalling sites. It is important to keep in mind that our procedure only identified the most pronounced changes in replication dynamics and

generated results possibly underestimated the true impact of senataxin-deficiency on replication progression.

In *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* mutants, we identified 9, 30 and 143 high-penetrance replication fork stalling sites, respectively. In each senataxin mutant, stalling sites affecting leftward and rightward moving forks were approximately evenly represented (leftward forks stalled – *sen1Δ:* n = 5, *dbl8Δ*: n = 13, *sen1Δdbl8Δ*: n = 74; rightward forks stalled – *sen1Δ*: n = 4, *dbl8Δ*: n = 17, *sen1Δdbl8Δ*: n = 69).

In accordance with our preliminary inference, stalling sites identified in *sen1Δ* and *dbl8Δ* mutants did not colocalise and the vast majority of them were also present in *sen1Δdbl8Δ* double mutant (Figure 3.2F). Interestingly, disruption of both senataxin genes seemed to have an additive effect, as *sen1Δdbl8Δ* double mutant displayed 108 unique stalling sites, which were not detected in *sen1Δ* and *dbl8Δ* mutants (Figure 3.2F and 2G). Such observation further supported the notion of partial functional overlap between *sen1Δ* and *dbl8Δ* mutations. We also identified two sites which affected replication fork progression in *dbl8Δ*, but not in *sen1Δdbl8Δ* mutant (Figure 3.2H). We theorised that, at these rare sites, deviations in Pu-Seq profiles were masked in *sen1Δdbl8Δ* mutants due to simultaneous impairment of both converging replication forks.

Two replication fork stalling sites identified in all three datasets represented loci with two closely positioned divergently oriented genes. One such locus, featuring divergently oriented histone protein genes *hta1*$^{+}$ and *htb1*$^{+}$, is depicted in (Figure 3.2I). While *hta1*$^{+}$, encoded by the bottom strand, stalls rightward-moving forks in *dlb8Δ* mutants, *htb1*$^{+}$, encoded by the top strand, impairs the fork movement in *sen1Δ* and *sen1Δdbl8Δ* cells. Insufficient data resolution (300bp bins) does not allow adequate distinction of distinct replication fork stalling events at *hta1*$^{+}$ and *htb1*$^{+}$ loci. Consequently, our custom analytical pipeline identified *hta1*$^{+}$ and *htb1*$^{+}$ genes as a singular replication fork stalling site (Figure 3.2I).

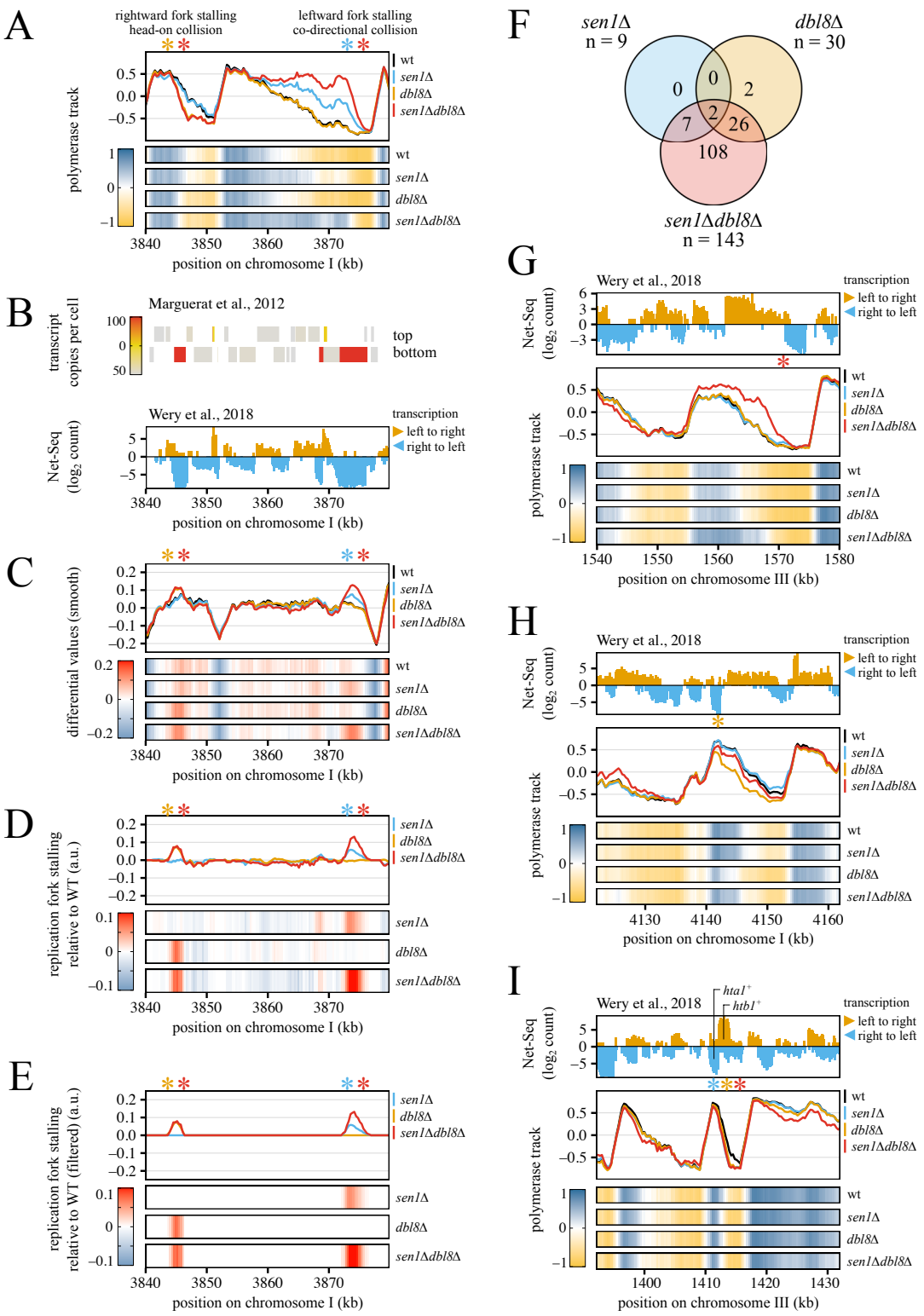

Figure 3.2 – Pu-Seq analysis of senataxin-deficient mutants. (A-E) Representative locus showing stalling of leftward and rightward moving forks in senataxin-deficient mutants. **(A)** Polε track values represented as lines or heatmaps. Positive and negative values indicate predominant Polε activity on top (rightward-moving forks) and bottom (leftward-moving forks) DNA strands, respectively. Asterisks indicate replication fork stalling sites in *sen1Δ*,

dbl8Δ and sen1*Δdbl8Δ* mutants. **(B)** RNA-Seq and Net-Seq data showing transcriptional activity on top and bottom DNA strands. **(C)** Differential values represented as lines or heatmaps. Positive and negative differentials indicate replication termination and initiation, respectively. **(D)** Relative replication fork stalling represented as change in positive differential tracks in *sen1Δ*, *dbl8Δ* and sen1*Δdbl8Δ* mutants with respect to WT. **(E)** Filtered changes in positive differential tracks. Only peaks characterised by areas and maxima above 0.98 quantiles were considered. **(F)** Venn diagram showing colocalization of replication fork stalling sites identified in *sen1Δ*, *dbl8Δ* and sen1*Δdbl8Δ* mutants. **(G)** Representative locus showing impairment of replication fork movement only present in sen1*Δdbl8Δ* double mutant. **(H)** Representative locus showing impairment of replication fork movement only present in *dbl8Δ* mutant. **(I)** Representative locus showing impairment of replication fork movement present in all three mutants: *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ*. (G-I) Net-Seq transcription profiles and Polε tracks represented as lines or heatmaps.

Characterisation of replication fork stalling sites in senataxin-deficient mutants

Using the current annotation of the *S. pombe* genome (Lock *et al*., 2019), we determined the genes most likely associated with replication fork stalling sites in *sen1Δ and dbl8Δ* cells, as well as 108 stalling sites uniquely identified in *sen1Δdbl8Δ* mutant (*Methods*, Tables 5-7). Out of 147 identified genes, 145 were protein coding and 2 non-coding (*prl53*$^{+}$, *snoU14*$^{+}$), indicating that observed defects in replication predominantly correlated with RNAP2-mediated transcription.

Notably, RNAP3-transcribed genes, such as 5S rRNA, *srp7*$^{+}$, U6 snRNA and tRNA, termination of which requires Sen1^{+} activity (Rivosecchi *et al*., 2019), did not display any changes in replication fork progression in any senataxin-deficient mutant (Figure 3.3). Since Pu-Seq is based on analysis of relatively short DNA fragments, our experiments did not reliably cover repetitive genomic regions such as 18S/28S rRNA loci which display significant Dbl8^{+} enrichment (Rivosecchi *et al*., 2019). Out of five RNAP2-transcribed loci (*adh1*$^{+}$, *SPAC27E2.11c*, *fba1*$^{+}$, *tef3*$^{+}$ and *act1*$^{+}$), which also display high levels of Dbl8^{+}, *adh1*$^{+}$ and *fba1*$^{+}$ (Rivosecchi *et al*., 2019), only *adh1*$^{+}$ and *fba1*$^{+}$ showed replication defects in *dbl8Δ* mutant (*Methods*, Table 6). Interestingly, depletion of both senataxin paralogs caused changes in replication fork movement at all five sites (*Methods*, Table 7).

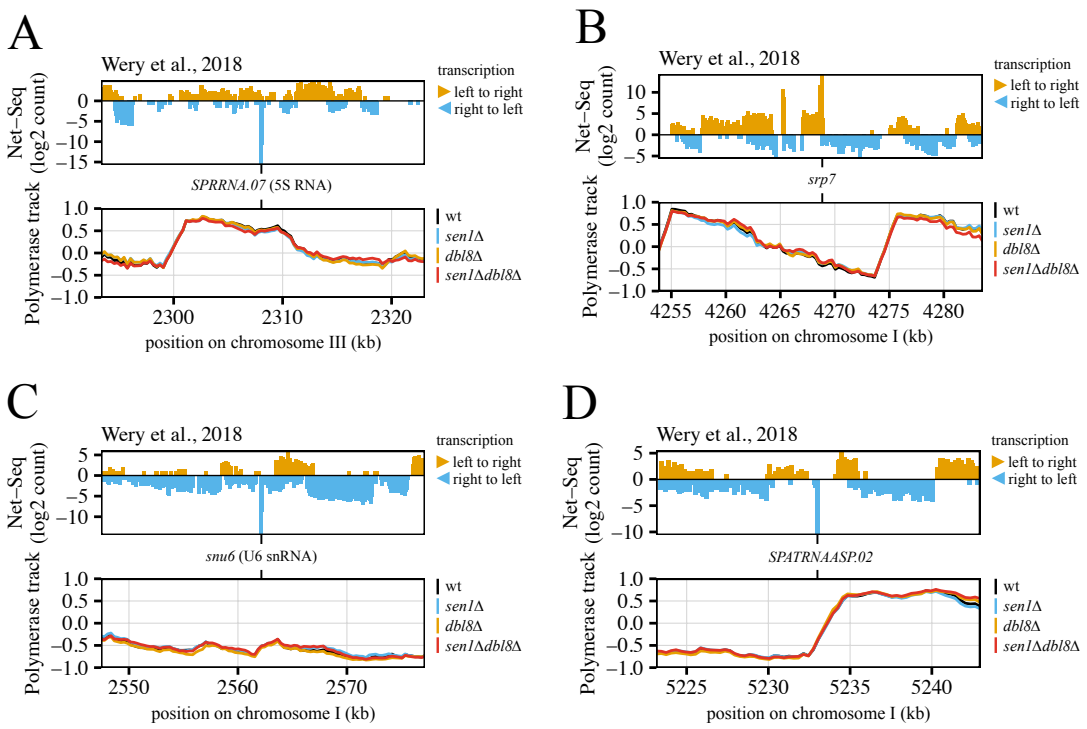

Figure 3.3 – Representative RNAP3-transcribed genes regulated by Sen1⁺. Net-Seq and Pu-Seq profiles of Sen1⁺ targets **(A)** 5S RNA representative *SPRRNA.07*, **(B)** signal recognition particle component *srp7⁺*, **(C)** small nuclear RNA U6 *snu6⁺* and **(D)** representative of tRNA genes *SPATRNAASP.02*.

According to publicly available gene expression profiles, out of 147 stalling associated genes, 16 displayed periodic gene expression. Out of 16 differentially expressed genes, 11 were upregulated in G2 phase, and 5 transcripts were most abundant during G1/S transition and S phase (Rustici *et al.*, 2004). Such analysis indicated that identified stalling-associated loci did not represent S phase-specific genes.

Initial inspection of Pu-Seq data suggested that replication fork stalling in *sen1Δ* and *dbl8Δ* mutants represented a consequence of codirectional and head-on RTCs, respectively (Figure 3.2A). To test such notion, we utilised publicly available Net-Seq data (Wery *et al.*, 2018) and analysed transcriptional activity 5 kb upstream and downstream of transcription start sites of 147 genes associated with replication fork stalling (Figure 3.4). In accordance with our preliminary analysis, all replication fork stalling events in *sen1Δ* and *dbl8Δ* mutants colocalised with codirectionally and head-on oriented highly transcribed genes, respectively (Figure 3.4). Notably, replication

defects uniquely identified in *sen1Δdbl8Δ* mutants were associated with either codirectionally or head-on oriented transcription and did not manifest in mutants depleted of only one of the two senataxin helicases (Figure 3.4). Such findings indicated the existence of two distinct types of replication perturbations and/or their resolutions(s). One that requires specific action of Sen1^{+} or Dbl8^{+}, and another that can be resolved and/or prevented by either of the two helicases. Since senataxin function has been implicated in resolution of R-loops, we also retrieved publicly available strand specific DNA/RNA immunoprecipitation followed by sequencing (DRIPc-Seq) data (Hartono *et al*., 2018) and tested whether, under unperturbed conditions, identified loci displayed increased R-loop levels. Indeed, a majority of analysed genes displayed increased DRIPc-Seq signal (Figure 3.4). However, since R-loops are generally over-represented at highly transcribed loci (Figure 3.5), the biological significance of such correlation was questionable.

To demonstrate that replication fork progression defects observed in senataxin-deficient mutants did not represent a feature common to all highly expressed RNAP2-transcribed genes, we analysed relative replication fork stalling proximal to TSS of all highly expressed protein-coding genes which, according to our analytical pipeline, were not associated with replication defects. In all senataxin-deficient mutants, relative replication fork stalling levels at these 280 control loci were mostly unrecognisable (Figure 3.5). Notably, replication of these 280 loci was mediated by either leftward- or rightward-moving forks with distinct probabilities and only rarely represented natural centres of replication termination zones (Figure 3.6). Such findings implied that transcription-linked replication fork impediments observed in senataxin-deficient mutants did not represent a consequence of generic transcriptional activity. Rather, genes prone to replication perturbations exhibited a unique behaviour and, in theory, were characterised by specific, yet unknown, features. Additionally, these results further demonstrated functionality of the analytical pipeline used to determine replication fork stalling accompanying senataxin-deficiency. Notably, 280 control loci also displayed high levels of R-loops (Figure 3.5). Collectively, these observations indicated that replication fork perturbations observed in senataxin-deficient mutants could not be simply explained by transcriptional activity nor accumulation of RNA/DNA structures (Figure 3.5).

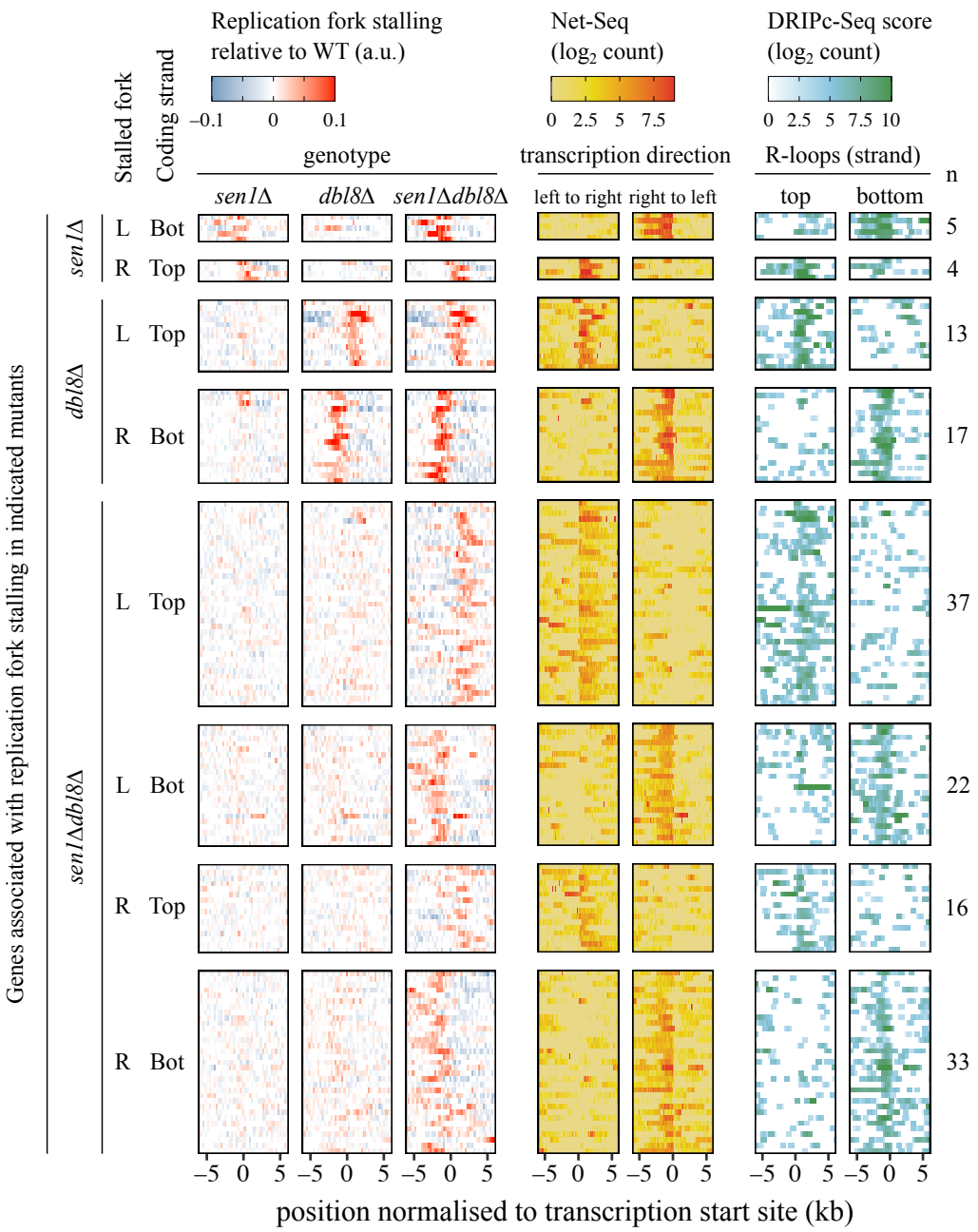

Figure 3.4 – Transcription and R-loop levels at replication fork stalling sites identified in senataxin-deficient mutants. Relative replication fork stalling in *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* mutants (left panel), transcriptional activity in unperturbed WT cells represented by Net-Seq data (Wery *et al.*, 2018) (middle panel) and R-loop levels in unperturbed WT cells represented by DRIPc-Seq data (Hartono *et al.*, 2018) (right panel) around genes causing replication fork stalling in *sen1Δ* and *dbl8Δ* mutants and genes causing replication fork stalling in *sen1Δdbl8Δ* mutants only. Data are clustered by direction of stalled replication fork (L – Leftward, R – Rightward) and coding strands of respective stalling-associated genes (Top, Bot – Bottom). Transcription direction and R-loop strand-specificity are indicated. Numbers of identified genes are indicated (n). Positions are normalised to transcription start site (0).

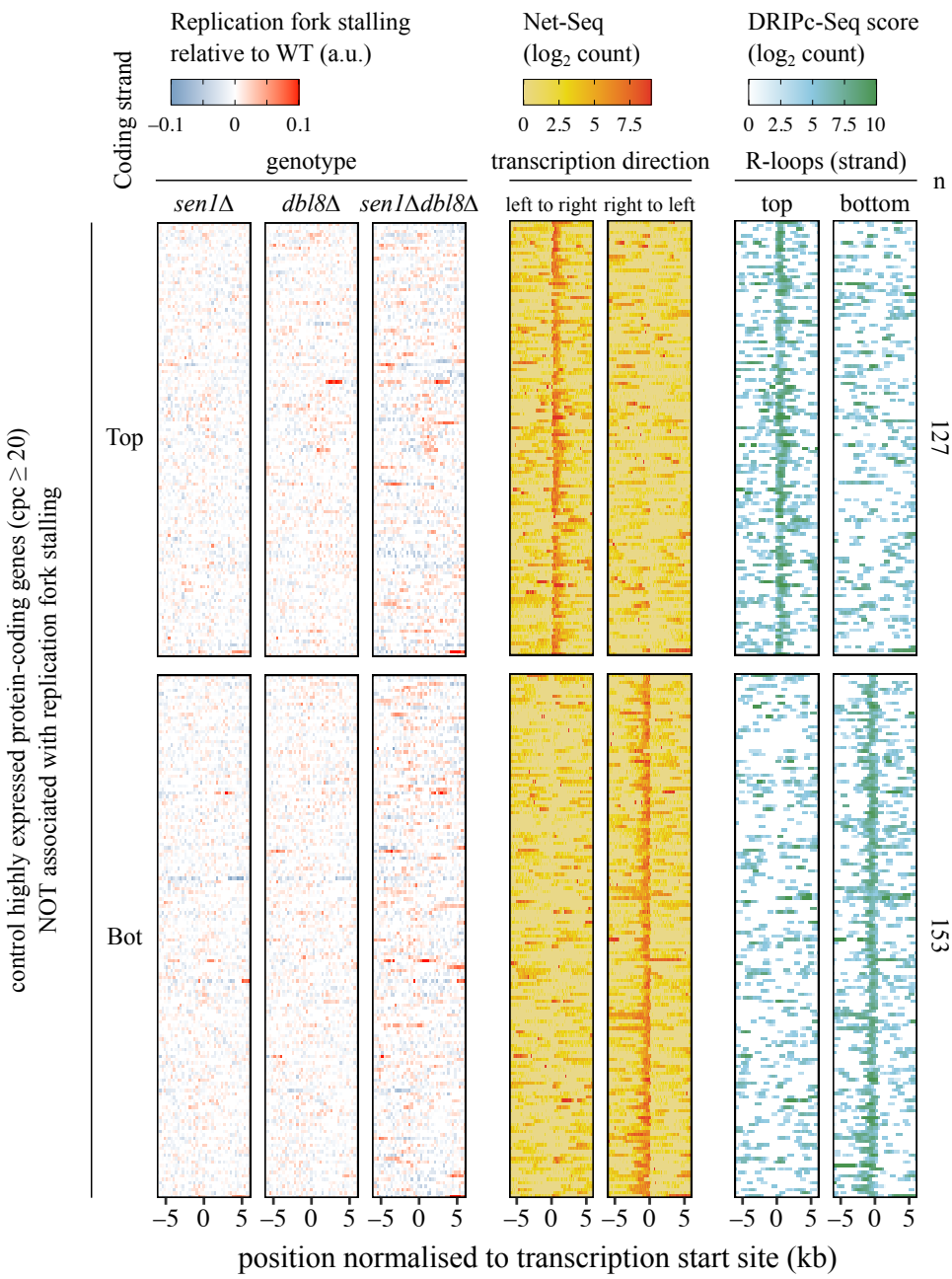

Figure 3.5 – Relative replication fork stalling and R-loop levels at highly expressed protein-coding loci not identified by the analytical pipeline. Relative replication fork stalling in *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* mutants (left panel), transcriptional activity in unperturbed WT cells represented by Net-Seq data (Wery *et al.*, 2018) (middle panel) and R-loop levels in unperturbed WT cells represented by DRIPc-Seq data (Hartono *et al.*, 2018) (right panel) around control highly expressed (cpc ≥ 20) protein-coding genes which were not associated with replication defects in senataxin-deficient mutants. Data are clustered by coding strands of respective genes (Top, Bot – Bottom). Transcription direction and R-loop strand-specificity are indicated. Numbers of identified genes are indicated (n). Positions are normalised to transcription start site (0).

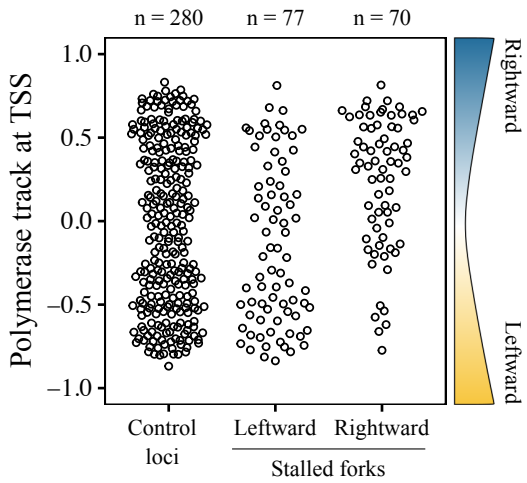

Figure 3.6 – Replication fork directionality at control and stalling-associated genes identified in senataxin-deficient mutants. Distributions of polymerase tracks at TSS positions of highly expressed protein-coding genes which were not associated with replication fork stalling and genes associated with replication defects in senataxin-deficient mutants. Directionality of stalled replication forks identified in senataxin-deficient mutants is indicated. Directionality of fork progression deduced from polymerase tracks is indicated. Individual points represent distinct TSS. Numbers of analysed loci are indicated.

In order to verify that no underlying patterns were lost in the heatmap-based analysis, we summarised the data presented in Figure 3.4 and Figure 3.6 (Figure 3.7). At this point, we also analysed analogous data covering relative replication fork stalling levels in senataxin-deficient mutants over-expressing bacterial RNase A, which has been shown to suppress the accumulation of most R-loops (Hartono *et al*., 2018).

The summary analysis of replication fork stalling at identified stalling sites and control loci revealed two additional insights of potential relevance. Firstly, expression of RNase A did not have an apparent impact on manifestation of replication defects in any of the senataxin-deficient mutants, further indicating that accumulation of R-loops is unlikely to be the cause of observed replication perturbations (Figure 3.7). Secondly, stalling sites identified in *sen1Δ* mutants were associated with accumulation of RNA/DNA structures on both strands (Figure 3.7). Provided that the over-expression of RNase A did not rescue replication defects in *sen1Δ* mutants, we speculated that increased DRIPc-Seq signal represented a consequence, rather than the cause of impaired replication fork progression. Apart from the aforementioned observations, the summary analysis (Figure 3.7) was in line with our previous inference that replication defects specific for senataxin-deficient mutants were not simply

consequential to transcriptional activity and/or accumulation of R-loop structures (Figure 3.4 and Figure 3.5).

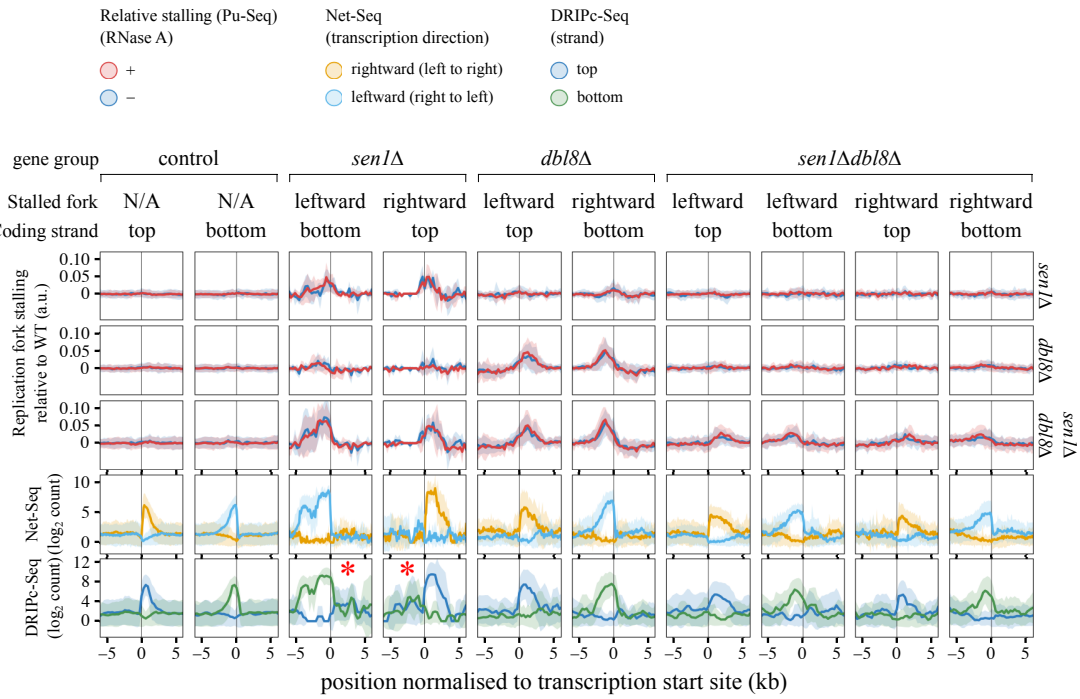

Figure 3.7 – Summary of replication fork stalling and wt R-loop levels at stalling-causing and control loci. Loci associated with replication fork stalling in *sen1Δ* and *dbl8Δ* mutants, loci associated with replication fork stalling in *sen1Δdbl8Δ* mutants only and control highly expressed protein-coding loci which are not associated with replication defects in any senataxin-deficient mutant. Top panel – summary of relative replication fork stalling in *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* mutants expressing or repressing bacterial RNase A. Middle panel – summary of transcriptional activity in unperturbed WT cells represented by Net-Seq data (Wery *et al.*, 2018). Bottom panel – summary of R-loop levels in unperturbed WT cells represented by DRIPc-Seq data (Hartono *et al.*, 2018). Solid lines and shaded regions represent means ± standard deviations. Data are clustered by directionality of stalled replication forks (L – Leftward, R – Rightward, N/A – no stalling) and coding strands of a respective stalling-causing or control genes. RNase A expression, transcription direction and R-loop strand-specificity are indicated. Red asterisks indicate R-loop accumulation on both strands.

