DISCOVERY, CHARACTERIZATION AND MECHANISM OF RNA

AND cDNA-MEDIATED DNA DOUBLE-STRAND BREAK REPAIR

A Dissertation
Presented to
The Academic Faculty

by

Havva Keskin

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Biological Science

Georgia Institute of Technology
May 2017

DISCOVERY, CHARACTERIZATION AND MECHANISM OF RNA

AND cDNA-MEDIATED DNA DOUBLE-STRAND BREAK REPAIR

Approved by:

Dr. Francesca Storici, Advisor
School of Biological Science
Georgia Institute of Technology

Dr. Yuhong Fan
School of Biological Science
Georgia Institute of Technology

Dr. Yury Chernoff
School of Biological Science
Georgia Institute of Technology

Dr. Milo Fasken
Department of Biochemistry
Emory University

Dr. Kirill Lobachev
School of Biological Science
Georgia Institute of Technology

Date Approved: 04.05.2017

For my beloved family and friends

ACKNOWLEDGEMENTS

I would like to thank all the staff members in School of Biological Science, especially Kevin Roman.

I would also like to thank all my friends for their support.

Lastly, I would like to thank to my bellowed family, who always support me to come here in USA to study and encourage me to do best in my life. I would like to especially thank to my mom who came to USA help me to finish this thesis. I would like to thank to my lovely husband for being helpful, supportive and understanding during my Ph.D journey. And I would like to thank my son, Kerem, for entering my world and make it beautiful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AGS	Aicardi Goutieres Syndrome
AI	Artificial intron
cDNA	Complementary DNA
DNA	Deoxyribonucleic acid
ds	double strand
DSB	Double strand break
dsDNA	Double strand DNA
GFP	Green florescent protein
HEK	Human embryonic kidney cells
HO	Homothallic-switching endonuclease
HR	Homologous recombination
LINE-1	Long intersperse element-1
LTR	Long terminal repeats
MAT	Mating type locus
mg	milligrams
mL	Milliliters
mM	Millimolar
mtDNA	Mitochondrial DNA
NAT	Nourseothricin
NHEJ	Non-homologous end joining
nmol	Nanomole
NTD	N-terminal domain

Oligo	Oligonucleotide
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PFA	Foscarnate
pGAL1	Galactose inducible promoter
PIP	PCNA interacting peptide
qPCR	Quantitative real time PCR
RNA	Ribonucleic acid
RNase H	Ribonuclease H
rNMP	Ribonucleotide
RPA	Replication protein A
RT	Reverse transcriptase
SAGA	Spt-Ada-Gcn5-acetyltransferase
ss	single strand
SSA	Single strand annealing
ssDNA	single strand DNA
ssRNA	single strand RNA
Ty	Transposon of Yeast
w/v	Weight-to-volume
°	Degrees
°C	Celsius degrees
Δ	Deletion
μL	Microliters
μM	Micromolar

SUMMARY

A double-strand break (DSB) is one of the most deleterious DNA lesions and its repair is crucial for genome stability. Even if a single DSB is not repaired precisely, this could cause mutations, chromosomal rearrangements, cell death, and apoptosis. The safest mechanism to repair a DSB is homologous recombination (HR). HR requires an identical or nearly identical DNA template, such as a sister chromatid or a homologous chromosome to retrieve the missing genetic information and accomplish error-free repair. In special cases, HR can occur between RNA molecules, such as RNA molecules in RNA viruses. However, very little is known about RNA-DNA HR. Previously, it was demonstrated that synthetic RNA-containing molecules can serve as templates for repairing defective or broken homologous chromosomal DNA in yeast, human and bacterial cells, but it remained unclear whether cellular RNA transcripts can recombine with genomic DNA. Here, we investigated whether yeast cells can use transcript RNA as a template to repair a chromosomal DSB either directly or indirectly, if the RNA is converted first into a DNA copy, cDNA. We developed a system to detect HR between chromosomal DNA and transcript RNA in budding yeast, *Saccharomyces cerevisiae*. We focused on repair of a chromosomal DSB occurring either in a homologous but remote locus (*trans*) or in the same transcript-generating locus (*cis*) in yeast. We proved that transcript RNA can repair a DSB indirectly, via cDNA. Moreover, we found that cDNA repair is much more frequent in the *trans* than in the *cis* system. Interestingly, in the absence of Ribonuclease H1 and H2 (RNases H1 and H2), we could detect DSB repair even in conditions that strongly inhibit cDNA formation, suggesting direct DSB repair by

transcript RNA. In contrast to DSB repair by cDNA, the direct DSB repair by transcript RNA is more efficient in the *cis* than in the *trans* system, despite the higher abundance of the transcript in the *trans* system. These results suggest that the vicinity of the transcript RNA to the break site in the *cis* system may facilitate DSB repair. DSB repair by transcript RNA in *cis* is promoted by the HR protein Rad52 but not Rad51, in agreement with the demonstration that the yeast and human Rad52 proteins efficiently catalyze annealing of RNA to a DSB-like DNA end *in vitro*. We also showed that yeast cells expressing hypomorphic mutants of RNase H2, which correspond to the human RNase H2 mutants that are associated with the neuroimmunological disease, Aicardi Goutieres (AGS) syndrome, have increased frequency of DSB repair by cDNA, significantly higher than in wild-type RNase H2 cells. In addition, we showed that in contrast to DSB repair by single strand DNA (ssDNA) oligonucleotides (oligos), RNA templated DSB repair is not dependent on factors that are major players in DNA end resection. This result could be explained by a mechanism in which transcript RNA repairs a DSB in conditions of limited end resection via an inverse strand exchange reaction. Our study provides proof and initial characterization of a new mechanism of DNA repair and HR mediated by RNA in yeast, and unravels novel aspects in the complex relationship between RNA and DNA in genome stability.

CHAPTER 1

INTRODUCTION

1.1 Double-strand break and its repair

Double-strand breaks (DSBs) are among the most dangerous DNA lesions, which can be generated by endogenous or exogenous sources, and can lead to mutations and genome rearrangement if not properly repaired [1]. Their repair is crucial for cell survival and genome stability [2]. There are two predominant mechanisms for DSB repair: non-homologous end-joining (NHEJ) and homologous recombination (HR) [3, 4]. NHEJ does not depend on a homologous template and it is predominant in the G_0/G_1 phase of the cell cycle in yeast [5]. Once a DSB is formed, Ku proteins (Ku70/80) bind to the broken ends [6] and lead to ligation of the broken ends by Dnl4-Lif1 (**Figure 1.1A**) [7]. This process is mostly error prone because of the possibility of introducing small deletions and insertions at the site of the DSB [1].

HR is the repair mechanism in which homologous or homeologous DNA molecules interact with each other and exchange genetic information. HR is a major repair pathway in the S and G_2 phases of the cell cycle allowing for most accurate repair templated by sister chromatids in mitotic cells [8]. Because HR uses a template with identical or nearly identical sequence for repair of DSBs, HR is considered a less error prone mechanism than NHEJ [9]. After a DSB is formed, the broken DNA ends are resected in a 5' to 3' direction, resulting in 3' overhang single-stranded DNA (ssDNA) ends that are used for homology search to find an homologous intact sequence [10]. Four nucleases, MRX

complex (Mre11-Rad50-Xrs2), Exo1, Dna2, and Sae2 and one helicase, Sgs1, are mainly

involved in the end resection process [11]. Resection inhibits NHEJ and triggers the HR

mechanism for DSB repair [12]. Replication Protein A (RPA) binds to the resected

ssDNA and prevents the formation of secondary structures. Following RPA binding,

Rad51 binds to ssDNA and displaces RPA with the help of Rad52 to form the Rad51

filament for homology search and DNA strand invasion for repair (**Figure 1.1B**) [13].
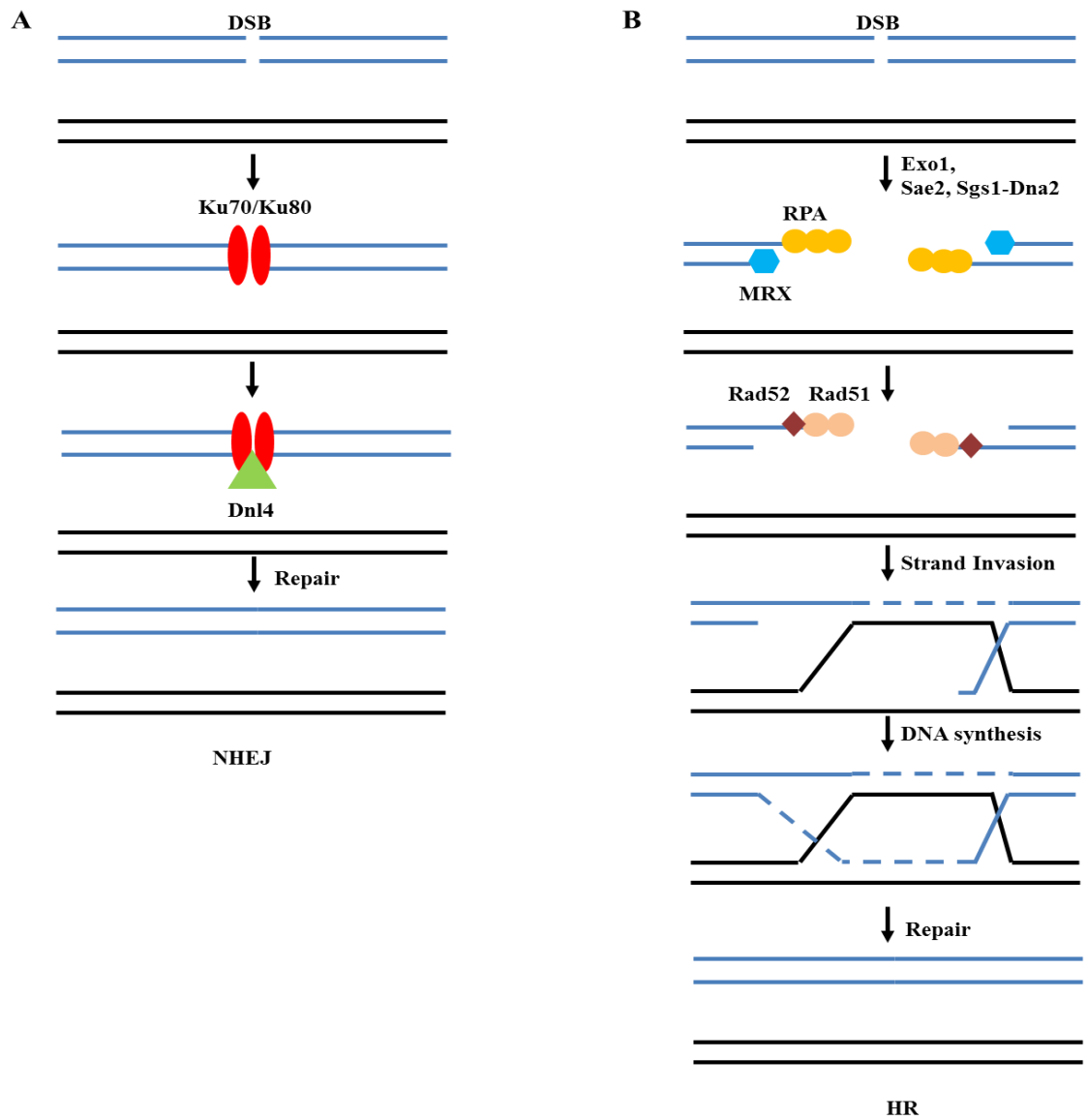
Figure 1.1 Major mechanisms for DSB repair. A) NHEJ mechanism. Ku70/Ku80 are in red. Dnl4 is in light green. **B)** HR mechanism. RPA is in yellow. MRX is in light blue. Rad51 is in light orange. Rad52 is in dark red. Repaired DNA is in blue and homologous DNA is in black. (**Modified from [14]**)

1.2 Evidence in support of indirect RNA-mediated DNA DSB repair

RNA is an abundant molecule in the cell and could be another source of homologous template for DNA DSB repair via HR in addition to a sister chromatid or a homologous chromosome. Reverse transcriptase (RT) enzymes utilize RNA as a template to create complementary DNA (cDNA) and can transfer genetic information back to DNA in retroviruses, retrotransposons and telomeres [15, 16]. Derr *et al*. and Curcio *et al.* showed that the retrotransposon of yeast (Ty) can reverse transcribe not only Ty RNA but also other cellular RNA to produce cDNA by RT enzyme [17-19]. Such cDNA can be integrated into DNA, recombine with homologous DNA [17-19] or be captured at the site of a DSB via NHEJ [20, 21]. In mammalian cells, the Long Interspersed Element-I (LINE-1) can incorporate endogenous mRNA sequences at the site of a DSB in addition to LINE-1 mRNA [22, 23].

Moore *et al.* studied retrotransposon activity by expressing Homothallic-switching endonuclease (HO) to induce a DSB at the Mating Type (MAT) locus in budding yeast deficient for both *HML* and *HMR* donor sequences that repair *MAT* in normal condition [20]. They showed that short Ty sequences could be captured at the site of the HO DSB and this event was not dependent on the recombination protein Rad52 [20]. In addition,

Teng *et al*. developed a system, in which a *his3* gene with an artificial intron (AI) was fused with a Ty1 element on a plasmid. Following induction of the DSB by HO, they found that the *Ty1/HIS3* cDNA was integrated at the site of the HO DSB [21] (**Figure 1.2A**).

LINE-1 is a retrotransposon in mammalian cells and has endonuclease cleavage activity. This endonuclease activity of LINE-1 generates a 3'-OH end that can be used as a primer for RT to make cDNA and integrate LINE-1 cDNA into the genome [24]. Moorish *et al*. showed that LINE-1 with its RT activity can integrate its cDNA at sites of DNA lesions even in the absence of the NHEJ mechanism in Chinese hamster ovary cells [23, 24]. LINE-1 elements transpose other retroelements or cellular RNA as well as themselves [22]. Moreover, LINE-1 can use free DNA ends as primers for insertion of cDNA and DSB repair (**Figure 1.2B**).
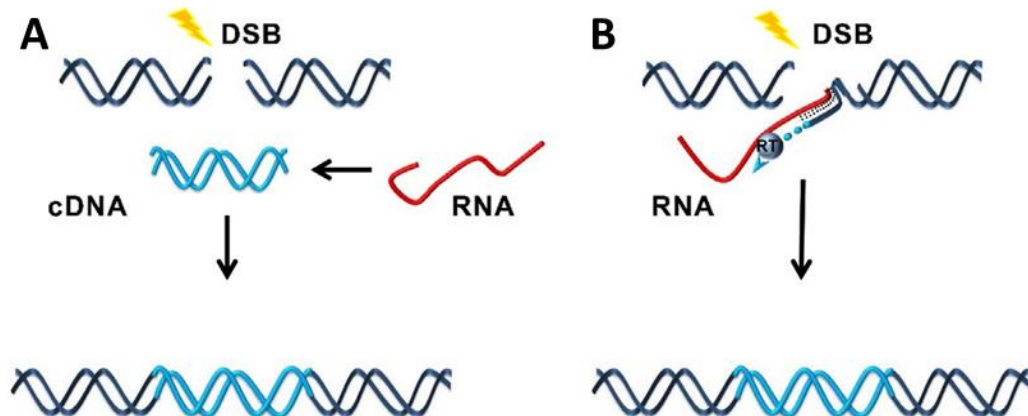
Figure 1.2 Model for DSB repair by cDNA insertion. A) RNA-mediated, non-templated DSB repair **B)** RNA-templated DSB repair by cDNA insertion **(by Havva Keskin in [25])**

Ty elements generate cDNA molecules by RT in yeast. This cDNA can recombine with homologous Ty elements by HR [26, 27] and induction of a DSB within the Ty DNA sequence can increase recombination between Ty cDNA and genomic Ty DNA [28]. Ty not only has the capability of reverse transcribing its own RNA but can also RT other cellular RNAs in the yeast cells. The study by Derr *et al.* showed the ability of Ty elements to generate cDNA using a *HIS3* reporter gene interrupted by an artificial intron (AI) in its antisense orientation [17]. The *his3*-AI cassette was placed on a vector under the control of a galactose inducible promoter (*pGAL1*). Upon expression of this cassette in galactose medium, the *his3* antisense mRNA generated His$^+$ colonies. Half of the His$^+$ colonies were due to integration of cDNA into genomic DNA and other half was due to the HR event between the cDNA and the plasmid [17]. These findings support the possibility that cDNA could be a template for DNA DSB repair when carrying homology to the broken DNA ends.

1.3 Signs of RNA-DNA recombination

RNA can guide genetic and epigenetic modifications in DNA. RNA can guide site-specific DNA cleavage in adaptive bacterial immunity [29]. Small RNAs have a role in DSB repair by guiding molecules directing chromatin modifications or recruiting proteins to sites of DNA damage [30, 31]. Transcriptional gene silencing can be induced by microRNAs [32]. *In vitro* work showed that the bacterial DNA recombination protein RecA promotes pairing between duplex DNA and single-strand RNA resulting in formation of a three strand structure called an R-loop [33, 34]. Artificial RNA molecules can direct genome modifications and rearrangements when injected into the protozoa,

Oxytricha trifallax [35]. An RNA-templated insertion was postulated from analysis of DNA sequences repaired via NHEJ after DSB induction by zinc-finger nucleases in *Drosophila* cells [36].

Furthermore, discovery of a viral genome representing an RNA-DNA chimera inferred a recombination between RNA and DNA viruses [37]. In addition, recent work showed that Rad51 promotes the formation of RNA-DNA hybrids in yeast *S. cerevisiae* [38]. Nevertheless, these findings do not provide direct proof that RNA can directly exchange genetic information with DNA.

1.4 Genomic DNA modification by RNA-oligonucletides

Work by Storici *et al*. showed that synthetic RNA containing DNA oligonucleotides (oligos) or RNA only oligos can repair a DSB in chromosomal DNA of yeast *S. cerevisiae* [39]. RNA only or RNA-containing oligos were transformed into yeast cells to repair a broken *leu2* locus to detect RNA-templated DNA repair. DSB repair by RNA oligos was not affected by deletion of the *SPT3* gene, which is essential for Ty1 and Ty2 transcription and transposition activity [39]. Genetic modification of DNA guided by short RNA oligos was also found in *Escherichia coli* and human embryonic kidney cells (HEK-293) [40, 41]. However, a big question remained as to whether not only synthetic RNA or RNA-containing molecules, but also endogenous cellular RNA molecules could engage in RNA-DNA HR.

1.5 RNA-DNA hybrids

RNA-DNA hybrids can form during replication and transcription [42]. DNA primase generates RNA primers during lagging strand synthesis, resulting in a RNA-DNA hybrid [43]. During transcription, a short RNA-DNA hybrid forms at the transcription bubble [43]. However, formation of long RNA-DNA hybrids during transcription generates structures known as R-loops. An R-loop consists of an RNA-DNA hybrid duplex and a displaced ssDNA loop. R-loops can be a source of genome instability including mutations, recombination, chromosome rearrangement and chromosome loss [42]. R-loops are more sensitive to lesions, transcription-associated mutagenesis and transcription-associated recombination [44]. Moreover, R-loops and RNA-DNA hybrids have been implicated in human diseases. The main concern is that R-loops are associated with chromosomal breakage [45]. It has been shown that XPF and XPG, nucleotide excision repair endonucleases, process R-loops into DSBs. In addition, accumulation of R-loops and RNA-DNA hybrids were linked to Aicardi-Goutieres syndrome (AGS), Fragile X syndrome, and Friedreich's ataxia [45]. Considering the negative effects that RNA-DNA hybrids can have on genome stability, cells possess numerous mechanisms to prevent formation of RNA-DNA hybrids. These include RNA-DNA helicases, like Sen1 in yeast and Senataxin in humans [46, 47], the Pif1 helicase [48, 49], topoisomerases [50, 51], the THO/TREX complex, which keeps RNA away from DNA during transcription [42], and Ribonuclease (RNase) H enzymes [52]. Two RNase H enzymes, RNase H1 and RNase H2, directly cut the RNA strand of RNA-DNA hybrids and prevent their accumulation [52].

1.6 The function of Ribonuclease (RNase) H enzymes

RNase H enzymes catalyze the cleavage of an RNA strand in RNA-DNA hybrids [53, 54] and play roles in DNA replication, transcription, recombination, and repair [55]. There are two main types of RNase H. RNase H1/I has only one subunit. In eukaryotes, RNase H2/II has three different subunits, the catalytic subunit (Rnh201 in yeast, and RNase H2A in humans) and two additional subunits (Rnh202 and Rnh203 in yeast, and RNase H2B and RNase H2C in humans), which are necessary for catalysis. RNase H1/I and RNase H2/II have different cleavage specificity. RNase H1/I requires a stretch of at least 4 ribonucleotides (rNMPs) in a DNA duplex in order to cleave RNA efficiently (**Figure 1.3**) [52]. RNase H2/II cleaves at long RNA-DNA hybrids as well as at single rNMPs embedded in a DNA duplex (**Figure 1.3**) [56].
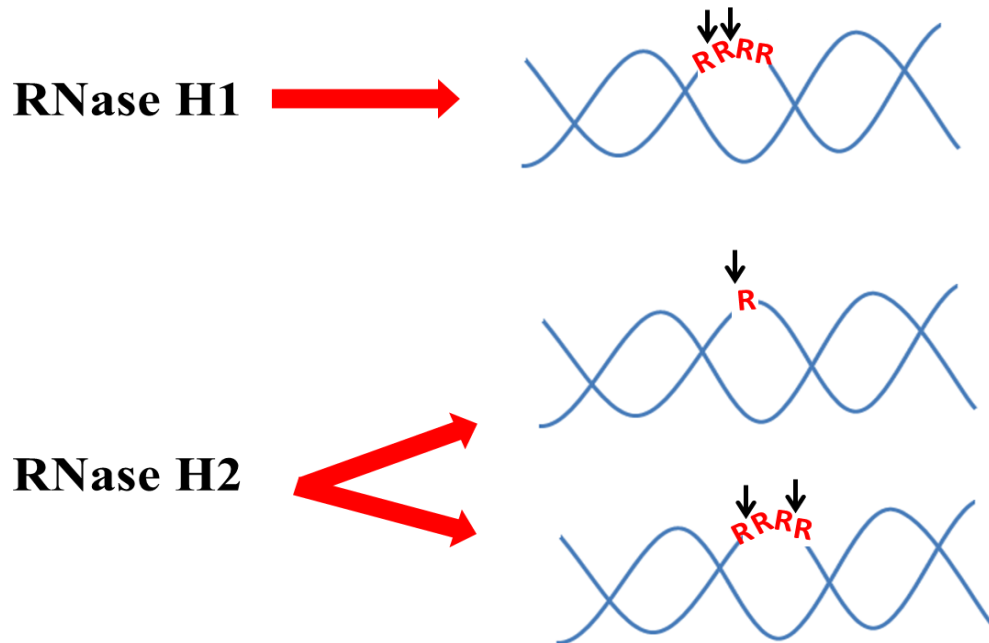
Figure 1.3 Cleavage specificity of RNase H1 and RNase H2. Ribonucleotides are in red as 'R'. DNA is in blue. Arrows represent sites of cleavage by RNase H1 or RNase H2 (**Modified from [52]**)

RNase H1 has been implicated in mitochondrial DNA (mtDNA) replication during mouse development [57]. Null mutations in RNase H1 or RNase H2 genes are embryonic lethal in mice [57, 58]. Mutations in any of the human RNase H2 subunits are associated with AGS, which is a severe neurological and inflammatory disorder the mainly affects the brain, the immune system, and the skin in humans [52, 59, 60]. AGS can be also caused by mutations in TREX1 (3' to 5' exonuclease) [61], SAMHD1 (dNTP triphosphatase) [62], ADAR1 (RNA-editing enzyme) [63], or IFIH1 (Interferon Induced with helicase C domain 1) [63]. The molecular mechanisms that cause AGS are still unclear; however, the seven genes that are mutated in AGS patients are directly or indirectly related to nucleic acid modification/degradation. The possible accumulation of RNA-DNA hybrids in defective RNase H2 cells could be a trigger for the disease [64]. AGS patients with a defect in RNase H2 could have an increased level of cDNA in the form of RNA-DNA hybrids, which could play a role in activating the immune system.

1.7 Research Goals

With the scope of better understanding the relationship between RNA and DNA in the context of genome stability and a particular focus on exploring the possibility that RNA may have a positive role in DNA repair, we set up the following research goals.

1.7.1 To detect transcript-RNA-templated DSB repair in yeast cells

Previous study showed that synthetic RNA or RNA-containing oligos can be used as a template for repairing mutated or broken homologous chromosomal DNA in yeast,

9

human and bacterial cells. Here, we hypothesized that endogenous RNA molecules, RNA transcripts, can serve as a template for DNA DSB repair in yeast cells. We aimed to develop two experimental yeast systems to examine whether an RNA transcript from a yeast marker gene could repair a DSB induced i) in the same DNA locus that generates the repairing transcript RNA (in *cis*) or ii) in a homologous sequence but located in a different locus from the one generating the repairing transcript RNA (in *trans*) in yeast *S. cerevisiae*.

1.7.2 To explore the effect of RNase H mutations on cDNA- and direct RNA-templated DSB repair

In the course of this study, we found that the frequency of DNA repair by cDNA and transcript RNA was highly increased in the absence of the catalytic subunit of RNase H2 and without RNase H1. We, therefore, first planned to test whether null alleles of the other two RNase H2 subunits would also impair the frequency of DSB repair by RNA, and whether binding of RNase H2 to the DNA clamp PCNA (proliferative cell nuclear antigen) was required for RNA-templated DSB repair. Moreover, we planned to examine the effect of mutations in the RNase H2 gene that are associated with the AGS syndrome on the frequency of direct and indirect RNA-templated DSB repair in yeast.

1.7.3 To characterize the mechanism of transcript-RNA-templated DNA DSB repair

Our results uncovered that transcript RNA can be used as a template to repair a DNA DSB in yeast cells. Here, we aimed to characterize the mechanism of transcript-driven DNA repair. Our major goal was to identify factors regulating transcript RNA-templated

DNA DSB repair. Because we know very little about the mechanism of DSB repair by

transcript RNA, we first explored the involvement of factors functioning in HR and DNA

end processing.

CHAPTER 2

TRANSCRIPT-RNA-TEMPLATED DNA RECOMBINATION AND REPAIR

The study in Chapter 2 consists of the work published in *Nature* 515 (2014), 436-439.

Havva Keskin[1], Ying Shen[1†], Fei Huang[2], Mikir Patel[2], Taehwan Yang[1], Katie Ashley[1], Alexander V. Mazin[2] & Francesca Storici[1]

[1]Department of Biochemistry and Molecular Biology, Drexel University College of Medicine, Philadelphia, PA 19102, USA.

[2]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA

[†]Present address: Division of Computational Biomedicine, Boston University School of Medicine, Boston, Massachusetts 02118, USA.

2.1 Abstract

Homologous recombination is a molecular process that has multiple important roles in DNA metabolism, both for DNA repair and genetic variation in all forms of life [13]. Generally, homologous recombination involves the exchange of genetic information between two identical or nearly identical DNA molecules [13]; however, homologous recombination can also occur between RNA molecules, as shown for RNA viruses [65]. Previous research showed that synthetic RNA oligonucleotides can act as templates for DNA double-strand break (DSB) repair in yeast and human cells [39, 66], and artificial long RNA templates injected in ciliate cells can guide genomic rearrangements [67]. Here we report that endogenous transcript RNA mediates homologous recombination with chromosomal DNA in yeast *Saccharomyces cerevisiae*. We developed a system to detect the events of homologous recombination initiated by transcript RNA following the repair of a chromosomal DSB occurring either in a homologous but remote locus, or in the same transcript-generating locus in reverse-transcription-defective yeast strains. We found that RNA–DNA recombination is blocked by ribonucleases H1 and H2. In the presence of H-type ribonucleases, DSB repair proceeds through a complementary DNA intermediate, whereas in their absence, it proceeds directly through RNA. The proximity of the transcript to its chromosomal DNA partner in the same locus facilitates Rad52-driven homologous recombination during DSB repair. We demonstrate that yeast and human Rad52 proteins efficiently catalyze annealing of RNA to a DSB-like DNA end in vitro.

Our results reveal a novel mechanism of homologous recombination and DNA repair in which transcript RNA is used as a template for DSB repair. Thus, considering the abundance of RNA transcripts in cells, RNA may have a marked impact on genomic stability and plasticity.

2.2 Materials and Methods

2.2.1 Experimental design to explore transcript-RNA-templated chromosomal DSB repair in yeast.

In the experimental design to explore transcript-RNA-templated chromosomal DSB repair it is critical to discriminate repair of the DSB by transcript RNA from repair by the DNA region that generates the transcript. Also, translation of the repairing transcript mRNA should not produce the functional His3 protein. Moreover, it is essential that DSB repair would not restore the *HIS3* marker sequence by simple end ligation via non-homologous end-joining (NHEJ). To satisfy these requirements, the DNA region that generates the transcript was constructed to contain a his3 allele on chromosome III consisting of a yeast *HIS3* gene interrupted by an artificial intron in the antisense orientation (*mhis3-AI cassette*), which was previously used to study reverse transcription in yeast [68, 69]. The antisense *his3* RNA is not translated into the functional His3 protein. Moreover, after intron splicing, the transcript RNA sequence has no intron, while the DNA region that generates the transcript retains the intron; thus they are distinguishable. We developed two experimental yeast cell systems, *trans* and *cis* (**Figure 2.1a, b and Figure A.1**) in strains YS-289, 290 and YS-291, 292, respectively (**Table A.1**). In both systems, transcription of the antisense *his3* RNA and expression of the homothallic switching endonuclease are regulated by the galactose-inducible promoter (*pGAL1*). In addition, these yeast cell systems are auxotrophic for histidine (His⁻) and thus do not grow on media without histidine. Upon induction of the homothallic switching endonuclease DSB, the broken *his3* allele of the *trans* and *cis* cell systems can, in principle, only be repaired to a functional *HIS3* allele by recombination with a homologous template. Alternative mechanism of *HIS3* repair by ligation of the broken ends via NHEJ is inefficient in this system (<0.1 out of 10^7 viable cells) (data not shown), as the *HIS3* gene is disrupted by a long sequence with the homothallic switching endonuclease site (*trans* system) or an intron and the homothallic switching endonuclease

14

site (*cis* system) (**Figure A.1b, c**). To impair DSB repair by cDNA deriving from the *his3* antisense, we deleted the *SPT3* or the *DBR1* gene. *SPT3* encodes for a subunit of the SAGA and SAGA-like transcriptional regulatory complexes and its null allele reduces Ty reverse transcriptase function over 100-fold [66, 70, 71]. DBR1 encodes for the RNA debranching enzyme Dbr1 and its null allele in yeast cells impairs cDNA formation and diminishes Ty transposition up to a factor of tenfold [72, 73]. As further proof that we can detect DSB repair by transcript RNA independently of cDNA, we performed the *trans* and *cis* assays with and without RNase H functions in the presence of foscarnet (phosphonoformic acid, PFA), an inhibitor of the HIV reverse transcriptase, which blocks Ty reverse transcription in yeast [74] (and data not shown).

2.2.2 Yeast strains.

The yeast strains used in this work are listed in **Table A.1** and derive from the FRO-767 strain [66], which contains the site for the site-specific homothallic switching endonuclease in the middle of the *LEU2* gene on chromosome III. A gene cassette carried on plasmid pSM50 (refs [68, 69]) containing the *his3* gene disrupted by an artificial intron and regulated in the antisense orientation by the galactose inducible promoter *pGAL1* and containing the *URA3* marker gene (*pGAL1-mhis3-AI-URA3*) was integrated into the *leu2* locus of strain FRO-767 after DSB induction at the homothallic switching endonuclease site by the gene collage technique with no PCR amplification . The *URA3* gene was then replaced with the *ADE3* gene generating strain FRO-1073 [75]. To build the strains of the *trans* system, an homothallic switching endonuclease site was integrated into the endogenous *HIS3* locus on chromosome XV of FRO-1073 exactly in the same position in which the artificial intron was inserted in the *pGAL1-mhis3-AI* cassette using the delitto perfetto method, as described previously [75, 76], to generate FRO-1075, 1080.The correct sequence and insertion position of the homothallic switching endonuclease site was confirmed by sequence analysis. For constructing strains of the *cis*

system, first the *his3* gene disrupted by the homothallic switching endonuclease site of FRO-1075 and 1080 was replaced with a *TRP1* gene to generate YS-164, 165, and then an homothallic switching endonuclease cutting site was integrated into the artificial intron in the *his3* cassette on chromosome III to generate strain YS-172, 174. To be cautious to avoid any possibility of transcription from Ty into the *pGAL1-mhis3-AI* cassette in both the trans and *cis* systems, the Ty2 element located upstream of the *leu2* locus on chromosome III, YCLWTy2-1, was deleted following the delitto perfetto method to generate YS-289, 290 (*trans* system) and YS-291, 292 (*cis* system). These new strain constructs were verified by PCR and sequence analysis to confirm correct constructions. However, no difference in the frequency of His$^+$ cells was observed between the strains with the YCLWTy2-1 and those without it for the strains of both the *trans* and *cis* systems (data not shown). Deletion mutants for the *trans* YS-289, 290, and the *cis* 291, 292 strains contain either the kanMX4, hygMX4, natMX4 and/or the *Kluyveromyces lactis URA3* (*KlURA3*) marker gene in place of the open reading frame or the promoter of the gene(s) of choice. All gene disruptions were confirmed by colony PCR. Strains HK-396, 400 and HK-391, 394 were constructed using the delitto perfetto method by deleting the first 23 base pairs on the 59 end of the artificial intron via insertion of the CORE cassette, and then by popping out the CORE cassette with a pair of oligonucleotides. These constructs were confirmed by sequence analysis. Strain HK-404, 407 was obtained by deleting the SPT3 gene with kanMX4 from HK-391, 394. The FRO-1092, 1093 strain is *rad52Δ* and has only one *his3* allele, the endogenous allele on chromosome XV that has been inactivated by the homothallic switching endonuclease site.

2.2.3 Standard genetic, molecular biology techniques and plasmids.

Yeast genetic methods and molecular biology analyses were done as described [66, 75, 76]. The BDG606 vector [77] and the BDG283 control vector (a gift from D. Garfinkel),

used to verify a direct role of transcript RNA in DSB repair (**Table A.4**), are centromeric plasmids with the *URA3* marker. BDG606 contains the *pGAL1-mhis3-AI* cassette fused to Ty (*pGTy1-H3his3-AI/Cen-URA3*) and BDG283contains only *pGAL1*. The plasmids used for the complementation assay with RNase H2 are YEp195SpGAL, which is a high-copy expression plasmid containing the *URA3* selectable marker [78], YEp195SpGAL containing the wild-type *RNH201* gene (YEp195SpGAL-RNH201) inserted by gap repair, and YEp195SpGAL-rnh201-D39A constructed by in vitro mutagenesis (Quick Change Mutagenesis Kit, Stratagene, La Jolla, CA) of YEp195 SpGAL-*RNH201* and confirmed by sequence analysis. To confirm occurrence of the homothallic switching endonuclease DSB following incubation in the 2% galactose medium, the percentage of G2 arrested cells was determined right before adding galactose and after 8-h incubation in galactose as previously described [40] (**Figure A.2c**). All primers used for strain and plasmid constructions, PCR verifications and sequence analyses are available upon request. Samples for sequencing were submitted to Eurofins MWG Operon. The Southern blot experiment was done as follows. Cells from colonies growing on rich medium containing yeast extract, peptone and 2% (w/v) dextrose (YPD) or His⁻ media were grown on YPD overnight (O/N). Genomic DNA was extracted as described [79] and digested with either BamHI or NarI restriction enzyme. After digestion, column purification was applied by using QIA quick PCR Purification Kit (Qiagen). DNA was run in a 0.8% agarose gel. Following electrophoresis and Southern blotting chromosomal regions containing the *HIS3* gene were detected using a [a-32P] ATP (PerkinElmer)-labelled (Prime-ItRmTRandom Primer Labelling Kit, Agilent Technologies) 250-base-pair *HIS3*-specific probe. Membrane was exposed to a phosphor screen for 3 days. Images were taken with Typhoon Trio1 (GE Healthcare) and obtained with Image Quant (GE Healthcare).

2.2.4 *Trans* and *cis* assays using patches or liquid cultures.

Yeast cells of the chosen strains were patched on YPD and grown at 30 °C for 1 day. The cells were then replicaplated on medium containing yeast extract, peptone and 2%(w/v) galactose (YPGal) or YPGal containing phosphonoformic acid (PFA, 2.5 mg ml^{-1}) to turn on transcription of the *his3* antisense on chromosome III and expression of the homothallic switching endonuclease. As a control, cells were also replica-plated from the YPD medium on synthetic complete medium plates lacking histidine (SC-His^{-}) and grown for 3 days at 30 °C. We never detected a single His^{+} colony from any of the *trans* and *cis* strains used in this study following replica-plating from the YPD medium on SC-His^{-} (not shown). After 2 days' incubation on YPGal medium, these cells were replica-plated onto SC-His^{-} and grown for 3 days at 30 °C to form visible colonies. At this stage, plates were photographed and photo files stored. For experiments using the BDG606 and BDG283 plasmids, cells were replica-plated from SC-Ura^{-} onto SC-Ura^{-}Gal medium, and were then replica-plated onto SC-Ura^{-}His^{-}. As a control, cells were also replica-plated from the SC-Ura^{-} medium onto SC-Ura^{-}His^{-} and grown for 3 days at 30 °C. For the experiments in liquid culture, flasks with 50 ml of liquid medium containing yeast extract, peptone and 2.7% (v/v) lactic acid (YPLac) were inoculated with yeast cells of the chosen strains and incubated in a 30 °C shaker for 24 h. The density of the cultures was determined by counting cells using a hemocytometer and counting under a microscope. Generally, 10^{7} or, in rare cases, 10^{8} cells (we note that survival is very low on galactose medium) were then plated on YPGal medium, or YPGal medium containing PFA (2.5mg ml^{-1}) for experiments using PFA to obtain from 1 to, ~500 His^{+} colonies per plate after the replica-plating on His^{-} medium, and grown for 2 days at 30 °C. Two aliquots of 10^{4} cells were plated, each on one YPGal medium plate, or YPGal medium containing PFA (2.5 mgml^{-1}) for experiments using PFA plate, to measure the cell survival after galactose treatment. After 2 days' incubation on YPGal medium, cells were replica-plated on His^{-} plates and grown for 3 days at 30 °C. The frequency of DSB repair

18

was calculated by dividing the number of His$^+$ colonies on SC-His$^-$ medium by the number of colonies on YPGal medium. The survival was calculated by dividing the number of colonies on YPGal medium by the number of cells plated on the same medium. For experiments using the BDG606 and BDG283 plasmids, cells were treated as described above except that they were plated from YPLac on SC-Ura$^-$Gal medium in different dilutions, and were then replica-plated on SC-Ura$^-$His$^-$. The frequency of His$^+$ colonies was calculated by dividing the number of His$^+$ colonies on SC-Ura$^-$His$^-$ medium by the number of colonies on SC-Ura$^-$Gal medium. The survival was calculated by dividing the number of colonies on SC-Ura$^-$Gal medium by the number of cells plated on the same medium.

2.2.5 Oligonucleotide transformation

Transformation by oligonucleotides (1nmol) was performed as described [66]. Induction of the homothallic switching endonuclease DSB was done by incubating cells in 2% galactose medium for 3 h.

2.2.6 Transposition assay

Yeast cells of the chosen strains transformed with BDG102 (empty plasmid) or BDG598 (*pGTy-H3mhis3-AI*) plasmid [80] (containing a Ty transposon fused to the *his3* gene, which is in the antisense orientation and disrupted by an artificial intron; both Ty and the *his3* antisense are regulated by the galactose inducible promoter) were patched on SC-Ura$^-$ and grown overnight at 30 °C. Cells were then replica-plated on synthetic medium lacking uracil with 2% (w/v) galactose (SC-Ura$^-$Gal) and grown for 48 or 96 h at 30 °C or 22 °C, respectively. As control, cells were also replica-plated on SC-His$^-$ to determine the background of His$^+$ clones. After the incubation in galactose, cells were replica-plated on SC-His$^-$ and grown for 3 days at 30 °C to form visible colonies. At this stage, plates were photographed and photo files stored. For the experiments in liquid culture, strains with

BDG102 or BDG598 were grown in 5 ml SC-Ura$^-$ liquid medium or in 10 ml of YPLac liquid medium in a 30 °C shaker for 24 h. Then, 1×10^6 cells were transferred from the SC-Ura$^-$ liquid medium into 5 ml SC-Ura$^-$ or 5 ml SC-Ura$^-$Gal liquid medium and incubated for 48 or 96 h at 30 °C or 22 °C, respectively. After 24 h, YPLac cultures were split in half. One-half was kept growing for additional 48 h at 30 °C, while galactose was directly added to the other half to reach 2% and cells were then incubated for 48 h at 30 °C. From glucose and YPLac cultures grown at 22 °C or 30 °C, 10^7 or 10^8 cells were plated on SC-His$^-$Ura$^-$ medium, respectively, and were grown for 2 days at 30 °C. From galactose cultures grown at 22 °C or 30 °C, 10^5 or 10^6 cells were plated on SC-His$^-$Ura$^-$ medium, respectively, and were grown for 2 days at 30 °C. Two aliquots of 5×10^2 cells were plated each on one SC-Ura$^-$ medium plate, to measure the cell survival after glucose, YPLac or galactose treatment. The rate of formation of His$^+$ cells was calculated using the maximum-likelihood method described in ref. [81].

2.2.7 Quantitative real-time PCR

RNA was isolated from the chosen yeast strains of the *trans* and *cis* systems using a protocol adapted from a method described previously [82]. RNA was converted in to cDNA using QuantiTect ReverseTranscription Kit (Qiagen). SYBR Green qPCR Mix (BioRad) was used for analyzing RNA expression in 96-well plates (Applied Biosystems). The total volume in each well was 20 ml, which consisted of 10 ml of SYBR Green qPCR Mix, 4 ml of nuclease-free water, 2 ml of primers and 4 ml of cDNA. The cDNA levels were determined using an ABI Prism 7000 RT–PCR machine (Applied Biosystems). ACT1.F and ACT1.R, HIS3.F2 and HIS3.R2 primers were used in this study (**Table A.2a**). ACT1 primers were used for normalization. Values for each sample were normalized with ACT1, and then a second normalization was performed by subtracting normalized values of each time point from the control normalized value per each gene [83]. As a negative control, CEN16.F and CEN16.R primers were used to

20

show that there is minimal or no qPCR product derived from a chromosomal region that is not transcribed (A. El Hage, personal communication) (data not shown).

2.2.8 Rad52 in vitro annealing assay

In vitro assays using yeast or human Rad52 were performed as described [84, 85] (and references therein), with all DNA and RNA concentrations expressed in moles of molecules. All oligonucleotide sequences (no. 211, no. 501, no. 508 and no. 509) are shown in **Table A.2a**. A single nucleotide mismatch was incorporated into the dsDNA (relative to ssDNA or RNA) to reduce the spontaneous Rad52-independent annealing. Tailed dsDNA (no. 508 and no. 509) (0.4 nM) was incubated in the absence or presence of yeast or human RPA (2 nM)in a buffer containing 25 mM Tris acetate, pH7.5, 100 mg/ml[-1] BSA, and 1mM DTT (dithiothreitol) for 5 min at 37 °C. Then yeast or human Rad52 (1.35 nM) was added to the mixture containing either yeast or human RPA, respectively, and incubation continued for 10 min. Annealing reactions were initiated by adding 32P-labelled ssRNA (no. 501) or ssDNA (no. 211) (0.3 nM). Aliquots were withdrawn at indicated time points and deproteinized by incubating samples in stop solution containing 1.5% SDS, 1.4 mg/ml[-1] proteinase K, 7% glycerol and 0.1% bromophenol blue for 15 min at 37 °C. Samples were analyzed by electrophoresis in 10% (17:1 acrylamide:bisacrylamide) polyacrylamide gels in 1X TBE (90mMTrisborate, pH 8.0, 2mM EDTA) at 150V for 1 h and were quantified using a Storm 840 Phosphorimager and Image Quant 5.2 software (GE Healthcare).

2.2.9 Data presentation and statistics

Graphs were made using GraphPad Prism 5 (Graphpad Software). The results are each expressed as a median and 95% confidence interval (in brackets), or alternatively the range when number of repeated experiments was, 6. Statistically significant differences between the His$^+$ frequencies were calculated using the nonparametric two-tailed Mann–

Whitney U-test [86]. All P values obtained using the Mann–WhitneyU-test were then

adjusted by applying the false discovery rate method to correct for multiple hypothesis

testing [87] (**Table A.1**).

2.3 Results

To investigate the capacity of transcript RNA to recombine with genomic DNA, we

sought to discover whether a chromosomal DSB could be repaired directly by

endogenous RNA in yeast *S. cerevisiae* cells. We designed a strategy by which we could

induce a DSB in the *HIS3* marker gene and monitor precise repair of the DSB by a

homologous transcript messenger RNA by restoration of *HIS3* function resulting in

histidine prototrophic (His$^+$) cells (see Methods). We developed two experimental yeast

cell systems, *trans* and *cis*, in strains YS-289, 290 and YS-291, 292, respectively (**Table

A.1**). The *trans* system is designed to test the ability of a spliced (intron-less) antisense

his3 transcript from chromosome III to repair a DSB in a different *his3* allele on

chromosome XV, which contains an engineered homothallic switching endonuclease

cutting site (**Figure 2.1a and Figure A.1a, b**). The *cis* system is designed to test the

capacity of the spliced antisense *his3* transcript from chromosome III to repair a

homothallic-switching-endonuclease-induced DSB located inside the intron of the same

his3 locus (**Figure 2.1b and Figure A.1c**).
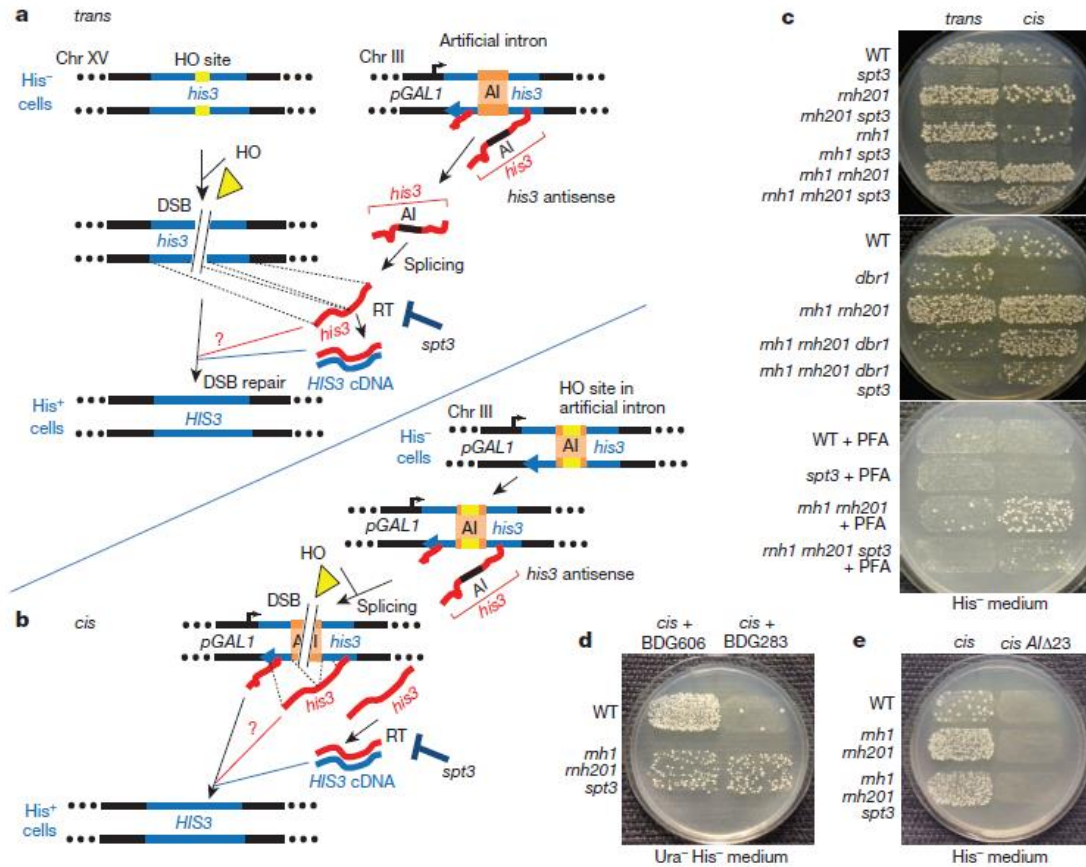
Figure 2.1 Repair of a chromosomal DSB by transcript RNA. a, b, Scheme of the *trans* (a) and *cis* (b) cell systems used to detect DSB repair by transcript RNA. AI, artificial intron; HO, homothallic switching endonuclease; pGAL1, galactose-inducible promoter; RT, reverse transcriptase. Yellow triangles, cleavage activity by HO homothallic switching endonuclease; red question marks, hypothesis for transcript-RNA-templated DSB repair mechanism. c–e, Examples of replica-plating results (n 56) from galactose medium to histidine dropout medium demonstrating the ability of various yeast strains (relevant genotypes shown) of the *trans* and *cis* systems to generate histidine prototrophic colonies in the absence of SPT3, or DBR1 function, or with phosphonoformic acid (PFA) (c), in the presence of the plasmid carrying the pGAL1-mhis3-AI cassette (BDG606) or the control (BDG283) (d), or when the artificial intron has a 23-base-pair deletion (AID23) (e). WT, wild type.

