INVESTIGATION OF LONG NON-CODING RNA AND CHROMATIN INTERACTIONS IN HeLa CELLS

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Molecular Biology and Genetics

by
Melis ATBİNEK

July 2022
İZMIR

ACKNOWLEDGEMENTS

First, I would like to express my gratitude to my PI, Prof. Dr. Bünyamin AKGÜL. I feel so fortunate that he has trusted me with this project and let me join his lab in a short notice. I want to express my appreciation to TUBITAK – COST for funding this project (119Z611). I am also grateful for my committee members Assoc. Prof. Özden Yalçın ÖZUYSAL and Assoc. Prof. Yavuz OKTAY for their time and support for my thesis.

This endeavor would not have been possible without our lovely Dr. İpek ERDOĞAN VATANSEVER. She is a huge support for our lab, and I see her as my guardian angel in science. I cannot imagine a lab without her. Also, I am extremely grateful to Dr. Özge TUNCEL who supports us anytime. She has an amazing visionary in mechanics and chemistry. It is a great opportunity to be able to work with her.

Thanks, should also go to my lab members Dilek Cansu GÜRER, Ayşe Bengisu GELMEZ, Bilge YAYLAK, Yusuf Cem ÇİFTÇİ, Vahide İlayda KAÇAR and Şirin Elife CEREN who have supported me through this journey and let me occupy RNA isolation for two weeks. And for the remaining of my colleagues, Buket SAĞLAM was a great dance partner for our short breaks in between experiments and thesis writing. Azime AKÇAÖZ ALASAR was both a great mentor and support through this period. Last but most definitely not least, I would like to thank my partner in crime, Merve KARA, for accompanying me during my Master's in which we faced many difficulties but had also a lot of fun.

Words cannot express my gratitude to my one and only Cemoşko, Deniz Cemre TURGUT. I must admit there is not any other person who can calm me down like she does. Also, her knowledge in science and her vision of art helped me complete my Master's study overall. I am extremely grateful for my boyfriend, Onur AKGÜNEL, who faced all the craziness and did not stop supporting me for a second. Finally, I am incredibly lucky to have the world's most perfect parents Hamit ATBİNEK and Ayşe Füsun BEŞGÜRBÜZ ATBİNEK. Your support can make any person a superhero. Also, Piko, our little marshmallow, she helped me to get through my thesis with unlimited support.

ABSTRACT

INVESTIGATION OF LONG NON-CODING RNA AND CHROMATIN INTERACTIONS IN HeLa CELLS

The DNA in the cells is surrounding histone proteins to form nucleosomes. The structure is packed further into chromatin. The chromatin structure is dynamic and flexible. It is regulated by many factors including long non-coding RNAs (lncRNAs). LncRNAs are a class of non-coding RNAs, transcripts that do not encode protein. They are longer than 200 nucleotides and might contain a polyA tail and a 5' cap. Thus, they are localized in the nucleus. lncRNAs interact with chromatin in two ways, indirect and direct. Direct interaction occurs via two mechanisms: R-loop and triplex formation. These interactions affect the folding of chromatin inducing gene expression under various cellular conditions. LncRNAs interacting with chromatin regulating genes are found in HEK cells. Thus, it is hypothesized that lncRNA – chromatin interactions may differ in cancerous cells as well. In this study, the iMARGI method is optimized to be used in adenocarcinoma HeLa cells. The chromatin digestion and incubation conditions are adjusted to give optimal results for HeLa cells. iMARGI is a recently developed method employed to investigate such interactions in a genome-wide manner. iMARGI allows the isolation of all lncRNAs interacting with the whole genome. The interacting RNA – DNA molecules are pulled down with streptavidin conjugated beads after linker ligation. The chimeric molecules are amplified with PCR forming lncRNA – chromatin libraries of HeLa cells. In the future, new libraries can be formed after inducing apoptosis in HeLa cells. Identification of lncRNAs involved in chromatin remodeling in apoptotic conditions can facilitate new therapeutic methods for fighting tumor initiation and development.