Characterisation of genes associated with replication fork stalling

Next, we aimed to test whether 147 genes associated with replication fork stalling shared any underlying feature(s) that could explain their role in observed impediments of replication fork progression. We decided to address four fundamental characteristics including gene ontology (GO), transcript levels, gene length and GC-content. As a control group, we used 280 highly transcribed protein-coding genes

which, according to our analysis, were not associated with any impediments of replication fork progression. The vast majority of 147 stalling-causing genes as well as 280 control genes were involved in fundamental cellular processes such as gene expression, intermediary metabolism, biogenesis, signalling and transport (Figure 3.8A). Considering the similarity of two GO distributions, we concluded that 147 stalling-associated genes did not form a distinct functional cluster, and therefore were unlikely to be governed by a specific subset of cellular regulome.

In accordance with our previous analysis, 147 stalling-associated genes displayed relatively high transcript levels. However, considering that 280 highly expressed control genes were not associated with any replication defects, in line with our previous conclusions, steady state expression on its own could not sufficiently explain changes in replication fork progression specific to senataxin-deficient mutants (Figure 3.8B).

Interestingly, while genes associated with replication fork stalling in *sen1Δ* and *dbl8Δ* mutants did not show significant differences in gene length, stalling-associated genes uniquely identified in *sen1Δdbl8Δ* double mutant were significantly longer than highly expressed control genes (Figure 3.8C). Even though difficult to interpret, such observation indicated that gene length could play a role in replication defects consequential to simultaneous depletion of both senataxin paralogs, Sen1^{+} and Dbl8^{+}.

Next, we analysed GC contents of stalling-associated and control genes. When compared with the control group of highly expressed protein-coding genes, 39 stalling-associated genes identified in *sen1Δ* and *dbl8Δ* mutants did not show any significant difference in GC content. Interestingly, 108 stalling-associated genes unique to *sen1Δdbl8Δ* mutants displayed significantly lower GC content. However, considering that distributions of GC contents reflected respective distributions of gene expression levels (Figure 3.8A and 3.8C), it is likely that the observed difference represented a correlation, which, in the context of Sen1^{+} and Dbl8^{+} function(s), did not bear any biological relevance.

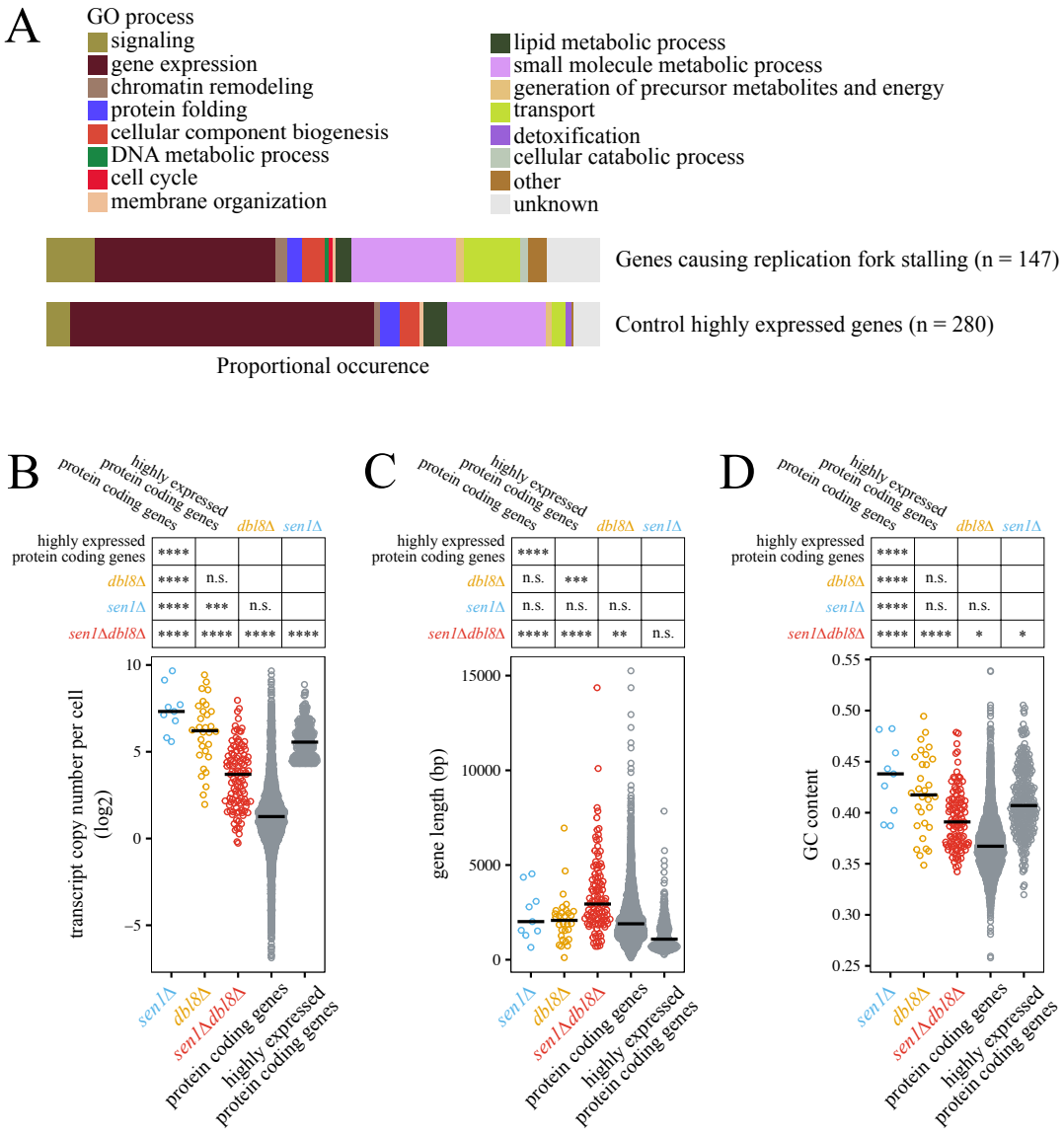

Figure 3.8 – Characterisation of genes associated with replication fork stalling sites in senataxin-deficient mutants. (A) Gene ontology (GO) analysis of 147 genes associated with replication fork stalling in senataxin-deficient mutants (*sen1Δ*, *dbl8Δ*, *sen1Δdbl8Δ*) mutants and 280 highly expressed (cpc ≥ 20) protein-coding genes which did exhibit replication defects. Proportional occurrences of GO processes were constructed using current GO annotation and PomBase's Quick Little Tool (QuiLT) (Lock *et al*., 2019; The Gene Ontology Consortium, 2019). **(B)** Transcript levels based on RNA-Seq data (Marguerat *et al*., 2012). **(C)** Lengths of respective genes according to the current annotation of *S. pombe* genome (Lock *et al*., 2019). **(D)** GC contents of respective genes according to the current genome build (Wood *et al*., 2002). (B-C) Analysis of genes associated with replication fork stalling in *sen1Δ* (n = 9) and *dbl8Δ* (n = 30) mutants and genes associated with replication fork stalling specific to *sen1Δdbl8Δ* mutant (n = 108). Also displayed are characteristics of RNAP2-transcribed genes (5062 protein-coding genes and non-coding genes *prl53* and *snoU14*) (n = 5064) and highly expressed (cpc ≥ 20) protein-coding genes (n = 280). Points represent values corresponding to individual genes.

Horizontal lines represent median values. Statistical significance was determined by Wilcoxon rank sum test. * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$; **** $p \leq 0.0001$.

In order to test whether stalling-associated genes displayed local extremes in GC content, we retrieved sequences of all identified stalling-associated genes and control highly expressed protein-coding genes and analysed GC content across equally binned gene bodies (Figure 3.9A and 3.9B). In accordance with the analysis of average GC content, stalling-associated genes which were only identified in *sen1Δdbl8Δ* cells showed marginally lower GC distributions across gene bodies (Figure 3.9A and 3.9B).

In a similar manner, we also analysed GC skew which has been previously correlated with R-loop formation (Ginno *et al.*, 2013, 2012). Similarly to GC content, average GC-skew across stalling-associated genes which were only identified in *sen1Δdbl8Δ* mutants was noticeably lower than GC-skew profiles of control highly expressed genes. GC-profiles of stalling-associated genes identified in *sen1Δ* and *dbl8Δ* mutants where comparable with the control (Figure 3.9C and 3.9D). Consequently, we inferred that GC content and GC skew do not play a significant role in transcription associated impairments of replication fork progression in senataxin-deficient mutants.

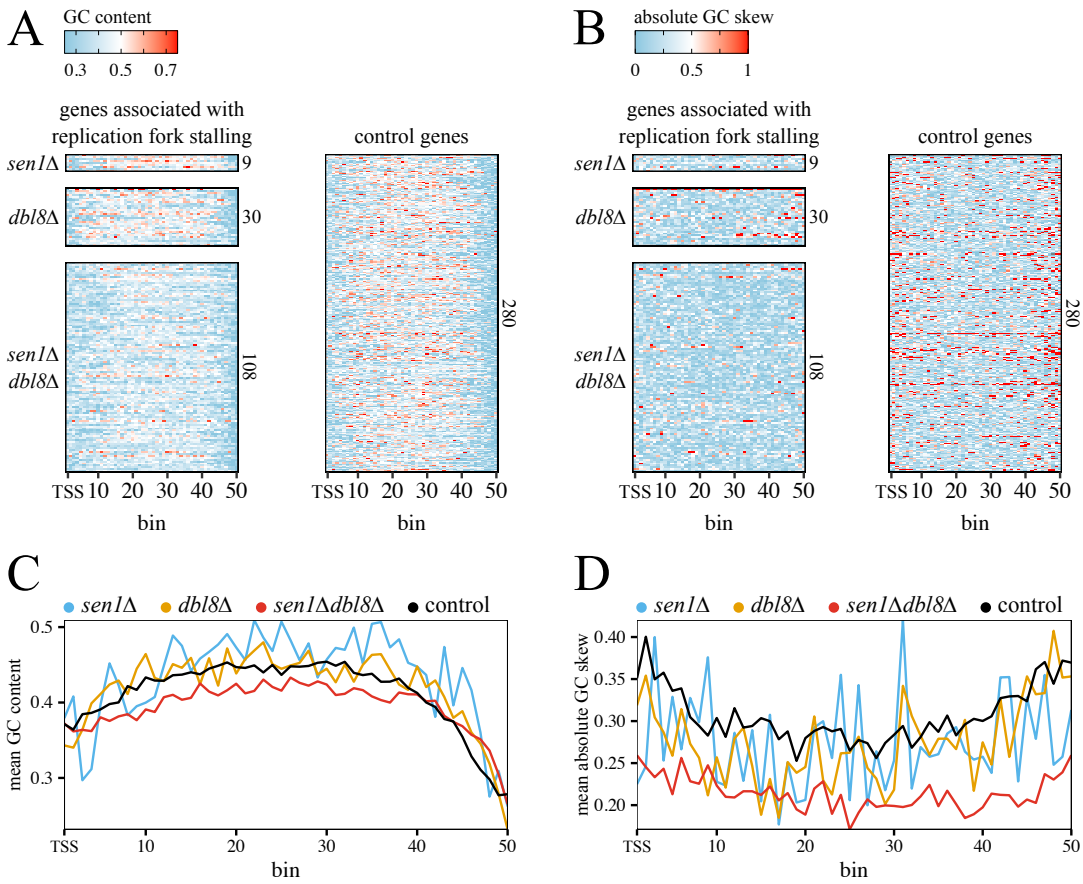

Figure 3.9 – GC content distribution across genes associated or not with replication fork stalling in senataxin-deficient mutants. (A, B) Heatmap representation of GC content (A) and absolute GC skew (C) across genes associated with replication fork stalling in *sen1Δ* and *dbl8Δ* mutants, genes associated with replication fork stalling in *sen1Δdbl8Δ* mutants only and control highly expressed protein-coding genes. Numbers of genes in respective gene groups are indicated. TSS – transcription start site. GC skew was calculated as GC_{Skew} = (G – C) / (G + C), where G and C represent numbers of guanosine and cytosine nucleotides in a given bin. Absolute GC skew values are presented. **(C-D)** Mean GC content and GC skew distributions calculated from values presented in (A, B).

Discussion 3

Across a wide range of experimental models, senataxin RNA/DNA helicases have been implicated in regulation of transcription and resolution of R-loops, triple-stranded RNA/DNA structures, accumulation of which has been linked to compromised genome integrity and consequential cellular pathologies (Alzu *et al.*, 2012; Appanah *et al.*, 2020; Crossley *et al.*, 2019; García-Muse & Aguilera, 2019; Helmrich *et al.*, 2011; Mischo *et al.*, 2011; Stirling *et al.*, 2012). Even though RTCs and their impact on genome integrity have been studied extensively (García-Muse &

Aguilera, 2016; Hamperl & Cimprich, 2016), little is known about the proteins involved in the resolution of RTCs. We postulated that senataxin helicases could represent RTC-suppressing factors. Therefore, in an attempt to resolve a significant gap in understanding of senataxin RNA/DNA helicases and their role in canonical replication, we analysed the impact of senataxin-deficiency on replication dynamics in *S. pombe*.

While *S. cerevisiae* and human cells express only one senataxin helicase, which, in both species, is indispensable for survival, *S. pombe* genome encodes two senataxin paralogs, Sen1^{+} and Dbl8^{+}, both of which are not required for viability (Lemay *et al*., 2016). To address replication dynamics in senataxin-deficient mutants, we constructed *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* strains carrying a thiamine-regulatable RNase A construct. In accordance with the literature (Lemay *et al*., 2016), *dbl8Δ* cells retained WT-like growth characteristics, whereas *sen1Δ* and *sen1Δdbl8Δ* mutants developed minor growth retardation. In response to DNA damage inducing agents including CPT, UV and HU, *sen1Δ* and *dbl8Δ* mutants were mostly unaffected.

Interestingly, *sen1Δdbl8Δ* cells displayed minor sensitivity to HU, but not CPT and UV, indicating a yet uncharacterised genetic interaction between *sen1Δ* and *dbl8Δ* mutations, which only manifests under conditions imposing replication stress. Although potentially relevant, the role of senataxin helicases in DNA damage response was not our primary target, and we decided not to further address the observation.

Notably, in all senataxin-deficient mutants and WT cells, induction of RNase A expression was accompanied by reduced cell fitness under normal conditions as well as in response to DNA damaging agents. Even though over-expression of RNase A is standardly used in studies addressing R-loop functions, to our best knowledge, detrimental effects of RNase A over-expression on cellular physiology has never been publicly communicated. Considering that the mechanism behind cellular defects observed in cells over-producing RNase A is unclear, we argue that conclusions presented in studies utilising the ectopic RNase A expression systems should be interpreted with caution.

Commonly, Pu-Seq experiments utilise strains expressing the L591G-mutated Polδ catalytic subunit, Cdc6^{L591G}, which reduces base selectivity of Polδ, and thus increases the frequency of misincorporated rNMPs. Notably, we did not recover a *sen1Δdbl8Δ* mutant expressing Cdc6^{L591G}, indicating a lethal genetic interaction.

Synthetic lethality associated with the *cdc6*^{L591G} allele is not uncommon and has been observed in different mutants across various projects developed in the lab. In general, however, mechanisms underlying these genetic interactions are currently not clear. We were able to construct a *sen1Δdbl8Δ* mutant expressing Cdc6^{L591M}, an alternative to Cdc6^{L591G} characterised by lower rate of rNMP misincorporation.

Standardly, Pu-Seq analysis combines information on Polδ and Polε activities, and thus provides comprehensive insights into genome-wide patterns of replication fork movement and polymerase usage. Unfortunately, due to technical complications accompanying newly developed Pu-Seq library preparation protocol, we were limited to the analysis of Polε activities only. Nevertheless, using calculated Polε track values, we addressed movement and distribution of bidirectional replication forks across the genome in WT and senataxin-deficient mutants.

In all three senataxin-deficient mutants, analysis of two independent Pu-Seq experiments revealed apparent changes in distributions of leftward and rightward moving replication forks. Given the character of observed disturbances, we inferred that identified loci represented sites, which, in senataxin-deficient cells, displayed increased occurrence of replication fork stalling.

To address the data in a comprehensive and unbiased manner, we developed an analytical pipeline which identified replication fork stalling sites in the respective senataxin-deficient mutants. According to our analysis, *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* mutants displayed 9, 30 and 143 replication fork stalling sites, respectively. Replication fork stalling sites identified in *sen1Δ* and *dbl8Δ* mutants did not colocalise, indicating that Sen1^{+} and Dbl8^{+} operate at distinct genomic sites. Notably, mutants depleted of both senataxins, *sen1Δdbl8Δ*, displayed 108 replication fork stalling sites, which were absent in cells depleted of *sen1Δ* or *dbl8Δ* only. Additionally, frequently, but not always, replication fork perturbations identified in *sen1Δdbl8Δ* double-mutants were significantly more pronounced than their equivalents in either *sen1Δ* or *dbl8Δ* single-mutants. Collectively, such observations implied a separation of Sen1^{+} and Dbl8^{+} activities with a degree of functional crosstalk of an unknown character.

When correlated with transcriptional activity, perturbations of replication fork progression detected in senataxin-deficient mutants colocalised with highly transcribed genes. Employing publicly available RNA-Seq and Net-Seq data, we

determined that impaired progression of replication forks specific for *sen1Δ* and *dbl8Δ* mutants strongly correlated with co-directional and head-on oriented transcriptional activity, respectively. Such observations further indicated a degree of functional divergence of the two senataxin paralogs and suggested a role of Sen1$^+$ and Dbl8$^+$ in resolution and/or prevention of co-directional and head-on RTCs, respectively. 108 replication fork stalling sites uniquely identified in *sen1Δdbl8Δ* double-mutant, however, did not correlate with co-directional or head-on oriented transcription. We speculate that, in certain contexts, functions of Sen1$^+$ and Dbl8$^+$ are, to a degree, redundant. Alternatively, depletion of both senataxin helicases is accompanied by local alterations of DNA topology which represents and obstacle for progressing replication forks. Notably, the finding that codirectional RTCs can be associated with impaired replication fork movement contrasts with the established consensus that signs of genomic instability are exclusively associated with head-on RTCs (Hamperl *et al*., 2017; Hamperl & Cimprich, 2016; Pomerantz & O'Donnell, 2010; Rudolph *et al*., 2007).

Interestingly, with the exception of non-coding *prl53* and *snoU14*, all stalling sites identified in *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* mutants colocalised with highly transcribed protein-coding genes. In case of Sen1$^+$, such findings were highly unexpected as Sen1$^+$ activity has not been previously linked to RNAP2-transcribed loci (Legros *et al*., 2014; Rivosecchi *et al*., 2019). Additionally, a subset of RNAP3-transcribed genes, unperturbed termination of which is dependent on Sen1$^+$, did not show any replication fork progression defects in *sen1Δ* mutants. Unlike Sen1$^+$, based on ChIP-qPCR experiments, it has been shown that Dbl8$^+$ is enriched at a subset of protein-coding loci including *adh1$^+$*, *SPAC27E2.11c*, *fba1$^+$*, *tef3$^+$* and *act1$^+$* (Rivosecchi *et al*., 2019). Of these, however, only *adh1$^+$* and *fba1$^+$* exhibited replication fork progression defects in *dbl8Δ* cells (*Methods*, Table 6). Interestingly, all three remaining loci, *SPAC27E2.11c*, *tef3$^+$* and *act1$^+$*, exhibited altered patterns of replication fork movement in *sen1Δdbl8Δ* mutants (*Methods*, Table 7). Notably, we determined that additional 280 highly expressed protein-coding genes were not associated with any replication defects in any of the three senataxin-deficient mutants, *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ*. Based on such results, we inferred that replication fork progression defects characteristic for senataxin-depleted cells were not a common feature of

highly transcribed loci but were rather associated with a specific subset RNAP2-transcribed protein-coding genes.

Since physical association of Sen1$^+$ and Dbl8$^+$ with RNAP2 has never been established (Rivosecchi *et al*., 2019), we hypothesised that the two senataxins paralogs perform their functions in stable or transient association with the replisome. Indeed, it has been reported that Sen1p interacts with the replisome in *S. cerevisiae*, and, based on a protein alignment, that Sen1p motif responsible for such interaction is conserved in respective *S. pombe* orthologues, Sen1$^+$ and Dbl8$^+$ (Appanah *et al*., 2020). It has been argued that, in *S. cerevisiae*, the role of Sen1p in promotion of replication progression is separate from its function in regulation of transcription (Appanah *et al*., 2020). Considering that identified replication defects characteristic for senataxin-deficient mutants do not correspond to transcription sites previously reported to be governed by Sen1$^+$ and Dbl8$^+$, it is tempting to speculate that, similarly to Sen1p in *S. cerevisiae*, Sen1$^+$ and Dbl8$^+$ helicases associate with the replisome and operatively resolve transcription-based obstacles to promote replication fork progression.

In an attempt to determine the common parameter(s) of genes associated with replication fork stalling, we addressed fundamental characteristics of stalling-associated genes including gene ontology, length, GC content and GC skew.

According to the gene ontology classification, genes associated with replication fork stalling in senataxin-deficient mutants did not form a functional cluster or a gene group, and thus were unlikely governed by specific regulatory elements.

Interestingly, in terms of gene lengths, 108 stalling-associated genes uniquely identified in *sen1Δdbl8Δ* mutants displayed significant difference when compared with 280 control highly expressed protein-coding genes. Even though biological relevance of such observation is yet to be determined, it indicates that functional crosstalk and/or redundancy of Sen1$^+$ and Dbl8$^+$ helicases could be manifesting on long highly expressed protein-coding loci. Interestingly, the longest genes encoded by human genome are associated with frequent RTCs, accumulation of RNA/DNA hybrids and signs of genomic instability (Helmrich *et al*., 2011). Unlike in *S. pombe*, however, transcription of the longest genes in human cells exceeds the length of the cell cycle, which necessitates some form of interference between replication and transcription machineries (Helmrich *et al*., 2011). In this respect, analogy between replication

defects characteristic for *sen1Δdbl8Δ* mutants and genomic instability observed at long human genes is yet to be validated.

We established that stalling-associated genes did not show an unusual GC content and GC-skew, the latter of which has been previously correlated with increased R-loop levels (Ginno *et al*., 2013, 2012). Accordingly, based on publicly available DRIPc-Seq data, when compared to a control group of highly expressed protein-coding genes, stalling-associated loci did not exhibit increased levels of R-loops. Interestingly; however, 9 replication fork stalling sites identified in *sen1Δ* mutants displayed unusual accumulation of R-loops on both strands upstream of identified replication fork stalling sites. Even though biological relevance of such observation is currently unclear, in theory, it could be indicative of uncommon topological structures and/or undergoing processes such as homologous recombination.

Collectively, we failed to provide a satisfactory explanation for the gene-specific impairments of replication fork progression identified in *sen1Δ*, *dbl8Δ* and *sen1Δdbl8Δ* cells, suggesting an involvement of other factors and/or molecular mechanisms.

Chapter 4 – Replication dynamics of ectopic heterochromatin domains

Background 4

Eukaryotic genome exists in a form of chromatin, a nucleoprotein complex which embodies genomic DNA, a plethora of structural and regulatory proteins and RNA molecules. Chromatin represents a highly sophisticated platform which regulates all aspects of DNA metabolism, including DNA replication, transcription and DNA repair (Audia & Campbell, 2016; Bannister & Kouzarides, 2011; Li & Reinberg, 2011; Yadav *et al.*, 2018). The most fundamental building block of chromatin fibre, the nucleosome core particle, comprises of ca. 147 base pairs of DNA duplex wrapped around a hetero-octameric assembly of four core histone proteins (H2A, H2B, H3 and H4) (Luger *et al.*, 1997). Individual nucleosome core particles separated by variable lengths of linker DNA form a primary chromatin fibre with a diameter of 11 nm (Li & Reinberg, 2011). Interaction of nucleosome core particles with the linker histone H1 nucleates formation of a more compact 30 nm chromatin fold, which is believed to accommodate either a solenoid or a zig-zag architecture (Li & Reinberg, 2011). It is assumed that 30nm chromatin fibres further fold into context dependent higher-order assemblies, the concrete architecture of which is still being experimentally addressed (Li & Reinberg, 2011).

Perhaps the most significant feature of chromatin is its high degree of modularity. The highly modular nature of chromatin is mainly attributed to introduction of alternative histone variants, DNA methylation, and wide range of post-translational modifications (PTMs) of histones, also known as histone marks (Audia & Campbell, 2016; Bannister & Kouzarides, 2011; Martire & Banaszynski, 2020; Moore *et al.*, 2013). Fundamentally, the scientific community differentiates between two states of chromatin, which are commonly referred to as euchromatin and heterochromatin. While euchromatin is transcriptionally active and accommodates more accessible fold, heterochromatin represents a compact environment, which inhibits the activity of embedded gene promoters (Li & Reinberg, 2011).

In *S. pombe* and other eukaryotic species, stable (constitutive) heterochromatin associates with pericentromeric regions, telomeres and a fraction of rDNA repeats (Allshire & Ekwall, 2015; Allshire & Madhani, 2018; Bi, 2014).

Additionally, in *S. pombe* and other yeast species, constitutive heterochromatin mediates repression of a silent mating type cassette at the mating type locus (Allshire & Ekwall, 2015; Bi, 2014). Eukaryotic cells also feature highly modular (facultative) heterochromatin. In vegetative *S. pombe* cells, facultative heterochromatin is believed to play a role in repression of genes specifically expressed during meiosis (Zofall *et al*., 2012). In higher eukaryotes, facultative heterochromatin suppresses self-propagating retrotransposons, contributes to genetic stability of low complexity satellite repeats and plays an important role in cell differentiation and inactivation of the X chromosome (Cramer *et al.*, 2016; Janssen *et al.*, 2018; Peng & Karpen, 2007; Żylicz & Heard, 2020).

According to the current paradigm, chromatin state is mainly dictated by post-translational modifications of unstructured N-terminal tails of histones H3 and H4 and, in higher eukaryotes, DNA methylation (Allshire & Madhani, 2018). The most studied PTMs of histones associated with transcriptionally active chromatin include acetylation of histone H3 at lysine 4 (H3K9ac), methylation of histone H3 at lysine 4 (H3K4me) and acetylation of histone H4 at lysine 16 (H4K16ac)(Audia & Campbell, 2016; Shogren-Knaak *et al*., 2006). Histone marks predominantly characteristic for heterochromatin include di- or tri-methylation of histone H3 at lysine 9 (H3K9me2/3) and trimethylation of histone H3 at lysine 27 (H3K27me3) (Allshire & Ekwall, 2015; Audia & Campbell, 2016; Janssen *et al*., 2018).

In general, heterochromatin domains are established and maintained by concerted actions of histone deacetylases (HDACs), histone methyltransferases (HMTs), chromatin remodelling factors and other regulators (Audia & Campbell, 2016; Bannister & Kouzarides, 2011). In *S. pombe*, HDACs, including Clr3^{+}, Clr6^{+} and Sir2^{+}, contribute to deacetylation of various lysine residues of histone H3 (Wirén *et al*., 2005). HDAC-mediated removal of acetyl moieties creates a substrate for the HMT Clr4^{+}, which catalyses deposition of H3K9me2/3. Introduction of H3K9me2/3 creates a docking platform for chromodomain-featuring factors including Clr4^{+} itself, and three members of the HP1 protein family, Chp1^{+} and Chp2^{+} and Swi6^{+} (Ekwall *et al*., 1995; Ishida *et al*., 2012; Ivanova *et al.*, 1998; Maksimov *et al*., 2018). Interaction of Clr4^{+} with established H3K9me2/3 allows self-propagated spread and maintenance of heterochromatin (Ivanova *et al*., 1998; Zhang *et al.*, 2008). Chp1^{+} functions as a

component of RNA-induced initiation of transcriptional gene silencing (RITS) complex, which also embodies Tas3$^+$ and the argonaut protein Ago1$^+$. It is believed that Chp1$^+$ promotes association of the RITS machinery with heterochromatin and contributes to establishment of the H3K9me2/3 signature and deposition of Swi6$^+$ (Ishida *et al*., 2012; Partridge *et al.*, 2002; Schalch *et al.*, 2011). Chp2$^+$ facilitates recruitment of the SHREC complex, which embodies activities of the HDAC Clr3$^+$ and the chromatin remodelling factor Mit1$^+$ (Maksimov *et al*., 2018).

Chromatin-bound Swi6$^+$ is recognised by factors featuring a chromoshadow domain. Notably, Swi6$^+$ itself carries a chromoshadow domain, which promotes Swi6$^+$ dimerization and contributes to formation of a compact heterochromatin architecture (Canzio *et al*., 2011; Cowieson *et al.*, 2000). Other examples of chromoshadow domain factors characterised in *S. pombe* are the MCM helicase loader Cdc18$^+$ (the orthologue of Cdc6) and the regulatory subunit of the DDK kinase Dfp1$^+$ (Hayashi *et al*., 2009; Li, *et al.*, 2011). Interestingly, in *S. pombe*, binding of Cdc18$^+$ and Dfp1$^+$ to Swi6$^+$ contributes to early replication of the heterochromatin-embedded mating type locus and pericentromeric regions, but not telomeres (Hayashi *et al*., 2009; Li *et al*., 2011).