In both the *trans* and *cis* cell systems, the spliced antisense *his3* transcript RNA can serve as a homologous template to repair the broken *his3* DNA and restore its function. However, given the abundance of Ty retrotransposons in yeast cells, the spliced antisense *his3* RNA could potentially be reverse transcribed by the Ty reverse transcriptase in the cytoplasm to cDNA that could then recombine with the homologous broken *his3* sequence or be captured by non-homologous end joining at the homothallic switching endonuclease break site to produce His$^+$ cells [17, 70, 88]. To distinguish DSB repair mediated by the transcript RNA template from repair mediated by the cDNA template, we performed the *trans* and *cis* assays in two yeast strains that contained either a wild-type *SPT3* gene or its null allele, which prevents Ty transcription and strongly reduces Ty transposition and transpositional recombination [66, 70, 71]. In both assays, cells containing wild-type *SPT3* produced numerous His$^+$ colonies after DSB induction (**Figure 2.1c and Table 2.1a**). As expected, the frequency of His$^+$ colonies in the *trans* system was significantly higher than that in the *cis* system because the *his3* transcript is continuously generated in the presence of galactose. In contrast, production of the full *his3* transcript is immediately terminated upon DSB formation in the *cis* system. This frequency difference is not specific to the particular genomic loci in which the DSBs are induced, as transformation by DNA oligonucleotides (HIS3.F and HIS3.R) designed to repair the broken *his3* gene produced the same frequency of His$^+$ colonies in the two systems (**Tables A.2a and A.3**), demonstrating that the homothallic switching endonuclease DSB stimulates homologous recombination in the *trans* and *cis* systems equally well. Notably, almost all the His$^+$ colonies are dependent on *SPT3* function, indicating that the DSB in *his3* is repaired exclusively via the cDNA pathway (**Figure 2.1c and Table 2.1a**). This finding demonstrates that if an actively transcribed gene is broken, it can be repaired using a cDNA template derived from its intact transcript. Moreover, these data also support the model in which reverse-transcribed products from any sort of RNA can be a significant source of genome modification at DSB sites [89].

24

For RNA to recombine with DNA, an intermediate step that is probably required is the formation of an RNA–DNA heteroduplex. We therefore deleted the genes coding for ribonuclease (RNase) H1 (*RNH1*) and/or the catalytic subunit of RNaseH2 (*RNH201*), which both cleave the RNA strand of RNA–DNA hybrids . Remarkably, while deletion of *RNH1* slightly increased the frequency of His$^+$ colonies in the *trans* system, deletion of *RNH201* increased the frequency of His$^+$ colonies in both the *trans* and *cis* systems, and combined deletion of *RNH1* and *RNH201* resulted in an even stronger increase of His$^+$ colonies in both systems. Moreover, we detected His$^+$ colonies in *rnh1 rnh201* cells in the absence of *SPT3* (**Figure 2.1c and Table 2.1a**). Notably, there were more His$^+$ colonies in *cis*-system *rnh1 rnh201 spt3* than in *trans*-system, and the frequency of His$^+$ colonies observed in the *rnh1 rnh201 spt3* relative to *spt3* cells was much higher in *cis* (>69,000) than in *trans* (>6,400) (**Figure 2.1c and Table 2.1a**). If DSB repair in *rnh1 rnh201 spt3* cells were due to cDNA, we would expect a higher His$^+$ frequency in the *trans* than in the *cis* system, as observed in wild-type cells. The fact that the His$^+$ frequency is higher in the *cis* system suggests that DSB repair is not mediated by cDNA but instead by RNA or predominantly RNA. To further examine the possibility that residual cDNA rather than transcript RNA is responsible for *his3* correction in *cis*-system *rnh1 rnh201 spt3* cells, we introduced a *trans* system directly into these cells and into the control *cis* wild-type cells. When wild-type cells of the *cis* system were transformed with a low-copy-number plasmid carrying the *pGAL1-mhis3-AI* cassette, where AI represents an artificial intron (BDG606; see Methods), they displayed a large (a factor of 4,000) increase in the His$^+$ frequency following DSB induction in *his3* compared to the same cells transformed with the control empty vector (BDG283). In contrast, BDG606 in *cis*-system *rnh1 rnh201 spt3* cells did not significantly increase the His$^+$ frequency (**Figure 2.1d and Table A.4**). These results argue against the role of residual cDNA in template-dependent DSB repair in *cis*-system *rnh rnh201 spt3* cells and support a predominant, direct template function of the *cis*-system *his3* transcript RNA in these cells. Overall, these data support the

conclusion that a transcript RNA can directly repair a DSB in *cis*-system *rnh1 rnh201* and *rnh1 rnh201 spt3* cells. The physical proximity of the *his3* transcript to its own *his3* DNA during transcription could facilitate annealing of the broken DNA ends to the transcript. This possibility is consistent with the fact that closer donor sequences repair DSBs more efficiently [90, 91] and that mature transcript RNAs are exported rapidly to the cytoplasm or degraded after completion of transcription [92].

To confirm that inactivation of RNases H1 and H2 allows for direct transcript RNA repair of a DSB in homologous DNA, we conducted a complementation test in the *cis* system using a vector expressing either a catalytically inactive mutant of *RNH201*, *rnh201* (D39A) [93], or wild-type *RNH201*. Results showed that when wild-type RNH201 was expressed from the plasmid in *rnh1 rnh201 spt3* cells, there were no His[+] colonies following DSB induction (**Figure A.2a**). Deletion of *SPT3* is a well-established and robust method to suppress reverse transcription and formation of cDNA in yeast [66, 70, 94]. However, to prove that the increased frequency of His[+] detected in the *cis*-relative to the *trans*-system *rnh1 rnh201 spt3* background was not solely linked to *SPT3* deletion, we impaired cDNA formation by deleting the *DBR1* gene, which codes for the RNA debranching enzyme Dbr1 [72, 73], or by using the reverse transcriptase inhibitor foscarnet (phosphonoformic acid) [74]. Results shown in **Figure 2.1c** and **Table A.5a** support our conclusion that RNA transcripts can directly repair a DSB in chromosomal DNA without being first reverse transcribed into cDNA in *rnh1 rnh201* cells.

Efficient generation of His[+] colonies in cis wild-type, *rnh1 rnh201*, or *rnh1 rnh201 spt3* cells requires transcription and splicing of the antisense *his3* and DSB formation in the *his3* gene. Deletion of *pGAL1* (the galactose-inducible promoter) upstream of *his3* on chromosome III, deletion of the homothallic switching endonuclease gene, or growing cells in glucose medium, in which homothallic switching endonuclease is repressed,

drastically decreased His$^+$ frequency (**Figure A.2b, c and Table A.5b, c**). Similarly, yeast wild-type, *rnh1 rnh201* and *rnh1 rnh201 spt3* cells of the *cis* system containing a 23-base-pair truncation of the artificial intron in *his3* lacking the 59 splice site (**Table A.1 and Figure A.1c**) produced no His$^+$ colonies following DSB induction (**Figure 2.1e and Table A.5d**), yet these cells were efficiently repaired by HIS3.F and HIS3.R synthetic oligonucleotides indicating that the DSB occurred in these cells (**Table A.3**).

Table 2.1 Frequencies of cDNA and transcript RNA-templated DSB repair in *trans* and in *cis*

a	*trans*			*cis*		
Genotype	His$^+$ freq.		Survival	His$^+$ freq.		Survival
WT	12,300	(10,000-14,600)	1.1%	2,100	(1,800-2,700)	0.7%
spt3	<0.1	(0-8)	8%[1]	<0.1	(0-0)	4.8%
rnh201	33,000	(30,400-42,200)	0.7%	15,800	(11,800-18,300)	0.6%
rnh201 spt3	<0.1	(0-5)	8%	<0.1	(0-0)	7%
rnh1	20,610	(17,100-23,900)	0.8%	1,780	(1,200-2,600)	0.5%
rnh1 spt3	<0.1	(0-5)	9%	<0.1	(0-10)	4.5%
rnh1 rnh201	69,000	(58,600-76,500)	1%	75,000	(57,900-82,100)	0.5%
rnh1 rnh201 spt3	642	(590-800)	11%	6,920	(5,840-7,900)	6%

b	*cis*			*cis*		
Genotype	His$^+$ freq.		Survival	Genotype	His$^+$ freq.		Survival
WT	1,640	(1,200-1,850)	1%	*rnh1 rnh201 rad51*	74,540	(55,130-87,530)	0.09%
rad52	<0.1	(0-0)	0.2%	*rnh1 rnh201 spt3*	7,560	(5,720-11,300)	7.5%
rad51	5,700	(4,170- 8,150)	0.4%	*rnh1 rnh201 spt3 rad52*	520	(300-1,100)	0.3%
rnh1 rnh201	74,600	(64,900-84,000)	0.6%	*rnh1 rnh201 spt3 rad51*	31,560	(12,910-39,220)	0.6%
rnh1 rnh201 rad52	1,520	(970-2,580)	0.1%				

a, Frequencies of His$^+$ colonies per 10^7 viable cells for yeast strains of the *trans* and *cis* cell system following 48 h of galactose treatment are shown as median and 95% CI (in parentheses). Percentage of cell survival after incubation in galactose is also shown. For the *trans* system there were 26 repeats for WT, 12 repeats for *spt3, rnh201, rnh201 spt3, rnh1, rnh1 spt3;* 24 repeats for *rnh1 rnh201, rnh1 rnh201 spt3*. For the *cis* system there were 26 repeats for WT, 12 repeats for *spt3, rnh201, rnh201 spt3, rnh1, rnh1 spt3;* 24 repeats for *rnh1 rnh201;* 18 repeats for *rnh1 rnh201 spt3*. The significance of comparisons between the strains in the *trans* and the *cis* systems, and between different strains of the *trans* or the *cis* system was calculated using the Mann-Whitney U test (**Table A.7a**). **b**, Frequencies of His$^+$ colonies per 10^7 viable cells for different *rad52* and *rad51* mutant strains of the *cis* system following 48 h of galactose treatment are shown as median and 95% CI (in parentheses). For the *cis* system there were 12 repeats for WT, *rnh1 rnh201 spt3*, *rnh1 rnh201 rad52, rnh1 rnh201 spt3 rad52;* 6 repeats for *rad52, rnh1 rnh201, rad51, rnh1 rnh201 rad51, rnh1 rnh201 spt3 rad51*. Percentage of cell survival after incubation in galactose is also shown. The significance of comparisons between the strains in the *cis* systems were calculated using the Mann-Whitney U test (**Table A.7b**). [1]Cells with the *spt3*-null allele have higher survival than wild-type *SPT3* cells after DSB induction because they spend more time in G2 (data not shown; see also **Figure A.2c**).

Next, to examine whether DSB repair frequencies at the *his3* locus in the *trans* and *cis* systems correlate with the expression level of antisense *his3* transcript, we performed quantitative real-time PCR (qPCR). The qPCR data showed that with increased time of incubation in galactose medium (from 0.25 to 8 h) the *trans* strains had significantly more *his3* RNA than the *cis* strains in all backgrounds, including the *rnh1 rnh201 spt3* strain. Furthermore, the levels of *his3* transcript dropped significantly from 0.25 to 8 h in galactose in *cis* but not in *trans* strains, except for the *cis* strain in which the homothallic switching endonuclease gene was deleted (**Figure A.2d**). These results are expected in

the *cis* strains because as soon as the homothallic switching endonuclease DSB is made, a full *his3* transcript cannot be generated. Therefore, these data corroborate the conclusion that the higher frequency of His+ colonies obtained in *cis-* than in *trans*-system *rnh1 rnh201 spt3* cells (**Figure 2.1c and Table 2.1a**) is not due to more abundant and/or more stable transcript but rather to the proximity of the transcript to the target DNA.

PCR analysis of ten random His⁺ colonies from each of the *trans-* and the *cis*-system *rnh1 rnh201 spt3* backgrounds, and Southern blot analysis of three samples from each background showed that the *his3* locus that was originally disrupted by the homothallic switching endonuclease site (trans background), or by the intron with the homothallic switching endonuclease site (*cis* background), was indeed corrected to an intact *HIS3* sequence. No integration of the *HIS3* gene at the homothallic switching endonuclease site or elsewhere in the genome was detected in tested clones (20 of 20), excluding possible mechanisms of repair via capture of cDNA by end joining or via transposition (**Figure 2.2a and Figures A.3 and 4a–c**). We also excluded the possibility that double deletion of *RNH1* and *RNH201* resulted in increased level of Ty transposition. In fact, results presented in **Table A.6** show transposition rates a factor of 3–14 lower in null *rnh1 rnh201* than in wild-type cells. This could be due to an increase of non-productive Ty RNA–DNA substrates for the Ty integrase, resulting in abortive integrations and/or titration of the enzyme. Sequence analysis of 24 random His⁺ colonies from the *cis* system *rnh1 rnh201 spt3* background revealed that all 24 clones had the same precise sequence as the spliced antisense *his3* transcript and did not present a typical end joining pattern with small insertion, deletion or substitution mutations (**Figure A.1c and Table A.2b**). These results, together with our observation of no His⁺ colony formation in cells unable to splice the intron in *his3* (**Figure 2.1e and Table A.5d**), strongly support a homologous recombination mechanism of DSB repair by transcript RNA in *cis*-system *rnh1 rnh201spt3* cells.
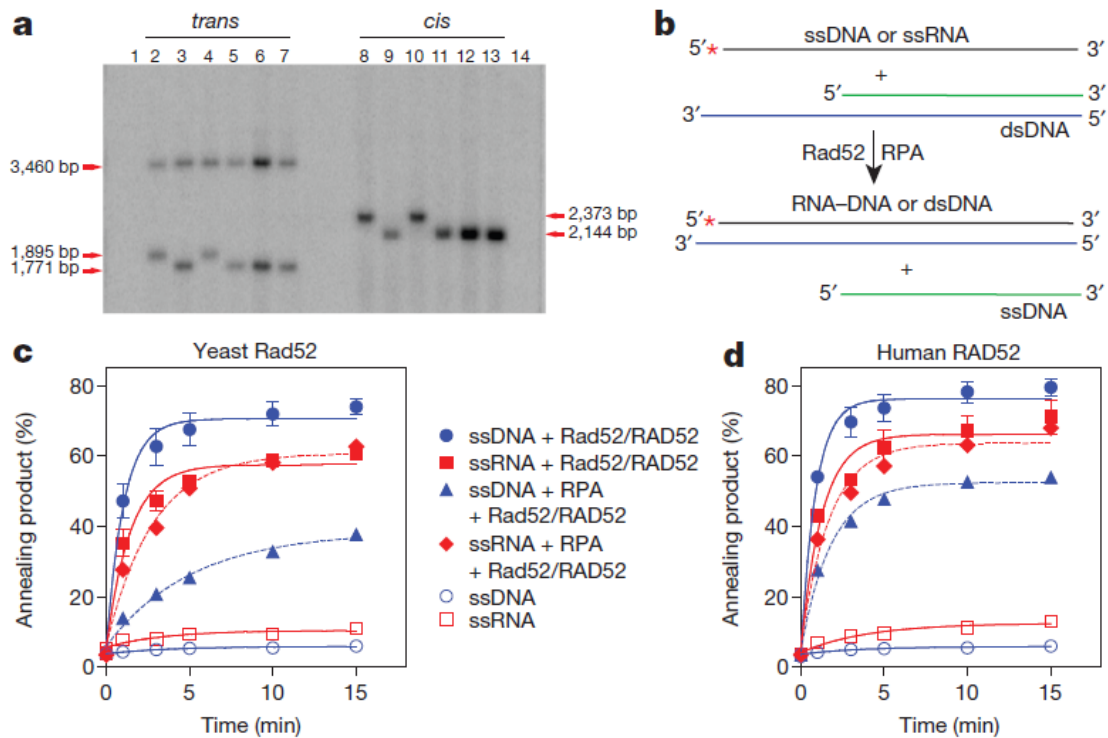
Figure 2.2 Transcript-templated DSB repair follows a homologous recombination mechanism. a, Southern blot analysis of yeast genomic DNA derived from trans wild-type His$^+$ (lane 2) or His$^+$ (lane 3), *rnh1 rnh201 spt3* His$^+$ (lane 4) or His+ (lanes 5–7) cells, digested with BamHI restriction enzyme and hybridized with the HIS3 probe, or derived from *cis* wild-type His$^+$ (lane 8) or His$^+$ (lane 9), *rnh1 rnh201 spt3* His$^+$ (lane 10) or His$^+$ (lanes 11–13) cells, digested with NarI restriction enzyme and hybridized with the *HIS3* probe (**Figure A.4a, c**). Lanes 1 and 14, 1-kilobase DNA ladder visible in the ethidium-bromide-stained gel (**Figure A.4b**). Size of digested DNA bands is indicated by red arrows. bp, base pairs. b, Experimental scheme of Rad52-promoted annealing between RNA and DNA in vitro. Asterisk denotes 32P label. ssDNA (named no. 211) or ssRNA (no. 501) oligonucleotides are in black; DNA oligonucleotides no. 508 and no. 509, forming double-stranded DNA (dsDNA), are in blue and green, respectively. Sequences of oligonucleotides no. 201, no. 501, no. 508 and no. 509 are

shown in **Table A.2a. c, d**, The kinetics of annealing promoted by yeast Rad52 (c) and human RAD52 (d). Nucleoprotein complexes were assembled between dsDNA (no. 508 and no. 509) with an ssDNA protruding tail (0.4nM) and either yeast or human Rad52 (1.35nM) in the presence (dashed lines) or absence (solid lines) of yeast or human RPA (2nM). Annealing was initiated by addition of 32P-labelled ssRNA or ssDNA (0.3nM). The kinetics of protein-free annealing reactions are indicated by open squares and circles. The error bars represent the standard error of the mean, n=4. For the significance of comparisons between the last two time points we used the two-tailed Mann–Whitney U-test. P values are given in **Table A.7c**.

Previous studies showed the ability of Escherichia coli RecA to promote pairing between duplex DNA and single-strand RNA *in vitro* [33, 34]. Recent work suggests that Rad51 (the homologous protein to bacterial RecA) can promote formation of RNA–DNA hybrids in yeast [38]. Here we show that transcript-RNA-directed chromosomal DNA repair is stimulated by the function of Rad52 but not Rad51 recombination protein [95]. Rad52 is important for homologous recombination both via single-strand annealing and via strand invasion [13, 95]. DSB repair by transcript RNA was reduced over 14-fold in *cis*-system *rnh1 rnh201 spt3 rad52* but was increased by a factor of 4 in *cis*-system *rnh1 rnh201 spt3 rad51* compared to *rnh1 rnh201 spt3* cells (**Table 2.1b**). Notably, our in vitro experiments demonstrate that both yeast and human Rad52 efficiently promote annealing of RNA to a DSB-like DNA end (**Figure 2.2b–d and Figure A.4d–h**). Importantly, Rad52 catalyses the reaction with RNA at nearly the same rate as the reaction with single-stranded DNA (ssDNA) of the same sequence. Moreover, in our experiments replication protein A (RPA), a ubiquitous ssDNA binding protein [13], caused a moderate inhibition of Rad52-promoted annealing between complementary ssDNA molecules, but not between ssRNA and ssDNA molecules. Thus, in the presence

31

of RPA, the annealing between ssRNA and ssDNA proceeded with higher efficiency than

the reaction between ssDNA molecules (**Figure 2.2b–d and Figure A.4d–g**).
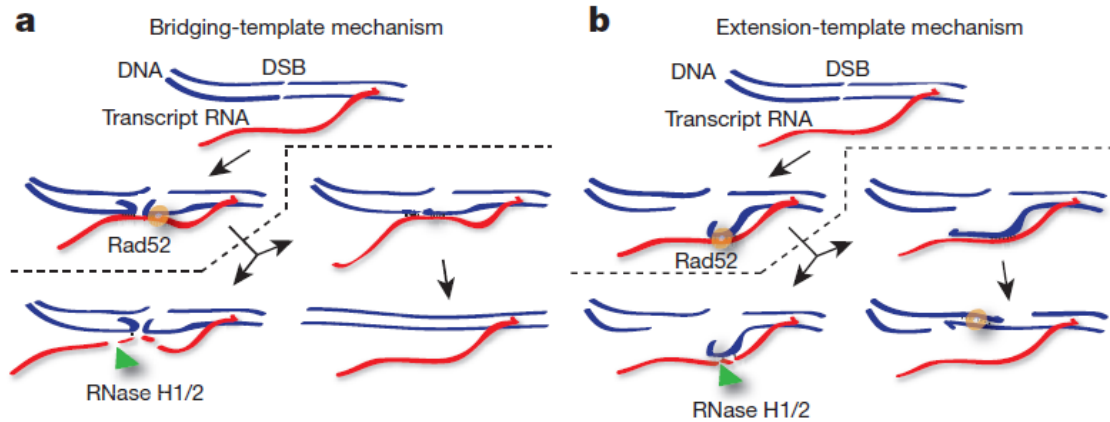
Figure 2.3 Models of transcript-RNA-templated DSB repair in *cis*. An actively

transcribed DNA region experiencing a DSB uses its own transcript RNA as a bridging

(a) or an extension (b) template for repair. The small black lines indicate initial annealing

between the transcript RNA and the DSB end(s), and between the two DSB ends. Orange

circles, Rad52; green triangles, RNase H1 and H2 (H1/2).

In vivo, cDNA and/or RNA-dependent DSB repair may be especially important in the

absence of functional Rad51 that prevents repair by the uncut sister chromatid via strand

invasion [96]. Indeed, our results show that deletion of RAD51 increases the frequency of

repair by cDNA and/or RNA (**Table 2.1b**). Hence, considering the bias observed for

DSB repair in *cis* versus *trans* systems when Ty reverse transcription was impaired, we

propose a model that in the absence of H-type RNase function, transcript RNA mediates

DSB repair preferentially in *cis* systems via a Rad52-facilitated annealing mechanism. In

this mechanism, the transcript may provide a template that either bridges broken DNA

ends to facilitate precise re-ligation or initiate single-strand annealing via a reverse-transcriptase dependent extension of the broken DNA ends (**Figure 2.3**). The reverse transcriptase activity could be provided by a replicative DNA polymerase3, minimal Ty reverse transcriptase, or both. The current view in the field is that RNA–DNA hybrids formed by the annealing of transcript RNA with complementary chromosomal DNA either in *cis* or in *trans* systems are mainly a cause of DNA breaks, DNA damage and genome instability [97]. Here we demonstrate that under genotoxic stress, transcript RNA is recombinogenic and can efficiently and precisely template DNA repair in the absence of H-type RNase function in yeast. In the central dogma of molecular biology, the transfer of genetic information from RNA to DNA is considered to be a special condition, which has been restricted to retro-elements [98] and telomeres [99]. Our data show that the transfer of genetic information from RNA to DNA occurs with an endogenous generic transcript (*his3* antisense), and is thus a more general phenomenon than previously anticipated. In addition, in vitro RNA–DNA annealing was markedly promoted not only by yeast but also human RAD52, suggesting that transcript-RNA-templated DNA repair could occur in human cells. RNA transcripts could template DNA damage repair at highly transcribed loci, in cells that do not divide (lack sister chromatids), or have more stable RNA–DNA heteroduplexes, like those defective in RNASEH2 in patients with Aicardi–Goutie`res syndrome [100]. Our findings lay the groundwork for future exploration of RNA-driven DNA recombination and repair in different cell types.

2.4 Acknowledgements

CHAPTER 3

DEFECTS IN RNASE H2 STIMULATE DNA BREAK REPAIR BY RNA REVERSE TRANSCRIBED INTO cDNA

The study in Chapter 3 consists of the work published in *microRNA 4* (2015)*, 109-116.

Havva Keskin, and Francesca Storici.[1]

[1]School of Biology, Georgia Institute of Technology, Atlanta, GA 30332

3.1 Abstract

Eukaryotic ribonucleases (RNase) H1 and H2 are endonucleases that cleave RNA in a double-stranded RNA-DNA molecule. RNase H2 can also cleave a single ribonucleotide embedded in DNA duplex. While the activity of RNase H1 and H2 has been extensively characterized *in vitro*, still much is unclear about the specific targets of these enzymes *in vivo*. We recently demonstrated that yeast cells can repair a double-strand break (DSB) in DNA by homologous recombination (HR) using antisense (non-coding) RNA, either directly or indirectly after converting RNA into cDNA. In wild-type RNase H1 and/or H2 cells, repair by cDNA dominates, whereas in the absence of RNase H1 and H2 functions cDNA and, in particular, direct transcript-RNA repair mechanisms are markedly stimulated. Here we found that null alleles of any of the three RNase H2 subunits stimulate DSB repair by cDNA significantly more than a null allele of RNase H1. These results show that RNase H2 is the preferred RNase H enzyme to target cDNA in yeast. Targeting of cDNA by RNase H2 does not require RNase H2 interaction with the DNA clamp proliferating cell nuclear antigen (PCNA). Moreover, yeast RNase H2 orthologous mutants of two common RNase H2 defects associated with Aicardi Goutières syndrome (AGS) in humans, displayed elevated cDNA-driven repair of a DSB when combined with each other or with RNase H1 null mutation. Our findings support the hypothesis that defective RNase H2 alleles have higher level of cDNA derived from either coding or non-coding RNA in the form of RNA-cDNA hybrids.

3.2 Introduction

Ribonucleases (RNases) H type 1 and 2 are endonucleases that catalyze the cleavage of RNA in RNA-DNA hybrid duplexes in prokaryotes, archaea and eukaryotes [52]. While RNase H1 has one subunit, eukaryotic RNase H2 has three subunits: a catalytic subunit (Rnh201 in yeast, and RNase H2A in humans) and two auxiliary subunits (Rnh202 and Rnh203 in yeast, and RNase H2B and RNase H2C in humans).

The substrate specificity of RNases H1 and H2 is different. RNase H1 requires a substrate with an RNA stretch containing at least four ribonucleotides in a DNA duplex to allow cleavage [52], while RNase H2 cleaves even a single ribonucleotide embedded in DNA both *in vitro* [52] and *in vivo* [101, 102]. RNase H1 and H2 remove RNA primers during lagging strand synthesis [103-105] and cleave the RNA strand of R-loop structures originated by strand invasion of duplex DNA by nascent or mature, coding or non-coding transcript RNA that form extended RNA-DNA hybrids [50, 52, 106-111]. While both RNase H1 and H2 remove R-loops, recent studies discovered that mitochondrial R-loops are preferentially targeted by RNase H1 [108]. Differently, R-loops associated with replication fork collapse are primarily removed by RNase H2, which may be brought to the target during the process of DNA replication thanks to the presence of a PCNA interacting peptide (PIP) box on the second subunit of RNase H2 [107].

In addition to R-loops, another source of extended RNA DNA hybrids, derived from either coding or non-coding RNA, are RNA-cDNA intermediates of reverse transcription

generated by reverse transcriptase (RT) enzymes of retroelements [112, 113]. Yeast cells contain numerous transposons (Tys) with long-terminal repeats (LTRs). Yeast Tys express an RT enzyme that converts Ty RNA into cDNA in the cytoplasm of yeast cells within Ty virus-like particles (VLPs) [112, 114, 115]. Ty cDNA is released into the cell nucleus, and it is integrated in the yeast genome by the Ty integrase protein or via HR at sites of preexisting LTRs [112]. Little is known about the activity of RNase H1 and H2 on RNA-DNA hybrids of cDNA generated by RT activity. El Hage *et al.* recently showed that most of the RNA-DNA hybrids with the sequence of the Ty1 transposon found in yeast strains with defects in both RNase H1 and H2 were likely derived from cDNA rather than chromosomal R-loops [108]. Studies on RT of yeast Tys and insects R2 retrotransposon showed that i) not only RNA originating from retroelements could be reverse transcribed but potentially any RNA [116], such as the non-coding antisense RNA deriving from the yeast *HIS3* marker gene, and that ii) RNA could mediate recombination with DNA and modify genomic DNA once converted into cDNA via RT [17-19]. Additional studies in yeast revealed involvement of cDNA in HR [26, 117, 118] and it was suggested that different types of RT products including single-strand DNA and RNA-DNA hybrids could be engaged in recombination [115, 119]. Moreover, work in mammals showed that Long INterspersed Elements (LINEs) can be captured at sites of DNA damage, and that retrotransposition of LINEs can carry fragments at their ends that are derived from RT of endogenous RNA [22, 23, 120]. In the work by Keskin *et al.*, we found that knockout of both RNase H1 (*rnh1Δ*) and the catalytic subunit of RNase H2 (*rnh201Δ*) allows not only detection of direct DSB repair by non-coding antisense transcript RNA, but also results in a strong increase of DSB repair by this transcript RNA

that is converted into cDNA by Ty RT [121]. These results suggest that, in the absence of RNase H1 and/or H2, cDNA and/or RNA-cDNA derived from reverse transcription of any coding or noncoding RNA in addition to Ty RNA could be abundant and/or more stable than when RNase H function is normal.

The presence of RNA-DNA hybrids has been proposed as a cause of the congenital immune-mediated neurodevelopmental syndrome of Aicardi Goutières (AGS) in patients with defects in RNase H2 [122, 123]. Majority of AGS patients have a defect in any of the three subunits of RNase H2 [64, 122, 124, 125]. Notably, defects in RNase H1 have not been found in AGS, suggesting that these substrates could be preferentially targeted by RNase H2 [107]. Here, using our system of DSB repair by RNA and cDNA in yeast, we investigated how null alleles of each of the three RNase H2 subunits, as well as two RNase H2 mutants that are orthologous of known AGS defects of RNase H2 in humans (RNase H2A G37S and RNase H2C R69W) impact the frequency of DSB repair.

3.3 Materials and methods

3.3.1 Yeast Strains

Strains used in this study are derivatives of FRO-767 [66] and are shown in Suppl. Table **1**. The *delitto perfetto* method [75, 126, 127] was used to generate the *rnh201*-G42S, *rnh203*-K46W and *rnh202-PIP* mutations by using RNH201.G42S, RNH203.K46W, and 202PIP.F and 202PIP.R oligonucleotides, respectively (**Table B.2**). All mutations were confirmed by sequence analysis of PCR products obtained from amplification of a DNA region surrounding the mutation sites.

3.3.2 Patch and Fluctuation Assays

Experiments were done as previously described [121]. Briefly, in the patch assay, yeast cells were patched and grown on YPD (rich medium) for 1 day at 30°C. Then, cells were replica-plated on galactose containing medium (YPGal) for 2 days at 30 °C. After incubation on YPGal, cells were replica-plated on histidine minus medium (SC-His-) and grown for 3 days at 30 °C. After colonies were grown on SC⁻His⁻, plates were photographed. In the fluctuation assay, cells were inoculated in 50 ml liquid lactose containing medium (YPLac) and incubated for 24 h at 30 °C in a shaker. Next day, cells were counted and 10^7 or 10^8 cells were plated on YPGal medium. For the cell survival, 10^4 cells were plated on YPGal medium. After 2 days incubation at 30 °C, cells were replica-plated on SC⁻His⁻ and grown for 3 days 30 °C. We also replica-plated the cells plated on YPD to SCHis- as a negative control (no DSB) (**Table B.3**). The DSB repair frequency and the survival were calculated as previously described [121].

3.3.3 Data Presentation and Statistics

GraphPad Prism 5 (GraphPad Software, La Jolla, CA) was used to make graphs and conduct statistical analysis of the data. The results are each expressed as median and 95% confidence limits in Tables, and as mean and 95% confidence limits in graphs. The nonparametric two-tailed Mann-Whitney *U*-test [86] was used to calculate statistical significant differences between the His+ frequencies, and *P*-values are shown in **Table B.4**.

3.4 Results

3.4.1 The Experimental System

Following our finding that null mutants of *RNH1* and/or *RNH201* promote repair of DNA

DSBs via HR by RNA and cDNA, and that such repair mechanism is abolished when the

RNH201-null mutation is complemented by expression of wild-type *RNH201* from a

vector [121], we set up to test how different mutants of RNase H2 affect the frequency of

DSB repair by RNA and cDNA. To allow detection of DSB repair directly by transcript

RNA or indirectly by cDNA we utilized an experimental system we recently developed

in yeast *Saccharomyces cerevisiae* [121]. In this system, the *HIS3* marker gene is

interrupted by an artificial intron in the antisense orientation and contains the homothallic

switching (HO) endonuclease cutting site in the middle of the intron sequence. The

expression of the HO nuclease and the transcription of the *his3* antisense are regulated by

the galactose inducible promoter (*pGAL1*). Yeast strains are auxotrophic for histidine

(His$^-$), therefore they do not grow on media without histidine (His$^-$). After induction of

the HO DSB, the broken *his3* allele can be repaired to functional *HIS3* producing His$^+$

cells only by recombination with a homologous sequence [121]. The spliced *his3*

antisense RNA, which is a non-coding RNA, can serve as template for DSB repair of

broken *his3* DNA either directly, or indirectly after being converted into RNA-cDNA

hybrids and/or double-stranded (ds) cDNA by Ty RT within the VLPs [121] (**Figure

3.1**).

3.4.2 Null Alleles of RNase H2 Stimulate DSB Repair by cDNA, and the *rnh203*-null Mutant is the Most Effective

We first examined the capacity of the spliced *his3* antisense to repair the HO DSB in its DNA gene by RNA either directly or indirectly using null mutants of the three RNase H2 subunits. We utilized strains YS-291, 292 containing wild-type RNase H function, *rnh201*-null cells [121], and constructed *rnh202*-null and *rnh203*-null cells (**Table B.1**). The DSB repair assay was performed by culturing yeast cells in patches on solid medium (patch assay), or in liquid medium (fluctuation assay) (see Materials and Methods). Results shown in **Figure 3.2A and 3.2B** and **Table 3.1** demonstrate that deletion of *RNH202* or *RNH203*, similarly to deletion of *RNH201*, strongly increases the frequency of DSB repair at the *his3* locus. However, we note that *rnh203*-null displays a reproducible higher frequency of His$^+$ colonies than *rnh201*-null and *rnh202*-null both in the patch and in the fluctuation assays (**Figure 2.2A, B**).
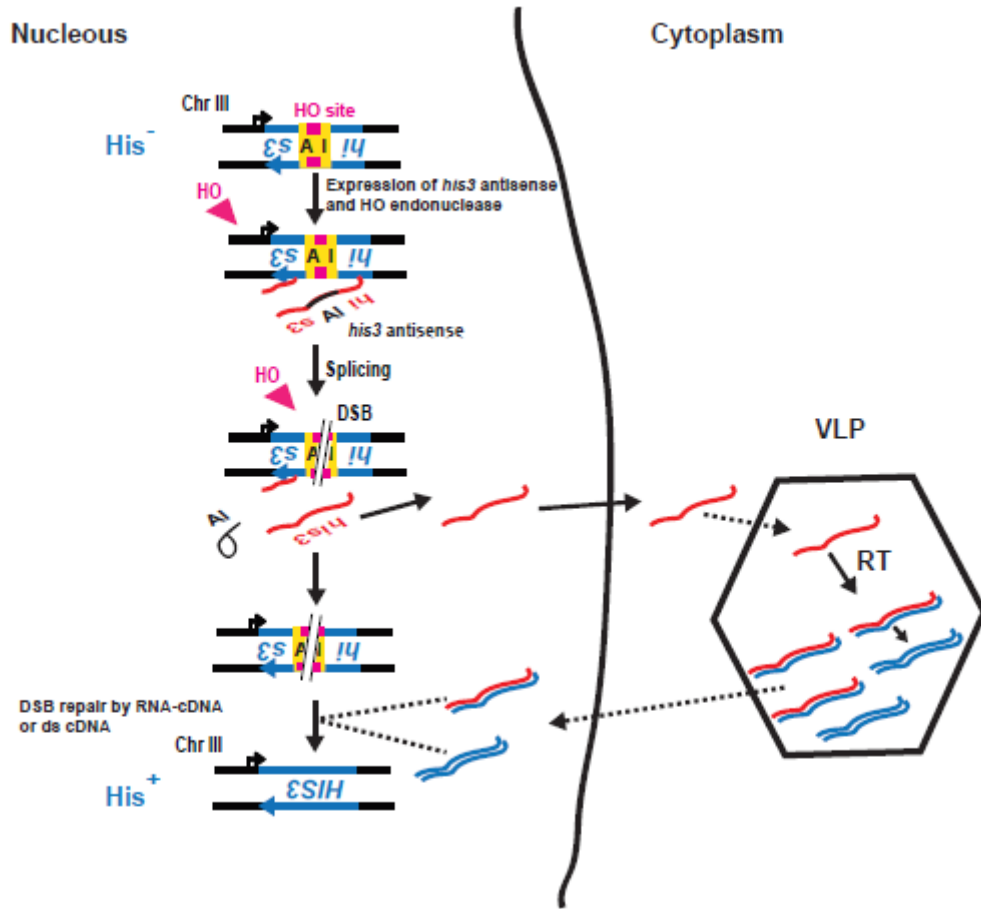
Figure 3.1 Scheme of the system to assay DSB repair by cDNA in yeast. A sketch of

the locus on chromosome III is shown containing one copy of the *his3* gene disrupted by

an artificial intron (AI, yellow), which contains the site for the homothallic switching

(HO) endonuclease (pink). In this configuration the disrupted *his3* gene is non-functional

and yeast cells are unable to grow on medium lacking histidine (His⁻cells). In the

presence of galactose, the *his3* antisense is transcribed (red wavy line) and the HO

endonuclease is also expressed (pink triangle). Upon splicing, the intron (AI, shown as

black short line) is removed as a lariat from the antisense RNA. The newly formed HO

protein makes a DSB in *his3* DNA. Following transcription and splicing in the nucleus,

the *his3* antisense RNA can be converted into cDNA in the form of RNA-cDNA (red and

43

blue parallel lines) or ds cDNA (blue parallel lines) by Ty RT within the Ty VLPs in the

cytoplasm of yeast cells. When the content of VLPs is released into the nucleus, RNA-

cDNA and ds cDNA molecules can serve as templates to repair the DSB in *his3* and

reconstitute a function *HIS3* gene, which produces histidine prototrophic (His$^+$) cells.
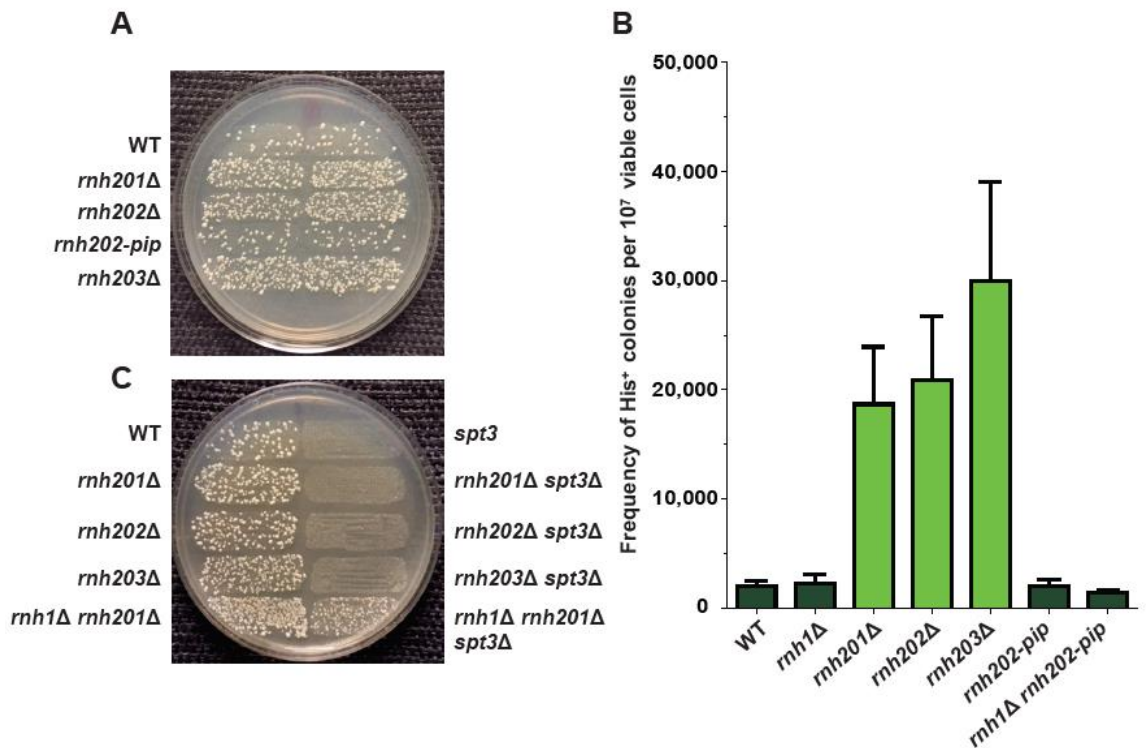
**Figure 3.2 Strong stimulation of DSB repair by cDNA in null mutants of each RNase
H2 subunit.** (**A**) Replica-plating results of the patch assay for the indicated yeast

genotypes. Two repeats are shown for each genotype. Shown are yeast colonies on His-

medium, demonstrating the ability of the indicated yeast strains to form histidine

prototrophic (His$^+$) colonies. (**B**) Results of fluctuation assay with the indicated yeast

genotypes. Data are represented in a histogram graph as frequency of His$^+$ colonies per

10^7 viable cells. Mean and 95% CI are shown; n = 6-12. Bar colors from dark to light

green match the degree of activity (from highest to lowest, respectively) of WT and RNase H1 and/or H2 mutants to suppress DSB repair by cDNA. These data are also presented in **Table 3.1**. (**C**) Replica-plating results of the patch assay for the indicated yeast genotypes containing wild-type or null *SPT3* gene. Yeast colonies on His⁻ medium are shown, demonstrating the ability of the indicated yeast strains to form histidine prototrophic (His$^+$) colonies.

Table 3.1 Frequencies of cDNA templated DSB repair in mutants of RNase H2 subunits.

Genotype	His$^+$ freq.		Survival
WT	2,160	(1,485-2,780)	1%
rnh11'	2,077	(1,390-3,090)	1%
rnh2011'	25,325	(22,570-29,855)	1.2%
rnh2021'	30,050	(17,210-44,040)	0.9%
rnh2031'	50,000	(46,330-51,660)	0.7%
rnh202-pip	2,044	(1,340-2,670)	0.75%
rnh11' rnh202-pip	1,390	(1,240-1,620)	0.9%

Frequencies of His$^+$ colonies per 10^7 viable cells for the indicated yeast strains following 48 h of galactose treatment are shown as median and 95% CI (in parentheses). Percentage of cell survival after incubation in galactose is also shown. There were 12 repeats for WT and *rnh202-pip,* 6 repeats for *rnh1, rnh201, rnh202, rnh203,* and *rnh1 rnh202-pip.* The significance of comparisons between different strains of the system was calculated using the Mann-Whitney *U*-test (**Table B.4A**).

DSB repair in *rnh201Δ*, *rnh202Δ*, and *rnh203Δ* cells was mainly driven by ds cDNA or RNA-cDNA, rather than directly by RNA, because the frequency of His[+] colonies dropped dramatically when the *SPT3* gene was deleted in these cells (**Figure 3.2C**). Spt3, which is a component of the Spt-Ada-Gcn5-acetyltransferase (SAGA) complex of transcription activation in yeast, is required for normal activation of Ty transcription [128]. In *spt3*-null cells, Ty RT expression is severely compromised [71, 121]; thus conversion of RNA to cDNA is impeded. Only in the double *rnh1Δ rnh201Δ* mutant the number of His[+] colonies remain high in the absence of *SPT3*, due to direct DSB repair by transcript RNA and not by cDNA [121] (**Figure 3.2C**). Overall, these data show that RNase H2 suppresses DSB repair by cDNA, suggesting that the reverse transcribed products of *his3* antisense RNA contain abundant RNA-cDNA hybrids.

We then tested whether the suppressive function of RNase H2 on *his3* cDNA was dependent on the interaction of RNase H2 with PCNA. We mutated the PCNA interacting peptide box that is present in Rnh202 [52, 107, 129] to make *rnh202*-FF346, 347AA (*rnh202-pip*). The PIP-box is a highly conserved region in eukaryotic RNASEH2B proteins [129]. The two Phe residues (FF) in the PIP-box are also well conserved, and mutation of these two Phe residues to Ala in human RNASEH2B greatly decreases the amount of PCNA interacting with RNase H2 [129], but does not affect RNase H2 activity [52]. Results presented in **Figure 3.2A, B** and **Table 3.1** clearly show that cells expressing RNase H2 with defective PIP-box have no increased frequency of DSB repair by cDNA compared to cells with wild-type RNase H2.

3.4.3 RNase H2 has Stronger Activity on cDNA than RNase H1

When comparing the frequency of His+ in RNase H2 defective cells with that obtained in RNase H1 defective cells, we found that null alleles of RNase H2 stimulated DSB repair by cDNA much more than a null RNase H1 mutant (**Figures 3.2B and 3.3, Tables 3.1 and 3.2**). The highest stimulation of DSB repair in *his3* by cDNA was seen when both RNase H1 and H2 were non-functional ([121], **Figure 3.3 and Table 3.2**). However, as in null RNase H1 and H2 cells both indirect and direct DSB repair by RNA occur [121], we do not have an accurate measure of cDNA repair without the contribution of direct RNA-templated DSB repair in these cells.

3.4.4 Yeast Orthologous of Human RNase H2A-G37S and RNase H2C-R69W Defects Associated with AGS Stimulate DSB Repair by cDNA

Defects in any of the three subunits of RNase H2 are associated with AGS in humans [100]. The specific reason why RNase H2 defects cause AGS is still uncertain. Here we examined the effect of two AGS mutations in our DSB repair assay in yeast. We constructed yeast strains containing the *rnh201*-G42S or the *rnh203*-K46W mutation, which correspond to RNase H2A-G37S and RNase H2C-R69W defects found in AGS patients, respectively [52]. These AGS mutations were made either in wild-type or *rnh1*-null cells, and we also made cells containing both of these mutations (**Table B.1**). While the *rnh201*-G42S or the *rnh203*-K46W mutation did not significantly alter DSB repair by cDNA, each of these mutations significantly increased the frequency of His+ cells of *rnh1*-null cells (**Figure 3.3 and Table 3.2**). On the contrary, the *rnh202-pip* mutation did not increase the His+ frequency of *rnh1*-null cells (**Figure 3.2B and Table 3.1**).