Keywords: *LncRNA, Chromatin, iMARGI, Cervical cancer*

ÖZET

HeLa HÜCRELERİNDE UZUN KODLANMAYAN RNA VE KROMATİN ETKİLEŞİMLERİNİN ARAŞTIRILMASI

DNA hücre içerisinde histon proteinlerinin etrafına sarılı halde bulunur ve bu yapılar katlanarak nükleozomları oluşturur. Ardından tekrar katlanma gerçekleşerek kromatin yapısı oluşturulur. Bu yapının regülasyonu genlerin okunması için çok önemlidir. Hücrenin durumuna göre katlanmalarda değişiklikler gerçekleşir. Kromatin regülasyonunda uzun kodlamayan RNA'lar (ukmRNA) görev alır. ukmRNA'lar protein kodlama potansiyeli olmayan 200 nükleotitten daha uzun transkriptlerdir. Bu moleküller polyA kuyrukları ve 5' başlarını içerir. Böylece, çekirdekte yer alırlar. ukRNA'lar kromatinle dolaylı ya da direkt yoldan etkileşir. Bu direkt etkileşimler sırasında R-loop ya da triplex yapıları oluşur. Önceki çalışmalarda HEK hücrelerinde ukmRNA'ların kromatin katlanmasında görev alan DNA'lar ile etkileştiği görülmüştür. Kanser hücrelerinde de bu etkileşimler görülebilir. Bu bağlamda iMARGI metodu kullanarak adenokarsinom HeLa hücrelerinde bütün kromatinle etkileşen tüm ukmRNA'ların belirlenmesi hedeflenmiştir. Bu metot kapsamında birbiriyle etkileşim halinde olan DNA ve RNA'lar biyotin içeren bir linker molekülü tarafından birleştirildi. Yeni oluşturulan kompleks streptavidin – biotin etkileşimi ile elde edilmiştir. Sonuç örneği PCR ile çoğaltılarak HeLa hücrelerinde kromatinle etkileşen RNA'ların kütüphanesi oluşturulabilecektir. Gelecekte apoptoz indüklenmiş hücrelerden kütüphaneler oluşturabilir. Bu kütüphanelerden bulunan ukmRNA'lar tedavi amaçlı kullanılma potansiyeline sahiptir.

Anahtar Kelimeler: *UkmRNA, Kromatin, iMARGI, Servikal kanser*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table **Page**

CHAPTER 1

INTRODUCTION

1.1 Chromatin

DNA is packed around histone proteins forming the chromatin structure. Due to electrostatic forces negatively charged DNA is easily bound to positively charged histones forming nucleosomes. They contain two copies of each histone protein: H2A, H2B, H3 and H4 (Mariño-Ramírez et al. 2005).

DNA in between nucleosomes is called linker DNA and is protected by another histone protein, H1 (Tremethick 2007). Neighbor nucleosome interaction leads up to chromatin fiber formation as seen in Figure 1.1-1. Chromatin fibers came together to form a higher level of chromatin organization, chromosomes (Olins and Olins 2003). Chromosomes are formed during cellular division for the equal and complete segregation of genetic materials (Hirano 2015).

