RECQ-LIKE HELICASE *SGS1* COUNTERACTS DNA:RNA HYBRID INDUCED GENOME INSTABILITY

by

Carolina Novoa

B.Sc., The University of British Columbia, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Genome Science and Technology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2016

Abstract

Dividing cells are constantly under threat from both endogenous and exogenous DNA damaging stresses that can lead to mutations and structural variations in DNA. One contributor to genome instability is three-stranded DNA:RNA hybrid structures called R-loops. Though R-loops are known to induce DNA damage and DNA replication stress, it is unclear whether they are recognized and processed by an established DNA repair pathway prior to inducing DNA breaks. Canonically, DNA repair proteins work downstream of R-loop-induced DNA damage to stimulate repair and suppress genome instability. Recently, the possibility that some DNA repair pathways actively destabilize R-loops, thus preventing unscheduled DNA damage has emerged. Here we identify the helicase *SGS1* as a suppressor of R-loop stability. Our data reveals that *SGS1* depleted cells accumulate R-loops. In addition, we define a role for transcription in genome instability of cells lacking *SGS1*, which is consistent with an R-loop based mechanism. Hyper-recombination in *SGS1* mutants is dependent on transcript length, transcription rate, and active DNA replication. Also, rDNA instability in *sgs1*Δ can be suppressed by ectopic expression of RNaseH1, a protein that degrades DNA:RNA hybrids. Interestingly, R-loops are known to form at rDNA loci. We favour a model in which *SGS1* contributes to the stabilization of stalled replication forks associated with transcription complexes, and unresolved DNA:RNA hybrids. Finally, we showed that knockdown of the human Sgs1 orthologue BLM in HCT116 cells also led to the accumulation of more R-loops than control HCT116 cells. In summary, our data supports the idea that some DNA repair proteins involved in replication fork stabilization might also prevent and process R-loops.

Preface

This thesis is the work of Carolina Novoa. She and Dr. Peter C. Stirling were responsible for the conception of this project. The experimental work and analysis of the data were carried out by Carolina Novoa except for the DRIP-ChIP experiments and analysis, which were performed by Dr. Maria Aristizabal from the laboratory of Dr. Michael Kobor (**Figure 15**). Christelle Keong also contributed to this work by performing yeast growth curves (**Figure 4**).

Table of contents

List of tables

List of figures

List of abbreviations

ALF: A-like faker

BSA: Bovine serum albumin

CIN: Chromosome instability

DIC: Differential interference contrast filter

D-loop: Displacement loop

DRIP: DNA:RNA imunoprecipitation

DSB: DNA double-strand break

FA: Fanconi anemia

Gal: Galactose

GI: Genome instability

HR: Homologous recombination

HU: Hydroxyurea

LEU: Leucine

MMS: Methyl methanesulfonate

NER: Nucleotide excision repair

PBS: Phosphate buffered saline

rDNA: Ribosomal DNA

RF: Replication fork

SC: Synthetic complete yeast medium

ssDNA: Single stranded DNA

ts: temperature sensitive

URA: Uracil

YFP: Yellow fluorescent protein

YPD: Yeast peptone-dextrose medium

5-FOA: 5-Fluoroorotic

Acknowledgements

I would like to thank my supervisor Dr. Peter Stirling for being a supportive mentor and an inspiring scientist. Peter's enthusiasm for research and yeast genetics in particular has created a nurturing and stimulating environment in the lab, which has kept me engaged, excited and accountable for my work. I am also sincerely grateful to Peter for allowing me to work in a challenging and interesting project in the growing transcription-induced genome instability field. I am extremely grateful to have had the opportunity to be part of the Stirling lab!

I would also like to thank my committee members for pushing me to consider the most important research questions of my field of study and their relevance to my project. This helped me better understand the limitations of the approaches I used, and the significance of my research project. I also want to thank my committee for supporting the scientific, professional and personal decisions I have made as a graduate student.

Finally, I would like to thank my friends and lab mates for their sincere advise, company and endless kindness. You have been my main support network during my time as a graduate student, and I can only hope that the connections we have made during the last three and a half years will keep growing.

Chapter 1: INTRODUCTION

1.1 Overview and literature summary

R-loops and Genome Instability

Genome instability (GI) refers to the increase in mutation rate and genomic structural variants in a cell population over time (Pikor, Thu, Vucic, & Lam, 2013). This process is a main contributing factor to cancer development because it allows mutations to occur at a sufficiently high frequency for cells to acquire the genetic changes required for tumorigenesis, and tumor progression and evolution (Stratton, Campbell, & Futreal, 2009). GI is also a distinguishing feature of many hereditary and sporadic malignancies (Negrini, Gorgoulis, & Halazonetis, 2010). In addition, GI has been exploited by chemo- and radio-therapies which use genotoxicity to push cancer cells with impaired genome maintenance to lethality. Nevertheless, details of the many pathways that can regulate genome stability and their roles in cancer genome integrity are not fully understood (Paulsen et al., 2009; Stirling et al., 2011). Further investigation of GI mechanisms will likely contribute to the development of targeted agents directed at unique molecular features of cancer cells, eventually improving upon the effects of non-specific cytotoxic drugs.

Previously understudied sources of GI include DNA:RNA hybrid containing R-loops. These structures come about when newly made transcripts re-anneal to the DNA template strand displacing the non-template strand (**Figure 1**). R-loops can arise during normal transcription even though their unscheduled formation and accumulation is potentially harmful and genotoxic. Furthermore, R-loop formation can occur in homeostasis with other cellular processes. R-loop occupancy has been detected and profiled in yeast and human wild type genomes, where they presumably are not causing genome instability (Chan, 2014; El Hage, Webb, Kerr, & Tollervey,

2014; Ginno, Lott, Christensen, Korf, & Chedin, 2012). In addition, R-loops have been shown to have a normal role in the regulation of gene expression through several mechanisms including facilitation of PolII pausing and termination (Skourti-Stathaki, Proudfoot, & Gromak, 2011), modulation of epigenetic marks (Ginno et al., 2012) and blockage of transcription promoters (Sun, Csorba, Skourti-Stathaki, Proudfoot, & Dean, 2013). Moreover, R-loop formation contributes to the diverse antibody repertoire produced by activated B-cells. Specifically, R-loops provide the enzyme Activation-Induced Cytidine Deaminase (AID) with the ssDNA substrate required for class switch recombination at the immunoglobin heavy chain (Costantino & Koshland, 2015).

In contrast with the previous examples, R-loop formation can also be detrimental to cells. R-loops have been associated with GI by several groups showing that mutants which accumulate DNA:RNA hybrids also present increased chromosome instability (CIN), recombination, and replication defects (Costantino & Koshland, 2015; Sollier & Cimprich, 2015; Stirling et al., 2012a; Wahba, Amon, Koshland, & Vuica-Ross, 2011). The dichotomised effects of R-loops continue to present a challenge in the characterization of the GI they are believed to induce. So far, it has been suggested that R-loops lead to GI by two major mechanisms:

Induction of single and double strand breaks: DNA damage has been associated with the exposure of single stranded DNA (ssDNA) in the R-loop to environmental threats and DNA processing enzymes (Aguilera & Garcia-Muse, 2012; Gan et al., 2011; Huertas & Aguilera, 2003). In this model, ssDNA is susceptible to lesions and transcription-associated mutagenesis or transcription-associated recombination. Also, Sollier *et al.* demonstrated in human cells that R-loops are actively processed into DNA double stand breaks (DSBs) by the nucleotide excision

repair (NER) endonucleases XPF and XPG (Sollier et al., 2014). Therefore, it is possible that DSBs result from the direct processing of R-loops.

Impairment of DNA replication: R-loops can encounter a replication fork head-on or co-directionally and impede its progression, causing stalling of DNA polymerase, subsequent impaired replication and DSBs. Although both types of collisions can disrupt or arrest replication forks (Brambati, Colosio, Zardoni, Galanti, & Liberi, 2015), several lines of evidence indicate that head-on replication-transcription collisions are more detrimental to fork and genome stability (Kim, Abdulovic, Gealy, Lippert, & Jinks-Robertson, 2007; Prado & Aguilera, 2005).

Eukaryotic cells have mechanisms that respond to replication-transcription conflicts. However, these mechanisms can be error prone thus making obstruction of replication via R-loop accumulation a threat to genome stability. Stalled-fork restarting mechanisms include uncoupling of leading and lagging strand synthesis and replication across the non-transcribed strand (Zeman & Cimprich, 2014). Those pathways might require template switching and/or trans-lesion DNA synthesis and/or fork reversal for DNA replication across DNA:RNA hybrids (Gomez-Gonzalez, Felipe-Abrio, & Aguilera, 2009; Zeman & Cimprich, 2014), which can lead to single and double strand breaks, recombination events, and genome instability. DNA breaks result from the action of nuclease digestion of reversed or stalled forks, or accumulation of ssDNA gaps. Persistent or unrepaired replication-transcription collisions will also lead to forks collapse due to the action of DNA cutting enzymes (Zeman & Cimprich, 2014). Alternatively unreplicated DNA regions will lead to dsDNA breaks during cell division. In summary, R-loop persistence can cause severe damage to the DNA through impairment of DNA replication.

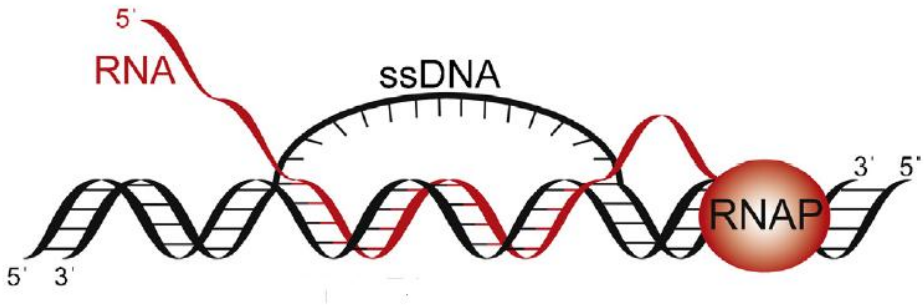

Figure 1 R-loop structure. This excerpt from (Hamperl & Cimprich, 2014)) illustrates the structure of a DNA:RNA hybrid. The RNA strand (red) is generated by RNA polymerase (RNAP) while the DNA:RNA hybrid arises by annealing back of the RNA to the template strand and displacing the non-template strand.

R-loop protective mechanisms

The best understood players in R-loop metabolism are those involved in RNA processing. Normal transcript elongation, termination, splicing, packaging, nuclear export and RNA degradation have all been shown to suppress R-loop formation (Gomez-Gonzalez et al., 2011; Li & Manley, 2005; Mischo et al., 2011; Stirling et al., 2012a; Wahba et al., 2011). In addition, there are factors that actively remove R-loops. At least five different mechanisms are thought to regulate R-loop formation (**Figure 2**):

Ribonuclease H (RNaseH) enzymes, which specifically degrade the RNA moiety of DNA:RNA hybrids are major and direct R-loop regulators (Cerritelli & Crouch, 2009). Eukaryotic cells contain two versions: RNaseH1 (RNH1) and RnaseH2 (RNH201 in yeast). Rnh1 functions as a monomer and can be found in the mitochondria and the nucleus while Rnh201 functions as part of a trimeric complex with regulatory subunits Rnh202 and Rnh203 (RNASEH2A, B and C in humans). The only known function of Rnh1 is to remove R-loops, therefore this protein has been widely used experimentally to test for the presence of DNA:RNA hybrids. Unlike RNaseH1, RNaseH2 can also remove RNA primers during Okazaki fragment synthesis, and errant ribonucleotides misincorporated during DNA replication (Cerritelli & Crouch, 2009).

DNA:RNA helicases, which unwind hybrids, are the second type of direct R-loop regulator. Examples of this category are Aquarius, DHX9 and Senataxin/*SEN1* (Mischo et al., 2011; Skourti-Stathaki et al., 2011; Sollier et al., 2014). A temperature sensitive (*ts*) *SEN1* mutant, which also has reduced helicase activity, accumulates R-loops. *SEN1* is also part of the NRD transcription termination complex (Sen1, Nrd1 and Nab3) where it facilitates termination of some genes and most noncoding RNAs. Moreover, *SEN1* genetically interacts with *SGS1* and *MUS81* (Mischo et al., 2011).

Topoisomerase I, which relaxes DNA negative supercoiling and torsional stress behind RNA Pol I and II has also been shown to reduce R-loops (El Hage, French, Beyer, & Tollervey, 2010; Tuduri et al., 2009; Yang et al., 2014). Interestingly, topoisomerase II has also recently been implicated in R-loop mitigation (Yeo et al., 2016), suggesting that regulators of DNA topological strain may generally counteract R-loop accumulation.

mRNA biogenesis and processing proteins that prevent R-loop formation, likely do so by binding to RNA as it emerges from RNA polymerases. For example, several proteins of the conserved elongation and export complexes THO/TREX, human SRSF1 splicing factor, and yeast mRNA export protein Npl3, have been implicated in R-loop prevention (Castellano-Pozo, Garcia-Muse, & Aguilera, 2012; Dominguez-Sanchez, Barroso, Gomez-Gonzalez, Luna, & Aguilera, 2011; Gomez-Gonzalez et al., 2011; Huertas & Aguilera, 2003; Li & Manley, 2005; Santos-Pereira et al., 2013; Stirling et al., 2012b; Wahba et al., 2011) .

The R-loop associated GI of THO/TREX complex is conserved from yeast to humans. In *S. cerevisiae*, the THO complex consists of four nuclear proteins Hrp1, Tho2, Mft1, and Thp2 and is associated with TREX, which contains Tex1, Sub2 and Yra1. Mutations affecting THO/TREX

have transcriptional elongation and mRNA export defects, and consequent genomic instability (DNA damage, recombination defects), which is directly associated with R-loop formation (Gavalda, Santos-Pereira, Garcia-Rubio, Luna, & Aguilera, 2016; Huertas & Aguilera, 2003). A distinctive phenotype of yeast THO mutants is their transcription-associated hyperrecombination phenotype.

Suppressors of proteins that promote R-loop formation. For example, Srs2 is a Rad51 regulating proteins that prevent R-loop formation. Rad51 is a DNA repair protein that has been proposed to facilitate the formation of R-loops *in trans* or away from the point of transcription. Deletion of Rad51 in budding yeast also results in reduced accumulation of R-loops. (Wahba & Koshland, 2013).

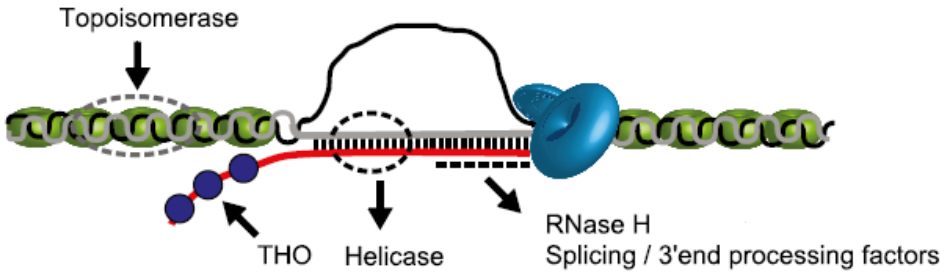

Figure 2 Known R-loop regulatory enzymes This schematic illustrates the mode of action of the best understood R-loop surveillance factors. Topoisomerases prevent DNA:RNA hybrid formation by reducing negative supercoiling behind RNA polymerase (light blue torus). THO complex factors (blue circles), which in yeast include *THP2* and *MFT1* bind to newly made RNA and prevent it from reannealing to the DNA duplex. Helicases such as yeast *SEN1* unwind the DNA:RNA portion of the R-loop. Finally, RNaseH enzymes degrade RNA, which hybridizes to the template DNA strand. Adapted from (Skourti-Stathaki & Proudfoot, 2014).

It is unknown if all or a specific set of R-loops are recognized as aberrant structures and processed by any of the known DNA repair pathways. However, several proteins of the Fanconi Anemia (FA) pathway including BRCA1 (Hatchi et al. 2015), BRCA2 (Bhatia et al. 2014), FANCM, FANCD2 and FANCA (Garcia-Rubio et al. 2015; Schwab et al. 2015) have been

implicated in the regulation of R-loops. In addition, other DNA repair proteins like XPF and XPG (Sollier et al., 2014) have been linked to R-loop processing. Finally, the DSB response protein kinase ATM has been shown to have a role in signaling at DSBs that involves R-loops (Britton et al., 2014; Tresini et al., 2015).

