UNIVERSITY OF CALIFORNIA SAN DIEGO

A Review of R-Loop Genome-Wide Profiling via R-ChIP

A thesis submitted in partial satisfaction of the requirements
for the degree of Master of Science

in

Biology

by

Kevin Kai Sun

Committee in charge:

Professor Xiangdong Fu, Chair
Professor Dong-er Zhang, Co-chair
Professor Gen-Sheng Feng

2021

The thesis of Kevin Kai Sun is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT OF THE THESIS

A Review of R-Loop Genome-Wide Profiling via R-ChIP

by

Kevin Kai Sun

Master of Science in Biology

University of California San Diego, 2021

Professor Xiangdong Fu, Chair

Professor Dong-Er Zhang, Co-Chair

R-loops, DNA/RNA hybrids synthesized during transcription, have been causally associated with genome instability. Their dysregulation has been linked to the onset of myriad pathologies, most notably tumorigenesis. Though methodologies for R-loop isolation and quantification have been developed, it was not until recently that R-loops could be extracted reliably in vivo. With the newly developed R-ChIP technique, it is possible to isolate R-loop-protein complexes from K562 cells with high efficiency, which, when coupled with mass spectroscop y and deep sequencing, can provide details on R-loop-protein interactions at specific genomic loci. Further study can compare the results of our DNA/RNA hybrid interactome to that of DNA/RNA hybrid interactomes characterized by older S9.6-based methodologies. These newly identified interactions may shed important light on the mechanisms by which R-loops induce disease.

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Xiangdong Fu for his support as the chair of this committee. His wise words and guidance have proven an invaluable source of inspiration.

I would like to thank my mentor Xuan Zhang for her patient and helpful guidance in imparting her understanding of the subject and technical knowledge.

I would like to acknowledge the rest of the members of the Fu lab for their help and support.

Chapter 1

Introduction

R-loops are triple-stranded DNA/RNA hybrids that form during transcription, when an RNA strand exiting from the RNA channel of DNA polymerase II binds to a section of the complementary DNA strand, generating a displacement loop where both strands of a piece of DNA are separated for a stretch by the strand of RNA (1). These nucleic acid structures have been found to play impactful roles in various biological processes, including promotion of class-switch recombination (CSR) in activated B cells (2), impairment of mRNA elongation via RNA polymerase (3), impairment of DNA replisome action (4), termination of transcription (5), and recruitment of transcription factors (6). These functions have led researchers to find that R-loops have a large impact on genome (in)stability caused by DNA damage through transcription and replication aberrancies (7). Further research has shown that the genome instability phenotype of cells that acquire excess R-loops and undergo its related DNA damages are remarkably similar to the phenotype of cancer cells (7). In fact, recent studies have linked R-loop abnormalities to a multitude of diseases, including neurological diseases such as amyotrophic lateral sclerosis type 4 (ALS4) and ataxia oculomotor apraxia type 2 (AOA2), as well as other prominent diseases such as Prader-Willi syndrome, myelodysplastic syndrome, and cancer (8).

The detection and isolation of R-loops has long been a difficult process and was traditionally done with an *in vitro* approach using monoclonal antibody S9.6, a protein with high binding affinity to DNA/RNA hybrids (9). Fragmented genomic DNA would be exposed to S9.6, and DNA/RNA hybrids would be subsequently captured and deep-sequenced using DNA:RNA immunoprecipitation sequencing [DRIP-seq], or its derivatives (10). However, recent studies have found that S9.6 has a weak, but still significant, affinity for dsRNA as well (Additionally,

S9.6 has a targeting bias for R-loops based on their nucleotide composition, with certain sequences having higher binding affinities than others (12). In light of these flaws, researchers have started to pivot away from S9.6-based assays in favor of enzymes such as RNASEH1.