Moreover, the double *rnh201*-G42S *rnh203*-K46W mutant significantly increased His$^+$ frequency compared to wildtype, *nh201*-G42S or *rnh203*-K46W cells, indicating that these AGS mutations do increase DSB repair by cDNA (**Figure 3.3 and Table 3.2**).
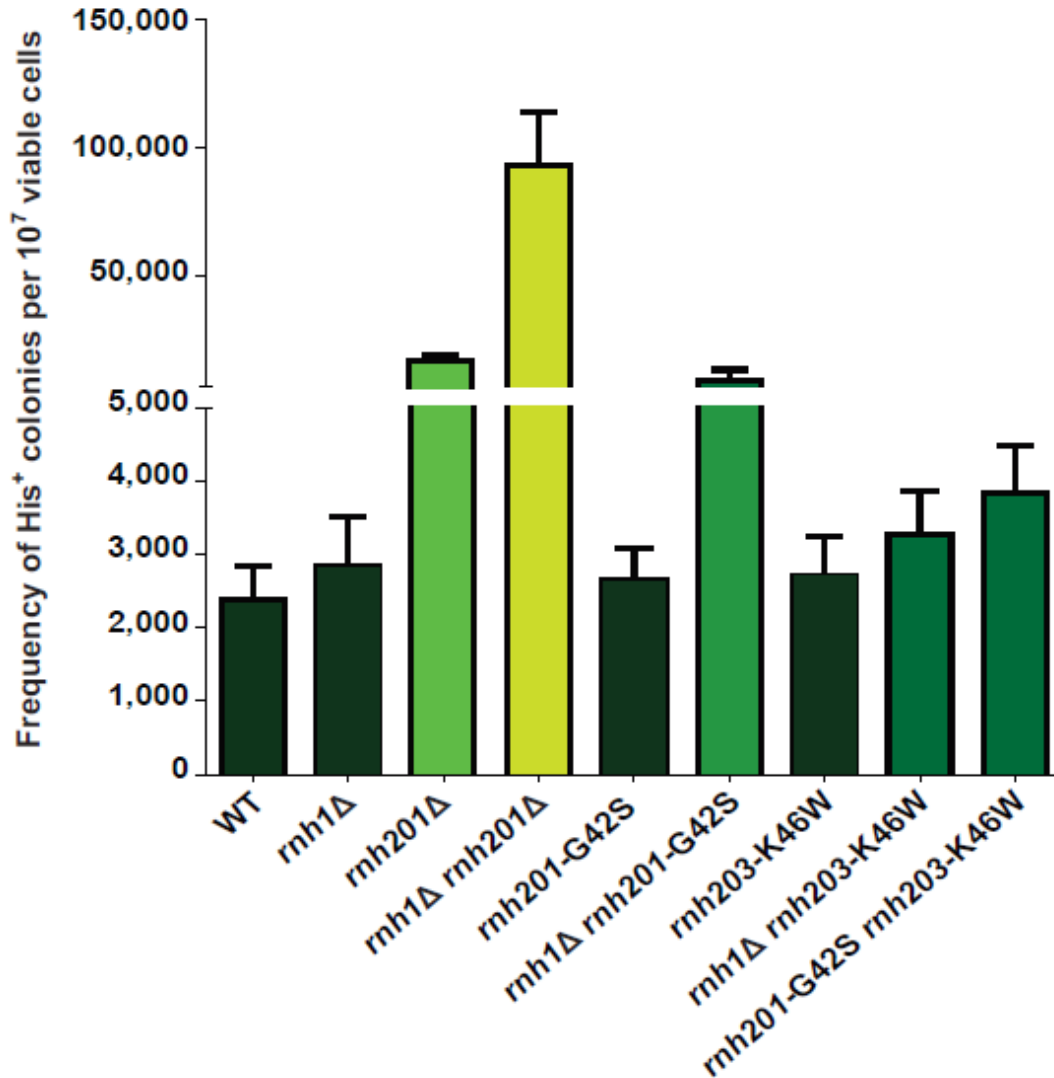
Fig. 3.3 Yeast AGS orthologous mutants of RNase H2 stimulate DSB repair by cDNA. Results of fluctuation assay with the indicated yeast genotypes. Data are represented in a histogram graph as frequency of His$^+$ colonies per 10^7 viable cells. Mean and 95% CI are shown; n= 12-18. Bar colors from dark to light green match the degree of

activity (from highest to lowest, respectively) of WT and RNase H1 and/or H2 mutants to suppress DSB repair by cDNA. These data are also presented in **Table 3.2**.

Table 3.2 Frequencies of cDNA-templated DSB repair in AGS mutants.

Genotype	His+ freq.		Survival
WT	2,470	(1,920-2,840)	1.4%
rnh1ti	2,710	(2,170-3,510)	1%
rnh201ti	16,790	(15,230-19,040)	2.2%
rnh1ti rnh201ti	84,770	(71,950-113,650)	1.2%
rnh201-G42S	2,430	(2,210-3,080)	0.9%
rnh1ti rnh201-G42S	5,820	(4,570-12,810)	0.9%
rnh203-K46W	2,560	(2,165-3,240)	1.5%
rnh1ti rnh203-K46W	2,950	(2,650-3,870)	0.9%
rnh201-G42S *rnh203*-K46W	3,660	(3,190-4,470)	1.3%

Frequencies of His+ colonies per 10^7 viable cells for the indicated yeast strains following 48 h of galactose treatment are shown as median and 95% CI (in parentheses). Percentage of cell survival after incubation in galactose is also shown. There were 18 repeats for WT, *rnh1, rnh201, rnh201*-G42S, *rnh1 rnh201*-G42S, *rnh203*-K46W, *rnh1 rnh203*-K46W, and *rnh201*-G42S *rnh203*-K46W; 12 repeats for *rnh1 rnh201*. The significance of comparisons between different strains of the repair assay was calculated using the Mann-Whitney *U*-test (**Table 3.4B**).

3.5 Discussion

In this study, we show that not only deletion of the catalytic subunit of RNase H2 (*rnh201Δ*) [121], but also deletion of *RNH202* or *RNH203* markedly increases the frequency of DSB repair by cDNA in yeast cells. These data are consistent with the fact

that deletion of any RNase H2 subunit eliminates RNase H2 activity in yeast extracts and that all three subunits are required for RNase H2 cleavage activity [130]. Moreover, we noted that *rnh203Δ* cells have significantly higher frequency of His$^+$ colonies than *rnh201Δ* or *rnh202Δ* cells in our DSB repair assays by cDNA. It is possible that in the absence of Rnh203, cDNA binding by Rnh202, which among the three RNase H2 subunits is the only one capable to bind an RNA-DNA hybrid substrate *in vitro* [93], prevents in part binding and cleavage of cDNA by RNase H1. Therefore, absence of Rnh203 not only abolishes RNase H2 activity *in vivo* but also may in part interfere with RNase H1 function on cDNA.

Compared to *rnh1Δ* mutant, null alleles of any RNase H2 subunit displayed stronger stimulation of cDNA-driven repair. In addition, the fact that PCNA interaction is dispensable for the role of RNase H2 in impeding DSB repair by cDNA supports a function of RNase H2 outside of DNA replication and repair, such as cleavage of RNA-cDNA products of reverse transcription. It is likely that RNA-cDNA molecules are abundantly formed as intermediates during the reverse transcription of coding or non-coding RNA into dscDNA. In fact, our work and the recent study by El Hage *et al.* suggested the presence of copious RNA-cDNA hybrids in RNase H1 and H2 defective cells [108, 121]. Defects in RNase H1 and H2 do not increase DSB repair frequency via HR by DNA molecules. The frequency of DSB repair by DNA oligos in *rnh1Δ rnh201Δ* cell is the same as that obtained in wild-type cells [121]. Thus, we believe that the stimulatory effect on the frequency of DSB repair by cDNA in RNase H1 and H2

defective cells, and in particular in RNase H2 defective cells, is due to specific lack of

cleavage activity of the enzyme/s on RNA-cDNA hybrid molecules (**Figure 3.4**).
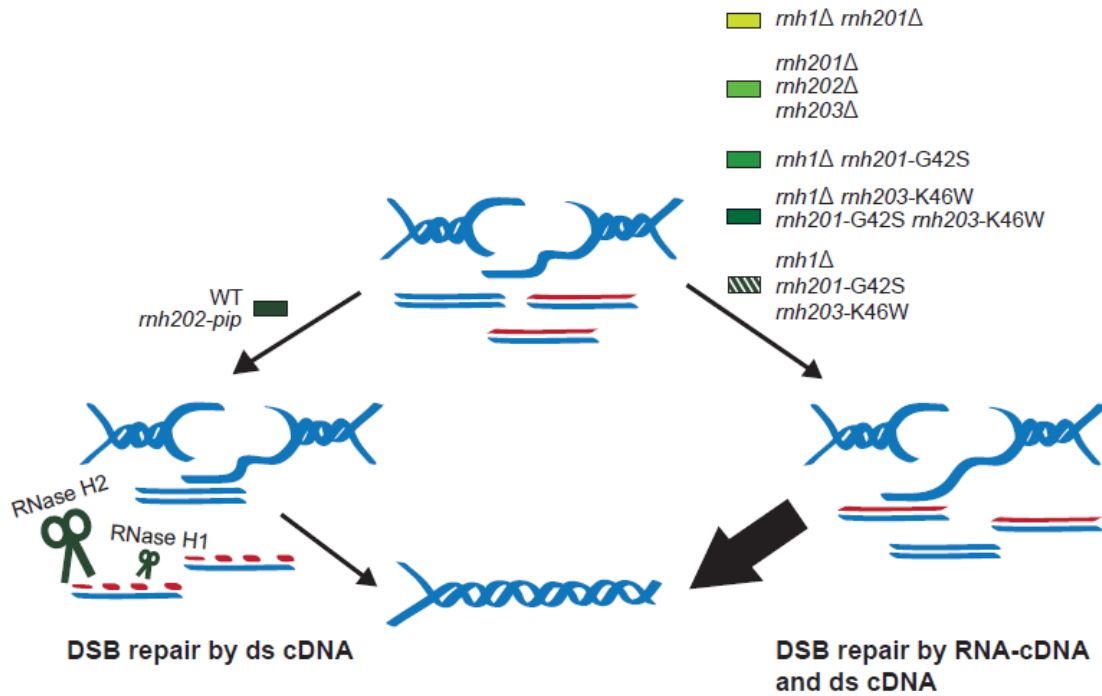
Figure 3.4 Model for cDNA templated DSB repair in RNase H1 and H2 wild-type

and mutants of this study. Two alternative ways for DSB repair by cDNA are shown.

The cDNA can be present as ds cDNA or RNA-cDNA. DNA strands are shown in blue

and RNA strands in red. RNase H2 and H1 are shown as dark green scissors. RNase H2

is shown as bigger scissors because it has stronger activity on cDNA than RNase H1 in

our assays. Box colors from dark to light green match the degree of activity (from highest

to lowest, respectively) of WT and RNase H1 and/or H2 mutants to suppress DSB repair

by cDNA. The white-striped dark green box points out that the indicated mutants lost

some activity on cDNA, although this activity was not significantly different from WT

activity in our assays. In WT RNase H2 and H1 cells, as well as in *rnh202-pip* mutant

cells the DSB is repaired by ds cDNA because the RNA strand of the RNA-cDNA hybrid

is degraded (dotted red line). The various RNase H2 and/or H1 mutants significantly

increase the DSB repair frequency by cDNA to different extent because they have

impaired cleavage of RNA-cDNA molecules, which are then more abundant and/or more

stable templates for DSB repair by HR. The more efficient DSB repair in RNase H2

and/or H1 mutants is indicated by the large black arrow.

Because mutations in human RNase H2 constitute the most frequent genetic defects in

AGS patients, we investigated the impact of two common AGS mutations on DSB repair

by cDNA, by making the orthologous changes in the DNA of yeast RNase H2. While the

single *rnh201*-G42S or *rnh203*-K46W mutation did not significantly increase DSB repair

by cDNA in our DSB repair assay, each of these mutations in combination with *rnh1*-null

allele did increase cDNA-driven DSB repair. The stronger stimulation of DSB repair by

cDNA observed in *rnh1 rnh201*-G42S compared to *rnh1 rnh203*-K46W cells ($P =$

0.0046, **Table 3.4B**) is likely due to the much lower activity of Rnh201-G42S compared

to Rnh203-K46W on RNA-DNA substrates. The AGS mutants RNase H2A-G37S and

RNase H2C-R69W have reduced catalytic activity *in vitro*, with RNase H2A-G37S being

the least active [129]. The corresponding yeast mutant of human RNase H2A-G37S

(Rnh201-G42S) also has very little activity *in vitro*, while yeast Rnh203-K46W retains

about 70% specific activity relative to the wild-type protein [107, 131]. It would be

interesting to examine additional AGS mutants and/or other RNase H2 defects with

differential cleavage activities on long RNA-DNA hybrid substrates for their *in vivo*

capacity to stimulate DSB repair by cDNA. This experiment will help to verify whether the frequency of DSB repair by RNA-cDNA observed in yeast cells expressing different variants of RNase H2 inversely correlates with the activity of these proteins on RNA-DNA hybrids.

When the *rnh201*-G42S and *rnh203*-K46W mutations were combined together in the same cells, DSB repair by cDNA was significantly increased over the frequency obtained in wild-type cells, suggesting that each mutation does slightly reduce RNase H2 function, and only an additive effect of these alleles becomes evident in our assay. Our results indicate that these AGS mutations may likely increase abundance/stability of RNA-cDNA forms (**Figure 3.4**).

While defects in RNase H2 constitute the majority of AGS associated mutations, other factors have been involved in the disease, including defects in the DNA exonuclease TREX1, the Sam domain and HD domain containing protein (SAMHD1), adenosine deaminase acting on RNA (ADAR1) and the cytosolic double-stranded RNA receptor gene (IFIH1) [132, 133]. Notably, all these factors either directly or indirectly play a role in retroelement metabolism. ADAR1 edits of ds RNA of Alu sequences and could have a role in their degradation. IFIH1 is important for sensing dsRNA [134]. SAMHD1, which is an deoxynucleoside triphosphate triphosphohydrolase has a role in inhibiting reverse transcription of the human immunodeficiency virus type 1 (HIV-1) [135], and TREX1 degrades single-stranded DNA derived from retroelements [61]. Therefore, defects in these genes may lead to accumulation of nucleic acids derived from reverse transcription that could activate the immune response in AGS patients [122, 123]. In line with this

hypothesis, fresh results by Lim *et al.* revealed that fibroblasts derived from AGS patients with defects in TREX1, RNASEH2A, RNASEH2B, and SAMHD1 have an excessive amount of RNA-DNA hybrids [136]. RNA-DNA hybrids can frequently occur in cells. They can form on chromosomal DNA at telomeres [137], as a consequence of transcription in R-loops [110, 138], with primer synthesis in DNA replication [104], following binding of microRNAs to target DNA in transcriptional gene silencing [32], or during RNA-driven break repair [66, 121]. Moreover, as highlighted in this work, RNA-DNA hybrids can also form extrachromosomally as intermediate structures in the process of reverse transcription and generation of cDNA [15]. The marked impact of RNase H2 on cDNA-driven repair observed in our study reveals a major role of RNase H2 in targeting cDNA in the form of RNA-cDNA hybrids in yeast. Our results support a model in which RNA-cDNA molecules generated by reverse transcription of RNA derived from retroelements and potentially from any coding or non-coding RNA could be abnormally abundant in AGS patients and activate an immune response.

3.6 Acknowledgements

CHAPTER 4

TRANSCRIPT RNA SUPPORTS PRECISE REPAIR OF ITS OWN

DNA GENE

The study in Chapter 3 consists of the work published in *RNA Biology*, *2* (2016)*,* 157-165.

Havva Keskin, Chance Meers and Francesca Storici.[1]

[1]School of Biology, Georgia Institute of Technology, Atlanta, GA 30332

4.1 Abstract

The transfer of genetic information from RNA to DNA is considered an extraordinary process in molecular biology. Despite the fact that cells transcribe abundant amount of RNA with a wide range of functions, it has been difficult to uncover whether RNA can serve as a template for DNA repair and recombination. An increasing number of experimental evidences suggest a direct role of RNA in DNA modification. Recently, we demonstrated that endogenous transcript RNA can serve as a template to repair a DNA double-strand break (DSB), the most harmful DNA lesion, not only indirectly via formation of a DNA copy (cDNA) intermediate, but also directly in a homology driven mechanism in budding yeast. These results point out that the transfer of genetic information from RNA to DNA is more general than previously thought. We found that transcript RNA is more efficient in repairing a DSB in its own DNA (in *cis*) than in a homologous but ectopic locus (in *trans*). Here, we summarize current knowledge about the process of RNA-driven DNA repair and recombination, and provide further data in support of our model of DSB repair by transcript RNA in *cis*. We show that a DSB is precisely repaired predominately by transcript RNA and not by residual cDNA in conditions in which formation of cDNA by reverse transcription is inhibited. Additionally, we demonstrate that defects in ribonuclease (RNase) H stimulate precise DSB repair by homologous RNA or cDNA sequence, and not by homologous DNA sequence carried on a plasmid. These results highlight an antagonistic role of RNase H in RNA-DNA recombination. Ultimately, we discuss several questions that should be addressed to better understand mechanisms and implications of RNA-templated DNA repair and recombination.

4.2 Materials and methods

4.2.1 Yeast strains and plasmids

The background strain used to develop all strains used in this study is the haploid FRO-767 strain (*leu2*::HOcs, matαΔ::*hisG*, *ho*Δ, *hml*Δ::*ADE1*, *hmr*Δ::*ADE1*, *ade1*, *leu2*–3,112, *lys5*, *trp1*::*hisG*, *ura3*–52, *ade3*::GAL::HO) [121]. BDG283 and BDG998 vectors (gifts from D. Garfinkel) were transformed into YS-291, 292 (WT) and YS-486, 487 (*spt3 rnh1 rnh201*) strains. BDG283 contains only *pGAL1* and BDG998 contains the *pGAL1-mhis3-AI* cassette, and both plasmids are centromeric with the *URA3* marker [17]. YCp50pK and phis3.210 vectors are also centromeric with the *URA3* marker. YCp50pK was constructed by cloning a SalI/EcoRI fragment with the kanMX4 gene from pFA6a-kanMX4 plasmid52 into the EcoRI/SalI sites of YCp50.53. To construct the phis3.210 vector, a 210-bp fragment of HIS3 was amplified by PCR from genomic DNA using forward primer 50-ACAGTGCTAAGT-AAGCTTATCTTCCCAGAAAAAGAGGC-30 (HindIII site underlined) and reverse primer 50-ATTGAGTTCCTA-AAGCTT-TACCACCGCTCTGGAAAGTG-30 (HindIII site underlined). The PCR product was digested with HindIII enzyme and was ligated into the YCp50pK vector, which was also digested with HindIII within the kanMX4 gene. The resulting plasmid was sequenced to confirm the correct 210-bp *HIS3* insert. Both YCp50pK and phis3.210 were transformed into YS-291, 292 (WT), YS-444, 445 (*rad52*), YS-424, 426 (*rnh1 rnh201*), YS-490, 491 (*rad52 rnh1 rnh201*), YS-440, 441 (*spt3*), and YS-486, 487 (*spt3 rnh1 rnh201*) strains. Genetic methods and standard media were described previously [127].

4.2.2 Fluctuation assay of DSB repair

All strains carrying a plasmid with the *URA3* marker gene were maintained on Ura$^-$ medium. Fluctuation assays of DSB repair at the *his3* locus were done as previously described [121]. Briefly, yeast cells were grown in 50-ml lactose containing medium (YPLac), and incubated for 24 hours at 30 °C. Next day, cells were counted and 10^7 or 10^8 cells were plated on galactose medium (YPGal) or SC-Ura$^-$Gal medium. 10^4 cells were also plated on YPGal or SC-Ura$^-$Gal medium to calculate survival. After 2 d incubation, cells were replica plated on His$^-$ or Ura$^-$His$^-$ medium, and after 3 d His$^+$ or Ura$^+$His$^+$ colonies were counted. Repair frequency and survival were calculated as previously described [121]. Without galactose induction, no or rare His$^+$ clones are obtained, as discussed in reference [121].

Table 4.1 Statistical analysis (P-values) of the data.

A	
Genotype of *cis* system	*P*-value
WT + BDG283 *vs.* WT + BDG998	< 0.0001
WT + BDG283 *vs. spt3 rnh1 rnh201* + BDG283	< 0.0001
WT + BDG998 *vs. spt3 mh1 mh201* + BDG998	< 0.0001
spt3 rnh1 rnh201 + BDG283 *vs. spt3 mh1 mh201* + BDG998	0.4356

B	
WT + YCp50pK *vs.* WT + phis3.210	< 0.0001
WT + YCp50pK *vs. rnh1 rnh201* + YCp50pK	< 0.0001
WT + YCp50pK *vs. spt3 rnh1 mh201* + YCp50pK	< 0.0001
WT + YCp50pK *vs. rnh1 rnh201 rad52* + YCp50pK	< 0.0001
WT + phis3.210 *vs. spt3* + phis3.210	0.5425
WT + phis3.210 *vs. rnh1 rnh201* + phis3.210	< 0.0001
WT + phis3.210 *vs. spt3 mh1 mh201* + phis3.210	0.1938
WT + phis3.210 *vs. rnh1 rnh201 rad52* + phis3.210	0.0001
rnh1 mh201 + YCp50pK *vs. mh1 mh201* + phis3.210	< 0.0001
rnh1 mh201 + YCp50pK *vs. spt3 rnh1 rnh201* + YCp50pK	< 0.0001
rnh1 mh201 + YCp50pK *vs. mh1 mh201 rad52* + YCp50pK	< 0.0001
rnh1 mh201 + phis3.210 *vs. spt3 rnh1 rnh201* + phis3.210	< 0.0001
rnh1 mh201 + phis3.210 *vs. mh1 mh201 rad52* + phis3.210	0.0001
spt3 + phis3.210 *vs. mh1 mh201* + phis3.210	0.0012
spt3 + phis3.210 *vs. mh1 mh201 rad52* + phis3.210	0.0004
spt3 + phis3.210 *vs. mh1 mh201 spt3* + phis3.210	0.0830
spt3 rnh1 rnh201 + YCp50pK *vs. rnh1 rnh201 rad52* + YCp50pK	< 0.0001
spt3 rnh1 rnh201 + phis3.210 *vs. rnh1 rnh201 rad52* + phis3.210	0.0001
mh1 mh201 rad52 + YCp50pK *vs. mh1 mh201 rad52* + phis3.210	0.4990

Mann-Whitney U-test was applied to determine whether a statistical significant difference exists between pairs of gene correction frequencies obtained in DSB repair assays. A, Comparison of frequencies presented in **Table 4.1**. Two groups in a pair were considered to be significantly different when adjusted P-values were less than 0.05. B,

Comparison of frequencies presented in **Table 4.2**. Two groups in a pair were considered to be significantly different when adjusted P-values were less than 0.05.

4.2.3 Data presentation and statistics

Statistical analysis was calculated by using GraphPad Prism 5 (GraphPad Software, La Jolla, CA). Median and 95% confidence limits were expressed for each data sample. Statistical significance differences were calculated by using the nonparametric 2-tailed Mann-Whitney U-test, and all P-values of frequency comparisons are shown in **Table 4.3**.

4.2.4 Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

4.3 Results and discussion

4.3.1 Transfer of genetic information from RNA to DNA: Theory and supporting evidence

Can RNA transfer genetic information to DNA beyond the special cases of retroviruses, retrotransposons and telomere synthesis? [15, 16] Can RNA recombine with DNA either directly or indirectly if converted into cDNA? Studies on reverse transcription mediated by retrotransposons of yeast (Tys), or of insects (R2), have shown that not only RNA originating from retroelements could be reverse transcribed but potentially any RNA [116], such as the RNA deriving from the yeast HIS3 marker gene, and that RNA could

mediate recombination with DNA and modify genomic DNA once converted into cDNA

via reverse transcription [17-19]. It was found that not only Ty cDNA, but also HIS3

cDNA could recombine with homologous or homeologous (partially homologous) DNA

[19], integrate into genomic DNA if fused to transposon sequences, or be captured at sites

of chromosomal DSBs via non-homologous end joining (NHEJ) [20, 21]. Additional

studies in yeast revealed involvement of cDNA in homologous recombination (HR) [26,

117, 118], and it was suggested that different types of reverse transcription products

including ssDNA and RNA-DNA hybrids could be engaged in recombination [119].

Further work in mammalian cells showed that Long INterspersed Elements (LINEs) can

be captured at sites of DNA damage, and that retrotransposition of LINEs can carry

fragments at their ends that are derived from reverse transcription of endogenous mRNA

[22, 23, 120]. There has been a series of hypotheses and speculations that RNA can work

as a template in DNA recombination and repair [139]. Recombination mediated by

reverse transcripts of cellular RNAs with homologous DNA has been suggested to

explain the paucity of introns in yeast genomic DNA, while end-joining-driven insertions

of cDNA products could explain the abundance of pseudogenes in multicellular

eukaryotes [140, 141]. Indeed mRNA-mediated intron losses were shown to occur in

yeast mitochondrial DNA [142] (and references therein). Murakami et al. suggested a

mechanism of RNA-directed DNA repair in mitochondria facilitated by the reverse

transcriptase activity of DNA polymerase gamma [143], whereas possible mechanisms of

DSB repair in nuclear DNA by RNA have been proposed by Trott and Porter [144]. The

discovery of a widespread type of viral genome representing a chimera between an RNA

and a DNA virus has inferred the occurrence of RNA-DNA recombination between two quite different virus groups [37, 145].

From work in plants, Xu et al. proposed a direct or indirect RNA-templated DSB repair mechanism via gene conversion to explain the observed high frequency of gene homozygosity in rice [146]. Furthermore, a recent study reported that DSBs in neurons of young adult mice can be part of normal brain functions such as learning, as long as the DSBs are controlled and repaired in short time [147]. Could RNA serve as template for DNA repair of these physiological DSBs in neurons? It has been proposed that flow of information from RNA to DNA could lead to DNA recoding events in the nervous system and could be the basis for permanent storage of long term memories [148, 149]. Considering the abundance of RNA in cells, the flow of genetic information from RNA to DNA could strongly affect genome stability, either by increasing or decreasing it, depending on the circumstances. Different experimental insights suggest that mechanisms of RNA-driven DNA modification might be more common than is currently recognized. Evidence of RNA-derived insertions came from analysis of sequences at DSB sites in fruit fly and mammalian cells. An exon–exon junction sequence was found from the analysis of DNA sequences repaired via NHEJ after DSB induction by zinc-finger nucleases in Drosophila cells [36], suggesting a direct or indirect RNA-templated insertion mechanism. Work in human cells revealed presence of murine sequences derived from murine RNA that was co-transfected into the human cells together with the DNA of the I-SceI DSB-inducing vector [89]. More recently, exonic RNA insertions were detected in knock-in mouse experiments at sites of DNA DSBs generated using the

CRISPR/Cas9 system [150]. Overall, these studies showed that insertions of RNA derived sequences can result in an error-prone form of DNA repair, which may play a role in genetic disorders and evolution.

4.3.2 Is there experimental proof for RNA-DNA recombination and RNA-mediated DNA repair that is homology driven?

Can RNA directly mediate genetic DNA modifications in a homology-driven manner? Can RNA repair a DSB in homologous DNA sequences? Experiments in budding yeast showed that not only short ribonucleotide tracts carried within synthetic DNA oligonucleotides (oligos) but also RNA-only oligos can precisely repair a DSB in homologous DNA, serving as direct templates for DNA synthesis at the chromosomal level, and transferring genetic information also in conditions in which Ty reverse transcription is repressed [40, 66, 151]. The capacity of short RNA patches to directly modify DNA was also found in the bacterium Escherichia coli [39, 41], and RNA oligos could precisely repair a DSB in the green fluorescent protein gene in human embryonic kidney (HEK-293) cells [39]. As a model to explain the occurrence of transgenerational inheritance of genomic DNA rearrangements in ciliated protozoa, Angeleska et al. proposed a mechanism in which RNA molecules, single- or double-stranded (ss or ds), act as template catalyst to guide specific recombination events [152]. The model for RNA-templated DNA rearrangements was then tested using long synthetic RNA sequences injected into the ciliate *Oxytricha trifallax* and the RNA templates were found to mediate correct and precise DNA rearrangements [35, 153]. In addition, mutations carried on the artificial RNA templates were transferred to the homologous endogenous

DNA sequences suggesting a process of RNA-guided DNA repair in *O. trifallax* [35].

Models of RNA-DNA HR are supported by biochemical studies, showing the ability of

the E. coli recombinase RecA to promote pairing between duplex DNA and ssRNA in

vitro [33, 34, 154, 155].

Moreover, recent work suggests that the eukaryotic RecA homolog, Rad51, can also

promote formation of RNA-DNA hybrids in yeast [38]. Beyond the demonstration that

synthetic RNA molecules introduced into cells can mediate HR with DNA, our recent

work showed that endogenous transcript RNA can be a template for DSB repair and HR

in yeast [121]. We provided experimental evidence that the transfer of genetic

information from RNA to DNA occurs with an endogenous generic transcript, and is thus

a broader phenomenon than previously anticipated.

4.3.3 Transcript RNA mediates DSB repair in a homology-driven manner

We developed a system to explore the prospects of an endogenous RNA transcript ability

to serve as a template for the repair of DSBs, casting a new light on the roles of RNA in

the DNA damage response [121]. Our strategy is based on the induction of a DSB located

inside a nonfunctional *his3* marker gene, and successive DSB repair via an endogenous

spliced transcript RNA resulting in histidine prototrophic (His$^+$) cells. We engineered *cis*

and *trans* systems granting the possibility to evaluate the effects of localization and

continuous productions of the transcript RNA. The *cis* system transcribes an antisense

his3 sequence with an artificial intron inserted in the antisense orientation that upon

galactose induction results in a spliced antisense *his3* transcript that can facilitate repair

of a DSB located inside the artificial intron resulting in a functional *HIS3* locus. The

artificial intron (105 bp) contains the site for the HO endonuclease (124 bp); in total, a 229-bp insert disrupts the *HIS3* gene. Likewise, the trans system is based on dual *his3* loci, in which, one locus is the endogenous *HIS3* gene on chromosome XV but disrupted by the cutting site of the HO endonuclease, and the other locus is located on chromosome III and serves to produce an antisense *his3* transcript with an artificial intron inserted in the antisense orientation that upon galactose induction produces a *his3* antisense transcript that can aid in the repair of the DSB generated at the HO site of the endogenous *HIS3* (**Figure 4.1**). Considering the abundance of retrotransposons in the yeast genome [156], we sought to eliminate the reverse transcription activities associated with retroelements to explore the ability of RNA to serve directly as a template for repair rather than through the cDNA intermediates of retrotransposition. To this end, we created an *spt3*-null mutant, which prevents normal Ty transcription and reduces Ty transposition [71]. As a result, in *spt3*-null mutant yeast, no His[+] colonies are observed suggesting that cDNA-mediated repair is the major pathway of repair in transposition proficient cells [121]. This indicates that any actively transcribed gene can be repaired using a reverse transcribed cDNA template. Because an RNA-DNA heteroduplex is a probable prerequisite for RNA to recombine directly with DNA, we sought to facilitate stable formation of RNA-DNA hybrids by deletion of RNase H1 (*RNH1*) and the catalytic subunit of RNase H2 (*RNH201*) genes, which both code for nonsequence-specific endonucleases that cleave RNA backbone of RNA-DNA hybrids [52]. Deletion of both *RNH1* and *RNH201* results in a 5-fold increase of His[+] colonies in *trans* and a 35-fold increase in *cis*. Surprisingly, the *spt3 rnh1 rnh201* genotype results in more than 69,000 His[+] colonies than in *spt3* single mutant, and even more intriguingly, the *cis* system of the

65

spt3 rnh1 rnh201 cells yields 10-fold more His+ colonies than the *trans* system, which

continuously produces transcript for repair [121]. Furthermore, deletion of the *RAD52*

gene, which codes for an important homologous recombination protein facilitating the

annealing of complementary ssDNA, results in a strong reduction of His+ colonies in *spt3*

rnh1 rnh201 cells [121]. A complementary *in vitro* study suggests that yeast and human

Rad52 can promote the annealing of RNA to DNA, and in the presence of RPA, even

more efficiently than DNA to DNA [121]. Thus, we propose a model that upon the

occurrence of a DSB in a transcribed DNA, *Rad52* promotes the annealing of RNA to

DNA, and, in the absence of RNases H, RNA serves as a template bridging the broken

DNA ends to promote precise re-ligation, or allowing extension of the broken end via

reverse transcription [121]. Given the prerequisite that our assay requires a spliced

mRNA to display a phenotype, we could be missing repair by unspliced mRNA, thus

RNA-templated DNA modifications may have a substantial impacts on genomic stability.
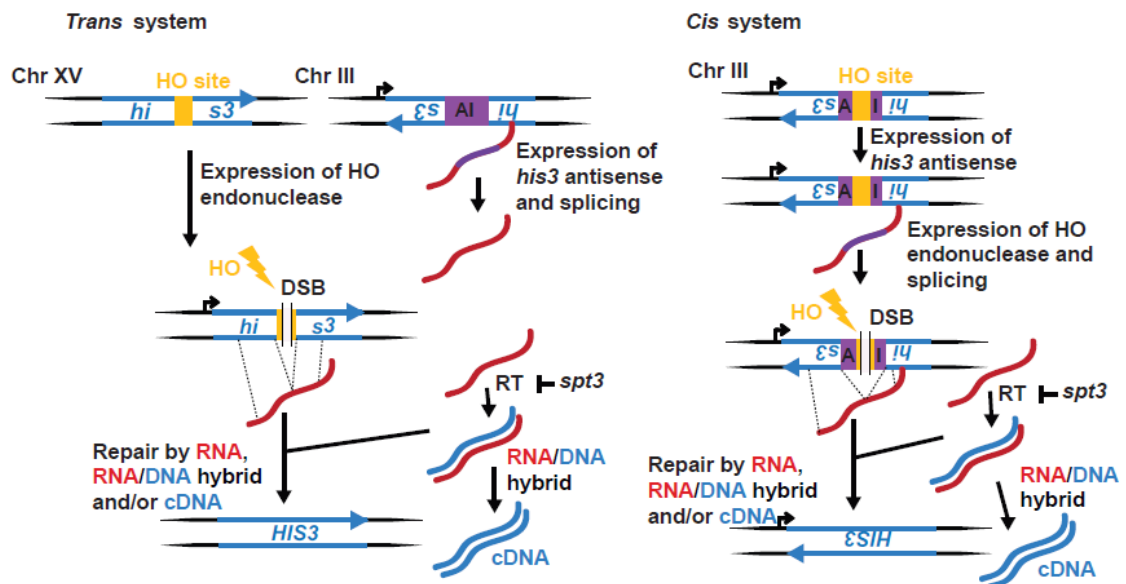
Figure 4.1 Scheme of the *trans* and *cis* systems. HO, homothallic switching endonuclease (yellow); AI, artificial intron (purple); right turn arrow, *pGAL1*; yellow lightning bolt, cleavage activity by HO; RT, reverse transcriptase.

4.3.4 DNA self-repair by transcript RNA

Our results of DSB repair in the *cis* and *trans* systems showed that the frequency of His$^+$ colonies in *cis spt3 rnh1 rnh201* cells was >69,000-fold higher than in *cis spt3* cells, and >10 –fold higher than in *trans spt3 rnh1 rnh201* cells [121]. Is this high frequency of His$^+$ colonies in *cis spt3 rnh1 rnh201* cells due to the RNA functioning as homologous template to mediate a precise re-ligation of the broken DSB ends? Alternatively, is this repair templated by cDNA due to residual Ty activity? We showed that the DSB repair at the *his3* locus in *cis spt3 rnh1 rnh201* cells was predominately mediated by transcript RNA rather than cDNA [121]. Here, we corroborate our finding that transcript RNA can directly serve as a template for repair of a DSB occurring in the same DNA that generated the transcript *in spt3 rnh1 rnh201* cells of the *cis* system.

We examined the effect of an extra copy of the *his3* allele, disrupted by the artificial intron in the antisense orientation (*mhis3-AI*) carried on a yeast centromeric plasmid (BDG998) (**Figure 4.2A**), on the frequency of His$^+$ colonies following DSB induction in wild-type and *spt3 rnh1 rnh201* backgrounds of the *cis* system. We transformed wild-type and *spt3 rnh1 rnh201* cells with low copy number plasmid BDG998 or with the control empty plasmid (BDG283), which carry the *URA3* marker gene (**Figure 4.2A**),

and selected for colonies able to grow on medium lacking uracil (Ura$^+$ colonies). We then

performed the fluctuation assay as described in Materials and Methods and in [121]. In

wild-type cells, the His$^+$ frequency was strongly increased in the presence of the BDG998

plasmid (**Table 4.1 and Figure 4.3**) compared to BDG283-containing cells. This was

expected because not only the *his3* antisense transcribed from the chromosomal *his3*

copy, but also the one transcribed from the *his3* copy carried on the BDG998 plasmid can

be converted into cDNA by Ty reverse transcriptase and provide additional copies for

DSB repair. Moreover, differently from the chromosomal copy, the plasmid copy of *his3*

can continue to be transcribed in galactose medium because it does not contain the site

for the HO endonuclease within the artificial intron, thus, it can generate lots of cDNA

molecules. In contrast, there is no significant difference in the frequency of His$^+$ colonies

between *spt3 rnh1 rnh201* cells containing BDG283 and BDG998 (**Table 4.1 and Figure

4.3**). If cDNA would be the major template for *his3* repair in *spt3 rnh1 rnh201* cells we

would expect higher frequency of His$^+$ colonies also when these cells contain BDG998

than in cells containing BDG283. These data suggest that even if there is residual cDNA

in *cis spt3 rnh1 rnh201* cells, cDNA does not play a major role in DSB repair of the *his3*

locus. Rather, it is the transcript RNA from the chromosomal locus that mediates, in *cis*,

most of DSB repair to restore the function of its broken *his3* gene on the chromosome.

**4.3.5 Defects in RNase H activity stimulate homology-driven DSB repair by cDNA
and RNA, but not by plasmid DNA**

Our findings show that absence of RNase H1 and/or H2 activity in wild-type or null-*spt3*

cells results in increased frequency of His$^+$ colonies after DSB induction not only in the

cis but also in the *trans* system compared to wild-type RNase H cells [121]. These results

indicate that absence of RNase H function activates DSB repair by transcript RNA, and

also stimulates DSB repair by cDNA. Following reverse transcription of RNA into

cDNA, cDNA can be present as RNA-DNA hybrid, ssDNA, and/or dsDNA. Previously,

we showed that DSB repair by ssDNA oligos was not increased in *rnh1 rnh201* cells

compared to RNase H wild-type cells [121]. Moreover, our recent work indicates that

defective RNase H2 alleles have higher level of cDNA in the form of RNA-cDNA

hybrids [157]. Here, we examined whether the RNase H defect is specific to stimulate

DSB repair of the broken *his3* locus via HR only by RNA and/or cDNA, or it can also

stimulate DSB repair by gene conversion using as template for HR a truncated *his3* copy

carried on a dsDNA plasmid. We transformed wild-type, *rad52*, *rnh1 rnh201*, *rnh1*

rnh201 rad52, *spt3* and *spt3 rnh1 rnh201* strains of the *cis* system with a plasmid

carrying an internal 210-bp segment of the *HIS3* gene sequence (phis3.210) or with the

control empty plasmid (YCp50pK) (**Figure 4.2B**). To determine the frequency of His$^+$

colonies following DSB induction at the *his3* chromosomal locus for all these strains, we

conducted the fluctuation assay of DSB repair. Depending on the genotype of the strains,

cells containing the control vector YCp50pK can repair the DSB in the chromosomal *his3*

allele by using as template for HR the RNA, RNA-DNA hybrid and/or cDNA derived

from the chromosomal *his3* locus, while cells containing phis3.210, in addition to the

RNA, RNA-DNA hybrid and/or cDNA derived from the chromosomal *his3* locus, can

also repair the DSB in *his3* by using as template the DNA of the truncated *his3* allele

carried on phis3.210 cDNA, and/or potentially the RNA, RNA-DNA hybrid and/or

cDNA derived from the transcription of this *his3* plasmid allele (**Table 4.2 and Figure**

69

4.4). In wild-type cells, there is a factor of 50 increase in the His$^+$ frequency in the presence of phis3.210 compared to YCp50pK (**Table 4.2**). As expected, upon deletion of the *RAD52* gene, which is required for any mechanism of DNA-DNA HR in yeast [158], no His$^+$ colonies are detected with either plasmid. In *rnh1 rnh201* cells carrying YCp50pK, the His$^+$ frequency is more than a factor of 20 higher than in wild-type cells, due to elevated repair by cDNA and RNA, in agreement with our previous findings [121, 156]. The *rnh1 rnh201* mutations in cells carrying phis3.210 result in less than 2-fold increase of the His$^+$ frequency compared to wildtype cells carrying the same plasmid (**Table 4.2**). Such increase can be explained by the fact that in this background the DSB can be repaired not only by the truncated *his3* on the plasmid, but also by RNA and cDNA derived from the chromosomal *his3* copy, as well as by cDNA derived from the his3 allele on phis3.210. There could also be some repair in *trans* by the RNA derived from the *his3* allele on phis3.210, although we expect this to be minimal compared to repair by cDNA. However, clearly, defects in RNase H1 and H2 do not stimulate DSB repair by the DNA of the *his3* copy on the plasmid. In fact, in *spt3* mutant cells, in which there is no or very little cDNA, there is no difference in the frequency of His$^+$ colonies between *spt3* and *spt3 rnh1 rnh201* cells carrying phis3.210 (**Tables 4.2 and 4.3**). While there could be some repair in *trans* by the RNA derived from the *his3* allele on phis3.210 in *spt3 rnh1 rnh201* cells, we expect this to be minimal as shown in Keskin *et al*. 2014 [121]. Differently, there is a remarkable difference (more than a factor of 60,000) in the frequency of His$^+$ colonies between *spt3* and *spt3 rnh1 rnh201* cells carrying YCp50pK due to repair by RNA. Deletion of *RAD52* in *rnh1 rnh201* cells prevents repair by the *his3* copy on the plasmid and by cDNA, while, as previously shown [121], it reduces, but

not abolishes RNA repair either in the presence of phis3.210 or YCp50pK (**Table 4.2**).

Overall, these results demonstrate that absence of RNase H activity does not stimulate

DSB repair via DNA-DNA HR, while it strongly activates RNA-DNA HR, and HR

between DNA and cDNA, in which the cDNA is most likely an RNA-DNA hybrid.
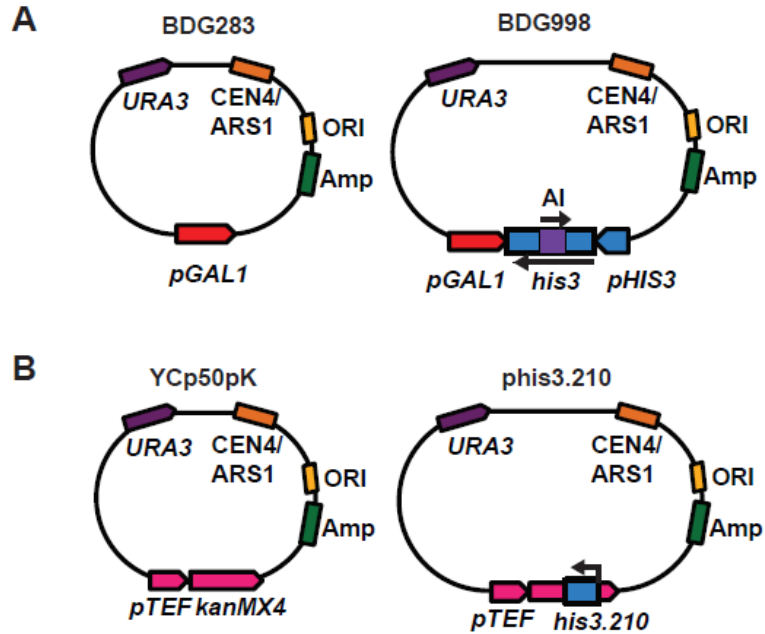
Figure 4.2 Scheme of the plasmids introduced in the *cis* system. A) BDG283 and

BDG998. GAL1 promoter, *pGAL1* (red); *his3* promoter and open-reading frame, *pHIS3*

and *his3* (blue); *AI*, artificial intron (purple). The arrows indicate the orientation of the *AI*

and that of the *his3* gene. Other parts of the plasmids are also shown. B) YCp50pK and

phis3.210. The kanMX4 gene with the *pTEF* promoter are in pink; 210-bp fragment of

HIS3 sequence, *his3*.210 (blue) is inserted in the kanMX4 gene. The orientation of the

his3 fragment is indicated by an arrow. Other parts of the plasmids are also shown.

Table 4.2 Transcript RNA-templated repair is the major mechanism for precise DSB repair in *spt3 rnh1 rnh201* cells in *cis* system.

Genotype of *cis* system	His$^+$ freq.		Survival
WT + BDG283	33	(20–45)	9%
WT + BDG998	4,130	(2,680–6,190)	10%
spt3 rnh1 rnh201 + BDG283	870	(706–960)	22%
spt3 rnh1 rnh201 + BDG998	890	(850–980)	27%

Frequencies of His$^+$ colonies per 10^7 viable cells for yeast strains of the *cis* system of the indicated genotypes and containing either the control empty vector BDG283 or vector BDG998, following 48 h of galactose treatment are shown as median and 95% CI (in parentheses). Percentage of cell survival after incubation on galactose is also shown. There were 9–12 repeats for each strain. The significance of comparisons between different strains of the system was calculated using the Mann-Whitney U-test and it is shown in **Table 4.3A**. **Figure 4.3** serves as graphical guide for all results presented in this table.

4.3.6 What's next?

Our recent findings raise a multitude of unanswered questions. We have shown that a transcript RNA can facilitate the repair of a DSB via a direct or indirect cDNA intermediate pathway. What are the players involved in this newly discovered mechanism of DNA repair? What factors mediate the increasing amount of repair in *cis* versus *trans* in *spt3 rnh1 rnh201* cells? Based on the localization of the transcript, nearby its DNA gene, the *cis* system is more prone to the generation of an RNA-DNA hybrid at the *his3*

locus. If so, can reverse transcriptase enter the nucleus and facilitate reverse transcription at the site of a DSB? Can other polymerases use RNA as a template in DSB repair in vivo? What is the real efficiency of transcript-templated DNA repair? Our assay is limited by the detection of a phenotype, His$^+$ cells, which originate only if the RNA template repairs the DSB after splicing of the artificial intron. If transcript RNA mediates DSB repair before splicing, there is no phenotype detected in our assay. Therefore, it is quite possible that we are underestimating the frequency of DSB repair by template transcript RNA. Does DSB repair by template transcript RNA occur in mammalian cells and in other cell types? We showed that transcript RNA-templated DNA repair occurs in dividing yeast cells. Can RNA template DSB repair in non-dividing cells? For example, highly transcribed genes in non-dividing cells, in which no sister chromatid is available, could be vulnerable; thus, these genes could be liable to RNA-templated DNA repair.
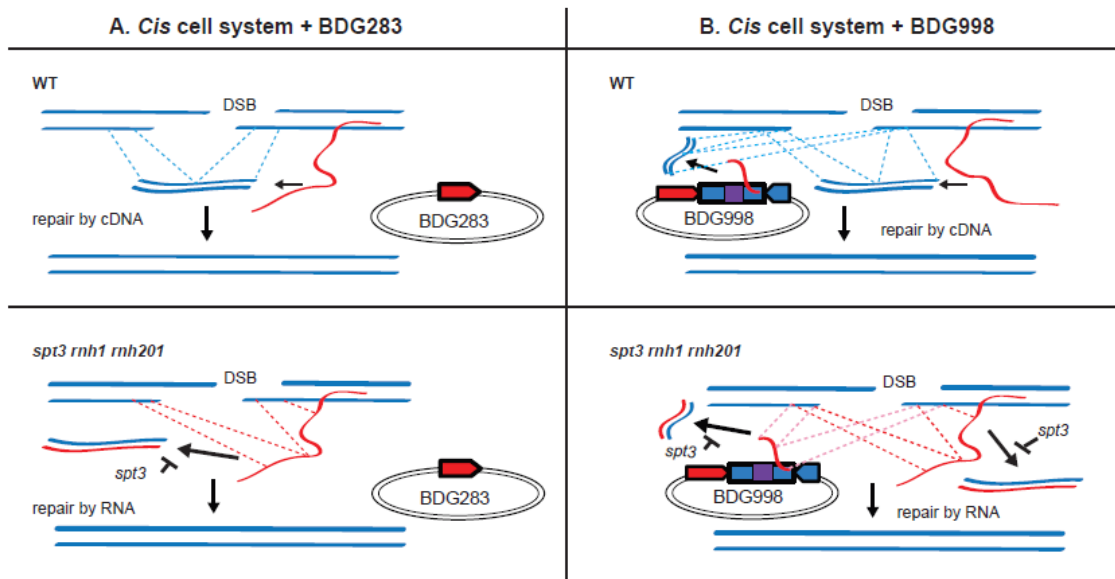
Figure 4.3 Templates for DSB repair in *his3* locus to generate a functional *HIS3* gene in a *trans-cis* system. This figure reflects the results of **Table 4.1**. Only repair mechanisms resulting in functional restoration of *HIS3* are shown. Repair may also proceed by canonical NHEJ or HR with sister chromatid but does not result in functional *HIS3*. Regions of homology to the DSB site in *his3* are shown as dashed lines. The *spt3-*null mutation results in inhibition of reverse transcription by Ty retroelements. Relevant genotypes are shown in the top left corner of each panel. Donor molecules that can serve as template for DSB repair are shown as solid blue lines for cDNA and dsDNA, red and blue lines for RNA-DNA hybrid, and red lines for transcript-RNA. A) Repair of a DSB in *cis* system in the presence of BDG283. B) Repair of a DSB in *cis* system in the presence of BDG998. DSB repair in *trans* templated by the spliced RNA from the transcription of *his3* on BDG998 is also possible in cells containing *rnh1 rnh201* mutations, although this is inefficient.