It is possible that other DNA repair proteins regulate R-loop induced GI. To explore this idea, our collaborators conducted a cytological screen using the DNA:RNA specific antibody S9.6 in 60 of the 143 DNA repair and replication deletion mutants with a CIN phenotype in yeast. They identified 18 genes with ≥ 2 fold increase in the percentage of cells containing S9.6 foci thus uncovering previously unidentified candidate R-loop suppressors (Chan, 2014) (**Figure 3**). *SGS1* stood out as one of the strongest hits in the screen. Several accepted R-loop surveillance factors were also detected demonstrating the robustness of the assay. For instance, *TOP1* has been previously implicated in R-loop suppression through its role in DNA topology (El Hage et al., 2010; Tuduri et al., 2009) and *MPH1* is the orthologue of *FANCM* which was recently implicated in R-loop removal (Schwab et al., 2015).

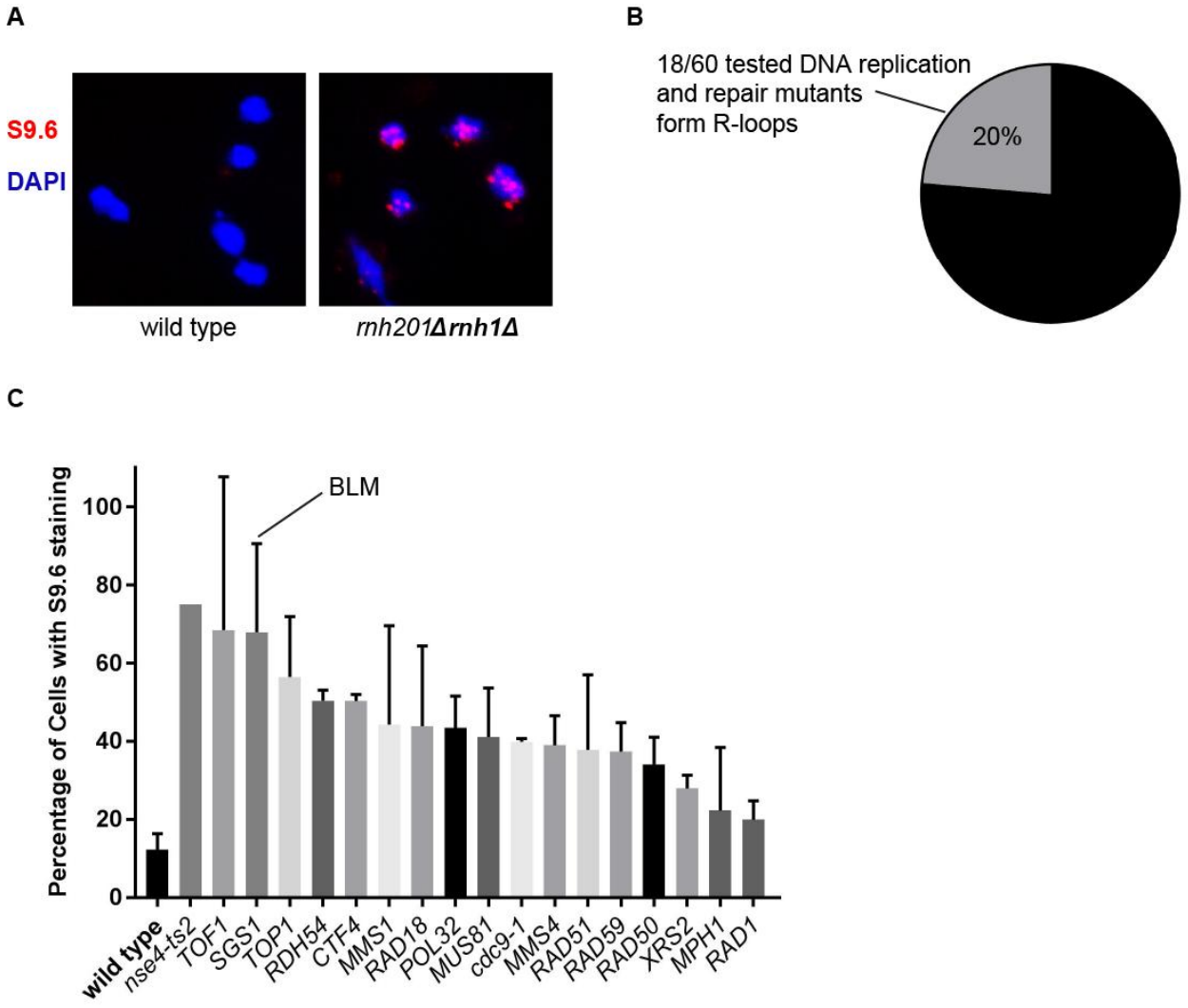

Figure 3 R-loop accumulation levels in DNA replication and repair CIN yeast mutants. (**A**) Sample chromosome spread stained with the S9.6 monoclonal antibody in wild type and *rnh1*Δ*rnh201*Δ control cells. (**B**) Pie charts showing the number and percentage of hits in the S9.6 antibody staining screen of 60 CIN mutants. (**C**) DNA repair mutants with ≥20% of cells containing S9.6 foci.

SGS1 and the RECQ-like family of helicases

SGS1 is a helicase of the highly conserved RECQ family. Unlike humans which encode five family members (BLM, WRN, RECQ4, RECQ1 and RECQ5), yeast only encodes *SGS1* and *HRQ1* (Bochman, Paeschke, Chan, & Zakian, 2014), with *SGS1* being far better characterized. Sgs1and BLM are structurally and functionally very similar, with much of what is known about

BLM's function being first described for its orthologue in *S. cerevisiae*, making *SGS1* a powerful system to study BLM helicase. Defects in BLM give rise to Bloom's syndrome, an autosomal recessive disorder characterized by cancer predisposition, reduced fertility and immunodeficiency. Sgs1/BLM have been shown to play a critical role in homologous recombination (HR) at multiple steps, including end-resection, displacement loop (D-loop) formation, branch migration and double Holliday junction dissolution (Bohm & Bernstein, 2014; Larsen & Hickson, 2013).

In addition, recent evidence has revealed a role of Sgs1/BLM in the stabilisation and repair of replication forks damaged during a perturbed S-phase (Bohm & Bernstein, 2014; Larsen & Hickson, 2013). For example, Sgs1 contributes to replication fork stability by dissolving fold back structures and Holliday junctions formed at collapse forks (Hegnauer et al., 2012). Sgs1 also recruits the intro-S phase checkpoint kinase Rad53 to stalled forks, and *SGS1* and *BLM* deficient cells are sensitive to replication stalling agents like hydroxyurea (HU), aphidicolin and Mehtyl Methanesulfonate (MMS) (Spies, 2013).

Finally, BLM plays a role in the suppression and/or resolution of ultra-fine anaphase DNA bridges that form between sister-chromatids during mitosis (Bohm & Bernstein, 2014; Larsen & Hickson, 2013).

Sgs1/BLM's functions fit well with the established models of R-loop processing. R-loops are thought to be open DNA structures with flap extremities, and we speculate that they might be recognized and processed by Sgs1/BLM *in vivo*. Sgs1/BLM can act on a wide array of substrates *in vitro* including Holliday junctions, D-loops, G-quadruplexes, DNA:RNA hybrids, R-loops and single-stranded overhangs (Larsen & Hickson, 2013; Popuri et al., 2008). Moreover, *SGS1* deficient cells show synthetic phenotypes with RNaseH2 deletion strains in yeast, suggesting a

functional cooperation between the two proteins (Chon et al., 2013; Kim & Jinks-Robertson, 2011).

The most accepted R-loop mediated GI mechanism relies on the potential capacity of DNA:RNA hybrid structures to stall or block replication fork progression. The observation that factors of DSB repair, and tumor suppressors BRCA1, BRCA2 (FANCD1) and FANCM partially prevent the accumulation of R loops in human cells, suggests that DNA damage response factors may participate in this process. Therefore, it is possible that replication restart and HR core proteins like Sgs1/BLM may contribute to the removal of a fraction of R loops. Besides, BLM protein is critical for the activation of the FA pathway. FANC and BLM collaborate during mitosis (Naim & Rosselli, 2009), and in response to stalled forks (Pichierri, Franchitto, & Rosselli, 2004).

1.2 Thesis objectives

The main objective of this project was to investigate a potential role for the DNA repair player *SGS1* in the prevention and/or processing of R-loops. Our working hypothesis was that the helicase *SGS1* counteracts R-loops at least partially by promoting resumption of stalled replication forks at R-loop prone regions.

To test this idea, we first aimed to determine whether *SGS1*'s function is associated with R-loop induced GI such us the hyper-recombination phenotypes of THO complex and *SEN1* mutants. Our strategy involved testing how *SGS1* deletion affects THO complex mutants (*mft1*Δ and *thp2*Δ), and whether *SGS1* alone is able to reduce THO complex mutants' hyper-recombination phenotype.

Our second aim was to establish whether R-loops contribute to the GI phenotypes characteristic of *SGS1* deficient cells. Our approach consisted in investigating if *sgs1*Δ GI defects are sensitive

to RNaseH1 (*RNH1*), a known R-loop suppressor protein, and/or associated with transcription, which also modulates R-loop formation.

Chapter 2: MATERIALS & METHODS

2.1 Yeast strains and growth analysis

Yeast strains and plasmids used in this study are listed in Tables 1 and 2. All strains are derived from the S288C background, with the exception of the rDNA instability strains, which are SK1. Strains were cultured according to standard conditions in Yeast Peptone Dextrose (YPD) or Synthetic Complete medium (SC) lacking the appropriate amino acid when nutritional selection was required.

Growth curves were performed by diluting logarithmic phase YPD cultures to OD600 0.01, and distributing them -at least in triplicate- in 96-well plates. OD600 measurements were taken every 30 minutes for 24 hours at 30°C using a TECAN M200 reader. Strain fitness was assessed by comparing the area under the curve of the growth curves under investigation and using a student's t-test. The expected fitness term for double mutants captured the fitness value predicted from the multiplication of the associated single mutant effects. Drug sensitivity for MMS (0.0025%) and HU (50µg/ml) was determined from liquid growth curves as described above.

Strains	Genotype	Source
BY4741	*MATa his3Δ1 leu2Δ0 met15Δ ura3Δ0*	Peter Stirling
BY4742	*MATαhis3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*	Peter Stirling
PSYL746	*MATa trp1Δ*	This study
PSYL747	*MATα trp1Δ*	This study
PSYL762	*MATa sgs1Δ::KanMX*	Peter Stirling
PSYL 763	*MATα sgs1Δ::KanMX*	Peter Stirling
	MATa leu2Δ0 sen1-1- *ts::KanMX*	Ts-Collection
	MATα leu2Δ0 sen1-1-*ts::KanMX*	This study
2241	*MATα leu2Δ0 mft1Δ::KanMX, trp1Δ*	Phil Hieter
2234	*MATa leu2Δ0 thp2Δ::KanMX, trp1Δ*	Phil Hieter
PSY778	*MATα rnh1Δ:: KanMX*	Peter Stirling
PSY779	*MATa rnh1Δ:: KanMX*	Peter Stirling
PSY780	*MATa rnh201Δ:: KanMX*	Peter Stirling
PSY781	*MATα rnh201Δ:: KanMX*	Peter Stirling
PSYL230	*MATα rnh1Δ::NatMX, rnh201Δ:: KanMX, sgs1Δ::KanMX*	This study
PSYL286	*MATα rnh1Δ::NatMX, sgs1Δ::KanMX*	This study
PSYL287	*MATa rnh1Δ::NatMX, sgs1Δ::KanMX*	This study
PSYL290	*MATα rnh201Δ::G418, sgs1Δ::KanMX*	This study
PSYL291	*MATa rnh201Δ::G418, sgs1Δ::KanMX*	This study
PSYL641	*MATα sen1-1 ts::KanMX, sgs1Δ::KanMX*	This study
PSYL324	*MATa leu2Δ0 mft1Δ::KanMX, sgs1Δ::KanMX*	This study
PSYL325	*MATα leu2Δ0 mft1Δ::KanMX, sgs1Δ::KanMX, trp1Δ*	This study
PSYL326	*MATa leu2Δ0 thp2::KanMX, sgs1Δ::KanMX*	This study
PSYL327.1	*MATα leu2Δ0 thp2::KanMX, sgs1Δ::KanMX, trp1Δ*	This study
PSYL339	*MATα leu2Δ0, sgs1Δ::KanMX, trp1Δ*	This study
PSYL343	*MATa leu2Δ0, sgs1Δ::KanMX, trp1Δ*	This study
PSYL493	*leu2Δ0, Rad52::YFP::URA3*	This study
PSYL515	*leu2Δ0 sgs1Δ::KanMX, Rad52::YFP::URA*	This study
PSYL516	*leu2Δ0 sgs1Δ::KanMX, rnh1Δ::KanMX, Rad52::YFP::URA3*	This study
PSYL517	*leu2Δ0 sgs1Δ::KanMX, rnh201Δ::KanMX, Rad52::YFP::URA3*	This study
PSYL636	*MATa tsa1Δ::KanMX, Rad52::YFP::URA3*	This study
PSYL637	*MATα tsa1Δ::KanMX, Rad52::YFP::URA3*	This study
PSY1232	*MATα rnh1Δ:: KanMX, rnh201Δ:: KanMX, Rad52::YFP::URA3*	Peter Stirling
PSYL646	*MATa rnh1Δ::NatMX, rnh201::G418, sgs1Δ::KanMX, Rad52::YFP::URA3*	This study

Strains	Genotype	Source
PSYL654	*MATa rmi1Δ::KanMX, Rad52::YFP::URA3*	This study
PSYL655	*MATα rmi1Δ::KanMX, Rad52::YFP::URA3*	This study
PSYL656	*MATa top3Δ::KanMX, Rad52::YFP::URA3*	This study
PSYL657	*MATa top3Δ::KanMX, Rad52::YFP::URA3*	This study
PSY1195	*MATα leu2Δ0 rnh201Δ::G418,Rad52::YFP::URA3*	Peter Stirling
PSYL336	*MATa his4 ura3 leu2 ho::LYS2 lys2, 25SRNA::URA3*	Phil Hieter
PSYL742	*MATa his4 ura3 leu2 ho::LYS2 lys2 25SRNA::URA3, sgs1Δ::KanMX*	This study
PSYL337	*MATa his4 ura3 leu2 ho::LYS2 lys2 25SRNA::URA3, sin3Δ:: NatMX*	Phil Hieter
PSY748	*MATa* sen1-1-*ts::KanMX, trp1Δ*	This study
PSY749	*MATa* sen1-1-*ts::KanMX, trp1Δ*	This study
PSY750	*MATa rnh1Δ:: KanMX, trp1Δ*	This study
PSY751	*MATa rnh1Δ:: KanMX, trp1Δ*	This study
PSY752	*MATa rnh201Δ:: KanMX, trp1Δ*	This study
PSY753	*MATa rnh201Δ:: KanMX, trp1Δ*	This study
PSY754	*MATa rnh1Δ::NatMX, rnh201Δ:: KanMX, trp1Δ*	This study
PSY755	*MATa rnh1Δ::NatMX, rnh201Δ:: KanMX, trp1Δ*	This study

Table 1 *S. cerevisiae* strains used in this study

2.2 Recombination analysis

Recombination events in plasmids L, LYΔNS, LNA, pARSHLB-IN, pARSHLB-OUT, pARSCLB-IN, pARSBLB-IN, L-lacZ and GL-lacZ (**Tables 2**) were scored by counting Leucine prototophic (LEU+) colonies. Recombination frequencies were obtained from the average value of all tests performed with 3-9 independent transformants (3 tests each). The entire endogenous *LEU2* gene was completely absent in all strains used in this study, so that LEU+ recombinants could only arise by recombination between the *LEU2* repeats of the plasmid-borne recombination constructs.

Cell viability was measured in parallel to recombination events by growing transformants in SC medium lacking the amino acids matching the auxotrophic marker in the recombination plasmid. In assays where yeast strains were transformed with a recombination and an over-expression vector, recombination and viability plates maintained both plasmids. Recombination frequencies are expressed as a proportion of viable cells plated.

2.3 Ribosomal DNA (rDNA) instability assay

Yeast strains with *URA3* marker inserted into the rDNA locus were used in this assay (Wahba et al., 2011). rDNA instability was calculated, as described for the recombination assay, from the frequency of colonies appearing in 5-Fluoroorotic (5-FOA) plates, a counter-selection for the *URA3* gene, which most often resulted from the excision of *URA3* from rDNA. Cell viability was measured by growing test strains on SC minus uracil (-URA) plates.

The rDNA instability assay was also conducted using strains carrying a human RNaseH over-expression plasmid. Cells were dilution streaked out on SC-URA-LEU plates to select for

presence of *URA3* in the rDNA, as well as the *LEU2*-marked RNaseH or control vector. Cells were grown in parallel in 5-FOA-LEU to quantify rDNA instability.

2.4 A-like faker assay

Absolute frequencies of chromosome III loss were quantified in *MAT*α haploid knockout collection strains using the ALF assay described in (Novoa, (In press, 2017, 16 pages)). This assay measures the presence of MATa-like cells in *MAT*α strains by a mating test involving selection of prototrophic mated products after exposure to a MATα tester strain.