RNASEH1 is an endonuclease that catalyzes cleavage of RNA-DNA hybrids and is ubiquitous in nearly all living organisms (13). RNASEH1 was chosen for its simple singular polypeptide structure, allowing for ease of modification (14). In eukaryotes, it is typically involved in DNA replication, and the RNase H family of proteins is known to target and modify DNA/RNA hybrids (15). A derivative of DRIP-seq, DRIVE-seq (DNA:RNA *in vitro* Enrichment) utilizes a catalytically inactive RNASEH1 mutant to perform affinity pulldowns *in vitro*; however, it is very inefficient, possibly because of missing cofactors that would improve its targeting function *in vivo* (10). In order to more accurately capture R-loop interactions within human cells *in vivo,* the Fu lab developed the technique R-ChIP (Figure 1) (14). For the purposes of R-ChIP, a stable cell line expressing mutant *RNASEH1* must be established, with a nuclear localization signal (NLS) at the N terminus and a V5 recognition tag at the C terminus for immunoprecipitation (IP). The mutants generated in this experiment are the D210N mutant, which is catalytically inactive but binding competent; and the WKKD mutant, which has inactive catalytic and binding sites, thus serving as the control that is unable to recognize R-loops (14). Chromatin immunoprecipitation can then be done, followed by library construction and sequencing analysis of the V5-captured DNA/RNA hybrids (14). This technique can capture R-loops with better resolution than DRIP-seq, allowing for more precise pinpointing of R-loop formation locations (14, 16). Additionally, R-ChIP can identify human R-loop loci more accurately than DRIP-seq (14). Thus, R-loops captured by this method can be analyzed to yield information vital to human health.

Here, I attempt to give an impartial review of R-ChIP, as we attempted to use R-ChIP to identify R-loops found in human cells, including those that may have been previously missed by DRIP-seq-related methodologies. R-ChIP is able to provide us with the first detailed look into mapping human R-loop associated sites *in vivo*, verify protein factors previously found to associate with R-loops *in vitro*, as well as offer new insight into R-loop-protein interactions that may be loci-specific. It is our hope that the information we find helps elucidate the role of R-loops in biological processes and their mechanisms for disease onset.