Perhaps the best characterised mechanism of heterochromatin foundation emerged from studies addressing epigenetic regulation of centromeric functions in fission yeast. In *S. pombe*, each centromere features a central segment flanked by outer repeats, which are composed of several kb long *dh* and *dg* elements. Each central segment features a central core, which is characterised by deposition of the histone H3 variant CenH3/CENP-A (Cnp1$^+$) flanked by the inner-most repeats (Pidoux & Allshire, 2004). Notably, to allow assembly of functional kinetochores, only outer repeats undergo heterochromatinisation (Allshire & Ekwall, 2015).

In *S. pombe*, establishment of pericentromeric heterochromatin is dependent on short interfering RNA (siRNA) machinery (Martienssen & Moazed, 2015; Volpe *et al*., 2002). Nucleation of pericentromeric heterochromatin starts with S phase-specific bidirectional transcription of outer centromeric repeats, forward and reverse transcripts of which anneal and form double-stranded RNA (dsRNA) species (Volpe *et al*., 2002). In *S. pombe*, process of dsRNA formation is further facilitated by the action of RNA-directed RNA polymerase Rdp1$^+$, which synthesises complementary strands of nascent single-stranded *dh* and *dg* transcripts (Sugiyama *et al.*, 2002). It is believed

that Rdp1^{+} amplifies the siRNA signal and, in *S. pombe*, represents an indispensable heterochromatin-forming factor (Sugiyama *et al*., 2005; Volpe *et al*., 2002). dsRNA species homologous to centromeric repeats are subsequently processed by dicer (RNase III class ribonuclease, Dcr1^{+} in *S. pombe*) into short (21-24 bp) siRNA molecules, which are incorporated by the RITS complex. Guided by the accommodated siRNA, the RITS complex localises to pericentromeric regions and recruits the Clr4^{+} complex (CLRC), which catalyses deposition of initial H3K9me2/3 (Verdel *et al*., 2004; Zofall & Grewal, 2006). It is also believed that RITS and CLRC complexes attract Rdp1^{+}, which amplifies the generation of dsRNAs, and thus contributes to the self-reinforcing nature of heterochromatin (Sugiyama *et al*., 2005).

Interestingly, the mating type locus (*mat2/3*) also features a centromere-homologous (*cenH*) site, bidirectional transcription of which can drive heterochromatin formation via siRNA-mediated recruitment of Clr4^{+} (Hall *et al*., 2002). In addition to the siRNA-mediated pathway, heterochromatin at *mat2/3* locus can be established via a low-efficiency mechanism involving ATF/CREB family transcription factors Atf1^{+} and Pcr1^{+} (Jia *et al.*, 2004). At telomeres, shelterin complex subunit Taz1^{+} recruits Clr4^{+}, which introduces H3K9me2/3, and thus propagates heterochromatinisation (Kanoh, *et al.*, 2005). Small non-coding RNA interacting with Piwi proteins (piRNA) have been implicated in heterochromatin-dependent silencing of transposon elements in *Drosophila melanogaster* and *Caenorhabditis elegans* (Das *et al*., 2008; Sienski *et al.*, 2012). In higher eukaryotes, many factors have been associated with repressive chromatin folds, however, the precise mechanisms governing heterochromatin formation are poorly understood (Fodor *et al.*, 2010).

It is widely assumed that, due to its highly compact structure, heterochromatin represents an obstacle for replication fork progression. Experimental evidence supporting this notion, however, is scarce, and mainly based on studies addressing replication of human pericentromeric regions, which are largely composed of short (ca. 171 bp) arguably difficult to replicate AT-rich α satellite repeats (Aze *et al.*, 2016; Collins *et al*., 2002; Mendez-Bermudez *et al*., 2018). In order to test whether heterochromatin represents a structure which blocks or slows replisome progression in a sequence-independent manner, we aspired to utilise polymerase usage

sequencing (Pu-Seq) and analyse replication dynamics of ectopically heterochromatinised non-repetitive regions.

Results 4

In order to analyse replication dynamics of heterochromatinised regions, we aimed to combine Pu-Seq methodology and a previously established system, which allows induction of artificial heterochromatin domains. The heterochromatinisation system utilises anhydrotetracycline (ahTET)-regulatable tethering of artificial $TetR^{Off}$-$Clr4^{+}$ construct, which is composed of chromodomain-depleted $Clr4^{+}$ histone methyltransferase and the "OFF" variant of tetracycline repressor ($TetR^{Off}$), onto manufactured genomic loci featuring arrays of tetracycline operator (tetO) sequences (Audergon *et al*., 2015; Ragunathan, Jih, & Moazed, 2015). In the "OFF" version of the system, binding of $TetR^{Off}$-$Clr4^{+}$ to tetO arrays is inhibited by ahTET. The $TetR^{Off}$-$Clr4^{+}$ construct, as well as the original tetO site, composed of four tetO repeats ($tetO_4$) integrated into *ura4*$^{+}$ locus together with *ade6*$^{+}$ marker (*ade6*$^{+}$-$tetO_4$), were kindly provided by R. Allshire and colleagues (Audergon *et al*., 2015). To further improve the system, we constructed an additional tetO site which contained seven tetO repeats ($tetO_7$) and kanamycin resistance (KanR) marker integrated into chromosome 1 at position 3,325,162 (KanR-$tetO_7$).

For the pilot Pu-Seq experiment, we constructed strains which carried *ade6*$^{+}$-$tetO_4$ and KanR-$tetO_7$ sites, *TetR*^{Off}*-clr4*$^{+}$, *rnh201-RED*, which abrogates nucleotide excision repair (Naiman *et al*., 2021), and either *cdc6*^{L591G} or *cdc20*^{M630F}, which compromise base selectivity of Polδ and Polε, respectively (Figure 4.1). In the absence of ahTET, constructed strains displayed increased sensitivity to geneticin (G418) and reduced growth in the absence of supplementary adenine (Figure 4.1). We argued that observed growth defects were consequential to KanR and *ade6*$^{+}$ silencing, which implied successful induction of heterochromatin at both *ade6*$^{+}$-$tetO_4$ and KanR-$tetO_7$ loci. Notably, ectopic heterochromatinisation on its own was associated with minor growth defect, arguably due to silencing of tetO-proximal genes required for normal growth (Figure 4.1). Overall, this preliminary experiment indicated that the constructed system was functional.

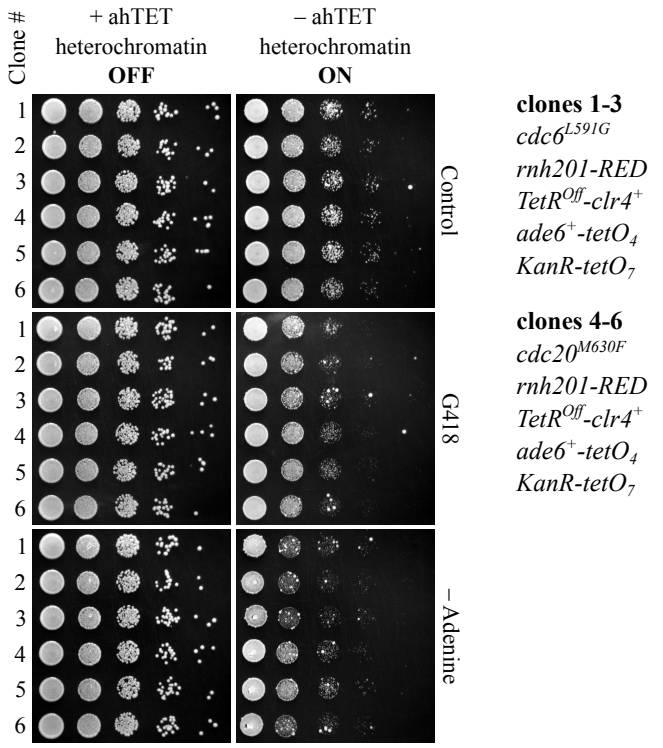

Figure 4.1 – Verification of the heterochromatin induction system. Proliferation of cells with listed genotypes under indicated conditions. Cells were 10-fold serially diluted and spotted onto EMM plates containing indicated additives: ahTET (2.5 µg/mL), G418 (200 µg/mL). "– Adenine" signifies the lack of supplementary Adenine. Control plates contained uracil, leucine and adenine (225 mg/L each). Spot tests were imaged 3 days after plating.

Unfortunately, we later realised that all reported strains were constructed in the *ade6-704* genetic background. The endogenous *ade6-704* allele, which only carries T645A point mutation, represented an almost perfect copy of the *ade6*$^{+}$ gene which was used to label *ade6*$^{+}$-tetO$_4$ integration at *ura4*$^{+}$ locus. The presence of identical sequences made subsequent Pu-Seq analysis, which is based on genomic mapping of short sequencing reads, unreliable. Nevertheless, to acquire potentially insightful preliminary data, we decided to proceed with the Pu-Seq analysis of the KanR-tetO$_7$ locus.

We utilised the newly developed Pu-Seq procedure (discussed in Chapter 1) and analysed Polδ and Polε activities across the region spanning the KanR-tetO$_7$ construct under conditions repressing (+ ahTET) or inducing (- ahTET) heterochromatin formation. In this experiment, we utilised a classic Pu-Seq analysis approach, which calculates polymerase usage of Polδ and Polε on the top and the bottom strands. For instance, Polδ usage on the top strand at coordinate "i", Usageδ(i)^{top}, is calculated as Usageδ(i)^{top} = δ(i)^{top} / [δ(i)^{top} + ε(i)^{top}], where δ(i)^{top} and ε(i)^{top} represent numbers of normalised sequencing reads at position "i" in mutants expressing Cdc6^{L591G} and Cdc20^{M630F}, respectively. Replication initiation and

termination sites were deduced from differential polymerase usage values as described previously (Daigaku *et al.*, 2015).

According to the pilot experiment, ectopic heterochromatinisation dramatically altered local replication profiles (Figure 4.2). According to our interpretation, formation of heterochromatin resulted in activation of additional origins, which were in close proximity to the KanR-tetO$_7$ site (Figure 4.2). We theorised that activation of extra origins could result from recruitment of origin firing factors such as DDK kinase or the Cdc6-orthologue Cdc18$^+$, which have been shown to associate with heterochromatin at pericentromeric regions and mating type locus in *S. pombe* (Hayashi *et al.*, 2009; Li *et al.*, 2011). Alternatively, unusual replication initiation pattern could represent a compensation for replication fork defects arising either from heterochromatinisation or TetR^{Off}-Clr4$^+$ binding to tetO$_7$ array.

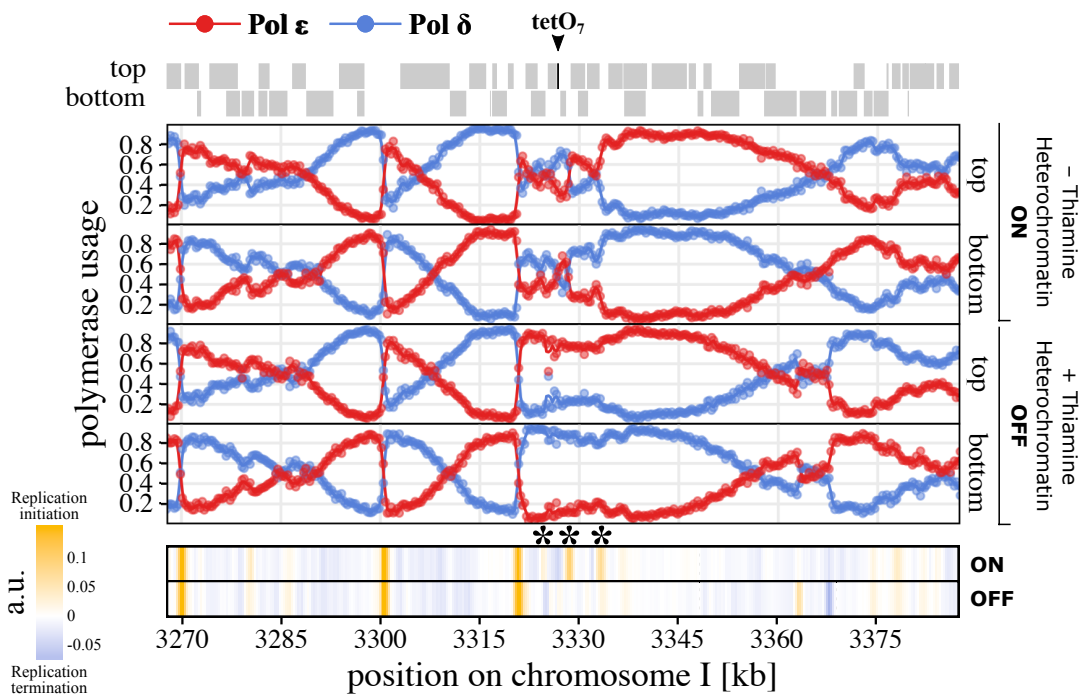

Figure 4.2 – Replication profiles of an ectopic heterochromatin domain. Pilot Pu-Seq data demonstrating changes in replication progression accompanying induction of synthetic heterochromatin at KanR-tetO$_7$ locus (chromosome 1, position 3,325,162). Top panel – genomic annotation displaying genes encoded by the top and the bottom DNA strands. Middle panel – Polε and Polδ usage tracks. Bottom panel – heatmap representation of replication initiation and termination sites. Position of tetO$_7$ array is indicated. Asterisks indicate replication initiation sites induced by heterochromatinisation.

The fact that induced heterochromatin stimulated origin activity complicated the interpretation of the acquired data, and thus, to be able to understand the impact of heterochromatin on replication progression in more detail, we decided to expand the existing experimental system. To test the possibility that chromatin-bound TetR^{Off}-Clr4^{+} represents a potent replication block, we constructed catalytically inactive chromodomain-depleted TetR^{Off}-Clr4^{H410K} construct and integrated it into chromosome 1 at position 5,230,932 by Cre-Lox mediated cassette exchange (Figure 4.3A). The catalytically dead *clr4^{H410K}* allele was kindly suggested by R. Allshire. To be sure our experiments were not introducing any locus-specific bias, we also de-novo generated functional chromodomain-depleted TetR^{Off}-Clr4^{+} and, using Cre-Lox mediated cassette exchange, integrated it into the same locus as the inactive TetR^{Off}-Clr4^{H410K} variant (Figure 4.3A).

We also aimed to test whether firing of additional origins, which accompanied ectopic heterochromatinisation, was consequential to the recruitment of DDK kinase and/or the MCM loader Cdc18^{+}. In *S. pombe*, DDK kinase is composed of the catalytic subunit Hsk1^{+} and the regulatory moiety Dfp1^{+}, which facilitates interaction with the HP1-orthologue Swi6^{+}. The most straightforward experiment to do would be to test whether disruption of Swi6^{+}-mediated recruitment of Cdc18^{+} and/or Dfp1^{+} supresses origin firing at ectopically-heterochromatinised loci. Inconveniently, it is well documented that disruption of *dfp1^{+}*, *hsk1^{+}* or *cdc18^{+}* results in defective replication and consequential cellular pathologies (Kelly *et al*., 1993; Ogino *et al*., 2001; Patel *et al*., 2008). S. Forsburg kindly provided us with the *dfp1-3A* and *cdc18-I43A* mutants, which impair association with Swi6^{+} (Hayashi *et al*., 2009; Li *et al*., 2011). At the moment, however, we do not yet know, whether *dfp1-3A* and *cdc18-I43A* alleles can be introduced into cells carrying genetic background required for Pu-Seq.

As a potential alternative, we decided to construct mutants expressing TetR^{Off}-Dfp1^{+} and determine whether artificial DDK tethering onto loci of interest results in changes in replication progression analogous to those observed in cells inducing heterochromatin (Figure 4.3A). In order to avoid changes in Dfp1^{+} expression, which also compromise genome stability (Patel *et al*., 2008; Takeda *et al*., 1999), we fused the endogenous copy of *dfp1^{+}* with the N-terminal TetR^{Off} tag. *TetR^{Off}-dfp1^{+}* strain was prepared using PCR based gene targeting (Bähler *et al*., 1998) using a custom *TetR^{Off}-*

dfp1⁺-BleoR DNA fragment generated by overlap extension PCR. Constructed *TetR^{Off}-dfp1⁺* mutants retained WT-like cellular and nuclear morphology (Figure 4.3C) and, based on personal observation, did not show any growth defects. Collectively, brief examination of *TetR^{Off}-dfp1⁺* cells indicated that the N-terminal TetR^{Off}-tagging of Dfp1⁺ did not disrupt normal DDK function.

Moreover, we included additional two tetO$_7$ sites integrated into chromosome 2 at position 1,389,186 (tetO$_7$-NatR) and chromosome 3 at position 1,609,353 (*ura4⁺*-tetO$_7$) (Figure 4.3B). Both constructs were produced by Cre-Lox mediated cassette exchange (Watson *et al*., 2008).

In all three genetic backgrounds (*TetR^{Off}-clr4⁺, TetR^{Off}-clr4^{H410K}*, *TetR^{Off}-dfp1⁺*) we introduced three tetO$_7$ sites (KanR-tetO$_7$, tetO$_7$-NatR, *ura4⁺*-tetO$_7$), *rnh201-RED* and either *cdc6^{L591G}* or *cdc20^{M630F}*, which are required for Pu-Seq. To test the functionality of the system, in all generated strains, we utilised RT-qPCR and measured relative transcript levels of KanR, NatR and *ura4⁺*, the selection markers associated with distinct tetO$_7$ sites, under conditions repressing (+ ahTET) or inducing (- ahTET) heterochromatin formation. As expected, in the absence of ahTET, relative transcript levels of the respective tetO$_7$-associated selection markers were significantly decreased in cells expressing TetR^{Off}-Clr4⁺, but not TetR^{Off}-Clr4^{H410K} and TetR^{Off}-Dfp1⁺ (Figure 4.3D).

Even though the obtained RT-qPCR results indicated that the system was functional, further examination was necessary. Due to the lack of time and equipment malfunction, however, we have not managed to perform additional validation experiments. Planned experiments included ChIP-qPCR analysis of TetR^{Off} binding to the target tetO$_7$ sites, ChIP-qPCR analysis of H3K9me2 to further confirm successful heterochromatin formation and ChIP-Seq analysis of H3K9me2 to define the spread of induced heterochromatin. Following these analyses, we would perform Pu-Seq experiment to determine replication dynamics around constructed tetO$_7$ sites in cells expressing TetR^{Off}-Clr4⁺, TetR^{Off}-Clr4^{H410K} or TetR^{Off}-Dfp1⁺.

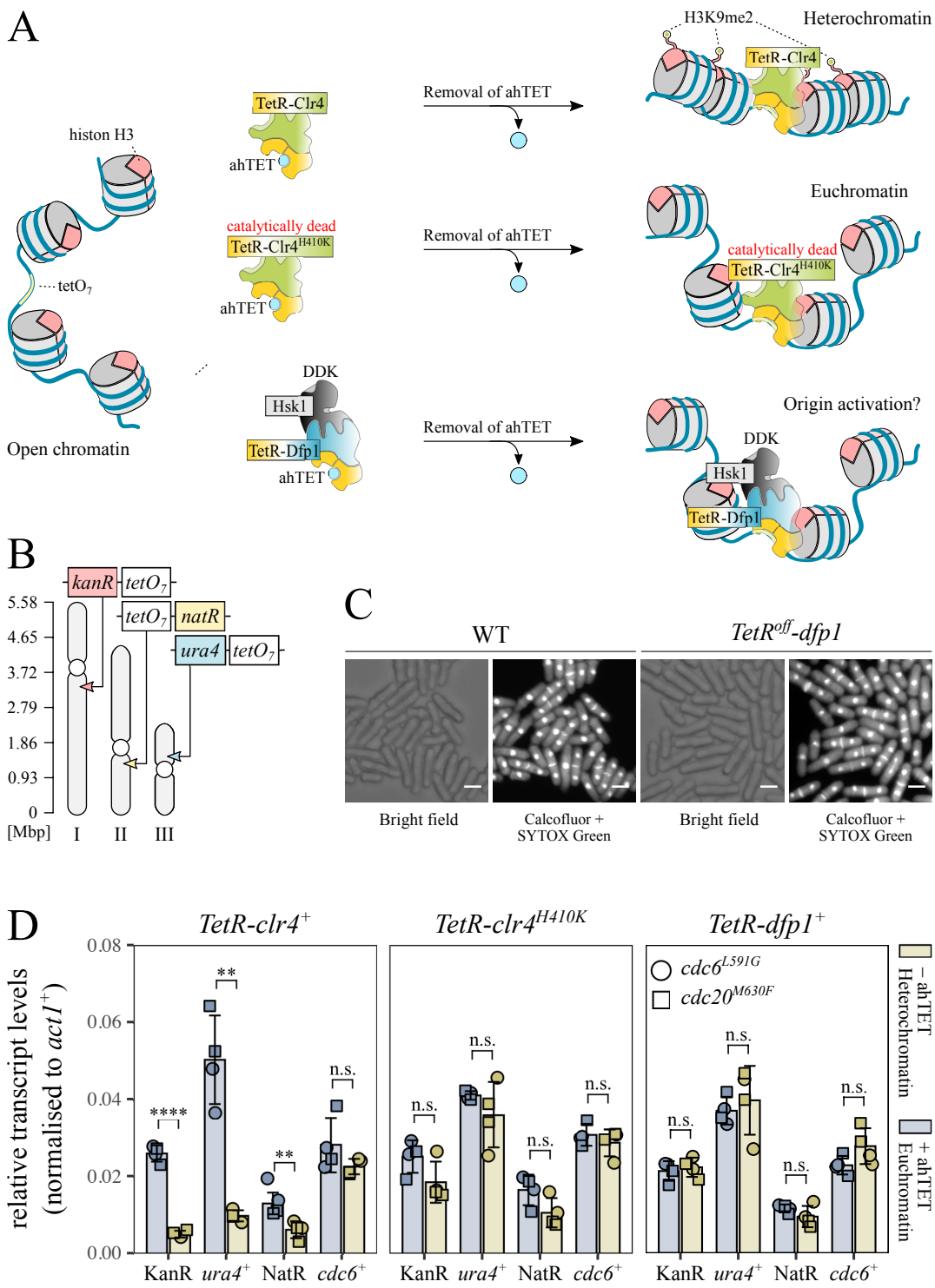

Figure 4.3 – Outline of the improved heterochromatin induction system. (A) Schematic diagram of expected TetR^{Off}-Clr4^{+}, TetR^{Off}-Clr4^{H410K} and TetR^{Off}-Dfp1^{+} functions. **(B)** Schematic diagram of *tetO$_7$-KanR*, *tetO$_7$-NatR* and *ura4^{+}-tetO$_7$* chromosomal integrations. **(C)** Microscopy images of WT and *TetR^{Off}-dfp1^{+}* cells fixed with 70% ethanol and stained with Calcofluor (5 µg/mL) and SYTOX Green (5 µM). Scale bar represents 5 µm. **(D)** RT-qPCR analysis of relative KanR, NatR and *ura4^{+}* transcript levels normalised to *act1^{+}*. *cdc6^{+}* locus was used as a control. All tested strains carried

one of the three TetR^{Off}-tagged constructs indicated in (A), all three tetO$_7$ sites indicated in (B), rnh201-RED and either *cdc6^{L591G}* or *cdc20^{M630F}*. Bars represent means of 4 independent experiments. Error bars represent means ± respective standard deviations. Points represent results of individual experiments. Statistical significance was determined by the unpaired two-sample t-test. **** $p \leq 0.0001$; ** $p \leq 0.01$; n.s. = nonsignificant.

Discussion 4

It is widely assumed that heterochromatin represents an endogenous barrier, which affects replication fork progression. Experimental evidence supporting such claim, however, is underwhelming and, in our opinion, additional investigation is required. In this chapter, we summarise our efforts to combine Pu-Seq methodology with TetR-tetO based molecular system allowing targeted induction of synthetic heterochromatin. In our preliminary experiment, we induced formation of ectopic heterochromatin by tethering of the TetR^{Off}-Clr4^{+} construct to the KanR-tetO$_7$ locus on chromosome 1, and utilised Pu-Seq to analyse consequential changes in replication dynamics. According to this preliminary experiment, ectopic heterochromatinisation resulted in local activation of dormant or low-efficiency origins. Such observation is in accordance with studies, which show that heterochromatin at pericentromeric regions and silent mating type locus recruits the MCM loader Cdc18^{+} and the regulatory subunit of DDK kinase Dfp1^{+} (Hayashi *et al*., 2009; Li *et al*., 2011). Dramatic changes in replication profiles at heterochromatinised KanR-tetO$_7$ locus prevented us from establishing whether compact chromatin architecture affects replication fork speed, but led us to build a more elaborate system, which involves two additional tetO$_7$ sites (tetO$_7$-NatR, *ura4^{+}*-tetO$_7$), newly constructed *TetR^{Off}-clr4^{+}*,catalytically inactive *TetR^{Off}-clr4^{H410K}* and *TetR^{Off}-Ddp1^{+}*. Furthermore, we acquired mutants expressing Dfp1-3A and Cdc18-I43A, which exhibit reduced heterochromatin binding capacity (Hayashi *et al*., 2009; Li *et al*., 2011). At the moment; however, we are not yet sure whether *dfp1-3A* and *cdc18-I43A* mutations can be introduced into cells carrying genetic background required for Pu-Seq.

Using the improved system we expect to demonstrate the following points. 1) Tethering of TetR^{Off}-Clr4^{+}, but not TetR^{Off}-Clr4^{H410K}, induces heterochromatinisation and consequential activation of dormant origins of replication in close proximity to respective tetO$_7$ sites. 2) Tethering of TetR^{Off}-Dfp1^{+} increases local origin activity,

similarly to TetR^{Off}-Clr4^{+}. 3) Induction of local heterochromatin by TetR^{Off}-Clr4^{+} binding to tetO$_7$ sites in cells carrying *dfp1-3* and/or *cdc18-I43A* is not accompanied by any changes in replication progression.

So far, we have only demonstrated that tethering of TetR^{Off}-Clr4^{+} to all three tetO$_7$ sites results in reduced expression of associated selection markers. Such observation indicates transcriptional silencing, and therefore successful formation of heterochromatin. It is evident that further system validating experiments are required. These include: 1) ChIP-qPCR analysis of ahTET-regulatable binding of all TetR^{Off}-tagged constructs to respective tetO$_7$ sites, 2) ChIP-qPCR analysis of H3K9me2 following the tethering of TetR^{Off}-tagged constructs and 3) ChIP-Seq analysis of TetR^{Off}-Clr4^{+}-dependent H3K9me2 to determine the spread of synthetic heterochromatin.

Even though our work on this project is not over yet, we believe we generated a robust molecular tool kit, which allows comprehensive analysis of replication progression across artificial heterochromatin domains at three distinct genomic sites. We hope that further utilisation of this system will lead to better understanding of how compact heterochromatin architecture influences replication progression. Additionally we argue that the created system will be useful in future studies addressing processes of DNA metabolism in the context of compact chromatin environment.

Chapter 5 – Characterisation of Cdm1$^+$ function in fission yeast

Background 5

In eukaryotes, polymerase delta (Polδ) represents a core constituent of DNA synthesis machinery. It has been firmly established that Polδ is responsible for OF synthesis during canonical replication (Daigaku *et al*., 2015; Garbacz *et al*., 2018; Miyabe *et al*., 2011) and plays an essential part in HR and DNA-repair (McVey *et al.*, 2016; Mocquet *et al*., 2008; Prindle & Loeb, 2012).

In mammals and fission yeast *S. pombe*, the Polδ complex is composed of 4 distinct subunits: p125/Cdc6$^+$, p50/Cdc1$^+$, p68/Cdc27$^+$ and p12/Cdm1$^+$ (Hughes *et al*., 1999; Iino & Yamamoto, 1997; Lee *et al*., 1991; Liu *et al*., 2000; MacNeill *et al*., 1996; Reynolds *et al*., 1998; Jian Zhang *et al*., 1995). It has been generally accepted that Polδ operates as a heterotetramer (Lancey *et al*., 2020), however, a 5-subunit Polδ architecture featuring a p12/Cdm1$^+$ dimer has been also suggested (Khandagale *et al.*, 2019). While the largest p125/Cdc6$^+$ subunit embodies 5'-3' polymerase and 3'-5' exonuclease activities, the remaining p50/Cdc1$^+$, p68/Cdc27$^+$ and p12/Cdm1$^+$ subunits facilitate structural and regulatory roles (Lancey *et al*., 2020). It has been shown that p125/*cdc6$^+$*, p50/*cdc1$^+$* and p68/*cdc27$^+$* are indispensable for cell survival, whereas the biological role of p12/*cdm1$^+$* is non-essential. Strikingly, p12/Cdm1$^+$ embodies a PCNA interacting protein (PIP) motif, suggesting a possibility that p12/Cdm1$^+$ function manifests via interaction with PCNA (Li *et al*., 2006). Interestingly, budding yeast does encode a p12/Cdm1$^+$ counterpart and features a 3-subunit Polδ composed of Pol3p, Pol31p and Pol32p, which are the orthologues of p125/Cdc6$^+$, p50/Cdc1$^+$ and p68/Cdc27$^+$, respectively (Jain *et al*., 2019). Currently, it is not clear whether p12/Cdm1$^+$ function is completely absent in *S. cerevisiae* or is substituted for by other factors or mechanisms.