Our results of RNA repairing a DSB indirectly, via cDNA, shed light on the possibility of any RNA molecule being a target for reverse transcription by endogenous retrotransposon activity. If so, what factors mediate this reverse transcription? How abundant is the cDNA generation of endogenous RNA molecules? The Saccharomyces cerevisiae genome contains 5 classes of retroelements known as Tys, with Ty1 being the most abundant and well-studied. Is one class more prone to the generation of cDNA by endogenous RNA molecules? What can these factors tell us about other endogenous retroelements and retroviral infections? Retrotransposons are ubiquitous and plentiful in plant genomes, in some cases accounting for over 50% for the nuclear genome [159].

Mammalian genomes are no strangers to retroelements with »3 million transposable elements in the human genome and 90% of those being retrotransposons [160]. Given the copious amounts of retroelements found throughout various genomes and the relative abundant amounts of RNA in contrast to DNA, could RNA-templated DNA repair be playing a significant role in genome stability and modification?

Our work has provided fundamental preliminary data and resulted in the development of unique tools to study DNA repair via HR directly by RNA in the yeast model system. While inactivation of RNase H function allowed us to discover the capacity of cells to use transcript RNA in DSB repair, it is possible that RNA-DNA HR occurs also in RNase H wild-type cells. Mechanisms and functions of RNA-DNA HR are mostly unknown. Further studies are needed to illuminate the implications RNA-DNA HR may have on genome integrity.

Table 4.3 Effect of RNase H1 and H2-null mutations on DSB repair frequency by homologous cDNA, RNA-DNA hybrid, RNA and/or plasmid dsDNA.

Genotype of *cis* system	His$^+$ freq.		Survival
WT + YCp50pK	2,500	(2,300–2,830)	1.9%
WT + phis3.210	133,000	(104,000–151,000)	2%
rad52 + YCp50pK	<0.1	(0–0)	0.2%
rad52 + phis3.210	<0.1	(0–0)	0.16%
rnh1 rnh201 + YCp50pK	52,300	(47,700–63,300)	1.4%
rnh1 rnh201 + phis3.210	248,000	(226,000–309,000)	1.2%
rnh1 rnh201 rad52 + YCp50pK	930	(790–1,300)	0.07%
rnh1 rnh201 rad52 + phis3.210	1,000	(590–1,200)	0.08%
spt3 + YCp50pK	<0.1	(0–0)	7%
spt3 + phis3.210	134,000	(96,000–180,000)	4%
spt3 rnh1 rnh201 + YCp50pK	6,300	(5,900–7,200)	10%
spt3 rnh1 rnh201 + phis3.210	99,000	(92,000–119,000)	7%

Frequencies of His$^+$ colonies per 10^7 viable cells for yeast strains of the *cis* system of the indicated genotypes and containing the indicated plasmid, following 48 h of galactose treatment are shown as median and 95% CI (in parentheses). Percentage of cell survival after incubation on galactose is also shown. There were 6–12 repeats for each strain. The significance of comparisons between different strains of the system was calculated using the Mann-Whitney U-test and it is shown in **Table 4.3B**. **Figure 4.4** serves as graphical guide for all results presented in this table.
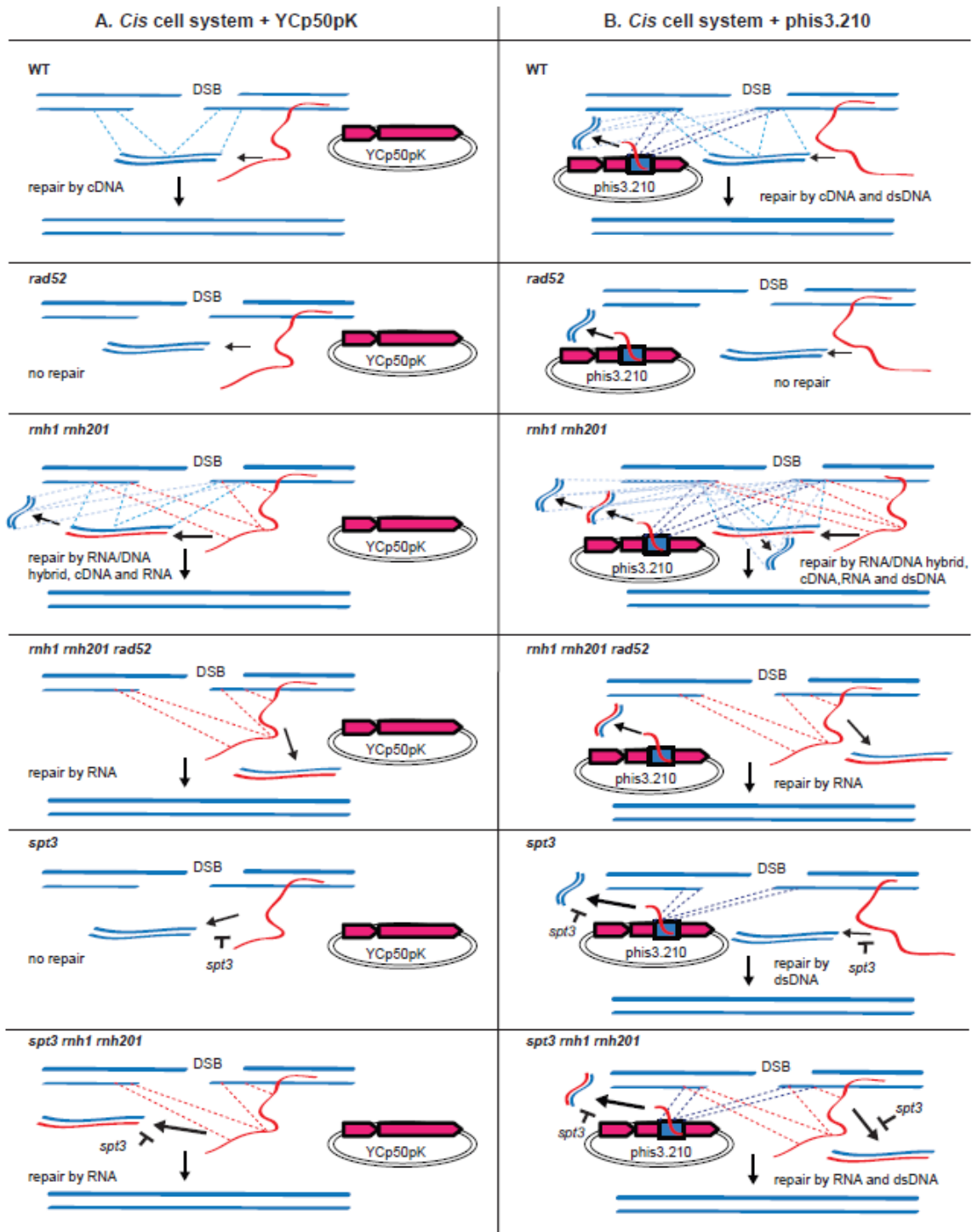
Figure 4.4 Templates for DSB repair in *his3* locus to generate a functional *HIS3*

gene in *cis* system. This figure reflects the results of **Table 4.2**. Only repair mechanisms

resulting in functional restoration of *HIS3* are shown. DSB repair in *his3* may also proceed by canonical NHEJ or HR with sister chromatid but does not result in functional *HIS3*. Regions of homology to the DSB site in *his3* are shown as dashed lines. The *spt3*-null mutation results in inhibition of reverse transcription by Ty retroelements. Relevant genotypes are shown in the top left corner of each panel. Donor molecules that can serve as template for DSB repair are shown as solid blue lines, for cDNA and dsDNA, red and blue lines for RNA-DNA hybrid, and red lines for transcript-RNA. A) Repair of a DSB in *cis* system in the presence of YCp50pK. B) Repair of a DSB in *cis* system in the presence of phis3.210, which contains 210 bp of *HIS3* (blue rectangle). DSB repair in *trans* templated by the RNA from the transcription of *his3* on phis3.210 is also possible in cells containing *rnh1 rnh201* mutations. Due to inefficient DSB repair by RNA in *trans*, we did not show the dashed lines for this template in the panels.

4.4 Acknowledgments

CHAPTER 5

RAD52-INVERSE STRAND EXCHANGE DRIVES RNA-TEMPLATED DNA DOUBLE-STRAND BREAK REPAIR

The study in Chapter 5 consists of the work in revision

Olga M. Mazina[1‡], Havva Keskin[2‡], Kritika Hanamshet[1], Francesca Storici[2*], Alexander V. Mazin[1*§].

[1]Department of Biochemistry and Molecular Biology, Drexel University College of Medicine, Philadelphia, PA 19102, USA.

[2]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA

‡ Equal contribution

5.1 Summary

Recent data show that RNA can serve as a template for DNA double-strand break repair in yeast cells. Rad52, a member of the homologous recombination pathway, appeared to be an important player in this process. However, the exact mechanism of how Rad52 contributes to RNA-dependent DSB repair remained unknown. Here, we report a novel activity of yeast and human Rad52: inverse strand exchange, in which Rad52 forms a complex with dsDNA and promotes strand exchange with homologous ssRNA or ssDNA. We show that in eukaryotes, inverse strand exchange between homologous dsDNA and RNA is a unique activity of Rad52; neither Rad51 recombinase, nor the yeast Rad52 paralog Rad59 has this activity. In accord with our *in vitro* results, our experiments in budding yeast provide evidence that Rad52-inverse strand exchange plays an important role in RNA-templated DSB repair *in vivo*.

5.2 Introduction

Homologous recombination (HR) is a high fidelity process that uses homologous DNA sequences as a template to repair damaged DNA [161-163]. In eukaryotes, HR is carried out by the Rad52 epistasis group of proteins [163]. In this group, Rad51 plays a key role by promoting a search for homologous dsDNA-template and forming DNA joint molecules that provide both the template and the primer for DNA polymerase during repair of DNA double-strand breaks (DSB) [164]. However, we recently demonstrated that transcript RNA can serve as template for DSB repair via HR in yeast cells either indirectly, if converted into cDNA, or directly [121]. Direct RNA-templated DSB repair is efficient in the absence of ribonuclease (RNase H) function, and in *cis*, that is when the RNA is used as template to repair a break occurring in its own DNA gene [121]. Currently, little is known about the enzymatic machinery that executes RNA-templated DSB repair. Our results from budding yeast implicated Rad52, but not Rad51, in this RNA-directed DSB repair mechanism [121]. The role of Rad52 in RNA-dependent DSB repair is also consistent with data from human cells, which show an RNA-dependent localization of Rad52 at sites of DSBs [165]. However, the exact mechanism of how Rad52 contributes to RNA-dependent DSB repair remains to be elucidated.

It is known that recombinases of the Rad51/RecA family form a nucleoprotein filament on ssDNA and promote DNA strand exchange with homologous dsDNA. However, in addition to this canonical or "forward" reaction, *E. coli* RecA was shown to form a nucleoprotein filament on dsDNA. The filament can promote strand exchange with either homologous ssDNA or ssRNA [33, 34]. This unconventional pairing process was called the "inverse" DNA strand exchange reaction [33].

Previously, it was shown that Rad52, an important member of the HR pathway [166], promotes annealing either of two complementary ssDNA molecules [167-169] or of ssDNA with complementary ssRNA [121]. Here, we tested whether Rad52 also carries "inverse strand exchange" activity between homologous dsDNA and ssRNA, which could also account for the role of Rad52 in RNA-dependent DNA repair identified in our genetic experiments. Our current results demonstrate that both human and yeast Rad52 efficiently promotes inverse strand exchange between dsDNA and homologous ssRNA or ssDNA. We show that in eukaryotes, inverse RNA strand exchange is a unique activity of Rad52; neither Rad51 recombinase, nor the yeast Rad52 paralog Rad59 carries this activity. Our experiments in yeast *Saccharomyces cerevisiae* cells support the biological significance of inverse RNA strand exchange. Taken together, our biochemical and genetic data indicate that inverse RNA strand exchange promoted by Rad52 may play a central role in RNA-dependent DSB repair.

5.3 Experimental Procedures

5.3.1 Proteins, DNA and RNA.

Human Rad52, hRad52$_{1-209}$ NTD, hRad51, and RPA proteins were purified as described [170-173]. The deoxyribonucleotides (**Table C.1**) were purchased from IDT Inc. and further purified by electrophoresis in polyacrylamide gels containing 50% urea as described[174]. Oligoribonucleotides of an HPLC-purified grade were purchased from IDT Inc. Duplex or tailed dsDNA substrates were prepared by annealing of equimolar (molecules) amounts of indicated complementary oligonucleotides, as described [174]. When indicated, oligonucleotides were 5'-end labeled with ^{32}P using T4 polynucleotide

kinase (New England Biolabs). All DNA and RNA concentrations are expressed in moles of molecules.

5.3.2 Inverse DNA and RNA strand exchange promoted by hRad52.

Nucleoprotein complexes were assembled by incubating hRad52 (900 nM) with ^{32}P-labeled dsDNA (no. 1/ no. 2; 68.6 nM) or 3'-tailed DNA (no. 1/ no. 117, or indicated otherwise; 68.6 nM) in buffer A containing 25 mM Tris-Acetate, pH 7.5, 100 µg/ml BSA, 2 mM magnesium acetate, and 2 mM DTT for 15 min at 37 °C. The reactions were initiated by addition of ssDNA (no. 2; 205.8 nM) or ssRNA (no. 2R; 205.8 nM). Variations to these conditions are indicated in figure legend 1a and 1b). Aliquots (10 µl) were withdrawn at indicated time points and DNA or RNA samples were deproteinized by incubation in 1% SDS, 1.6 mg/ml proteinase K, 6% glycerol and 0.01% bromophenol blue for 15 min at 37 °C. Samples were analyzed by electrophoresis in 10% polyacrylamide gels (acrylamide:bis-acrylamide, 17:1) in 1x TBE buffer (89 mM Tris, 89 mM boric acid and 1 mM EDTA, pH 8.3); the gels were processed as described [174] and the reaction yield was determined using a Storm 840 Phosphor Imager (GE Healthcare).

When human or yeast RPA (1 µM) were used, they were pre-incubated with ssDNA (no. 2; 411.6 nM) or ssRNA (no. 2R; 411.6 nM) in buffer A (40 µl of reaction mixture) for 15 min at 37 °C. Separate reaction mixture (40 µl) containing hRad52 (1.8 µM) and labeled 3'-tailed DNA (137.2 nM) in buffer A was incubated for 15 min at 37 °C. Inverse strand exchange reaction (80 µl) were initiated by addition of mixtures containing RPA and

ssDNA or RNA to the hRad52 nucleoprotein complexes. The final concentration of RPA was 500 nM, which corresponding to a stoichiometry of 1 RPA trimer per 30 nt of ssRNA or ssDNA, including ssDNA tail of 3'-tailed DNA.

5.3.3 Inverse DNA and RNA strand exchange promoted by hRad52$_{1-209}$ NTD.

The conditions were the same as for hRad52, except of a 10-fold molar excess of ssRNA (no. 2R; 63-mer, 686 nM) or ssDNA (no. 2; 63-mer, 686 nM) were used and the concentration of hRad52**$_{1-209}$** was 1 µM.

5.3.4 Inverse DNA and RNA strand exchange promoted by yRad52.

The conditions for yRad52 inverse DNA and RNA strand-exchange were the same as for hRad52, except of a 10-fold molar excess of ssRNA (no. 2R; 63-mer, 686 nM) was used. The concentration of yeast RPA in the reaction with ssRNA was 1.5 µM to maintain the stoichiometry of 1 yeast RPA trimer per 30 nt of ssRNA.

5.3.5 Inverse DNA and RNA strand exchange promoted by yRad59.

To assemble nucleoprotein complexes yRad59 (3.5 µM) was incubated with ^{32}P-labeled 3'-tailed DNA (no. 1/ no. 117; 68.6 nM) in buffer A for 15 min at 37 °C. The reactions were initiated by addition of ssDNA (no. 2; 63-mer, 686 nM) or ssRNA (no. 2R; 63-mer 686 nM) and carried out for 1 h. The reaction products were deproteinized and analyzed, as described for inverse DNA or RNA strand exchange promoted by hRad52.

5.3.6 Inverse DNA and RNA strand promoted by hRad51.

To form nucleoprotein filament hRad51 (2.15 µM) was incubated with labeled 3'-tailed

DNA (no. 1/ no. 117; 68. 6 nM) in buffer containing 25 mM Tris-acetate, pH 7.5, 100

µg/ml BSA, 3 mM magnesium acetate, 2 mM ATP, and 2 mM DTT for 15 min at 37 °C.

Afterward, the concentration of magnesium acetate was increased to 10 mM. Inverse

DNA strand exchange reactions were initiated by addition of 7-fold molar excess of a 63-

mer ssDNA (no. 2; 480.2 nM) or ssRNA (no. 2R, 63-mer; 480.2 nM). The products were

deproteinized and analyzed as described for inverse strand exchange promoted by

hRad52.

5.3.7 P1 nuclease assay.

hRad52 protein (900 nM) was incubated with 13.4 ng of ^{32}P-labeled 63-mer ssDNA (no.

2; 68.6 nM) or with 26.8 ng of a 63-mer dsDNA (no. 1/ no. 2; 68.6 nM) in 9 µl of buffer

A for 15 min at 37 °C, followed by the addition of 0.4 units of P1 nuclease (USBiological

Life Science) in 1 µl. Reactions were carried out for 10 min at 37 °C, then quenched by

addition of SDS to 1%, proteinase K to 1.6 mg/ml, glycerol to 6% and bromophenol blue

to 0.01% followed by 15 min incubation at 37 °C. The DNA products were analyzed by

electrophoresis in 10% polyacrylamide gels (acrylamide:bis-acrylamide, 17:1) in 1x TBE

buffer (89 mM Tris, 89 mM boric acid and 1 mM EDTA, pH 8.3).

5.3.8 Yeast strains, plasmids and genetics methods.

All the strains used in this study are FRO-767 [121] derivatives and are shown in Table

S2. Plasmids YEP-NAT, YEP-NAT-ScRAD52-327 and YEP-NAT-hRAD52-209 are

episomal vectors containing the *URA3* and the nourseothricin (*NAT*) resistance marker

genes, and the *GAL1* promoter, and are described in [96]. YEP-NAT-ScRAD52 was

constructed like YEP-NAT-ScRAD52-327 but using a PCR product with the full length

of the yeast *RAD52* gene. The sequence of the YEP-NAT-ScRAD52 vector was verified

by sequencing. YEP-NAT is the control empty vector, YEP-NAT-ScRAD52-327

contains the first 327 codons of yeast *RAD52* gene, lacking the C-terminal region for

Rad51 binding, expressed under the *GAL1* promoter. YEP-NAT-hRAD52-209 contains

the first 209 residues from the cDNA of human Rad52 isoform # (NM_002879)

expressed under the *GAL1* promoter. YEP-NAT-ScRAD52 contains full length yeast

RAD52 gene expressed under the *GAL1* promoter. Plasmid transformation was done as

described [96]. YEp-NAT, YEP-NAT-ScRAD52, YEP-NAT-ScRAD52-327 and YEP-

NAT-hRAD52-209 were transformed in strains CM-95, 96 (WT), CM-100, 101 (*rnh1Δ*

rnh201Δ) and CM-107, 108 (*rnh1Δ rnh201Δ spt3Δ*) which were generated by

introducing the yeast 2-micron plasmid following the procedure described in[175], to

stabilize the YEp vectors used. Yeast genetic methods and molecular biology analyses

were done as described [66, 75, 121]. All primers used for strain and plasmid

constructions, PCR verifications and sequence analyses are available upon request.

Samples for sequencing were submitted to Eurofins MWG Operon.

5.3.9 Assay to calculate the frequency of DSB repair by RNA.

To determine the frequency of His$^+$ colonies in the strains of the *cis* system following

induction of DSB, we conducted a fluctuation experiment as previously described [121].

Briefly, yeast cells were inoculated in 50 ml lactic acid containing media (YPLac) and

incubated in a shaker for 24h at 30 °C. Cells were then counted and 10^7, or in some cases,

10^8 cells were plated on galactose containing medium (YPGal) to turn on transcription of the *his3* antisense on chromosome III and expression of the homothallic switching endonuclease. In addition, 10^4 cells were plated on YPGal medium to determine cell survival on galactose. Cells were incubated for 48 hours at 30 °C and then replica-plated on synthetic complete medium lacking histidine (SC-His⁻) and grown for 3 days at 30 °C. The frequency of His⁺ colonies following DSB induction was calculated by dividing the number of His⁺ colonies obtained on SC-His⁻ medium by the number of colonies obtained on YPGal medium. The survival was calculated by dividing the number of colonies obtained on YPGal medium by the number of cells plated on the same medium. For experiments using plasmids YEP-NAT, YEP-NAT-ScRAD52, YEP-NAT-ScRAD52-327 and YEP-NAT-hRAD52-209, 10^7 or 10^8 cells were plated on medium lacking uracil and containing galactose (Ura⁻Gal) and 10^3 or 10^4 cells were plated on Ura⁻Gal medium to determine the cell survival. After 48 hours of incubation at 30 °C, cells were replica-plated on SC-His⁻ medium.

For experiments without induction of the DSB, cells were grown on 50 ml YPLac overnight at 30 °C shaker. Next day, cells were counted and 10^8 cells were plated on glucose containing medium (YPD) and incubated for 24 hours at 30 °C. After incubation, cells were replica-plated on SC-His⁻ medium. In addition, 10^3 cells were also plated on YPD for cell survival. Results obtained in glucose are shown in **Table C.4**.

5.3.10 Assay of DSB repair by oligonucleotide transformation.

Transformation by oligonucleotide HIS3.F (80mer, 5'-

ACCAATGCACTCAACGATTAGCGACCAGCCGGAATGCTTGGCCAGAGCATGT

ATCATATGGTCCAGAAACCCTATACCTG) (1nmol) was performed as described

[66]. Induction of the homothallic switching endonuclease DSB was done by incubating

cells in 2% galactose medium for 3 h.

5.3.11 Data presentation and statistics.

For conducting statistical analysis, GraphPad Prisim 5 software was used. *In vitro*

experiments were repeated at least three times; standard deviations (SD) are presented on

the graphs. Results of genetic experiments in yeast cells are expressed as median and

95% confidence interval is shown in parenthesis, or alternatively the range when number

of repeated experiments was less than 6. The nonparametric two-tailed Mann-Whitney-U

test [86] was used to calculate differences between His$^+$ frequencies and P values that are

presented in **Table C.5**.

5.4 Results

5.4.1 Rad52 promotes inverse DNA strand exchange

First, we tested whether human Rad52 (hRad52) can promote inverse DNA strand

exchange between homologous dsDNA and ssDNA. hRad52 nucleoprotein complex was

formed with 3'-tailed dsDNA (no. 1, 63-mer/ no. 117, 94-mer) (**Table C.1**), in which

oligo no.1 was ^{32}P-labeled, and then inverse DNA strand exchange was initiated by

addition of homologous ssDNA (no.2, 63-mer) (**Figure 5.1A**). We found that hRad52

promotes inverse strand exchange with remarkably high efficiency; the initial rate of

inverse reaction was approximately 10-fold higher than that of forward DNA strand

exchange promoted by hRad52 with the same DNA substrates (**Figure 5.1B,C**). The inverse reaction required both DNA sequence homology (**Figure C.1A**) and hRad52 protein (**Figure 5.1C**). Forward and inverse DNA strand exchange reactions promoted by hRad52 show different requirements for Mg^{2+} concentrations. The inverse reaction occurs across a much broader (1-20 mM) range of Mg^{2+} concentrations (**Figure C.1B**) than the forward reaction (0.1-1 mM) [176, 177] (Mazina, Bugreev and Mazin, unpublished observations). hRad52 was significantly more efficient in promoting inverse DNA strand exchange than hRad51 under conditions that were optimal for both proteins (**Figure 5.1D**). The initial rate for hRad52-promoted reaction is about 6-fold higher than that for hRad51. In contrast to RecA that requires a large (10-fold) excess of ssDNA for inverse DNA strand exchange [33], the hRad52 reaction was efficient at an equimolar ssDNA:tailed dsDNA ratio and reached the maximal rate at a 3-fold excess of ssDNA (**Figure C.1C**). Inverse DNA strand exchange does not involve melting of dsDNA, as no ssDNA intermediate sensitive to P1 nuclease was detected (**Figure C.1D**). Furthermore, we found that yeast Rad52 (yRad52) also promotes inverse DNA strand exchange, indicating an evolutionary conservation of this activity (**Figure 5.1E**). Taken together, our current results demonstrate that human and yeast Rad52 possess inverse DNA strand exchange activity. This activity appears to be distinct and stronger than the forward DNA strand exchange activity of these Rad52 orthologs.
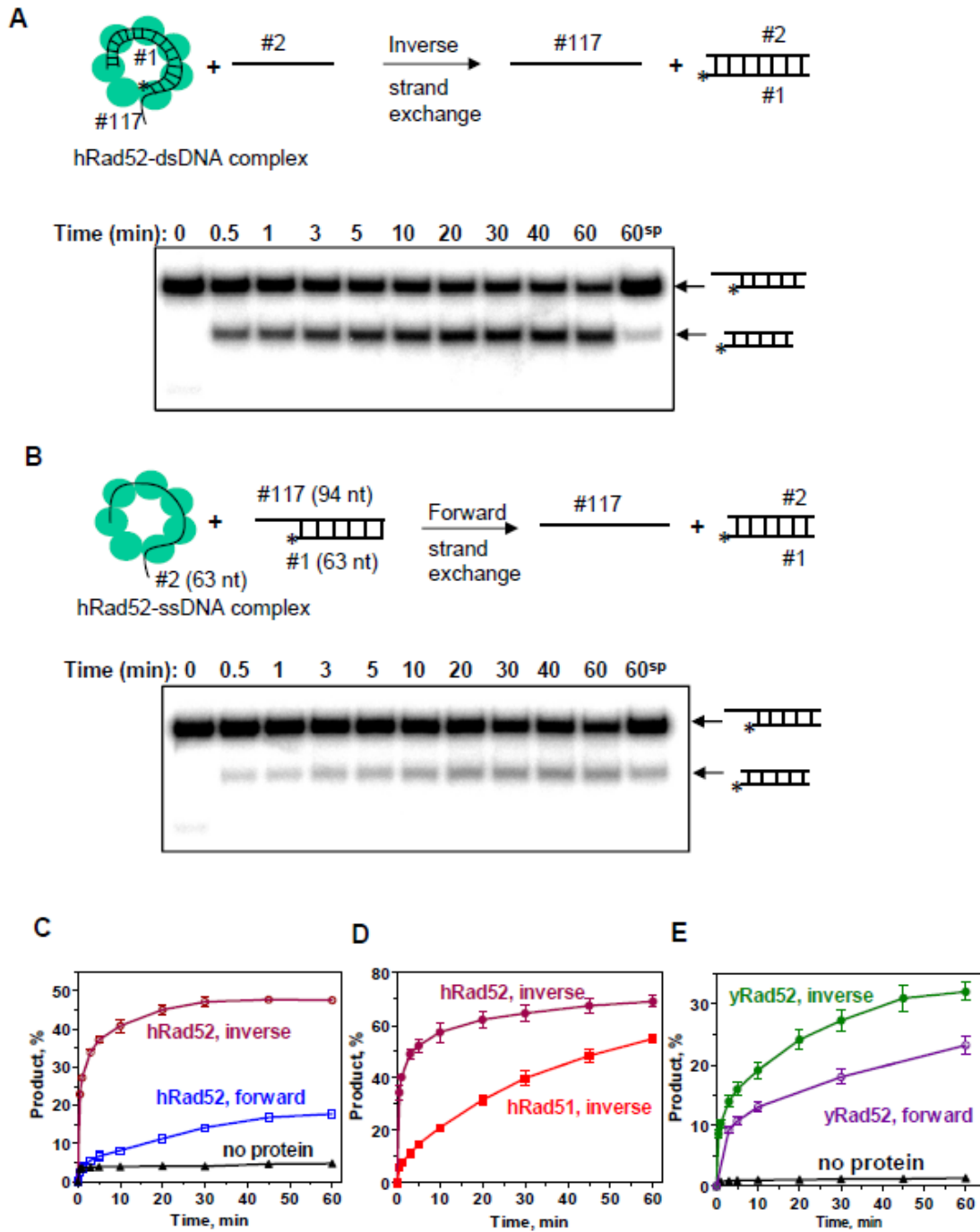
Figure 5.1 Rad52 promotes inverse DNA strand exchange with high efficiency. A,

Top: the scheme of inverse strand exchange. Asterisk represents ^{32}P-label.

Oligonucleotide sequences are shown in Table S1. hRad52 (900 nM) was incubated with

the 3'-tailed DNA (no. 1/ no. 117; 68.6 nM) followed by addition of ssDNA (no. 2; 68.6

nM). Bottom: analysis of the reaction products by electrophoresis in a polyacrylamide gel. **B**, Top: the scheme of the forward DNA strand exchange. hRad52 (900 nM) was incubated with ssDNA (no. 2; 68.6 nM) for 15 min at 37 °C, then the reaction was initiated by adding the 3'-tailed DNA (no. 1/ no. 117; 68.6 nM). Bottom: analysis of the reaction products by electrophoresis in a polyacrylamide gel. **C**, Data from A and B were plotted as a graph. In "no protein" control hRad52 was substituted by storage buffer. **D**, hRad52 promotes inverse DNA strand exchange more efficiently than hRad51. The reaction conditions were as in panel A, except that a three-fold excess of ssDNA (205.8 nM) and seven-fold excess of ssDNA (480.2 nM) were used in reactions with Rad52 and Rad51, respectively. **E**, Yeast Rad52 promotes inverse DNA strand exchange. The DNA substrates and conditions for forward and inverse reactions were the same as for hRad52 in panels **B** and **D**, respectively. In "no protein" reaction, yRad52 was substituted with storage buffer. The experiments were repeated at least three times, error bars indicate standard deviation (SD). (**by A. V. Mazin**)

5.4.2 Inverse RNA strand exchange is a unique activity of Rad52 in eukaryotes

Our recent data indicate that in yeast cells RNA can serve as a template for DSB repair via HR and that Rad52 plays a significant role in this process. In a *rad52*-null mutant the frequency of DSB repair by RNA was reduced by a factor of ten [121]. These data prompted us to test whether human and yeast Rad52 can carry out inverse strand exchange between tailed dsDNA that mimics processed DNA ends and homologous ssRNA. We found that hRad52 promotes inverse strand exchange between 3'-tailed dsDNA (no. 1, 63-mer/ no. 117, 94-mer) and ssRNA (no. 2R, 63-mer) (**Figure 5.2A**).

Under standard conditions (a 3-fold molar excess of ssRNA or ssDNA), the inverse reaction with ssRNA showed a 4-5-fold lower initial rate than with ssDNA, but the extents of the reactions were similar (53% and 69% for ssRNA and ssDNA, respectively) (**Figure 5.2B**). A 10-fold molar excess of ssRNA further stimulated inverse strand exchange (**Figure C.2A**). The reaction required homologous RNA (**Figure C.2B**). We found that inverse RNA strand exchange activity is evolutionarily conserved, as yeast Rad52 can also promote exchange between dsDNA and ssRNA, albeit with lower efficiency (**Figure 5.2C**).

In Rad52, the N-terminal domain (NTD), spanning approximately half of the protein, is responsible for its ssDNA annealing, DNA strand exchange and protein multimerization. The Rad52 C-terminal domain carries the nuclear localization site and regions involved in interaction with Rad51 and RPA [166, 178-180]. We found that the hRad52$_{1-209}$ NTD is sufficient to promote inverse RNA strand exchange efficiently (**Figure 5.3A, B**). We then examined the ability of hRad51 recombinase to promote inverse strand exchange between tailed dsDNA and ssRNA. We found no significant activity under tested conditions either in the presence of a 7-fold or even 100-fold excess of ssRNA (**Figure C.3A**). Under the same conditions, hRad51 was active in promoting inverse strand exchange with ssDNA (a 7-fold excess) (**Figure C.3A**). Similarly, yeast Rad51 (yRad51) promotes inverse strand exchange with ssDNA, but is incapable of using ssRNA in this reaction (**Figure C.3B-D**). We also tested inverse strand exchange activity of yeast Rad59 (yRad59), which shares homology with the Rad52 NTD [181]. We found that yRad59 promotes inverse strand exchange with ssDNA, but not with ssRNA (**Figure**

5.3C,D), even though it promotes both ssDNA-ssDNA and ssDNA-RNA annealing

(**Figure 5.3E**). Thus, we find that Rad52 is unique among eukaryotic HR proteins in

promoting inverse strand exchange between dsDNA and ssRNA.
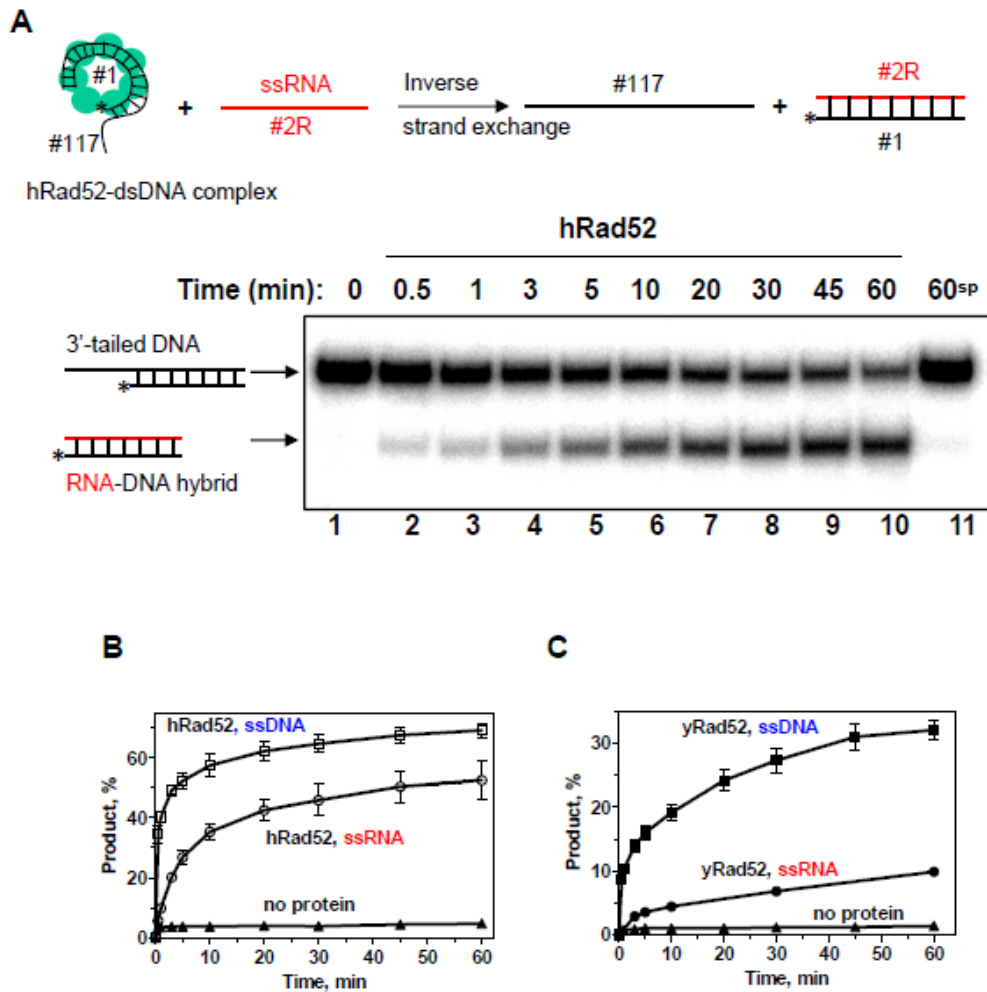
Figure 5.2 Rad52 promotes inverse strand exchange between 3'-tailed dsDNA and homologous ssRNA. A, The reaction with hRad52 was conducted as in Fig 1D, except that ssDNA was replaced with ssRNA (no. 2R, 205.8 nM). The reaction products were analyzed by electrophoresis in a polyacrylamide gel. **B**, Data from panel A are plotted as a graph. The DNA inverse exchange graph from Fig. 1D is shown for comparison. **C**,

Yeast Rad52 promotes inverse strand exchange with ssDNA or ssRNA. The reaction

conditions were as in panel A, except that a 10-fold excess of ssRNA (no. 2R, 686 nM)

was used. The experiments were repeated at least three times, error bars indicate SD. (**by**

A. V. Mazin)
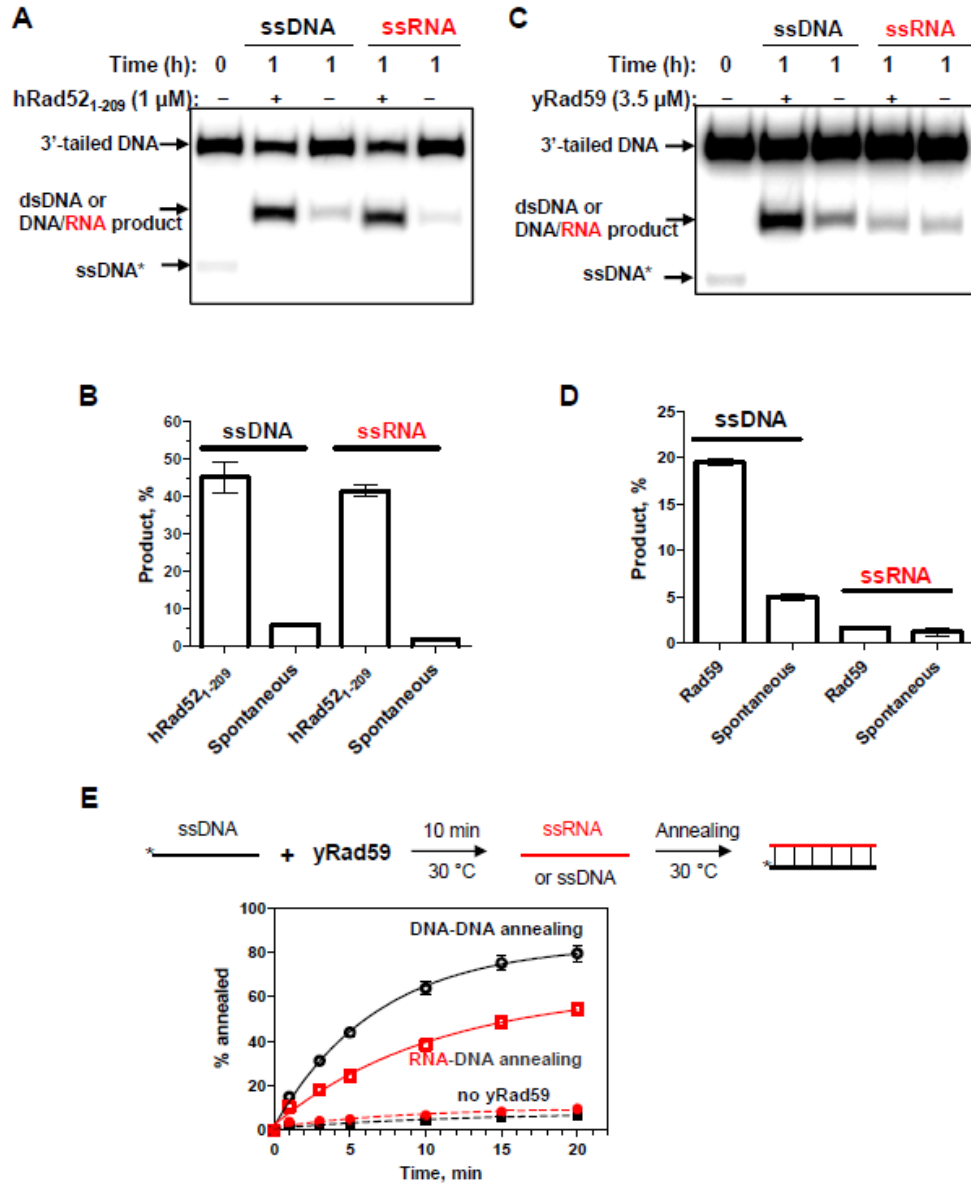
Figure 5.3 Different specificity in inverse strand exchange promoted by hRad52$_{1-209}$

NTD and yRad59. A, The reactions were carried out in the presence of hRad52$_{1-209}$ (1

µM), 3'-tailed DNA (no. 117/ no. 1; 68.6 nM) and ssRNA (no. 2R, 686 nM) or ssDNA (no. 2, 686 nM) for 1 h at 37 °C; the products were analyzed by electrophoresis in polyacrylamide gels. **B**, Graphical representation of the data from pane a. **C**, Rad59 promotes inverse strand exchange with ssDNA, but not with RNA. To form nucleoprotein complexes, Rad59 (3.5 µM) was incubated with ^{32}P-labeled 3'-tailed DNA (68.6 nM) for 15 min at 37 °C. The reactions were initiated by addition of free ssDNA (no. 2, 63-mer, 686 nM) or ssRNA (no. 2R, 63-mer 686 nM) and carried out for 1 h; the products were analyzed by electrophoresis in a polyacrylamide gel. **D**, The data from B are shown as a graph. **E**, Rad59 promotes annealing between ssDNA and ssRNA. Top: Experimental scheme. Asterisk represents ^{32}P-label. Rad59 (125 nM) was incubated with a 48-mer ^{32}P-labelled ssDNA (no. 65, 5 nM) for 10 min at 30°C. To initiate annealing reactions complementary 48-mer ssDNA (no. 64, 5 nM) or ssRNA (no. 64R, 5 nM) were added. In controls, protein storage buffer was added instead of Rad59. The products of annealing reactions were analyzed by electrophoresis in polyacrylamide gels. The experiments were repeated at least three times, error bars indicate SD. (**by A. V. Mazin**)

5.4.3 RPA stimulates inverse RNA strand exchange promoted by Rad52

In vivo, RPA, a ubiquitous ssDNA binding protein [182], plays an essential role in DSB repair and physically interacts with Rad52 [183]. Therefore, we tested the effect of RPA on inverse DNA strand exchange promoted by hRad52 between tailed dsDNA and homologous ssDNA or ssRNA (**Figure 5.4A**). We found that RPA inhibited the initial rate of inverse strand exchange with ssDNA by approximately two-fold (**Figure 5.4B**). In contrast, under the same conditions RPA stimulated the rate of inverse strand exchange

with ssRNA also by approximately two-fold (**Figure 5.4C**). The stimulation appeared to be species-specific; hRad52 was stimulated only by human RPA, whereas yeast RPA inhibited the reaction promoted by hRad52 (**Figure 5.4D**), but stimulated the yRad52-promoted reaction (**Figure 5.4E**). These data indicate that physical interaction of RPA with Rad52 is important for stimulation of inverse strand exchange between dsDNA and ssRNA, rather than by destabilization of DNA duplex. This conclusion was further strengthened by the observation that inverse RNA strand exchange promoted by the hRad52$_{1-209}$ NTD, which lacks the RPA binding region, was not stimulated by human RPA (**Figure 5.4F**). It is possible that RPA stimulates inverse RNA strand exchange by inducing a favorable conformation in Rad52. A partial inhibition of inverse DNA strand exchange by RPA could be due to formation of stable RPA-ssDNA complexes that hinder ssDNA binding to Rad52-dsDNA complexes. In addition, yRPA may inhibit inverse RNA strand exchange promoted by hRad52 by competing for binding to tailed dsDNA.
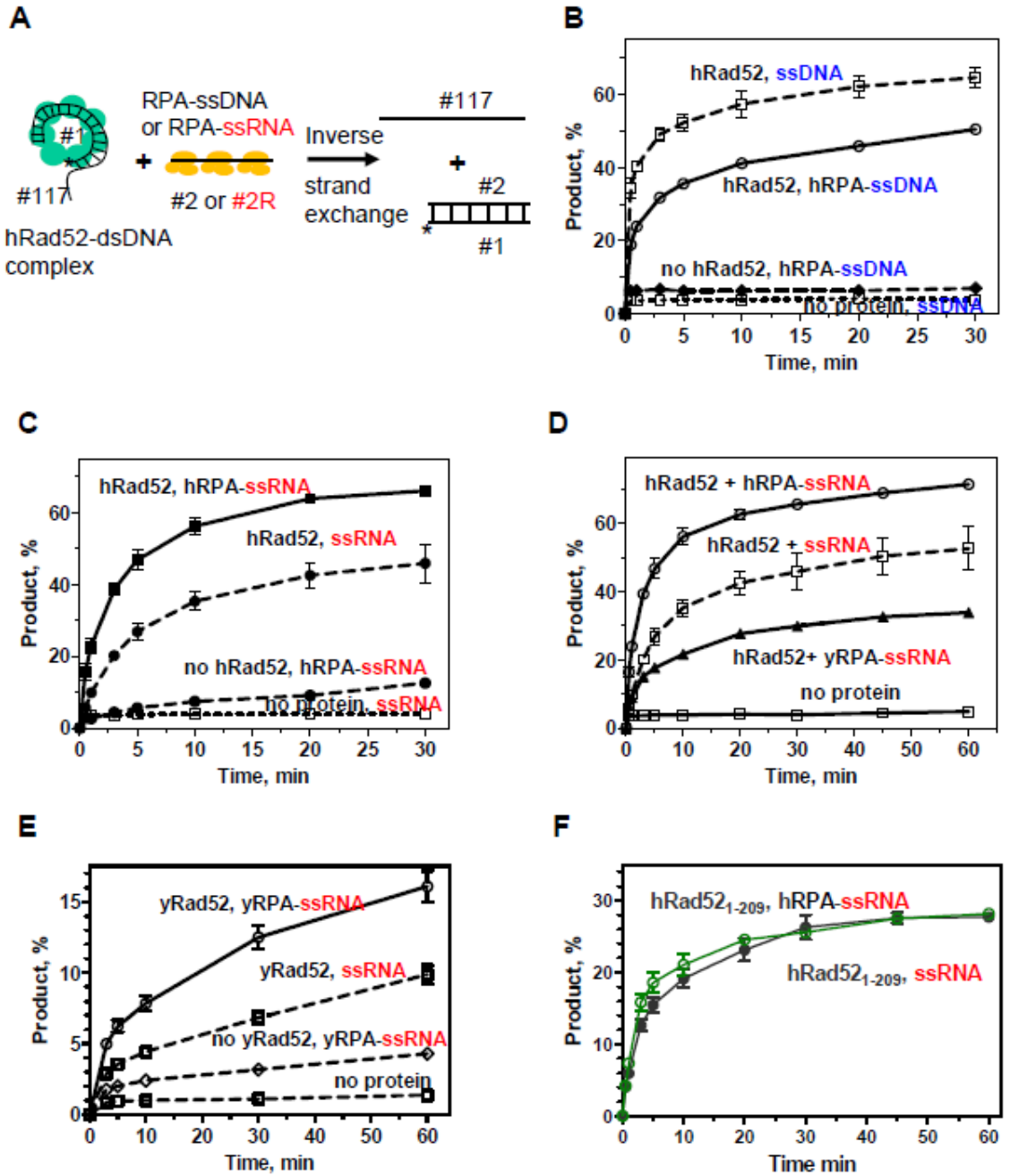
Figure 5.4 RPA stimulates Rad52-promoted inverse RNA strand exchange in a species-specific manner. A, Experimental scheme. Asterisk represents ^{32}P-label. **B**, Human RPA inhibits inverse DNA strand exchange, but **C**, stimulates inverse RNA strand exchange promoted by hRad52. **D**, Yeast RPA does not stimulate inverse RNA strand exchange promoted by hRad52. **E**, Yeast RPA stimulates inverse RNA strand

exchange promoted by yRad52. **F**, Human RPA does not stimulate inverse RNA strand exchange promoted by hRad52$_{10\text{-}209}$ NTD. The reactions were initiated by adding ssDNA (no. 2; 205.8 nM) or ssRNA (no. 2R; 205.8 nM) that were pre-incubated with RPA (500 nM) to the hRad52-tailed dsDNA complexes. In "no Rad52" or "no protein" controls, hRad52 and RPA were substituted with their storage buffer. The experiments were repeated at least three times, error bars indicate SD. (**by A. V. Mazin**)