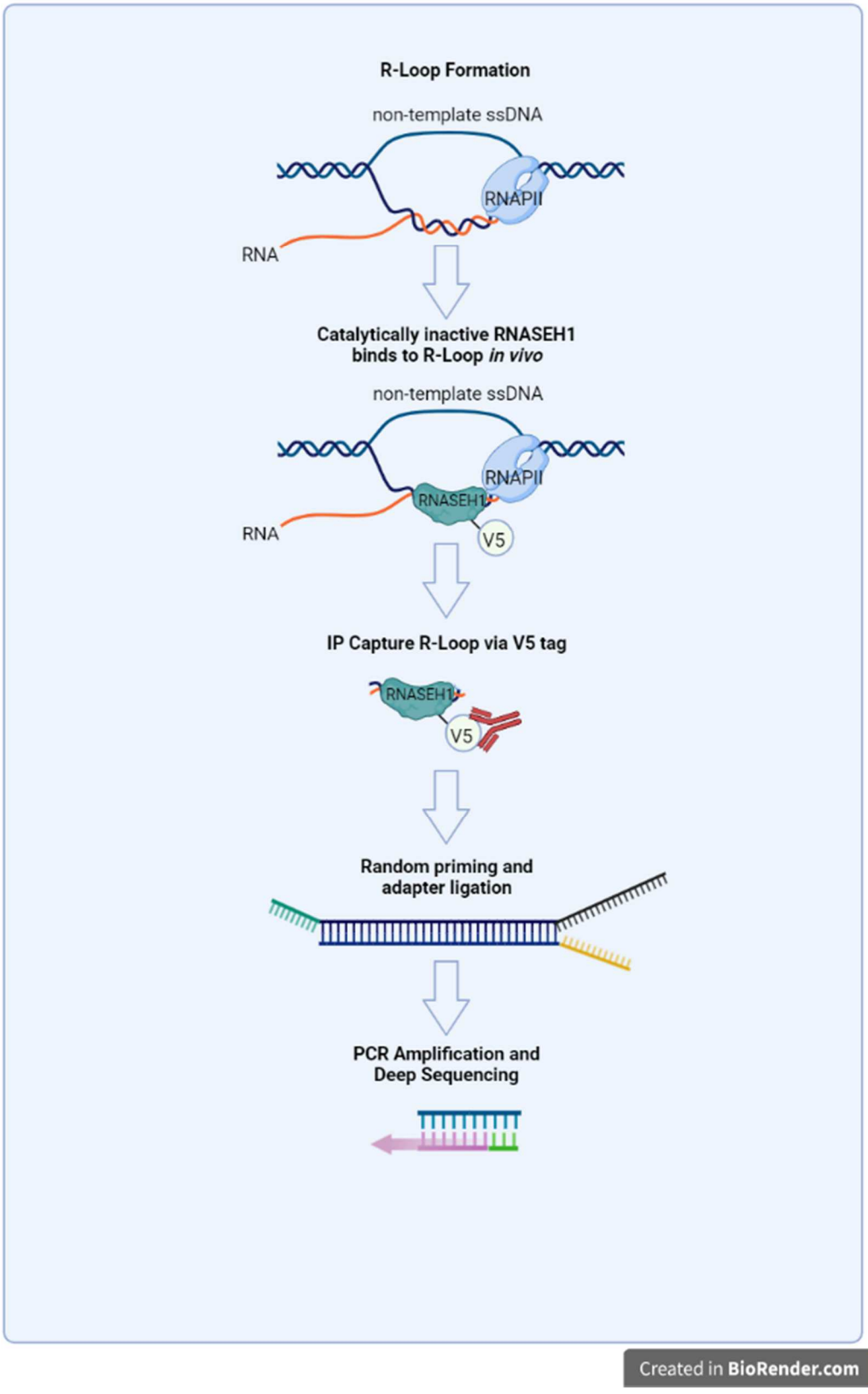

Figure 1. Schematic showing major steps of the R-ChIP workflow.

Chapter 2

Materials and Methods

2.1 Establishment of stable K562 cell lines expressing RNASEH1 mutants

Stable K562 cell lines with RNASEH1 mutant (D210N/WKKD) were obtained from a common laboratory stock (Dr. Xiangdong Fu's lab). K562 cells were cultured in 10-cm culture dishes with RPMI1640 with 10% FBS, 2 mM sodium pyruvate and 1 X penicillin-streptomycin. Cells were maintained and passaged every 2-3 days. For experiments, cells were seeded at 20-30% confluency and harvested at 70-80% confluency (cell count usually about 1×10^7 cells per plate).

2.2 Cross-linking and Harvesting of K562 cell lines

K562 cells were counted with a hemocytometer and about 1×10^7 cells were transferred to a 15mL conical, spun at 300g for 5 min at 4 °C, and pellet washed with 10mL cold PBS twice, before finally being resuspended in 10mL cold PBS. 270 μL of 37% (vol/vol) formaldehyde was then added to cells and resulting mixture was incubated at room temperature for 10min with constant mixing. Afterwards, 1 mL of 1.375M glycine was added to inactivate formaldehyde. Cells were then spun down at 600g for 5 minutes at 4 °C, and supernatant was aspirated and discarded. Cells were then washed with 10mL cold PBS three times, as done previously. Afterwards, the supernatant was aspirated and discarded, and pellet was resuspended in 1 mL cold PBS and transferred to 1.5 mL Lo-Bind tube. A final spin at 600g for 5 minutes at 4 °C was performed, and supernatant discarded leaving only the pellet.