According to *in vitro* studies employing reconstituted human Polδ complex, Polδ can operate in two functionally distinct forms, Polδ_4 embodying all 4 subunits and Polδ_3 which lacks p12 (Meng *et al*., 2010; Podust *et al*., 2002; Xie *et al*., 2002). It has been reported that p12 increases the rate of phosphodiester bond formation which, in turn, enhances the rate of polymerisation but, at the same time, decreases

proofreading capacity. Thus, while p12-lacking Polδ_3 shows lower rate of polymerisation but higher fidelity, Polδ_4 exhibits higher rate of polymerisation (estimated to 4.6-fold higher than Polδ_3) at the cost of reduced accuracy (Huang, Akashi, *et al*., 2010; Meng *et al*., 2010, 2009; Podust *et al*., 2002). Upon encounter of the 5' end of the blocking oligonucleotide in an *in vitro* primer extension assay, Polδ_3 idles, only displacing 1-5 nucleotides, whereas Polδ_4 facilitates long-track strand displacement synthesis (Lin *et al*., 2013). 5' flaps generated by idling Polδ_3 represent a preferable substrate for the flap endonuclease FEN1, an essential OF processing factor, making Polδ_3 a likely candidate for the enzyme catalysing the lagging strand synthesis during canonical replication (Balakrishnan & Bambara, 2013; Lin *et al*., 2013). Similarly to OF maturation, nick translation synthesis during nucleotide excision repair (NER) is orchestrated by Polδ and FEN1, suggesting an additional role of Polδ_3 in DNA repair (Lin *et al*., 2013; Mocquet *et al*., 2008). As a complement to Polδ_3, Polδ_4 featuring strand displacement synthesis activity is believed to facilitate extension of the invading strand during HR (Lin *et al*., 2013). In support of this notion, the Bloom syndrome RecQ like helicase BLM directly interacts with p12 and stimulates Polδ_4 strand displacement synthesis *in vitro* and *in vivo* (Selak *et al*., 2008). Accordingly, human A549 cells depleted of p12 by CRISPR-Cas9 develop decreased HR efficiency and show increased sensitivity to poly (ADP-ribose) polymerase (PARP) inhibitors (Zhang *et al*., 2019), which are known to induce lethality in HR-deficient cells (Bryant *et al*., 2005; Farmer *et al*., 2005).

In agreement with predicted catalytic divergence of Polδ_3 and Polδ_4, p12 is targeted for degradation during S phase and as a consequence of genotoxic stress triggered by UV, methyl methanesulfonate (MMS) and aphidicolin (Darzynkiewicz *et al*., 2015; Terai *et al*., 2013; Zhang *et al*., 2013, 2007; Hong Zhao *et al*., 2014). Degradation of p12 observed in cells undergoing replication as well as cells with compromised genome integrity is dependent on CRL4^{CDT2} E3-ubiquitin ligase which recognises a PIP-degron motif situated on the N-terminus of p12 (Terai *et al*., 2013; Zhang *et al*., 2013). Interestingly, UV-induced p12 degradation is abrogated in ATR-deficient cells, indicating a role of checkpoint signalling in regulation of p12 levels during DNA-damage response (Zhang *et al*., 2007).

Strikingly, expression of p12 is reduced in virtually all tested small cell lung cancer (SCLC) and some non-small cell lung cancer (NSCLC) tissues (Huang *et al.*, 2010). It has been shown that further diminishment of p12 levels in a subset of SCLC cell lines including Calu6, ACC-LC- 319 and PC-10 results in decreased colony forming capacity and increased genomic instability manifested by formation of lobed nuclei and increased occurrence of chromosomal pathologies (Huang, *et al.*, 2010a; 2010b). Furthermore, it has been reported that adequate proliferation rate is dependent on significant upregulation of p12 in murine aortic endothelial (MAE) cells exposed to the fibroblast growth factor-2 (FGF2) (Dell'Era *et al*., 2005). Interestingly, despite the apparent defect in HR, p12-defficient A549 cells retain WT-like characteristics when grown under normal conditions and only develop growth defects when treated with DNA crosslinking agents Cisplatin and Mitomycin C (Zhang *et al*., 2019). Currently, it is unclear whether the discrepancy between phenotypes associated with the loss of p12 in different cell lines represents a true biological phenomenon or experimental artefact.

In *S. pombe*, transcription of *cdm1*$^{+}$, the orthologue of p12, is cell-cycle regulated (Rustici *et al*., 2004). According to J. Bähler and colleagues, *cdm1*$^{+}$ falls into the Ace2^{+}-regulated cluster of genes which exhibit increased expression in late mitosis and G_1 (Rustici *et al*., 2004). Additionally, transcription of *cdm1*$^{+}$ is upregulated during the process of spore maturation following meiotic division (Mata *et al.*, 2002). Differential regulation of Cdm1^{+} protein levels, however, has not been experimentally addressed and it is yet to be determined whether oscillations in *cdm1*$^{+}$ mRNA represent a biologically relevant phenomenon. Additionally, to our best knowledge, it is not known whether Cdm1^{+} undergoes degradation during replication and DNA damage response.

Similarly to p12-defficient A549 cells, *S. pombe* mutants depleted of Cdm1^{+} do not develop defects in growth or cell morphology (Reynolds *et al*., 1998). Interestingly, over-expression of Cdm1^{+} rescues temperature-sensitive strains carrying defective alleles of essential Polδ subunits (*cdc6-121*, *cdc1-P13*, *cdc27-P11*) and the DNA replication protein Cdc24^{+} (*cdc24-M38*) (Reynolds *et al*., 1998). Even though potentially relevant, the molecular details behind such observation are currently unclear. Furthermore, it has been established that Cdm1^{+}-deficient cells are not

sensitive to DNA-damaging agents including MMS, hydroxyurea (HU), bleomycin (BM) and UV (Reynolds *et al*., 1998). We reason that closer characterisation of Cdm1$^+$ would improve our understanding of how and, more importantly, why Polδ activity is modulated. Thus, in accordance with this notion, we set out to investigate the role of Cdm1$^+$ in replication and DNA repair.

Results 5

General characterisation of cdm1Δ mutant.

In order to establish that Cdm1$^+$ is indeed not required for cell survival and DNA-damage response, as postulated in (Reynolds *et al*., 1998), we *de novo* constructed *cdm1Δ* mutant by PCR-based gene targeting (Bähler *et al*., 1998). *cdm1Δ* cells were viable and did not display impaired growth (Figure 4.1A) or defects in cell morphology (personal observation). Additionally, *cdm1Δ* cells did not exhibit increased sensitivity to HU, camptothecin (CPT), MMS, UV, Mitomycin C and Cisplatin (Figure 5.1A).

Provided that replication mutants are frequently characterised by apparent cellular and nuclear defects (Hayles *et al*., 2013), the seemingly WT-like characteristics of *cdm1Δ* mutant indicated that Cdm1$^+$ does not play an important role in canonical replication. Nevertheless, we wanted to explore whether replication is indeed unaffected in Cdm1$^+$-deficient cells. To analyse replication dynamics in *cdm1Δ* mutant we aimed to utilise Pu-Seq which provides readout on origin firing, replication fork directionality and polymerase usage across the genome (Daigaku *et al*., 2015). Employing standard crossing, we combined deletion of *cdm1$^+$* with *rnh201-RED* allele, which impairs RER , and either *cdc6L591G* or *cdc20M630F*, which respectively impair selectivity of Polδ or Polε, leading to higher frequency misincorporated rNTPs (Daigaku *et al*., 2015). Based on the *in vitro* studies addressing the enzymatic properties of Polδ_3 and Polδ_4 complexes, in theory, the lack of Cdm1$^+$ could result in slower replication fork progression. If such notion proved to be correct, increased firing of dormant/low-efficiency origins would be expected. Pu-Seq analysis, however, did not reveal any

changes in polymerase usage and origin firing across the genome (Figure 5.1B), indicating that Cdm1⁺ indeed does not play a role in canonical replication.

To test the possibility that Cdm1⁺ plays a role in spore maturation, as suggested by Mata *et al*., (2002), we sporulated diploids carrying two (*cdm1*$^{+/+}$), one (*cdm1*$^{\Delta/+}$) or none (*cdm1*$^{\Delta/\Delta}$) *cdm1* allele(s). Virtually all asci originating from *cdm1*$^{\Delta/+}$ and *cdm1*$^{\Delta/\Delta}$ diploids contained 4 spores of regular shapes and sizes, indicating that morphogenesis of spores is largely unaffected in Cdm1-defficient cells (Figure 5.1C).

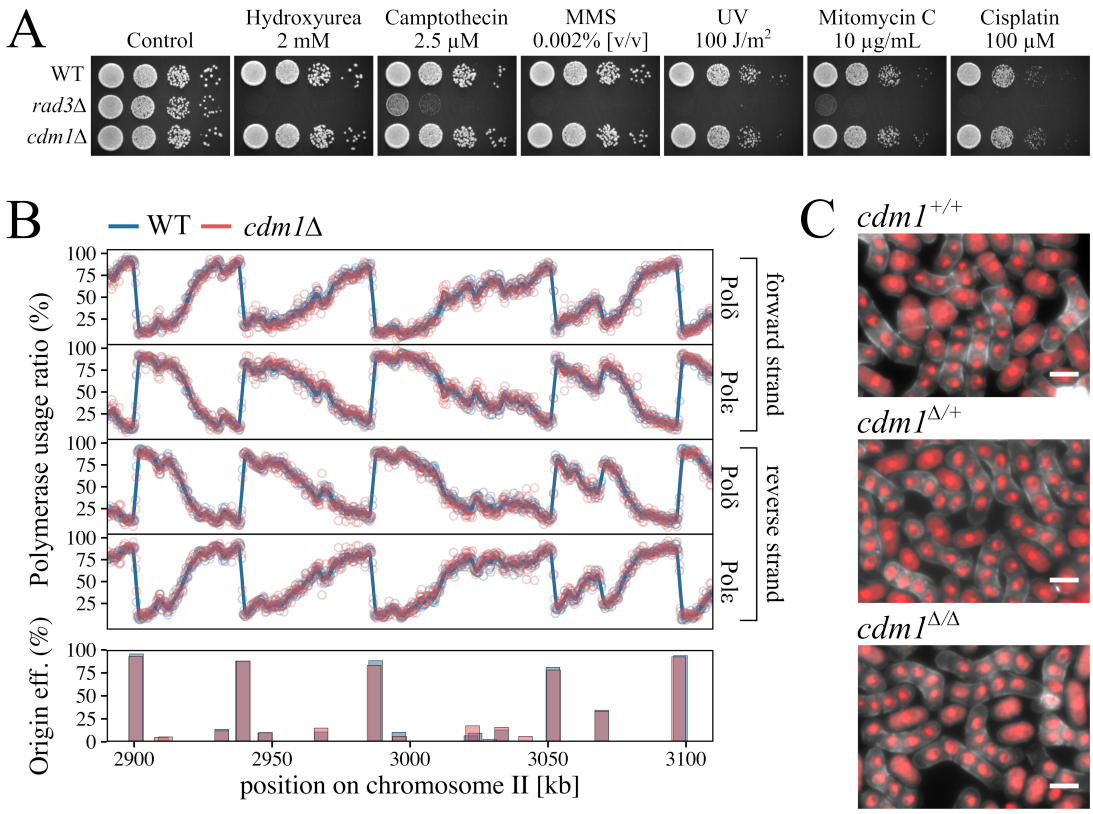

Figure 5.1 – General characterisation of *cdm1Δ* mutant cells. (A) Proliferation of *cdm1Δ* cells under normal and genotoxic conditions. Exponentially growing WT, *rad3Δ* and *cdm1Δ* cells were 10-fold serially diluted and spotted onto YEA plates containing listed genotoxic drugs at indicated doses. Cells subjected to UV-irradiation were first plated on YEA plate and then UV-irradiated. ATR-deficient *rad3Δ* mutant represents a positive control. Experiment was repeated two times. **(B)** Polymerase usage sequencing analysis of *cdm1Δ* mutant. Representative locus on chromosome II (2,890,000-3,110,000) is shown. Upper panel – Polδ and Polε usage on forward and reverse strands calculated as described in the *Methods* section. Circles and lines represent individual data points and simple moving average (window of 7 bins), respectively. Lower panel – estimated origin firing efficiencies calculated as described in the *Methods* section. Experiment was performed once. **(C)** Meiotic outcome of Cdm1-defficient diploids. Spore tetrads originating from diploids carrying two (*cdm1*$^{+/+}$), one

(*cdm1*$^{\Delta/+}$) or none (*cdm1*$^{\Delta/\Delta}$) *cdm1* allele(s) were fixed using 70% ethanol and stained with Calcofluor and SYTOX Green. Representative composite images of Calcofluor (greys) and SYTOX Green (red) signals are shown. Scale bar represents 5 µm.

Cell-cycle control of cdm1^{+} *expression*

Multiple lines of evidence indicate that the expression of small Polδ subunit is under cell-cycle control (Darzynkiewicz *et al*., 2015; Rustici *et al*., 2004; Zhang *et al*., 2013; Hong Zhao *et al*., 2014). First, we aimed to validate that *cdm1*$^{+}$ transcription is upregulated in M/G_1, as has been reported by J. Bähler and colleagues (Rustici *et al*., 2004). We synchronised cells in G_2 using analogue-sensitive CDK kinase Cdc2asM17 (Aoi *et al*., 2014) (Figure 5.2A and 5.2B) and, in a course of two consecutive cell cycles, estimated relative transcript levels of all 4 Polδ subunits (*cdc6*$^{+}$, *cdc1*$^{+}$, *cdc27*$^{+}$ and *cdm1*$^{+}$). As a positive control, we measured RNA levels of ribonucleotide reductase large subunit (RNR, *cdc22*$^{+}$), a well characterised cell-cycle regulated gene (Gordon & Fantes, 1986). Similarly to *cdc22*$^{+}$, transcription of which is upregulated prior to S phase (Gordon & Fantes, 1986), *cdm1*$^{+}$ transcript levels displayed apparent oscillation with a clear peak at 20 min and marginal elevation at 140 min (Figure 5.2C). In agreement with the published data (Rustici *et al*., 2004), elevation of *cdm1*$^{+}$ expression was observed in cells undergoing mitosis (Figure 5.2B and 5.2C). Since, under standard conditions, G_1 is characteristically brief in *S. pombe* (Carlson *et al.*, 1999), temporal resolution of the time-course was not sufficient to determine whether the wave of *cdm1*$^{+}$ transcription was mitosis-specific or overlapped with both, M and G_1 phases. Also in agreement with the literature (Rustici *et al*., 2004), expression of remaining Polδ subunits (*cdc1*$^{+}$, *cdc27*$^{+}$, *cdc6*$^{+}$*)* did not display any apparent changes throughout the time-course (Figure 5.2C).

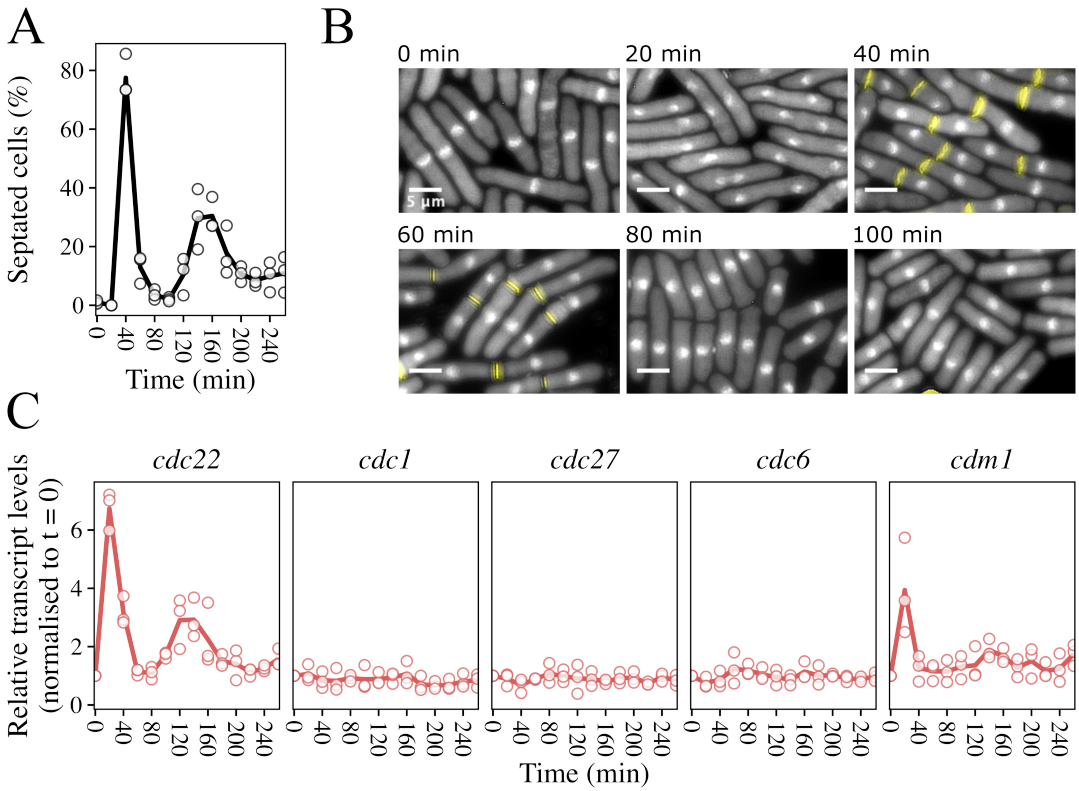

Figure 5.2 – Cell cycle regulation of *cdm1* transcription. (A) Percentage of septated cells. Individual points represent values from 3 independent experiments. Solid line represents the mean. **(B)** Representative images of ethanol-fixed cells stained with Calcofluor (5 µg/mL) and SYTOX Green (5 µM) at indicated time-points. Greys – SYTOX Green; yellow – Calcofluor. Cells were fixed by 70% ethanol and stained with Calcofluor (5 µg/mL) and SYTOX Green (5 µM). **(C)** Relative transcript levels analysed by RT-qPCR. Genes encoding 4 Polδ subunits (*cdc1*$^+$, *cdc27*$^+$, *cdc6*$^+$, *cdm1*$^+$) and positive control (*cdc22*$^+$) were analysed. Transcript levels were normalised to *act1* and 0-timepoint. Individual points represent values from 3 independent experiments. Solid line represents the mean.

Next, we aimed to test whether reported cell-cycle regulated stability of human Cdm1$^+$-orthologue p12 (Darzynkiewicz *et al*., 2015; Zhang *et al*., 2013; Zhao *et al*., 2014) is conserved in *S. pombe*. In order to do so, we employed PCR-based gene targeting (Bähler *et al*., 1998) and C-terminally tagged endogenous *cdm1*$^+$ allele with green fluorescent protein (GFP, *cdm1*$^+$*-GFP*). Cdm1$^+$-GFP construct could be detected by western blot (Figure 5.3). Unfortunately, we could not detect Cdm1$^+$-GFP by microscopy, indicating dysfunctional GFP folding (data not shown).

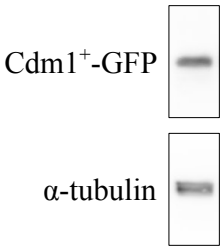

Figure 5.3 – Western blot analysis of the Cdm1⁺-GFP construct. Proteins extracted from exponentially growing cells were resolved on 10% acrylamide gel and detected by western blot as described in the *Methods* section.

We decided to generate the *cdm1⁺-nGFP* mutant, which carried an endogenous *cdm1⁺* allele N-terminally tagged with GFP (Figure 5.4A). According to live cell imaging performed by T. Etheridge, *cdm1⁺-nGFP* cells displayed clear nuclear GFP fluorescence (Figure 5.4B). Strikingly, Cdm1⁺-nGFP signal was rapidly diminished in S-phase cells (Figure 5.4B), indicating that cell-cycle control of Cdm1⁺ protein levels is conserved in *S. pombe*. Protein levels of Cdm1⁺-nGFP re-appeared in G_2 phase (Figure 5.4B), suggesting that Cdm1⁺ operates in G_2 and/or mitosis. Additionally, we wanted to implement a single-molecule tracking microscopy, which would give us clues about Cdm1⁺ mobility and possible interactions with chromatin (Etheridge *et al*., 2014). Thus, we constructed a mutant carrying endogenous *cdm1⁺* allele N-terminally tagged with photoconvertible mEos3.2 tag. According to a pilot experiment performed by T. Etheridge, Cdm1⁺ displayed very high mobility in unperturbed cells (data not shown). More elaborate analysis, however, has not been carried out.

Next, we aimed to test whether Cdm1⁺-nGFP is targeted for degradation in response to DNA-damage. In order to do so, T. Etheridge performed live cell imaging of *cdm1⁺-nGFP* cells exposed to 0.08% MMS. Similarly to human Cdm1⁺-orthologue p12 (Darzynkiewicz *et al*., 2015; Terai *et al*., 2013; Zhang *et al*., 2013, 2007; Zhao *et al*., 2014), Cdm1⁺-nGFP levels were diminished in MMS-treated cells (Figure 5.4C). Interestingly, unlike the rapid depletion of Cdm1⁺-nGFP observed in S phase, DNA-damage induced Cdm1⁺-nGFP degradation occurred in a slow and continuous manner (Figure 5.4C).

To test whether the role of $CRL4^{CDT2}$ and ATR in degradation of small Polδ subunit is conserved in *S. pombe*, we constructed *cdt2Δ*, *rad3Δ* and *tel1Δ* strains carrying *cdm1-nGFP* allele. While *rad3Δ* and *tel1Δ* mutants originated from the laboratory strain collection, *cdt2Δ* allele was *de-novo* constructed by PCR-based gene

targeting (Bähler *et al*., 1998). Since *cdt2Δ* cells are characterised by severe defects in cellular and nuclear morphology, we co-deleted *spd1* (*spd1Δ* was provided by Cong Liu), disruption of which is known to suppress some of the *cdt2Δ* phenotypes (Liu *et al*., 2005). In agreement with the studies addressing regulation of p12 in human cells, deletion of *cdt2*$^{+}$, but not *rad3*$^{+}$ and *tel1*$^{+}$, resulted in stabilisation of Cdm1^{+}-nGFP in S phase (Figure 5.4D). In MMS-treated cells, degradation of Cdm1-nGFP was abrogated in *cdt2Δ* and *rad3Δ*, but not *tel1Δ* mutants (Figure 5.4D).

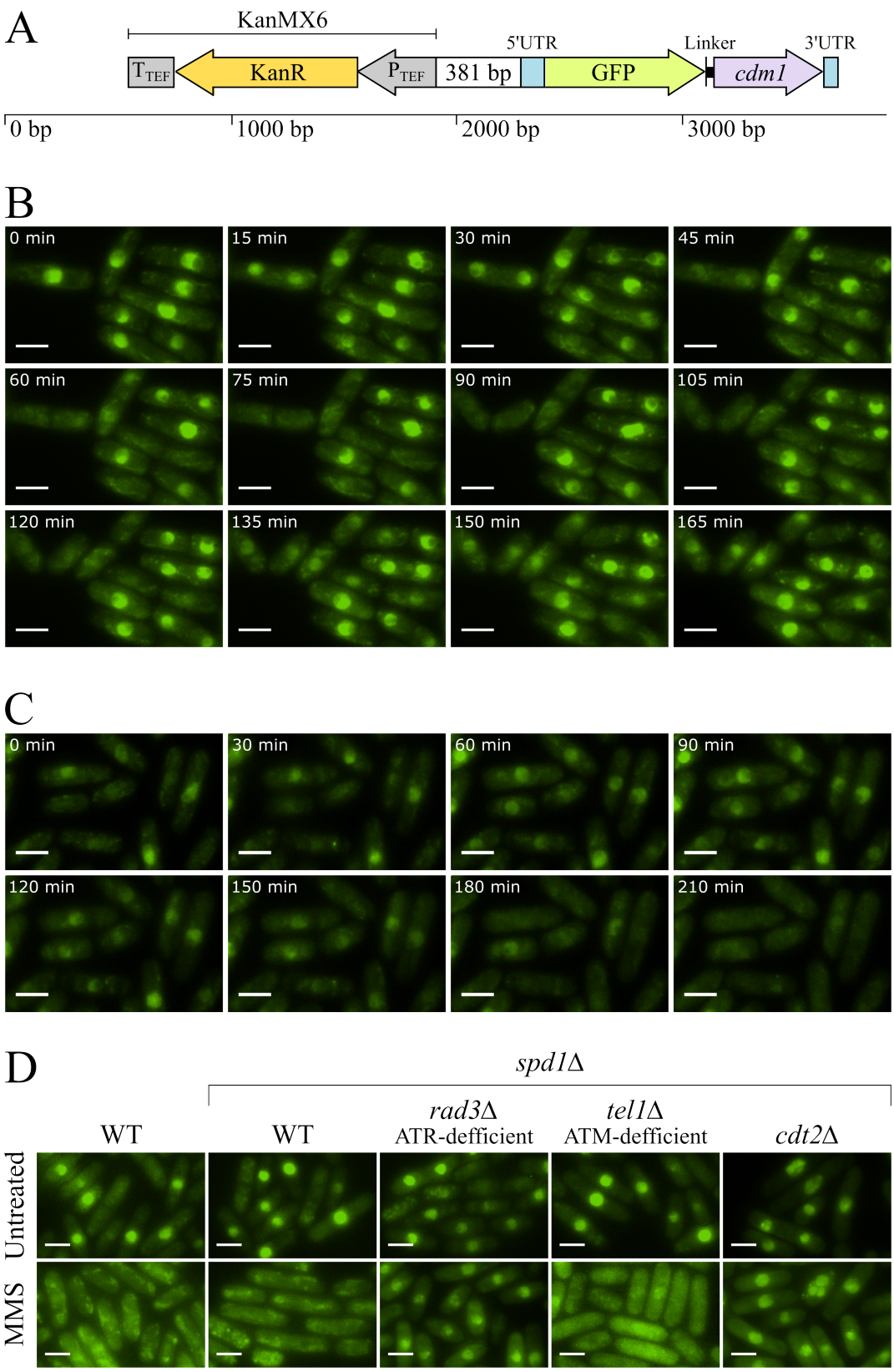

Figure 5.4 – Construction and initial characterisation of *cdm1⁺-nGFP* mutant cells. (A) Schematic illustration of endogenous *cdm1⁺-nGFP* locus. KanMX6 – kanamycin resistance cassette MX6; T_{TEF} – TEF terminator; KanR – kanamycin resistance (aminoglycoside 3'-phosphotransferase); P_{TEF} – TEF promoter; 381 bp – region upstream of *cdm1⁺*; 5'UTR – 5'

untranslated region; GFP – green fluorescent protein, 3'UTR – 3' untranslated region. **(B, C)** Unperturbed (B) and MMS (0.08) treated (C) *cdm1⁺-nGFP* cells. S phase – binucleated cells with daughter nuclei positioned on the opposite poles of the cell. Representative images at indicated time points are shown. **(D)** Cdm1⁺-nGFP signal in *rad3Δ*, *tel1Δ* and *cdt2Δ* cells before and 4 h after MMS (0.08%) treatment. Representative images are shown. Scale bar represents 5 µm.

To identify the counterpart of human p12 PIP-degron motif, which is recognised by CRL4^{CDT2}, we aligned p12 and Cdm1⁺ protein sequences (Figure 5.5A). Based on the p12/Cdm1⁺ alignment, we identified Cdm1⁺ Lysine-26 to be the most likely candidate for Cdm1⁺ PIP-degron. To test whether substitution of Lysine-26 for Alanine abrogates Cdm1⁺ degradation, we constructed *cdm1^{R26A}-nGFP* mutant. Remarkably, Cdm1^{R26A}-nGFP was no longer degraded during S phase (Figure 5.5B). Collectively, the aforementioned results indicate that differential regulation of p12 and Cdm1⁺ protein levels is evolutionary conserved.

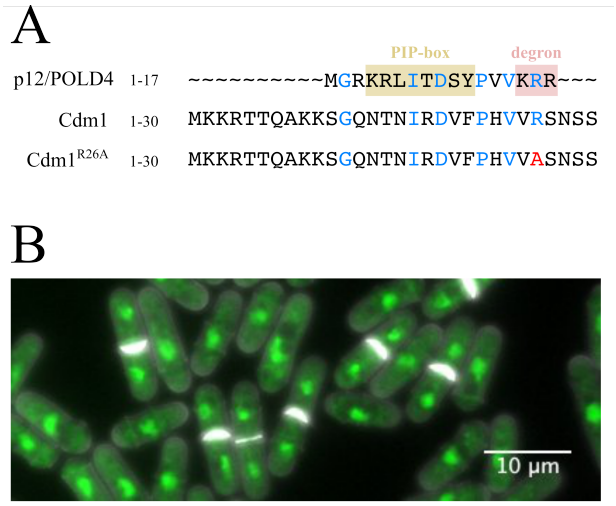

Figure 5.5 – Identification of Cdm1 PIP-degron motif. (A) Alignment of N-terminal parts of p12, Cdm1⁺ and Cdm1^{R26A} constructed by Clustal Omega multiple sequence alignment, EMBL-EBI (Madeira *et al*., 2019). p12 PIP-box and degron are indicated. Blue letters signify conserved residues. R26A mutation is indicated by red "A". **(B)** Representative image of *cdm1^{R26A}-nGFP* cells. Composite image of GFP (green) and Calcofluor (greys) is shown.

To ensure credibility of our findings, we wanted to demonstrate that Cdm1⁺-nGFP and Cdm1⁺-nmEos3.2 proteins retained their biological function. It had been previously reported that over-expression of Cdm1⁺ rescues temperature-sensitive

mutants *cdc27-p11* and *cdc24-m38*. We reasoned that temperature-dependent phenotypes of *cdc27-p11* and *cdc24-m38* should be also rescued by over-production of Cdm1$^+$-nGFP and Cdm1$^+$-nmEos3.2, provided that GFP and mEos3.2 tags do not impair biological function of Cdm1$^+$. In both mutants, *cdc27-p11* and *cdc24-m38*, we ectopically expressed Cdm1$^+$, Cdm1$^+$-nGFP and Cdm1$^+$-nmEos3.2. When grown at restrictive temperature (36°C) *cdc27-p11* mutant was rescued by over-expression of Cdm1$^+$ and partially rescued by over-expression of Cdm1$^+$-nmEos3.2 (Figure 5.6A). Growth of *cdc27-p11* cells incubated at 36°C, however, was not supported by over-production of Cdm1$^+$-nGFP (Figure 5.6A). Over-expression of Cdm1$^+$, Cdm1$^+$-nGFP and Cdm1$^+$-nmEos3.2 did not rescue temperature-dependent phenotype of *cdc24-m38* (Figure 5.6B). Based on such results, we concluded that Cdm1$^+$-nGFP and Cdm1$^+$-nmEos3.2 were functionally compromised and therefore not suitable for further experiments.