2.5 Yeast live-cell imaging of Rad52-foci

Cells expressing Rad52-Yellow Fluorescent Protein (YFP) were grown to logarithmic phase. Then, ~100µl of OD600 0.1-0.3 yeast cells were added to concanavalin-A (0.25mg/ml) coated slides and imaged on a Leica dmi8 inverted fluorescence microscope (Stirling et al., 2012a). For chemical treatment, 6-azauracil was added at 100µg/mL for 240 minutes prior to imaging.

For Rnh1 over-expression (pYX253 and pKJM1011 plasmids, **Table2**), strains were grown overnight in SC-LEU + 2% raffinose liquid media. The next day, cultures were diluted to OD600 ~0.2 and shaken at 30°C for 2 hours in SC-LEU + 2% raffinose liquid media. Then, galactose was added to the media to a final concentration of 2% for induction of gene expression. Strains were shaken for additional 4 hours to an OD600 ~1.0, and cells were prepared for imaging.

2.6 Chromosome spreads

Chromosome spreads were performed as previously described (Chan, 2014). Slides were incubated with the mouse monoclonal antibody S9.6 directed to DNA:RNA hybrids, and available in the hybridoma cell line HB-8730. S9.6 was obtained as a gift from the laboratories of Doug Koshland and Phil Hieter or from Kerafast (Boston, MA). The primary antibody was diluted

1:1000 in blocking buffer (5% BSA, 0.2% milk, 1× phosphate buffer saline (PBS)) for a final concentration 2µg/ml. The secondary Cy3-conjugated goat anti-mouse antibody (No. 115-165-003) was obtained from Jackson ImmunoResearch (West Grove, PA) and diluted 1:1000 in blocking buffer. For each sample, at least 60 nuclei were visualized, and the nuclear fluorescent signal was quantified using ImageJ (Schneider, Rasband, & Eliceiri, 2012). Each mutant was assayed in quadruplicate. For comparisons purposes, the S9.6 median fluorescence intensity of the wild type strain of each experiment was used for normalization. Mutants were compared to wild type by the unpaired t test.

2.7 Cell culture

HCT116 (CIN^-, microsatellite instability (MIN^+)) colorectal carcinoma cells were a gift from the laboratory of Dr. Phil Hieter. Specifically, we obtained parental population BV8 (p53 +/+) and BV73 ($BLM^{-/-}$). Cell culture reagents were purchased from STEMCELL Technologies unless otherwise specified. Human colorectal carcinoma HCT116 cells were cultured in McCoy's 5a medium supplemented with 10% fetal bovine serum (FBS), 100U/ml penicillin, 100µg/ml streptomycin and 2mM Glutamine. Cells were incubated at 37°C and humidified in a 5% CO_2 atmosphere. Cells were passaged into glass coverslips at 70 to 80% confluency by washing the cells with PBS followed by incubation with 15% media volume Trypsin/EDTA for 2 minutes at 37°C. After trypsinization cells were re-suspended in fresh medium and centrifuged at 300 rcf for 3 minutes. Cells were re-suspended once more in fresh medium and allowed to grow for a maximum of 2 days on squared glass coverslips prior to fixation and imaging.

2.8 Western blot

Protein extraction from HCT116 cells was performed on ice according to standard procedures. Briefly, cultured cells were washed twice with cold PBS. Then, cells were treated with 100 µl of RIPA buffer containing proteinase inhibitor cocktail -for every $1x10^6$ cells-, and scraped from culture dish. Cell lysates were incubated on ice for 30 minutes, centrifuged (10,000 rcf, 5 minutes) and the supernatant retained. Equal amounts of protein (quantified by Bradford Assay reagent, Bio-Rad) were run on SDS-PAGE gels (8%), transferred to nitrocellulose membranes and probed with the indicated antibodies.

2.9 Cell fixation and immunostaining

Cells grown on coverslips were washed once with 1x PBS at room temperature. Cells were fixed to the glass coverslips by treating them with -20°C methanol for 10 minutes. After fixation, cells were rinsed once for 1 minute with 1x PBS and permeabilized for 1 minute with acetone. Then, cells were washed 3x for 1 minute with 1x PBS. After, slides were incubated overnight at 37°C either with 500µl of RNH1 (5U) from New England Biolabs (Ipswich, MA) or with 4% bovine serum albumin (BSA). Both RNH1 and BSA were diluted in PBS. Next, slides were incubated with blocking buffer (3% BSA, 0.1% Tween 20 in 4x saline sodium citrate) for 30 minutes. Slides were drained well and incubated with antibodies S9.6 (10µg/ml) from Kerafast (Boston, MA) and Nucleolin (1µg/ml) from Abcam (Cambridge, UK) at $4C^0$ overnight. The next day, cell were washed 3x for 5 minutes with PBS on a rocker. Secondary antibodies were added and slides were incubated for 1 hour at room temperature. Cells were washed 5x for 5 minutes with 1x PBS on a rocker. Finally, cells were incubated with DAPI (0.5 µg/ml) for 30 minutes for nuclear staining. Coverslips were washed 3x with PBS and mounted on clean glass slides containing two drops of Fluorosave TM Reagent (CalBiochem).

2.10 Microscopy

Image acquisition was done using either a Nikon A1-siConfocal TIRF microscope (Nikon Instruments Inc.) or Leica dmi8 inverted fluorescence microscope (Leica Biosystems). A 60x oil apochromat objective was used for analysis of all HCT116 coverslips and a 100x oil Plan Neofluar objective was used for imaging of chromosome spreads and live yeast.

Quantification of S9.6 nuclear staining for HCT116 cells was done using ImageJ (Schneider et al., 2012). For each sample, at least 25 cells were visualized. For comparisons purposes, the S9.6 median fluorescence intensity of wild type HCT116 cells was used for normalization in each experiment. HCT116 $BLM^{-/-}$ cells were compared to wild type by the unpaired t test.

Quantification of Rad52 foci was done by analysis of both DIC and fluorescence images using Metamorph (Molecular devices).

Plasmids	Description	Source
pRS316	YCp vector based on the *URA3* marker	(Mischo et al., 2011)
pRS316L	pRS316 containing two direct repeats of *LEU2* gene sharing 600pb of homology	Gift from Andres Aguilera (Prado & Aguilera, 1995)
pRS316LYΔNS	pRS316L containing the complete YIp5 sequence (pBR322 containing the *URA3* gene), except a 1.92kb *Sph*I-*Nsi*I deletion, inserted at the *Bgl*II site located in between the repeats	Gift from Andres Aguilera (Prado, Piruat, & Aguilera, 1997)
pRS314	YCp vector based on the *LEU2* gene	(Mischo et al., 2011)
pRS314L	pRS314 containing two 598 bp *Cla*I–*Eco*RV *LEU2* fragments repeated in direct orientation, separated by 31 bp of polylinker	(Mischo et al., 2011)
pRS314LYΔNS	pRS314-LY with the 1.92 kb *Sph*I–*Nsi*I fragment of YIp5 deleted	(Mischo et al., 2011)
pRS314-LNA	pRS314-L with the 2.16 kb *Bam*HI–*Sma*I fragment of YIp5 inserted at *Nru*I, leaving the *Sma*I this study site distal to the *Eco*RV–*Bgl*II *LEU2* fragment	(Mischo et al., 2011)
pCM189	Centromeric plasmid containing *tetO* promoter and *URA3* marker	(Gari, Piedrafita, Aldea, & Herrero, 1997)
pCM189RNH1	*RNH1* ORF cloned into pCM189 *Not*I	(Castellano-Pozo et al., 2013)
pYX253	Galactose inducible, *LEU2*	(Bell et al., 2011)
pKJM 1011	Galactose inducible, *LEU2*, *RNH1*	(Bell et al., 2011)
p416-pGal-RNH1	*CEN6, URA3, full-length yeast RNH1*	Gift from Phil Hieter. Originally from (Cho, Kim, & Jinks-Robertson, 2015)
pYes2.1	2μ Galactose expression plasmid with *URA3* and Amp markers	Gift from Dr. Elizabeth Conibear
pYES2-*SGS1*	pYes2 plasmid containing full-length *SGS1* under Gal promoter	Gift from Dr. Allan Morgan (Mankouri, Craig, & Morgan, 2002)
MoBY-YMR190C	MoBY-ORF collection, CEN, *URA3*	Gift from Phil Hieter
pARSHLB-IN	Centromeric plasmid containing 0.6 lb *LEU2* direct-repeat transcribed from *HHF2* promoter. ARS facing towards gene	Gift from Phil Hieter. Originally from (Prado & Aguilera, 2005)
pARSHLB-OUT	Centromeric plasmid containing 0.6 lb *LEU2* direct-repeat transcribed from *HHF2* promoter. ARS facing away from gene	Gift from Phil Hieter. Originally from (Prado & Aguilera, 2005)
pARSCLB-IN	Centromeric plasmid containing 0.6 lb *LEU2* direct-repeat transcribed from *CLN2* promoter. ARS facing towards gene	Gift from Andres Aguilera
pARSBLB-IN	Centromeric plasmid containing 0.6 lb *LEU2* direct-repeat transcribed from *CLB2* promoter. ARS facing towards gene	Gift from Andres Aguilera
L-lacZ	0.6kb internal *LEU2* repeated sequences, under *LEU2* promoter	Gift from Andres Aguilera (Gonzalez-Aguilera et al., 2008)
GL-lacZ	0.6kb internal *LEU2* repeated sequences, under *GAL1* promoter	Gift from Andres Aguilera (Gonzalez-Aguilera

Plasmids	Description	Source
		et al., 2008)
pGDP-empty	Plasmid marked with LEU2, 2µ	Gift from Douglas Koshland
pGPD-RNH1	Human RNaseH1 cloned into plasmid marked with LEU2, 2µ	Gift from Douglas Koshland

Table 2 Plasmids used in this study

Chapter 3: RESULTS

3.1 *SGS1* interaction with R-loop suppressors

Genetic interaction between *SGS1* and R-loop surveillance genes

Based on the idea that replication restart and HR core proteins like Sgs1 can contribute to the removal of a fraction of R-loops, we investigated the genetic interaction of *SGS1* with genes known to play a role in R-loop metabolism such as *RNH1*, *RNH201*, *SEN1*, *MFT1* and *THP2*. For this purpose, we created double and triple mutants and monitored their growth and drug sensitivity over a period of 24 hours. We detected a decrease in fitness in the double mutants lacking *SGS1* and with dysfunctional *SEN1*, *RNH201*, *MFT1* or *THP2*. The decrease in fitness was significantly higher than expected in the *sgs1*Δ*rnh201*Δ, *sgs1*Δ*rnh1*Δ*rnh201*Δ, *sgs1*Δ*sen1-1-ts* and *sgs1*Δ*mft1*Δ strains but not in the *sgs1*Δ*thp2*Δ strain (**Figure 4**, A). Interestingly, the *sgs1*Δ*rnh1*Δ*rnh201*Δ triple mutant showed sensitivity to growth in HU and MMS, which are known to enhance DNA replication stress (**Figure 4** B, C), while *sgs1*Δ*sen1-1-ts* was sensitive to HU (**Figure 4** B, C). This shows that *SGS1* is required for the survival of deletion and temperature sensitive R-loop surveillance mutants. Cells that lack R-loop surveillance factors are not able to compensate for the lack of *SGS1*.

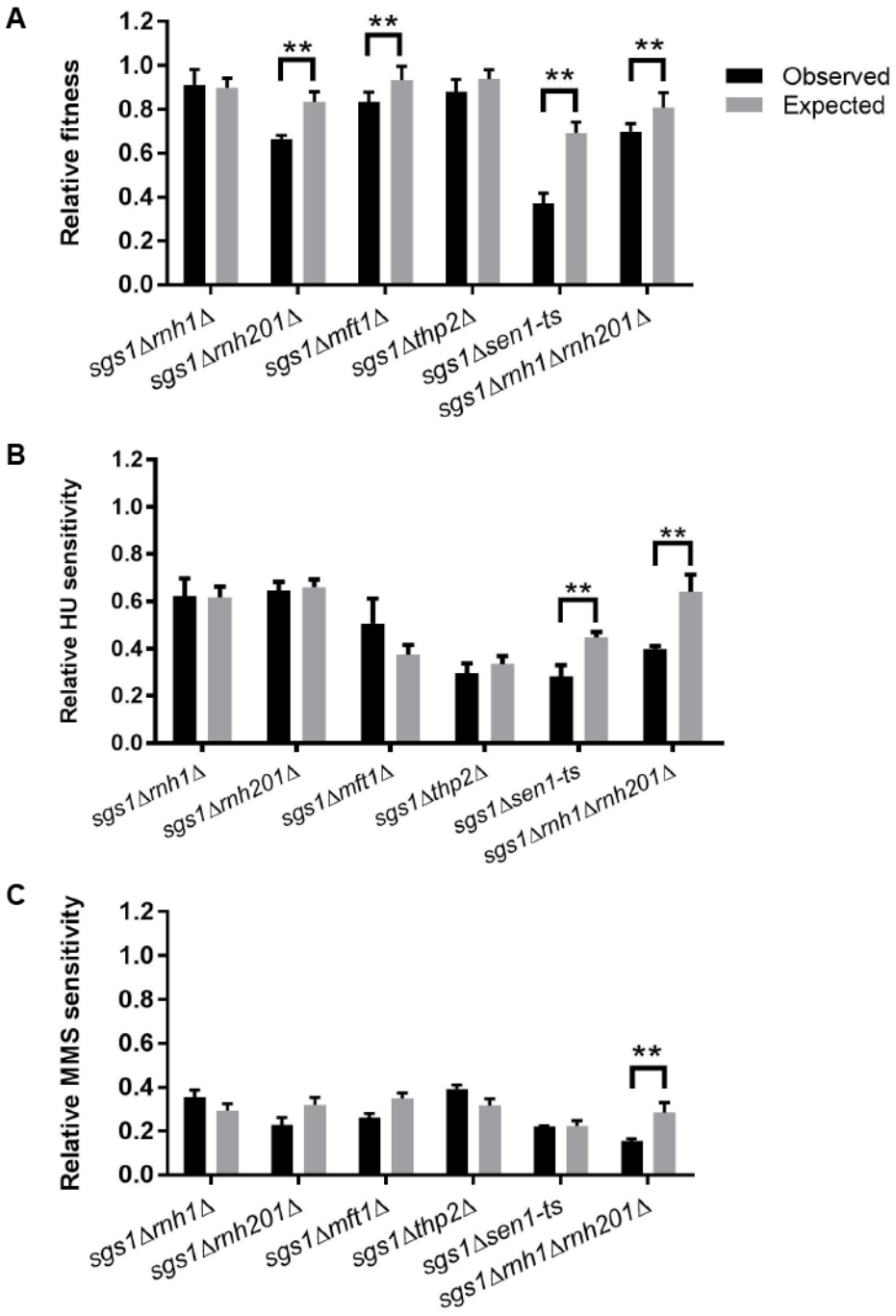

Figure 4 *SGS1*and R-loop suppressor factors double and triple mutants exhibit growth defects. Relative fitness based on 24 hour OD 600 measurements of cultured yeast mutant strains at 30°C. Starting dilution of OD 600 = 0.01 for each replicate (**A**) Cells were grown on YPD, (**B**) YPD plus 50mM hydroxyurea (HU), or (**C**) YPD plus 0.0025% methyl methanesulfonate (MMS). (**P<0.01by student's t-test).

Effects of *SGS1* deletion on CIN phenotypes of R-loop surveillance genes (*RNH201*, *RNH1*, *MFT1* and *SEN1*)

In addition to growth defects, we explored the effect of *SGS1* deletion on CIN phenotypes of mutants of R-loop surveillance factors. First, we tested whether *sgs1*Δ exacerbated the R-loop associated hyper-recombination phenotype previously shown in THO complex mutants *mft1*Δ and *thp2*Δ (Huertas & Aguilera, 2003). For this, we employed a plasmid borne recombination substrate that carries two truncated regions of *LEU2* overlapping by 600 nucleotides of homologous sequence (LNA) (Prado et al., 1997) (**Figure 5**, A). We measured recombination as the frequency of colonies growing in plates missing Leucine. Deletion of *SGS1* and either *MFT1* or *THP2* resulted in the restoration of *LEU2* at a much higher frequency than in either *SGS1* or THO complex single deletion mutants (**Figure 5**, B). Our results support previous observations that R-loop mediated damage is repaired via HR. More importantly, the combined effect of both *SGS1* and THO complex mutation appears to be synergistic, potentially supporting a unique role for Sgs1 in R-loop resolution.