2.3 Cell Lysis and Nuclei Isolation

Cell lysis buffer (10 mM Tris-Cl pH 8.0, 10 mM NaCl, and 0.5% Igepal CA-630) was prepared with 1X protease inhibitor cocktail and RiboLock RNase inhibitor (1 μL/mL final concentration). Cell pellet was resuspended in 1 mL cell lysis buffer and incubated on ice for 10 minutes, then spun at 700g for 5 minutes at 4 °C. Supernatant is discarded and 1 mL lysis buffer is used to resuspend pellet, which is then spun again at 700g for 5 minutes at 4 °C. Nuclear lysis buffer (50 mM Tris-Cl pH 8.0, 10 mM EDTA pH 8.0, and 1% SDS) was prepared with 1X protease inhibitor cocktail and RiboLock RNase inhibitor (1 μL/mL final concentration). Pellet was resuspended with 0.5 mL of nuclear lysis buffer and incubated for 10 minutes on ice.

2.4 Nucleic Acid Fragmentation

Nuclear suspension was sonicated in seven cycles using a probe sonicator (Sonifier Cell Disruptor; Branson, model no. 185). Each cycle consisted of 10 seconds of sonication, followed by a 1-minute resting period on ice. Fragmentation of nucleic acids was confirmed via standard DNA purification followed by a 2% agarose gel electrophoresis in 1X TAE buffer.

2.5 Immunoprecipitation and Western Blotting

For immunoprecipitation, 10 μg of anti-V5 antibody per sample was conjugated to protein A/G magnetic beads (Thermo Fisher) via incubation with gentle agitation for 1 hr at 4°C. Cellular debris was removed via centrifugation at 13000 rpm for 15 min at 4°C. The supernatant was mixed with antibody conjugated beads and incubated for 2 hr at 4°C with gentle agitation.

Beads were subsequently washed three times with lysis buffer and boiled for 5 min in 170 μL Laemmli sample buffer. The supernatant was collected, and 20 μL was used for western blotting. Western blotting was performed according to standard protocol. Briefly, proteins were resolved in 10% SDS-PAGE gels and transferred to a nitrocellulose membrane. After blocking for 1 hr with 5% non-fat milk in TBST buffer (50 mM Tris-HCl, 150 mM NaCl and 0.05% Tween 20 pH 7.6), the membrane was incubated with primary antibody for 1 hr at room temperature, followed by HRP-conjugated secondary antibody for 1 hr at room temperature. The membrane was finally incubated with ECL substrate (Thermo Fisher) for 5 minutes and imaged via autoradiography.

2.6 Reversal of Cross-linking and Recovery of IPed DNA

Remaining elution underwent reversal of crosslinking via alternating agitation overnight at 65°C in a Thermomixer. Samples were subsequently treated with appropriate amounts of RNase A and proteinase K, and IPed DNA was recovered with multiple liquid-liquid extractions of phenol and phenol:chloroform:isoamyl alcohol followed by 100% ethanol and sodium acetate precipitation. Pellet was washed twice with 100% ethanol, air dried at room temperature, and dissolved in 25 μL TE buffer. Concentration of DNA was then measured with a Qubit fluorometer.

2.7 qPCR Quality Check

10 μL per sample was taken and each diluted four-fold with molecular biology-grade water. qPCR on typical R-loop forming regions (Transcription Start Sites [TSS] regions of *JUN, NEAT1, CLSPN, PMS2)* was performed in triplicate for each 1 μL of diluted sample in 10 μL reaction

mixtures with FastStart Universal SYBR Green Master. qPCR was done with an initial denaturing step at 98°C for 10 minutes, followed by 40 cycles of denaturation at at 98°C for 15s, annealing and extension at 60°C for 45s, and ramping up from 60°C to 95°C at a rate of 2°C/s. Enrichment was confirmed via comparison of the TSS, expected R-loop formation regions, to corresponding TTS.

2.8 Primer Extension and Cleanup

DNA was mixed with primer extension mixture (11 µL DNA per sample, final concentrations of 1X phi29 DNA polymerase buffer, 0.2 mg/mL BSA, 0.15 mM dNTP mix, 1 µM N9 random primer [5′- /invddt/CAAGCAGAAGACGGCATACGAGNNNNNNNNN-3′] in 20 µL) for a total reaction mixture volume of 19 µL and incubated for 5 min at 95°C, followed by 5 min at 25°C. 1 µL of 10 U/µl phi29 DNA polymerase is then added and incubated in a thermal cycler for 20 min at 30°C, followed by 10 min at 65°C. The extension product was subsequently purified with PureLink PCR Micro Kit and eluted with 20 µL elution buffer.