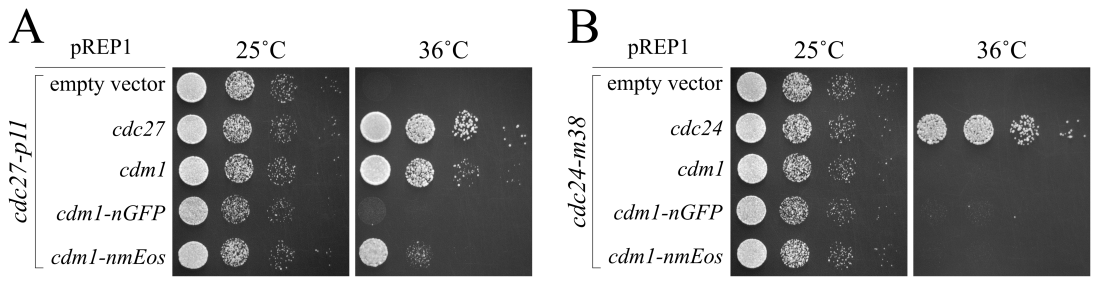

Figure 5.6 – Functionality of Cdm1$^+$-nGFP and Cdm1$^+$-nmEos3.2. (A, B) *cdm1$^+$*, *cdm1$^+$-nGFP*, *cdm1$^+$-nmEos3.2* were over-produced from pREP1 expression vector in leucine-autotrophic *cdc27-p11* (A) and *cdc24-m38* (B) mutants. Cells were 10-fold serially diluted, spotted onto leucine-deficient EMM plates and incubated at 25°C or 36°C.

Discussion 5

A considerable body of evidence indicates that p12, the small subunit of Polδ complex, is cell-cycle regulated and modulates enzymatic properties of Polδ in human cells. According to the current consensus, p12-lacking Polδ_3 facilitates canonical replication, whereas p12-embodying Polδ_4 is important for HR during G2 and/or mitosis. It has been demonstrated that the p12 function contributes to the overall genome stability in highly aggressive cancer cell lines and FGF2-treated MAE cells,

however, underlying DNA-repair contexts requiring p12 have remained unclear. To provide new insights into the biology of small Polδ subunit, we aimed to utilise powerful genetics of *S. pombe* and characterise the function of Cdm1$^+$, the fission yeast orthologue of p12.

In the first part of the project, we aimed to reproduce some of the published experiments addressing Cdm1$^+$ function and provide a general characterisation of *cdm1Δ* mutant. In agreement with the original study characterising Cdm1$^+$, we demonstrated that Cdm1$^+$ is not required for viability under normal conditions and Cdm1$^+$-deficient cells are not sensitive to HU, CPT, MMS and UV. In addition to the aforementioned stressors, we also tested sensitivity of *cdm1Δ* cells to Mitomycin C and Cisplatin, both of which reduce colony forming capacity in human A549 cells depleted of p12. Unlike in p12-defficient A549 cells, Mitomycin C and Cisplatin do not impair growth of *cdm1Δ* cells. Such finding could indicate that requirement of the small Polδ subunit differ between human and *S. pombe* under certain conditions. It is also possible, however, that sensitivity to Mitomycin C and Cisplatin observed in p12-defficient A549 cells is conditioned by inherent chromosomal instability characteristic for A549 cell line or represents a technical consequence of experimental design. In accordance with the predicted role of Cdm1$^+$ outside of S phase, we demonstrate that canonical replication is unaffected in *cdm1Δ* cells. Even though it has been shown that *cdm1$^+$* is upregulated during the late stages of meiotic cycle, we did not detect any abnormalities in morphology of spore-tetrads originating from Cdm1$^+$-deficient diploids. However, since we have not tested whether the frequency of meiotic recombination events is altered in *cdm1Δ* cells, the involvement of Cdm1$^+$ in meiosis cannot be ruled out. Additionally, we provide an evidence that transcription of *cdm1$^+$* is upregulated during mitosis, as had been previously reported by J. Bähler and colleagues (Rustici *et al.*, 2004). Overall, such findings indicate that Cdm1$^+$ is dispensable for fundamental cellular processes, such as canonical replication, and likely operates in a specific, yet unknown, contexts.

In the second part of the project, we aspired to provide an evidence that function of small Polδ subunit is conserved between human and *S. pombe*. We constructed a set of tag mutants expressing Cdm1$^+$ C-terminally tagged with GFP (*cdm1$^+$-GFP*) and Cdm1$^+$ N-terminally tagged with GFP (*Cdm1$^+$-nGFP*) or mEos3.2

(*Cdm1⁺-nmEos3.2*). Based on immunodetection and/or microscopy data, only Cdm1⁺-nGFP and Cdm1⁺-nmEos3.2 were suitable for experimental work. Our initial experiments indicate that, similarly to p12, Cdm1⁺-nGFP is targeted for degradation in cells undergoing replication and cells treated with MMS. In further analogy to p12 in human cells, degradation of Cdm1⁺-nGFP depends on Cdt2⁺, the component of CRL^{CDT2} E3 ubiquitin ligase and the PIP-box degron motif situated on the N-terminus of Cdm1⁺. In addition to CRL^{CDT2}, MMS-induced depletion of Cdm1⁺-nGFP requires functional ATR, which is also required for p12 degradation in UV-irradiated human cells. Collectively, such findings imply that regulatory nodes governing protein levels of p12 and Cdm1⁺ are conserved.

Even though Cdm1⁺-nGFP construct retained nuclear localisation and degron function, we found that Cdm1⁺ activity was highly compromised in *cdm1⁺-nGFP* strains. We suspect that, in Cdm1⁺-nGFP construct, the large GFP domain blocks the N-terminal PIP-box of Cdm1⁺ and hinders the interaction with PCNA, which can be crucial for Cdm1⁺ function. Interestingly, activity of Cdm1⁺-nmEos3.2 construct was only partially impaired. Thus, we reason that fully functional N-terminally tagged Cdm1⁺ variant can be constructed by implementing different tags and/or linker peptides.

In the third part of the project, we planned to decipher the concrete biological role of Cdm1⁺. Unfortunately, due to technical complications and the lack of time, we have not been able to comply with the plan and mechanistic insights into Cdm1⁺ function have not been thoroughly addressed. Nevertheless, we believe that the work presented in this chapter represents a solid basis for an interesting study. To outline the future work, we propose following hypotheses and experiments.

First, $Pol\delta_4$ has been characterised by higher rate of replication and competence in strand-displacement synthesis, which is essential for HR. We theorise that $Pol\delta_4$ is required for completion of under-replicated regions, replication of which takes place in G_2 and must be completed prior to mitosis. It would be interesting to utilise RTS1 replication fork barrier (RFB) system (Lambert *et al.*, 2005) and test whether Cdm1⁺ plays a role in the restart of blocked replication forks. Similarly, it could be insightful to test whether deletion of *cdm1⁺* impairs genome stability in senataxin

mutants *sen1Δ* and *dbl8Δ*, which, according to our unpublished data, display increased occurrence of replication fork stalling events at highly transcribed loci.

Second, it is assumed that p12-lacking Polδ_3 complex that operates during S phase exhibits high fidelity and enzymatic properties ideal for processing and maturation of OF in human cells. Interestingly, no phenotypes have been reported in human cells ectopically expressing non-degradable p12 variant. Accordingly, we did not observe any defects in cells expressing non-degradable Cdm1^{R26A}-nGFP. Such findings, however, are most likely irrelevant, as N-terminal GFP abrogates Cdm1^{+} function. We believe, it would be informative to construct and characterise an untagged *cdm1^{R26A}* mutant. Of particular interest, Pu-Seq analysis of *cdm1^{R26A}* cells could provide robust insights into how Cdm1^{+} levels modulate activity of Polδ across the genome in living cells.

Conclusions

DNA replication represents one of the most important biological processes with relevance to physiological mechanisms of reproduction and development, as well as pathophysiological conditions such as cancer and neurodegeneration. Thus, it is not surprising that biomedical research addressing mechanistic and regulatory aspects of replication progression is regarded as one of the most influential fields of molecular biology. Throughout the past decades, principles and mechanisms underlying DNA replication have been characterised to a great detail. However, in accordance with Albert Einstein's remark, "The more I learn, the more I realize how much I don't know," accumulated knowledge have formulated new questions and initiated discussions of well-established theorems. The aim of this thesis was to introduce new methodological and theoretical concepts, which would improve our repertoire of analytical tools and further refine our understanding of DNA replication and its regulation.

Scientific progress is accompanied by establishment of novel hypotheses and replacement of outdated perspectives. The original polymerase usage sequencing (Pu-Seq) protocol was established in January 2015 and, at that time, created a new platform for studies addressing replication dynamics. A few years later, personal experience as well as evidence provided by other research groups suggested changes in the original Pu-Seq protocol, which should be implemented to improve accuracy, processivity and cost-efficiency. In chapter 1, we introduced an alternative Pu-Seq procedure, which is built on principles underlying GLOE-Seq and RHII-HydEn-seq protocols. According to preliminary data, the updated Pu-Seq methodology is highly reproducible, requires less time for completion, provides better cost-efficiency and allows simultaneous processing of up to 12 samples. Notably, we encountered a technical issue of a yet unknown origin, which compromised the quality of a subset of prepared Pu-Seq libraries. We believe that occasional sub-optimal performance is linked to suboptimal input genomic DNA prepared by a quicker DNA extraction method; however, additional investigation is necessary. We believe that, when fully optimised, the newly developed Pu-Seq protocol has a potential to become a new standard among methodologies detecting misincorporated ribonucleotides.

Scientific method, which allows us to acquire empirical knowledge, is partly based on scepticism, questioning and experimental validation. In chapter 2, we presented results of our recent publication (Zach & Carr, 2021), which aimed to test a previously postulated theory that, in a concentration-dependent manner, Polδ interferes with canonical leading strand replication. We constructed an elaborate set of *S. pombe* strains characterised by either 2- or 4-fold increased expression of all Polδ subunits and utilised Pu-Seq to determine whether such an increase in Polδ production alter normal replication dynamics. In contrast to the original *in vitro* observation, our experiments indicate that up to 4-fold increased Polδ expression has no detectable impact on replication dynamics. Even though we cannot rule out that Polδ interferes with the leading strand synthesis at very low frequencies, we argue that, at least in our experimental system, Polδ mediated perturbation of the leading strand replication likely does not represent a physiologically relevant phenomenon.

It is becoming increasingly clear that distinct cellular processes can interfere with each other in many ways. For instance, molecular machines facilitating replication and transcription compete for the same DNA template. It is widely accepted that conflicts between replisomes and RNA polymerases can lead to DNA damage and promote genome instability. In chapter 3 we investigated functions of Sen1$^+$ and Dbl8$^+$, senataxin RNA/DNA helicases in *S. pombe*, in prevention and/or resolution of replication/transcription collisions (RTCs). We provide a conclusive evidence that Sen1$^+$ and Dbl8$^+$ participate in suppression of replication fork stalling at a subset of highly expressed protein-coding loci. According to our analysis, Sen1$^+$ and Dbl8$^+$ play a role in codirectional and head-on RTCs, respectively. Interestingly, we also determined that simultaneous disruption of both senataxin helicases results in significant increase in replication fork stalling events across the genome, potentially indicating a yet uncharacterised functional redundancy of Sen1$^+$ and Dbl8$^+$. Notably, we postulate that RTCs do not represent a simple consequence of ongoing transcription, and that many highly transcribed genes do not show any signs of impaired replication fork mobility in senataxin-deficient strains. We explored several mechanisms potentially responsible for observed locus-specific occurrence of replication fork stalling, including accumulation of R-loops, gene length, high GC content, GC-skew and gene ontology classification. We did not find any conclusive correlations, however, and thus it is still

unclear why only a subset of highly expressed genes exhibits defective replication fork progression in senataxin-deficient strains. We are convinced that further development of the project will lead to elucidation of novel genome stability maintenance mechanisms related to RTCs.

Even though not strictly common, certain widely recognised biological models are based on educated assumptions, rather than robust empirical evidence. A good example of a primarily assumption-based model is the notion that condensed heterochromatin structure represents an endogenous obstacle impairing progression of replication forks. In chapter 4, we aspire to generate a molecular system, which allows comprehensive analysis of replication of heterochromatinised domains. We successfully constructed strains which carry genetic background required for Pu-Seq analysis and, at the same time, allow induction of synthetic heterochromatin at three distinct genomic positions. Our preliminary data, targeting a single ectopically heterochromatinised locus, indicate that, at least in *S. pombe*, condensed chromatin induces firing of dormant origins, as has been previously indicated. Even though we have not managed to conclusively test whether heterochromatin causes retardation of replication fork progression, we believe that further development and utilisation of presented system could provide additional insights into the replication of compact heterochromatin domains.

While biomedical research tends to focus on a relatively small subset of factors, many genes are understudied and remain without assigned molecular function. According to *in vitro* experiments, the small Polδ subunit p12/Cdm1^{+} functions as an important modulator of Polδ activity. Interestingly; however, only few studies attempted to address p12/Cdm1^{+} function in cellular systems. In chapter 5, we aim to establish *S. pombe* as a versatile experimental model to decipher the function of Cdm1^{+}. We confirm previous reports that Cdm1^{+}-deficient cells do not develop any apparent phenotypes, and that *cdm1^{+}* transcript levels oscillate in a cell cycle dependent manner. Additionally, we provide an evidence that Cdm1^{+} is targeted for degradation by S phase-specific Cdt2^{+}-dependent E3-ubiquitin ligase, which recognises the Cdm1^{+} PIP-degron motif, arginine-26. The presented data indicate that, similarly to p12 (human Cdm1^{+} orthologue), Cdm1^{+} likely functions as a cell cycle dependent modulator of Polδ activity. We believe that further development of this

project will generate new insights into regulation of Polδ activity with respect to cell cycle progression.

In summary, this thesis addressed diverse set of replication-related topics and provided new insights into biology of replicative polymerases, senataxin RNA/DNA helicases and their involvement in suppression of RTCs, replication of condensed heterochromatin domains and functional modulation of the main lagging strand replicase Polδ. This thesis also introduced an updated Pu-Seq procedure with outstanding performance. We believe that presented data represent a valuable source of novel information, which significantly contributes to our understanding of unperturbed and non-canonical eukaryotic replication. Additionally, presented work introduced several original techniques, which will be useful in addressing experimentally challenging biological phenomena, such as replication of highly compact heterochromatic domains and functional interference between replisomes and transcription machineries.

References

Abid Ali, F., Douglas, M. E., Locke, J., Pye, V. E., Nans, A., Diffley, J. F. X., & Costa, A. (2017). Cryo-EM structure of a licensed DNA replication origin. *Nature Communications*, *8*(1), 2241. https://doi.org/10.1038/s41467-017-02389-0

Achuthankutty, D., Thakur, R. S., Haahr, P., Hoffmann, S., Drainas, A. P., Bizard, A. H., … Mailand, N. (2019). Regulation of ETAA1-mediated ATR activation couples DNA replication fidelity and genome stability. *Journal of Cell Biology*, *218*(12), 3943–3953. https://doi.org/10.1083/jcb.201905064

Ade, C., Roy-Engel, A. M., & Deininger, P. L. (2013). Alu elements: an intrinsic source of human genome instability. *Current Opinion in Virology*, *3*(6), 639–645. https://doi.org/10.1016/j.coviro.2013.09.002

Airoldi, G., Guidarelli, A., Cantoni, O., Panzeri, C., Vantaggiato, C., Bonato, S., … Bassi, M. T. (2010). Characterization of two novel SETX mutations in AOA2 patients reveals aspects of the pathophysiological role of senataxin. *Neurogenetics*, *11*(1), 91–100. https://doi.org/10.1007/s10048-009-0206-0

Al Mamun, M., Albergante, L., Moreno, A., Carrington, J. T., Blow, J. J., & Newman, T. J. (2016). Inevitability and containment of replication errors for eukaryotic genome lengths spanning megabase to gigabase. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(39), E5765-74. https://doi.org/10.1073/pnas.1603241113

Allshire, R. C., & Ekwall, K. (2015). Epigenetic Regulation of Chromatin States in *Schizosaccharomyces pombe*. *Cold Spring Harbor Perspectives in Biology*, *7*(7), a018770. https://doi.org/10.1101/cshperspect.a018770

Allshire, R. C., & Madhani, H. D. (2018). Ten principles of heterochromatin formation and function. *Nature Reviews. Molecular Cell Biology*, *19*(4), 229–244. https://doi.org/10.1038/nrm.2017.119

Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., … Liberi, G. (2012). Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell*, *151*(4), 835–846. https://doi.org/10.1016/j.cell.2012.09.041

Anheim, M., Monga, B., Fleury, M., Charles, P., Barbot, C., Salih, M., … Koenig, M. (2009). Ataxia with oculomotor apraxia type 2: clinical, biological and genotype/phenotype correlation study of a cohort of 90 patients. *Brain : A Journal of Neurology*, *132*(Pt 10), 2688–2698. https://doi.org/10.1093/brain/awp211

Aoi, Y., Kawashima, S. A., Simanis, V., Yamamoto, M., & Sato, M. (2014). Optimization of the analogue-sensitive Cdc2/Cdk1 mutant by in vivo selection eliminates physiological limitations to its use in cell cycle analysis. *Open Biology*, *4*(7). https://doi.org/10.1098/rsob.140063

Appanah, R., Lones, E. C., Aiello, U., Libri, D., & De Piccoli, G. (2020). Sen1 Is Recruited to Replication Forks via Ctf4 and Mrc1 and Promotes Genome Stability. *Cell Reports*, *30*(7), 2094-2105.e9. https://doi.org/10.1016/j.celrep.2020.01.087

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, *408*(6814), 796–815. https://doi.org/10.1038/35048692

Araki, H. (2010). Cyclin-dependent kinase-dependent initiation of chromosomal DNA replication. *Current Opinion in Cell Biology*, *22*(6), 766–771. https://doi.org/10.1016/j.ceb.2010.07.015

Aria, V., & Yeeles, J. T. P. (2018). Mechanism of Bidirectional Leading-Strand Synthesis Establishment at Eukaryotic DNA Replication Origins. *Molecular Cell*, *73*(2), 199-211.e10. https://doi.org/10.1016/j.molcel.2018.10.019

Arias, E. E., & Walter, J. C. (2005). Replication-dependent destruction of Cdt1 limits DNA replication to a single round per cell cycle in Xenopus egg extracts. *Genes & Development*, *19*(1), 114–126. https://doi.org/10.1101/gad.1255805

Arias, E. E., & Walter, J. C. (2006). PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. *Nature Cell Biology*, *8*(1), 84–90. https://doi.org/10.1038/ncb1346

Arigo, J. T., Eyler, D. E., Carroll, K. L., & Corden, J. L. (2006). Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Molecular Cell*, *23*(6), 841–851. https://doi.org/10.1016/j.molcel.2006.07.024

Arndt, K. M., & Reines, D. (2015). Termination of Transcription of Short Noncoding RNAs by RNA Polymerase II. *Annual Review of Biochemistry*, *84*, 381–404. https://doi.org/10.1146/annurev-biochem-060614-034457

Audergon, P. N. C. B., Catania, S., Kagansky, A., Tong, P., Shukla, M., Pidoux, A. L., & Allshire, R. C. (2015). Epigenetics. Restricted epigenetic inheritance of H3K9 methylation. *Science (New York, N.Y.)*, *348*(6230), 132–135. https://doi.org/10.1126/science.1260638

Audia, J. E., & Campbell, R. M. (2016). Histone Modifications and Cancer. *Cold Spring Harbor Perspectives in Biology*, *8*(4), a019521. https://doi.org/10.1101/cshperspect.a019521

Aze, A., Sannino, V., Soffientini, P., Bachi, A., & Costanzo, V. (2016). Centromeric DNA replication reconstitution reveals DNA loops and ATR checkpoint suppression. *Nature Cell Biology*, *18*(6), 684–691. https://doi.org/10.1038/ncb3344

Bacolla, A., & Wells, R. D. (2004). Non-B DNA conformations, genomic rearrangements, and human disease. *The Journal of Biological Chemistry*, *279*(46), 47411–47414. https://doi.org/10.1074/jbc.R400028200

Bähler, J., Wu, J. Q., Longtine, M. S., Shah, N. G., McKenzie, A., Steever, A. B., … Pringle, J. R. (1998). Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast (Chichester, England)*, *14*(10), 943–951. https://doi.org/10.1002/(SICI)1097-0061(199807)14:10<943::AID-YEA292>3.0.CO;2-Y

Balakrishnan, L., & Bambara, R. A. (2013). Flap endonuclease 1. *Annual Review of Biochemistry*, *82*(1), 119–138. https://doi.org/10.1146/annurev-biochem-072511-122603

Ball, H. L., Myers, J. S., & Cortez, D. (2005). ATRIP binding to replication protein A-single-stranded DNA promotes ATR-ATRIP localization but is dispensable for Chk1 phosphorylation. *Molecular Biology of the Cell*, *16*(5), 2372–2381. https://doi.org/10.1091/mbc.e04-11-1006

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

Baretić, D., Jenkyn-Bedford, M., Aria, V., Cannone, G., Skehel, M., & Yeeles, J. T. P. (2020). Cryo-EM Structure of the Fork Protection Complex Bound to CMG at a Replication Fork. *Molecular Cell*, *78*(5), 926-940.e13. https://doi.org/10.1016/j.molcel.2020.04.012

Baxter, J., & Diffley, J. F. X. (2008). Topoisomerase II inactivation prevents the completion of DNA replication in budding yeast. *Molecular Cell*, *30*(6), 790–802. https://doi.org/10.1016/j.molcel.2008.04.019

Bębenek, A., & Ziuzia-Graczyk, I. (2018). Fidelity of DNA replication-a matter of proofreading. *Current Genetics*, *64*(5), 985–996. https://doi.org/10.1007/s00294-018-0820-1

Becherel, O. J., Yeo, A. J., Stellati, A., Heng, E. Y. H., Luff, J., Suraweera, A. M., … Lavin, M. F. (2013). Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genetics*, *9*(4), e1003435. https://doi.org/10.1371/journal.pgen.1003435

Beeharry, N., Rattner, J. B., Caviston, J. P., & Yen, T. (2013). Centromere fragmentation is a common mitotic defect of S and G2 checkpoint override. *Cell Cycle (Georgetown, Tex.)*, *12*(10), 1588–1597. https://doi.org/10.4161/cc.24740

Bermejo, R., Lai, M. S., & Foiani, M. (2012). Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. *Molecular Cell*, *45*(6), 710–718. https://doi.org/10.1016/j.molcel.2012.03.001

Bermudez, V. P., Lindsey-Boltz, L. A., Cesare, A. J., Maniwa, Y., Griffith, J. D., Hurwitz, J., & Sancar, A. (2003). Loading of the human 9-1-1 checkpoint complex onto DNA by the checkpoint clamp loader hRad17-replication factor C complex in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 1633–1638. https://doi.org/10.1073/pnas.0437927100

Besnard, E., Babled, A., Lapasset, L., Milhavet, O., Parrinello, H., Dantec, C., … Lemaitre, J.-M. (2012). Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. *Nature Structural & Molecular Biology*, *19*(8), 837–844. https://doi.org/10.1038/nsmb.2339

Bi, X. (2014). Heterochromatin structure: lessons from the budding yeast. *IUBMB Life*, *66*(10), 657–666. https://doi.org/10.1002/iub.1322

Blackford, A. N., & Jackson, S. P. (2017). ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Molecular Cell*, *66*(6), 801–817. https://doi.org/10.1016/j.molcel.2017.05.015

Bleichert, F., Botchan, M. R., & Berger, J. M. (2017). Mechanisms for initiating cellular DNA replication. *Science (New York, N.Y.)*, *355*(6327), 25–54. https://doi.org/10.1126/science.aah6317

Blow, J. J., & Ge, X. Q. (2009). A model for DNA replication showing how dormant origins safeguard against replication fork failure. *EMBO Reports*, *10*(4), 406–412. https://doi.org/10.1038/embor.2009.5

Bontekoe, C. J. M., Bakker, C. E., Nieuwenhuizen, I. M., van der Linde, H., Lans, H., de Lange, D., … Oostra, B. A. (2001). Instability of a (CGG)98 repeat in the Fmr1 promoter. *Human Molecular Genetics*, *10*(16), 1693–1699. https://doi.org/10.1093/hmg/10.16.1693

Brambati, A., Zardoni, L., Achar, Y. J., Piccini, D., Galanti, L., Colosio, A., … Liberi, G. (2018). Dormant origins and fork protection mechanisms rescue sister forks arrested by transcription. *Nucleic Acids Research*, *46*(3), 1227–1239. https://doi.org/10.1093/nar/gkx945

Braun, K. A., Lao, Y., He, Z., Ingles, C. J., & Wold, M. S. (1997). Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochemistry*, *36*(28), 8443–8454. https://doi.org/10.1021/bi970473r

Brüning, J.-G., & Marians, K. J. (2020). Replisome bypass of transcription complexes and R-loops. *Nucleic Acids Research*, *48*(18), 10353–10367. https://doi.org/10.1093/nar/gkaa741

Brustel, J., Kozik, Z., Gromak, N., Savic, V., & Sweet, S. M. M. (2018). Large XPF-dependent deletions following misrepair of a DNA double strand break are prevented by the RNA:DNA helicase Senataxin. *Scientific Reports*, *8*(1), 3850. https://doi.org/10.1038/s41598-018-21806-y

Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., … Helleday, T. (2005). Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature*, *434*(7035), 913–917. https://doi.org/10.1038/nature03443

Burgers, P. M. J., & Kunkel, T. A. (2017). Eukaryotic DNA Replication Fork. *Annual Review of Biochemistry*, *86*(1), 417–438. https://doi.org/10.1146/annurev-biochem-061516-044709

Burwinkel, B., & Kilimann, M. W. (1998). Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *Journal of Molecular Biology*, *277*(3), 513–517. https://doi.org/10.1006/jmbi.1998.1641

Byun, T. S., Pacek, M., Yee, M., Walter, J. C., & Cimprich, K. A. (2005). Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes & Development*, *19*(9), 1040–1052. https://doi.org/10.1101/gad.1301205

Canzio, D., Chang, E. Y., Shankar, S., Kuchenbecker, K. M., Simon, M. D., Madhani, H. D., … Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Molecular Cell*, *41*(1), 67–81. https://doi.org/10.1016/j.molcel.2010.12.016

Carlson, C. R., Grallert, B., Stokke, T., & Boye, E. (1999). Regulation of the start of DNA replication in *Schizosaccharomyces pombe*. *Journal of Cell Science*, *112 (Pt 6*(6), 939–946. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10036243

Carr, A. M., & Lambert, S. (2013). Replication stress-induced genome instability: the dark side of replication maintenance by homologous recombination. *Journal of Molecular Biology*, *425*(23), 4733–4744. https://doi.org/10.1016/j.jmb.2013.04.023

Carroll, K. L., Pradhan, D. A., Granek, J. A., Clarke, N. D., & Corden, J. L. (2004). Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. *Molecular and Cellular Biology*, *24*(14), 6241–6252. https://doi.org/10.1128/MCB.24.14.6241-6252.2004

Carson, C. T., Orazio, N. I., Lee, D. V., Suh, J., Bekker-Jensen, S., Araujo, F. D., … Weitzman, M. D. (2009). Mislocalization of the MRN complex prevents ATR signaling during adenovirus infection. *The EMBO Journal*, *28*(6), 652–662. https://doi.org/10.1038/emboj.2009.15

Cayrou, C., Coulombe, P., Vigneron, A., Stanojcic, S., Ganier, O., Peiffer, I., … Méchali, M. (2011). Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. *Genome Research*, *21*(9), 1438–1449. https://doi.org/10.1101/gr.121830.111

Chen, Y.-H., Keegan, S., Kahli, M., Tonzi, P., Fenyö, D., Huang, T. T., & Smith, D. J. (2019). Transcription shapes DNA replication initiation and termination in human cells. *Nature Structural & Molecular Biology*, *26*(1), 67–77. https://doi.org/10.1038/s41594-018-0171-0

Chen, Y.-Z., Bennett, C. L., Huynh, H. M., Blair, I. P., Puls, I., Irobi, J., … Chance, P. F. (2004). DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *American Journal of Human Genetics*, *74*(6), 1128–1135. https://doi.org/10.1086/421054

Chilkova, O., Stenlund, P., Isoz, I., Stith, C. M., Grabowski, P., Lundström, E.-B., … Johansson, E. (2007). The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. *Nucleic Acids Research*, *35*(19), 6588–6597. https://doi.org/10.1093/nar/gkm741

Chinchilla, K., Rodriguez-Molina, J. B., Ursic, D., Finkel, J. S., Ansari, A. Z., & Culbertson, M. R. (2012). Interactions of Sen1, Nrd1, and Nab3 with multiple phosphorylated forms of the Rpb1 C-terminal domain in *Saccharomyces cerevisiae*. *Eukaryotic Cell*, *11*(4), 417–429. https://doi.org/10.1128/EC.05320-11

Clausen, A. R., Lujan, S. A., Burkholder, A. B., Orebaugh, C. D., Williams, J. S., Clausen, M. F., … Kunkel, T. A. (2015). Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. *Nature Structural & Molecular Biology*, *22*(3), 185–191. https://doi.org/10.1038/nsmb.2957

Cohen, S., Puget, N., Lin, Y.-L., Clouaire, T., Aguirrebengoa, M., Rocher, V., … Legube, G. (2018). Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations. *Nature Communications*, *9*(1), 533. https://doi.org/10.1038/s41467-018-02894-w