5.4.4 Rad52 promotes inverse RNA strand exchange with blunt-ended dsDNA

Canonical DSB repair mechanisms by HR require extensive processing of DNA ends by exonucleases. Here, we wanted to test the effect of dsDNA end resection on the hRad52-promoted inverse strand exchange with ssRNA or ssDNA. We found that hRad52 is capable of promoting the inverse reaction between blunt-end duplex DNA (no. 1, 63-mer/ no. 2, 63-mer) and homologous ssRNA (no. 2R, 63-mer) (**Figure 5.5A,B**). The rate and the extent of the reaction were significantly reduced compared with the reaction utilizing a 31-nt tailed dsDNA (**Figure 5.5B**). However, addition of hRPA greatly stimulated reaction with blunt-ended DNA nearly to the level observed for tailed dsDNA (**Figure 5.5C**). Also, hRad52 promoted inverse strand exchange between blunt-end dsDNA and ssDNA (no. 2, 63-mer) (**Figure 5.5D**). However, no stimulation by RPA was observed for this reaction (unpublished observation). Utilization of blunt-ended dsDNA by Rad52 in inverse strand exchange may have important biological implications obviating the need for dsDNA end resection during DSB repair *in vivo*.
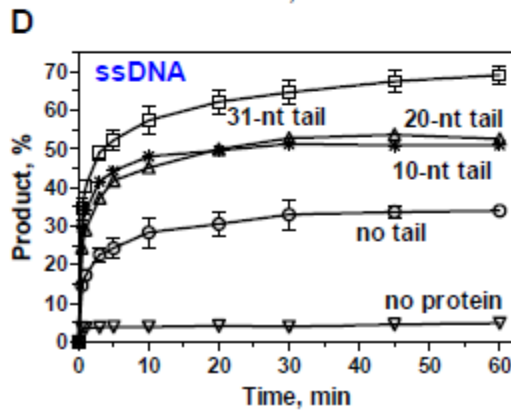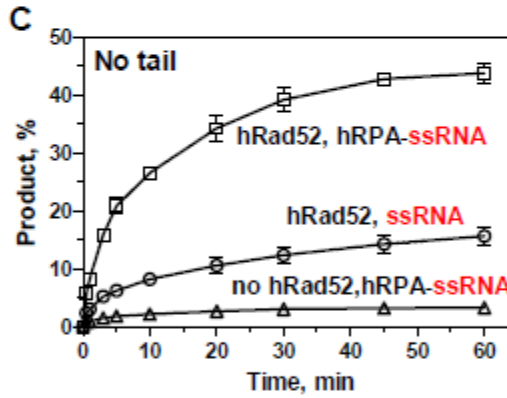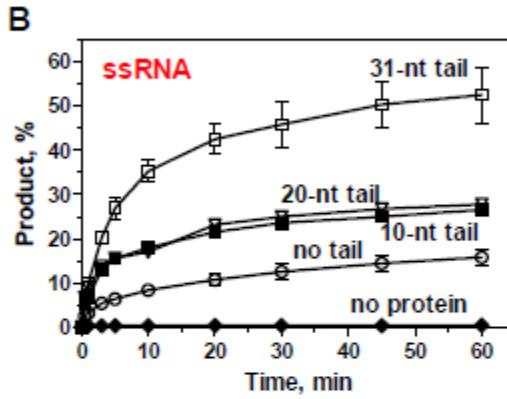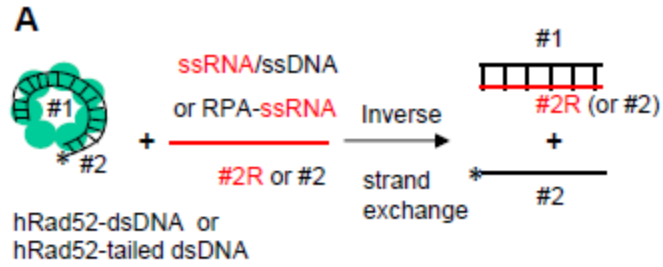
Figure 5.5 Rad52 promotes inverse RNA or DNA strand exchange with blunt-ended dsDNA. A, Experimental scheme. Asterisk represents ^{32}P-label. **B**, The kinetics of

inverse DNA strand exchange promoted by hRad52 (900nM) between labeled 63 bp

dsDNA substrates (68. 6 nM) containing either no ssDNA tails (no. 1/no. 2) or 3'-ssDNA

tails of different length: 10-nt tail (no. 1/ no. 518), 20-nt tail (no. 1/ no. 519), and 31-nt

tail DNA (no. 1/ no. 117). The reactions were initiated by adding ssRNA (no. 2R, 205.8

nM). **C**, RPA (432.2 nM) stimulates inverse strand exchange promoted by hRad52

between blunt-ended dsDNA (no. 1/ no. 2; 68. 6 nM) and ssRNA (no. 2R; 205.8 nM). **D,**

The kinetics of inverse DNA strand exchange promoted by hRad52 (900nM) between

dsDNA (68.6 nM) containing either no 3'-ssDNA tails (no. 1/no. 2), or 10-nt tail (no. 1/

no. 518), 20-nt tail (no. 1/ no. 519), and 31-nt tail DNA (no. 1/ no. 117) and ssDNA

(no.2, 205.8 nM). The experiments were repeated at least three times, error bars indicate

SD. (**by A. V. Mazin**)

5.4.5 Overexpression of Rad52 or Rad52 NTD stimulates RNA-dependent DSB repair in yeast cells

Next, using the specific features of Rad52-promoted inverse RNA strand exchange

identified in this study, we wanted to test the relevance of this reaction to RNA-directed

DSB repair *in vivo*. The system, which we developed in yeast to study DSB repair by

RNA in *cis*, consists of a defective *his3* gene expressed from the galactose inducible

promoter, *pGAL1*, in its antisense orientation, and disrupted by an artificial intron (*AI*).

This *AI* can only be spliced from the antisense transcript of *his3* (**Figure C.4**). The *AI*

contains the site for the homothallic switching HO endonuclease. The expression of HO,

also from a *pGAL1* promoter, generates a DSB in *his3* within the *AI*. Following induction

of the *his3* antisense RNA and the DSB by galactose, only repair of the DSB by the

spliced *his3* antisense RNA can restore the functional sequence of the *HIS3* gene and produce histidine prototrophic (His$^+$) cells.

To corroborate the importance of Rad52 function in RNA-dependent DSB repair, we overexpressed yRad52, or either the yeast or human Rad52 NTD, in strains carrying our *cis* system. As a reminder, Rad52 NTD retains catalytic activities of Rad52 including inverse DNA/RNA strand exchange (**Figure 5.3A,B**), but lacks the Rad51 and RPA binding domains [96, 178, 180]. We showed previously that in the absence of RNase H activity DSB repair proceeds using RNA template directly, whereas in its presence it proceeds through a cDNA intermediate [121]. Therefore, we tested the effect of Rad52 or Rad52 NTD overexpression in wild-type yeast cells, in a strain defective in RNase H activity, or in a strain that is both RNase H defective and also carries a null-mutation in the *SPT3* gene that activates reverse transcription in yeast and is thus required for cDNA formation (**Table C.2**) [121]. In all these strains, we observed a significant increase in the frequency of DSB repair by cDNA and RNA upon overexpression yRad52, or either the yeast or human Rad52 NTD (**Table 5.1A**). Strains with deleted endogenous *RAD52* gene showed the largest response; *e.g.,* a 68-fold increase was observed when hRad52$_{1\text{-}209}$ NTD was expressed in *rnh1 rnh201 rad52* cells (**Table 5.1A**). Importantly, the fact that overexpression of the hRad52$_{1\text{-}209}$ NTD stimulated DSB repair by RNA in all studied yeast strains including *rnh1 rnh201* and *rnh1 rnh201 spt3* cells suggests that hRad52 could catalyze DSB repair by RNA in human cells.

5.4.6 Rad59 is not required for RNA-dependent DSB repair

Previously, we found that deletion of the *yRAD51* recombinase gene did not reduce the frequency of DSB repair by transcript RNA [121]. Instead, the frequency of DSB repair by RNA in *rnh1 rnh201 spt3 rad51* cells was significantly elevated compared to that in *rnh1 rnh201 spt3* cells. We proposed that suppression of DSB repair through recombination with sister chromatids, resulted in channeling the broken DNA substrate into the RNA-dependent pathway [121]. Here, we examined whether the yRad59 protein, which shares homology with Rad52 NTD and has partial functional overlap with Rad52 [184, 185], is required for DSB repair by RNA in *cis*. Remarkably, we found that the frequency of DNA repair by RNA is increased by a factor of 4 and 5.7 in *rnh1 rnh201* and *rnh1 rnh201 spt3* cells when the *RAD59* gene is deleted, respectively (**Table 5.1B**). This finding parallels the results in *rad51*-null cells [121], suggesting that RNA-templated DSB repair does not require yRad59 or yRad51. On the contrary, in a control experiment using a ssDNA oligonucleotide as a template for DSB repair in *his3* in the *rad59*-null strains, we found that the repair frequency by ssDNA was significantly reduced by a factor of three in *rnh1 rnh201 spt3 rad59* cells, and a factor of nine in *rnh1 rnh201 rad59* cells compared to *rnh1 rnh201 spt3* and *rnh1 rnh201* cells, respectively (**Table C.3**), as previously shown in *rad59*-null mutant cells [96]. These results suggest that yRad59 has an important role in DNA-templated, but not in RNA-templated, DSB repair. It is relevant to note that while both Rad52 [121] and yRad59 (**Figure 5.3E**) can promote RNA-DNA annealing, only Rad52 has the inverse RNA strand exchange activity (**Figure 5.2**). No such activity was observed for Rad51 or Rad59 (**Figure C.3; Figure 5.3D**). Thus, our findings in yeast and our biochemical data support inverse RNA strand

exchange activity of both yeast and human Rad52 as a unique activity in eukaryotes which can contribute to the mechanism of DSB repair directed by RNA.

5.4.7 RNA-dependent DSB repair is independent of SAE2 and EXO1

While DNA end resection is an essential step in DSB repair via a single-strand annealing (SSA) mechanism [13], we show that inverse strand exchange occurs between ssDNA or ssRNA and homologous duplex DNA that is either non- or minimally resected (**Figure 5.5**). To determine whether the process of resection of broken DNA ends is essential for DSB repair by transcript RNA via HR in yeast cells, we tested the effect of null mutations in *SAE2* and *EXO1* genes, which code for two major factors important for efficient DNA end resection [186-188], in our *cis* system. In the absence of *SAE2* or *EXO1*, the frequency of DSB repair by RNA (in the *rnh1 rnh201 spt3* background) was either increased or not changed, respectively (**Table 5.1B**), suggesting that efficient resection is not required or is even an obstacle for DSB repair by RNA in *cis*. In a control experiment testing DSB repair and using a ssDNA oligonucleotide in an SSA assay [96] in the same strains of the *cis* system, we found that in the *sae2*-null mutant the frequency of His$^+$ colonies was significantly reduced, and in *exo1*-null mutant the frequency was also significantly reduced, although to a lesser extent than in *sae2*-null cells (**Table C.3**). These results support an RNA-dependent mechanism of DSB repair mediated by Rad52 that catalyzes a reaction in which RNA invades a broken dsDNA that is minimally or not resected (**Figure 5.6**).

Table 5.1 Effect of *RAD52* overexpression and lack of *RAD59*, *SAE2*, or *EXO1* on the frequency of RNA-templated DSB repair in *cis* in yeast cells

A	*cis*		
Genotype	His+ freq.		Survival
WT + YEP	600	(500-700)	1.7%
WT + ScRAD52	1,075	(900-1,500)	0.3%
WT + ScRAD52-327	2,600	(2,100-3,200)	0.11%
WT + hRAD52-209	2,100	(1,700-2,500)	0.2%
rnh1 rnh201 + YEP	28,000	(24,000-30,000)	1.2%
rnh1 rnh201 + ScRAD52	57,000	(55,000-60,000)	0.15%
rnh1 rnh201 + ScRAD52-327	86,200	(78,900-95,600)	0.11%
rnh1 rnh201 + hRAD52-209	107,000	(100,000-128,000)	0.16%
rnh1 rnh201 spt3 + YEP	1,600	(1,400-2,000)	2.2%
rnh1 rnh201 spt3 + ScRAD52	2,000	(1,600-2,400)	0.8%
rnh1 rnh201 spt3 + ScRAD52-327	10,700	(8,700-14,000)	0.16%
rnh1 rnh201 spt3 + hRAD52-209	7,800	(6,000-9,700)	0.6%
rad52 + YEP	<10	(0-0)	0.03%
rad52 + ScRAD52	500	(400-700)	0.2%
rad52 + ScRAD52-327	391	(246-1,740)	0.05%
rad52 + hRAD52-209	185	(120-590)	0.04%
rnh1 rnh201 rad52 + YEP	700	(670-1,050)	0.016%
rnh1 rnh201 rad52 + ScRAD52	29,000	(25,000-30,000)	0.13%
rnh1 rnh201 rad52 + ScRAD52-327	9,530	(6,900-13,900)	0.05%
rnh1 rnh201 rad52 + hRAD52-209	48,000	(45,000-54,000)	0.03%

B	*cis*		
Genotype	His+ freq.		Survival
WT	1,400	(1,300-1,900)	1.7%
rad59	2,300	(1,600-2,700)	0.3%
sae2	13,000	(10,000-18,000)	1%
exo1	500	(400-1,000)	2.8%
rnh1 rnh201	70,000	(60,000-119,000)	0.8%
rnh1 rnh201 rad59	285,000	(245,000-344,000)	0.1%
rnh1 rnh201 sae2	549,000	(411,000-652,000)	1%
rnh1 rnh201 exo1	32,000	(28,000-36,000)	2%
rnh1 rnh201 spt3	11,300	(9,500-13,700)	13%
rnh1 rnh201 spt3 rad59	65,000	(45,000-80,000)	0.5%
rnh1 rnh201 spt3 sae2	43,000	(39,000-50,000)	16%
rnh1 rnh201 spt3 exo1	15,000	(14,000-17,000)	13%

Frequencies of His+ colonies per 10^7 viable cells for yeast strains of the *cis* system

following 48 h of galactose treatment are shown as median and 95% CI (in parentheses).

Percentage of cell survival after incubation in galactose is also shown. **A**, Effect of

RAD52 overexpression on DSB repair by cDNA and transcript RNA. There were 12-18

repeats for all the strains. The significance of comparisons between different strains of

the *cis* system was calculated using the Mann-Whitney-*U* test (**Table C.5A**). **B**, Effect of

RAD59, *SAE2* or *EXO1* null mutations on DSB repair by cDNA and transcript RNA.

Percentage of cell survival after incubation in galactose is also shown. There were 6-18

repeats for each strain. The significance of comparisons between different strains of the

cis system was calculated using the Mann-Whitney-*U* test (**Table C.5B**).

DNA with DSB
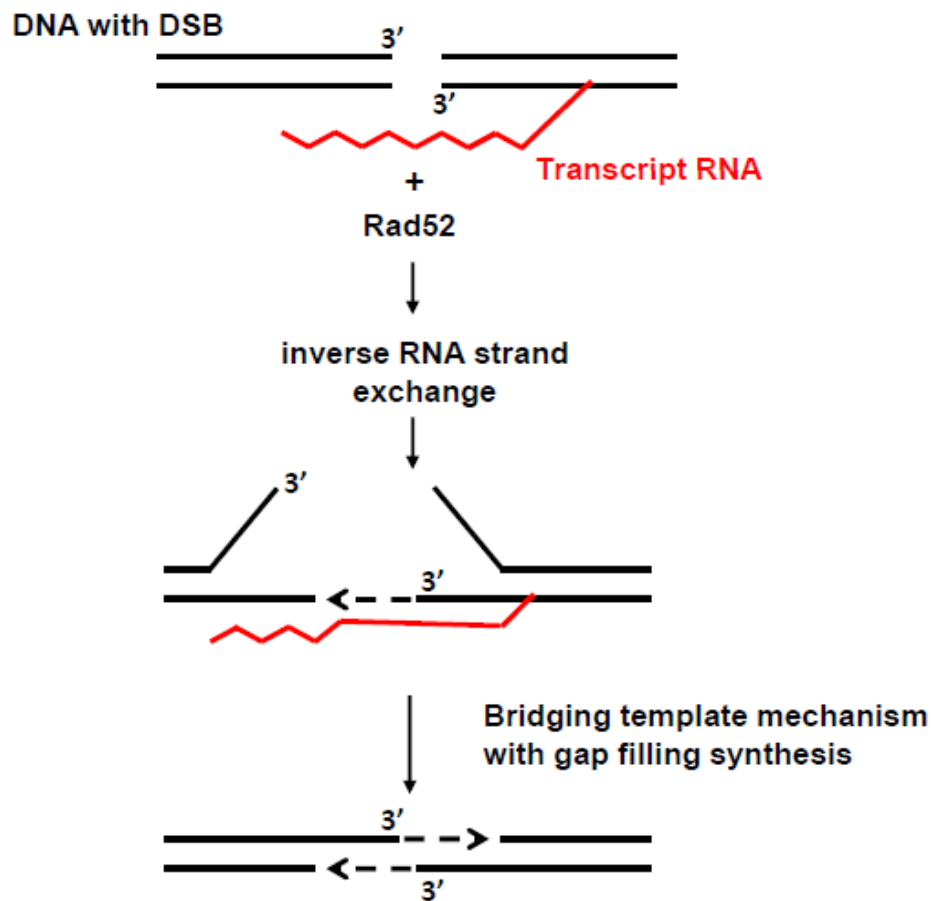
**Figure 5.6 Proposed mechanism of RNA-dependent DSB repair via Rad52 inverse
RNA strand exchange.** Rad52 forms a complex with DSB ends either blunt-ended or
minimally processed by exonucleases/helicases, and then promotes inverse RNA strand

exchange with homologous RNA transcript. The RNA transcript provides a template for guiding end joining or for a short gap filling synthesis (bridging template mechanism). Short DNA synthesis on RNA templates can be carried out by DNA polymerases, which have limited reverse transcriptase activity or by reverse transcriptases [121]. The single-stranded tails are removed by flap nucleases, the gaps are filled in, and any remaining nicks are sealed by DNA ligases restoring the original DNA sequence in an error-free manner. (**by A. V. Mazin**)

5.5 Discussion

Our current *in vitro* and *in vivo* findings on the mechanism of RNA-templated DNA double-strand break repair bring a new perspective to the complex relationship between RNA and DNA in the context of genome stability (summarized in [189]). Recent work revealed an important function of Rad52 in RNA-dependent DSB repair [121]. Here, we describe a novel activity of Rad52, inverse strand exchange that may be responsible for this function. Our *in vitro* results demonstrate that (i) both yeast and human Rad52 promote inverse strand exchange much more efficiently than the forward reaction, in contrast to Rad51 that is more efficient in forward reaction; (ii) Rad52 promotes inverse strand exchange much more efficiently than Rad51 or yRad59; (iii) Rad52 is unique among eukaryotic proteins, as it can utilize both ssDNA and ssRNA in inverse strand exchange; (iv) the hRad52$_{1-209}$ NTD retains the inverse strand exchange activity; (v) the reaction with ssRNA is stimulated by RPA; and (vi) Rad52 can use non resected duplex DNA as a substrate in inverse strand exchange.

Our *in vivo* data in yeast cells corroborate these findings. While both RNA-DNA annealing and inverse RNA strand exchange activities of Rad52 can contribute to RNA-directed DSB repair, our genetic data support a mechanism of RNA-directed DSB repair driven by Rad52-mediated inverse strand exchange. An important watershed between RNA-DNA annealing and inverse RNA strand exchange mechanism lies on the structure of the DNA substrate. While annealing requires that both DNA and RNA being in a single-stranded form, in inverse strand exchange Rad52 may utilize unresected duplex DNA. In accord with the prerequisite of Rad52-promoted inverse strand exchange, RNA-templated DSB repair in yeast cells does not require the function of key resection factors like Sae2 and Exo1. Differently, the frequency of DSB repair by an ssDNA oligo, which follows an SSA mechanism [96], was significantly reduced. These results suggest that RNA-templated DSB repair does not depend on prompt and efficient end resection of broken dsDNA ends, which is consistent with the role of inverse RNA strand exchange in this process. In addition, yRad59 that has RNA-DNA annealing, but not inverse RNA strand exchange activity, is required for DSB repair by an ssDNA oligo, but cannot substitute for Rad52 in RNA-directed DSB repair in yeast cells.

Interestingly, Chakraborty *et al*. recently showed that non-homologous end joining proteins preferentially associate with transcribed sequences following DSB induction and facilitate an error free mechanism of DSB repair in transcribed DNA in mammalian cells [190], supporting an RNA-guided DSB repair mechanism occurring prior to extensive end resection at the DSB ends. Here, we propose that Rad52 inverse RNA strand exchange can contribute to RNA-directed DSB repair in conditions of limited end

resection by generating a heteroduplex between RNA and homologous DNA at the site of DSBs, in which RNA serves as a bridging template guiding DSB repair without or with a short gap filling synthesis (**Figure 5.6**). This mechanism may be especially efficient for DSB repair with reduced end resection, which is encountered in cells that are in the G_1 stage of the cell cycle [191].

It was demonstrated that while Rad52 inactivation alone does not show any significant deficiency in DSB repair in mammalian cells [192], it causes synthetic lethality in combination with mutations in several other HR proteins, including BRCA1 and BRCA2 [193], defects of which are associated with various types of cancer [194]. These data indicate an essential back-up function of Rad52, which may complement the BRCA-dependent HR mechanism in mammals. We suggest that the novel Rad52 inverse strand exchange activities described in the current study may contribute to this back-up function. Thus, our findings may also help to identify new therapeutic targets for cancer.

5.6 Author Contributions

O.M.M. conducted most of the experiments *in vitro*; K.H. purified hRad52$_{1-209}$ NTD and conducted some of the experiments with it; H.K. conducted all *in vivo* experiments and conducted all statistical analysis of *in vivo* data. A.V.M. together with O.M.M., and F.S. with H.K. designed and analyzed *in vitro* and *in vivo* experiments. A.V.M., O.M.M. and F.S. wrote the manuscript with input and suggestions from all authors.

5.7 Acknowledgements

CHAPTER 6

CONCLUSION

Most genomic DNA is transcribed into RNA [195, 196]. Considering the high genomic coverage by coding and non-coding transcript RNA, it is easy to speculate that transcript RNA could be a template for DNA repair, and in particular for DSB repair. Previous findings showed that synthetic RNA-containing oligos can transfer genetic information from RNA sequences to chromosomal or plasmid DNA in yeast, human and bacteria cells [39, 40, 66, 151]. The experiments in yeast cells showed that synthetic RNA only or RNA-containing DNA oligos transformed into yeast cells could precisely repair a broken *leu2* marker gene to generate Leu$^+$ transformants [66]. Deletion of the *SPT3* gene, which is important to promote formation of cDNA in yeast cells, did not affect repair frequency by RNA oligos, strongly suggesting a direct repair of broken DNA ends by template RNA sequences. A question remained whether RNA transcripts generated in the yeast cells could also serve as a template for DSB repair. To address this question, we developed two experimental systems in yeast *S. cerevisiae*. We examined whether an RNA transcript could directly repair a DSB induced i) in the same DNA locus generating the repairing transcript (in *cis*), or ii) in a homologous but different locus from the one generating the repairing transcript (in *trans*) (**Figure 6.1A and B**).
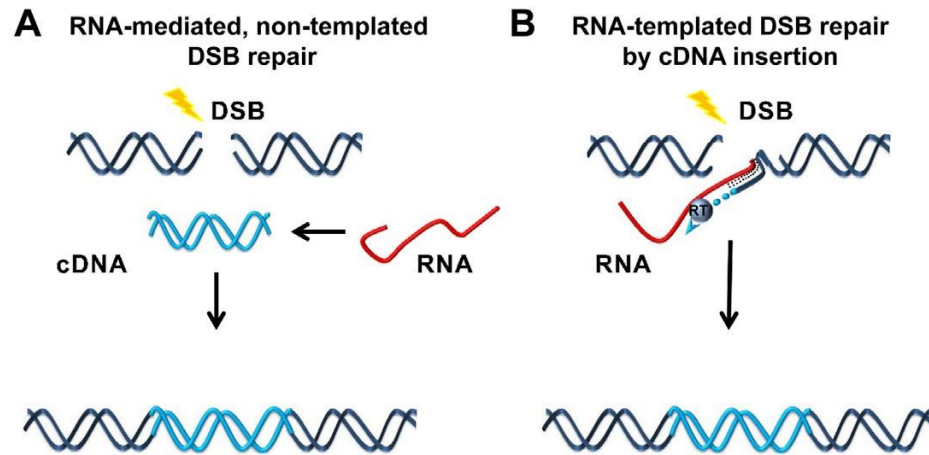
Figure 6.1 Model for RNA-mediated DSB repair (by Havva Keskin in [25]) A) RNA-templated DSB repair in *cis*. **B)** RNA-templated DSB repair in *trans*. DNA is in blue. RNA is in red. RNA polymerase is in yellow pink circle.

6.1 DNA repair by RNA is occurs in the absence of RNases H and *SPT3*

Initial results revealed that transcript RNA could repair the DSB only indirectly, if converted first into DNA copy, cDNA, by the RT function of the yeast retrotransposon Ty. Only when the genes coding for RNases H, which cleaves the RNA strand of RNA-DNA hybrid duplexes, were deleted, DSB repair was detected, even under conditions of minimal Ty RT function. We found that in the absence of *RNH1*, *RNH201* and *SPT3*, DSB repair by transcript RNA was much more frequent in the *cis* system than in the *trans* system, even though the *trans* system generated more *HIS3* RNA than the *cis* system. These data suggest that DSB repair might be due to the proximity of the transcript to the target broken DNA. The work in Ruff *et al.* supports our hypothesis. In this study, it is shown that DNA donor molecules carried in the vicinity of a DSB site in a target DNA gene via fusion with an aptamer sequence specific to the nuclease that generates the DSB mediate gene editing more efficiently than donors that are not fused to the aptamer both

in yeast and human cells [90]. In addition, we showed that in cells lacking both RNase H and *SPT3,* DSB repair was stimulated only by the RNA template and not by the cDNA or a homologous DNA sequence carried on a plasmid, showing that lack of RNase H function promotes RNA-DNA HR but not DNA-DNA HR (**Table 4.3 and Figure 4.4**).

6.2 RNA-mediated DSB repair in the absence of a *bona fide* RT function

To further confirm that His$^+$ cells obtained in the *rnh1 rnh201 spt3* cells of the *cis* system are due to direct RNA repair *vs* repair by residual cDNA that could be present in the yeast cells, it would be relevant to examine DSB repair by RNA in cells that are devoid of Tys. We have initiated work with a yeast strain that lacks Ty transposons from *S. paradoxus,* and thus has no RT function [197]. We built the *cis* system on a centromeric plasmid (**Figure 6.2**), and transformed it into *S. cerevisiae* cells with Tys or Ty-less *S. paradoxus* cells. The *cis* system is under control of the *pGAL1* promoter, and the DSB is generated by expression of the I-SceI endonuclease, which is also regulated by the *pGAL1* promoter. In preliminary results, we observe His$^+$ cells in wild-type or single *RNase H1* or *RNase H2* null mutants only in the presence of functional *SPT3*, confirming DSB repair by cDNA in these backgrounds. We can only detect His$^+$ in the absence of *SPT3* when both RNase H1 and H2 are non functional (**Figure 6.3**). In contrast, in *S. paradoxus*, we cannot detect any His$^+$ cells in wild-type, single *RNase H1* or *RNase H2* null mutants, likely because there is no cDNA formation in these cells. We can, however, detect His$^+$ colonies in the absence of both RNase H1 and H2. These results support a mechanism of direct RNA-templated DSB repair in the absence of RNase H function (**Figure 6.3**). In addition, these data indicate that direct RNA-templated DSB repair can

112

be observed in a yeast species different from *S. cerevisiae*. More tests are underway to corroborate these results.
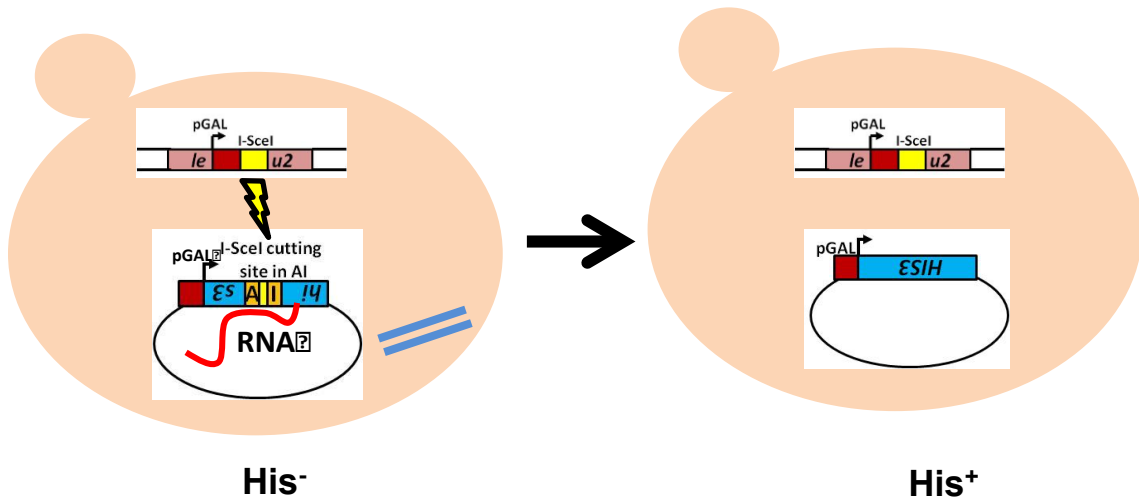
Figure 6.2 Experimental model for RNA-mediated DSB repair in Ty-less strain.

pGAL (galactose inducible promoter) is in red. I-SceI is in yellow. *Leu2* is in light pink.

HIS3 is in blue. Artificial intron (AI) is in orange. Transcript RNA is in red.
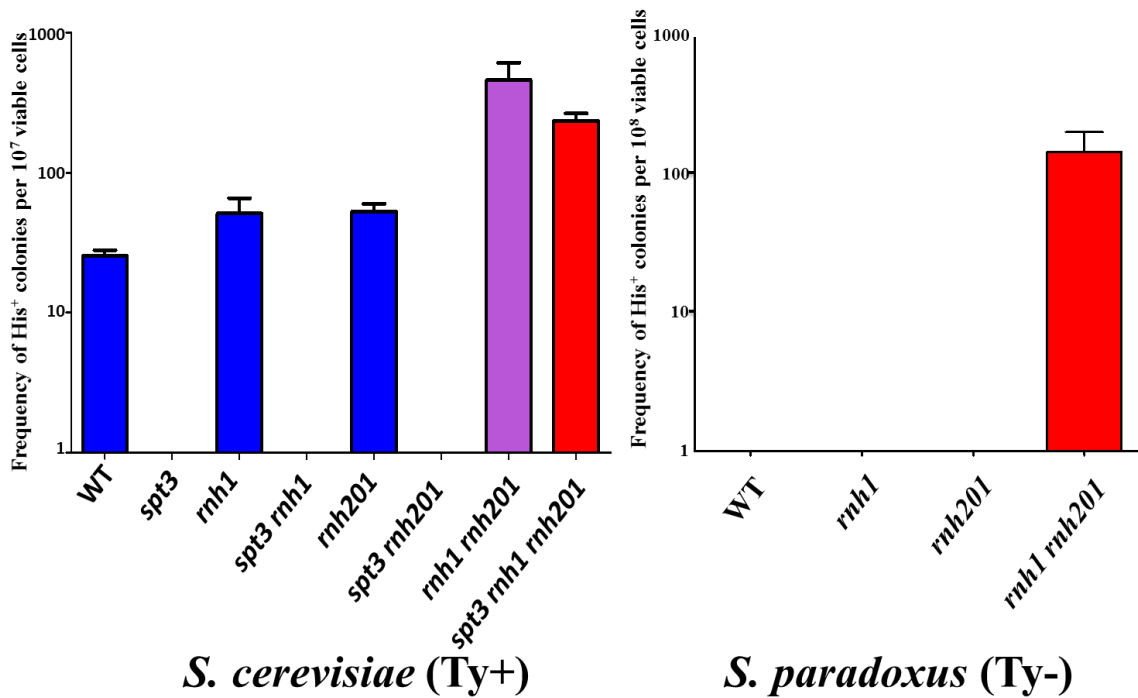
Figure 6.3 Repair frequency for RNA-mediated DSB repair in *S. cerevisiae* and *S. paradoxus* Ty-less strain. Blue bars represent cDNA-mediated DNA repair. Purple bar represents cDNA and RNA-mediated DNA repair. Red bar represents direct RNA-templated DNA repair. Strains genotypes are in X axis and repair frequency is on Y axis.

6.3 Absence of RNase H stimulates cDNA-templated DSB repair

Interestingly, we showed that the loss of RNase H stimulates RNA-templated DSB repair as well as cDNA-templated repair (**Figure 6.4**). These results suggest that cDNA and/or RNA-cDNA hybrid molecules derived from reverse transcription are more abundant and/or more stable when there is no RNase H function. Mutations in any of the RNase H2 subunit are associated with AGS disease [136]. The possible accumulation of RNA-DNA hybrids in defective RNase H2 cells could be a trigger for the disease [64]. AGS patients with defects in RNase H2 could have an increased level of cDNA in the form of RNA-DNA hybrids, which could play a role in activating the immune system. We showed that in our yeast *cis* system, deletion of any RNase H2 subunit increases DSB repair frequency by cDNA. Moreover, yeast AGS orthologous mutations combined with *RNH1* deletion resulted in increased repair of a DNA DSB. These findings suggest that RNA-DNA hybrids and cDNA in AGS patients could be abundant. In support our hypothesis, Lim *et al*. detected RNA-DNA hybrids in AGS patients having mutations in *TREX1* (AGS1), *RNase H2* (AGS2, 3, 4) and *SAMDH1* (AGS5) [136].
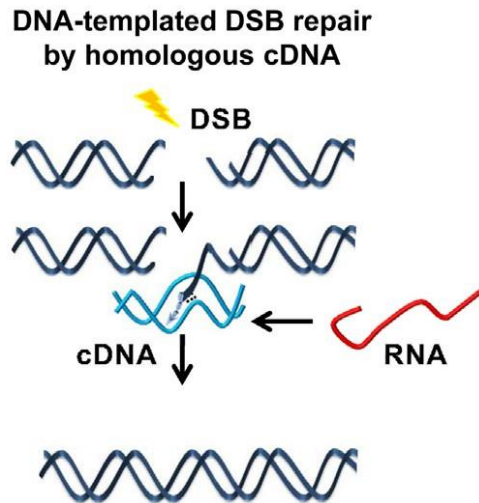
Figure 6.4 Model for cDNA-mediated DSB repair (by Havva Keskin in [25]. DNA is in blue. RNA is in red. cDNA is in light blue.

6.4 Molecular mechanism of RNA-mediated DSB repair

6.4.1 RNA-mediated DSB repair is dependent on *RAD52* but not *RAD51* or *RAD59*

Following the results showing the evidence of RNA-templated DNA DSB repair, its molecular mechanism needed to be determined. We found that RNA-transcript mediated DSB repair is dependent on Rad52, but not Rad51 (**Figure 2.1b**). These data suggest that transcript RNA stimulates the DSB repair via Rad52 mediated annealing (**Figure 2.3**). Furthermore, we studied the effect of overexpression of full-length yeast Rad52 (yRad52) and yeast and human Rad52 N terminal Domain (NTD), which have the activity of annealing but do not have Rad51 and RPA binding domains, in yeast cells carrying the *cis* system. Upon overexpression of yRad52, yeast or human Rad52 NTD, we observed increased DSB repair frequency in wild-type, *rnh1 rnh201*, and *rnh1 rnh201 spt3* strains as well as in null-*rad52* mutant strains in the presence or absence of RNases H function.

These data underline the importance of Rad52 in RNA-templated DNA repair. Moreover, increasing DSB repair frequency by overexpression of human Rad52 NTD in *rad52 rnh1 rnh201* cells suggests that human Rad52 could catalyze RNA-templated DNA repair not only in yeast but possibly also in human cells. In addition, we tested the effect of yeast Rad59, which has homology and overlapping function with Rad52 [184, 185], in RNA-templated DNA repair. Interestingly, deletion of the *RAD59* gene increased the frequency of RNA-templated DSB repair (**Table 5.1**). This result was similar to the deletion of *RAD51* suggesting that Rad59 and Rad51 are not required for RNA-templated DSB repair. In the *in vitro* collaborative work with Dr. A. Mazin using human and yeast Rad52 purified proteins, we showed that Rad52 catalyzes annealing between RNA and DNA strands. This data also support our *in vivo* findings, suggesting that transcript RNA-templated DNA repair could occur in human cells. We also studied the null mutants of *SAE2* and *EXO1,* which are important genes for efficient DNA-end resection, whether or not end resection is required in RNA-templated DSB repair. Deletion of *SAE2* or *EXO1* increased or did not change the DSB repair frequency by RNA, respectively. On the contrary, these deletions reduced the frequency of DSB repair by DNA oligos significantly (**Table A.3**). These data suggest that RNA-templated DSB repair does not require efficient DNA end resection.

6.4.2 End-resection is not required for RNA-templated DSB repair

If DNA end resection is not required for RNA-tamplated DSB repair, in order to repair the DSB, RNA should then invade a broken duplex DNA rather than simply annealing to a single-stranded resected DNA end. Inverse strand exchange reaction happens when a

recombinase protein like the Rad51 homolog RecA protein forms a nucleoprotein

filament on dsDNA to catalyze the exchange reaction with either ssDNA or ssRNA, as

shown in experiments *in vitro* [33, 34]. In more recent work in collaboration with Dr.

Mazin, we found that human and yeast Rad52, but not Rad51 or Rad59, efficiently

promote inverse strand exchange between dsDNA and ssDNA, and also between dsDNA

and ssRNA. These data suggest that inverse strand exchange could be a better mechanism

for RNA-templated DNA DSB repair (**Figure 5.6**).

The work of this study demonstrates a new and unique way to repair DNA DSBs using

endogenous RNA transcript as template for HR that possibly requires limited end resection

(**Figure 6.5**).
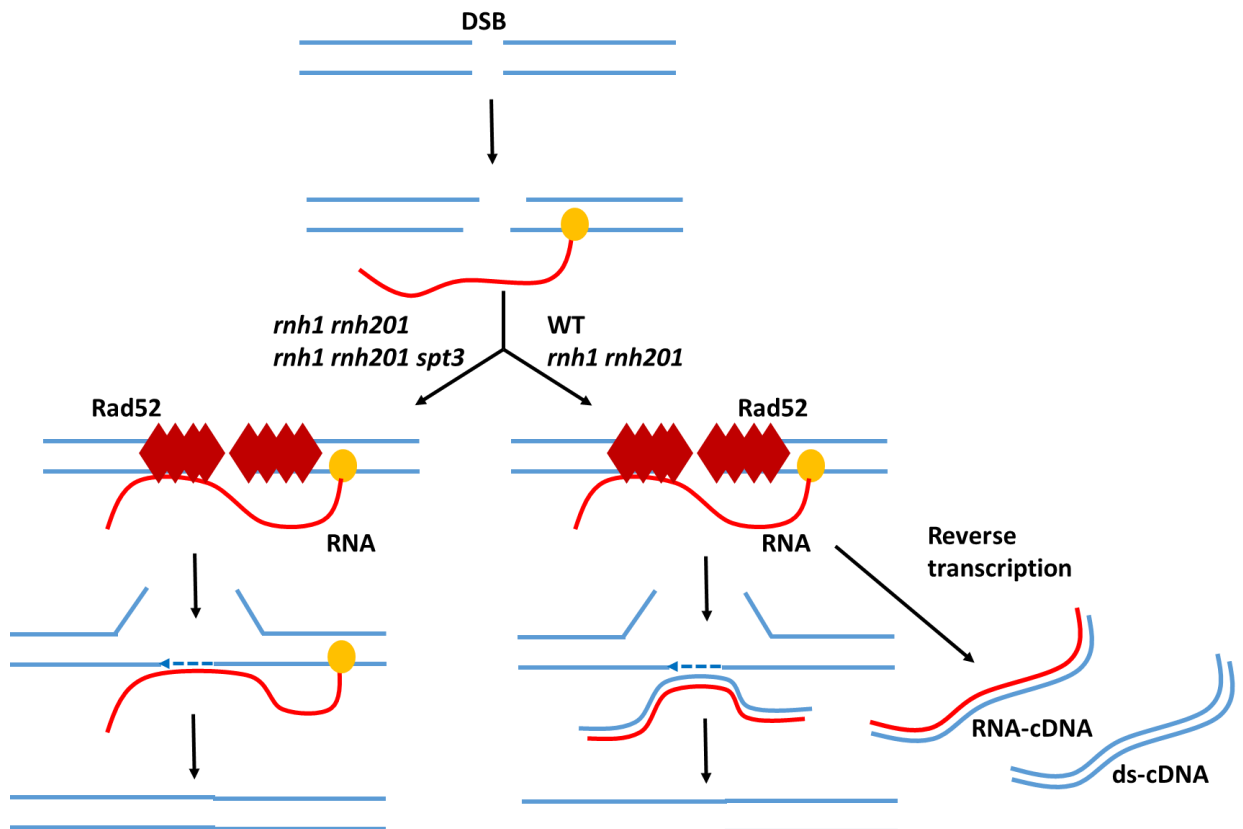
Figure 6.5 Current model for RNA, cDNA and RNA-cDNA-mediated DSB repair.

DNA is in blue. RNA is in red. RNA-cDNA hybrid is in blue and red and ds cDNA is in blue. RNA polymerase is in yellow. Rad52 is in dark red.

6.5 Research directions

Details of the direct DSB repair mechanism still need to be determined. For example, it would be interesting to unravel which DNA polymerases and helicase function/s can use the transcript RNA as template or substrate during DSB repair. In our transcript RNA-templated DSB repair assay, we did not detect DSB repair when we deleted the 5'-splice site of the artificial intron in the *his3* marker, which is crucial for splicing (**Table A.5d**), demonstrating that splicing is essential in our system to detect DSB repair by transcript RNA. We could study the effect of hypomorphic mutations in factors important for RNA metabolism, like RNA splicing factors or RNA export factors on RNA-templated DNA repair.

To better understand the involvement of Rad52 in RNA-tamplated DSB repair, we could study the effect of specific mutants of the Rad52 recombinase protein. For example, we could generate mutations that are known to affect binding of Rad52 to ssDNA, like human R55A and yeast R70A mutant of Rad52 [198], and determine how these mutants expressed in our *cis* system affect the frequency of RNA-templated DSB repair in yeast cells that are *RAD52* wild-type or *rad52*-null. This experiment could tell us whether ssDNA-binding domain of Rad52 is also important to bind ssRNA. Parallel *in vitro*

studies using same mutant forms of yeast and/or human Rad52 could help us to better understand the activity of Rad52 in RNA-templated DSB repair.

Work of our lab by Katz *et al.* [199] showed that a nick, a single-strand break in DNA, stimulates gene correction by ssDNA oligos in yeast and human cells. It would be interesting to determine whether a nick activates HR between DNA and RNA, and to examine also whether RNA-DNA HR could occur without a specific DSB or nick induction in DNA.

NHEJ and HR are the major mechanisms to repair DNA DSBs. NHEJ is not restricted in any phase of the cell cycle but this process is often error prone [2]. In contrast, HR is mostly an error free mechanism using a sister chromatid for repair in mitotic cells but it can only happen in S/G_2 phase of the cell cycle when sister chromatids are available [2]. Because of this restriction, the G_0 or G_1 phase of the cell cycle has apparently no mechanism for accurate repair of DSBs, and templates alternative to sister chromatids would be essential to maintain genome stability. Given the importance of having a template for DSB repair in the G_0 or G_1 phase of the cell cycle, and considering the abundance of RNA transcripts in cells, it would be exciting to detect indirect or direct repair of DNA DSBs by endogenous RNA transcripts in cells that cannot replicate such as non-dividing cells, or terminally differentiated cells. Moreover, our yeast systems to study DSB repair by RNA could be translated to other organisms/or cell types, such as human cells, to determine whether RNA-templated DSB repair is conserved in mammalian systems. To implement such translational studies, we could use the modular

CRISPR/Cas9 system [200] to create efficient DSBs in human DNA at desired genetic

loci, and determine whether we could detect RNA and/or cDNA-templated DNA DSB

repair in these cells.