2.9 dA Tailing and Cleanup

30 µL of dA tailing mixture (20 µL of DNA from end of **Section 2.8**, 1X NEB Buffer 2, 0.2 mM dATP, 0.16 U/µL Klenow fragment [3′→5′ exo-]) was prepared per sample, which were then incubated in a thermal cycler for 30 min at 37°C. The tailing product was finally purified with PureLink PCR Micro Kit and eluted with 13 µL elution buffer.

2.10 Adapter Ligation

Oligo A (5′-/Phos/GATCGGAAGAGCGTC GTGTAGGGAAAGAGTGT-3′) and Oligo B (5′-AGACGTGTGCTCTTCCGATCT-3′) were dissolved in 1X annealing buffer to a final concentration of 40 μM. 30 μL of each oligo were mixed in a 1:1 molar ratio to anneal adapters on a thermal cycler with these parameters: 2 min at 95°C, ramping down at 0.1 °C/s from 95 to 25°C. This resulted in a final adapter concentration of 20 μM, which was then diluted to 2 μM for use. 20 μL of ligation mixture (13 μL of DNA from end of **Section 2.9**, final concentrations of 1X ligation buffer, 0.1 μM adapter, 80 U/μL T4 DNA ligase) was prepared per sample and incubated for 2 hours at room temperature.

2.11 PCR Amplification and Library Prep

100 μL PCR reactions (16 μL of DNA sample, 1X Phusion HF buffer, 0.4 mM dNTP mix, 0.02 U/μL Phusion High-Fidelity DNA Polymerase, 0.2 μM Barcode primer [5′-AATGATACGGCGACCACCGAGATCTACACNNNNNACACTCTTTCCCTACACGACGC TCTTCCGATCT-3′], 0.2 μM PCR primer [5′-CAAGCAGAAGACGGCATACGAG −3′], 55 μL molecular-grade H_2O) were prepared per sample. PCR was performed on samples with an initial denaturation at 98°C, followed by 12 cycles of 10 sec denaturation at 98°C, 15 sec annealing at 60°C, and 15 sec extension at 72°C. PCR products were resolved on, and 130-350bp fragments were extracted from, a gel using a PureLink Quick Gel Extraction Kit.

2.12 Sequencing and Data Analysis

Deep sequencing was performed on Illumina HiSeq 2500 per manufacturer's instruction, and data analyzed as outlined by Chen et. al. (2017). Briefly, adapter sequences were removed via Cutadapt and poor base-calls trimmed in FastQC. Filtered reads were aligned to pre-built Bowtie 2, and only uniquely mapped reads with high (<30) mapping quality were kept. Potential PCR duplicates were removed with Picard. The resulting reads (.bam) were separated into two categories based on which strand they were mapped (Watson or Crick). Narrow peaks called separately for strand-specific reads with the MACS v.2. For visualization, .bam files were converted into bigwig files via genomeCoverageBed and bedGraphToBigWig (from UCSC Genome Browser utilities) and sequencing coverage normalized (usually to reads per million).