Collins, N., Poot, R. A., Kukimoto, I., García-Jiménez, C., Dellaire, G., & Varga-Weisz, P. D. (2002). An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. *Nature Genetics*, *32*(4), 627–632. https://doi.org/10.1038/ng1046

Cornacchia, D., Dileep, V., Quivy, J.-P., Foti, R., Tili, F., Santarella-Mellwig, R., … Buonomo, S. B. C. (2012). Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *The EMBO Journal*, *31*(18), 3678–3690. https://doi.org/10.1038/emboj.2012.214

Corradi, N., Pombert, J.-F., Farinelli, L., Didier, E. S., & Keeling, P. J. (2010). The complete sequence of the smallest known nuclear genome from the microsporidian Encephalitozoon intestinalis. *Nature Communications*, *1*(6), 77. https://doi.org/10.1038/ncomms1082

Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M. R., & Berger, J. M. (2011). The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nature Structural & Molecular Biology*, *18*(4), 471–477. https://doi.org/10.1038/nsmb.2004

Coster, G., & Diffley, J. F. X. (2017). Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. *Science*, *357*(6348), 314–318. https://doi.org/10.1126/science.aan0063

Cowieson, N. P., Partridge, J. F., Allshire, R. C., & McLaughlin, P. J. (2000). Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Current Biology : CB*, *10*(9), 517–525. https://doi.org/10.1016/s0960-9822(00)00467-x

Cramer, D., Serrano, L., & Schaefer, M. H. (2016). A network of epigenetic modifiers and DNA repair genes controls tissue-specific copy number alteration preference. *ELife*, *5*(NOVEMBER2016), 1–20. https://doi.org/10.7554/eLife.16519

Crosland, M. W. J., & Crozier, R. H. (1986). Myrmecia pilosula, an Ant with Only One Pair of Chromosomes. *Science (New York, N.Y.)*, *231*(4743), 1278. https://doi.org/10.1126/science.231.4743.1278

Crossley, M. P., Bocek, M., & Cimprich, K. A. (2019). R-Loops as Cellular Regulators and Genomic Threats. *Molecular Cell*, *73*(3), 398–411. https://doi.org/10.1016/j.molcel.2019.01.024

Dai, J., Chuang, R.-Y., & Kelly, T. J. (2005). DNA replication origins in the *Schizosaccharomyces pombe* genome. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(2), 337–342. https://doi.org/10.1073/pnas.0408811102

Daigaku, Y., Etheridge, T. J., Nakazawa, Y., Nakayama, M., Watson, A. T., Miyabe, I., … Carr, A. M. (2017). PCNA ubiquitylation ensures timely completion of unperturbed DNA replication in fission yeast. *PLoS Genetics*, *13*(5), e1006789. https://doi.org/10.1371/journal.pgen.1006789

Daigaku, Y., Keszthelyi, A., Müller, C. A., Miyabe, I., Brooks, T., Retkute, R., … Carr, A. M. (2015). A global profile of replicative polymerase usage. *Nature Structural & Molecular Biology*, *22*(3), 192–198. https://doi.org/10.1038/nsmb.2962

Darzynkiewicz, Z., Zhao, H., Zhang, S., Lee, M. Y. W. T., Lee, E. Y. C., & Zhang, Z. (2015). Initiation and termination of DNA replication during S phase in relation to cyclins D1, E and A, p21WAF1, Cdt1 and the p12 subunit of DNA polymerase δ revealed in individual cells by cytometry. *Oncotarget*, *6*(14), 11735–11750. https://doi.org/10.18632/oncotarget.4149

Das, P. P., Bagijn, M. P., Goldstein, L. D., Woolford, J. R., Lehrbach, N. J., Sapetschnig, A., … Miska, E. A. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Molecular Cell*, *31*(1), 79–90. https://doi.org/10.1016/j.molcel.2008.06.003

Das, S. P., Borrman, T., Liu, V. W. T., Yang, S. C. H., Bechhoefer, J., & Rhind, N. (2015). Replication timing is regulated by the number of MCMs loaded at origins. *Genome Research*, *25*(12), 1886–1892. https://doi.org/10.1101/gr.195305.115

Davé, A., Cooley, C., Garg, M., & Bianchi, A. (2014). Protein Phosphatase 1 Recruitment by Rif1 Regulates DNA Replication Origin Firing by Counteracting DDK Activity. *Cell Reports*, *7*(1), 53–61. https://doi.org/10.1016/j.celrep.2014.02.019

De Amicis, A., Piane, M., Ferrari, F., Fanciulli, M., Delia, D., & Chessa, L. (2011). Role of senataxin in DNA damage and telomeric stability. *DNA Repair*, *10*(2), 199–209. https://doi.org/10.1016/j.dnarep.2010.10.012

De Temmerman, N., Seneca, S., Van Steirteghem, A., Haentjens, P., Van der Elst, J., Liebaers, I., & Sermon, K. D. (2008). CTG repeat instability in a human embryonic stem cell line carrying the myotonic dystrophy type 1 mutation. *Molecular Human Reproduction*, *14*(7), 405–412. https://doi.org/10.1093/molehr/gan034

Deans, A. J., & West, S. C. (2011). DNA interstrand crosslink repair and cancer. *Nature Reviews. Cancer*, *11*(7), 467–480. https://doi.org/10.1038/nrc3088

Dell'Era, P., Nicoli, S., Peri, G., Nieddu, M., Ennas, M. G., & Presta, M. (2005). FGF2-induced upregulation of DNA polymerase-delta p12 subunit in endothelial cells. *Oncogene*, *24*(6), 1117–1121. https://doi.org/10.1038/sj.onc.1208359

DeMarini, D. J., Winey, M., Ursic, D., Webb, F., & Culbertson, M. R. (1992). SEN1, a positive effector of tRNA-splicing endonuclease in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, *12*(5), 2154–2164. https://doi.org/10.1128/mcb.12.5.2154

DePamphilis, M. L. (1993). Origins of DNA replication in metazoan chromosomes. *The Journal of Biological Chemistry*, *268*(1), 1–4. https://doi.org/10.1016/s0021-9258(18)54104-1

Dewar, J. M., Budzowska, M., & Walter, J. C. (2015). The mechanism of DNA replication termination in vertebrates. *Nature*, *525*(7569), 345–350. https://doi.org/10.1038/nature14887

Dewar, J. M., & Walter, J. C. (2017). Mechanisms of DNA replication termination. *Nature Reviews. Molecular Cell Biology*, *18*(8), 507–516. https://doi.org/10.1038/nrm.2017.42

DiNardo, S., Voelkel, K., & Sternglanz, R. (1984). DNA topoisomerase II mutant of *Saccharomyces cerevisiae*: topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, *81*(9), 2616–2620. https://doi.org/10.1073/pnas.81.9.2616

Donnianni, R. A., Zhou, Z.-X., Lujan, S. A., Al-Zain, A., Garcia, V., Glancy, E., … Symington, L. S. (2019). DNA Polymerase Delta Synthesizes Both Strands during Break-Induced Replication. *Molecular Cell*, *76*(3), 371-381.e4. https://doi.org/10.1016/j.molcel.2019.07.033

Douglas, M. E., Ali, F. A., Costa, A., & Diffley, J. F. X. (2018). The mechanism of eukaryotic CMG helicase activation. *Nature*, *555*(7695), 265–268. https://doi.org/10.1038/nature25787

Dueva, R., & Iliakis, G. (2020). Replication protein A: a multifunctional protein with roles in DNA replication, repair and beyond. *NAR Cancer*, *2*(3), 1437–1454. https://doi.org/10.1093/narcan/zcaa022

Duursma, A. M., Driscoll, R., Elias, J. E., & Cimprich, K. A. (2013). A role for the MRN complex in ATR activation via TOPBP1 recruitment. *Molecular Cell*, *50*(1), 116–122. https://doi.org/10.1016/j.molcel.2013.03.006

Ekwall, K., Javerzat, J. P., Lorentz, A., Schmidt, H., Cranston, G., & Allshire, R. (1995). The chromodomain protein Swi6: a key component at fission yeast centromeres. *Science (New York, N.Y.)*, *269*(5229), 1429–1431. https://doi.org/10.1126/science.7660126

Engel, S. R., Dietrich, F. S., Fisk, D. G., Binkley, G., Balakrishnan, R., Costanzo, M. C., … Cherry, J. M. (2014). The reference genome sequence of *Saccharomyces cerevisiae*: then and now. *G3 (Bethesda, Md.)*, *4*(3), 389–398. https://doi.org/10.1534/g3.113.008995

Etheridge, T. J., Boulineau, R. L., Herbert, A., Watson, A. T., Daigaku, Y., Tucker, J., … Carr, A. M. (2014). Quantification of DNA-associated proteins inside eukaryotic cells using single-molecule localization microscopy. *Nucleic Acids Research*, *42*(19), e146. https://doi.org/10.1093/nar/gku726

Fachinetti, D., Bermejo, R., Cocito, A., Minardi, S., Katou, Y., Kanoh, Y., … Foiani, M. (2010). Replication Termination at Eukaryotic Chromosomes Is Mediated by Top2 and Occurs at Genomic Loci Containing Pausing Elements. *Molecular Cell*, *39*(4), 595–605. https://doi.org/10.1016/j.molcel.2010.07.024

Fairman-Williams, M. E., Guenther, U.-P., & Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Current Opinion in Structural Biology*, *20*(3), 313–324. https://doi.org/10.1016/j.sbi.2010.03.011

Fang, D., Lengronne, A., Shi, D., Forey, R., Skrzypczak, M., Ginalski, K., … Lou, H. (2017). Dbf4 recruitment by forkhead transcription factors defines an upstream rate-limiting step in determining origin firing timing. *Genes & Development*, *31*(23–24), 2405–2415. https://doi.org/10.1101/gad.306571.117

Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N. J., Johnson, D. A., Richardson, T. B., … Ashworth, A. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature*, *434*(7035), 917–921. https://doi.org/10.1038/nature03445

Feng, W., & D'Urso, G. (2001). *Schizosaccharomyces pombe* cells lacking the amino-terminal catalytic domains of DNA polymerase epsilon are viable but require the DNA damage checkpoint control. *Molecular and Cellular Biology*, *21*(14), 4495–4504. https://doi.org/10.1128/MCB.21.14.4495-4504.2001

Fodor, B. D., Shukeir, N., Reuter, G., & Jenuwein, T. (2010). Mammalian Su(var) genes in chromatin control. *Annual Review of Cell and Developmental Biology*, *26*, 471–501. https://doi.org/10.1146/annurev.cellbio.042308.113225

Forsburg, S. L. (2013). The CINs of the centromere. *Biochemical Society Transactions*, *41*(6), 1706–1711. https://doi.org/10.1042/BST20130146

Forsburg, S. L., & Shen, K.-F. (2017). Centromere Stability: The Replication Connection. *Genes*, *8*(1). https://doi.org/10.3390/genes8010037

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

Furnari, B., Rhind, N., & Russell, P. (1997). Cdc25 mitotic inducer targeted by chk1 DNA damage checkpoint kinase. *Science (New York, N.Y.)*, *277*(5331), 1495–1497. https://doi.org/10.1126/science.277.5331.1495

Gan, H., Yu, C., Devbhandari, S., Sharma, S., Han, J., Chabes, A., … Zhang, Z. (2017). Checkpoint Kinase Rad53 Couples Leading- and Lagging-Strand DNA Synthesis under Replication Stress. *Molecular Cell*, *68*(2), 446-455.e3. https://doi.org/10.1016/j.molcel.2017.09.018

Gansauge, M.-T., Gerber, T., Glocke, I., Korlevic, P., Lippik, L., Nagel, S., … Meyer, M. (2017). Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Research*, *45*(10), e79. https://doi.org/10.1093/nar/gkx033

Garbacz, M. A., Lujan, S. A., Burkholder, A. B., Cox, P. B., Wu, Q., Zhou, Z.-X., … Kunkel, T. A. (2018). Evidence that DNA polymerase δ contributes to initiating leading strand DNA replication in *Saccharomyces cerevisiae*. *Nature Communications*, *9*(1), 858. https://doi.org/10.1038/s41467-018-03270-4

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

García-Muse, T., & Aguilera, A. (2019). R Loops: From Physiological to Pathological Roles. *Cell*, *179*(3), 604–618. https://doi.org/10.1016/j.cell.2019.08.055

Gardiner, K. (1995). Human genome organization. *Current Opinion in Genetics & Development*, *5*(3), 315–322. https://doi.org/10.1016/0959-437x(95)80045-x

Garrido-Ramos, M. A. (2017). Satellite DNA: An Evolving Topic. *Genes*, *8*(9). https://doi.org/10.3390/genes8090230

Gellon, L., Kaushal, S., Cebrián, J., Lahiri, M., Mirkin, S. M., & Freudenreich, C. H. (2019). Mrc1 and Tof1 prevent fragility and instability at long CAG repeats by their fork stabilizing function. *Nucleic Acids Research*, *47*(2), 794–805. https://doi.org/10.1093/nar/gky1195

Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I., & Chédin, F. (2013). GC skew at the 5’ and 3’ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. *Genome Research*, *23*(10), 1590–1600. https://doi.org/10.1101/gr.158436.113

Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I., & Chédin, F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Molecular Cell*, *45*(6), 814–825. https://doi.org/10.1016/j.molcel.2012.01.017

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., … Oliver, S. G. (1996). Life with 6000 genes. *Science (New York, N.Y.)*, *274*(5287), 546, 563–567. https://doi.org/10.1126/science.274.5287.546

Gong, Z., Kim, J.-E., Leung, C. C. Y., Glover, J. N. M., & Chen, J. (2010). BACH1/FANCJ acts with TopBP1 and participates early in DNA replication checkpoint control. *Molecular Cell*, *37*(3), 438–446. https://doi.org/10.1016/j.molcel.2010.01.002

Gopalakrishnan, V., Simancek, P., Houchens, C., Snaith, H. A., Frattini, M. G., Sazer, S., & Kelly, T. J. (2001). Redundant control of rereplication in fission yeast. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 13114–13119. https://doi.org/10.1073/pnas.221467598

Gordon, C. B., & Fantes, P. a. (1986). The *cdc22* gene of *Schizosaccharomyces pombe* encodes a cell cycle-regulated transcript. *The EMBO Journal*, *5*(11), 2981–2985. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16453725

Goren, A., Tabib, A., Hecht, M., & Cedar, H. (2008). DNA replication timing of the human beta-globin domain is controlled by histone modification at the origin. *Genes & Development*, *22*(10), 1319–1324. https://doi.org/10.1101/gad.468308

Gottipati, P., Cassel, T. N., Savolainen, L., & Helleday, T. (2008). Transcription-associated recombination is dependent on replication in Mammalian cells. *Molecular and Cellular Biology*, *28*(1), 154–164. https://doi.org/10.1128/MCB.00816-07

Guerrero-Puigdevall, M., Fernandez-Fuentes, N., & Frigola, J. (2021). Stabilisation of half MCM ring by Cdt1 during DNA insertion. *Nature Communications*, *12*(1), 1746. https://doi.org/10.1038/s41467-021-21932-8

Guiblet, W. M., Cremona, M. A., Harris, R. S., Chen, D., Eckert, K. A., Chiaromonte, F., … Makova, K. D. (2021). Non-B DNA: a major contributor to small- and large-scale variation in nucleotide substitution frequencies across the genome. *Nucleic Acids Research*, *49*(3), 1497–1516. https://doi.org/10.1093/nar/gkaa1269

Guilliam, T. A., & Doherty, A. J. (2017). PrimPol-Prime Time to Reprime. *Genes*, *8*(1). https://doi.org/10.3390/genes8010020

Guilliam, T. A., & Yeeles, J. T. (2021). The eukaryotic replisome tolerates leading-strand base damage by replicase switching. *The EMBO Journal*, *40*(5), e107037. https://doi.org/10.15252/embj.2020107037

Guilliam, T. A., & Yeeles, J. T. P. (2020a). An updated perspective on the polymerase division of labor during eukaryotic DNA replication. *Critical Reviews in Biochemistry and Molecular Biology*, *55*(5), 469–481. https://doi.org/10.1080/10409238.2020.1811630

Guilliam, T. A., & Yeeles, J. T. P. (2020b). Reconstitution of translesion synthesis reveals a mechanism of eukaryotic DNA replication restart. *Nature Structural & Molecular Biology*, *27*(5), 450–460. https://doi.org/10.1038/s41594-020-0418-4

Hall, I. M., Shankaranarayana, G. D., Noma, K.-I., Ayoub, N., Cohen, A., & Grewal, S. I. S. (2002). Establishment and maintenance of a heterochromatin domain. *Science (New York, N.Y.)*, *297*(5590), 2232–2237. https://doi.org/10.1126/science.1076466

Hamperl, S., Bocek, M. J., Saldivar, J. C., Swigut, T., & Cimprich, K. A. (2017). Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell*, *170*(4), 774-786.e19. https://doi.org/10.1016/j.cell.2017.07.043

Hamperl, S., & Cimprich, K. A. (2016). Conflict Resolution in the Genome: How Transcription and Replication Make It Work. *Cell*, *167*(6), 1455–1467. https://doi.org/10.1016/j.cell.2016.09.053

Hartley, G., & O'Neill, R. (2019). Centromere Repeats: Hidden Gems of the Genome. *Genes*, *10*(3), 223. https://doi.org/10.3390/genes10030223

Hartono, S. R., Malapert, A., Legros, P., Bernard, P., Chédin, F., & Vanoosthuyse, V. (2018). The Affinity of the S9.6 Antibody for Double-Stranded RNAs Impacts the Accurate Mapping of R-Loops in Fission Yeast. *Journal of Molecular Biology*, *430*(3), 272–284. https://doi.org/10.1016/j.jmb.2017.12.016

Hashimoto, Y., Tsujimura, T., Sugino, A., & Takisawa, H. (2006). The phosphorylated C-terminal domain of Xenopus Cut5 directly mediates ATR-dependent activation of Chk1. *Genes to Cells : Devoted to Molecular & Cellular Mechanisms*, *11*(9), 993–1007. https://doi.org/10.1111/j.1365-2443.2006.00998.x

Hatchi, E., Skourti-Stathaki, K., Ventz, S., Pinello, L., Yen, A., Kamieniarz-Gdula, K., … Livingston, D. M. (2015). BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Molecular Cell*, *57*(4), 636–647. https://doi.org/10.1016/j.molcel.2015.01.011

Hayano, M., Kanoh, Y., Matsumoto, S., Renard-Guillet, C., Shirahige, K., & Masai, H. (2012). Rif1 is a global regulator of timing of replication origin firing in fission yeast. *Genes & Development*, *26*(2), 137–150. https://doi.org/10.1101/gad.178491.111

Hayashi, M. T., Takahashi, T. S., Nakagawa, T., Nakayama, J., & Masukata, H. (2009). The heterochromatin protein Swi6/HP1 activates replication origins at the pericentromeric region and silent mating-type locus. *Nature Cell Biology*, *11*(3), 357–362. https://doi.org/10.1038/ncb1845

Hayles, J., Wood, V., Jeffery, L., Hoe, K.-L., Kim, D.-U., Park, H.-O., … Nurse, P. (2013). A genome-wide resource of cell cycle and cell shape genes of fission yeast. *Open Biology*, *3*(5), 130053. https://doi.org/10.1098/rsob.130053

Hazelbaker, D. Z., Marquardt, S., Wlotzka, W., & Buratowski, S. (2013). Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. *Molecular Cell*, *49*(1), 55–66. https://doi.org/10.1016/j.molcel.2012.10.014

Heffernan, T. P., Ünsal-Kaçmaz, K., Heinloth, A. N., Simpson, D. A., Paules, R. S., Sancar, A., … Kaufmann, W. K. (2007). Cdc7-Dbf4 and the human S checkpoint response to UVC. *The Journal of Biological Chemistry*, *282*(13), 9458–9468. https://doi.org/10.1074/jbc.M611292200

Heller, R. C., Kang, S., Lam, W. M., Chen, S., Chan, C. S., & Bell, S. P. (2011). Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. *Cell*, *146*(1), 80–91. https://doi.org/10.1016/j.cell.2011.06.012

Helmrich, A., Ballarino, M., & Tora, L. (2011). Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Molecular Cell*, *44*(6), 966–977. https://doi.org/10.1016/j.molcel.2011.10.013

Hizume, K., & Araki, H. (2019). Replication fork pausing at protein barriers on chromosomes. *FEBS Letters*, *593*(13), 1449–1458. https://doi.org/10.1002/1873-3468.13481

Howes, T. R. L., & Tomkinson, A. E. (2012). DNA ligase I, the replicative DNA ligase. *Sub-Cellular Biochemistry*, *62*, 327–341. https://doi.org/10.1007/978-94-007-4572-8_17

Hu, J., & Xiong, Y. (2006). An evolutionarily conserved function of proliferating cell nuclear antigen for Cdt1 degradation by the Cul4-Ddb1 ubiquitin ligase in response to DNA damage. *The Journal of Biological Chemistry*, *281*(7), 3753–3756. https://doi.org/10.1074/jbc.C500464200

Hu, L., Kim, T. M., Son, M. Y., Kim, S.-A., Holland, C. L., Tateishi, S., … Hasty, P. (2013). Two replication fork maintenance pathways fuse inverted repeats to rearrange chromosomes. *Nature*, *501*(7468), 569–572. https://doi.org/10.1038/nature12500

Huang, Q. M., Akashi, T., Masuda, Y., Kamiya, K., Takahashi, T., & Suzuki, M. (2010). Roles of POLD4, smallest subunit of DNA polymerase delta, in nuclear structures and genomic stability of human cells. *Biochemical and Biophysical Research Communications*, *391*(1), 542–546. https://doi.org/10.1016/j.bbrc.2009.11.094

Huang, Q. M., Tomida, S., Masuda, Y., Arima, C., Cao, K., Kasahara, T.-A., … Suzuki, M. (2010). Regulation of DNA polymerase POLD4 influences genomic instability in lung cancer. *Cancer Research*, *70*(21), 8407–8416. https://doi.org/10.1158/0008-5472.CAN-10-0784

Hughes, P., Tratner, I., Ducoux, M., Piard, K., & Baldacci, G. (1999). Isolation and identification of the third subunit of mammalian DNA polymerase delta by PCNA-affinity chromatography of mouse FM3A cell extracts. *Nucleic Acids Research*, *27*(10), 2108–2114. https://doi.org/10.1093/nar/27.10.2108

Ibuki, Y., & Toyooka, T. (2015). Evaluation of chemical phototoxicity, focusing on phosphorylated histone H2AX. *Journal of Radiation Research*, *56*(2), 220–228. https://doi.org/10.1093/jrr/rru105

Iino, Y., & Yamamoto, M. (1997). The *Schizosaccharomyces pombe cdc6* gene encodes the catalytic subunit of DNA polymerase delta. *Molecular & General Genetics : MGG*, *254*(1), 93–97. https://doi.org/10.1007/s004380050395

Ilves, I., Petojevic, T., Pesavento, J. J., & Botchan, M. R. (2010). Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Molecular Cell*, *37*(2), 247–258. https://doi.org/10.1016/j.molcel.2009.12.030

Im, J.-S., Ki, S.-H., Farina, A., Jung, D.-S., Hurwitz, J., & Lee, J.-K. (2009). Assembly of the Cdc45-Mcm2-7-GINS complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(37), 15628–15632. https://doi.org/10.1073/pnas.0908039106

Irony-Tur Sinai, M., Salamon, A., Stanleigh, N., Goldberg, T., Weiss, A., Wang, Y.-H., & Kerem, B. (2019). AT-dinucleotide rich sequences drive fragile site formation. *Nucleic Acids Research*, *47*(18), 9685–9695. https://doi.org/10.1093/nar/gkz689

Ishida, M., Shimojo, H., Hayashi, A., Kawaguchi, R., Ohtani, Y., Uegaki, K., … Nakayama, J.-I. (2012). Intrinsic nucleic acid-binding activity of Chp1 chromodomain is required for heterochromatic gene silencing. *Molecular Cell*, *47*(2), 228–241. https://doi.org/10.1016/j.molcel.2012.05.017

Ivanova, A. V., Bonaduce, M. J., Ivanov, S. V., & Klar, A. J. S. (1998). The chromo and SET domains of the Clr4 protein are essential for silencing in fission yeast. *Nature Genetics*, *19*(2), 192–195. https://doi.org/10.1038/566

Jain, R., Rice, W. J., Malik, R., Johnson, R. E., Prakash, L., Prakash, S., … Aggarwal, A. K. (2019). Cryo-EM structure and dynamics of eukaryotic DNA polymerase δ holoenzyme. *Nature Structural & Molecular Biology*, *26*(10), 955–962. https://doi.org/10.1038/s41594-019-0305-z

Jallepalli, P. V., Brown, G. W., Muzi-Falconi, M., Tien, D., & Kelly, T. J. (1997). Regulation of the replication initiator protein p65cdc18 by CDK phosphorylation. *Genes & Development*, *11*(21), 2767–2779. https://doi.org/10.1101/gad.11.21.2767

Jamonnak, N., Creamer, T. J., Darby, M. M., Schaughency, P., Wheelan, S. J., & Corden, J. L. (2011). Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA (New York, N.Y.)*, *17*(11), 2011–2025. https://doi.org/10.1261/rna.2840711

Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. *Trends in Biochemical Sciences*, *36*(1), 19–29. https://doi.org/10.1016/j.tibs.2010.07.008

Janssen, A., Colmenares, S. U., & Karpen, G. H. (2018). Heterochromatin: Guardian of the Genome. *Annual Review of Cell and Developmental Biology*, *34*, 265–288. https://doi.org/10.1146/annurev-cellbio-100617-062653

Jia, S., Noma, K., & Grewal, S. I. S. (2004). RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. *Science (New York, N.Y.)*, *304*(5679), 1971–1976. https://doi.org/10.1126/science.1099035

Kamada, K., Kubota, Y., Arata, T., Shindo, Y., & Hanaoka, F. (2007). Structure of the human GINS complex and its assembly and functional interface in replication initiation. *Nature Structural & Molecular Biology*, *14*(5), 388–396. https://doi.org/10.1038/nsmb1231

Kang, H. Y., Choi, E., Bae, S. H., Lee, K. H., Gim, B. S., Kim, H. D., … Seo, Y. S. (2000). Genetic analyses of *Schizosaccharomyces pombe dna2*$^{+}$ reveal that *dna2* plays an essential role in Okazaki fragment metabolism. *Genetics*, *155*(3), 1055–1067. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10880469

Kanke, M., Kodama, Y., Takahashi, T. S., Nakagawa, T., & Masukata, H. (2012). Mcm10 plays an essential role in origin DNA unwinding after loading of the CMG components. *The EMBO Journal*, *31*(9), 2182–2194. https://doi.org/10.1038/emboj.2012.68

Kanoh, J., Sadaie, M., Urano, T., & Ishikawa, F. (2005). Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres. *Current Biology : CB*, *15*(20), 1808–1819. https://doi.org/10.1016/j.cub.2005.09.041

Kaplan, D. L., Davey, M. J., & O'Donnell, M. (2003). Mcm4,6,7 uses a "pump in ring" mechanism to unwind DNA by steric exclusion and actively translocate along a duplex. *The Journal of Biological Chemistry*, *278*(49), 49171–49182. https://doi.org/10.1074/jbc.M308074200

Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., & Proudfoot, N. J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes & Development*, *22*(8), 1082–1092. https://doi.org/10.1101/gad.463408

Kazadi, D., Lim, J., Rothschild, G., Grinstein, V., Laffleur, B., Becherel, O., … Basu, U. (2020). Effects of senataxin and RNA exosome on B-cell chromosomal integrity. *Heliyon*, *6*(3), e03442. https://doi.org/10.1016/j.heliyon.2020.e03442

Kellner, V., & Luke, B. (2020). Molecular and physiological consequences of faulty eukaryotic ribonucleotide excision repair. *The EMBO Journal*, *39*(3), e102309. https://doi.org/10.15252/embj.2019102309

Kelly, T. J., Martin, G. S., Forsburg, S. L., Stephen, R. J., Russo, A., & Nurse, P. (1993). The fission yeast *cdc18*$^{+}$ gene product couples S phase to START and mitosis. *Cell*, *74*(2), 371–382. https://doi.org/10.1016/0092-8674(93)90427-r

Kesti, T., Flick, K., Keränen, S., Syväoja, J. E., & Wittenberg, C. (1999). DNA polymerase epsilon catalytic domains are dispensable for DNA replication, DNA repair, and cell viability. *Molecular Cell*, *3*(5), 679–685. https://doi.org/10.1016/S1097-2765(00)80361-5

Keszthelyi, A., Daigaku, Y., Ptasińska, K., Miyabe, I., & Carr, A. M. (2015). Mapping ribonucleotides in genomic DNA and exploring replication dynamics by polymerase usage sequencing (Pu-seq). *Nature Protocols*, *10*(11), 1786–1801. https://doi.org/10.1038/nprot.2015.116

Khandagale, P., Peroumal, D., Manohar, K., & Acharya, N. (2019). Human DNA polymerase delta is a pentameric holoenzyme with a dimeric p12 subunit. *Life Science Alliance*, *2*(2), e201900323. https://doi.org/10.26508/lsa.201900323

Kim, H. D., Choe, J., & Seo, Y. S. (1999). The *sen1*$^{+}$ gene of *Schizosaccharomyces pombe*, a homologue of budding yeast SEN1, encodes an RNA and DNA helicase. *Biochemistry*, *38*(44), 14697–14710. https://doi.org/10.1021/bi991470c

Kirstein, N., Buschle, A., Wu, X., Krebs, S., Blum, H., Kremmer, E., … Schepers, A. (2021). Human ORC/MCM density is low in active genes and correlates with replication time but does not delimit initiation zones. *ELife*, *10*(Mcm), 1–30. https://doi.org/10.7554/eLife.62161