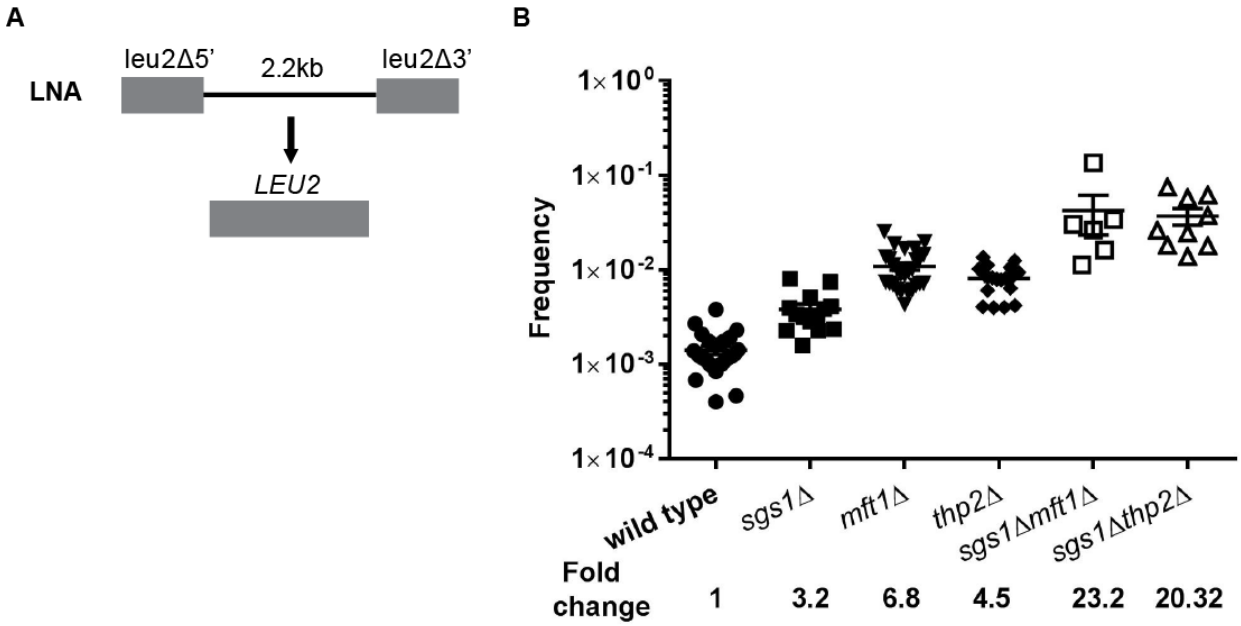

Figure 5 *SGS1* deletion further increases recombination frequency of THO complex deletion mutants. (**A**) Schematic of recombination substrate LNA (Prado et al., 1997). (**B**) Scatter plot with mean and standard error (whiskers) of recombination tests from four independent colonies.

Next, the established A-like Faker assay for disruption of the *MAT* locus in chromosome III (Yuen et al., 2007) confirmed the effect of deletion of *SGS1* on THO complex mutants. In this assay, *MATα* haploid yeast knockout mutants were tested for elevated frequency of loss of the *MATα* locus, which results in dedifferentiation to the MATa-mating type, thus the name of the assay. The presence of these cells was detected by prototrophic selection of mated products after exposure to a *MATα* tester strain (**Figure 6**, A). Double deletion mutants of *SGS1* and the THO complex genes *MFT1* or *THP2* exhibited dramatic increases in the frequency of chromosome III loss (**Figure 6**, B). Similarly, mutations in *SEN1* or deletion of *RNH1* and *RNH201* led to large synergistic increases in ALF frequency (**Figure 6**, B).

The results of the recombination and ALF assays are not surprising because *SGS1* plays a major role in the HR pathway. Moreover, *SGS1* prevents mitotic crossing-over events during repair of DSBs. If R-loops lead to CIN in the studied mutants via DSBs, which can be repaired by HR, it

would be expected that cells known to form R-loops would show increased levels of recombination and *MAT* type conversion. Nevertheless, the magnitude of the increase in CIN in the tested double mutants suggests that *SGS1* deletion might result in putative R-loop increase and subsequent GI.

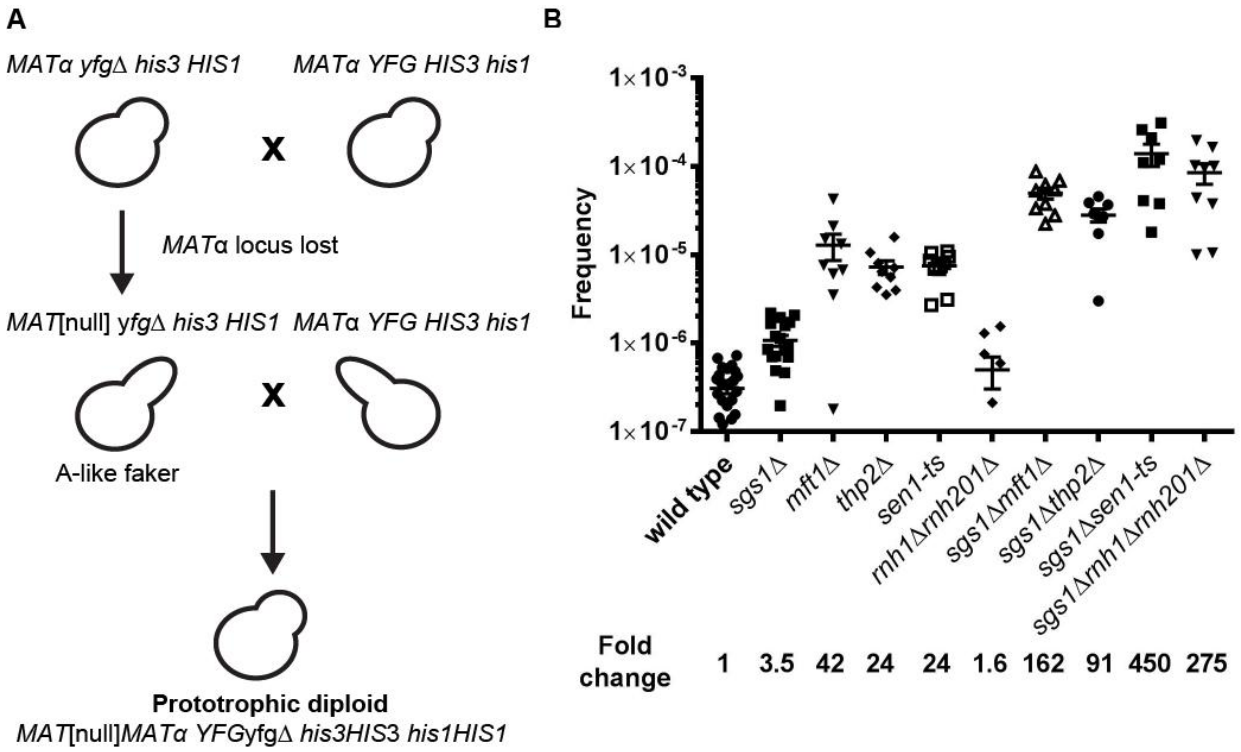

Figure 6 *SGS1* deletion in mutants of R-loop regulators aggravates *MAT* locus instability (**A**) Schematic of ALF assays. A MAT*α* query strain deleted for any gene of interest (*yfg*Δ = your favorite gene) loses the *MATα* locus (black arrow at left) leading to a *MAT*a-like phenotype. The frequency of that event was captured by scoring mating events. Only prototrophic diploids resulting from the cross between the tester (*YFG*) and inquiry (*yfg*Δ) strain are able to grow on synthetic dextrose agar plates (minimal media) because the *MAT[null]* strain is rescued by mating with the tester strain which lacks only the *HIS1* gene. (**B**) Scatterplot of frequency of *MAT* locus instability in various mutants. The mean and standard error (whiskers) of at least nine mating type conversion tests from independent colonies were analyzed.

The third and last assay we employed to explore GI in double deletion mutants of R-loop regulating genes and *SGS1* was quantification of DNA repair centres. Yeast cells accumulate Rad52 and other proteins into distinct subnuclear foci in response to DNA damage, which can be visualized when Rad52 is tagged with a fluorescent protein (Alvaro, Lisby, & Rothstein, 2007).

We measured Rad52 foci formation in the following strains: *sgs1*Δ, *rnh201*Δ, *sgs1*Δ*rnh201*Δ and *sgs1*Δ*rnh1*Δ*rnh201*Δ. We observed that wild type, *sgs1*Δ, and *rnh201*Δ cells carry 3.4%, 8% and 9.4% Rad52-YFP foci, respectively. (**Figure 7**, B). A previous screen of gene deletions affecting the levels of spontaneous Rad52 foci (Alvaro et al., 2007) identified *SGS1* as a hit. They showed 22% of *sgs1*Δ cells form spontaneous Rad52-YFP foci. Therefore, we were not able to replicate this result in our assay despite of re-making and re-testing the *sgs1*Δ strain. A potentially important difference between the two assays is that we used deletion haploids while Alvaro *et al.* tested hybrid diploid cells homozygous for gene deletions. Importantly, they observed that in wild type yeast foci occur in about 5% percent of unsynchronized cells (Alvaro et al., 2007), which is similar to our own finding for that strain.

In spite of our lower single-mutant values, we detected that loss of *SGS1* in *rnh201*Δ and *rnh1*Δ*rnh201*Δ strains caused a large increase in DNA damage. *sgs1*Δ*rnh201*Δ and *sgs1*Δ*rnh1*Δ*rnh201*Δ strains exhibited foci in 18.9% and 26.2% of cells, respectively.

In summary, this section showed that a functional interaction exists between *SGS1* and R-loop regulators like *MFT1*, *SEN1* and RNaseH genes.

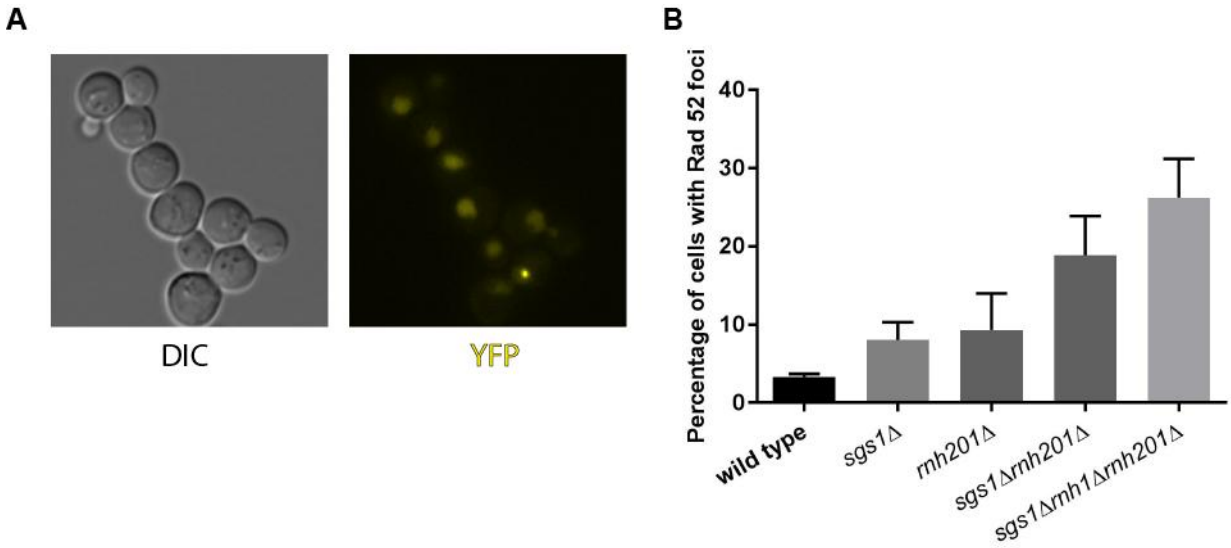

Figure 7 Spontaneous DNA damage in *SGS1* and *RNaseH* double and triple deletion mutant strains. (**A**) Differential interference contrast filter (DIC) and YFP fluorescence images of wild type Rad52-YFP::*URA3* cells. Post image acquisition, DIC images were used to identify and count focused cells. Subsequently, focused cells were examined for the presence of Rad52 foci on the matched YFP image. (**B**) DNA damage in *sgs1*Δ, *rnh1*Δ and *rnh201*Δ. Bars contain averages and standard deviations from at least 3 independent experiments. At least 100 cells were analysed per strain per experiment.

Effects of *SGS1* over-expression on the recombination phenotype of R-loop surveillance gene *MFT1*

Replication fork stalling and/or blockage has been shown to occur in yeast THO mutants at locations where R loops are preferentially accumulated (Wellinger, Prado, & Aguilera, 2006). At the same time, Sgs1 has several roles in fork restart (Bernstein et al., 2009). In light of both observations, we tested whether Sgs1 alone is able to reduce the R-loop induced hyper-recombination phenotype observed in THO complex mutant *mft1*Δ.

SGS1 over-expression from a *GAL1* inducible high copy plasmid rescued the mild recombination phenotypes of *mft1*Δ cells bearing the L substrate (**Figure 8**). This suggests that *SGS1* alone might be able to rescue the GIN phenotypes of THO mutants. As expected *SGS1* over-expression also rescued *sgs1*Δ cells.

SGS1 over-expression has been previously reported to have detrimental effects (Frei & Gasser, 2000; Sinclair, Mills, & Guarente, 1997). In our hands, *SGS1* over-expression from the *GAL1* inducible promoter resulted in increased recombination frequencies in wild type cells (**Figure 8**). The effect of Sgs1 might also be dose dependent.

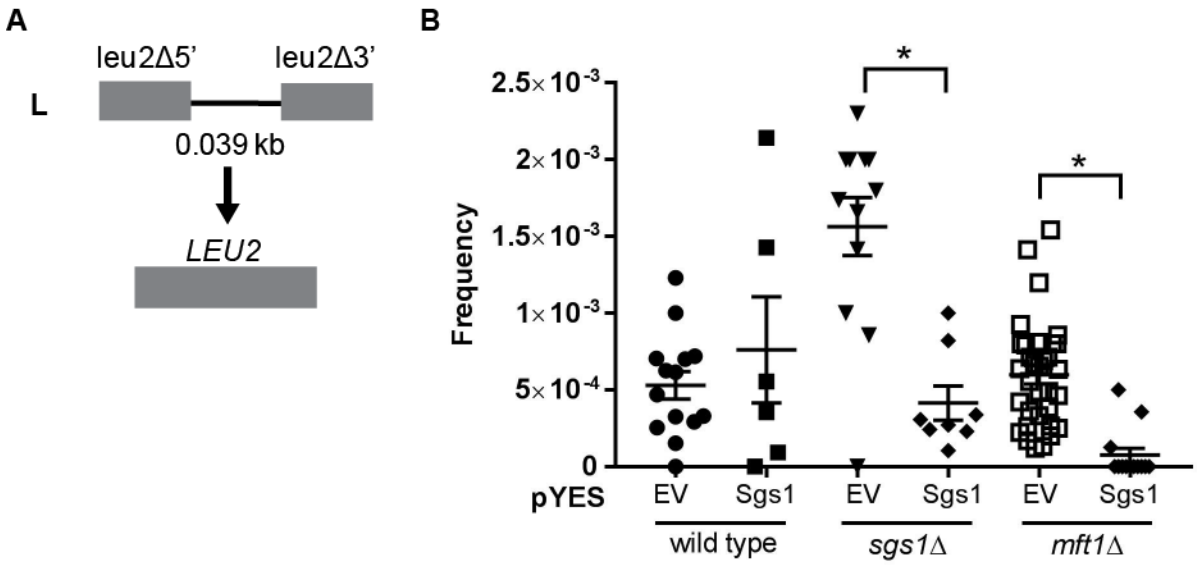

Figure 8 *SGS1* over-expression rescues recombination phenotypes of THO complex mutant (**A**) Schematic of recombination substrate L (Prado et al., 1997). (**B**) Recombination frequencies in wild type, *sgs1*Δ and *mft1*Δ strains. The mean and standard error are indicated. (*P<0.05 by student's t-test)

3.2 R-loop dependent genome instability in Sgs1 deficient cells

DNA damage and GI in *sgs1*Δ are partially transcription dependent

Hyper-recombination in THO/TREX-2 and *SEN1* mutants show clear transcription dependence, as the instability increases with greater transcript length and transcription rate but decreases when the R loop-forming RNA is removed either by RNaseH activity or ribozyme directed RNA cleavage. We investigated if a similar transcription dependency was true for the GI observed in cells deficient in Sgs1.

To start, we tested if *sgs1*Δ shows a similar transcript length dependent recombination phenotype using recombination substrates L and LYΔNS (**Figure 9**, A). The former carries 39 bps of nucleotides of homologous sequence between two truncated regions of *LEU2* while the latter carries 3700 bps between the same repeat sequences. As expected, *sgs1*Δ had higher recombination frequencies than wild type for both L and LYΔNS. Importantly, *sgs1*Δ had a larger fold change in recombination between the long (LYΔNS) and short transcript (L) (**Figure 9**, B).