APPENDIX A

SUPPLEMENTARY MATERIALS FOR CHAPTER 2

Table A.1 Yeast strains used in this study

Strain	Relevant genotype	Source
FRO-767	*hoΔ hmlΔ::ADE1 matαΔ::hisG hmrΔ::ADE1 ade1 leu2::HOcs lys5 trp1::hisG ura3-52 ade3::GAL::HO*	3
FRO-1072	*hoΔ hmlΔ::ADE1 Δmatα::hisG hmrΔ::ADE1 ade1 leu2::pGAL1-mhis3AI-URA3 lys5 trp1::hisG ura3-52 ade3::GAL::HO*	this study
FRO-1073	*hoΔ hmlΔ::ADE1 matαΔ::hisG hmrΔ::ADE1 ade1 leu2::pGAL1mhis3AI-ADE3 lys5 trp1::hisG ura3-52 ade3::GAL::HO*	this study
FRO-1074	FRO-1073 *his3::CORE-UH*	this study
FRO-1075,1080 (*trans*)	FRO-1073 *his3::HOcs*	this study
FRO-1092, 1093	*hoΔ hmlΔ::ADE1 matαΔ::hisG hmrΔ::ADE1 ade1 leu2::pGAL1-ADE3 lys5 trp1::hisG ura3-52 ade3::GAL::HO his3::HOcs rad52Δ::kanMX4*	this study
YS-164, 165	FRO-1075, 1080 (*HIS3::HOcs*)::*TRP1*	this study
YS-166, 167	YS-164, 165 *pGAL1-mhis3AI::CORE*	this study
YS-172, 174 (*cis*)	YS-166, 167 *pGAL1-mhis3AI::CORE*	this study
YS-275, 276	FRO-1075, 1080, *YCLWTy2-1::CORE*	this study
YS-278, 281	YS-172, 174 *YCLWTy2-1::CORE*	this study
YS-289, 290 (*trans*) WT	YS-275, 276 *YCLWTy2-1Δ*	this study
YS-291, 292 (*cis*) WT	YS-278, 281 *YCLWTy2-1Δ*	this study
YS-414, 415 (*trans*)	YS-289, 290 *rnh1Δ::kanMX4*	this study
YS-416, 417 (*cis*)	YS-291, 292 *rnh1Δ::kanMX4*	this study
YS-410, 411 (*trans*)	YS-289, 290 *rnh201Δ::hygMX4*	this study
YS-412, 413 (*cis*)	YS-291, 292 *rnh201Δ::hygMX4*	this study
YS-428, 429 (*trans*)	YS-289, 290 *spt3Δ::kanMX4*	this study
YS-440, 441 (*cis*)	YS-291, 292 *spt3Δ::kanMX4*	this study
YS-444, 445 (*cis*)	YS-291, 292 *rad52Δ::kanMX4*	this study
YS-446, 447 (*cis*)	YS-291, 292 *rad51Δ::kanMX4*	this study
HK-76, 77 (*trans*)	YS-289, 290 *dbr1Δ::kanMX4*	this study
HK-72, 73 (*cis*)	YS-291, 292 *dbr1Δ::kanMX4*	this study
YS-520, 521 (*trans*)	YS-414, 415 *spt3Δ::hygMX4*	this study
YS-522, 524 (*cis*)	YS-416, 417 *spt3Δ::hygMX4*	this study
YS-452, 453 (*trans*)	YS-410, 411 *spt3Δ::hygMX4*	this study
YS-464, 465 (*cis*)	YS-412, 413 *spt3Δ::hygMX4*	this study
YS-422, 423 (*trans*)	YS-289, 290 *rnh1Δ::NAT rnh201Δ::hygMX4*	this study
YS-424, 426 (*cis*)	YS-291, 292 *rnh1Δ::NAT rnh201Δ:: hygMX4*	this study
YS-476, 477 (*trans*)	YS-289, 290 *rnh1Δ::NAT rnh201Δ::hygMX4 spt3Δ::kanMX4*	this study
YS-486, 487 (*cis*)	YS-291, 292 *rnh1Δ::NAT rnh201Δ::hygMX4 spt3Δ::kanMX4*	this study
YS-490, 491 (*cis*)	YS-424, 426 *rad52Δ::kanMX4*	this study
YS-492, 493 (*cis*)	YS-424, 426 *rad51Δ::kanMX4*	this study
HK-78, 79 (*trans*)	YS-422, 423 *dbr1Δ::kanMX4*	this study
HK-74, 75 (*cis*)	YS-424, 426 *dbr1Δ:: kanMX4*	this study
HK-213, 215 (*trans*)	YS-422, 423 *dbr1Δ::KlURA3*	this study
HK-217, 219 (*cis*)	YS-424, 426 *dbr1Δ::KlURA3*	this study
HK-136, 137 (*trans*)	YS-422, 423 *spt3Δ::KlURA3*	this study
HK-138, 139 (*cis*)	YS-424, 426 *spt3Δ::KlURA3*	this study
HK-194, 197 (*cis*)	HK-138, 139 *rad52Δ::kanMX4*	this study
HK-180, 184 (*cis*)	HK-138, 139 *rad51Δ::kanMX4*	this study
HK-112, 113 (*trans*)	HK-78, 79 *spt3Δ::KlURA3*	this study
HK-110, 111 (*cis*)	HK-74, 75 *spt3Δ::KlURA3*	this study
YS-526, 527 (*cis*)	YS-291 *pGAL1Δ::KlURA3*	this study
YS-528, 529 (*cis*)	YS-424, 426 *pGAL1Δ::KlURA3*	this study
YS-530, 531 (*cis*)	YS-486, 487 *pGAL1Δ::KlURA3*	this study
YS-532, 533 (*cis*)	YS-291, 292 *ade3::GAL::hoΔ::KlURA3*	this study
YS-534, 535 (*cis*)	YS-424, 426 *ade3::GAL::hoΔ::KlURA3*	this study
YS-536, 537 (*cis*)	YS-486, 487 *ade3::GAL::hoΔ::KlURA3*	this study
HK-9, 10 (*cis*)	YS-291, 292 + Yep195spGAL	this study
HK-11, 12 (*cis*)	YS-291, 292 + Yep195spGAL-*RNH201*-WT	this study
HK-13, 14 (*cis*)	YS-291, 292 + Yep195spGAL-*rnh201*-D39A	this study
HK-15, 16 (*cis*)	YS-440, 441 + Yep195spGAL	this study
HK-17, 18 (*cis*)	YS-440, 441 + Yep195spGAL-*RNH201*-WT	this study
HK-19, 20 (*cis*)	YS-440, 441 + Yep195spGAL-*rnh201*-D39A	this study
HK-21, 22 (*cis*)	YS-424, 426 + Yep195spGAL	this study
HK-23, 24 (*cis*)	YS-424, 426 + Yep195spGAL-*RNH201*-WT	this study
HK-25, 26 (*cis*)	YS-424, 426 + Yep195spGAL-*rnh20*- D39A	this study
HK-27, 28 (*cis*)	YS-486, 487 + Yep195spGAL	this study
HK-29, 30 (*cis*)	YS-486, 487 + Yep195spGAL-*RNH201*-WT	this study
HK-31, 32 (*cis*)	YS-486, 487 + Yep195spGAL-*rnh201*-D39A	this study
YS-301	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 trp5(ΔCC1001-2; G1017→A)*	43
YS-305	YS-301 *rnh201Δ::kanMX4*	43
KK-72	YS-305 *rnh1Δ::hygMX4*	this study
TY-32, 52	YS-301 + BDG102 (empty vector)	this study
TY-17, 53	YS-301 + BDG598 (*pGTy-H3mHIS3AI*)	this study
TY-36, 66	KK-72 + BDG102 (empty vector)	this study
TY-22, 67	KK-72 + BDG598 (*pGTy-H3mHIS3AI*)	this study
HK-386, 388 (*cis*)	YS-291, 292 *mhis3AI::CORE*	this study
HK-382, 384 (*cis*)	YS-424, 426 *mhis3AI::CORE*	this study
HK-396, 400 (*cis*)	HK-386, 388 *AIΔ23*	this study

S. cerevisiae strains used in this study. Strains specifically used for the *trans* or *cis* assay are indicated.

Table A.2 Oligonucleotides used in this study and sequence patterns of the *HIS3*

region repaired by transcript RNA or via non-homologous end-joining

a

Name	Size	Sequence	Experiment
HIS3.F	80	5'ACCAATGCACTCAACGATTAGCGACCAGCCGGAATGCTTG GCCAGAGCATGTATCATATGGTCCAGAAACCCTATACCTG	Transformation
HIS3.R	80	5'CAGGTATAGGGTTTCTGGACCATATGATACATGCTCTGGC CAAGCATTCCGGCTGGTCGCTAATCGTTGAGTGCATTGGT	Transformation
His3.F2	20	5' CCTGTTCTGCTACTGCTTCT	qRT-PCR
His3.R2	20	5' CGATCTCTTTAAAGGGTGGT	qRT-PCR
ACT1.F	20	5' TTGGATTCCGGTGATGGTGT	qRT-PCR
ACT1.R	20	5' CGGCCAAATCGATTCTCAAA	qRT-PCR
CEN16.F	20	5' TGAGCAAACAATTTGAACAG	qRT-PCR
CEN16.R	18	5' CCGATTTCGCTTTAGAAC	qRT-PCR
His3.2	20	5' GAGAGCAATCCCGCAGTCTT	Colony PCR
His3.5	20	5' ATGACAGAGCAGAAAGCCCT	Colony PCR
HO.F	20	5' AACCACTCTACAAAACCAAA	Colony PCR
INTRON.F	20	5' GTATGTTAATATGGACTAAA	Colony PCR
S3.1	20	5' TTAAAGAGGCCCTAGGGGCC	Southern blot
S3.2	20	5' CTACATAAGAACACCTTTGG	Southern blot
S3.3	20	5' TTTGCGCCTTTGGATGAGGC	Southern blot
S3.4	20	5' TTGGGCGAGGTGGCTTCTCT	Southern blot
211	48	5' GAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA TTTGAAT	Rad52 Annealing
501	48	5' CUUCGUAAAUAGUCCCAAUAACAGAGUACUCGCCUAUG UAUAAACUUA	Rad52 Annealing
508	53	5' ATTCAA ATATGTATCCGCTAATGAGACAATAACCCTGATAA ATGCTTCACTAG	Rad52 Annealing
509	32	5' TTATTGTCTCATTAGCGGATACATATTTGAAT	Rad52 Annealing

b

Pattern	Tract of *HIS3* or *his3* sequence next to the HO site insertion
HIS3	5' -CATATGATACATGCTCTGGCCAAGCATTCCGGCTGGTCGCT-
his3::HO	5' -CATATGATACATGCTCTGGC--*HO*--CAAGCATTCCGGCTGGTCGCT-
A	5' -CATATGATACATGCTCTGGC**GGTA**CATTCCGGCTGGTCGCT-
B	5' -CATATGATACATGCTCTGGC**GGTC**CATTCCGGCTGGTCGCT-
C	5' -CATATGATACATGCTCTGGC**GGTG**CATTCCGGCTGGTCGCT-

a, Name, size and sequence of the oligonucleotides used in this study are described. The specific experiments in which the oligonucleotides were used are indicated. b, Sequence patterns of the *HIS3* region repaired by transcript RNA or via non-homologous end-joining. All 24 His$^+$ *cis*-system *rnh1 rnh201 spt3* clones that were sequenced perfectly matched the wild-type *HIS3* sequence. In contrast, when we examined the sequence of the rare His$^+$ clones that we could obtain (~10 out of 10^7 viable cells) from a strain that had the homothallic switching endonuclease site in *his3* on chromosome (Chr) XV (the construct is identical to that described in Extended Data Fig. 1b) and was *rad52*-null (FRO-1092, 1093), 29 out of 29 His$^+$ samples had replaced 4 nucleotides (CAAG) of *his3* next to the homothallic switching endonuclease site with a new sequence. Differences from the wild-type *HIS3* gene are in bold. A–C, patterns of the *HIS3* region from spontaneous His$^+$ revertants. Among the 29 sequenced *HIS3* regions, 25 displayed pattern A, 3 displayed pattern B and 1 displayed pattern C. The four bases inconsistent with the wild-type *HIS3* affected two codons, causing a silent mutation (GCCRGCG: AlaRAla) and a missense mutation (AAGRGTA, GTC or GTG:LysRVal).

Table A.3 His⁺ frequency in the *trans* and *cis* systems following transformation by HIS3.F and HIS3.R oligonucleotides

Genotype	*trans*				*cis*			
	No Oligo		HIS3.F + HIS3.R		No Oligo		HIS3.F + HIS3.R	
WT	2.3	(0-8)	1.6×10^6	$(1.4 \times 10^6 - 1.9 \times 10^6)$	<0.1	(0-0)	1.5×10^6	$(946,000 - 2.3 \times 10^6)$
rnh1 rnh201	8	(0-56)	1×10^6	$(1.1 \times 10^5 - 1.9 \times 10^6)$	165	(63-275)	845,500	$(669,000 - 1 \times 10^6)$
rnh1 rnh201 spt3	1.7	(1-2)	215,480	(196,000-235,000)	49	(25-78)	225,300	(156,000-326,700)
rnh1 rnh201 spt3 AlΔ23	ND		ND		<0.1	(0-0)	798,370	$(610,100 - 1 \times 10^6)$

Frequency of His⁺ transformant colonies per 10^7 viable cells for wild-type (WT), *rnh1 rnh201*, and *rnh1 rnh201 spt3* mutant strains after transformation with HIS3.F and HIS3.R oligonucleotides in both *trans* and *cis* systems is shown as median and 95% confidence interval (in brackets). There were four or eight repeats for each of the strains transformed with these oligonucleotides. The significance of comparisons between the strains in the *trans* and the *cis* systems, and between different strains of the *trans* or *cis* system, that is between-group and within-group analysis, were calculated using the Mann–Whitney U-test (Supplementary Table 1d). The strains used in this experiment were YS-289, YS-290, YS-291, YS-292, YS-422, YS-423, YS-424, YS-426, YS-476, YS-477, YS-486, YS-487 and HK-404, HK-407. ND, not determined.

*We note that when the trans- and cis-system *rnh1 rnh201* or *rnh1 rnh201 spt3* strains were transformed using exogenous HIS3.F and HIS3.R synthetic oligonucleotides following DSB induction, the frequencies of His1 colonies were similar to each other in the *trans*- and *cis*-system *rnh1 rnh201* or *rnh1 rnh201 spt3* cells. In contrast, when no oligonucleotides were added, the few His⁺ colonies were 20- to 28-fold more numerous

in *cis*- than in *trans*-system *rnh1 rnh201* or *rnh1 rnh201 spt3* cells, respectively, probably

originating from DSB repair by the *his3* antisense transcript.

Table A.4 His⁺ frequencies in the presence of plasmid BDG283 or BDG606 in *cis*

strains

Galactose			
	cis		
Genotype	Ura⁺His⁺ freq.		Survival
WT + BDG283	36	(27-45)	9%
WT + BDG606	157,000	(143,020-193,000)	9%
rnh1 rnh201 spt3 + BDG283	820	(720-900)	25%
rnh1 rnh201 spt3 + BDG606	815	(680-900)	25%
Glucose			
	cis		
Genotype	Ura⁺His⁺ freq.		Survival
WT + BDG283	<0.01	(0-0)	56%
WT + BDG606	<0.01	(0-0)	50%
rnh1 rnh201 spt3 + BDG283	0.28	(0.04-0.45)	93%
rnh1 rnh201 spt3 + BDG606	8	(0-24)	80%

Frequencies of Ura⁺His⁺ colonies per 10^7 viable cells for yeast strains of the *cis* cell

system transformed with plasmid BDG283 or BDG606 following 48 h galactose or

glucose treatment are shown as median and 95% confidence interval (in brackets).

Percentage of cell survival after incubation in galactose or glucose is shown. There were

six repeats for all the strains. The significance of comparisons between strains was

calculated using the Mann–Whitney U-test (supplementary table 1e).

Table A.5 His⁺ frequencies for strains with *dbr1*-null, grown in the presence of PFA, with and without the *pGAL1* promoter, grown in glucose, or containing the AIΔ23

a

Genotype	trans His⁺ freq.		Survival	cis His⁺ freq.		Survival
dbr1	1,330	(1,030-1,660)	1.6%	23	(0-47)	2%
rnh201 dbr1	2,130	(1,150-3,620)	2.6%	322	(122-453)	3%
rnh1 dbr1	2,455	(1,500- 3,250)	1.2%	18	(0-78)	2.5%
rnh1 rnh201 dbr1	7,420	(7,400-11,300)	1.7%	29,900	(26,900-33,200)	1.2%
WT + PFA	519	(400-1,300)	1.7%	112	(94-380)	0.9%
rnh1 rnh201 + PFA	4,120	(3,100-5,340)	0.9%	9,400	(7,290-20,800)	0.7%

b

Genotype	cis His⁺ freq.		Survival
WT	1,050	(600-1,460)	1%
rnh1 rnh201	62,100	(52,900-68,900)	0.7%
rnh1 rnh201 spt3	5,100	(3,660-6,660)	11%
pGAL1Δ	<1	(0-0)	0.4%
rnh1 rnh201 pGAL1Δ	540	(270-1,300)	0.4%
rnh1 rnh201 spt3 pGAL1Δ	630	(500-920)	0.8%

c

Genotype	trans His⁺ freq.		Survival	cis His⁺ freq.		Survival
WT	<0.01	(0-0)	16%	<0.01	(0-0)	19%
spt3	<0.01	(0-0)	96%	<0.01	(0-0)	93%
dbr1	<0.01	(0-0)	33%	<0.01	(0-0)	54%
rad52	ND		ND	<0.01	(0-0)	6%
rad51	ND		ND	<0.01	(0-0)	24%
pGAL1Δ	ND		ND	<0.01	(0-0)	67%
rnh1 rnh201	11	(5-25)	18%	21	(17-31)	11%
rnh1 rnh201 spt3	4	(2-14)	92%	9	(0.3-16)	76%
rnh1 rnh201 dbr1	<0.01	(0-0)	28%	1.5	(0-6)	34%
rnh1 rnh201 rad52	ND		ND	<0.01	(0-0)	23%
rnh1 rnh201 rad51	ND		ND	<0.01	(0-7)	17%
rnh1 rnh201 pGAL1Δ	ND		ND	0.9	(0-2)	45%
rnh1 rnh201 spt3 rad52	ND		ND	2	(0-5)	26%
rnh1 rnh201 spt3 rad51	ND		ND	2	(0-4)	50%
rnh1 rnh201 spt3 pGAL1Δ	ND		ND	<0.01	(0-0)	85%

d	*cis*		
Genotype	His+ freq.		Survival
WT	1,000	(840-1,240)	2%
rnh1 rnh201	43,100	(37,500-47,000)	1.7%
rnh1 rnh201 spt3	4,180	(3,310-5,550)	21%
AlΔ23	<0.1	(0-0)	1.7%
rnh1 rnh201 AlΔ23	<0.1	(0-0)	1.7%
rnh1 rnh201 spt3 AlΔ23	<0.1	(0-0)	15%

a, Frequencies of His+ colonies per 10^7 viable cells for yeast strains of the *trans* and *cis* cell systems following 48 h galactose treatment are shown as median and 95% confidence interval (in brackets). Percentage of cell survival after incubation in galactose is also shown. Eighteen repeats for dbr1 (in *trans*), 6 repeats for dbr1 (in *cis*); 6 repeats for *rnh201 dbr1, rnh1 dbr1*; 24 repeats for *rnh1 rnh201 dbr1* and 4 repeats for PFA data. The significance of comparisons between strains was calculated using the Mann–Whitney U-test (Supplementary Table 1a). **b**, Frequencies of His+ colonies per 10^7 viable cells for yeast strains of the *cis* cell system following 48 h galactose treatment are shown as median and 95% confidence interval (in brackets). Percentage of cell survival after incubation in galactose is also shown. There were 6 repeats for all the strains. The significance of comparisons between strains was calculated using the Mann–Whitney U-test (Supplementary Table 1f). **c**, Frequencies of His1 colonies per 10^7 viable cells for the indicated yeast strains following 24-h glucose treatment in both the trans and the *cis* cell systems are shown as median and 95% confidence interval (in brackets). Percentage of cell survival after growth in glucose is also shown. There were 8 repeats for each of the strains. The significance of comparisons between the strains in the *trans* and *cis* systems was calculated using the Mann–Whitney U-test (Supplementary Table 1g). ND, not determined. **d**, Frequencies of His+ colonies per 10^7 viable cells for yeast strains of the cis

cell system following 48 h galactose treatment are shown as median and 95%confidence interval (in brackets). Percentage of cell survival after incubation in galactose is also shown. There were six repeats for all the strains. The significance of comparisons between strains was calculated using the Mann–Whitney U-test (Supplementary Table 1h).

Table A.6 His+ rates in wild-type and *rnh1 rnh201* cells resulting from the transposition assay At 22 °C or 30 °C

22 °C	No gal (Ura$^-$)		+ gal (Ura$^-$Gal)	
Genotype	His+ rate (x10^{-7})	Survival	His+ rate (x10^{-3})	Survival
WT + BDG598	5.28 (0 – 141)	26%	2.68 (2.55 – 3.06)	15%
rnh1 rnh201 + BDG598	15.3 (16.3 – 42.4)	34%	0.78 (0.54 – 0.92)	27%
30 °C	No gal (Ura$^-$)		+ gal (Ura$^-$Gal)	
Genotype	His+ rate (x10^{-7})		His+ rate (x10^{-3})	
WT + BDG598	2.8* (0 – 7.37)	26%	0.58 (0.46 – 0.72)	15%
rnh1 rnh201 + BDG598	16.1 (5.31 – 24.2)	34%	0.04 (0.03 – 0.06)	27%
30 °C	No gal (YPLac)		+ gal (YPLac + gal)	
Genotype	His+ rate (x10^{-7})		His+ rate (x10^{-5})	
WT + BDG598	<0.1 (0 – 0)	26%	1.38 (0.52 – 2.38)	15%
rnh1 rnh201 + BDG598	15.1 (4.90 – 26.4)	34%	0.4 (0.30 – 0.60)	27%

Shown are rates of His+ colonies for wild-type (WT) and *rnh1 rnh201* yeast strains containing BDG598 following growth with no galactose with plasmid selection (Ura$^-$ medium) and without plasmid selection (YPLac medium) or galactose with plasmid selection (Ura$^-$Gal medium) and without plasmid selection (YPLac+gal medium) for 96 h

at 22 °**C**, or for 48 h at 30 °**C**. Data are presented as median and 95% confidence interval

(in brackets). Percentages of cell survival after growth without or with galactose are also

shown. There were 15 repeats for the strains incubated at 22 °**C** and 6 repeats for those

incubated at 30 °**C**. The significance of comparisons between strains was calculated using

the Mann–Whitney U-test (Supplementary Table 1i). The strains used in this experiment.

Table A.7 Statistical analysis (P-values and adjusted P-values) of the data a
l)

Genotype	P values		Adj. P values	
	trans	cis	trans	cis
WT *vs. spt3*	<0.0001	<0.0001	0.0003	0.0003
WT *vs. rnh1*	<0.0001	0.3708	0.0003	0.4194
WT *vs. rnh201*	<0.0001	<0.0001	0.0003	0.0003
WT *vs. rnh1 spt3*	<0.0001	<0.0001	0.0003	0.0003
WT *vs. rnh1 rnh201*	<0.0001	<0.0001	0.0003	0.0003
WT *vs. rnh1 rnh201 spt3*	<0.0001	<0.0001	0.0003	0.0003
WT *vs. dbr1*	<0.0001	<0.0001	0.0003	0.0003
WT *vs. rnh1 dbr1*	0.0009	0.0002	0.0018	0.0005
WT *vs. rnh201 dbr1*	0.0009	0.0003	0.0018	0.0007
WT *vs. rnh1 rnh201 dbr1*	0.5570	<0.0001	0.6210	0.0003
spt3 vs. rnh1 spt3	0.5970	NA	0.6561	NA
spt3 vs. rnh201 spt3	0.5419	NA	0.6079	NA
spt3 vs. rnh1 rnh201 spt3	<0.0001	<0.0001	0.0003	0.0003
spt3 vs. dbr1	<0.0001	0.9590	0.0003	0.9716
spt3 vs. rnh1 dbr1	0.0004	0.9590	0.0010	0.9716
spt3 vs. rnh201 dbr1	0.0004	0.0003	0.0010	0.0007
spt3 vs. rnh1 rnh201 dbr1	<0.0001	<0.0001	0.0003	0.0003
rnh1 vs. rnh201	0.0002	<0.0001	0.0005	0.0003
rnh1 vs. spt3	<0.0001	<0.0001	0.0003	0.0003
rnh1 vs. rnh1 spt3	<0.0001	<0.0001	0.0003	0.0003
rnh1 vs. rnh1 rnh201	<0.0001	<0.0001	0.0003	0.0003
rnh1 vs. rnh1 rnh201 spt3	<0.0001	<0.0001	0.0003	0.0003
rnh1 vs. dbr1	<0.0001	0.0009	0.0003	0.0018
rnh1 vs. rnh1 dbr1	0.0009	0.0009	0.0018	0.0018
rnh1 vs. rnh1 rnh201 dbr1	<0.0001	<0.0001	0.0003	0.0003
rnh201 vs. spt3	<0.0001	<0.0001	0.0003	0.0003
rnh201 vs. rnh201 spt3	<0.0001	<0.0001	0.0003	0.0003
rnh201 vs. rnh1 rnh201	0.0002	0.0002	0.0005	0.0005
rnh201 vs. rnh1 rnh201 spt3	<0.0001	<0.0001	0.0003	0.0003

rnh201 vs. dbr1	<0.0001	0.0009	0.0003	0.0018
rnh201 vs. rnh201 dbr1	0.0009	0.0009	0.0018	0.0018
rnh201 vs. rnh1 rnh201 dbr1	<0.0001	<0.0001	0.0003	0.0003
rnh1 rnh201 vs. rnh1 rnh201 spt3	<0.0001	<0.0001	0.0003	0.0003
rnh1 rnh201 vs. rnh1 rnh201 dbr1	<0.0001	<0.0001	0.0003	0.0003
dbr1 vs. rnh1 dbr1	0.0150	0.8640	0.0211	0.8987
dbr1 vs. rnh201 dbr1	0.0492	0.0192	0.0613	0.0265
dbr1 vs. rnh1 rnh201 dbr1	<0.0001	0.0002	0.0003	0.0005
WT *vs.* WT + PFA	0.0044	0.0017	0.0068	0.0033
WT *vs. rnh1 rnh201* + PFA	0.0249	0.0017	0.0339	0.0033
rnh1 rnh201 vs. rnh1 rnh201 + PFA	0.0043	0.0126	0.0067	0.0179
WT + PFA *vs. rnh1 rnh201* + PFA	0.0286	0.0286	0.0362	0.0362

II)

Genotype	P values trans vs. cis	Adj. P values trans vs. cis
WT	0.0002	0.0005
spt3	NA	NA
rnh1	<0.0001	0.0003
rnh201	<0.0001	0.0003
rnh1 rnh201	0.2921	0.3336
rnh1 rnh201 spt3	<0.0001	0.0003
dbr1	<0.0001	0.0003
rnh1 rnh201 dbr1	<0.0001	0.0003
WT + PFA	0.0571	0.0704
rnh1 rnh201 + PFA	0.0286	0.0362

Mann-Whitney *U*-test was applied to determine whether a statistical significant difference exists between pairs of gene correction frequencies or rates obtained in DSB repair or transposition assays or fold-change values obtained in the qRT-PCR experiment. All *P* values obtained using the Mann-Whitney *U*-test were then adjusted by applying the false discovery rate (FDR) method to correct for multiple hypothesis. **a**, Comparison of frequencies presented in Table 1a and Extended Data Table 5a. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05. I) Comparisons were between relative frequencies obtained in the *trans* or *cis* assay in different backgrounds, and II) between relative frequencies obtained in the *trans* and *cis* assays for each background. NA, not applicable because the frequencies were too low in both samples to allow meaningful comparison.

b

Genotype	*P* values *cis*	Adj. *P* values *cis*
WT *vs. rad51*	0.0011	0.0022
WT *vs. rnh1 rnh201*	0.0011	0.0022
WT *vs. rnh1 rnh201 spt3*	<0.0001	0.0003
WT *vs. rnh1 rnh201 rad52*	0.5427	0.6079
WT *vs. rnh1 rnh201 rad51*	0.0011	0.0022
WT *vs. rnh1 rnh201 spt3 rad52*	0.0277	0.0362
WT *vs. rnh1 rnh201 spt3 rad51*	0.0022	0.0039
rad51 vs. rnh1 rnh201	0.0022	0.0039
rad51 vs. rnh1 rnh201 spt3	0.6733	0.7331
rad51 vs. rnh1 rnh201 spt3 rad51	0.0043	0.0067
rnh1 rnh201 vs. rnh1 rnh201 spt3	0.0009	0.0018
rnh1 rnh201 vs. rnh1 rnh201 rad52	0.0009	0.0018
rnh1 rnh201 vs. rnh1 rnh201 rad51	1.0000	1.0000
rnh1 rnh201 vs. rnh1 rnh201 spt3 rad52	0.0009	0.0018
rnh1 rnh201 vs. rnh1 rnh201 spt3 rad51	0.0043	0.0067
rnh1 rnh201 spt3 vs. rnh1 rnh201 rad52	0.0001	0.0003
rnh1 rnh201 spt3 vs. rnh1 rnh201 rad51	0.0009	0.0018
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 rad52	<0.0001	0.0003
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 rad51	0.0037	0.0063
rnh1 rnh201 rad52 vs. rnh1 rnh201 spt3 rad52	0.0119	0.0170
rnh1 rnh201 rad51 vs. rnh1 rnh201 spt3 rad51	0.0043	0.0067

b, Comparison of frequencies presented in Table 1b. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05. I) Comparisons were between relative frequencies obtained in the *trans* or *cis* assay in different backgrounds, and II) between relative frequencies obtained in the *trans* and *cis* assays for each background.

c

Yeast	*P* values	Adj. *P* values
ssDNA *vs.* ssDNA+Rad52	0.0286	0.0362
ssDNA *vs.* ssDNA+Rad52+RPA	0.0286	0.0362
ssRNA *vs.* ssRNA+Rad52	0.0286	0.0362
ssRNA *vs.* ssRNA+Rad52+RPA	0.0286	0.0362
ssDNA+Rad52 *vs.* ssDNA+Rad52+RPA	0.0286	0.0362
ssRNA+Rad52 *vs.* ssRNA+Rad52+RPA	0.6857	0.7428
ssDNA *vs.* ssRNA	0.0286	0.0362
ssDNA+Rad52 *vs.* ssRNA+Rad52	0.0286	0.0362
ssDNA+Rad52+RPA *vs.* ssRNA+Rad52+RPA	0.0286	0.0362
Human	*P* values	Adj. *P* values
ssDNA *vs.* ssDNA+RAD52	0.0286	0.0362
ssDNA *vs.* ssDNA+RAD52+RPA	0.0286	0.0362
ssRNA *vs.* ssRNA+RAD52	0.0286	0.0362
ssRNA *vs.* ssRNA+RAD52+RPA	0.0286	0.0362
ssDNA+RAD52 *vs.* ssDNA+RAD52+RPA	0.0286	0.0362
ssRNA+RAD52 *vs.* ssRNA+RAD52+RPA	0.4857	0.5449
ssDNA *vs.* ssRNA	0.0286	0.0362
ssDNA+RAD52 *vs.* ssRNA+RAD52	0.0571	0.0704
ssDNA+RAD52+RPA *vs.* ssRNA+RAD52+RPA	0.0286	0.0362

c, Comparison of kinetics using yeast and human Rad52 (data obtained at 10 and 15 min were used) presented in Fig. 2c and d. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05.

133

	P values				Adj. P values			
	trans		cis		trans	cis	trans	cis
Genotype	No oligo	HIS3.F+R	No oligo	HIS3.F+R	No oligo	HIS3.F+R	No oligo	HIS3.F+R
WT *vs. rnh1 rnh201*	0.1913	0.0283	<0.0001	0.2141	0.2240	0.0362	0.0003	0.2482
WT *vs. rnh1 rnh201 spt3*	0.8846	0.0040	<0.0001	0.0040	0.9161	0.0064	0.0003	0.0064
rnh1 rnh201 vs. rnh1 rnh201 spt3	0.0286	0.0286	0.0286	0.0286	0.0362	0.0362	0.0362	0.0362
rnh1 rnh201 vs. rnh1 rnh201 spt3 AlΔ23	NA	NA	NA	0.0040	NA	NA	NA	0.0064

II)

	P values		Adj. P values	
	trans vs. cis		trans vs. cis	
Genotype	No Oligo	HIS3.F+R	No Oligo	HIS3.F+R
WT	NA	0.9591	NA	0.9716
rnh1 rnh201	0.0286	1.0000	0.0362	1.0000
rnh1 rnh201 spt3	0.0084	0.1091	0.0370	0.7397

d, Comparison of frequencies presented in Extended Data Table 3. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05. I) Comparisons were between relative frequencies obtained in the *trans* or *cis* assay in different backgrounds, and II) between relative frequencies obtained in the *trans* and *cis* assays for each background. NA, not applicable because the frequencies were too low in both samples to allow meaningful comparison.

e

Genotype	P values		Adj. P values	
	cis		*cis*	
	Galactose	Glucose	Galactose	Glucose
WT + BDG283 *vs.* WT + BDG606	0.0022	NA	0.0039	NA
WT + BDG283 *vs. rnh1 rnh201 spt3* + BDG283	0.0022	<0.0001	0.0039	0.0003
WT + BDG606 *vs. rnh1 rnh201 spt3* + BDG606	0.0022	<0.0001	0.0039	0.0003
rnh1 rnh201 spt3 + BDG283 *vs. rnh1 rnh201 spt3* + BDG606	0.9372	0.6631	0.9620	0.7254

e, Comparison of frequencies presented in Extended Data Table 4. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05. Comparison was between relative frequencies obtained in the *cis* assay in different backgrounds. NA, not applicable because data were not available for comparison.

f

Genotype	P values	Adj. P values
	cis	*cis*
WT *vs. pGAL1Δ*	<0.0001	0.0003
rnh1 rnh201 vs. rnh1 rnh201 pGAL1Δ	0.0022	0.0039
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 pGAL1Δ	0.0022	0.0039

f, Comparison of frequencies presented in Extended Data Table 5b. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05. I) Comparison was between relative frequencies obtained in the *cis* assay in different backgrounds.

Genotype	P values		Adj. P values	
	trans	*cis*	*trans*	*cis*
WT *vs. spt3*	NA	NA	NA	NA
WT *vs. rnh1 rnh201*	<0.0001	<0.0001	0.0003	0.0003
WT *vs. rnh1 rnh201 spt3*	<0.0001	<0.0001	0.0003	0.0003
WT *vs. dbr1*	NA	NA	NA	NA
WT *vs. rad52*	NA	NA	NA	NA
WT *vs. rad51*	NA	NA	NA	NA
WT *vs. rnh1 rnh201 dbr1*	NA	<0.0001	NA	0.0003
spt3 vs. rnh1 rnh201	<0.0001	<0.0001	0.0003	0.0003
spt3 vs. rnh1 rnh201 spt3	<0.0001	<0.0001	0.0003	0.0003
spt3 vs. dbr1	NA	NA	NA	NA
spt3 vs. rnh1 rnh201 dbr1	NA	NA	NA	NA
dbr1 vs. rnh1 rnh201	<0.0001	<0.0001	0.0003	0.0003
dbr1 vs. rnh1 rnh201 dbr1	NA	NA	NA	NA
rnh1 rnh201 vs. rnh1 rnh201 spt3	0.1079	0.0058	0.1283	0.0086
rnh1 rnh201 vs. rnh1 rnh201 dbr1	<0.0001	0.0058	0.0003	0.0086
rnh1 rnh201 vs. rad52	NA	<0.0001	NA	0.0003
rnh1 rnh201 vs. rnh1 rnh201 rad52	NA	0.8318	NA	0.8691
rnh1 rnh201 vs. rnh1 rnh201 spt3 rad52	NA	0.0058	NA	0.0086
rnh1 rnh201 vs. rad51	NA	NA	NA	NA
rnh1 rnh201 vs. rnh1 rnh201 rad51	NA	0.0058	NA	0.0086
rnh1 rnh201 vs. rnh1 rnh201 spt3 rad51	NA	0.0058	NA	0.0086
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 rad52	NA	0.0591	NA	0.0721
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 rad51	NA	0.0591	NA	0.0721

II)

Genotype	P values trans vs. cis	Adj. P values trans vs. cis
WT	NA	NA
spt3	NA	NA
dbr1	NA	NA
rnh1 rnh201	0.0727	0.0873
rnh1 rnh201 spt3	0.7986	0.8420
rnh1 rnh201 dbr1	NA	NA

g, Comparison of frequencies presented in Extended Data Table 5c. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05. I) Comparisons were between relative frequencies obtained in the *trans* or *cis* assay in different backgrounds, and II) between relative frequencies obtained in the *trans* and *cis* assays for each background. NA, not applicable because the frequencies were too low in both samples to allow meaningful comparison, or because data were not available for comparison.

Genotype	P values	Adj. P values
	cis	*cis*
WT *vs. rnh1 rnh201*	0.0050	0.0076
WT *vs. rnh1 rnh201 spt3*	0.0050	0.0076
WT *vs. AI∆23*	<0.0001	0.0003
rnh1 rnh201 vs. rnh1 rnh201 spt3	0.0022	0.0039
rnh1 rnh201 vs. rnh1 rnh201 AI∆23	<0.0001	0.0003
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 AI∆23	<0.0001	0.0003
AI∆23 vs. rnh1 rnh201 AI∆23	NA	NA
AI∆23 vs. rnh1 rnh201 spt3 AI∆23	NA	NA
rnh1 rnh201 AI∆23 vs. rnh1 rnh201 spt3 AI∆23	NA	NA

h, Comparison of frequencies presented in Extended Data Table 5d. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05. Comparisons were between relative frequencies obtained in the *cis* assays for each background shown. NA, not applicable because the frequencies were too low in both samples to allow meaningful comparison.

i

Genotype	P values		Adj. P values	
	No galactose	Galactose	No galactose	Galactose
WT *vs. rnh201 rnh1* (22 °C) (Ura⁻)	0.0021	<0.0001	0.0039	0.0003
WT *vs. rnh201 rnh1* (30 °C) (Ura⁻)	0.0250	0.0022	0.0339	0.0039
WT *vs. rnh201 rnh1* (30 °C) (YPLac)	0.0050	0.0152	0.0076	0.0212

i, Comparison of rates presented in Extended Data Table 6. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05. Comparisons were between rates of His⁺ colonies obtained in the WT *vs. rnh201 rnh1* strains containing BDG598.

j
I)

Genotype	P values		Adj. P values	
	trans	*cis*	*trans*	*cis*
	0.25 h *vs.* 8 h	0.25 h *vs.* 8 h	0.25 h *vs.* 8 h	0.25 h *vs.* 8 h
WT	0.7984	0.0002	0.8420	0.0005
spt3	0.9717	0.0022	0.9801	0.0039
rnh1 rnh201	0.0571	0.0286	0.0704	0.0362
rnh1 rnh201 spt3	0.7209	0.0002	0.7741	0.0005
rnh1 rnh201 spt3 pGAL1Δ	NA	0.0003	NA	0.0007
rnh1 rnh201 spt3 hoΔ	NA	0.0030	NA	0.0053

II)

Genotype	P values				Adj. P values			
	trans		*cis*		*trans*		*cis*	
	0.25 h	8 h	0.25 h	8 h	0.25 h	8 h	0.25 h	8 h
WT *vs. spt3*	0.0037	0.2318	0.0037	0.0200	0.0063	0.2674	0.0063	0.0274
WT *vs. rnh1 rnh201*	0.0485	0.0040	0.6828	0.5697	0.0608	0.0064	0.7397	0.6321
WT *vs. rnh1 rnh201 spt3*	0.7984	0.5737	0.1049	0.0019	0.8420	0.6335	0.1253	0.0037
WT *vs. rnh1 rnh201 spt3 pGAL1Δ*	NA	NA	0.0002	0.0006	NA	NA	0.0005	0.0014
WT *vs. rnh1 rnh201 spt3 hoΔ*	NA	NA	0.0650	0.0002	NA	NA	0.0785	0.0005
spt3 vs. rnh1 rnh201	0.8252	0.0180	0.0095	0.0095	0.8661	0.0250	0.0138	0.0138
spt3 vs. rnh1 rnh201 spt3	0.1388	0.9079	0.0027	0.7546	0.1642	0.9360	0.0048	0.8065
spt3 vs. rnh1 rnh201 spt3 pGAL1Δ	NA	NA	0.0007	0.0127	NA	NA	0.0016	0.0179
spt3 vs. rnh1 rnh201 spt3 hoΔ	NA	NA	0.0007	0.0007	NA	NA	0.0016	0.0016
rnh1 rnh201 vs. rnh1 rnh201 spt3	0.3677	0.0283	0.0007	0.0040	0.4179	0.0362	0.0016	0.0064
rnh1 rnh201 vs. rnh1 rnh201 spt3 pGAL1Δ	NA	NA	0.0040	0.0040	NA	NA	0.0064	0.0064
rnh1 rnh201 vs. rnh1 rnh201 spt3 hoΔ	NA	NA	0.1535	0.0040	NA	NA	0.1806	0.0064
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 pGAL1Δ	NA	NA	0.0002	0.0104	NA	NA	0.0005	0.0150
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 hoΔ	NA	NA	0.1949	0.0002	NA	NA	0.2271	0.0005
rnh1 rnh201 spt3 pGAL1Δ vs. rnh1 rnh201 spt3 hoΔ	NA	NA	0.0650	0.0002	NA	NA	0.0785	0.0005

j, Comparison of fold-change values presented in Extended Data Figure 2d. Two groups in a pair were considered to be significantly different when adjusted *P* values were less than 0.05. I) Comparisons were between fold-change values obtained at time point 0.25 h and 8 h for the *trans* or *cis* backgrounds, II) between fold-change values obtained in the different

139

backgrounds of the *trans* or *cis* system either at time point 0.25 h or 8 h, and III) between

fold-change values obtained in the *trans* and *cis* systems for each background either at 0.25

h or 8 h. NA, not applicable because data were not available for comparison.

a

trans system on Chr III

HIS3 STOP

[CTA]CATAAGAACACCTTTGGTGGAGGGAACATCGTTGGTACCATTGGGCGAGGTGGCTTCTCTTATGGCAACC

GCAAGAGCCTTGAACGCACTCTCACTACGGTGATGATCATTCTTGCCTCGCAGACAATCAACGTGGAGGGTAA

TTCTGCTAGCCTCTGCAAAGCTTTCAAGAAAATGCGGGATCATCTCGCAAGAGAGATCTCCTACTTTCTCCCTTT

GCAAACCAAGTTCGACAACTGCGTACGGCCTGTTCGAAAGATCTACCACCGCTCTGGAAAGTGCCTCATCCAA

AGGCGCAAATCCTGATCCAAACCTTTTTACTCCACGCACGGCCCCTAGGGCCTCTTTAAAAGCTTGACCGAGAG

CAATCCCGCAGTCTTCAGTGGTGTGATGGTCGTCTATGTGTAAGTCACCAATGCACTCAACGATTAGCGACCAG

CCGGAATGCTTGG[GTATGT]TAATATGGACTAAAGGAGGCTTTTCTGCAGGTCGACTCTAGAGGATCCCCGGGT

ACCGAGCTCGAATTTT[TACTAAC]AAATGGTATTATTTATAAC[AG]CCAGAGCATGTATCATATGGTCCAGAAACC

CTATACCTGTGTGGACGTTAATCACTTGCGATTGTGTGGCCTGTTCTGCTACTGCTTCTGCCTCTTTTTCTGGGA

AGATCGAGTGCTCTATCGCTAGGGGACCACCCTTTAAAGAGATCGCAATCTGAATCTTGGTTTCATTTGTAATA

CGCTTTACTAGGGCTTTCTGCTCTGT[CAT] *HIS3* ATG

b

trans system on Chr XV

HIS3 ATG

[ATG]ACAGAGCAGAAAGCCCTAGTAAAGCGTATTACAAATGAAACCAAGATTCAGATTGCGATCTCTTTAAAGG

GTGGTCCCCTAGCGATAGAGCACTCGATCTTCCCAGAAAAAGAGGCAGAAGCAGTAGCAGAACAGGCCACAC

AATCGCAAGTGATTAACGTCCACACAGGTATAGGGTTTCTGGACCATATGATACATGCTCTGGC[GGTACGGGG

ATCTAAATAAATTCGTTTTCAATGATTAAAATAGCATAGTCGGGTTTTTCTTTTAGTTTCAGCTTTCCGCAACAGT

ATAATTTTATAAACCCTGGTTTTGGTTTTGTAGAGTGGTTGTAC]AAGCATTCCGGCTGGTCGCTAATCGTTGAG

HO SITE

TGCATTGGTGACTTACACATAGACGACCATCACACCACTGAAGACTGCGGGATTGCTCTCGGTCAAGCTTTTAA

AGAGGCCCTAGGGGCCGTGCGTGGAGTAAAAAGGTTTGGATCAGGATTTGCGCCTTTGGATGAGGCACTTTC

CAGAGCGGTGGTAGATCTTTCGAACAGGCCGTACGCAGTTGTCGAACTTGGTTTGCAAAGGGAGAAAGTAGG

AGATCTCTCTTGCGAGATGATCCCGCATTTTCTTGAAAGCTTTGCAGAGGCTAGCAGAATTACCCTCCACGTTG

ATTGTCTGCGAGGCAAGAATGATCATCACCGTAGTGAGAGTGCGTTCAAGGCTCTTGCGGTTGCCATAAGAGA

AGCCACCTCGCCCAATGGTACCAACGATGTTCCCTCCACCAAAGGTGTTCTTATG[TAG] *HIS3* STOP

c

cis system on Chr III

HIS3 STOP

[CTA]CATAAGAACACCTTTGGTGGAGGGAACATCGTTGGTACCATTGGGCGAGGTGGCTTCTCTTATGGCAACC

GCAAGAGCCTTGAACGCACTCTCACTACGGTGATGATCATTCTTGCCTCGCAGACAATCAACGTGGAGGGTAA

TTCTGCTAGCCTCTGCAAAGCTTTCAAGAAAATGCGGGATCATCTCGCAAGAGAGATCTCCTACTTTCTCCCTTT

GCAAACCAAGTTCGACAACTGCGTACGGCCTGTTCGAAAGATCTACCACCGCTCTGGAAAGTGCCTCATCCAA

AGGCGCAAATCCTGATCCAAACCTTTTTACTCCACGCACGGCCCCTAGGGCCTCTTTAAAAGCTTGACCGAGAG

CAATCCCGCAGTCTTCAGTGGTGTGATGGTCGTCTATGTGTAAGTCACCAATGCACTCAACGATTAGCGACCAG

*deletion

CCGGAATGCTTGG[GTATGT]TAATATGGACTAAAGGAGGCTTTTCTGCAGGTCGACTCTAG[AACCACTCTACAA

AACCAAAACCAGGGTTTATAAAATTATACTGTTGCGGAAAGCTGAAACTAAAAGAAAAACCCGACTATGCTAT

HO SITE

TTTAATCATTGAAAACGAATTTATTTAGATCCCCGTAC]AGGATCCCCCGGGTACCGAGCTCGAATTTT[TACTAAC

AAATGGTATTATTTATAAC[AG]CCAGAGCATGTATCATATGGTCCAGAAACCCTATACCTGTGTGGACGTTAATC

ACTTGCGATTGTGTGGCCTGTTCTGCTACTGCTTCTGCCTCTTTTTCTGGGAAGATCGAGTGCTCTATCGCTAGG

GGACCACCCTTTAAAGAGATCGCAATCTGAATCTTGGTTTCATTTGTAATACGCTTTACTAGGGCTTTCTGCTCT
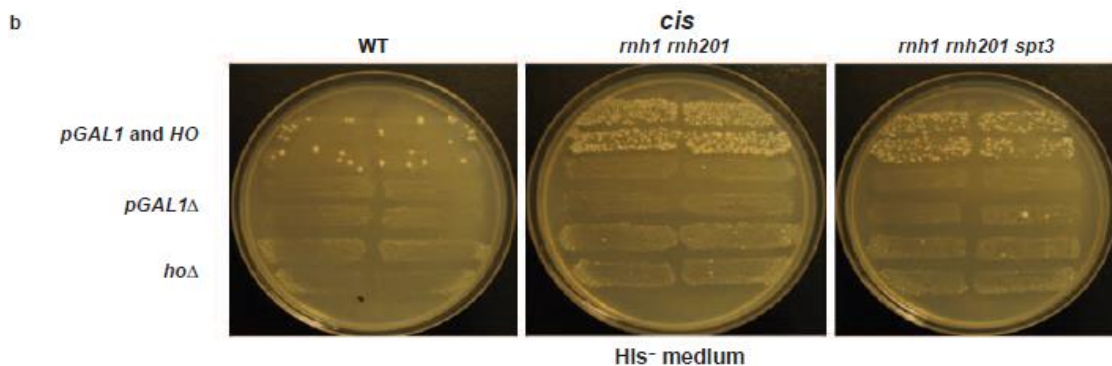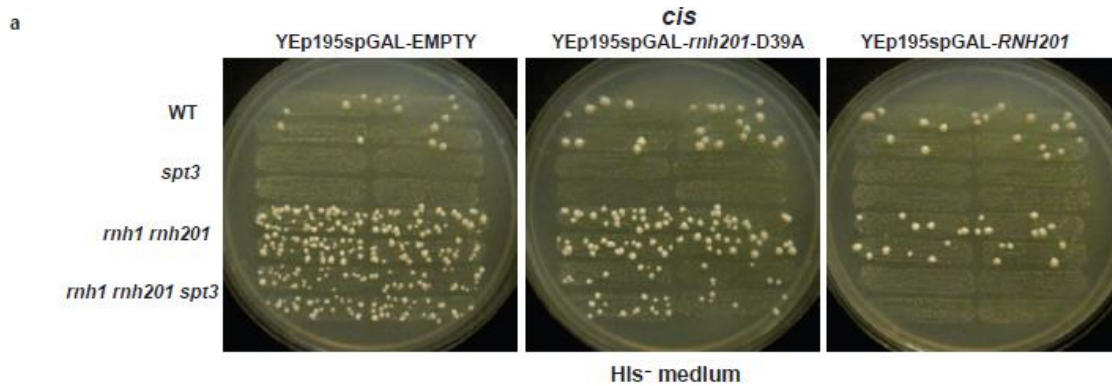
GT[CAT] *HIS3* ATG