Knott, S. R. V., Peace, J. M., Ostrow, A. Z., Gan, Y., Rex, A. E., Viggiani, C. J., … Aparicio, O. M. (2012). Forkhead transcription factors establish origin timing and long-range clustering in S. *cerevisiae*. *Cell*, *148*(1–2), 99–111. https://doi.org/10.1016/j.cell.2011.12.012

Knott, S. R. V, Viggiani, C. J., Tavaré, S., & Aparicio, O. M. (2009). Genome-wide replication profiles indicate an expansive role for Rpd3L in regulating replication initiation timing or efficiency, and reveal genomic loci of Rpd3 function in *Saccharomyces cerevisiae*. *Genes & Development*, *23*(9), 1077–1090. https://doi.org/10.1101/gad.1784309

Koh, K. D., Balachander, S., Hesselberth, J. R., & Storici, F. (2015). Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. *Nature Methods*, *12*(3), 251–257, 3 p following 257. https://doi.org/10.1038/nmeth.3259

Komarnitsky, P., Cho, E. J., & Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & Development*, *14*(19), 2452–2460. https://doi.org/10.1101/gad.824700

Kramara, J., Osia, B., & Malkova, A. (2018). Break-Induced Replication: The Where, The Why, and The How. *Trends in Genetics : TIG*, *34*(7), 518–531. https://doi.org/10.1016/j.tig.2018.04.002

Kunkel, T. A., Hamatake, R. K., Motto-Fox, J., Fitzgerald, M. P., & Sugino, A. (1989). Fidelity of DNA polymerase I and the DNA polymerase I-DNA primase complex from *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, *9*(10), 4447–4458. https://doi.org/10.1128/mcb.9.10.4447-4458.1989

Lambert, S., Watson, A., Sheedy, D. M., Martin, B., & Carr, A. M. (2005). Gross chromosomal rearrangements and elevated recombination at an inducible site-specific replication fork barrier. *Cell*, *121*(5), 689–702. https://doi.org/10.1016/j.cell.2005.03.022

Lancey, C., Tehseen, M., Raducanu, V.-S., Rashid, F., Merino, N., Ragan, T. J., … De Biasio, A. (2020). Structure of the processive human Pol δ holoenzyme. *Nature Communications*, *11*(1), 1109. https://doi.org/10.1038/s41467-020-14898-6

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., … International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. https://doi.org/10.1038/35057062

Langley, A. R., Gräf, S., Smith, J. C., & Krude, T. (2016). Genome-wide identification and characterisation of human DNA replication origins by initiation site sequencing (ini-seq). *Nucleic Acids Research*, *44*(21), 10230–10247. https://doi.org/10.1093/nar/gkw760

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

Larochelle, M., Robert, M.-A., Hébert, J.-N., Liu, X., Matteau, D., Rodrigue, S., … Bachand, F. (2018). Common mechanism of transcription termination at coding and noncoding RNA genes in fission yeast. *Nature Communications*, *9*(1), 4364. https://doi.org/10.1038/s41467-018-06546-x

Lee, C., Hong, B., Choi, J. M., Kim, Y., Watanabe, S., Ishimi, Y., … Cho, Y. (2004). Structural basis for inhibition of the replication licensing factor Cdt1 by geminin. *Nature*, *430*(7002), 913–917. https://doi.org/10.1038/nature02813

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L. J., … Cancer Genome Atlas Research Network. (2012). Landscape of somatic retrotransposition in human cancers. *Science (New York, N.Y.)*, *337*(6097), 967–971. https://doi.org/10.1126/science.1222077

Lee, M. Y., Jiang, Y. Q., Zhang, S. J., & Toomey, N. L. (1991). Characterization of human DNA polymerase delta and its immunochemical relationships with DNA polymerase alpha and epsilon. *The Journal of Biological Chemistry*, *266*(4), 2423–2429. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1703528

Lee, M. Y. W. T., Zhang, S., Wang, X., Chao, H. H., Zhao, H., Darzynkiewicz, Z., … Lee, E. Y. C. (2019). Two forms of human DNA polymerase δ: Who does what and why? *DNA Repair*, *81*(July), 102656. https://doi.org/10.1016/j.dnarep.2019.102656

Legros, P., Malapert, A., Niinuma, S., Bernard, P., & Vanoosthuyse, V. (2014). RNA processing factors Swd2.2 and Sen1 antagonize RNA Pol III-dependent transcription and the localization of condensin at Pol III genes. *PLoS Genetics*, *10*(11), e1004794. https://doi.org/10.1371/journal.pgen.1004794

Leimbacher, P.-A., Jones, S. E., Shorrocks, A.-M. K., de Marco Zompit, M., Day, M., Blaauwendraad, J., … Stucki, M. (2019). MDC1 Interacts with TOPBP1 to Maintain Chromosomal Stability during Mitosis. *Molecular Cell*, *74*(3), 571-583.e8. https://doi.org/10.1016/j.molcel.2019.02.014

Leman, A. R., & Noguchi, E. (2013). The replication fork: understanding the eukaryotic replication machinery and the challenges to genome duplication. *Genes*, *4*(1), 1–32. https://doi.org/10.3390/genes4010001

Lemay, J.-F., Marguerat, S., Larochelle, M., Liu, X., van Nues, R., Hunyadkürti, J., … Bachand, F. (2016). The Nrd1-like protein Seb1 coordinates cotranscriptional 3' end processing and polyadenylation site selection. *Genes & Development*, *30*(13), 1558–1572. https://doi.org/10.1101/gad.280222.116

Leonaitė, B., Han, Z., Basquin, J., Bonneau, F., Libri, D., Porrua, O., & Conti, E. (2017). Sen1 has unique structural features grafted on the architecture of the Upf1-like helicase family. *The EMBO Journal*, *36*(11), 1590–1604. https://doi.org/10.15252/embj.201696174

Li, G., & Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. *Current Opinion in Genetics and Development*, *21*(2), 175–186. https://doi.org/10.1016/j.gde.2011.01.022

Li, H., Xie, B., Zhou, Y., Rahmeh, A., Trusa, S., Zhang, S., … Lee, M. Y. W. T. (2006). Functional roles of p12, the fourth subunit of human DNA polymerase delta. *The Journal of Biological Chemistry*, *281*(21), 14748–14755. https://doi.org/10.1074/jbc.M600322200

Li, P.-C., Chretien, L., Côté, J., Kelly, T. J., & Forsburg, S. L. (2011). *S. pombe* replication protein Cdc18 (Cdc6) interacts with Swi6 (HP1) heterochromatin protein: region specific effects and replication timing in the centromere. *Cell Cycle (Georgetown, Tex.)*, *10*(2), 323–336. https://doi.org/10.4161/cc.10.2.14552

Lin, S. H. S., Wang, X., Zhang, S., Zhang, Z., Lee, E. Y. C., & Lee, M. Y. W. T. (2013). Dynamics of enzymatic interactions during short flap human Okazaki fragment processing by two forms of human DNA polymerase δ. *DNA Repair*, *12*(11), 922–935. https://doi.org/10.1016/j.dnarep.2013.08.008

Liu, C., Poitelea, M., Watson, A., Yoshida, S., Shimoda, C., Holmberg, C., … Carr, A. M. (2005). Transactivation of *Schizosaccharomyces pombe cdt2*$^{+}$ stimulates a Pcu4-Ddb1-CSN ubiquitin ligase. *The EMBO Journal*, *24*(22), 3940–3951. https://doi.org/10.1038/sj.emboj.7600854

Liu, L., Mo, J., Rodriguez-Belmonte, E. M., & Lee, M. Y. W. T. (2000). Identification of a fourth subunit of mammalian DNA polymerase delta. *The Journal of Biological Chemistry*, *275*(25), 18739–18744. https://doi.org/10.1074/jbc.M001217200

Liu, Q., Guntuku, S., Cui, X. S., Matsuoka, S., Cortez, D., Tamai, K., … Elledge, S. J. (2000). Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint. *Genes & Development*, *14*(12), 1448–1459. https://doi.org/10.1101/gad.14.12.1448

Liu, S., Opiyo, S. O., Manthey, K., Glanzer, J. G., Ashley, A. K., Amerin, C., … Oakley, G. G. (2012). Distinct roles for DNA-PK, ATM and ATR in RPA phosphorylation and checkpoint activation in response to replication stress. *Nucleic Acids Research*, *40*(21), 10780–10794. https://doi.org/10.1093/nar/gks849

Lock, A., Rutherford, K., Harris, M. A., Hayles, J., Oliver, S. G., Bähler, J., & Wood, V. (2019). PomBase 2018: user-driven reimplementation of the fission yeast database provides rapid and intuitive access to diverse, interconnected

information. *Nucleic Acids Research*, *47*(D1), D821–D827. https://doi.org/10.1093/nar/gky961

Lopez-Girona, A., Tanaka, K., Chen, X. B., Baber, B. A., McGowan, C. H., & Russell, P. (2001). Serine-345 is required for Rad3-dependent phosphorylation and function of checkpoint kinase Chk1 in fission yeast. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11289–11294. https://doi.org/10.1073/pnas.191557598

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

Lukhtanov, V. A. (2015). The blue butterfly Polyommatus (Plebicula) atlanticus (Lepidoptera, Lycaenidae) holds the record of the highest number of chromosomes in the non-polyploid eukaryotic organisms. *Comparative Cytogenetics*, *9*(4), 683–690. https://doi.org/10.3897/CompCytogen.v9i4.5760

Lygeros, J., Koutroumpas, K., Dimopoulos, S., Legouras, I., Kouretas, P., Heichinger, C., … Lygerou, Z. (2008). Stochastic hybrid modeling of DNA replication across a complete genome. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(34), 12295–12300. https://doi.org/10.1073/pnas.0805549105

Lynch, K. L., Alvino, G. M., Kwan, E. X., Brewer, B. J., & Raghuraman, M. K. (2019). The effects of manipulating levels of replication initiation factors on origin firing efficiency in yeast. *PLoS Genetics*, *15*(10), e1008430. https://doi.org/10.1371/journal.pgen.1008430

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

MacNeill, S. A., Moreno, S., Reynolds, N., Nurse, P., & Fantes, P. A. (1996). The fission yeast Cdc1 protein, a homologue of the small subunit of DNA polymerase delta, binds to Pol3 and Cdc27. *The EMBO Journal*, *15*(17), 4613–4628. https://doi.org/10.1002/j.1460-2075.1996.tb00839.x

Madeira, F., Park, Y. mi, Lee, J., Buso, N., Gur, T., Madhusoodanan, N., … Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, *47*(W1), W636–W641. https://doi.org/10.1093/nar/gkz268

Maksimov, V., Oya, E., Tanaka, M., Kawaguchi, T., Hachisuka, A., Ekwall, K., … Nakayama, J.-I. (2018). The binding of Chp2's chromodomain to methylated H3K9 is essential for Chp2's role in heterochromatin assembly in fission yeast. *PloS One*, *13*(8), e0201101. https://doi.org/10.1371/journal.pone.0201101

Mantiero, D., Mackenzie, A., Donaldson, A., & Zegerman, P. (2011). Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *The EMBO Journal*, *30*(23), 4805–4814. https://doi.org/10.1038/emboj.2011.404

Marahrens, Y., & Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science (New York, N.Y.)*, *255*(5046), 817–823. https://doi.org/10.1126/science.1536007

Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., & Bähler, J. (2012). Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell*, *151*(3), 671–683. https://doi.org/10.1016/j.cell.2012.09.019

Maric, M., Maculins, T., De Piccoli, G., & Labib, K. (2014). Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. *Science (New York, N.Y.)*, *346*(6208), 1253596. https://doi.org/10.1126/science.1253596

Marks, A. B., Fu, H., & Aladjem, M. I. (2017). Regulation of Replication Origins. *Advances in Experimental Medicine and Biology*, *1042*, 43–59. https://doi.org/10.1007/978-981-10-6955-0_2

Martienssen, R., & Moazed, D. (2015). RNAi and heterochromatin assembly. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a019323. https://doi.org/10.1101/cshperspect.a019323

Martin-Tumasz, S., & Brow, D. A. (2015). *Saccharomyces cerevisiae* Sen1 Helicase Domain Exhibits 5'- to 3'-Helicase Activity with a Preference for Translocation on DNA Rather than RNA. *The Journal of Biological Chemistry*, *290*(38), 22880–22889. https://doi.org/10.1074/jbc.M115.674002

Martire, S., & Banaszynski, L. A. (2020). The roles of histone variants in fine-tuning chromatin organization and function. *Nature Reviews. Molecular Cell Biology*, *21*(9), 522–541. https://doi.org/10.1038/s41580-020-0262-8

Masai, H., Taniyama, C., Ogino, K., Matsui, E., Kakusho, N., Matsumoto, S., … Arai, K.-I. (2006). Phosphorylation of MCM4 by Cdc7 kinase facilitates its interaction with Cdc45 on the chromatin. *The Journal of Biological Chemistry*, *281*(51), 39249–39261. https://doi.org/10.1074/jbc.M608935200

Mata, J., Lyne, R., Burns, G., & Bähler, J. (2002). The transcriptional program of meiosis and sporulation in fission yeast. *Nature Genetics*, *32*(1), 143–147. https://doi.org/10.1038/ng951

Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., … Elledge, S. J. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science (New York, N.Y.)*, *316*(5828), 1160–1166. https://doi.org/10.1126/science.1140321

McVey, M., Khodaverdian, V. Y., Meyer, D., Cerqueira, P. G., & Heyer, W.-D. (2016). Eukaryotic DNA Polymerases in Homologous Recombination. *Annual Review of Genetics*, *50*(1), 393–421. https://doi.org/10.1146/annurev-genet-120215-035243

Mendez-Bermudez, A., Lototska, L., Bauwens, S., Giraud-Panis, M.-J., Croce, O., Jamet, K., … Ye, J. (2018). Genome-wide Control of Heterochromatin Replication by the Telomere Capping Protein TRF2. *Molecular Cell*, *70*(3), 449-461.e5. https://doi.org/10.1016/j.molcel.2018.03.036

Méndez, J., Zou-Yang, X. H., Kim, S. Y., Hidaka, M., Tansey, W. P., & Stillman, B. (2002). Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. *Molecular Cell*, *9*(3), 481–491. https://doi.org/10.1016/S1097-2765(02)00467-7

Meng, X., Zhou, Y., Lee, E. Y. C., Lee, M. Y. W. T., & Frick, D. N. (2010). The p12 subunit of human polymerase delta modulates the rate and fidelity of DNA synthesis. *Biochemistry*, *49*(17), 3545–3554. https://doi.org/10.1021/bi100042b

Meng, X., Zhou, Y., Zhang, S., Lee, E. Y. C., Frick, D. N., & Lee, M. Y. W. T. (2009). DNA damage alters DNA polymerase delta to a form that exhibits increased discrimination against modified template bases and mismatched primers. *Nucleic Acids Research*, *37*(2), 647–657. https://doi.org/10.1093/nar/gkn1000

Mesner, L. D., Valsakumar, V., Cieslik, M., Pickin, R., Hamlin, J. L., & Bekiranov, S. (2013). Bubble-seq analysis of the human genome reveals distinct chromatin-mediated mechanisms for regulating early- and late-firing origins. *Genome Research*, *23*(11), 1774–1788. https://doi.org/10.1101/gr.155218.113

Michael, W. M., Ott, R., Fanning, E., & Newport, J. (2000). Activation of the DNA replication checkpoint through RNA synthesis by primase. *Science (New York, N.Y.)*, *289*(5487), 2133–2137. https://doi.org/10.1126/science.289.5487.2133

Mimura, S., Seki, T., Tanaka, S., & Diffley, J. F. X. (2004). Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control. *Nature*, *431*(7012), 1118–1123. https://doi.org/10.1038/nature03024

Mirkin, E. V, & Mirkin, S. M. (2007). Replication fork stalling at natural impediments. *Microbiology and Molecular Biology Reviews : MMBR*, *71*(1), 13–35. https://doi.org/10.1128/MMBR.00030-06

Mischo, H. E., Gómez-González, B., Grzechnik, P., Rondón, A. G., Wei, W., Steinmetz, L., … Proudfoot, N. J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. *Molecular Cell*, *41*(1), 21–32. https://doi.org/10.1016/j.molcel.2010.12.007

Miyabe, I., Kunkel, T. A., & Carr, A. M. (2011). The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genetics*, *7*(12), e1002407. https://doi.org/10.1371/journal.pgen.1002407

Miyabe, I., Mizuno, K., Keszthelyi, A., Daigaku, Y., Skouteri, M., Mohebi, S., … Carr, A. M. (2015). Polymerase δ replicates both strands after homologous recombination-dependent fork restart. *Nature Structural & Molecular Biology*, *22*(11), 932–938. https://doi.org/10.1038/nsmb.3100

Mizuno, K., Lambert, S., Baldacci, G., Murray, J. M., & Carr, A. M. (2009). Nearby inverted repeats fuse to generate acentric and dicentric palindromic chromosomes by a replication template exchange mechanism. *Genes & Development*, *23*(24), 2876–2886. https://doi.org/10.1101/gad.1863009

Mocquet, V., Lainé, J. P., Riedl, T., Yajin, Z., Lee, M. Y., & Egly, J. M. (2008). Sequential recruitment of the repair factors during NER: the role of XPG in initiating the resynthesis step. *The EMBO Journal*, *27*(1), 155–167. https://doi.org/10.1038/sj.emboj.7601948

Molinari, M., Mercurio, C., Dominguez, J., Goubin, F., & Draetta, G. F. (2000). Human Cdc25 A inactivation in response to S phase inhibition and its role in preventing premature mitosis. *EMBO Reports*, *1*(1), 71–79. https://doi.org/10.1093/embo-reports/kvd018

Mondol, T., Stodola, J. L., Galletto, R., & Burgers, P. M. (2019). PCNA accelerates the nucleotide incorporation rate by DNA polymerase δ. *Nucleic Acids Research*, *47*(4), 1977–1986. https://doi.org/10.1093/nar/gky1321

Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology*, *38*(1), 23–38. https://doi.org/10.1038/npp.2012.112

Moreira, M.-C., Klur, S., Watanabe, M., Németh, A. H., Le Ber, I., Moniz, J.-C., … Koenig, M. (2004). Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nature Genetics*, *36*(3), 225–227. https://doi.org/10.1038/ng1303

Moreno, S. P., Bailey, R., Campion, N., Herron, S., & Gambus, A. (2014). Polyubiquitylation drives replisome disassembly at the termination of DNA replication. *Science (New York, N.Y.)*, *346*(6208), 477–481. https://doi.org/10.1126/science.1253585

Müller, C. A., Boemo, M. A., Spingardi, P., Kessler, B. M., Kriaucionis, S., Simpson, J. T., & Nieduszynski, C. A. (2019). Capturing the dynamics of genome replication on individual ultra-long nanopore sequence reads. *Nature Methods*, *16*(5), 429–436. https://doi.org/10.1038/s41592-019-0394-y

Müller, C. A., Hawkins, M., Retkute, R., Malla, S., Wilson, R., Blythe, M. J., … Nieduszynski, C. A. (2014). The dynamics of genome replication using deep sequencing. *Nucleic Acids Research*, *42*(1), e3. https://doi.org/10.1093/nar/gkt878

Müller, C. A., & Nieduszynski, C. A. (2017). DNA replication timing influences gene expression level. *The Journal of Cell Biology*, *216*(7), 1907–1914. https://doi.org/10.1083/jcb.201701061

Muramatsu, S., Hirai, K., Tak, Y.-S., Kamimura, Y., & Araki, H. (2010). CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon}, and GINS in budding yeast. *Genes & Development*, *24*(6), 602–612. https://doi.org/10.1101/gad.1883410

Nahas, S. A., Duquette, A., Roddier, K., Gatti, R. A., & Brais, B. (2007). Ataxia-oculomotor apraxia 2 patients show no increased sensitivity to ionizing radiation. *Neuromuscular Disorders : NMD*, *17*(11–12), 968–969. https://doi.org/10.1016/j.nmd.2007.06.464

Naiman, K., Campillo-Funollet, E., Watson, A. T., Budden, A., Miyabe, I., & Carr, A. M. (2021). Replication dynamics of recombination-dependent replication forks. *Nature Communications*, *12*(1), 923. https://doi.org/10.1038/s41467-021-21198-0

Natsume, T., Müller, C. A., Katou, Y., Retkute, R., Gierliński, M., Araki, H., … Tanaka, T. U. (2013). Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Molecular Cell*, *50*(5), 661–674. https://doi.org/10.1016/j.molcel.2013.05.011

Nguyen, V Q, Co, C., & Li, J. J. (2001). Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature*, *411*(6841), 1068–1073. https://doi.org/10.1038/35082600

Nguyen, Van Q, Co, C., Irie, K., & Li, J. J. (2000). Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7. *Current Biology : CB*, *10*(4), 195–205. https://doi.org/10.1016/s0960-9822(00)00337-7

Ogino, K., Takeda, T., Matsui, E., Iiyama, H., Taniyama, C., Arai, K., & Masai, H. (2001). Bipartite binding of a kinase activator activates Cdc7-related kinase essential for S phase. *The Journal of Biological Chemistry*, *276*(33), 31376–31387. https://doi.org/10.1074/jbc.M102197200

Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., & Sugino, A. (1968). Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proceedings of the National Academy of Sciences of the United States of America*, *59*(2), 598–605. https://doi.org/10.1073/pnas.59.2.598

Pacek, M., & Walter, J. C. (2004). A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. *The EMBO Journal*, *23*(18), 3667–3676. https://doi.org/10.1038/sj.emboj.7600369

Pan, X., Lei, B., Zhou, N., Feng, B., Yao, W., Zhao, X., … Lu, H. (2012). Identification of novel genes involved in DNA damage response by screening a genome-wide *Schizosaccharomyces pombe* deletion library. *BMC Genomics*, *13*(1), 662. https://doi.org/10.1186/1471-2164-13-662

Parrilla-Castellar, E. R., & Karnitz, L. M. (2003). Cut5 is required for the binding of Atr and DNA polymerase alpha to genotoxin-damaged chromatin. *The Journal of Biological Chemistry*, *278*(46), 45507–45511. https://doi.org/10.1074/jbc.C300418200

Partridge, J. F., Scott, K. S. C., Bannister, A. J., Kouzarides, T., & Allshire, R. C. (2002). cis-acting DNA from fission yeast centromeres mediates histone H3 methylation and recruitment of silencing factors and cohesin to an ectopic site. *Current Biology : CB*, *12*(19), 1652–1660. https://doi.org/10.1016/s0960-9822(02)01177-6

Patel, P. K., Kommajosyula, N., Rosebrock, A., Bensimon, A., Leatherwood, J., Bechhoefer, J., & Rhind, N. (2008). The Hsk1(Cdc7) replication kinase regulates origin efficiency. *Molecular Biology of the Cell*, *19*(12), 5550–5558. https://doi.org/10.1091/mbc.e08-06-0645

Paulson, H. (2018). Repeat expansion diseases. *Handbook of Clinical Neurology*, *147*, 105–123. https://doi.org/10.1016/B978-0-444-63233-3.00009-9

Pavlov, Y. I., Frahm, C., Nick McElhinny, S. A., Niimi, A., Suzuki, M., & Kunkel, T. A. (2006). Evidence that errors made by DNA polymerase alpha are corrected by DNA polymerase delta. *Current Biology : CB*, *16*(2), 202–207. https://doi.org/10.1016/j.cub.2005.12.002

Peace, J. M., Ter-Zakarian, A., & Aparicio, O. M. (2014). Rif1 regulates initiation timing of late replication origins throughout the *S. cerevisiae* genome. *PloS One*, *9*(5), e98501. https://doi.org/10.1371/journal.pone.0098501

Peng, J. C., & Karpen, G. H. (2007). H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nature Cell Biology*, *9*(1), 25–35. https://doi.org/10.1038/ncb1514

Perez-Arnaiz, P., Bruck, I., & Kaplan, D. L. (2016). Mcm10 coordinates the timely assembly and activation of the replication fork helicase. *Nucleic Acids Research*, *44*(1), 315–329. https://doi.org/10.1093/nar/gkv1260

Perrino, F. W., & Loeb, L. A. (1990). Hydrolysis of 3'-terminal mispairs in vitro by the 3'----5' exonuclease of DNA polymerase delta permits subsequent extension by DNA polymerase alpha. *Biochemistry*, *29*(22), 5226–5231. https://doi.org/10.1021/bi00474a002

Petermann, E., Woodcock, M., & Helleday, T. (2010). Chk1 promotes replication fork progression by controlling replication initiation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(37), 16090–16095. https://doi.org/10.1073/pnas.1005031107

Petersen, J., & Russell, P. (2016). Growth and the Environment of *Schizosaccharomyces pombe*. *Cold Spring Harbor Protocols*, *2016*(3), pdb.top079764. https://doi.org/10.1101/pdb.top079764

Petryk, N., Kahli, M., D'Aubenton-Carafa, Y., Jaszczyszyn, Y., Shen, Y., Silvain, M., … Hyrien, O. (2016). Replication landscape of the human genome. *Nature Communications*, *7*, 10208. https://doi.org/10.1038/ncomms10208

Pidoux, A. L., & Allshire, R. C. (2004). Kinetochore and heterochromatin domains of the fission yeast centromere. *Chromosome Research : An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, *12*(6), 521–534. https://doi.org/10.1023/B:CHRO.0000036586.81775.8b

Podust, V. N., Chang, L.-S., Ott, R., Dianov, G. L., & Fanning, E. (2002). Reconstitution of human DNA polymerase delta using recombinant baculoviruses: the p12 subunit potentiates DNA polymerizing activity of the four-subunit enzyme. *The Journal of Biological Chemistry*, *277*(6), 3894–3901. https://doi.org/10.1074/jbc.M109684200

Pomerantz, R. T., & O'Donnell, M. (2010). Direct restart of a replication fork stalled by a head-on RNA polymerase. *Science (New York, N.Y.)*, *327*(5965), 590–592. https://doi.org/10.1126/science.1179595

Porrua, O., & Libri, D. (2013). A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nature Structural and Molecular Biology*, *20*(7), 884–891. https://doi.org/10.1038/nsmb.2592

Porrua, O., & Libri, D. (2015). Transcription termination and the control of the transcriptome: why, where and how to stop. *Nature Reviews. Molecular Cell Biology*, *16*(3), 190–202. https://doi.org/10.1038/nrm3943

Prado, F., & Aguilera, A. (2005). Impairment of replication fork progression mediates RNA polII transcription-associated recombination. *The EMBO Journal*, *24*(6), 1267–1276. https://doi.org/10.1038/sj.emboj.7600602

Převorovský, M., Oravcová, M., Zach, R., Jordáková, A., Bähler, J., Půta, F., & Folk, P. (2016). CSL protein regulates transcription of genes required to prevent catastrophic mitosis in fission yeast. *Cell Cycle (Georgetown, Tex.)*, *15*(22), 3082–3093. https://doi.org/10.1080/15384101.2016.1235100

Prindle, M. J., & Loeb, L. A. (2012). DNA polymerase delta in DNA replication and genome maintenance. *Environmental and Molecular Mutagenesis*, *53*(9), 666–682. https://doi.org/10.1002/em.21745

Quan, Y., Xia, Y., Liu, L., Cui, J., Li, Z., Cao, Q., … Lou, H. (2015). Cell-Cycle-Regulated Interaction between Mcm10 and Double Hexameric Mcm2-7 Is Required for Helicase Splitting and Activation during S Phase. *Cell Reports*, *13*(11), 2576–2586. https://doi.org/10.1016/j.celrep.2015.11.018

Quinet, A., Lemaçon, D., & Vindigni, A. (2017). Replication Fork Reversal: Players and Guardians. *Molecular Cell*, *68*(5), 830–833. https://doi.org/10.1016/j.molcel.2017.11.022

Ragunathan, K., Jih, G., & Moazed, D. (2015). Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science (New York, N.Y.)*, *348*(6230), 1258699. https://doi.org/10.1126/science.1258699

Randell, J. C. W., Bowers, J. L., Rodríguez, H. K., & Bell, S. P. (2006). Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Molecular Cell*, *21*(1), 29–39. https://doi.org/10.1016/j.molcel.2005.11.023

Randell, J. C. W., Fan, A., Chan, C., Francis, L. I., Heller, R. C., Galani, K., & Bell, S. P. (2010). Mec1 is one of multiple kinases that prime the Mcm2-7 helicase for phosphorylation by Cdc7. *Molecular Cell*, *40*(3), 353–363. https://doi.org/10.1016/j.molcel.2010.10.017

Razidlo, D. F., & Lahue, R. S. (2008). Mrc1, Tof1 and Csm3 inhibit CAG.CTG repeat instability by at least two mechanisms. *DNA Repair*, *7*(4), 633–640. https://doi.org/10.1016/j.dnarep.2008.01.009

Recolin, B., van der Laan, S., Tsanov, N., & Maiorano, D. (2014). Molecular mechanisms of DNA replication checkpoint activation. *Genes*, *5*(1), 147–175. https://doi.org/10.3390/genes5010147

Reijns, M. A. M., Kemp, H., Ding, J., de Procé, S. M., Jackson, A. P., & Taylor, M. S. (2015). Lagging-strand replication shapes the mutational landscape of the genome. *Nature*, *518*(7540), 502–506. https://doi.org/10.1038/nature14183

Remus, D., Blanchette, M., Rio, D. C., & Botchan, M. R. (2005). CDK phosphorylation inhibits the DNA-binding and ATP-hydrolysis activities of the Drosophila origin recognition complex. *The Journal of Biological Chemistry*, *280*(48), 39740–39751. https://doi.org/10.1074/jbc.M508515200

Reynolds, N., Watt, A., Fantes, P. A., & MacNeill, S. A. (1998). Cdm1, the smallest subunit of DNA polymerase d in the fission yeast *Schizosaccharomyces pombe*, is non-essential for growth and division. *Current Genetics*, *34*(4), 250–258. Retrieved from https://link.springer.com/content/pdf/10.1007/s002940050394.pdf