Chapter 3

Discussion

R-ChIP is a new and promising technique for extracting R-loops *in vivo*, with higher resolution and relevancy to human cells. R-ChIP has been demonstrated to accurately target DNA/RNA hybrids of size ranges resembling that of R-loop sizes identified via EM and bisulfate sequencing (2, 14, 17). R-loop regions mapped by R-ChIP also showed a common theme of G-rich clusters in non-template DNA, a key component of R-loop formation in which G quadruplex formation was thought to separate non-template DNA from template DNA (14, 17, 18). A majority (~60%) of R-loop regions mapped via R-ChIP were found to be near TSS promoter regions, with some (~16%) in intergenic and gene body regions, in contrast to those mapped via DRIP and its derivatives, which found a more even distribution of R-loops within TSS, gene body, and terminal regions (14). Notably, R-loop regions unique to DRIP and its derivatives demonstrated poor G/C skew enrichments and weak association with active chromatin marks (14). Additionally, it was found that elevated pausing of RNAPII at TSS promoted R-loop formation, though release of pausing did not lead to immediate remediation of R-loops (14). Another notable characteristic of R-ChIP mapped R-loops was the presence of a nearby free RNA end, thus leading to a proposed revision of how R-loop formation is thought to occur: newly emerged RNA from an elongating RNA polymerase only anneals to a sufficiently stable G/C-skewed DNA sequence and elongates further upstream until blocked by the G quadruplex originally thought to initiate R-loop formation (14, 18).

When comparing the profiles of R-loops mapped via R-ChIP to those of DRIP-related assays, only ~1/3 of peaks overlapped (14). This could be due to S9.6 binding to RNA/RNA hybrids as opposed to true R-loops (11); S9.6 binding to other DNA/RNA hybrids associated

with replication or immunoglobulin class-switching (2); or it could be that there exist R-loops that can only be targeted by S9.6 but not RNASEH1, and vice versa. Indeed, recent research using qDRIP, a variant of DRIP, has identified low-level DNA/RNA hybrid signals persisting at terminator regions after RNase H treatment, indicating there may exist R-loops that are resistant or are impervious to RNase H interaction (19). It should be noted that RNASEH1's function as an endonuclease may yield biased results towards the identification of unstable or transient R-loops. Consequently, a complementary approach to R-loop analysis combining both R-ChIP and DRIP may shed insight on R-loop stability at specific loci. It also brings up considerations of whether there exist additional R-loops that cannot be targeted by S9.6 nor RNASEH1 but have affinity for other unknown proteins. This could explain why R-ChIP found almost no R-loop formation spots in terminator regions, where Senataxin, a helicase, has been found to be more active than RNASEH1 (14, 20). These terminator regions may have transient R-loop activity that is hard to detect with current assays due to the dynamic action of Senataxin and other helicases, which may swiftly disassemble R-loops in these regions. A technique that uses Senataxin as the recognition protein instead of S9.6 or RNASEH1 might be able to identify even more novel R-loop associated sites and interactors.

Previous *in vitro* techniques utilizing S9.6 antibody, DRIP, have managed to characterize a DNA/RNA hybrid interactome of about 450 proteins (21), but considering the improved capabilities of R-ChIP, a novel characterization of the R-loop interactome would be helpful for expanding our knowledge of these enigmatic molecules. One possible way to accomplish this could be through UV crosslinking followed by capture of mutant RNASEH1, silver staining, western blot, and mass spectrometry (Figure 2). R-loop interactomes generated via R-ChIP could be compared to those generated via DRIP and its derivatives.

Despite the insight R-ChIP provides into R-loops, it is not without flaws. Establishing a stably expressed mutant RNASEH1 *in vivo* is a time-consuming and delicate task, since over-expression of RNase H enzymes can be cytotoxic, and this process may not be feasible in all cell types (13). Additionally, R-ChIP requires significant amounts of starting material, making its adaptation more limited.

Overall, R-ChIP is a significant new tool in the field R-loop research. Already, its use has led to the identification of high-risk splicing factor mutations in SRSF2 and U2AF1 in myelodysplastic syndromes (22). The information R-ChIP can provide in conjunction with other techniques will yield invaluable information on the nature of R-loops.