Figure A.1 DNA sequence of the *his3* loci in the *trans* and *cis* systems. a, Trans system on chromosome (Chr) III. *HIS3* ATG and STOP codons are boxed. The *HIS3* gene is disrupted by an insert (orange) carrying the artificial intron (AI). The consensus sequences of the AI are boxed. **b**, Trans system on chromosome XV. *HIS3* ATG and STOP codons are boxed. The *HIS3* gene is disrupted by an insert (yellow) containing the 124-base-pair homothallic switching endonuclease site (marked by lines). **c**, *Cis* system on chromosome III. *HIS3* ATG and STOP codons are shown. The *HIS3* gene is disrupted by an insert (orange) carrying the AI, which contains the 124 base pairs of the homothallic switching endonuclease site (yellow and marked by lines). The consensus sequences of the AI are boxed. *23-base-pair deletion of the AI, including the 59 splice site, made in some strains.

a

cis

YEp195spGAL-EMPTY   YEp195spGAL-*rnh201*-D39A   YEp195spGAL-*RNH201*

WT
spt3
rnh1 rnh201
rnh1 rnh201 spt3

His⁻ medium

b

cis

WT   *rnh1 rnh201*   *rnh1 rnh201 spt3*

pGAL1 and *HO*
pGAL1Δ
hoΔ

His⁻ medium

c

cis	0 h GAL			8 h GAL		
Genotype	G1	S	G2	G1	S	G2
WT	75	13	12	24	11	65
rnh1 rnh201	74	14	12	23	11	66
rnh1 rnh201 spt3	66	5	29	35	7	58
pGAL1Δ	83	7	10	27	7	66
rnh1 rnh201 pGAL1Δ	79	6	15	31	6	63
rnh1 rnh201 spt3 pGAL1Δ	73	5	22	28	7	65
hoΔ	90	5	5	69	14	16
rnh1 rnh201 hoΔ	91	4	5	66	19	15
rnh1 rnh201 spt3 hoΔ	70	9	21	59	14	27

d

143
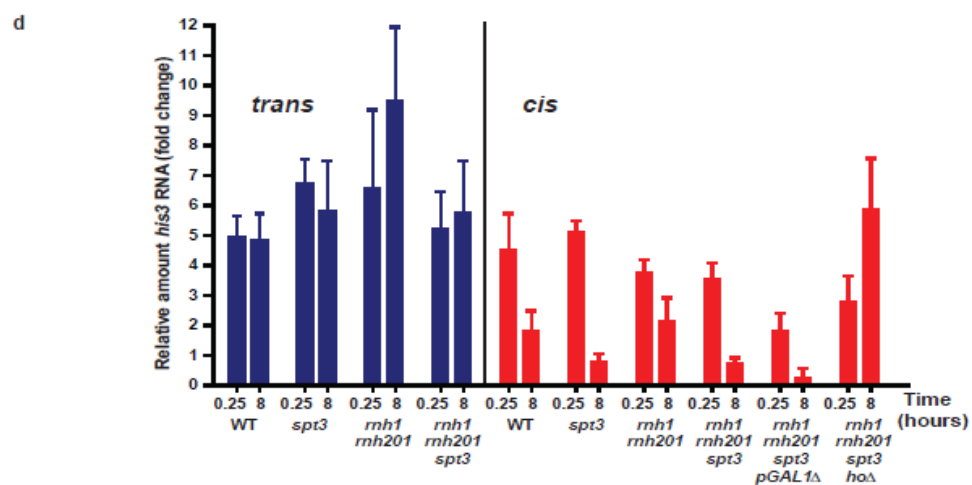
Figure A.2 Efficient transcript-RNA-directed gene modification is inhibited by *RNH201*, requires transcription of the template RNA and formation of a DSB in the target gene. a, Complementation of *rnh201* defect suppresses transcript-RNA-templated DSB repair in *cis*-system *rnh1 rnh201 spt3* cells. Wild-type (WT), *spt3*, *rnh1 rnh201*, *rnh1 rnh201 spt3* strains of the *cis* system were transformed by a control empty vector (YEp195spGAL-EMPTY), a vector expressing catalytically inactive from of RNase H2 (YEp195spGAL-*rnh201*-D39A) or a wild-type form of RNase H2 (YEp195spGAL-*RNH201*). All the vectors have the galactose-inducible promoter. Shown is an example of replica-plating results (n=6) from galactose medium to histidine dropout for the indicated strains and plasmids. **b**, Example of replica-plating results (n=6) from galactose medium to histidine dropout for the indicated strains of the *cis* system, which have functional *pGAL1* promoter and homothallic switching endonuclease (HO) gene, or have deleted *pGAL1* promoter (*pGAL1Δ*), or deleted HO gene (*hoΔ*). **c**, Table with percentages of cells in the G1, S or G2 stage of the cell cycle out of 200 random cells counted for the indicated strains of the cis system after 0 h and 8 h from galactose induction. If a homothallic switching endonuclease DSB is made in *his3*, yeast cells arrest in G2, thus a high percentage of G2-arrested cells indicates occurrence of the homothallic switching endonuclease DSB. We also note that strains with *spt3* mutation have a higher percentage of G2 cells than strains with wild-type *SPT3* before DSB induction (0 h GAL). **d**, Results of qPCR of *his3* RNA. Cells were grown in YPLac liquid medium O/N, and were collected and prepared for qPCR at 0, 0.25 or 8 h after adding galactose to the medium. *Trans*, blue bars; *cis*, red bars. Data are represented as a fold change value with respect to mRNA expression at time zero, as median and range of 6–8 repeats. The significance of

comparisons between fold changes obtained at 0.25 h versus those obtained at 8 h, fold

changes of different strains of the *trans* and *cis* systems, and between fold changes

obtained in the *trans* versus *cis* system for the same strains at the same time point was

calculated using the Mann–Whitney U-test and P values are presented in **Table A.7 jI, II**

and III, respectively. We note that an apparent higher level of *his3* RNA is detected at 8

h in galactose in both *trans-* and *cis*-system *rnh1 rnh201* cells relative to the other tested

genetic backgrounds. Our interpretation of these results is that *his3* RNA could be more

stable in *rnh1 rnh201* cells if present in the form of RNA–DNA heteroduplexes, and this

may explain the increased frequency of His[+] colonies observed in both *trans* and *cis* in

the *rnh1 rnh201* cells (**Figure 2.1c and Table 2.1a**).
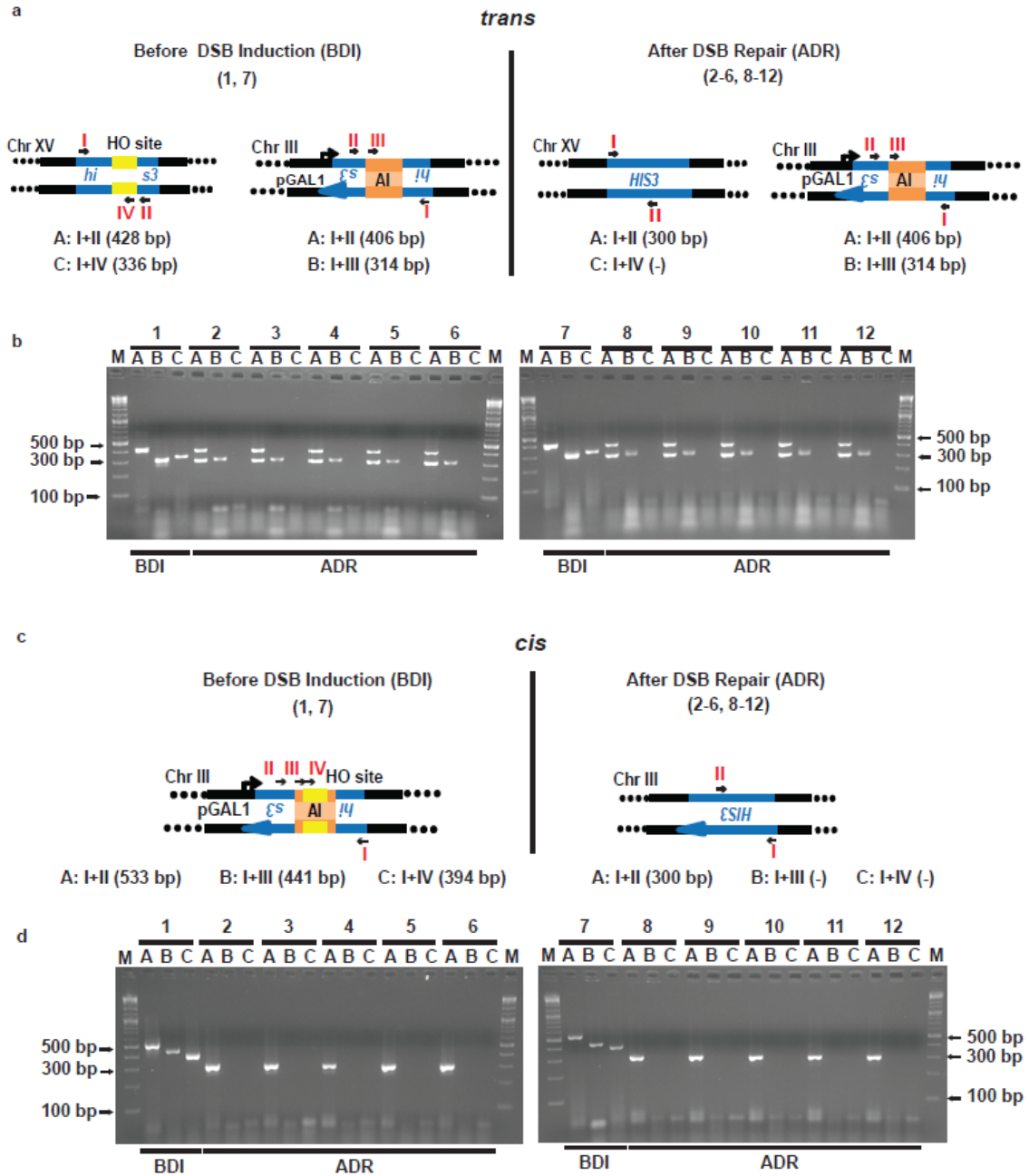
Figure A.3 Verification of *his3* repair in *trans*- and *cis*-system *rnh1 rnh201 spt3* cells via a homologous recombination mechanism using colony PCR. a, Scheme of the *trans* system before DSB induction (BDI, groups of lanes 1 and 7) and after DSB repair (ADR, groups of lanes 2–6 and 8–12) with the primers used in colony PCR shown as

small black arrows and named with roman numerals: I, HIS3.5; II, HIS3.2; III, INTRON.F; IV, HO.F. The primer pairs used for colony PCR are named A (I 1 II), B (I 1 III) and C (I1 IV), and base-pair sizes of the expected PCR products are shown in brackets. **b**, Photos of agarose gels with results of colony PCR reactions. M, 2-log DNA ladder marker; the 100-, 300- and 500-base-pair band sizes are indicated by arrows. Groups of lanes 1 and 7, two isolates of trans-system *rnh1 rnh201 spt3* mutants before DSB induction, each tested with primer pairs A, B and C. Groups of lanes 2–6 and 8–12, ten isolates of trans-system *rnh1 rnh201 spt3* mutants after DSB repair, each tested with primer pairs A, B and C. **c**, Scheme of the *cis* system before DSB induction (BDI, groups of lanes 1and 7) and after DSB repair (ADR, groups of lanes 2–6 and 8–12) with the primers used in colony PCR shown as small black arrows and named with roman numerals: I, HIS3.5; II, HIS3.2; III, INTRON.F; IV, HO.F. The primer pairs used for colony PCR are named A (I 1 II), B (I 1 III) and C (I 1 IV), and base-pair sizes of the expected PCR products are shown in brackets. **d**, Photos of agarose gels with results of colony PCR reactions. M, 2-log DNA ladder marker; the 100-, 300- and 500-base-pair band sizes are indicated by arrows. Groups of lanes 1 and 7, two isolates of *cis* system *rnh1 rnh201 spt3* mutants before DSB induction, each tested with primer pairs A, B and C. Groups of lanes 2–6 and 8–12, ten isolates of *cis*-system *rnh1 rnh201 spt3* mutants after DSB repair, each tested with primer pairs A, B and C.
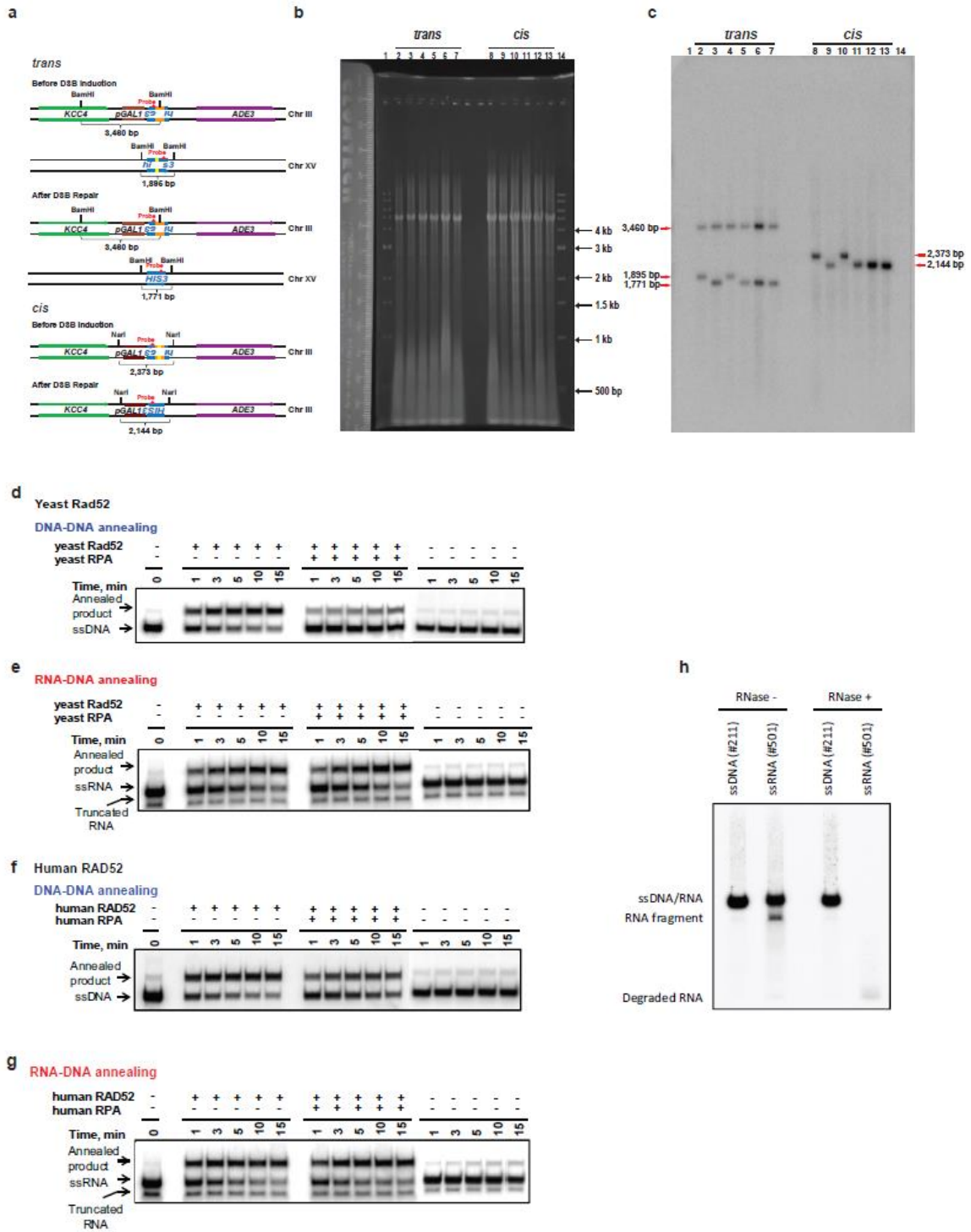
Figure A.4 RNA-templated DNA repair occurs via homologous recombination and requires Rad52. a, Scheme of the *trans* and *cis his3/HIS3* loci in His⁻ (before DSB induction) and His⁺ (after DSB repair) cells. The size of the BamHI (*trans*) or NarI (*cis*)

restriction digestion products and the position of the *HIS3* probe are shown. **b**, Photo of a ruler next to ethidium-bromide-stained agarose gel with marker and genomic DNA samples visible before Southern blot analysis. Lanes 1 and 14, 1-kilobase (kb) DNA ladder; 500-base-pair, 1-kb, 1.5-kb, 2-kb, 3-kb and 4-kb bands are indicated by arrows. *Trans* wild-type His⁻ (lane 2) or His⁺ (lane 3), *rnh1 rnh201 spt3* His⁻ (lane 4) or His⁺ (lanes 5–7) cells, digested with BamHI restriction enzyme. *Cis* wild-type His⁻ (lane 8) or His⁺ (lane 9), *rnh1 rnh201 spt3* His⁻ (lane 10) or His⁺ (lanes 11–13) cells, digested with NarI restriction enzyme. **c**, Southern blot analysis (same as in **Figure 2.2a**, but displaying the entire picture of the exposed membrane) of yeast genomic DNA derived from trans wild-type His⁻ (lane 2) or His⁺ (lane 3), *rnh1 rnh201 spt3* His⁻ (lane 4) or His⁺ (lanes 5–7) cells, digested with BamHI restriction enzyme and hybridized with the *HIS3* probe, or derived from *cis* wild-type His⁻ (lane 8) or His⁺ (lane 9), *rnh1 rnh201 spt3* His⁻ (lane 10) or His⁺ (lanes 11–13) cells, digested with NarI restriction enzyme and hybridized with the *HIS3* probe. Lanes 1 and 14, 1-kb DNA ladder visible in the ethidium-bromide-stained gel (**b**). Sizes of digested DNA bands are indicated. The annealing reactions were promoted by either yeast Rad52 (**d, e**) or human RAD52 (f, g) (1.35 nM) in the presence or absence of RPA (2 nM) (yeast or human RPA was used in the reaction with yeast or human Rad52, respectively). In control protein-free reactions, protein dilution buffers were added instead of the respective proteins. dsDNA containing a protruding ssDNA tail (no. 508 and no. 509) was incubated with RPA (when indicated) and then Rad52 was added to the mixture. To initiate the annealing reactions, 0.3nM 32P-labelled ssDNA (no. 211) or ssRNA (no. 501) were added. The reactions were carried out for the indicated periods of time, and the products of annealing reactions were deproteinized and analysed

149

by electrophoresis in 10% polyacrylamide gels in 13 TBE at 150V for 1 h. Visualization and quantification was accomplished using a Storm 840 Phosphorimager and ImageQuant 5.2 software (GE Healthcare). **e**, Treatment of RNA and DNA oligonucleotides with RNase.

ssDNA (no. 211) or RNA (no. 501) (3 mM) was incubated with 100 µg ml^{-1} (or 7Uml^{-1}) RNase (Qiagen) in buffer containing 50 mM Hepes, pH7.5 for 30 min at 37 °C, then 7% glycerol and 0.1% bromophenol blue were added to the samples and incubation continued for another 15 min at 37 °C before the samples were analysed by electrophoresis in a 10% (17:1 acrylamide:bisacrylamide) polyacrylamide gel at 150V for 1 h in 13 TBE buffer. The gel was quantified using a Storm 840 Phosphorimager. The RNA oligonucleotide, but not the DNA oligonucleotide, is completely degraded by RNase.

APPENDIX B

SUPPLEMENTARY MATERIALS FOR CHAPTER 3

Table B.1 Yeast strains used in this study

Strain	Relevant genotype	Source
YS-291, 292	*hoΔ hmlΔ::ADE1 mataΔ::hisG hmrΔ::ADE1 ade1 leu2::pGAL1mhis3AI::HOcs-ADE3 lys5 trp1::hisG ura3-52 ade3::GAL::HO (his3::HOcs)::TRP1 YCLWTy2-1Δ*	[29]
YS-416, 417	YS-291, 292 *rnh1Δ::kanMX4*	[29]
YS-412, 413	YS-291, 292 *rnh201Δ::hygMX4*	[29]
YS-424, 426	YS-291, 292 *rnh1Δ::NAT rnh201Δ::hygMX4*	[29]
HK-203, 205	YS-291, 292 *rnh201*-G42S	this study
HK-209, 211	YS-291, 292 *rnh203*-K46W	this study
HK-236, 238	HK-203, 205 *rnh201*-G42S *rnh1Δ::kanMX4*	this study
HK-244, 246	HK-209, 211 *rnh203*-K46W *rnh1Δ::kanMX4*	this study
ZYH-21	HK-205 *rnh203*-K46W	this study
ZYH-23	HK-209 *rnh201*-G42S	this study
HK-638, 639	YS-291, 292 *rnh202Δ::kanMX4*	this study
HK-642, 643	YS-291, 292 *rnh202-pip*	this study
HK-645, 647	YS-291, 292 *rnh203Δ*	this study
HK-666, 667, 668, 669	HK-642, 643 *rnh1Δ::hygMX4 rnh202-pip*	this study
HK-670, 672	HK-638, 639 *rnh202Δ::kanMX4 spt3Δ::hygMX4*	this study
HK-674, 676	HK-645, 647 *rnh203Δ spt3Δ::hygMX4*	this study

Shown are names, relevant genotype and source for the *S. cerevisiae* strains used in this study. *HOcs* indicates the HO cutting site.

Table B.2 Oligos used in this study

Name	Size	Sequence
202PIP.F	86	5'AAAACCAAAAGTAGCCATAGGAAAAGGGGCCATTGATGGA-**GCTGCT**-AAACGTAAGTAGCTAGTATCATAATTAAACAGCAATTTGA
202PIP.R	86	5' TCAAATTGCTGTTTAATTATGATACTAGCTACTTACGTTT-**AGCAGC**-TCCATCAATGGCCCCTTTTCCTATGGCTACTTTTGGTTTT
RNH201.G42S.80.F	80	5'ACAGAATGACTCTCCAATAATAATGGGTATCGATGAAGCT**A**GCAGAG GGCCCGTATTAGGGCCAATGGTCTACGCAGTAG
RNH201.G42S.80.R	80	5'CTACTGCGTAGACCATTGGCCCTAATACGGGCCCTCTGC**T**AGCTTCAT CGATACCCATTATTATTGGAGAGTCATTCTGT
RNH203.K42W.80.F	80	5'GATACTATTTACTTTCGTGGCAAGGAACTGAAGAGGGAA**T**GGTCTGCG ACGCCTTCCAGTAGCGATAACACAACTAGTAA
RNH203.K42W.80.R	80	5'TTACTAGTTGTGTTATCGCTACTGGAAGGCGTCGCAGAC**CA**TTCCCTCT TCAGTTCCTTGCCACGAAAGTAAATAGTATC

Name, size and sequence of the oligos used in this study for strain construction are described. Nucleotides that introduce the desired mutations are bolded.

Table B.3 His⁺ frequencies for strains grown in glucose

A

Genotype	His⁺ freq.		Survival
WT	<0.01	(0-0)	40%
rnh201Δ	<0.01	(0-0)	44%
rnh202Δ	<0.01	(0-0)	50%
rnh203Δ	<0.01	(0-0)	35%
rnh202-pip	<0.01	(0-0)	39%
rnh1Δ rnh202-pip	<0.01	(0-0)	37%

B

Genotype	His⁺ freq.		Survival
WT	<0.01	(0-0)	34%
rnh1Δ	<0.01	(0-0)	34%
rnh201Δ	<0.01	(0-0)	21%
rnh1Δ rnh201Δ	0.9	(0-2)	16%
rnh201-G42S	<0.01	(0-0)	19%
rnh1Δ rnh201-G42S	<0.01	(0-0)	16%
rnh203-K46W	<0.01	(0-0)	20%
rnh1Δ rnh203-K46W	<0.01	(0-0)	25%
rnh201-G42S rnh203-K46W	<0.01	(0-0)	13%

Frequencies of His⁺ colonies per 10^7 viable cells for the indicated yeast strains following 24 h of glucose treatment are shown as median and 95% CI (in parentheses). Percentage of cell survival after growth in glucose is also shown. There were 4 repeats for each of the strains.

Table B.4 Statistical analysis (*P*-values) of the data

A

Genotype	*P*-values
WT *vs. rnh1Δ*	0.6734
WT *vs. rnh201Δ*	0.0009
WT *vs. rnh202Δ*	0.0009
WT *vs. rnh202-pip*	0.7728
WT *vs. rnh203Δ*	0.0009
rnh1Δ vs. rnh201Δ	0.0022
rnh1Δ vs. rnh202Δ	0.0022
rnh1Δ vs. rnh203Δ	0.0022
rnh1Δ vs. rnh202-pip	0.5427
rnh1Δ vs. rnh1Δ rnh202-pip	0.0411
rnh201Δ vs. rnh202Δ	1.0000
rnh201Δ vs. rnh202-pip	0.0022
rnh201Δ vs. rnh203Δ	0.0022
rnh201Δ vs. rnh1Δ rnh202-pip	0.0022
rnh202Δ vs. rnh202-pip	0.0022
rnh202Δ vs. rnh203Δ	0.0043
rnh202Δ vs. rnh1Δ rnh202-pip	0.0022
rnh203Δ vs. rnh202-pip	0.0009
rnh203Δ vs. rnh1Δ rnh202-pip	0.0049
rnh202-pip vs. rnh1Δ rnh202-pip	0.4260

Mann-Whitney *U*-test was applied to determine whether a statistical significant difference exists between pairs of gene correction frequencies obtained from the DSB repair assays of this study. **A**, Comparison of frequencies presented in Fig. (**2B**) and Table **1**. Two groups in a pair were considered to be significantly different when *P*-values were less than 0.05.

B

Genotype	P-values
WT *vs. rnh1Δ*	0.4381
WT *vs. rnh201Δ*	< 0.0001
WT *vs. rnh1Δ rnh201Δ*	<0.0001
WT *vs. rnh201*-G42S	0.8371
WT *vs. rnh1Δ rnh201*-G42S	<0.0001
WT *vs. rnh203*-K46W	0.4289
WT *vs. rnh1Δ rnh203*-K46W	0.0413
WT *vs. rnh201*-G42S *rnh203*-K46W	0.0009
rnh201Δ vs. rnh1Δ	<0.0001
rnh201Δ vs. rnh1Δ rnh201Δ	<0.0001
rnh201Δ vs. rnh201-G42S	<0.0001
rnh201Δ vs. rnh1Δ rnh201-G42S	0.0005
rnh201Δ vs. rnh203-K46W	< 0.0001
rnh201Δ vs. rnh1Δ rnh203-K46W	<0.0001
rnh201Δ vs. rnh201-G42S *rnh203*-K46W	< 0.0001
rnh1Δ vs. rnh1Δ rnh201Δ	<0.0001
rnh1Δ vs. rnh201-G42S	0.8618
rnh1Δ vs. rnh1Δ rnh201-G42S	0.0005
rnh1Δ vs. rnh203-K46W	0.9369
rnh1Δ vs. rnh1Δ rnh203-K46W	0.2891
rnh1Δ vs. rnh201-G42S *rnh203*-K46W	0.0479
rnh1Δ rnh201Δ vs. rnh201-G42S	< 0.0001
rnh1Δ rnh201Δ vs. rnh1Δ rnh201-G42S	< 0.0001
rnh1Δ rnh201Δ vs. rnh203-K46W	< 0.0001
rnh1Δ rnh201Δ vs. rnh1Δ rnh203-K46W	< 0.0001
rnh1Δ rnh201Δ vs. rnh201-G42S *rnh203*-K46W	< 0.0001
rnh201-G42S vs. rnh1Δ rnh201-G42S	< 0.0001
rnh201-G42S vs. rnh203-K46W	0.7637
rnh201-G42S vs. rnh1Δ rnh203-K46W	0.1173
rnh201-G42S vs. rnh201-G42S *rnh203*-K46W	0.0031
rnh1Δ rnh201-G42S vs. rnh203-K46W	0.0004
rnh1Δ rnh201-G42S vs. rnh1Δ rnh203-K46W	0.0046
rnh1Δ rnh201-G42S vs. rnh201-G42S *rnh203*-K46W	0.0237
rnh203-K46W *vs. rnh1Δ rnh203*-K46W	0.1787
rnh203-K46W *vs. rnh201*-G42S *rnh203*-K46W	0.0109
rnh1Δ rnh203-K46W *vs. rnh201*-G42S *rnh203*-K46W	0.2114

B, Comparison of frequencies presented in Fig. (**3**) and Table **2**. Two groups in a pair were considered to be significantly different when *P*-values were less than 0.05.

APPENDIX C

SUPPLEMENTARY MATERIALS FOR CHAPTER 5

Table C.1 Related to Figures 5.1-5. Sequences of the oligonucleotides used in this study (by A. V. Mazin)

Number	Length, nt	Sequence (5'→3')
1	63	ACAGCACCAGATTCAGCAATTAAGCTCTAAGCCATCCGCAAAAATGACCTCT TATCAAAAGGA
2	63	TCCTTTTGATAAGAGGTCATTTTTGCGGATGGCTTAGAGCTTAATTGCTGAA TCTGGTGCTGT
2R	63	UCCUUUUGAUAAGAGGUCAUUUUUGCGGAUGGCUUAGAGCUUAAUUGCUGAA UCUGGUGCUGU (ribonucleotide)
64	48	GTCGACGACGTCTGAGTACTCATCTAGTGTGACATCATCGCATCGAGA
64R	48	GUCGACGACGUCUGAGUACUCAUCUAGUGUGACAUCAUCGCAUCGAGA (ribonucleotide)
65	48	TCTCGATGCGATGATGTCACACTAGATGAGTACTCAGACGTCGTCGAC
117	94	TCCTTTTGATAAGAGGTCATTTTTGCGGATGGCTTAGAGCT TAATTGCTGAATCTGGTGCTGTAGGTCAACATGTTGTAAATA TGCAGCTAAAG
176	63	CAACGGCATAAAGCTTGACGATTACATTGCTAGGACATGCTGTCTAGAGGAT CCGACTATCGA
176R	63	CAACGGCAUAAAGCUUGACGAUUACAUUGCUAGGACAUGCUGUCUAGAGGAU CCGACUAUCGA(ribonucleotide)
518	73	TCCTTTTGATAAGAGGTCATTTTTGCGGATGGCTTAGAGCTTAATTGCTGAA TCTGGTGCTGTAGGTCAACAT
519	83	TCCTTTTGATAAGAGGTCATTTTTGCGGATGGCTTAGAGCTTAATTGCTGAA TCTGGTGCTGTAGGTCAACATGTTGTAAATA

Table C.2 Related to Table 1. Yeast strains used in this study

Strain	Relevant genotype	Source
YS-291, 292	hoΔ hmlΔ::ADE1 mataΔ::hisG hmrΔ::ADE1 ade1 leu2::pGAL1mhis3AI::HOcs-ADE3 lys5 trp1::hisG ura3-52 ade3::GAL::HO (his3::HOcs)::TRP1 YCLWTy2-1Δ	(Keskin et al., 2014)
CM-95, 96	YS-291, 292 [Cir⁺]*	This study
HK-936,937	CM-95, 96 rad52Δ:: kanMX4	This study
HK-444, 446	YS-291, 292 sae2Δ:: kanMX4	(Keskin et al., 2014)
HK-475, 477	YS-291, 292 exo1Δ:: kanMX4	(Keskin et al., 2014)
HK-344, 345	YS-291, 292 rad59Δ:: kanMX4	This study
YS-424, 426	YS-291, 292 rnh1Δ::NAT rnh201Δ::hygMX4	(Keskin et al., 2014)
CM-100, 101	YS-424, 426 [Cir⁺]*	This study
HK-938, 939	CM-100, 101 rad52Δ:: kanMX4	This study
HK-448, 450	YS-424, 426 sae2Δ:: kanMX4	This study
HK-479, 481	YS-424, 426 exo1Δ:: kanMX4	This study
HK-346, 347	YS-424, 426 rad59Δ:: kanMX4	This study
HK-138, 139	YS-424, 426 spt3Δ::KIURA3 rnh1Δ::NAT rnh201Δ::hygMX4	(Keskin et al., 2014)
CM-107, 108	HK-138, 139 [Cir⁺]*	This study
HK-452, 454	HK-138, 139 sae2Δ:: kanMX4	This study
HK-483, 484, 485, 486	HK-138, 139 exo1Δ:: kanMX4	This study
HK-348, 349	HK-138, 139 rad59Δ:: kanMX4	This study

Shown are names, relevant genotype and source for the *S. cerevisiae* strains used in this study. *HOcs* indicates the HO cutting site. *[Cir⁺], the yeast 2-micron plasmid was introduced in these strains.

Table C.3 Related to Table 1. His⁺ frequency in the *cis* system following transformation by the HIS3.F oligonucleotide

	cis			
Genotype	No Oligo		HIS3.F	
WT	<0.1	(0-0.3)	137,000	(114,000-187,000)
rad59	<0.1	(0-0)	30,000	(26,000-34,000)
sae2	2.5	(2-5)	14,000	(8,000-18,000)
exo1	1	(1-1.3)	52,000	(44,000-64,000)
rnh1 rnh201	170	(92-200)	350,000	(280,000-450,000)
rnh1 rnh201 rad59	61	(6-160)	37,000	(29,000-45,000)
rnh1 rnh201 sae2	185	(130-240)	39,000	(20,000-67,000)
rnh1 rnh201 exo1	24	(6-41)	205,000	(135,000-350,000)
rnh1 rnh201 spt3	83	(57-130)	300,000	(250,000-375,000)
rnh1 rnh201 spt3 rad59	19	(2-46)	97,000	(70,000-157,000)
rnh1 rnh201 spt3 sae2	230	(140-440)	59,000	(30,000-110,000)
rnh1 rnh201 spt3 exo1	88	(51-125)	110,000	(95,000-130,000)

Frequency of His$^+$ transformant colonies per 10^7 viable cells for the indicated yeast genotypes after transformation with HIS3.F oligo in *cis* system is shown as median and 95% CI (in parentheses). There were 4-12 repeats for each of the strains transformed with these oligos. The significance of comparisons between the strains in the *cis* systems were calculated using the Mann-Whitney U test (**Table C.5C**).

Table C.4 Related to Table 1. His+ frequencies for *rad59, exo1 and sae2 mutant* strains grown in glucose

Genotype	*cis*		Survival
	His$^+$ freq.		
WT	<0.01	(0-0)	44%
rad59	<0.01	(0-0)	54%
exo1	<0.01	(0-0)	40%
sae2	<0.01	(0-0)	40%
rnh1 rnh201	1.5	(1-2)	30%
rnh1 rnh201 rad59	0.09	(0-0.3)	42%
rnh1 rnh201 exo1	0.6	(0-0.8)	38%
rnh1 rnh201 sae2	<0.01	(0-0.3)	29%
rnh1 rnh201 spt3	<0.01	(0-0)	59%
rnh1 rnh201 spt3 rad59	0.07	(0-0.15)	60%
rnh1 rnh201 spt3 exo1	<0.01	(0-0)	24%
rnh1 rnh201 spt3 sae2	<0.01	(0-0)	38%

Frequencies of His$^+$ colonies per 10^7 viable cells for yeast strains of the *cis* system following 24 h of glucose treatment are shown as median and 95% CI (in parentheses). There were 4 repeats for all the strains.

Table C.5 Related to Table 1.Statistical analysis (*P*-values) of the data

<u>A</u>

Genotype	*P*-value
WT + YEP *vs.* WT + ScRad52	< 0.0001
WT + YEP *vs.* WT + ScRad52-327	< 0.0001
WT + YEP *vs.* WT + hRAD52-209	< 0.0001
WT + ScRAD52 *vs.* WT + ScRAD52-327	0.0002
WT + ScRAD52-327 *vs.* WT + hRAD52-209	0.0783
rad52 + ScRad52 *vs. rad52* + ScRad52-327	0.3007
rad52 + ScRad52-327 *vs. rad52* + hRAD52-209	0.1028
rnh1 rnh201 + YEP *vs. rnh1 rnh201* + ScRad52	0.0004
rnh1 rnh201 + YEP *vs. rnh1 rnh201* + ScRad52-327	< 0.0001
rnh1 rnh201 + YEP *vs. rnh1 rnh201* + hRAD52-209	< 0.0001
rnh1 rnh201 + ScRad52 *vs. rnh1 rnh201* + ScRad52-327	0.0009
rnh1 rnh201 + ScRad52-327 *vs. rnh1 rnh201* + hRAD52-209	0.0020
rnh1 rnh201 spt3 + YEP *vs. rnh1 rnh201 spt3* + ScRad52	0.1935
rnh1 rnh201 spt3 + YEP *vs. rnh1 rnh201 spt3* + ScRad52-327	< 0.0001
rnh1 rnh201 spt3 + YEP *vs. rnh1 rnh201 spt3* + hRAD52-209	< 0.0001
rnh1 rnh201 spt3 + ScRad52 *vs. rnh1 rnh201 spt3* + ScRad52-327	0.0009
rnh1 rnh201 spt3 + ScRad52-327 *vs. rnh1 rnh201 spt3* + hRAD52-209	0.0304
WT + YEP *vs. rnh1 rnh201* + YEP	< 0.0001
WT + YEP *vs. rnh1 rnh201 rad52* + YEP	0.1441
WT + YEP *vs. rnh1 rnh201 spt3* + YEP	< 0.0001
WT + ScRAD52 *vs. rad52* + ScRAD52	0.0032
WT + ScRAD52 *vs. rnh1 rnh201* + ScRAD52	0.0009
WT + ScRAD52 *vs. rnh1 rnh201 rad52* + ScRAD52	0.0009
WT + ScRAD52 *vs. rnh1 rnh201 spt3* + ScRAD52	0.0076
WT + ScRAD52-327 *vs. rad52* + ScRAD52-327	0.0012
WT + ScRAD52-327 *vs. rnh1 rnh201* + ScRAD52-327	< 0.0001
WT + ScRAD52-327 *vs. rnh1 rnh201 rad52* + ScRAD52-327	0.0021
WT + ScRAD52-327 *vs. rnh1 rnh201 spt3* + ScRAD52-327	< 0.0001
WT + hRAD52-209 *vs. rad52* + hRAD52-209	< 0.0001
WT + hRAD52-209 *vs. rnh1 rnh201* + hRAD52-209	< 0.0001
WT + hRAD52-209 *vs. rnh1 rnh201 rad52* + hRAD52-209	< 0.0001
WT + hRAD52-209 *vs. rnh1 rnh201 spt3* + hRAD52-209	< 0.0001
rad52 + ScRad52 *vs. rnh1 rnh201* + ScRad52	0.0050
rad52 + ScRad52 *vs. rnh1 rnh201 rad52* + ScRad52	0.0050
rad52 + ScRad52 *vs. rnh1 rnh201 spt3* + ScRad52	0.0022
rad52 + ScRad52-327 *vs. rnh1 rnh201* + ScRad52-327	< 0.0001
rad52 + ScRad52-327 *vs. rnh1 rnh201 rad52* + ScRad52-327	< 0.0001
rad52 + ScRad52-327 *vs. rnh1 rnh201 spt3* + ScRad52-327	< 0.0001
rad52 + hRAD52-209 *vs. rnh1 rnh201* + hRAD52-209	< 0.0001
rad52 + hRAD52-209 *vs. rnh1 rnh201 rad52* + hRAD52-209	< 0.0001

rad52 + hRAD52-209 *vs. rnh1 rnh201 spt3* + hRAD52-209	< 0.0001
rnh1 rnh201 + YEP *vs. rnh1 rnh201 rad52* + YEP	< 0.0001
rnh1 rnh201 + YEP *vs. rnh1 rnh201 spt3* + YEP	<0.0001
rnh1 rnh201 + ScRAD52 *vs. rnh1 rnh201 rad52* + ScRAD52	0.0050
rnh1 rnh201 + ScRAD52 *vs. rnh1 rnh201 spt3* + ScRAD52	0.0050
rnh1 rnh201 + ScRAD52-327 *vs. rnh1 rnh201 rad52* + ScRAD52-327	< 0.0001
rnh1 rnh201 + ScRAD52-327 *vs. rnh1 rnh201 spt3* + ScRAD52-327	< 0.0001
rnh1 rnh201 + hRAD52-209 *vs. rnh1 rnh201 rad52* + hRAD52-209	< 0.0001
rnh1 rnh201 + hRAD52-209 *vs. rnh1 rnh201 spt3* + hRAD52-209	< 0.0001

B

Genotype	*P*-values
WT *vs. rad59*	0.0489
WT *vs. sae2*	0.0004
WT *vs. exo1*	0.0007
WT *vs. rnh1 rnh201*	< 0.0001
WT *vs. rnh1 rnh201 rad59*	0.0004
WT *vs. rnh1 rnh201 sae2*	0.0004
WT *vs. rnh1 rnh201 exo1*	0.0004
WT *vs. rnh1 rnh201 spt3*	< 0.0001
WT *vs. rnh1 rnh201 spt3 rad59*	0.0004
WT *vs. rnh1 rnh201 spt3 sae2*	0.0004
WT *vs. rnh1 rnh201 spt3 exo1*	< 0.0001
exo1 vs. sae2	0.0050
exo1 vs. rad59	0.0050
exo1 vs. rnh1 rnh201	0.0004
exo1 vs. rnh1 rnh201 exo1	0.0050
exo1 vs. rnh1 rnh201 sae2	0.0050
exo1 vs. rnh1 rnh201 rad59	0.0050
exo1 vs. rnh1 rnh201 spt3	0.0004
exo1 vs. rnh1 rnh201 spt3 exo1	0.0009
exo1 vs. rnh1 rnh201 spt3 sae2	0.0050
exo1 vs. rnh1 rnh201 spt3 rad59	0.0050
sae2 vs. rad59	0.0022
sae2 vs. rnh1 rnh201	0.0004
sae2 vs. rnh1 rnh201 exo1	0.0022
sae2 vs. rnh1 rnh201 sae2	0.0022
sae2 vs. rnh1 rnh201 rad59	0.0022
sae2 vs. rnh1 rnh201 spt3	0.2175
sae2 vs. rnh1 rnh201 spt3 exo1	0.4260
sae2 vs. rnh1 rnh201 spt3 sae2	0.0022
sae2 vs. rnh1 rnh201 spt3 rad59	0.0022
rad59 vs. rnh1 rnh201	0.0004
rad59 vs. rnh1 rnh201 exo1	0.0022
rad59 vs. rnh1 rnh201 sae2	0.0022
rad59 vs. rnh1 rnh201 rad59	0.0050
rad59 vs. rnh1 rnh201 spt3	0.0004
rad59 vs. rnh1 rnh201 spt3 exo1	0.0009
rad59 vs. rnh1 rnh201 spt3 sae2	0.0022
rad59 vs. rnh1 rnh201 spt3 rad59	0.0050
rnh1 rnh201 vs. rnh1 rnh201 exo1	0.0005
rnh1 rnh201 vs. rnh1 rnh201 sae2	0.0004
rnh1 rnh201 vs. rnh1 rnh201 rad59	0.0004

rnh1 rnh201 vs. rnh1 rnh201 spt3	< 0.0001
rnh1 rnh201 vs. rnh1 rnh201 spt3 exo1	< 0.0001
rnh1 rnh201 vs. rnh1 rnh201 spt3 sae2	0.0046
rnh1 rnh201 vs. rnh1 rnh201 spt3 rad59	0.4432
rnh1 rnh201 exo1 vs. rnh1 rnh201 sae2	0.0022
rnh1 rnh201 exo1 vs. rnh1 rnh201 rad59	0.0022
rnh1 rnh201 exo1 vs. rnh1 rnh201 spt3	0.0004
rnh1 rnh201 exo1 vs. rnh1 rnh201 spt3 exo1	0.0009
rnh1 rnh201 exo1 vs. rnh1 rnh201 spt3 sae2	0.0022
rnh1 rnh201 exo1 vs. rnh1 rnh201 spt3 rad59	0.0022
rnh1 rnh201 sae2 vs. rnh1 rnh201 rad59	0.0022
rnh1 rnh201 sae2 vs. rnh1 rnh201 spt3	0.0004
rnh1 rnh201 sae2 vs. rnh1 rnh201 spt3 exo1	0.0009
rnh1 rnh201 sae2 vs. rnh1 rnh201 spt3 sae2	0.0022
rnh1 rnh201 sae2 vs. rnh1 rnh201 spt3 rad59	0.0022
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 exo1	0.0188
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 sae2	0.0004
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 rad59	0.0004
rnh1 rnh201 spt3 exo1 vs. rnh1 rnh201 spt3 sae2	0.0009
rnh1 rnh201 spt3 exo1 vs. rnh1 rnh201 spt3 rad59	0.0009
rnh1 rnh201 spt3 sae2 vs. rnh1 rnh201 spt3 rad59	0.054

C

Genotype	*P*-values
WT *vs. exo1*	0.0002
WT *vs. sae2*	0.0044
WT *vs. rad59*	0.0044
WT *vs. rnh1 rnh201*	0.0003
WT *vs. rnh1 rnh201 exo1*	0.0826
WT *vs. rnh1 rnh201 sae2*	0.0044
WT *vs. rnh1 rnh201 rad59*	0.0002
WT *vs. rnh1 rnh201 spt3*	0.0004
WT *vs. rnh1 rnh201 spt3 exo1*	0.1770
WT *vs. rnh1 rnh201 spt3 sae2*	0.0091
WT *vs. rnh1 rnh201 spt3 rad59*	0.1535
exo1 vs. sae2	0.0040
exo1 vs. rad59	0.0040
exo1 vs. rnh1 rnh201	0.0001
exo1 vs. rnh1 rnh201 exo1	0.0002
exo1 vs. rnh1 rnh201 sae2	0.4606
exo1 vs. rnh1 rnh201 rad59	0.0030
exo1 vs. rnh1 rnh201 spt3	0.0001
exo1 vs. rnh1 rnh201 spt3 exo1	0.0002
exo1 vs. rnh1 rnh201 spt3 sae2	1.0000
exo1 vs. rnh1 rnh201 spt3 rad59	0.0030
sae2 vs. rad59	0.0294
sae2 vs. rnh1 rnh201	0.0029
sae2 vs. rnh1 rnh201 exo1	0.0084
sae2 vs. rnh1 rnh201 sae2	0.0294
sae2 vs. rnh1 rnh201 rad59	0.0084
sae2 vs. rnh1 rnh201 spt3	0.0029
sae2 vs. rnh1 rnh201 spt3 exo1	0.0084
sae2 vs. rnh1 rnh201 spt3 sae2	0.0294
sae2 vs. rnh1 rnh201 spt3 rad59	0.0084
rad59 vs. rnh1 rnh201	0.0029
rad59 vs. rnh1 rnh201 exo1	0.0040
rad59 vs. rnh1 rnh201 sae2	1.0000
rad59 vs. rnh1 rnh201 rad59	0.1535
rad59 vs. rnh1 rnh201 spt3	0.0029
rad59 vs. rnh1 rnh201 spt3 exo1	0.0040
rad59 vs. rnh1 rnh201 spt3 sae2	0.1143
rad59 vs. rnh1 rnh201 spt3 rad59	0.0040
rnh1 rnh201 vs. rnh1 rnh201 exo1	0.0922
rnh1 rnh201 vs. rnh1 rnh201 sae2	0.0029
rnh1 rnh201 vs. rnh1 rnh201 rad59	0.0001

rnh1 rnh201 vs. rnh1 rnh201 spt3	0.4739
rnh1 rnh201 vs. rnh1 rnh201 spt3 exo1	0.0002
rnh1 rnh201 vs. rnh1 rnh201 spt3 sae2	0.0029
rnh1 rnh201 vs. rnh1 rnh201 spt3 rad59	0.0004
rnh1 rnh201 exo1 vs. rnh1 rnh201 sae2	0.0040
rnh1 rnh201 exo1 vs. rnh1 rnh201 rad59	0.0002
rnh1 rnh201 exo1 vs. rnh1 rnh201 spt3	0.0922
rnh1 rnh201 exo1 vs. rnh1 rnh201 spt3 exo1	0.0104
rnh1 rnh201 exo1 vs. rnh1 rnh201 spt3 sae2	0.0040
rnh1 rnh201 exo1 vs. rnh1 rnh201 spt3 rad59	0.0104
rnh1 rnh201 sae2 vs. rnh1 rnh201 rad59	0.9333
rnh1 rnh201 sae2 vs. rnh1 rnh201 spt3	0.0029
rnh1 rnh201 sae2 vs. rnh1 rnh201 spt3 exo1	0.0040
rnh1 rnh201 sae2 vs. rnh1 rnh201 spt3 sae2	0.3429
rnh1 rnh201 sae2 vs. rnh1 rnh201 spt3 rad59	0.0081
rnh1 rnh201 rad59 vs. rnh1 rnh201 spt3	0.0001
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 exo1	0.0001
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 sae2	0.0029
rnh1 rnh201 spt3 vs. rnh1 rnh201 spt3 rad59	0.0002
rnh1 rnh201 spt3 exo1 vs. rnh1 rnh201 spt3 sae2	0.0081
rnh1 rnh201 spt3 exo1 vs. rnh1 rnh201 spt3 rad59	0.7209
rnh1 rnh201 spt3 sae2 vs. rnh1 rnh201 spt3 rad59	0.1091

Mann-Whitney-U test was applied to determine whether a statistical significant difference exists between pairs of gene correction frequencies obtained from the DSB repair assays of this study. Comparison of frequencies presented in **Table 5.1a** (**a**). Table 1b (**b**), and **Table C.3** (**c**). Two groups in a pair were considered to be significantly different when P-values were less than 0.05.