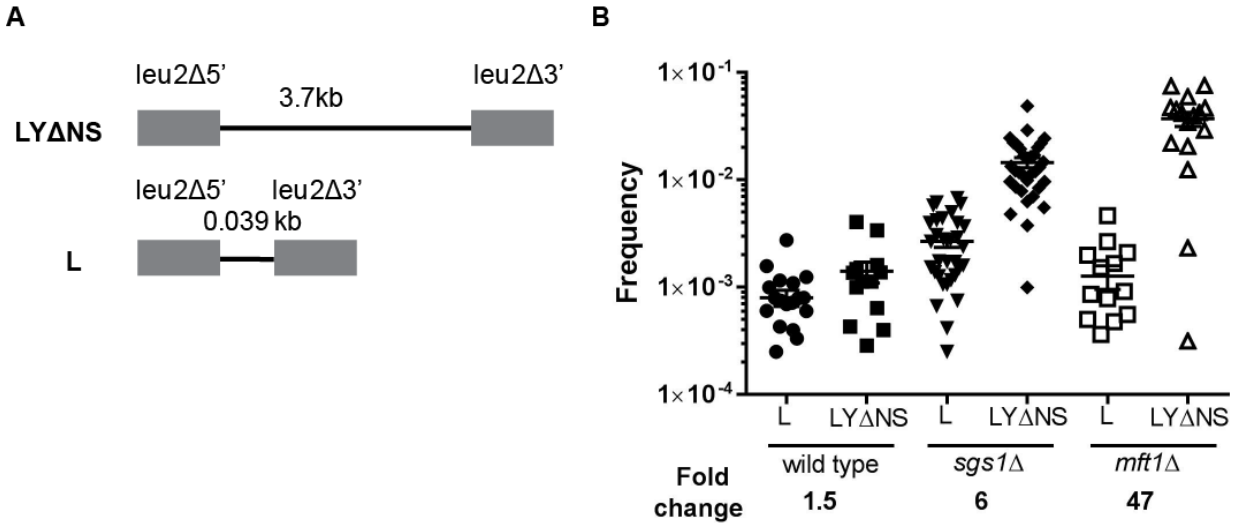

Figure 9 Transcript length affects recombination frequency in *sgs1*Δ cells. (**A**) Schemes of the recombination substrates L and LYΔNS (**B**) Recombination frequencies for wild type, *sgs1*Δ and *mft1*Δ strains with mean and standard error shown.

Another factor that modulates R-loops is transcription rate. We analyzed *sgs1*Δ recombination levels at low and high transcription using the GLacZ system (Mischo et al., 2011) and found that higher transcription levels led to a significant increase in recombination in *sgs1*Δ cells (**Figure 10** A,B). This increase was ~2 fold higher than the increase observed for wild type cells.

We also determined that recombination in *sgs1*Δ is induced by transcription only during S-phase, at which time collisions between transcription and replication conflicts are possible. To accomplish this we measured recombination frequencies using two versions of a *LEU2* gene

recombination construct. The first one with transcription of *LEU2* occurring in the opposite direction of replication (IN), and the second one with transcription and replication happening co-directionally (OUT) (Prado & Aguilera, 2005) (**Figure 11**, A). In addition, *LEU*2 transcription in the IN and OUT constructs was under the control of a promoter of either of the following cell-cycle specific genes: CLN2 G1-cyclin (CLB-IN), CLB2 G2-cyclin (BLB-IN), and H4-like histone gene HHF2 expressed in late G1/S phase (HHF-IN, HHF-OUT) (**Figure 11**, B). While transcription and replication took place at non-overlapping times in the CLB-IN and BLB-IN constructs and did not affect recombination, they would have occurred simultaneously in the HHF-IN construct. Transcription and replication machineries moving in opposite directions likely led to the increased recombination frequencies observed in the HHF-IN construct relative to the HHF-OUT construct (**Figure 11**, B).

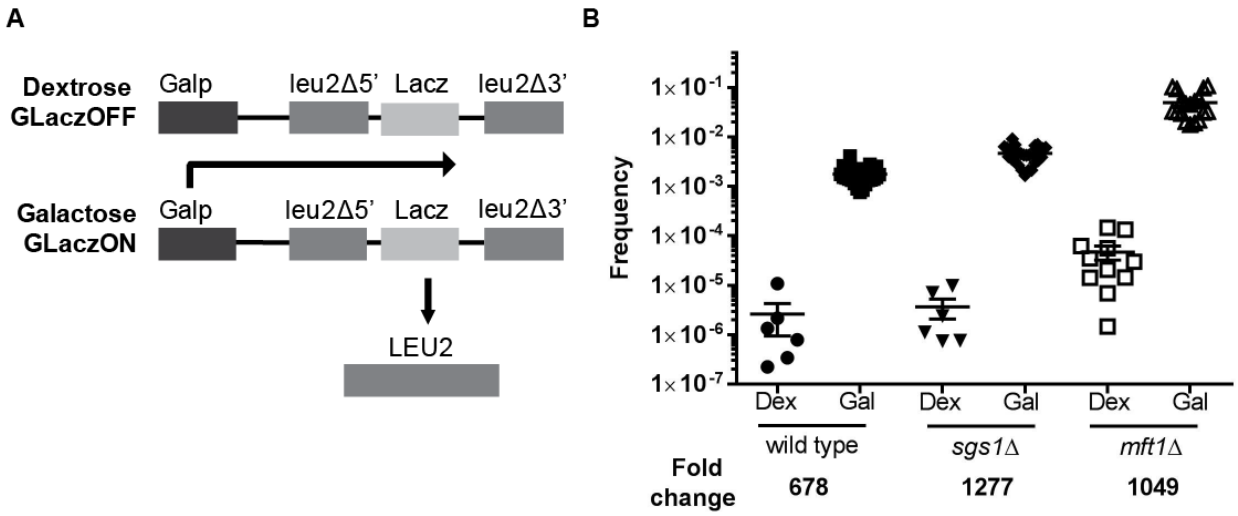

Figure 10 Recombination is induced by RNAPII-mediated transcription in s*gs1*Δ cells. (**A**) Schemes of the plasmids harbouring the recombination construct GLacz. (**B**). Transcription through the *LEU2* repeats was repressed in 2% glucose (GLacZ-OFF) and activated in 2% galactose 2% raffinose (GLacZ-ON). Recombination frequencies for wild type, *sgs1*Δ and *mft1*Δ strains with mean and standard error are indicted.

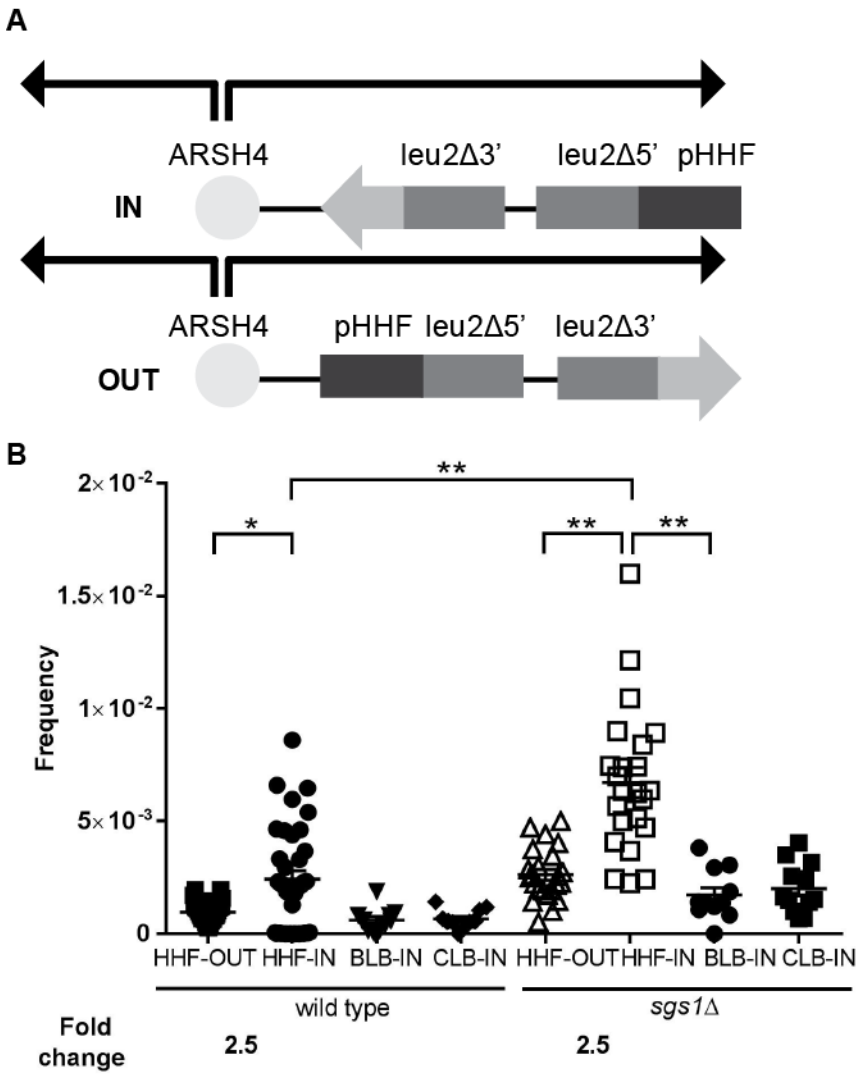

Figure 11 Replication-transcription conflicts are a required condition for transcription dependent recombination in *sgs1*Δ. (**A**) Schemes of the plasmids harbouring the recombination constructs IN and OUT. The arrows indicate the progression orientation of RNAPII transcription driven from the pHHF promoter and of the replication forks initiated at ARSH4. The *LEU2* direct repeats used in assay are under the control of the HHF2, CLN2, or CLB2 promoters, which are activated in G1/S, in the HHF-IN, G1 in the CLB-IN and G2/M in the BLB-IN constructs. (**B**). Recombination frequencies in wild type and *sgs1*Δ strains. Means and standard errors are indicated. (**P<0.01 by student's t-test)

Finally, we asked whether transcription inhibition would rescue any *sgs1*Δ GI phenotype. Specifically, we tested whether Rad52 marked damage in *sgs1*Δ or *sgs1*Δ*rnh201*Δ cells could be suppressed by the transcription inhibitor 6-Azauracil (6AU). We decided to also test the double

mutant because of the lower than reported percentage of *sgs1*Δ cells with Rad52 foci we had previously observed (**Figure 7**).

*sgs1*Δ*rnh201*Δ cells treated with 6-AU showed a reduction of damage marking foci (**Figure 12**). Surprisingly, the 6-AU treatment also suppressed Rad52 foci in *rnh201*Δ but not in *sgs1*Δ. No significant change was seen on the experimental negative control *tsa1*Δ. The damage observed in the *TSA1* mutant is believed to be caused largely by oxidative stress therefore transcription inhibition is not expected to affect it. Although it is unclear whether the effect observed in the double mutant was due to the lack of *rnh201*Δ only, it would be worth testing other transcription inhibitory drugs in a different yeast background or in diploids. As discussed previously, we did not see high levels of DNA damage in *sgs1*Δ cells, yet they showed a non-significant decrease in Rad52 foci in all of our experiments (**Figure 12**, also in data not shown). The dynamic range of the assay might also be obscuring a potential effect of 6-AU on the *SGS1* deletion strain.

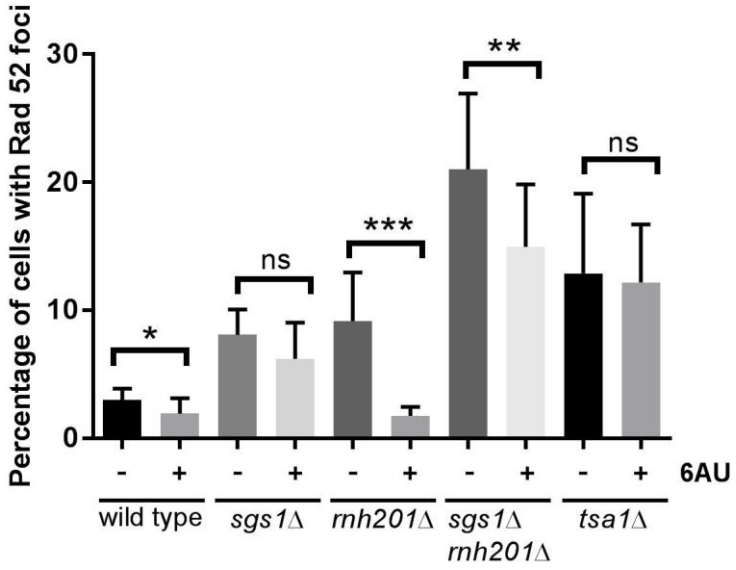

Figure 12 Spontaneous DNA damage after suppression of transcription Cells were treated with 50μg/ml 6AU for 4 hours. Average and standard deviation of at least three independent experiments are shown. At least 50 cells per treatment were counted. (**P<0.01, ***P<0.001 by Fisher's or Chi-square test)

Rad52 foci in *sgs1*Δ*rnh201*Δ double mutants can be reversed by the over expression of RnaseH1

We next tested whether the Rad52-YFP foci of *sgs1*Δ and *sgs1*Δ*rnh201*Δ strains were sensitive to RNaseH1 overexpression, and indicative of increased R-loop formation. Yeast *RNH1* was ectopically expressed in those strains from a *GAL* inducible promoter. Rad52-YFP foci formation was significantly reversed when *sgs1*Δ*rnh201*Δ cells over-expressed *RNH1*. Repression of Rad52 foci was also observed when *sgs1*Δ*rnh1*Δ*rnh201*Δ cells over expressed yeast RNaseH1 from the same plasmid (**Figure 13**). However, basal levels were not reached. This effect was not seen for either of the single mutants (**Figure 13**). Intriguingly, the percentage of cells with Rad52-YFP foci in the *rnh201*Δ strain was lower in this assay than previously reported in this document (**Figure 7, Figure 12**). A possible explanation for this difference is that the vectors used in this experiment might have affected processing of DSBs in *rnh201*Δ. An alternative hypothesis is that changing the sugar in the media could have affected the strains. However we did not see any differences in the levels of Rad52-YFP foci after *rnh201*Δ was cultured in dextrose or raffinose/galactose.

DNA damage in the crosses of *SGS1* and RnaseH enzyme mutants appeared to be dependent on *RNH1* overexpression and, by proxy, R-loop removal. However, the relationship between *SGS1*, *RNH201* and *RNH1* needs to be further investigated through other assays and GI phenotypes. Our and other groups' studies suggest a complicated functional interaction between the proteins studied in this section. For example, (Chon et al., 2013; Kim & Jinks-Robertson, 2011) showed that RNaseH1 overexpression does not complement *sgs1*Δ*rnh201*Δ synthetic growth defects.

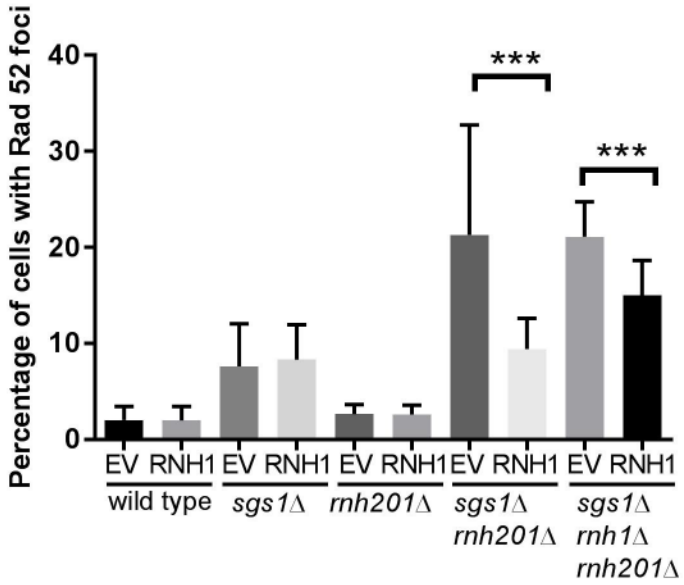

Figure 13 Rnh1 suppresses spontaneous DNA damage in *SGS1* and RNaseH double and triple deletion mutant strains. Cells transformed with empty vector (EV) or pGal-RNH1 plasmid (RNH1) were grown for 4 hours in induction media before imaging. Bars contain averages and standard deviations from at least 3 independent experiments. At least 100 cells were analysed per strain per experiment. (***P<0.001 by Fisher's or Chi-square test)

3.3 Sgs1 depletion shifts the profile of R-loops genome-wide

S9.6 Staining of Sgs1 deficient cells

Given our collaborator's observation that *SGS1* deletion mutant might accumulate R-loops, we validated their results by replicating their chromosome spreads assay but following a slightly different data analysis protocol. Briefly, yeast nuclei were isolated, fixed on a slide, and stained with the R-loop specific antibody S9.6. Then, measurements of S9.6 fluorescence intensity inside yeast nuclei were taken in each strain and the values were used as a proxy for R-loop content. We were able to confirm that *sgs1*Δ cells have higher DNA:RNA hybrid content than wild type cells (**Figure 14**). Our assay showed a close to 2-fold increase in S9.6 staining in *sgs1*Δ compared to the wild type strain. We used the *rnh1*Δ*rnh201*Δ double mutant as a positive

control. Similar to what others have reported, we saw a dramatic increase in R-loops levels after the deletion of both RNaseH enzymes from the yeast genome (**Figure 14**).