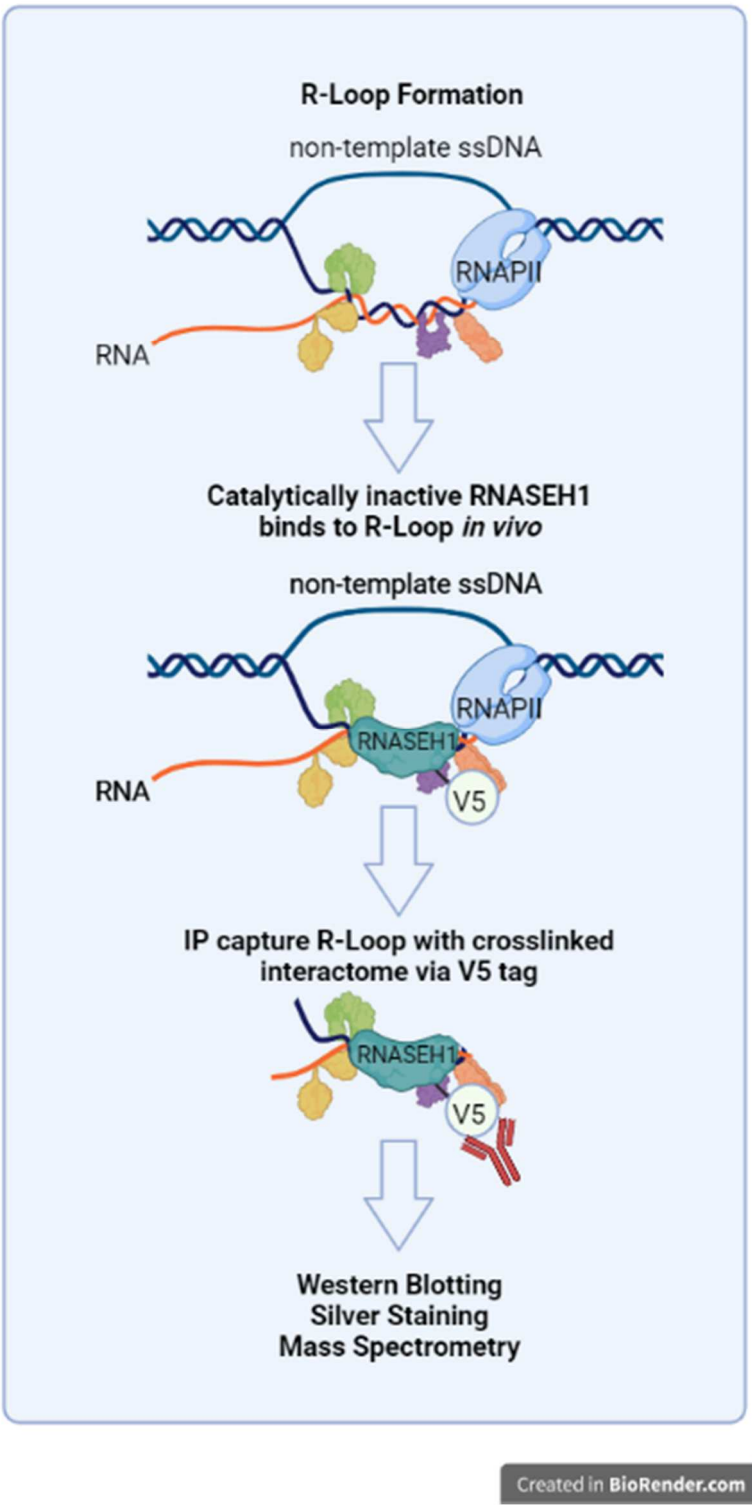

Figure 2. Schematic showing possible workflow of R-loop interactome characterization via R-ChIP.

Chapter 4

References

1. Thomas M, White RL, Davis RW. (1976). Hybridization of RNA to double-stranded DNA: formation of R-loops. Proceedings of the National Academy of Sciences of the United States of America, 73(7), 2294–2298. https://doi.org/10.1073/pnas.73.7.2294

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

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

4. Tuduri S, Crabbé L, Conti C, Tourrière H, Holtgreve-Grez H, Jauch A, Pantesco V, De Vos J, Thomas A, Theillet C, Pommier Y, Tazi J, Coquelle A, Pasero P. (2009). Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nat. Cell Biol., 11(11), 1315-1324. doi: 10.1038/ncb1984.