Rivera-Mulia, J. C., Buckley, Q., Sasaki, T., Zimmerman, J., Didier, R. A., Nazor, K., … Gilbert, D. M. (2015). Dynamic changes in replication timing and gene expression during lineage specification of human pluripotent stem cells. *Genome Research*, *25*(8), 1091–1103. https://doi.org/10.1101/gr.187989.114

Rivera-Mulia, J. C., Desprat, R., Trevilla-Garcia, C., Cornacchia, D., Schwerer, H., Sasaki, T., … Gilbert, D. M. (2017). DNA replication timing alterations identify common markers between distinct progeroid diseases. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(51), E10972–E10980. https://doi.org/10.1073/pnas.1711613114

Rivosecchi, J., Larochelle, M., Teste, C., Grenier, F., Malapert, A., Ricci, E. P., … Vanoosthuyse, V. (2019). Senataxin homologue Sen1 is required for efficient termination of RNA polymerase III transcription. *The EMBO Journal*, *38*(16), e101955. https://doi.org/10.15252/embj.2019101955

Rondón, A. G., Mischo, H. E., Kawauchi, J., & Proudfoot, N. J. (2009). Fail-safe transcriptional termination for protein-coding genes in *S. cerevisiae*. *Molecular Cell*, *36*(1), 88–98. https://doi.org/10.1016/j.molcel.2009.07.028

Rossi, S. E., Foiani, M., & Giannattasio, M. (2018). Dna2 processes behind the fork long ssDNA flaps generated by Pif1 and replication-dependent strand displacement. *Nature Communications*, *9*(1), 4830. https://doi.org/10.1038/s41467-018-07378-5

Rudolph, C. J., Dhillon, P., Moore, T., & Lloyd, R. G. (2007). Avoiding and resolving conflicts between DNA replication and transcription. *DNA Repair*, *6*(7), 981–993. https://doi.org/10.1016/j.dnarep.2007.02.017

Rustici, G., Mata, J., Kivinen, K., Lió, P., Penkett, C. J., Burns, G., … Bähler, J. (2004). Periodic gene expression program of the fission yeast cell cycle. *Nature Genetics*, *36*(8), 809–817. https://doi.org/10.1038/ng1377

Ryba, T., Battaglia, D., Chang, B. H., Shirley, J. W., Buckley, Q., Pope, B. D., … Gilbert, D. M. (2012). Abnormal developmental control of replication-timing domains in pediatric acute lymphoblastic leukemia. *Genome Research*, *22*(10), 1833–1844. https://doi.org/10.1101/gr.138511.112

Samadashwily, G. M., Raca, G., & Mirkin, S. M. (1997). Trinucleotide repeats affect DNA replication in vivo. *Nature Genetics*, *17*(3), 298–304. https://doi.org/10.1038/ng1197-298

Samel, S. A., Fernández-Cid, A., Sun, J., Riera, A., Tognetti, S., Herrera, M. C., … Speck, C. (2014). A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2-7 onto DNA. *Genes & Development*, *28*(15), 1653–1666. https://doi.org/10.1101/gad.242404.114

Sanchez, Y., Wong, C., Thoma, R. S., Richman, R., Wu, Z., Piwnica-Worms, H., & Elledge, S. J. (1997). Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25. *Science (New York, N.Y.)*, *277*(5331), 1497–1501. https://doi.org/10.1126/science.277.5331.1497

Santocanale, C, & Diffley, J. F. (1998). A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. *Nature*, *395*(6702), 615–618. https://doi.org/10.1038/27001

Santocanale, Corrado, Sharma, K., & Diffley, J. F. X. (1999). Activation of dormant origins of DNA replication in budding yeast. *Genes & Development*, *13*(18), 2360–2364. https://doi.org/10.1101/gad.13.18.2360

Sasaki, T., Rivera-Mulia, J. C., Vera, D., Zimmerman, J., Das, S., Padget, M., … Gilbert, D. M. (2017). Stability of patient-specific features of altered DNA replication timing in xenografts of primary human acute lymphoblastic leukemia. *Experimental Hematology*, *51*, 71-82.e3. https://doi.org/10.1016/j.exphem.2017.04.004

Schalch, T., Job, G., Shanker, S., Partridge, J. F., & Joshua-Tor, L. (2011). The Chp1-Tas3 core is a multifunctional platform critical for gene silencing by RITS. *Nature Structural & Molecular Biology*, *18*(12), 1351–1357. https://doi.org/10.1038/nsmb.2151

Schneider, C. a, Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Selak, N., Bachrati, C. Z., Shevelev, I., Dietschy, T., van Loon, B., Jacob, A., … Stagljar, I. (2008). The Bloom's syndrome helicase (BLM) interacts physically and functionally with p12, the smallest subunit of human DNA polymerase delta. *Nucleic Acids Research*, *36*(16), 5166–5179. https://doi.org/10.1093/nar/gkn498

Shechter, D., Costanzo, V., & Gautier, J. (2004). ATR and ATM regulate the timing of DNA replication origin firing. *Nature Cell Biology*, *6*(7), 648–655. https://doi.org/10.1038/ncb1145

Shell, S. M., Li, Z., Shkriabai, N., Kvaratskhelia, M., Brosey, C., Serrano, M. A., … Zou, Y. (2009). Checkpoint kinase ATR promotes nucleotide excision repair of UV-induced DNA damage via physical interaction with xeroderma pigmentosum group A. *The Journal of Biological Chemistry*, *284*(36), 24213–24222. https://doi.org/10.1074/jbc.M109.000745

Sheu, Y.-J., & Stillman, B. (2010). The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature*, *463*(7277), 113–117. https://doi.org/10.1038/nature08647

Shima, N., & Pederson, K. D. (2017). Dormant origins as a built-in safeguard in eukaryotic DNA replication against genome instability and disease development. *DNA Repair*, *56*(June), 166–173. https://doi.org/10.1016/j.dnarep.2017.06.019

Shoaib, M., Walter, D., Gillespie, P. J., Izard, F., Fahrenkrog, B., Lleres, D., … Sørensen, C. S. (2018). Histone H4K20 methylation mediated chromatin compaction threshold ensures genome integrity by limiting DNA replication licensing. *Nature Communications*, *9*(1), 3704. https://doi.org/10.1038/s41467-018-06066-8

Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., & Peterson, C. L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science (New York, N.Y.)*, *311*(5762), 844–847. https://doi.org/10.1126/science.1124000

Sienski, G., Dönertas, D., & Brennecke, J. (2012). Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell*, *151*(5), 964–980. https://doi.org/10.1016/j.cell.2012.10.040

Simon, A. C., Zhou, J. C., Perera, R. L., van Deursen, F., Evrin, C., Ivanova, M. E., … Pellegrini, L. (2014). A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. *Nature*, *510*(7504), 293–297. https://doi.org/10.1038/nature13234

Skourti-Stathaki, K., Proudfoot, N. J., & Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Molecular Cell*, *42*(6), 794–805. https://doi.org/10.1016/j.molcel.2011.04.026

Sriramachandran, A. M., Petrosino, G., Méndez-Lago, M., Schäfer, A. J., Batista-Nascimento, L. S., Zilio, N., & Ulrich, H. D. (2020). Genome-wide Nucleotide-Resolution Mapping of DNA Replication Patterns, Single-Strand Breaks, and Lesions by GLOE-Seq. *Molecular Cell*, *78*(5), 975-985.e7. https://doi.org/10.1016/j.molcel.2020.03.027

Steinmetz, E. J., Conrad, N. K., Brow, D. A., & Corden, J. L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature*, *413*(6853), 327–331. https://doi.org/10.1038/35095090

Steinmetz, E. J., Warren, C. L., Kuehner, J. N., Panbehi, B., Ansari, A. Z., & Brow, D. A. (2006). Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Molecular Cell*, *24*(5), 735–746. https://doi.org/10.1016/j.molcel.2006.10.023

Stirling, P. C., Chan, Y. A., Minaker, S. W., Aristizabal, M. J., Barrett, I., Sipahimalani, P., … Hieter, P. (2012). R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. *Genes & Development*, *26*(2), 163–175. https://doi.org/10.1101/gad.179721.111

Stodola, J. L., & Burgers, P. M. (2016). Resolving individual steps of Okazaki-fragment maturation at a millisecond timescale. *Nature Structural & Molecular Biology*, *23*(5), 402–408. https://doi.org/10.1038/nsmb.3207

Stodola, J. L., & Burgers, P. M. (2017). Mechanism of Lagging-Strand DNA Replication in Eukaryotes. *Advances in Experimental Medicine and Biology*, *1042*, 117–133. https://doi.org/10.1007/978-981-10-6955-0_6

Stroik, S., & Hendrickson, E. A. (2020). Telomere replication—When the going gets tough. *DNA Repair*, *94*(May), 102875. https://doi.org/10.1016/j.dnarep.2020.102875

Sugimoto, N., Maehara, K., Yoshida, K., Ohkawa, Y., & Fujita, M. (2018). Genome-wide analysis of the spatiotemporal regulation of firing and dormant replication origins in human cells. *Nucleic Acids Research*, *46*(13), 6683–6696. https://doi.org/10.1093/nar/gky476

Sugiyama, T., Cam, H., Verdel, A., Moazed, D., & Grewal, S. I. S. (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(1), 152–157. https://doi.org/10.1073/pnas.0407641102

Suraweera, A., Becherel, O. J., Chen, P., Rundle, N., Woods, R., Nakamura, J., … Lavin, M. F. (2007). Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. *The Journal of Cell Biology*, *177*(6), 969–979. https://doi.org/10.1083/jcb.200701042

Suraweera, A., Lim, Y., Woods, R., Birrell, G. W., Nasim, T., Becherel, O. J., & Lavin, M. F. (2009). Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. *Human Molecular Genetics*, *18*(18), 3384–3396. https://doi.org/10.1093/hmg/ddp278

Tada, S., Li, A., Maiorano, D., Méchali, M., & Blow, J. J. (2001). Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. *Nature Cell Biology*, *3*(2), 107–113. https://doi.org/10.1038/35055000

Takeda, T., Ogino, K., Matsui, E., Cho, M. K., Kumagai, H., Miyake, T., … Masai, H. (1999). A fission yeast gene, *him1*$^{+}$/*dfp1*$^{+}$, encoding a regulatory subunit for Hsk1 kinase, plays essential roles in S-phase initiation as well as in S-phase checkpoint control and recovery from DNA damage. *Molecular and Cellular Biology*, *19*(8), 5535–5547. https://doi.org/10.1128/MCB.19.8.5535

Tanaka, S., & Diffley, J. F. X. (2002). Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nature Cell Biology*, *4*(3), 198–207. https://doi.org/10.1038/ncb757

Tanaka, S., Nakato, R., Katou, Y., Shirahige, K., & Araki, H. (2011). Origin association of Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing. *Current Biology : CB*, *21*(24), 2055–2063. https://doi.org/10.1016/j.cub.2011.11.038

Taylor, J. H. (1977). Increase in DNA replication sites in cells held at the beginning of S phase. *Chromosoma*, *62*(4), 291–300. https://doi.org/10.1007/BF00327029

Taylor, M. R. G., & Yeeles, J. T. P. (2018). The Initial Response of a Eukaryotic Replisome to DNA Damage. *Molecular Cell*, *70*(6), 1067-1080.e12. https://doi.org/10.1016/j.molcel.2018.04.022

Terai, K., Shibata, E., Abbas, T., & Dutta, A. (2013). Degradation of p12 subunit by CRL4Cdt2 E3 ligase inhibits fork progression after DNA damage. *The Journal of Biological Chemistry*, *288*(42), 30509–30514. https://doi.org/10.1074/jbc.C113.505586

Thada, V., & Cortez, D. (2019). Common motifs in ETAA1 and TOPBP1 required for ATR kinase activation. *Journal of Biological Chemistry*, *294*(21), 8395–8402. https://doi.org/10.1074/jbc.RA119.008154

The Gene Ontology Consortium. (2019). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*, *47*(D1), D330–D338. https://doi.org/10.1093/nar/gky1055

Ticau, S., Friedman, L. J., Ivica, N. A., Gelles, J., & Bell, S. P. (2015). Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell*, *161*(3), 513–525. https://doi.org/10.1016/j.cell.2015.03.012

Tognetti, S., & Speck, C. (2016). Replicating repetitive DNA. *Nature Cell Biology*, *18*(6), 593–594. https://doi.org/10.1038/ncb3367

Tollervey, D. (2004). Molecular biology: termination by torpedo. *Nature*, *432*(7016), 456–457. https://doi.org/10.1038/432456a

Trenz, K., Smith, E., Smith, S., & Costanzo, V. (2006). ATM and ATR promote Mre11 dependent restart of collapsed replication forks and prevent accumulation of DNA breaks. *EMBO Journal*, *25*(8), 1764–1774. https://doi.org/10.1038/sj.emboj.7601045

Tsurimoto, T., & Stillman, B. (1989). Multiple replication factors augment DNA synthesis by the two eukaryotic DNA polymerases, alpha and delta. *The EMBO Journal*, *8*(12), 3883–3889. https://doi.org/10.1002/j.1460-2075.1989.tb08567.x

Tsutsui, Y., Morishita, T., Natsume, T., Yamashita, K., Iwasaki, H., Yamao, F., & Shinagawa, H. (2005). Genetic and physical interactions between *Schizosaccharomyces pombe* Mcl1 and Rad2, Dna2 and DNA polymerase alpha: evidence for a multifunctional role of Mcl1 in DNA replication and repair. *Current Genetics*, *48*(1), 34–43. https://doi.org/10.1007/s00294-005-0584-2

Ursic, D., Chinchilla, K., Finkel, J. S., & Culbertson, M. R. (2004). Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. *Nucleic Acids Research*, *32*(8), 2441–2452. https://doi.org/10.1093/nar/gkh561

Van, C., Yan, S., Michael, W. M., Waga, S., & Cimprich, K. A. (2010). Continued primer synthesis at stalled replication forks contributes to checkpoint activation. *The Journal of Cell Biology*, *189*(2), 233–246. https://doi.org/10.1083/jcb.200909105

van Deursen, F., Sengupta, S., De Piccoli, G., Sanchez-Diaz, A., & Labib, K. (2012). Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation. *The EMBO Journal*, *31*(9), 2195–2206. https://doi.org/10.1038/emboj.2012.69

Vasiljeva, L., & Buratowski, S. (2006). Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Molecular Cell*, *21*(2), 239–248. https://doi.org/10.1016/j.molcel.2005.11.028

Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., & Meinhart, A. (2008). The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nature Structural & Molecular Biology*, *15*(8), 795–804. https://doi.org/10.1038/nsmb.1468

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I. S., & Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. *Science (New York, N.Y.)*, *303*(5658), 672–676. https://doi.org/10.1126/science.1093686

Vesela, E., Chroma, K., Turi, Z., & Mistrik, M. (2017). Common Chemical Inductors of Replication Stress: Focus on Cell-Based Studies. *Biomolecules*, *7*(1), 19. https://doi.org/10.3390/biom7010019

Villa, F., Simon, A. C., Ortiz Bazan, M. A., Kilkenny, M. L., Wirthensohn, D., Wightman, M., … Labib, K. (2016). Ctf4 Is a Hub in the Eukaryotic Replisome that Links Multiple CIP-Box Proteins to the CMG Helicase. *Molecular Cell*, *63*(3), 385–396. https://doi.org/10.1016/j.molcel.2016.06.009

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B. J., & Grunstein, M. (2002). Histone acetylation regulates the time of replication origin firing. *Molecular Cell*, *10*(5), 1223–1233. https://doi.org/10.1016/s1097-2765(02)00702-5

Voineagu, I., Surka, C. F., Shishkin, A. A., Krasilnikova, M. M., & Mirkin, S. M. (2009). Replisome stalling and stabilization at CGG repeats, which are responsible for chromosomal fragility. *Nature Structural & Molecular Biology*, *16*(2), 226–228. https://doi.org/10.1038/nsmb.1527

Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. S., & Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science (New York, N.Y.)*, *297*(5588), 1833–1837. https://doi.org/10.1126/science.1074973

Walter, D., Hoffmann, S., Komseli, E.-S., Rappsilber, J., Gorgoulis, V., & Sørensen, C. S. (2016). SCF(Cyclin F)-dependent degradation of CDC6 suppresses DNA re-replication. *Nature Communications*, *7*, 10530. https://doi.org/10.1038/ncomms10530

Wang, D., & Gao, F. (2019). Comprehensive Analysis of Replication Origins in *Saccharomyces cerevisiae* Genomes. *Frontiers in Microbiology*, *10*(September), 2122. https://doi.org/10.3389/fmicb.2019.02122

Wang, G., & Vasquez, K. M. (2014). Impact of alternative DNA structures on DNA damage, DNA repair, and genetic instability. *DNA Repair*, *19*, 143–151. https://doi.org/10.1016/j.dnarep.2014.03.017

Wang, J., Gong, Z., & Chen, J. (2011). MDC1 collaborates with TopBP1 in DNA replication checkpoint control. *The Journal of Cell Biology*, *193*(2), 267–273. https://doi.org/10.1083/jcb.201010026

Wang, S., Han, Z., Libri, D., Porrua, O., & Strick, T. R. (2019). Single-molecule characterization of extrinsic transcription termination by Sen1 helicase. *Nature Communications*, *10*(1), 1545. https://doi.org/10.1038/s41467-019-09560-9

Wang, W., Klein, K. N., Proesmans, K., Yang, H., Marchal, C., Zhu, X., … Rhind, N. (2021). Genome-wide mapping of human DNA replication by optical replication mapping supports a stochastic model of eukaryotic replication. *Molecular Cell*, 1–14. https://doi.org/10.1016/j.molcel.2021.05.024

Wansink, D. G., Manders, E. E. M., van der Kraan, I., Aten, J. A., van Driel, R., & de Jong, L. (1994). RNA polymerase II transcription is concentrated outside replication domains throughout S-phase. *Journal of Cell Science*, *107 (Pt 6*(6), 1449–1456. https://doi.org/10.1242/jcs.107.6.1449

Ward, I. M., & Chen, J. (2001). Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. *The Journal of Biological Chemistry*, *276*(51), 47759–47762. https://doi.org/10.1074/jbc.C100569200

Ward, I. M., Minn, K., & Chen, J. (2004). UV-induced ataxia-telangiectasia-mutated and Rad3-related (ATR) activation requires replication stress. *The Journal of Biological Chemistry*, *279*(11), 9677–9680. https://doi.org/10.1074/jbc.C300554200

Wasserman, M. R., Schauer, G. D., O'Donnell, M. E., & Liu, S. (2019). Replication Fork Activation Is Enabled by a Single-Stranded DNA Gate in CMG Helicase. *Cell*, *178*(3), 600-611.e16. https://doi.org/10.1016/j.cell.2019.06.032

Watson, A. T., Garcia, V., Bone, N., Carr, A. M., & Armstrong, J. (2008). Gene tagging and gene replacement using recombinase-mediated cassette exchange in *Schizosaccharomyces pombe*. *Gene*, *407*(1–2), 63–74. https://doi.org/10.1016/j.gene.2007.09.024

WATSON, J. D., & CRICK, F. H. (1953). Genetical implications of the structure of deoxyribonucleic acid. *Nature*, *171*(4361), 964–967. https://doi.org/10.1038/171964b0

Wei, X., Samarabandu, J., Devdhar, R. S., Siegel, A. J., Acharya, R., & Berezney, R. (1998). Segregation of transcription and replication sites into higher order domains. *Science (New York, N.Y.)*, *281*(5382), 1502–1506. https://doi.org/10.1126/science.281.5382.1502

Wery, M., Gautier, C., Descrimes, M., Yoda, M., Vennin-Rendos, H., Migeot, V., … Morillon, A. (2018). Native elongating transcript sequencing reveals global anti-correlation between sense and antisense nascent transcription in fission yeast. *RNA (New York, N.Y.)*, *24*(2), 196–208. https://doi.org/10.1261/rna.063446.117

Wilmes, G. M., Archambault, V., Austin, R. J., Jacobson, M. D., Bell, S. P., & Cross, F. R. (2004). Interaction of the S-phase cyclin Clb5 with an "RXL" docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. *Genes & Development*, *18*(9), 981–991. https://doi.org/10.1101/gad.1202304

Wilson, K. A., Elefanty, A. G., Stanley, E. G., & Gilbert, D. M. (2016). Spatio-temporal re-organization of replication foci accompanies replication domain consolidation during human pluripotent stem cell lineage specification. *Cell Cycle (Georgetown, Tex.)*, *15*(18), 2464–2475. https://doi.org/10.1080/15384101.2016.1203492

Wirén, M., Silverstein, R. A., Sinha, I., Walfridsson, J., Lee, H. M., Laurenson, P., … Ekwall, K. (2005). Genomewide analysis of nucleosome density histone acetylation and HDAC function in fission yeast. *EMBO Journal*, *24*(16), 2906–2918. https://doi.org/10.1038/sj.emboj.7600758

Wlotzka, W., Kudla, G., Granneman, S., & Tollervey, D. (2011). The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *The EMBO Journal*, *30*(9), 1790–1803. https://doi.org/10.1038/emboj.2011.97

Wohlschlegel, J. A., Dwyer, B. T., Dhar, S. K., Cvetic, C., Walter, J. C., & Dutta, A. (2000). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science (New York, N.Y.)*, *290*(5500), 2309–2312. https://doi.org/10.1126/science.290.5500.2309

Wood, V., Gwilliam, R., Rajandream, M.-A., Lyne, M., Lyne, R., Stewart, A., … Cerrutti, L. (2002). The genome sequence of *Schizosaccharomyces pombe*. *Nature*, *415*(6874), 871–880. https://doi.org/10.1038/nature724

Xie, B., Mazloum, N., Liu, L., Rahmeh, A., Li, H., & Lee, M. Y. W. T. (2002). Reconstitution and characterization of the human DNA polymerase delta four-subunit holoenzyme. *Biochemistry*, *41*(44), 13133–13142. https://doi.org/10.1021/bi0262707

Yadav, T., Quivy, J.-P., & Almouzni, G. (2018). Chromatin plasticity: A versatile landscape that underlies cell fate and identity. *Science (New York, N.Y.)*, *361*(6409), 1332–1336. https://doi.org/10.1126/science.aat8950

Yamazaki, S., Ishii, A., Kanoh, Y., Oda, M., Nishito, Y., & Masai, H. (2012). Rif1 regulates the replication timing domains on the human genome. *The EMBO Journal*, *31*(18), 3667–3677. https://doi.org/10.1038/emboj.2012.180

Yang, W., & Gao, Y. (2018). Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. *Annual Review of Biochemistry*, *87*, 239–261. https://doi.org/10.1146/annurev-biochem-062917-012405

Yeeles, J. T. P., Janska, A., Early, A., & Diffley, J. F. X. (2017). How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. *Molecular Cell*, *65*(1), 105–116. https://doi.org/10.1016/j.molcel.2016.11.017

Yeo, A. J., Becherel, O. J., Luff, J. E., Graham, M. E., Richard, D., & Lavin, M. F. (2015). Senataxin controls meiotic silencing through ATR activation and chromatin remodeling. *Cell Discovery*, *1*, 15025. https://doi.org/10.1038/celldisc.2015.25

Zach, R., & Carr, A. M. (2021). Increased expression of Polδ does not alter the canonical replication program in vivo. *Wellcome Open Research*, *6*(May), 44. https://doi.org/10.12688/wellcomeopenres.16600.1

Zegerman, P., & Diffley, J. F. X. (2010). Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. *Nature*, *467*(7314), 474–478. https://doi.org/10.1038/nature09373

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

Zeng, Y., Forbes, K. C., Wu, Z., Moreno, S., Piwnica-Worms, H., & Enoch, T. (1998). Replication checkpoint requires phosphorylation of the phosphatase Cdc25 by Cds1 or Chk1. *Nature*, *395*(6701), 507–510. https://doi.org/10.1038/26766

Zhang, H., & Freudenreich, C. H. (2007). An AT-Rich Sequence in Human Common Fragile Site FRA16D Causes Fork Stalling and Chromosome Breakage in *S. cerevisiae*. *Molecular Cell*, *27*(3), 367–379. https://doi.org/10.1016/j.molcel.2007.06.012

Zhang, Jian, Tan, C. K., McMullen, B., Downey, K. M., & So, A. G. (1995). Cloning of the cDNAs for the small subunits of bovine and human DNA polymerase delta and chromosomal location of the human gene (POLD2). *Genomics*, *29*(1), 179–186. https://doi.org/10.1006/geno.1995.1229

Zhang, Jieqiong, Dewar, J. M., Budzowska, M., Motnenko, A., Cohn, M. A., & Walter, J. C. (2015). DNA interstrand cross-link repair requires replication-fork convergence. *Nature Structural & Molecular Biology*, *22*(3), 242–247. https://doi.org/10.1038/nsmb.2956

Zhang, Jing, Bellani, M. A., James, R. C., Pokharel, D., Zhang, Y., Reynolds, J. J., … Seidman, M. M. (2020). DONSON and FANCM associate with different replisomes distinguished by replication timing and chromatin domain. *Nature Communications*, *11*(1), 3951. https://doi.org/10.1038/s41467-020-17449-1

Zhang, K., Mosch, K., Fischle, W., & Grewal, S. I. S. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nature Structural & Molecular Biology*, *15*(4), 381–388. https://doi.org/10.1038/nsmb.1406

Zhang, S., Chao, H. H., Wang, X., Zhang, Z., Lee, E. Y. C., & Lee, M. Y. W. T. (2019). Loss of the p12 subunit of DNA polymerase delta leads to a defect in HR and sensitization to PARP inhibitors. *DNA Repair*, *73*(October 2018), 64–70. https://doi.org/10.1016/j.dnarep.2018.11.003

Zhang, S., Zhao, H., Darzynkiewicz, Z., Zhou, P., Zhang, Z., Lee, E. Y. C., & Lee, M. Y. W. T. (2013). A novel function of CRL4(Cdt2): regulation of the subunit structure of DNA polymerase δ in response to DNA damage and during the S phase. *The Journal of Biological Chemistry*, *288*(41), 29550–29561. https://doi.org/10.1074/jbc.M113.490466

Zhang, S., Zhou, Y., Trusa, S., Meng, X., Lee, E. Y. C., & Lee, M. Y. W. T. (2007). A novel DNA damage response: rapid degradation of the p12 subunit of dna polymerase delta. *The Journal of Biological Chemistry*, *282*(21), 15330–15340. https://doi.org/10.1074/jbc.M610356200

Zhao, Hong, Zhang, S., Xu, D., Lee, M. Y., Zhang, Z., Lee, E. Y. C., & Darzynkiewicz, Z. (2014). Expression of the p12 subunit of human DNA polymerase δ (Pol δ), CDK inhibitor p21(WAF1), Cdt1, cyclin A, PCNA and Ki-67 in relation to DNA replication in individual cells. *Cell Cycle (Georgetown, Tex.)*, *13*(22), 3529–3540. https://doi.org/10.4161/15384101.2014.958910

Zhao, Hui, & Piwnica-Worms, H. (2001). ATR-mediated checkpoint pathways regulate phosphorylation and activation of human Chk1. *Molecular and Cellular Biology*, *21*(13), 4129–4139. https://doi.org/10.1128/MCB.21.13.4129-4139.2001

Zhao, Hui, Watkins, J. L., & Piwnica-Worms, H. (2002). Disruption of the checkpoint kinase 1/cell division cycle 25A pathway abrogates ionizing radiation-induced S and G2 checkpoints. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(23), 14795–14800. https://doi.org/10.1073/pnas.182557299

Zhao, Q., Kirkness, E. F., Caballero, O. L., Galante, P. A., Parmigiani, R. B., Edsall, L., … Strausberg, R. L. (2010). Systematic detection of putative tumor suppressor genes through the combined use of exome and transcriptome sequencing. *Genome Biology*, *11*(11), R114. https://doi.org/10.1186/gb-2010-11-11-r114

Zhou, J. C., Janska, A., Goswami, P., Renault, L., Abid Ali, F., Kotecha, A., … Costa, A. (2017). CMG–Pol epsilon dynamics suggests a mechanism for the establishment of leading-strand synthesis in the eukaryotic replisome. *Proceedings of the National Academy of Sciences*, *114*(16), 4141–4146. https://doi.org/10.1073/pnas.1700530114

Zhou, Z.-X., Lujan, S. A., Burkholder, A. B., Garbacz, M. A., & Kunkel, T. A. (2019). Roles for DNA polymerase δ in initiating and terminating leading strand DNA replication. *Nature Communications*, *10*(1), 3992. https://doi.org/10.1038/s41467-019-11995-z

Zofall, M, & Grewal, S. I. S. (2006). RNAi-mediated heterochromatin assembly in fission yeast. *Cold Spring Harbor Symposia on Quantitative Biology*, *71*, 487–496. https://doi.org/10.1101/sqb.2006.71.059

Zofall, Martin, Yamanaka, S., Reyes-Turcu, F. E., Zhang, K., Rubin, C., & Grewal, S. I. S. (2012). RNA elimination machinery targeting meiotic mRNAs promotes facultative heterochromatin formation. *Science (New York, N.Y.)*, *335*(6064), 96–100. https://doi.org/10.1126/science.1211651

Zou, L. (2002). Regulation of ATR substrate selection by Rad17-dependent loading of Rad9 complexes onto chromatin. *Genes & Development*, *16*(2), 198–208. https://doi.org/10.1101/gad.950302

Zou, L., & Elledge, S. J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science (New York, N.Y.)*, *300*(5625), 1542–1548. https://doi.org/10.1126/science.1083430

Żylicz, J. J., & Heard, E. (2020). Molecular Mechanisms of Facultative Heterochromatin Formation: An X-Chromosome Perspective. *Annual Review of Biochemistry*, *89*, 255–282. https://doi.org/10.1146/annurev-biochem-062917-012655