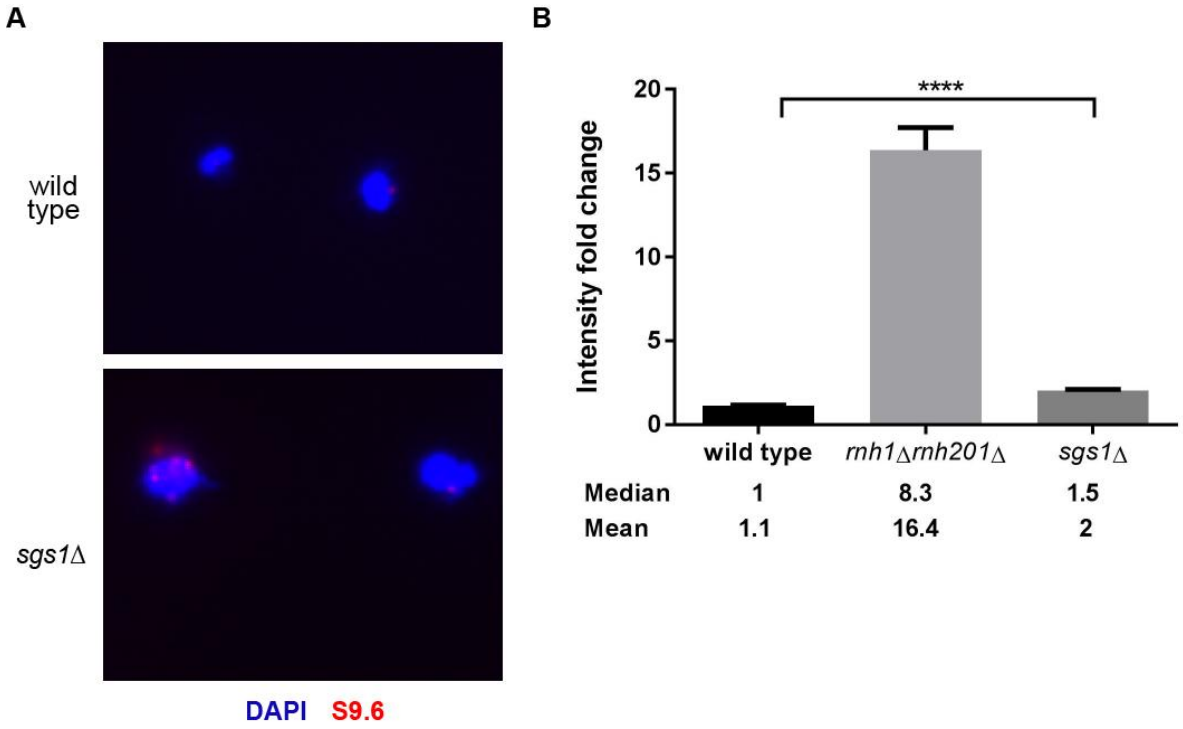

Figure 14 *SGS1* deficient *S. cerevisiae* accumulate R-loops Relative S9.6 antibody fluorescence intensity per fixed yeast nucleus. Error bars represent mean and standard error from four independent experiments. At least 61 nuclei were analyzed per strain. (****P<0.0001 by student's t-test).

Mapping of R-loops in Sgs1 deficient cells

Mapping the location of R-loops both in yeast and human wild type cells has shown that R-loops form at different densities throughout the genome under physiological conditions. This effort has also revealed features that promote R-loop formation, such as GC skew, and high levels of gene transcription (El Hage et al., 2014; Ginno et al., 2012; Wahba, Costantino, Tan, Zimmer, & Koshland, 2016). The acquired knowledge about the location of R-loops in the genome has also allowed researchers to collect evidence supporting a biological role of R-loops in transcription termination, and gene expression.

Using DNA:RNA immunoprecipitation followed by hybridization on a tiling microarray (DRIP-chip) (Chan, 2014), our collaborator from the Kobor lab collected data to map R-loops in Sgs1 deficient cells. The analysis of DRIP-chip data showed that *SGS1* deletion leads to a subtle overall increase in R-loop occupancy. It also suggested that DNA:RNA hybrids are enriched at long genes in *sgs1*Δ compared to wild type cells (Aristizabal, 2017). A small but significant increase in DNA:RNA hybrid signal over TY1 elements and telomeres was also identified (Aristizabal, 2017) in the *SGS1* mutant cells. In addition, R-loops at ribosomal DNA (rDNA) were found to be more abundant in *sgs1*Δ cells. This is consistent with the role of Sgs1 in rDNA replication termination (Mundbjerg et al., 2015).

While the analysis of the complete DRIP-chip dataset by our collaborator is ongoing, we wanted to immediately test the potential locus-specific effects of increased R-loops in the SGS1 mutant strains. Specifically, we chose to further investigate a possible role for Sgs1 in the regulation of rDNA R-loops by testing the sensitivity of that locus to *RNH1* overexpression when *SGS1* is absent. As mentioned before, *RNH1* is a well-accepted modulator of R-loops and it is often used to demonstrate the presence of *bona fide* R-loops.

We measured loss of a *URA3* marker inserted into the rDNA locus, as a proxy for rDNA instability. As previously reported, deletion of *SGS1* resulted in rDNA instability, which might occur through an R-loop mediated mechanism given the greater than wild type R-loop occupancy reported at that locus from the analysis of DRIP-chip data. Consistent with this idea, constitutive over-expression of human-RNaseH1 reversed rDNA instability in *sgs1*Δ cells (**Figure 15**). This results further links R-loops to instability at a specific locus where Sgs1 plays a role in mitigating R-loop accumulation.

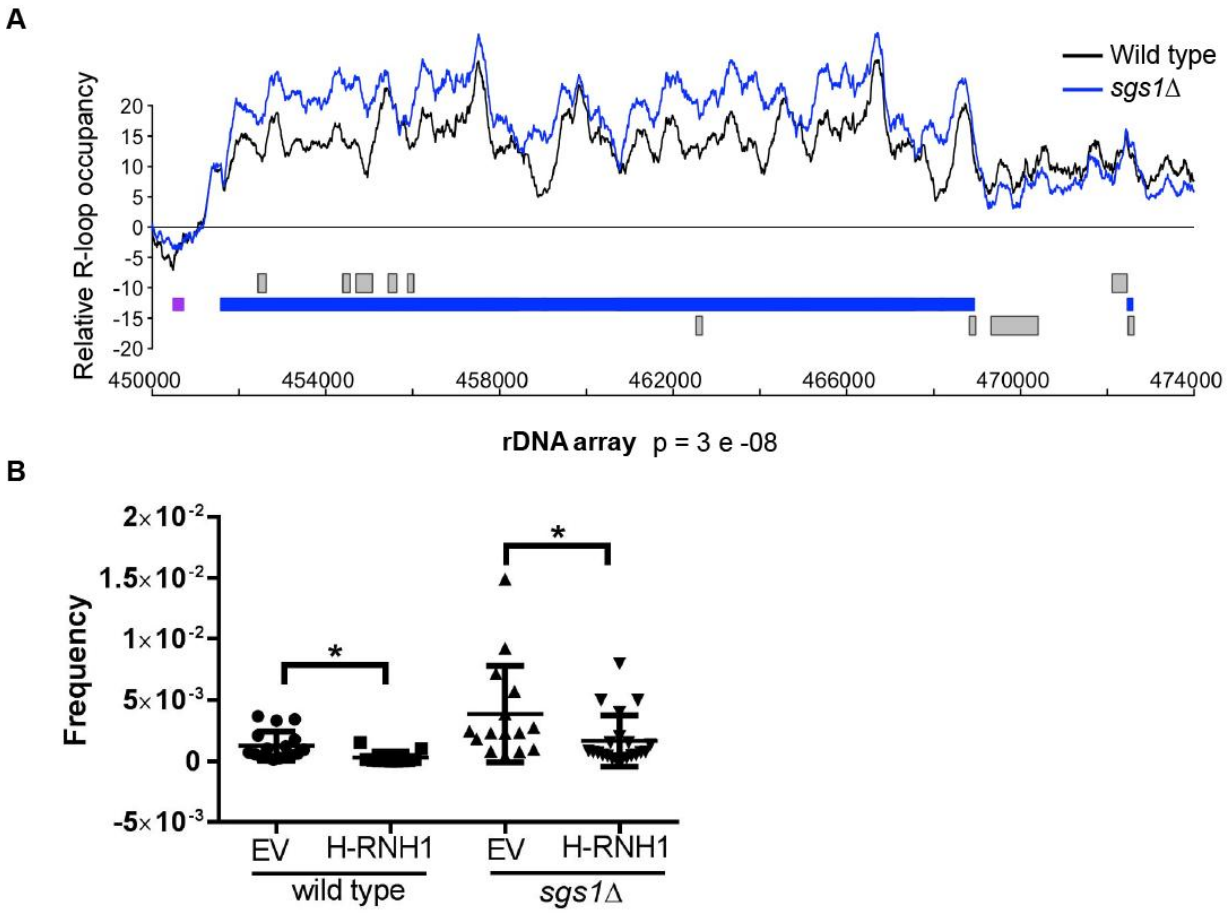

Figure 15 Deletion of *SGS1* shifts R-loop occupancy at the rDNA locus (**A**) R-loop occupancy at the yeast rDNA locus in *sgs1*Δ (blue) and wild type (black) yeast strains using custom R scripts (Schulze et al., 2009). The region is a hotspot for R-loop formation in wild type, but this is significantly enhanced in *sgs1*Δ. (**B**) Over-expression of human RNaseH1 rescues rDNA instability in *sgs1*Δ. Frequencies of excision of a *URA3* marker inserted into the rDNA locus as measured by the appearance of FOA resistant yeast colonies in wild type, and *sgs1*Δ strains. The average and standard deviation are indicated. (*P<0.05, by student's t-test)

S9.6 Staining of HCT116 BLM knockout cells

Sgs1 is a very well conserved protein which functions similarly to its human counterpart BLM in many settings. Therefore, we wondered whether BLM deficient mammalian cells also accumulate R-loops.

Using the S9.6 antibody and immunofluorescence (IF), we found that BLM knock-out in the human colorectal cancer cell line HCT116 leads to a significant increase in nuclear R-loops in comparison to

wild type cells (**Figure 16**). As a control for specificity, we found that treatment of BLM$^{-/-}$HCT116 cells fixed on glass slides with recombinant RNaseH1 results in a decrease in R-loop signal to wild type levels. This demonstrates that the S9.6 signal detected was specific to R-loops. In our analysis we excluded the S9.6 signal that co-localized with Nucleolin antibody, which specifically stains the nucleolus. The reasons for this were that high S9.6 signal was present in the nucleolus even in wild type HCT116 cells, and the nucleolar S9.6 signal was more resistant to RNaseH1 treatment. This has also been reported by other research groups and could be due to incomplete action of nucleases at the nucleolus, where DNA:RNA hybrids are abundant (Sollier et al., 2014).

In addition to demonstrating the presence of high levels of DNA:RNA hybrids in yeast as a consequence of *SGS1* deletion, we also confirmed by immunofluorescence that at least at the cellular level R-loops accumulate in a HCT116 human cell line deficient in BLM.

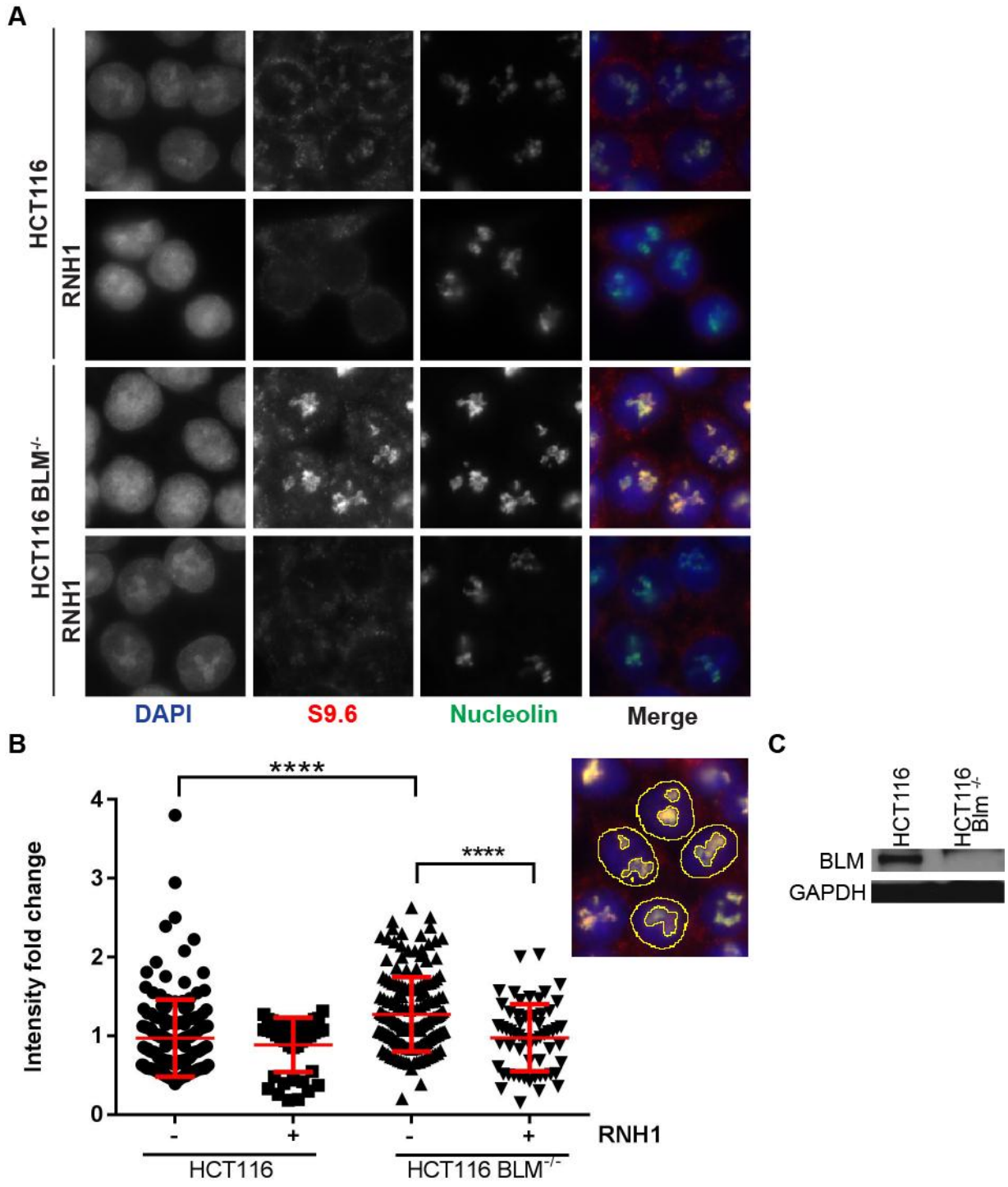

Figure 16 HCT116 BLM$^{-/-}$ cells accumulate R-loops. (**A**) Cells seeded on glass slides were fixed, treated with recombinant RNH1, and imaged (**B**) Relative S9.6 antibody fluorescence intensity per fixed nucleus. Mean and standard deviation from three independent experiments indicated in red. At least 30 cells were analyzed per condition per experiment. (****P<0.0001 by student's t-test) (**C**) Efficiency of BLM knockout.

Chapter 4: CONCLUDING CHAPTER

4.1 Conclusions and discussion

Synthetic genetic interaction experiments between *SGS1* and R-loop surveillance factors showed that *SGS1* is crucial for the survival of yeast known to accumulate R-loops. These findings are in alignment with previous identified interactions between *RNH201* and *SEN1*, and *SGS1*, and support the idea that DNA damage caused by R-loops is processed by homologous recombination repair proteins. The synergistic functional interactions between *RNH1*, *RNH201*, *MFT1*, *THP2* and *SEN1*, and *SGS1* seen in the GI assays also suggest that *SGS1* and R-loop suppressors have parallel roles in preventing instability.