5. Skourti-Stathaki K, Kamieniarz-Gdula K, Proudfoot NJ. (2014). R-loops induce repressive chromatin marks over mammalian gene terminators. Nature, 516(7531), 436-439. DOI: 10.1038/nature13787.

6. Chen PB, Chen HV, Acharya D, Rando OJ, Fazzio TG. (2015). R loops regulate promoter-proximal chromatin architecture and cellular differentiation. Nature Structural & Molecular Biology, 22(12), 999-1007. DOI: 10.1038/nsmb.3122.

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

8. Groh M, Gromak N. (2014). Out of balance: R-loops in human disease. PLoS genetics, 10(9), e1004630. https://doi.org/10.1371/journal.pgen.1004630

9. El Hage A, French SL, Beyer AL, Tollervey D. (2010). Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. Genes & development, 24(14), 1546–1558. https://doi.org/10.1101/gad.573310

10. Ginno PA, Lott PL, Christensen HC, Korf I, Chédin F. (2012). R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Molecular cell, 45(6), 814–825. https://doi.org/10.1016/j.molcel.2012.01.017

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

12. König F, Schubert T, & Längst G. (2017). The monoclonal S9.6 antibody exhibits highly variable binding affinities towards different R-loop sequences. PloS one, 12(6), e0178875. https://doi.org/10.1371/journal.pone.0178875

13. Shen W, Sun H, De Hoyos CL, Bailey JK, Liang XH, Crooke ST. (2017). Dynamic nucleoplasmic and nucleolar localization of mammalian RNase H1 in response to RNAP I transcriptional R-loops. Nucleic acids research, 45(18), 10672–10692. https://doi.org/10.1093/nar/gkx710

14. Chen L, Chen JY, Zhang X, Gu Y, Xiao R, Shao C, Tang P, Qian H, Luo D, Li H, Zhou Y, Zhang DE, & Fu XD. (2017). R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with

Transcriptional Pausing at Gene Promoters. Molecular cell, 68(4), 745–757.e5. https://doi.org/10.1016/j.molcel.2017.10.008

15. Cerritelli SM, Crouch RJ. (2009). Ribonuclease H: the enzymes in eukaryotes. The FEBS journal, 276(6), 1494–1505. https://doi.org/10.1111/j.1742-4658.2009.06908.x

16. Chen JY, Zhang X, Fu XD, & Chen L. (2019). R-ChIP for genome-wide mapping of R-loops by using catalytically inactive RNASEH1. Nature protocols, 14(5), 1661–1685. https://doi.org/10.1038/s41596-019-0154-6

17. Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N. (2004). Intracellular transcription of G-rich DNAs induces formation of G-loops, novel strctures containing G4 DNA. Genes Dev. 18, 1618-1629.

18. Stork CT, Bocek M, Crossley MP, Sollier J, Sanz LA, Chedin F, Swigut T, Cimprich KA. (2016). Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. eLife *5*. Published online August 23, 2016. https://doi.org/10.7554/eLife.17548.

19. Crossley M, Bocek M, Hamperl S, Swigut T, Cimprich K. (2020). qDRIP: a method to quantitatively assess RNA–DNA hybrid formation genome-wide. Nucleic Acids Research, 48(14), e84. Published online August 20, 2020. https://doi.org/10.1093/nar/gkaa500

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

21. Cristini A, Groh M, Kristiansen MS, Gromak N. (2018). RNA/DNA Hybrid Interactome Identifies DXH9 as a Molecular Player in Transcriptional Termination and R-Loop-Associated DNA Damage. Cell reports, 23(6), 1891–1905. https://doi.org/10.1016/j.celrep.2018.04.025

22. Chen L, Chen JY, Huang YJ, Gu Y, Qiu J, Qian H, Shao C, Zhang X, Hu J, Li H, He S, Zhou Y, Abdel-Wahab O, Zhang DE, Fu XD. The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. Mol Cell. 2018 Feb 1;69(3):412-425.e6. doi: 10.1016/j.molcel.2017.12.029. Epub 2018 Jan 27.