A RAD52, non-homologous ssDNA

Time (min): 0 5 10 20 30 45 60 60⁸ᴾ

→ 3'-tailed DNA #117/1*

→ ssDNA #1*

Figure C.1. Related to Figure 5.1. Requirements of the inverse DNA strand exchange promoted by hRad52. A, Inverse DNA strand exchange does not proceed

with nonhomologous ssDNA. The reaction conditions were as in Figure 1D, except that homologous ssDNA (no. 2; 205.8 nM) was substituted with non-homologous ssDNA (no. 176; 205.8 nM). **B**, Effect of magnesium acetate concentrations on Inverse DNA strand exchange. The DNA substrates and reaction conditions were the same as in Figure 1D. hRad52 (900 nM) was incubated with 3'-tailed DNA (no.1/no 117, 68.6 nM) in the presence of indicated magnesium acetate concentrations. Inverse DNA strand exchange was initiated by adding ssDNA (no. 2; 205.8 nM) and carried out for 5 min. **C**, Effect of the ssDNA concentrations on the efficiency of inverse DNA strand exchange. The reactions were carried out under the standard conditions in the presence of 3-tailed DNA (no. 117/ no. 1; 68.6 nM) and four concentrations of homologous ssDNA (no. 2): 68.6 nM, 137.2 nM, 205.8 nM, and 411.6 nM, which correspond to 1- (equimolar), 2-, 3-, and 6-fold excess (in molecules) over the 3'-tailed DNA substrate, respectively. **D,** No ssDNA intermediate is produced under conditions of inverse DNA strand exchange promoted by hRad52. Top: Experimental scheme. ssDNA (no. 2, 68.6 nM) or dsDNA (no. 1/no. 2, 68.6 nM) were incubated at 37 °C for 15 min with hRad52 protein (900 nM), and then treated with P1 nuclease (0.4 U per 10 μl of the reaction volume) at 37 °C for 10 min. In controls, hRad52 (lanes 1 and 3) or P1 nuclease (lanes 1, 3, 5 and 7) were omitted. Incubation with P1 nuclease causes degradation of free ssDNA (lane 2), but not free dsDNA (lane 4). Pre-incubation of hRad52 with ssDNA causes only partial protection of the ssDNA against P1 nuclease degradation (lane 8). The experiments were repeated at least three times, error bars indicate SD. (**by A. V. Mazin**)
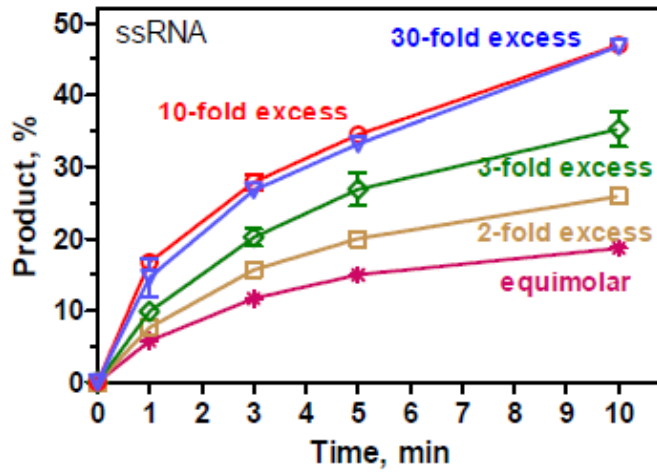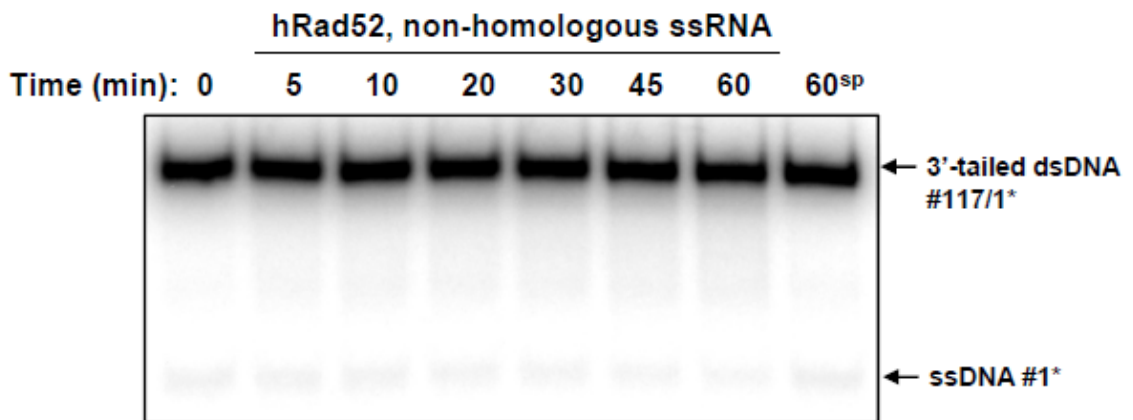
Figure C.2. Related to Figure 5.2. Effect of the ssRNA concentration (A) and homology (B) on hRad52-promoted inverse strand exchange. A, The reactions were performed using 1- (equimolar), 2-, 3-, 10-, and 30-fold excess of ssRNA over to the 3'-tailed dsDNA concentration (68.6 nM). **B**, Inverse RNA strand exchange does not proceed with non-homologous ssRNA. The reaction conditions were as in Fig. 2B, except

that homologous ssRNA (no. 2R; 205.8 nM) was substituted with non-homologous

ssRNA (no. 176R; 205.8 nM). Tailed dsDNA (#117/1) was used as a DNA substrate; the

asterisk on ssDNA #1 indicates 32P label. The experiments were repeated at least three

times, error bars indicate SD. (**by A. V. Mazin**)
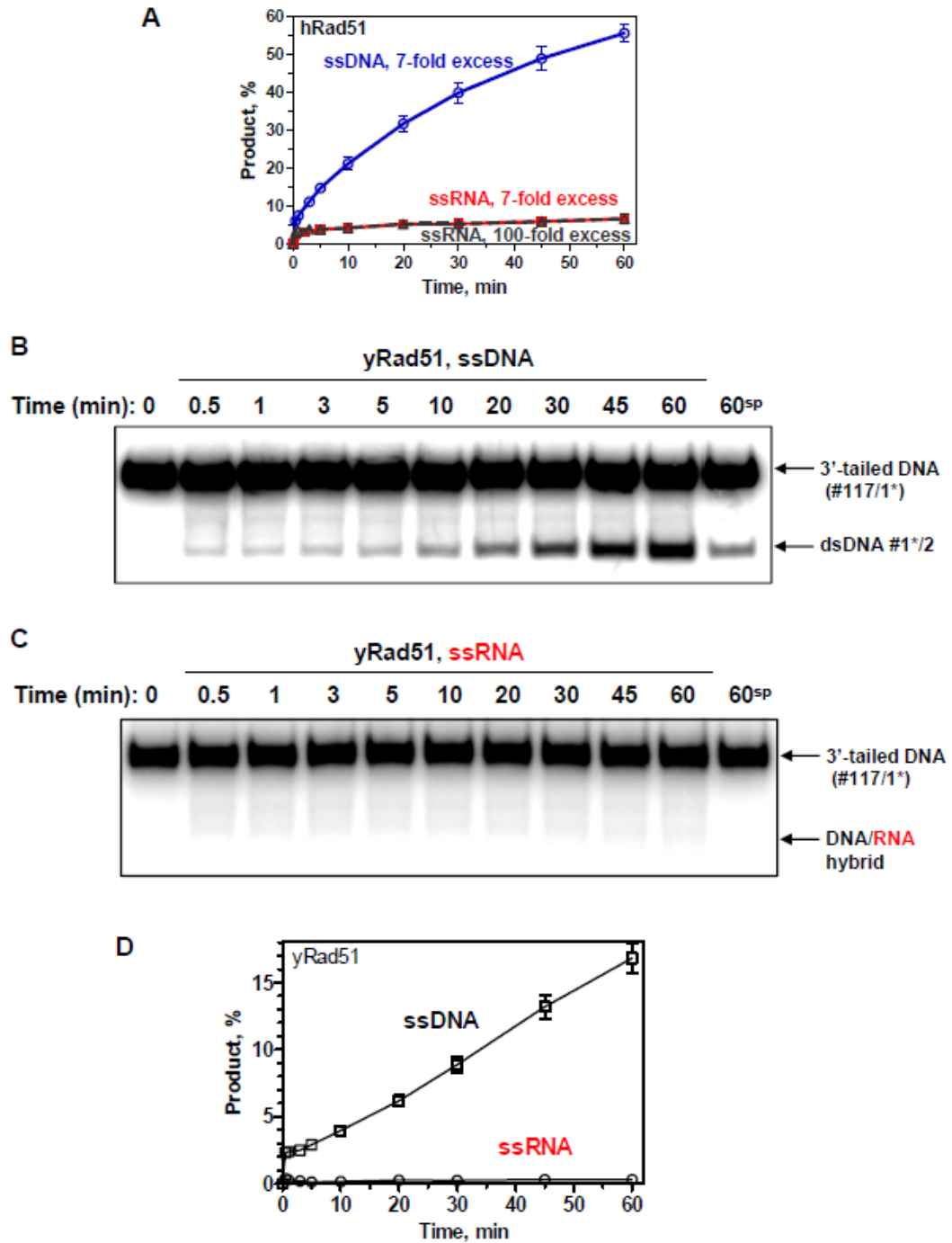
Figure C.3. Related to Figure 5.3. Human Rad51 and yeast Rad51 do not promote inverse strand exchange with ssRNA. A, hRad51 does not promote inverse RNA strand exchange even at high, 7- or 100-fold molar excess, ssRNA (no. 2R, 480.2 nM or 6.86 µM). The kinetics of hRad51-promoted DNA inverse exchange with ssDNA (7-fold

excess) from Figure 1D is shown for comparison. **B**, yRad51 promotes inverse DNA

strand exchange. The reaction was conducted as described for hRad51, except that a 10-

fold molar excess of ssDNA was used (no. 2, 686 nM). **C**, yRad51 does not promote

inverse RNA strand exchange. ssDNA was replaced with ssRNA (no. 2R, 686 nM). The

reaction products were analyzed by electrophoresis in a polyacrylamide gel. **D**, Data from

panels B and C were plotted as a graph. The experiments were repeated at least three

times, error bars indicate SD. (**by A. V. Mazin**)
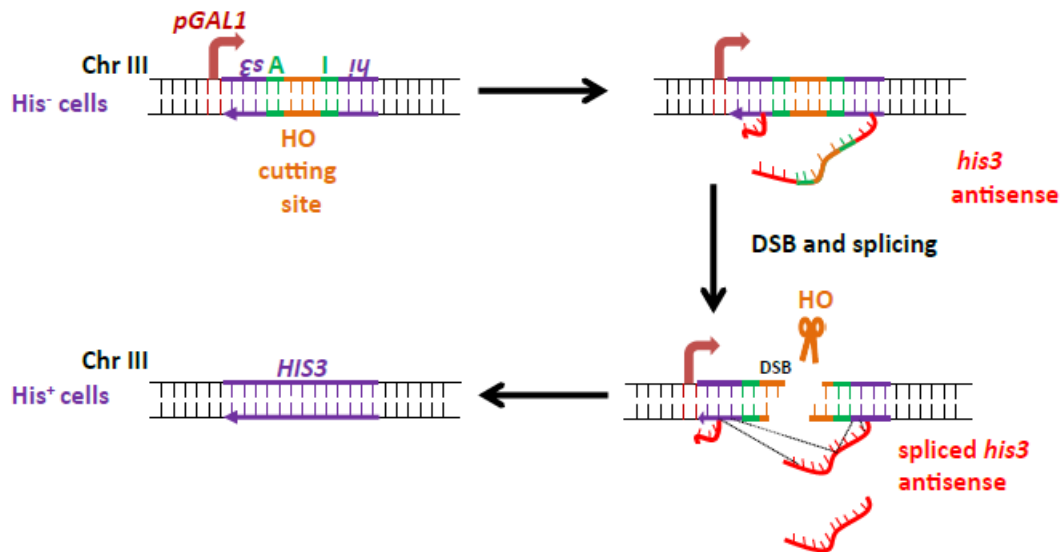
Figure C.4. Related to Table 1. Scheme of the *cis* assay of DSB repair by transcript

RNA. The inactive *his3* marker gene is in purple, the artificial intron (AI) in green and

the cutting site for the HO homothallic-switching-endonuclease in orange. Because the

his3 gene is inactive yeast cells cannot grow on medium without histidine, they are His⁻.

Upon plating on galactose-containing medium, the *pGAL1* promoter activates the

transcription the *his3* antisense RNA (in red) and the expression of the HO endonuclease.

After formation of the DSB by HO endonuclease the broken *his3* gene can be repaired to

functional *HIS3* gene giving His$^+$ cells only if the antisense *his3* RNA serves as template

after removal of the *AI* with the HO cutting site sequence by splicing.

REFERENCES

1.   Bhattacharjee, S. and S. Nandi, Choices have consequences: the nexus between DNA repair pathways and genomic instability in cancer. Clin Transl Med, 2016. **5**(1): p. 45.
2.   Symington, L.S. and J. Gautier, Double-strand break end resection and repair pathway choice. Annu Rev Genet, 2011. **45**: p. 247-71.
3.   Liu, T. and J. Huang, Quality control of homologous recombination. Cell Mol Life Sci, 2014. **71**(19): p. 3779-97.
4.   San Filippo, J., P. Sung, and H. Klein, Mechanism of eukaryotic homologous recombination. Annu Rev Biochem, 2008. **77**: p. 229-57.
5.   Chiruvella, K.K., Z. Liang, and T.E. Wilson, Repair of double-strand breaks by end joining. Cold Spring Harb Perspect Biol, 2013. **5**(5): p. a012757.
6.   Kass, E.M. and M. Jasin, Collaboration and competition between DNA double-strand break repair pathways. FEBS Lett, 2010. **584**(17): p. 3703-8.
7.   Ma, Y., et al., Repair of double-strand DNA breaks by the human nonhomologous DNA end joining pathway: the iterative processing model. Cell Cycle, 2005. **4**(9): p. 1193-200.
8.   Li, X. and W.D. Heyer, Homologous recombination in DNA repair and DNA damage tolerance. Cell Res, 2008. **18**(1): p. 99-113.
9.   Liu, T. and J. Huang, DNA End Resection: Facts and Mechanisms. Genomics Proteomics Bioinformatics, 2016. **14**(3): p. 126-30.
10.   Longhese, M.P., et al., Mechanisms and regulation of DNA end resection. Embo j, 2010. **29**(17): p. 2864-74.
11.   Mimitou, E.P. and L.S. Symington, Nucleases and helicases take center stage in homologous recombination. Trends Biochem Sci, 2009. **34**(5): p. 264-72.
12.   Huertas, P., DNA resection in eukaryotes: deciding how to fix the break. Nat Struct Mol Biol, 2010. **17**(1): p. 11-6.
13.   Heyer, W.D., K.T. Ehmsen, and J. Liu, Regulation of homologous recombination in eukaryotes. Annu Rev Genet, 2010. **44**: p. 113-39.
14.   Jekimovs, C., et al., Chemotherapeutic compounds targeting the DNA double-strand break repair pathways: the good, the bad, and the promising. Front Oncol, 2014. **4**: p. 86.
15.   Baltimore, D., Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. Cell, 1985. **40**(3): p. 481-2.
16.   Autexier, C. and N.F. Lue, The structure and function of telomerase reverse transcriptase. Annu Rev Biochem, 2006. **75**: p. 493-517.
17.   Derr, L.K., J.N. Strathern, and D.J. Garfinkel, RNA-mediated recombination in S. cerevisiae. Cell, 1991. **67**(2): p. 355-64.
18.   Curcio, M.J. and D.J. Garfinkel, Single-step selection for Ty1 element retrotransposition. Proc Natl Acad Sci U S A, 1991. **88**(3): p. 936-40.
19.   Derr, L.K. and J.N. Strathern, A role for reverse transcripts in gene conversion. Nature, 1993. **361**(6408): p. 170-3.
20.   Moore, J.K. and J.E. Haber, Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. Nature, 1996. **383**(6601): p. 644-6.
21.   Teng, S.C., B. Kim, and A. Gabriel, Retrotransposon reverse-transcriptase-mediated repair of chromosomal breaks. Nature, 1996. **383**(6601): p. 641-4.

22.   Esnault, C., J. Maestre, and T. Heidmann, Human LINE retrotransposons generate processed pseudogenes. Nat Genet, 2000. **24**(4): p. 363-7.
23.   Morrish, T.A., et al., DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nat Genet, 2002. **31**(2): p. 159-65.
24.   Ding, W., et al., L1 elements, processed pseudogenes and retrogenes in mammalian genomes. IUBMB Life, 2006. **58**(12): p. 677-85.
25.   Meers, C., H. Keskin, and F. Storici, DNA repair by RNA: Templated, or not templated, that is the question. DNA Repair (Amst), 2016. **44**: p. 17-21.
26.   Melamed, C., Y. Nevo, and M. Kupiec, Involvement of cDNA in homologous recombination between Ty elements in Saccharomyces cerevisiae. Mol Cell Biol, 1992. **12**(4): p. 1613-20.
27.   Nevo-Caspi, Y. and M. Kupiec, Induction of Ty recombination in yeast by cDNA and transcription: role of the RAD1 and RAD52 genes. Genetics, 1996. **144**(3): p. 947-55.
28.   Parket, A., O. Inbar, and M. Kupiec, Recombination of Ty elements in yeast can be induced by a double-strand break. Genetics, 1995. **140**(1): p. 67-77.
29.   Jinek, M., et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science, 2012. **337**(6096): p. 816-21.
30.   Wei, W., et al., A role for small RNAs in DNA double-strand break repair. Cell, 2012. **149**(1): p. 101-12.
31.   Francia, S., et al., Site-specific DICER and DROSHA RNA products control the DNA-damage response. Nature, 2012. **488**(7410): p. 231-5.
32.   Bayne, E.H. and R.C. Allshire, RNA-directed transcriptional gene silencing in mammals. Trends Genet, 2005. **21**(7): p. 370-3.
33.   Zaitsev, E.N. and S.C. Kowalczykowski, A novel pairing process promoted by Escherichia coli RecA protein: inverse DNA and RNA strand exchange. Genes Dev, 2000. **14**(6): p. 740-9.
34.   Kasahara, M., et al., RecA protein-dependent R-loop formation in vitro. Genes & Development, 2000. **14**(3): p. 360-365.
35.   Nowacki, M., et al., RNA-mediated epigenetic programming of a genome-rearrangement pathway. Nature, 2008. **451**(7175): p. 153-8.
36.   Bozas, A., et al., Genetic analysis of zinc-finger nuclease-induced gene targeting in Drosophila. Genetics, 2009. **182**(3): p. 641-51.
37.   Diemer, G.S. and K.M. Stedman, A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. Biol Direct, 2012. **7**: p. 13.
38.   Wahba, L., S.K. Gore, and D. Koshland, The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. Elife, 2013. **2**: p. e00505.
39.   Shen, Y., et al., RNA-driven genetic changes in bacteria and in human cells. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 2011. **717**(1–2): p. 91-98.
40.   Shen, Y. and F. Storici, Detection of RNA-Templated Double-Strand Break Repair in Yeast, in DNA Recombination, H. Tsubouchi, Editor. 2011, Humana Press. p. 193-204.
41.   Thaler, D.S., G. Tombline, and K. Zahn, Short-patch reverse transcription in Escherichia coli. Genetics, 1995. **140**(3): p. 909-15.

42. Aguilera, A. and T. Garcia-Muse, R loops: from transcription byproducts to threats to genome stability. Mol Cell, 2012. **46**(2): p. 115-24.

43. Westover, K.D., D.A. Bushnell, and R.D. Kornberg, Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. Cell, 2004. **119**(4): p. 481-9.

44. Skourti-Stathaki, K. and N.J. Proudfoot, A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev, 2014. **28**(13): p. 1384-96.

45. Nadel, J., et al., RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. Epigenetics Chromatin, 2015. **8**: p. 46.

46. Mischo, H.E., et al., Yeast Sen1 helicase protects the genome from transcription-associated instability. Mol Cell, 2011. **41**(1): p. 21-32.

47. Skourti-Stathaki, K., N.J. Proudfoot, and N. Gromak, Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol Cell, 2011. **42**(6): p. 794-805.

48. Boule, J.B., L.R. Vega, and V.A. Zakian, The yeast Pif1p helicase removes telomerase from telomeric DNA. Nature, 2005. **438**(7064): p. 57-61.

49. Zhang, D.H., et al., The human Pif1 helicase, a potential Escherichia coli RecD homologue, inhibits telomerase activity. Nucleic Acids Res, 2006. **34**(5): p. 1393-404.

50. El Hage, A., et al., Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. Genes Dev, 2010. **24**(14): p. 1546-58.

51. Yang, Y., et al., Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. Mol Cell, 2014. **53**(3): p. 484-97.

52. Cerritelli, S.M. and R.J. Crouch, Ribonuclease H: the enzymes in eukaryotes. Febs j, 2009. **276**(6): p. 1494-505.

53. Stein, H. and P. Hausen, Enzyme from calf thymus degrading the RNA moiety of DNA-RNA Hybrids: effect on DNA-dependent RNA polymerase. Science, 1969. **166**(3903): p. 393-5.

54. Kanaya, S. and M. Ikehara, Functions and structures of ribonuclease H enzymes. Subcell Biochem, 1995. **24**: p. 377-422.

55. Ohtani, N., et al., Identification of the genes encoding Mn2+-dependent RNase HII and Mg2+-dependent RNase HIII from Bacillus subtilis: classification of RNases H into three families. Biochemistry, 1999. **38**(2): p. 605-18.

56. Eder, P.S., R.Y. Walder, and J.A. Walder, Substrate specificity of human RNase H1 and its role in excision repair of ribose residues misincorporated in DNA. Biochimie, 1993. **75**(1-2): p. 123-6.

57. Cerritelli, S.M., et al., Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. Mol Cell, 2003. **11**(3): p. 807-15.

58. Reijns, M.A., et al., Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. Cell, 2012. **149**(5): p. 1008-22.

59. Rabe, B., Aicardi-Goutieres syndrome: clues from the RNase H2 knock-out mouse. J Mol Med (Berl), 2013. **91**(11): p. 1235-40.

60. Aicardi, J. and F. Goutieres, A progressive familial encephalopathy in infancy with calcifications of the basal ganglia and chronic cerebrospinal fluid lymphocytosis. Ann Neurol, 1984. **15**(1): p. 49-54.

61. Stetson, D.B., et al., Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell, 2008. **134**(4): p. 587-98.
62. Rice, G.I., et al., Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response. Nat Genet, 2009. **41**(7): p. 829-32.
63. Rice, G.I., et al., Mutations in ADAR1 cause Aicardi-Goutieres syndrome associated with a type I interferon signature. Nat Genet, 2012. **44**(11): p. 1243-8.
64. Crow, Y.J. and J. Rehwinkel, Aicardi-Goutieres syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity. Hum Mol Genet, 2009. **18**(R2): p. R130-6.
65. Sztuba-Solińska, J., et al., RNA-RNA Recombination in Plant Virus Replication and Evolution. Annual Review of Phytopathology, 2011. **49**(1): p. 415-443.
66. Storici, F., et al., RNA-templated DNA repair. Nature, 2007. **447**(7142): p. 338-41.
67. Nowacki, M., et al., RNA-mediated epigenetic programming of a genome-rearrangement pathway. Nature, 2008. **451**(7175): p. 153-158.
68. Curcio, M.J. and D.J. Garfinkel, Single-step selection for Ty1 element retrotransposition. Proceedings of the National Academy of Sciences of the United States of America, 1991. **88**(3): p. 936-940.
69. Dombroski, B.A., et al., An in vivo assay for the reverse transcriptase of human retrotransposon L1 in Saccharomyces cerevisiae. Molecular and Cellular Biology, 1994. **14**(7): p. 4485-4492.
70. Teng, S.-C., B. Kim, and A. Gabriel, Retrotransposon reverse-transcriptase-mediated repair of chromosomal breaks. Nature, 1996. **383**(6601): p. 641-644.
71. Boeke, J.D., C.A. Styles, and G.R. Fink, Saccharomyces cerevisiae SPT3 gene is required for transposition and transpositional recombination of chromosomal Ty elements. Mol Cell Biol, 1986. **6**(11): p. 3575-81.
72. Chapman, K.B. and J.D. Boeke, Isolation and characterization of the gene encoding yeast debranching enzyme. Cell, 1991. **65**(3): p. 483-492.
73. Karst, S.M., M.-L. Rütz, and T.M. Menees, The Yeast Retrotransposons Ty1 and Ty3 Require the RNA Lariat Debranching Enzyme, Dbr1p, for Efficient Accumulation of Reverse Transcripts. Biochemical and Biophysical Research Communications, 2000. **268**(1): p. 112-117.
74. Lee, B.-S., et al., Nucleotide Excision Repair/TFIIH Helicases Rad3 and Ssl2 Inhibit Short-Sequence Recombination and Ty1 Retrotransposition by Similar Mechanisms. Molecular and Cellular Biology, 2000. **20**(7): p. 2436-2445.
75. Stuckey, S., K. Mukherjee, and F. Storici, In vivo site-specific mutagenesis and gene collage using the delitto perfetto system in yeast Saccharomyces cerevisiae. Methods Mol Biol, 2011. **745**: p. 173-91.
76. Storici, F. and M.A. Resnick, The Delitto Perfetto Approach to In Vivo Site-Directed Mutagenesis and Chromosome Rearrangements with Synthetic Oligonucleotides in Yeast, in Methods in Enzymology, L.C. Judith and M. Paul, Editors. 2006, Academic Press. p. 329-345.
77. Dakshinamurthy, A., et al., BUD22 Affects Ty1 Retrotransposition and Ribosome Biogenesis in Saccharomyces cerevisiae. Genetics, 2010. **185**(4): p. 1193-1205.
78. Storici, F., et al., The flexible loop of human FEN1 endonuclease is required for flap cleavage during DNA replication and repair. The EMBO Journal, 2002. **21**(21): p. 5930-5942.

79.   Sherman, F., Fink, G.R.,& Hicks, J.B., Laboratory Course Manual for Methods in Yeast Genetics. Cold Spring Harbor Laboratory Press, 1986.

80.   Sharon, G., T.J. Burkett, and D.J. Garfinkel, Efficient homologous recombination of Ty1 element cDNA when integration is blocked. Molecular and Cellular Biology, 1994. **14**(10): p. 6540-6551.

81.   Lea, D.E. and C.A. Coulson, The distribution of the numbers of mutants in bacterial populations. J Genet, 1949. **49**(3): p. 264-85.

82.   Fasken, M.B., et al., Air1 Zinc Knuckles 4 and 5 and a Conserved IWRXY Motif Are Critical for the Function and Integrity of the Trf4/5-Air1/2-Mtr4 Polyadenylation (TRAMP) RNA Quality Control Complex. The Journal of Biological Chemistry, 2011. **286**(43): p. 37429-37445.

83.   Keskin, H., et al., Complex effects of flavopiridol on the expression of primary response genes. Cell Division, 2012. **7**: p. 11-11.

84.   Sugiyama, T., J.H. New, and S.C. Kowalczykowski, DNA annealing by Rad52 Protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(11): p. 6049-6054.

85.   Rossi, M., et al., Reconstituting the Key Steps of the DNA Double-Strand Break Repair In Vitro, in DNA Recombination, H. Tsubouchi, Editor. 2011, Humana Press. p. 407-420.

86.   Sokal, R.R., & Rohlf, F.J., Biometry: The Principles and Practice of Statistics in Biological Research 2nd edn. W.H. Freeman and Company, 1981.

87.   Storey, J.D. and R. Tibshirani, Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(16): p. 9440-9445.

88.   Moore, J.K. and J.E. Haber, Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. Nature, 1996. **383**(6601): p. 644-646.

89.   Onozawa, M., et al., Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(21): p. 7729-7734.

90.   Ruff, P., et al., Aptamer-guided gene targeting in yeast and human cells. Nucleic Acids Research, 2014. **42**(7): p. e61-e61.

91.   Rocha, P.P., J. Chaumeil, and J.A. Skok, Finding the Right Partner in a 3D Genome. Science (New York, N.Y.), 2013. **342**(6164): p. 1333-1334.

92.   Kohler, A. and E. Hurt, Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol, 2007. **8**(10): p. 761-773.

93.   Nguyen, T.A., et al., Analysis of subunit assembly and function of the Saccharomyces cerevisiae RNase H2 complex. Febs j, 2011. **278**(24): p. 4927-42.

94.   Boeke, J.D., C.A. Styles, and G.R. Fink, Saccharomyces cerevisiae SPT3 gene is required for transposition and transpositional recombination of chromosomal Ty elements. Molecular and Cellular Biology, 1986. **6**(11): p. 3575-3581.

95.   Symington, L.S., Role of RAD52 Epistasis Group Genes in Homologous Recombination and Double-Strand Break Repair. Microbiology and Molecular Biology Reviews, 2002. **66**(4): p. 630-670.

96. Storici, F., et al., Conservative Repair of a Chromosomal Double-Strand Break by Single-Strand DNA through Two Steps of Annealing. Molecular and Cellular Biology, 2006. **26**(20): p. 7645-7657.

97. Hamperl, S. and K.A. Cimprich, The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. DNA Repair, 2014. **19**(0): p. 84-94.

98. Crick, F., Central dogma of molecular biology. Nature, 1970. **227**(5258): p. 561-3.

99. Greider, C.W. and E.H. Blackburn, Identification of a specific telomere terminal transferase activity in tetrahymena extracts. Cell, 1985. **43**(2, Part 1): p. 405-413.

100. Crow, Y.J., et al., Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. Nat Genet, 2006. **38**(8): p. 910-916.

101. Shen, Y., et al., Mispaired rNMPs in DNA are mutagenic and are targets of mismatch repair and RNases H. Nat Struct Mol Biol, 2011. **19**(1): p. 98-104.

102. Koh, K.D., et al., Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. Nat Methods, 2015. **12**(3): p. 251-7, 3 p following 257.

103. Qiu, J., et al., Saccharomyces cerevisiae RNase H(35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease. Mol Cell Biol, 1999. **19**(12): p. 8361-71.

104. Kao, H.I. and R.A. Bambara, The protein components and mechanism of eukaryotic Okazaki fragment maturation. Crit Rev Biochem Mol Biol, 2003. **38**(5): p. 433-52.

105. Ruhanen, H., K. Ushakov, and T. Yasukawa, Involvement of DNA ligase III and ribonuclease H1 in mitochondrial DNA replication in cultured human cells. Biochim Biophys Acta, 2011. **1813**(12): p. 2000-7.

106. Groh, M. and N. Gromak, Out of balance: R-loops in human disease. PLoS Genet, 2014. **10**(9): p. e1004630.

107. Chon, H., et al., RNase H2 roles in genome integrity revealed by unlinking its activities. Nucleic Acids Res, 2013. **41**(5): p. 3130-43.

108. El Hage, A., et al., Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. PLoS Genet, 2014. **10**(10): p. e1004716.

109. Stuckey, R., et al., Role for RNA:DNA hybrids in origin-independent replication priming in a eukaryotic system. Proc Natl Acad Sci U S A, 2015. **112**(18): p. 5779-84.

110. Huertas, P. and A. Aguilera, Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol Cell, 2003. **12**(3): p. 711-21.

111. Wahba, L., et al., RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. Mol Cell, 2011. **44**(6): p. 978-88.

112. Wilhelm, F.X., M. Wilhelm, and A. Gabriel, Reverse transcriptase and integrase of the Saccharomyces cerevisiae Ty1 element. Cytogenet Genome Res, 2005. **110**(1-4): p. 269-87.

113. Beck, C.R., et al., LINE-1 retrotransposition activity in human genomes. Cell, 2010. **141**(7): p. 1159-70.

114. Lesage, P. and A.L. Todeschini, Happy together: the life and times of Ty retrotransposons and their hosts. Cytogenet Genome Res, 2005. **110**(1-4): p. 70-90.

115. Beauregard, A., M.J. Curcio, and M. Belfort, The take and give between retrotransposable elements and their hosts. Annu Rev Genet, 2008. **42**: p. 587-617.

116.   Luan, D.D., et al., Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell, 1993. **72**(4): p. 595-605.

117.   Nevo-Caspi, Y. and M. Kupiec, Transcriptional induction of Ty recombination in yeast. Proc Natl Acad Sci U S A, 1994. **91**(26): p. 12711-5.

118.   Sharon, G., T.J. Burkett, and D.J. Garfinkel, Efficient homologous recombination of Ty1 element cDNA when integration is blocked. Mol Cell Biol, 1994. **14**(10): p. 6540-51.

119.   Nevo-Caspi, Y. and M. Kupiec, cDNA-mediated Ty recombination can take place in the absence of plus-strand cDNA synthesis, but not in the absence of the integrase protein. Curr Genet, 1997. **32**(1): p. 32-40.

120.   Morrish, T.A., et al., Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres. Nature, 2007. **446**(7132): p. 208-12.

121.   Keskin, H., et al., Transcript-RNA-templated DNA recombination and repair. Nature, 2014. **515**(7527): p. 436-9.

122.   Crow, Y.J., et al., Therapies in Aicardi-Goutieres syndrome. Clin Exp Immunol, 2014. **175**(1): p. 1-8.

123.   Volkman, H.E. and D.B. Stetson, The enemy within: endogenous retroelements and autoimmune disease. Nat Immunol, 2014. **15**(5): p. 415-22.

124.   Crow, Y.J., et al., Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. Nat Genet, 2006. **38**(8): p. 910-6.

125.   Rice, G.I., et al., Synonymous mutations in RNASEH2A create cryptic splice sites impairing RNase H2 enzyme function in Aicardi-Goutieres syndrome. Hum Mutat, 2013. **34**(8): p. 1066-70.

126.   Storici, F., L.K. Lewis, and M.A. Resnick, In vivo site-directed mutagenesis using oligonucleotides. Nat Biotechnol, 2001. **19**(8): p. 773-6.

127.   Storici, F. and M.A. Resnick, Delitto perfetto targeted mutagenesis in yeast with oligonucleotides. Genet Eng (N Y), 2003. **25**: p. 189-207.

128.   Eisenmann, D.M., et al., SPT3 interacts with TFIID to allow normal transcription in Saccharomyces cerevisiae. Genes Dev, 1992. **6**(7): p. 1319-31.

129.   Chon, H., et al., Contributions of the two accessory subunits, RNASEH2B and RNASEH2C, to the activity and properties of the human RNase H2 complex. Nucleic Acids Res, 2009. **37**(1): p. 96-110.

130.   Jeong, H.S., et al., RNase H2 of Saccharomyces cerevisiae is a complex of three proteins. Nucleic Acids Res, 2004. **32**(2): p. 407-14.

131.   Rohman, M.S., et al., Effect of the disease-causing mutations identified in human ribonuclease (RNase) H2 on the activities and stabilities of yeast RNase H2 and archaeal RNase HII. Febs j, 2008. **275**(19): p. 4836-49.

132.   Lee-Kirsch, M.A., C. Wolf, and C. Gunther, Aicardi-Goutieres syndrome: a model disease for systemic autoimmunity. Clin Exp Immunol, 2014. **175**(1): p. 17-24.

133.   Rice, G.I., et al., Gain-of-function mutations in IFIH1 cause a spectrum of human disease phenotypes associated with upregulated type I interferon signaling. Nat Genet, 2014. **46**(5): p. 503-9.

134.   Oda, H., et al., Aicardi-Goutieres syndrome is caused by IFIH1 mutations. Am J Hum Genet, 2014. **95**(1): p. 121-5.

135. Lahouassa, H., et al., SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. Nat Immunol, 2012. **13**(3): p. 223-8.
136. Lim, Y.W., et al., Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Aicardi-Goutieres syndrome. Elife, 2015. **4**.
137. Forstemann, K. and J. Lingner, Telomerase limits the extent of base pairing between template RNA and telomeric DNA. EMBO Rep, 2005. **6**(4): p. 361-6.
138. Camps, M. and L.A. Loeb, Critical role of R-loops in processing replication blocks. Front Biosci, 2005. **10**: p. 689-98.
139. Storici, F., RNA-mediated DNA modifications and RNA-templated DNA repair. Curr Opin Mol Ther, 2008. **10**(3): p. 224-30.
140. Fink, G.R., Pseudogenes in yeast? Cell, 1987. **49**(1): p. 5-6.
141. Niu, D.K., W.R. Hou, and S.W. Li, mRNA-mediated intron losses: evidence from extraordinarily large exons. Mol Biol Evol, 2005. **22**(6): p. 1475-81.
142. Gargouri, A., The reverse transcriptase encoded by ai1 intron is active in trans in the retro-deletion of yeast mitochondrial introns. FEMS Yeast Res, 2005. **5**(9): p. 813-22.
143. Murakami, E., et al., Characterization of novel reverse transcriptase and other RNA-associated catalytic activities by human DNA polymerase gamma: importance in mitochondrial DNA replication. J Biol Chem, 2003. **278**(38): p. 36403-9.
144. Trott, D.A. and A.C. Porter, Hypothesis: transcript-templated repair of DNA double-strand breaks. Bioessays, 2006. **28**(1): p. 78-83.
145. Krupovic, M., Recombination between RNA viruses and plasmids might have played a central role in the origin and evolution of small DNA viruses. Bioessays, 2012. **34**(10): p. 867-70.
146. Xu, P.Z., et al., Genome-wide high-frequency non-Mendelian loss of heterozygosity in rice. Genome, 2007. **50**(3): p. 297-302.
147. Suberbielle, E., et al., Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-beta. Nat Neurosci, 2013. **16**(5): p. 613-21.
148. Mehler, M.F., Epigenetics and the nervous system. Ann Neurol, 2008. **64**(6): p. 602-17.
149. Mattick, J.S. and M.F. Mehler, RNA editing, DNA recoding and the evolution of human cognition. Trends Neurosci, 2008. **31**(5): p. 227-33.
150. Aida, T., et al., Cloning-free CRISPR/Cas system facilitates functional cassette knock-in in mice. Genome Biol, 2015. **16**: p. 87.
151. Shen, Y. and F. Storici, Generation of RNA/DNA hybrids in genomic DNA by transformation using RNA-containing oligonucleotides. J Vis Exp, 2010(45).
152. Angeleska, A., et al., RNA-guided DNA assembly. J Theor Biol, 2007. **248**(4): p. 706-20.
153. Fang, W. and L.F. Landweber, RNA-mediated genome rearrangement: hypotheses and evidence. Bioessays, 2013. **35**(2): p. 84-7.
154. Kirkpatrick, D.P. and C.M. Radding, RecA protein promotes rapid RNA-DNA hybridization in heterogeneous RNA mixtures. Nucleic Acids Res, 1992. **20**(16): p. 4347-53.
155. Kirkpatrick, D.P., B.J. Rao, and C.M. Radding, RNA-DNA hybridization promoted by E. coli RecA protein. Nucleic Acids Res, 1992. **20**(16): p. 4339-46.
156. Hani, J. and H. Feldmann, tRNA genes and retroelements in the yeast genome. Nucleic Acids Res, 1998. **26**(3): p. 689-96.

157. Keskin, H. and F. Storici, Defects in RNase H2 Stimulate DNA Break Repair by RNA Reverse Transcribed into cDNA. Microrna, 2015. **4**(2): p. 109-16.

158. Symington, L.S., R. Rothstein, and M. Lisby, Mechanisms and regulation of mitotic recombination in Saccharomyces cerevisiae. Genetics, 2014. **198**(3): p. 795-835.

159. Kumar, A. and J.L. Bennetzen, Plant retrotransposons. Annu Rev Genet, 1999. **33**: p. 479-532.

160. Bannert, N. and R. Kurth, Retroelements and the human genome: new perspectives on an old relation. Proc Natl Acad Sci U S A, 2004. **101 Suppl 2**: p. 14572-9.

161. Kowalczykowski, S.C., Initiation of genetic recombination and recombination-dependent replication. Trends Biochem Sci, 2000. **25**(4): p. 156-65.

162. Moynahan, M.E. and M. Jasin, Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. Nat Rev Mol Cell Biol, 2010. **11**(3): p. 196-207.

163. Krogh, B.O. and L.S. Symington, Recombination proteins in yeast. Annu Rev Genet, 2004. **38**: p. 233-71.

164. Sung, P., Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. Science, 1994. **265**(5176): p. 1241-3.

165. Wei, L., et al., DNA damage during the G0/G1 phase triggers RNA-templated, Cockayne syndrome B-dependent homologous recombination. Proc Natl Acad Sci U S A, 2015. **112**(27): p. E3495-504.

166. Mortensen, U.H., M. Lisby, and R. Rothstein, Rad52. Curr Biol, 2009. **19**(16): p. R676-7.

167. Mortensen, U.H., et al., DNA strand annealing is promoted by the yeast Rad52 protein. Proc Natl Acad Sci U S A, 1996. **93**(20): p. 10729-34.

168. Wu, Y., T. Sugiyama, and S.C. Kowalczykowski, DNA annealing mediated by Rad52 and Rad59 proteins. J Biol Chem, 2006. **281**(22): p. 15441-9.

169. Grimme, J.M., et al., Human Rad52 binds and wraps single-stranded DNA and mediates annealing via two hRad52-ssDNA complexes. Nucleic Acids Res, 2010. **38**(9): p. 2917-30.

170. Kagawa, W., et al., Homologous pairing promoted by the human Rad52 protein. J Biol Chem, 2001. **276**(37): p. 35201-8.

171. Singleton, M.R., et al., Structure of the single-strand annealing domain of human RAD52 protein. Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13492-7.

172. Sigurdsson, S., et al., Basis for avid homologous DNA strand exchange by human Rad51 and RPA. J Biol Chem, 2001. **276**(12): p. 8798-806.

173. Henricksen, L.A., C.B. Umbricht, and M.S. Wold, Recombinant replication protein A: expression, complex formation, and functional characterization. J Biol Chem, 1994. **269**(15): p. 11121-32.

174. Rossi, M.J., et al., Analyzing the branch migration activities of eukaryotic proteins. Methods, 2010. **51**(3): p. 336-46.

175. Ludwig, D.L. and C.V. Bruschi, The 2-micron plasmid as a nonselectable, stable, high copy number yeast vector. Plasmid, 1991. **25**(2): p. 81-95.

176. Bi, B., et al., Human and yeast Rad52 proteins promote DNA strand exchange. Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9568-72.

177. Arai, N., et al., Vital roles of the second DNA-binding site of Rad52 protein in yeast homologous recombination. J Biol Chem, 2011. **286**(20): p. 17607-17.

178. Mortensen, U.H., et al., A molecular genetic dissection of the evolutionarily conserved N terminus of yeast Rad52. Genetics, 2002. **161**(2): p. 549-62.

179. Seong, C., et al., Molecular anatomy of the recombination mediator function of Saccharomyces cerevisiae Rad52. J Biol Chem, 2008. **283**(18): p. 12166-74.

180. Hanamshet, K., O.M. Mazina, and A.V. Mazin, Reappearance from Obscurity: Mammalian Rad52 in Homologous Recombination. Genes (Basel), 2016. **7**(9).

181. Feng, Q., et al., Rad52 and Rad59 exhibit both overlapping and distinct functions. DNA Repair (Amst), 2007. **6**(1): p. 27-37.

182. Chen, R., et al., Dynamic binding of replication protein a is required for DNA repair. Nucleic Acids Res, 2016. **44**(12): p. 5758-72.

183. Park, M.S., et al., Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. J Biol Chem, 1996. **271**(31): p. 18996-9000.

184. Davis, A.P. and L.S. Symington, The yeast recombinational repair protein Rad59 interacts with Rad52 and stimulates single-strand annealing. Genetics, 2001. **159**(2): p. 515-25.

185. Davis, A.P. and L.S. Symington, The Rad52-Rad59 complex interacts with Rad51 and replication protein A. DNA Repair (Amst), 2003. **2**(10): p. 1127-34.

186. Mimitou, E.P. and L.S. Symington, Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. Nature, 2008. **455**(7214): p. 770-4.

187. Zhu, Z., et al., Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. Cell, 2008. **134**(6): p. 981-94.

188. Cejka, P., DNA End Resection: Nucleases Team Up with the Right Partners to Initiate Homologous Recombination. J Biol Chem, 2015. **290**(38): p. 22931-8.

189. Plosky, B.S., The Good and Bad of RNA:DNA Hybrids in Double-Strand Break Repair. Mol Cell, 2016. **64**(4): p. 643-644.

190. Chakraborty, A., et al., Classical non-homologous end-joining pathway utilizes nascent RNA for error-free double-strand break repair of transcribed genes. Nat Commun, 2016. **7**: p. 13049.

191. Symington, L.S., End resection at double-strand breaks: mechanism and regulation. Cold Spring Harb Perspect Biol, 2014. **6**(8).

192. Rijkers, T., et al., Targeted inactivation of mouse RAD52 reduces homologous recombination but not resistance to ionizing radiation. Mol Cell Biol, 1998. **18**(11): p. 6423-9.

193. Lok, B.H., et al., RAD52 inactivation is synthetically lethal with deficiencies in BRCA1 and PALB2 in addition to BRCA2 through RAD51-mediated homologous recombination. Oncogene, 2013. **32**(30): p. 3552-8.

194. Prakash, R., et al., Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. Cold Spring Harb Perspect Biol, 2015. **7**(4): p. a016600.

195. Baumann, M., J. Pontiller, and W. Ernst, Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview. Mol Biotechnol, 2010. **45**(3): p. 241-7.

196. Rando, O.J. and F. Winston, Chromatin and transcription in yeast. Genetics, 2012. **190**(2): p. 351-87.

197. Garfinkel, D.J., et al., Ty1 Copy Number Dynamics in Saccharomyces. Genetics, 2005. **169**(4): p. 1845-1857.

198.   Lloyd, J.A., D.A. McGrew, and K.L. Knight, Identification of residues important for DNA binding in the full-length human Rad52 protein. J Mol Biol, 2005. **345**(2): p. 239-49.

199.   Katz, S.S., F.S. Gimble, and F. Storici, To nick or not to nick: comparison of I-SceI single- and double-strand break-induced recombination in yeast and human cells. PLoS One, 2014. **9**(2): p. e88840.

200.   Yang, L., et al., CRISPR-Cas-mediated targeted genome editing in human cells. Methods Mol Biol, 2014. **1114**: p. 245-67.