Furthermore, it is possible that *SGS1* is acting upstream of R-loop induced damage by preventing R-loop formation considering that R-loops accumulate in *sgs1*Δ cells, and Sgs1 over-expression was found to suppress recombination in *mft1*Δ. Yet, further experiments are required to confirm the ability of Sgs1 to rescue R-loop induced GI as only one THO complex mutant was tested in this project. We predict that context and dose might be important considerations in the study of the effect of *SGS1* on R-loop surveillance mutants. We also don't expect *SGS1* overexpression to rescue R-loop induced GI for every R-loop surveillance factor. Sgs1 is a very versatile protein, which transiently partners with several proteins at the replisome (Rad53, Top2) to contribute to genome integrity. This suggests that *SGS1* overexpression might affect those interactions differently under diverse conditions (Bjergbaek, Cobb, Tsai-Pflugfelder, & Gasser, 2005). Moreover, Sgs1 forms a complex with *TOP3* and *RMI1*, known as STR, with functions in homologous recombination repair and active remodeling of stalled forks (Spies, 2013). Deletion of any of those genes leads to *sgs1Δ* like phenotypes showing that consideration of *SGS1*

physical interactions (stoichiometry, location, timing) is likely essential to the understanding of *SGS1*'s functional relationship to DNA:RNA hybrids.

This thesis also found that some of the instability observed in *sgs1*Δ is dependent on transcription, and RNH1 expression, both of which are well known modulators of R-loops. Specifically, we showed that the hyper-recombination phenotype characteristic of cell lacking *SGS1* was enhanced by long transcripts, high levels of transcription and active replication. These factors are permissive of replication-transcription collisions, which is a plausible mechanism of R-loop induced genome instability.

We also showed that rDNA instability in *sgs1*Δ is reversible by over-expressing human RNH1. Importantly, we found that R-loops levels in rDNA are higher in *sgs1*Δ than in wild type cells. Our data support a role for Sgs1 in mitigating transcription associated recombination and links instability in *sgs1*Δ cells to R-loop hotspots. Loss of Sgs1 slows replication through rDNA (Versini et al., 2003), and reduces rRNA transcription (Lee, Johnson, Yu, Prakash, & Prakash, 1999), therefore it is possible that *SGS1* maintains rDNA instability by preventing DNA:RNA hybrid formation in the rDNA locus. A similar mechanism has been proposed for TopI, which provides the major topoisomerase activity during rDNA transcription and is known to interact with Sgs1 (Lu et al., 1996). In strains lacking Top1, R-loops occur more frequently in rDNA, leading to Polymerase I arrests and pileups (El Hage et al., 2010).

R-loops also form in other repetitive DNA sequences including Ty elements and telomeres in wild type cells. *SGS1* has been implicated in the stability of those regions. For example, Ty1 mobility increases in *sgs1*Δ mutants due to increase recombination between extrachromosomal Ty1 cDNA molecules before or during integration (Bryk, Banerjee, Conte, & Curcio, 2001).

Since our collaborators DRIP data showed further R-loop accumulation in *sgs1*Δ, it will be interesting to determine if Ty1 mobility is sensitive to human *RNH1* overexpression in *sgs1Δ* cells.

Mechanism of Sgs1 action at R-loops

Our data suggest that replication-transcription collisions are a cause of enhanced instability in *sgs1*Δ cells and that R-loops accumulate at sites which become additionally difficult to replicate in *sgs1*Δ cells (*i.e.* repetitive elements). This adds to existing evidence suggesting that when a moving fork approaches an R-loop, replication stress arises. R-loops might impose mobility constraints in the DNA, and thus further increase the topological tension generated when forks converge with transcription. The structural consequence of the increased topological tension is speculative but this work supports the idea that *SGS1* counteracts the resulting GI.

However, these observations are not enough to determine the exact mechanisms in which Sgs1 counteracts R-loops. Given the many functions of Sgs1, this protein could reduce R-loop levels by several modes of action including its helicase or translocase activities or its non-catalytic protein binding sites. Sgs1 is mostly expressed in S-phase, where it travels with the replication fork and stabilizes it during stalling (Ashton & Hickson, 2010; Bohm & Bernstein, 2014; Chaudhury, Sareen, Raghunandan, & Sobeck, 2013). This fits well with our model, in which Sgs1 prevents R-loop induce damage by preventing R-loop and replication forks collisions. However, even if we propose that the important role of Sgs1 is in S-phase, Sgs1 could be acting directly on the R-loop structure, on the replication fork itself, or it could be playing a role in removing secondary structures that stabilize R-loops (**Figure 17, B**). For example, G-quadruplexes forming in the exposed non-template DNA of an R-loop (Duquette, Handa, Vincent, Taylor, & Maizels, 2004; Kim & Jinks-Robertson, 2011). Finally, Sgs1 could help

process R-loops by activating other helicases. This is plausible in yeast, and in human cells, BLM is known to physically and functionally coordinate its activity with helicases FANCM and FANCJ (Brosh, 2013; Deans & West, 2009; Meetei et al., 2003).

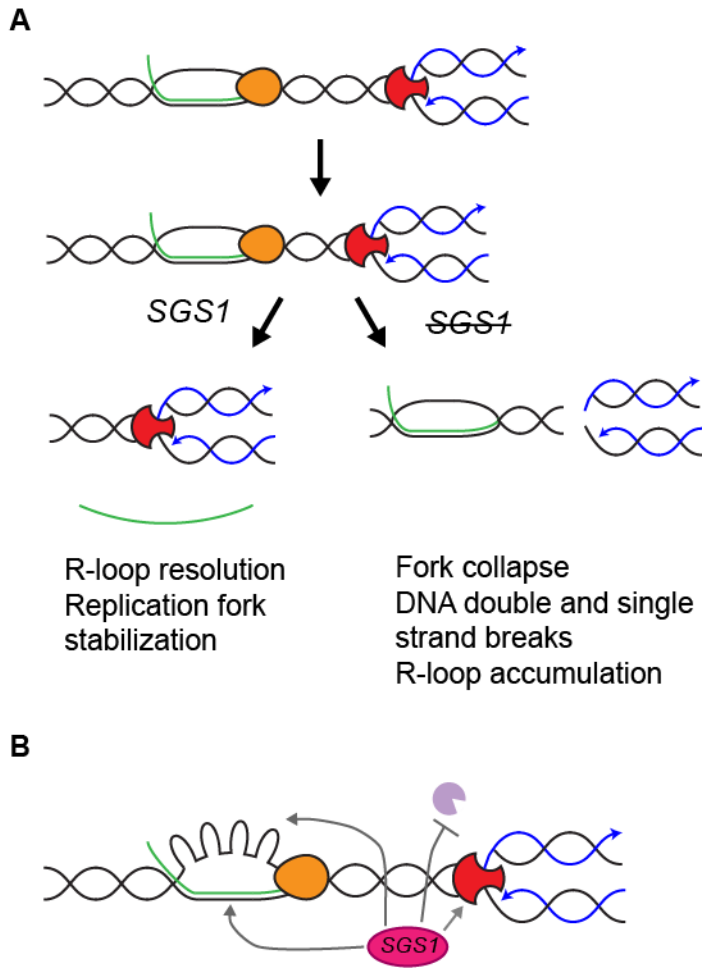

Figure 17 Model explaining how SGS1 counteracts R-loop induced genome instability(**A**) In the absence of the *SGS1*, DNA:RNA hybrids accumulate. Conflicts between replication and transcription result in activation of the DNA damage response, replication fork collapse DNA lesions, and genomic instability. In the presence of *SGS1*, stalled replication forks are stabilized. (**B**) *SGS1* might resolve replication blocks consisting of R-loops via its helicase activity acting directly on the DNA:RNA hybrid or the secondary structure on the displaced non-template strand. *SGS1*might also inhibit the action of nucleases on stalled forks. Finally, *SGS1* might aid in the restart of stalled forks by recruitment of other R-loop resolving proteins. Nascent DNA and RNA are indicated as blue and green solid lines, respectively; arrowheads are at the 3' ends of DNA strands. RNA polymerase is in orange and the replisome in red.

It would be interesting to repeat some of the experiments shown in this work with cells arrested at different cell cycle stages to begin to discriminate between the possible roles of *SGS1* in R-loop prevention. The response of *sgs1*Δ GI phenotypes to modulators of R-loops shown in this work might emerge stronger in S phase arrested cells as the role of *SGS1* in R-loop biology might be restricted to replicating DNA and fork restarting. Also, the response of fork defects in *sgs1*Δ to RNaseH1 overexpression and transcription repression should be tested. The rDNA locus could serve as a good model to better understand the mechanism by which *SGS1* might be counteracting R-loop induced GI. To address some of the technical challenges with these approaches, one of our strategies moving forward is to obtain constructs from a collaborator in which Sgs1 is fused to domains derived from cyclins that restrict Sgs1 expression to G1, S, or G2/M cell cycle phases.

R-loop removal by RNaseH in the context of *sgs1*Δ

Our results also suggest that other approaches are required to study Sgs1 to clarify the confounding aspects from our *RNH1* overexpression data. First, different effects where observed when full-length yeast or human *RNH1* were overexpressed in the *SGS1* deletion mutants for the rDNA instability assay. While reversibility of *sgs1*Δ's rDNA instability was seen after human RNaseH1 overexpression, no repression was seen when the yeast RNH1 overexpression vector was used (**Figure 18**).

Yeast-*RNH1* overexpression did reduce the accumulation of Rad52 foci in *sgs1Δrnh201Δ* but it did not reduce the recombination phenotype of *sgs1*Δ or *mft1*Δ (**Figure 19**). This demonstrates that yeast *RNH1* overexpression might not reverse R-loop formation under certain context even for mutants known to have transcription dependent GI. There might be several explanations for this. For example, foci are transient, while recombination events are fixed. It could be that *RNH1*

is able to reduce damage globally but not at the region tested in the recombination substrate. Another example where yeast *RNH1* overexpression did not reverse an R-loop dependent phenotype comes from the work of the laboratory of Dr. Crouch. Chon *et al.* (Chon et al., 2013) showed that RNaseH1 overexpression under the control of the strong triosephosphate isomerase (TPI) promoter did not restore the growth defect of a *sgs1*Δ*rnh201*Δ mutant strain to the rate of the single mutant *sgs1*Δ, suggesting that RNaseH1 cannot act on or have access to the RNaseH2 substrates that accumulate in the absence of Sgs1. Indeed, more recent work has demonstrated that yeast Rnh1only acts on a subset of endogenous DNA:RNA hybrids in the genome, although the regulation of this is unclear (Zimmer & Koshland, 2016). The human RNaseH1 construct may not be subject to the same regulatory mechanisms as yeast Rnh1 in yeast cells. A better understanding of the toxicity and specificity of yeast and human *RNH1*, and the mechanism of *RNH1* R-loop removal would allow better interpretation of our results.

Also, plasmids containing yeast and human *RNH1* sequences were used to overexpress that protein in the context of ALF assay. In that assay, neither version of the *RNH1* protein was able to reverse the ALF phenotype (**Figure 20**). This result can be explained by the fact that the ALF phenotype can take place due to several events including chromosome loss, which is not expected to be caused by R-loop accumulation. Perhaps the role of Sgs1 in R-loop metabolism is not global but dependent on genome context and requires assays that are sensitive to small increases in DSBs or HR events.

Given our data, we propose that Sgs1 contributes to preventing or resolving R-loops and we speculate that it does so by promoting resumption of stalled RFs at R-loop prone regions, through its known role in replication restart. However, we cannot disregard the possibility that *SGS1* might directly act on DNA:RNA hybrids to destabilize R-loop structures. *In vitro* evidence

suggests that BLM is certainly capable of unwinding R-loops (Popuri et al., 2008). Having taken into consideration the large number of R-loop prone sites, it is possible that more than one of the mentioned Sgs1 R-loop processing mechanisms are at play in the genome. If Sgs1 is having an indirect effect on R-loop metabolism by stabilizing stalled replication forks, Sgs1 creates time for other R-loop resolution factors to act. Indeed, the abundance of fork-protection factors emerging as R-loop regulators supports a generalized concept that functional DNA replication machinery is an important way to mitigate deleterious transcription-replication collisions (Helmrich, Ballarino, Nudler, & Tora, 2013). Our preliminary data on the HCT116 cell line also supports the idea that BLM might have a role similar to *SGS1* in the regulation of R-loops.

4.2 Significance

R-loops accumulate and lead to GI in response to lack of Sgs1, which suggests a role for yet another DNA repair proteins in R-loop metabolism. This study adds to a growing body of literature that implicates tumor suppressors orthologs and DNA repair proteins in R-loop biology (Bhatia et al., 2014; Hatchi et al., 2015; Schwab et al., 2015). This thesis also furthers our understanding of *SGS1* identifying R-loop surveillance as another pathway through which Sgs1 contributes to genomic integrity.

4.3 Future directions

More biochemical and genetics experiments are required to dissect the mechanisms by which Sgs1 might be preventing the formation of R-loops. A major challenge for the interpretation of our data was our inability to distinguish between a possible new role of Sgs1 and the consequences of the disruption of the canonical functions of Sgs1 in our deletion mutants. One way to further study Sgs1 in the context of R-loops would be to create a collection of mutants

lacking different protein domains. These mutants could be used to test which domains are required for R-loop accumulation.

Another approach to dissect Sgs1's possible R-loop function would be to test if partner DNA repair proteins might be involved in R-loop resolution. For example, enzymes of the STR (or BTR, BLM-Topo3-Rmi1, in humans) complex (Spies, 2013). Sgs1 might require the interaction with other proteins to process R-loops. Testing the effect of R-loop modulators in Sgs1 protein partners might help identify if other proteins are required for R-loop metabolism. This research would also pinpoint which Sgs1-protein interactions might be important in the context of R-loop prevention and therefore which Sgs1 led mechanisms might be at play during DNA:RNA hybrid processing. In addition it would be important to investigate exactly which Sgs1 phenotypes are *RNH1* reversible and under which conditions, as this would allow for a systematic mapping of the Sgs1'sactivities that are dependent on DNA:RNA accumulation.

The results of the suggested experiments together with the data present here should guide similar experiment in mammalian cell lines to investigate if BLM also has a similar R-loop counteracting role in humans. It would be interesting to see if this is true for other cell lines and patient samples.

Bibliography

Aguilera, A., & Garcia-Muse, T. (2012). R loops: from transcription byproducts to threats to genome stability. *Mol Cell, 46*(2), 115-124. doi: 10.1016/j.molcel.2012.04.009

Alvaro, D., Lisby, M., & Rothstein, R. (2007). Genome-wide analysis of Rad52 foci reveals diverse mechanisms impacting recombination. *PLoS Genet, 3*(12), e228. doi: 10.1371/journal.pgen.0030228

Aristizabal, Maria J. (2017). Unpublished observations.

Ashton, T. M., & Hickson, I. D. (2010). Yeast as a model system to study RecQ helicase function. *DNA Repair (Amst), 9*(3), 303-314. doi: 10.1016/j.dnarep.2009.12.007

Bell, D. W., Sikdar, N., Lee, K. Y., Price, J. C., Chatterjee, R., Park, H. D., . . . Myung, K. (2011). Predisposition to cancer caused by genetic and functional defects of mammalian Atad5. *PLoS Genet, 7*(8), e1002245. doi: 10.1371/journal.pgen.1002245

Bernstein, K. A., Shor, E., Sunjevaric, I., Fumasoni, M., Burgess, R. C., Foiani, M., . . . Rothstein, R. (2009). Sgs1 function in the repair of DNA replication intermediates is separable from its role in homologous recombinational repair. *EMBO J, 28*(7), 915-925. doi: 10.1038/emboj.2009.28

Bjergbaek, L., Cobb, J. A., Tsai-Pflugfelder, M., & Gasser, S. M. (2005). Mechanistically distinct roles for Sgs1p in checkpoint activation and replication fork maintenance. *EMBO J, 24*(2), 405-417. doi: 10.1038/sj.emboj.7600511

Bochman, M. L., Paeschke, K., Chan, A., & Zakian, V. A. (2014). Hrq1, a homolog of the human RecQ4 helicase, acts catalytically and structurally to promote genome integrity. *Cell Rep, 6*(2), 346-356. doi: 10.1016/j.celrep.2013.12.037

Bohm, S., & Bernstein, K. A. (2014). The role of post-translational modifications in fine-tuning BLM helicase function during DNA repair. *DNA Repair (Amst), 22*, 123-132. doi: 10.1016/j.dnarep.2014.07.007

Britton, S., Dernoncourt, E., Delteil, C., Froment, C., Schiltz, O., Salles, B., . . . Calsou, P. (2014). DNA damage triggers SAF-A and RNA biogenesis factors exclusion from chromatin coupled to R-loops removal. *Nucleic Acids Res, 42*(14), 9047-9062. doi: 10.1093/nar/gku601

Castellano-Pozo, M., Garcia-Muse, T., & Aguilera, A. (2012). R-loops cause replication impairment and genome instability during meiosis. *EMBO Rep, 13*(10), 923-929. doi: 10.1038/embor.2012.119

Castellano-Pozo, M., Santos-Pereira, J. M., Rondon, A. G., Barroso, S., Andujar, E., Perez-Alegre, M., . . . Aguilera, A. (2013). R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol Cell, 52*(4), 583-590. doi: 10.1016/j.molcel.2013.10.006

Cerritelli, S. M., & Crouch, R. J. (2009). Ribonuclease H: the enzymes in eukaryotes. *FEBS J, 276*(6), 1494-1505. doi: 10.1111/j.1742-4658.2009.06908.x

Chan, Y. A., Aristizabal, M. J., Lu, P. Y., Luo, Z., Hamza, A., Kobor, M. S., . . . Hieter, P. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. *PLoS Genet, 10*(4), e1004288. doi: 10.1371/journal.pgen.1004288

Chan, Y. Alina. (2014). *Unpublished observations*

Chaudhury, I., Sareen, A., Raghunandan, M., & Sobeck, A. (2013). FANCD2 regulates BLM complex functions independently of FANCI to promote replication fork recovery. *Nucleic Acids Res, 41*(13), 6444-6459. doi: 10.1093/nar/gkt348

Cho, J. E., Kim, N., & Jinks-Robertson, S. (2015). Topoisomerase 1-dependent deletions initiated by incision at ribonucleotides are biased to the non-transcribed strand of a highly activated reporter. *Nucleic Acids Res, 43*(19), 9306-9313. doi: 10.1093/nar/gkv824

Chon, H., Sparks, J. L., Rychlik, M., Nowotny, M., Burgers, P. M., Crouch, R. J., & Cerritelli, S. M. (2013). RNase H2 roles in genome integrity revealed by unlinking its activities. *Nucleic Acids Res, 41*(5), 3130-3143. doi: 10.1093/nar/gkt027

Costantino, L., & Koshland, D. (2015). The Yin and Yang of R-loop biology. *Curr Opin Cell Biol, 34*, 39-45. doi: 10.1016/j.ceb.2015.04.008

Dominguez-Sanchez, M. S., Barroso, S., Gomez-Gonzalez, B., Luna, R., & Aguilera, A. (2011). Genome instability and transcription elongation impairment in human cells depleted of THO/TREX. *PLoS Genet, 7*(12), e1002386. doi: 10.1371/journal.pgen.1002386

Duquette, M. L., Handa, P., Vincent, J. A., Taylor, A. F., & Maizels, N. (2004). Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev, 18*(13), 1618-1629. doi: 10.1101/gad.1200804

El Hage, A., French, S. L., Beyer, A. L., & Tollervey, D. (2010). Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev, 24*(14), 1546-1558. doi: 10.1101/gad.573310

El Hage, A., Webb, S., Kerr, A., & Tollervey, D. (2014). Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. *PLoS Genet, 10*(10), e1004716. doi: 10.1371/journal.pgen.1004716

Frei, C., & Gasser, S. M. (2000). The yeast Sgs1p helicase acts upstream of Rad53p in the DNA replication checkpoint and colocalizes with Rad53p in S-phase-specific foci. *Genes Dev, 14*(1), 81-96.

Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J. L., & Li, X. (2011). R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev, 25*(19), 2041-2056. doi: 10.1101/gad.17010011

Gari, E., Piedrafita, L., Aldea, M., & Herrero, E. (1997). A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae. *Yeast, 13*(9), 837-848. doi: 10.1002/(SICI)1097-0061(199707)13:9<837::AID-YEA145>3.0.CO;2-T

Gavalda, S., Santos-Pereira, J. M., Garcia-Rubio, M. L., Luna, R., & Aguilera, A. (2016). Excess of Yra1 RNA-Binding Factor Causes Transcription-Dependent Genome Instability, Replication Impairment and Telomere Shortening. *PLoS Genet, 12*(4), e1005966. doi: 10.1371/journal.pgen.1005966

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

Gomez-Gonzalez, B., Garcia-Rubio, M., Bermejo, R., Gaillard, H., Shirahige, K., Marin, A., . . . Aguilera, A. (2011). Genome-wide function of THO/TREX in active genes prevents R-loop-

dependent replication obstacles. *The EMBO journal*. doi: 10.1038/emboj.2011.206; 10.1038/emboj.2011.206

Gonzalez-Aguilera, C., Tous, C., Gomez-Gonzalez, B., Huertas, P., Luna, R., & Aguilera, A. (2008). The THP1-SAC3-SUS1-CDC31 complex works in transcription elongation-mRNA export preventing RNA-mediated genome instability. *Mol Biol Cell, 19*(10), 4310-4318. doi: 10.1091/mbc.E08-04-0355

Hamperl, S., & Cimprich, K. A. (2014). The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair (Amst), 19*, 84-94. doi: 10.1016/j.dnarep.2014.03.023

Helmrich, A., Ballarino, M., Nudler, E., & Tora, L. (2013). Transcription-replication encounters, consequences and genomic instability. *Nat Struct Mol Biol, 20*(4), 412-418. doi: 10.1038/nsmb.2543

Huertas, P., & Aguilera, A. (2003). Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol Cell, 12*(3), 711-721.

Kim, N., & Jinks-Robertson, S. (2011). Guanine repeat-containing sequences confer transcription-dependent instability in an orientation-specific manner in yeast. *DNA Repair (Amst), 10*(9), 953-960. doi: 10.1016/j.dnarep.2011.07.002

Larsen, N. B., & Hickson, I. D. (2013). RecQ Helicases: Conserved Guardians of Genomic Integrity. *Adv Exp Med Biol, 767*, 161-184. doi: 10.1007/978-1-4614-5037-5_8

Lee, S. K., Johnson, R. E., Yu, S. L., Prakash, L., & Prakash, S. (1999). Requirement of yeast SGS1 and SRS2 genes for replication and transcription. *Science, 286*(5448), 2339-2342.

Li, X., & Manley, J. L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell, 122*(3), 365-378. doi: 10.1016/j.cell.2005.06.008

Mankouri, H. W., Craig, T. J., & Morgan, A. (2002). SGS1 is a multicopy suppressor of srs2: functional overlap between DNA helicases. *Nucleic Acids Res, 30*(5), 1103-1113.

McLellan, J., O'Neil, N., Tarailo, S., Stoepel, J., Bryan, J., Rose, A., & Hieter, P. (2009). Synthetic lethal genetic interactions that decrease somatic cell proliferation in Caenorhabditis elegans identify the alternative RFC CTF18 as a candidate cancer drug target. *Mol Biol Cell, 20*(24), 5306-5313. doi: 10.1091/mbc.E09-08-0699

Mischo, H. E., Gomez-Gonzalez, B., Grzechnik, P., Rondon, A. G., Wei, W., Steinmetz, L., . . . Proudfoot, N. J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol Cell, 41*(1), 21-32. doi: 10.1016/j.molcel.2010.12.007

Mundbjerg, K., Jorgensen, S. W., Fredsoe, J., Nielsen, I., Pedersen, J. M., Bentsen, I. B., . . . Andersen, A. H. (2015). Top2 and Sgs1-Top3 Act Redundantly to Ensure rDNA Replication Termination. *PLoS Genet, 11*(12), e1005697. doi: 10.1371/journal.pgen.1005697

Naim, V., & Rosselli, F. (2009). The FANC pathway and BLM collaborate during mitosis to prevent micro-nucleation and chromosome abnormalities. *Nat Cell Biol, 11*(6), 761-768. doi: 10.1038/ncb1883

Novoa, C., Ang, S. J., Stirling, P.C.. . ((In press, 2017, 16 pages)). The A-Like Faker Assay for Measuring Yeast Chromosome III Stability. *Methods in Molecular Biology*

Paulsen, R. D., Soni, D. V., Wollman, R., Hahn, A. T., Yee, M. C., Guan, A., . . . Cimprich, K. A. (2009). A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol Cell, 35*(2), 228-239. doi: 10.1016/j.molcel.2009.06.021

Pichierri, P., Franchitto, A., & Rosselli, F. (2004). BLM and the FANC proteins collaborate in a common pathway in response to stalled replication forks. *EMBO J, 23*(15), 3154-3163. doi: 10.1038/sj.emboj.7600277

Pikor, L., Thu, K., Vucic, E., & Lam, W. (2013). The detection and implication of genome instability in cancer. *Cancer Metastasis Rev, 32*(3-4), 341-352. doi: 10.1007/s10555-013-9429-5

Popuri, V., Bachrati, C. Z., Muzzolini, L., Mosedale, G., Costantini, S., Giacomini, E., . . . Vindigni, A. (2008). The Human RecQ helicases, BLM and RECQ1, display distinct DNA substrate specificities. *J Biol Chem, 283*(26), 17766-17776. doi: 10.1074/jbc.M709749200

Prado, F., & Aguilera, A. (1995). Role of reciprocal exchange, one-ended invasion crossover and single-strand annealing on inverted and direct repeat recombination in yeast: different requirements for the RAD1, RAD10, and RAD52 genes. *Genetics, 139*(1), 109-123.

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

Prado, F., Piruat, J. I., & Aguilera, A. (1997). Recombination between DNA repeats in yeast hpr1delta cells is linked to transcription elongation. *EMBO J, 16*(10), 2826-2835. doi: 10.1093/emboj/16.10.2826

Santos-Pereira, J. M., Herrero, A. B., Garcia-Rubio, M. L., Marin, A., Moreno, S., & Aguilera, A. (2013). The Npl3 hnRNP prevents R-loop-mediated transcription-replication conflicts and genome instability. *Genes Dev, 27*(22), 2445-2458. doi: 10.1101/gad.229880.113

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods, 9*(7), 671-675.

Sinclair, D. A., Mills, K., & Guarente, L. (1997). Accelerated aging and nucleolar fragmentation in yeast sgs1 mutants. *Science, 277*(5330), 1313-1316.

Skourti-Stathaki, K., & Proudfoot, N. J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev, 28*(13), 1384-1396. doi: 10.1101/gad.242990.114

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. *Mol Cell, 42*(6), 794-805. doi: 10.1016/j.molcel.2011.04.026

Sollier, J., & Cimprich, K. A. (2015). Breaking bad: R-loops and genome integrity. *Trends Cell Biol, 25*(9), 514-522. doi: 10.1016/j.tcb.2015.05.003

Sollier, J., Stork, C. T., Garcia-Rubio, M. L., Paulsen, R. D., Aguilera, A., & Cimprich, K. A. (2014). Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell, 56*(6), 777-785. doi: 10.1016/j.molcel.2014.10.020

Stirling, P. C., Bloom, M. S., Solanki-Patil, T., Smith, S., Sipahimalani, P., Li, Z., . . . Hieter, P. (2011). The Complete Spectrum of Yeast Chromosome Instability Genes Identifies Candidate

CIN Cancer Genes and Functional Roles for ASTRA Complex Components. *PLoS genetics, 7*(4), e1002057. doi: 10.1371/journal.pgen.1002057

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

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

Stratton, M. R., Campbell, P. J., & Futreal, P. A. (2009). The cancer genome. *Nature, 458*(7239), 719-724. doi: 10.1038/nature07943

Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J., & Dean, C. (2013). R-loop stabilization represses antisense transcription at the Arabidopsis FLC locus. *Science, 340*(6132), 619-621. doi: 10.1126/science.1234848

Tresini, M., Warmerdam, D. O., Kolovos, P., Snijder, L., Vrouwe, M. G., Demmers, J. A., . . . Marteijn, J. A. (2015). The core spliceosome as target and effector of non-canonical ATM signalling. *Nature, 523*(7558), 53-58. doi: 10.1038/nature14512

Tuduri, S., Crabbe, L., Conti, C., Tourriere, H., Holtgreve-Grez, H., Jauch, A., . . . Pasero, P. (2009). Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat Cell Biol, 11*(11), 1315-1324. doi: 10.1038/ncb1984

Versini, G., Comet, I., Wu, M., Hoopes, L., Schwob, E., & Pasero, P. (2003). The yeast Sgs1 helicase is differentially required for genomic and ribosomal DNA replication. *EMBO J, 22*(8), 1939-1949. doi: 10.1093/emboj/cdg180

Wahba, L., Amon, J. D., Koshland, D., & Vuica-Ross, M. (2011). RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol Cell, 44*(6), 978-988. doi: 10.1016/j.molcel.2011.10.017; 10.1016/j.molcel.2011.10.017

Wahba, L., Costantino, L., Tan, F. J., Zimmer, A., & Koshland, D. (2016). S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev, 30*(11), 1327-1338. doi: 10.1101/gad.280834.116

Wahba, L., & Koshland, D. (2013). The Rs of biology: R-loops and the regulation of regulators. *Mol Cell, 50*(5), 611-612. doi: 10.1016/j.molcel.2013.05.024

Wellinger, R. E., Prado, F., & Aguilera, A. (2006). Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex. *Mol Cell Biol, 26*(8), 3327-3334. doi: 10.1128/MCB.26.8.3327-3334.2006

Yang, Y., McBride, K. M., Hensley, S., Lu, Y., Chedin, F., & Bedford, M. T. (2014). Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. *Mol Cell, 53*(3), 484-497. doi: 10.1016/j.molcel.2014.01.011

Yuen, K. W., Warren, C. D., Chen, O., Kwok, T., Hieter, P., & Spencer, F. A. (2007). Systematic genome instability screens in yeast and their potential relevance to cancer. *Proc Natl Acad Sci U S A, 104*(10), 3925-3930. doi: 10.1073/pnas.0610642104

Appendix A: Supplementary figures

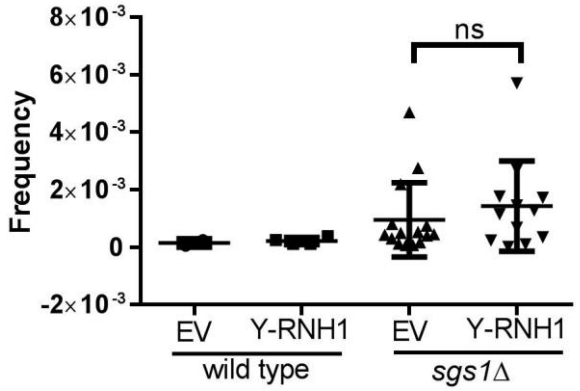

Figure 18 Over-expression of yeast RNaseH1 fails to rescue rDNA instability in *sgs1*Δ Frequencies of excision of a *URA3* marker inserted into the rDNA locus as measured by the appearance of FOA resistant yeast colonies in wild type, and *sgs1*Δ strains. Means and standard errors are indicated.

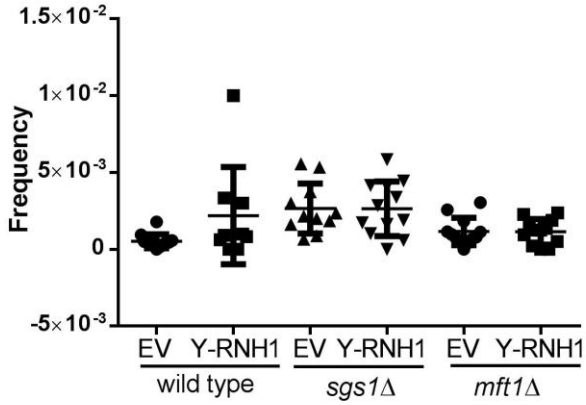

Figure 19 Over-expression of yeast RNaseH1 fails to rescue recombination in plasmid borne system Recombination frequencies for wild type, *sgs1*Δ and *mft1*Δ strains using the LYΔNS substrate. Strains tested were carrying a Gal expression plasmid either containing the full sequence of yeast RNaseH1(Y-RNH1) or no sequence at all (EV). Means and standard errors are shown.

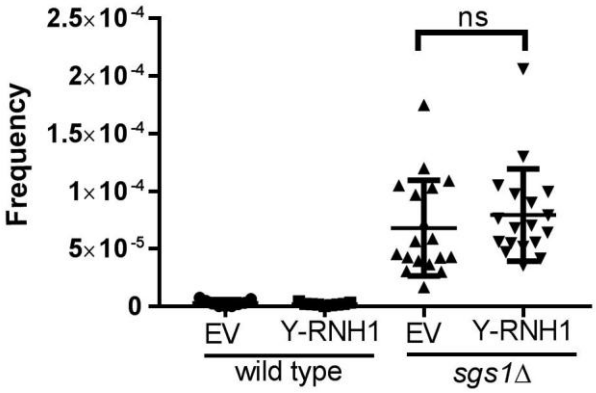

Figure 20 Overexpression of full length yeast RNaseH1 has no effect on *sgs1*Δ ALF phenotype Scatterplot of frequency of *MAT* locus instability in various mutants. The mean and standard error (whiskers) of at least nine mating type conversion tests from independent colonies were analyzed. Strains tested were carrying a Gal expression plasmid either containing the full sequence of yeast RNaseH1 (Y-RNH1) or no sequence at all (EV). Gene induction was achieved by incubating logarithmic phase cultures in 2% raffinose complemented with 2% galactose for 4 hours.