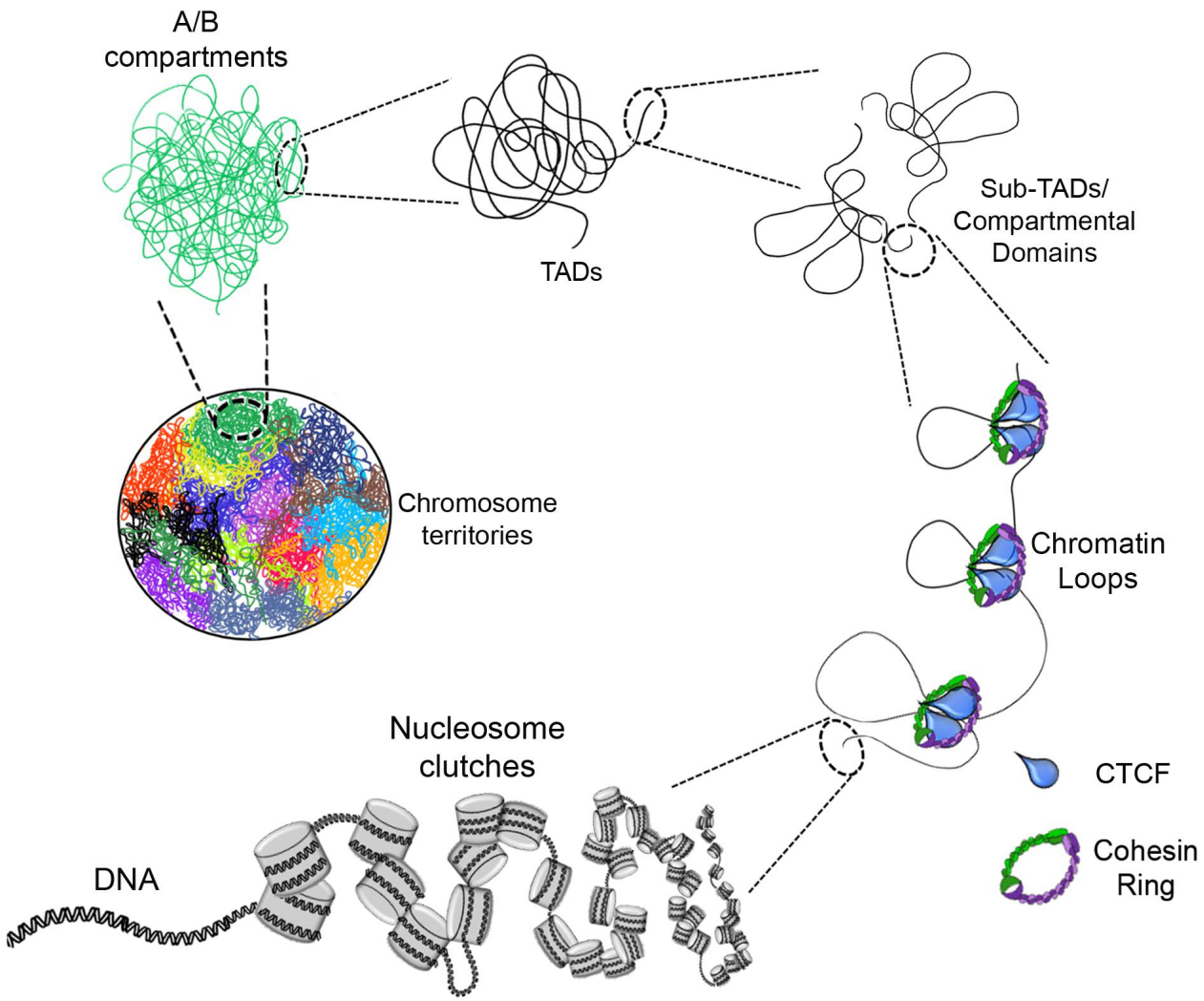

Figure 1.1 The different levels of DNA packing into chromosome territories. DNA is packed into smaller regions via nucleosome binding and their interactions. Further folding occurs with CTCF and Cohesin rings, creating chromatin loops. These domains may come together to form topologically associated domains (TADs). The compartments are also subcategorized as active or not, A/B (Source: Furlan-Magaril et al. 2009)

1.1.1 Structure & Architecture

Chromatin can be found in two different states in eukaryotes. A highly packed, condensed form where transcription is inactive, *heterochromatin*; and a less condensed form with active transcription sites, *euchromatin* (Lee et al. 2020). If exist in more permanent state heterochromatin is called constitutive heterochromatin (CH). It is seen in gene-poor regions, transposons, pericentromeric and telomeric sequences (Rea et al. 2000). Methylation on histone H3 is the marker of CH, implemented by histone methyltransferase (Peters et al. 2001). Facultative chromatin (FC) regions can switch into euchromatin form in some circumstances. Developmental genes are cited in these regions. Polycomb repressive complexes (PRC) are found in FC and mediate methylation (R. Cao et al. 2002). Nonetheless, sparse RNA synthesis occurs in heterochromatin regions which in return regulate heterochromatin maintenance (Chu et al. 2015).

Loosely packed, transcriptionally active sites are called euchromatin. Nucleosomes are located further away allowing wider regions for transcriptional machinery (Radman-Livaja and Rando 2010a). Nucleosomes are positioned according to the DNA sequence beneath them (Radman-Livaja and Rando 2010b). Euchromatin regions conclude actively transcribed genes and regulatory elements including promoters and enhancers (Morrison and Thakur 2021).

In the genome, function and structure are mechanically associated to allow genes located far away to interact (Sikorska and Sexton 2020). These domains are called topologically associated domains (TAD) (Figure 2-1). TADs are conserved among related species and stable throughout many cell divisions (Nora et al. 2012). They are acquired by ring-shaped protein complexes, cohesin (Hansen et al. 2018). The ring shape brings two chromatins together by surrounding them as seen in Figure 2-1. This event leads to long-range interaction between genes or enhancers and promoters. Nevertheless, there are also molecules for insulation of materials between neighbor TADs. CTCF, a zinc finger DNA binding protein, is an insulator element (Kim, Yu, and Kaang 2015). It prevents unwanted promoter enhancer interactions and specifies euchromatin/heterochromatin regions

(Ghirlando and Felsenfeld 2016). CTCFs are especially enriched at TAD boundaries. These boundaries are rich for CTCF binding sites (Dixon et al. 2012).

1.1.2 Regulation

The change in the folding of chromatin is believed to have a high impact on transcription choice. Thus, it makes the regulation of architecture crucial in gene regulation. Histones are key molecules involved in chromatin architecture and folding. Alterations on histones regulate chromatin overall. These post-transcriptional modifications can affect the interaction between adjacent nucleosomes (Bannister and Kouzarides 2011). Acetylation on lysine amino acids of histones neutralizes the positive charge on Lysine, minimizing the electrostatic attraction between DNA and histones (Parthun 2007). Phosphorylation on the other hand occurs on threonines, serines, and tyrosines. When these amino acids are phosphorylated via histone kinases, they become negatively charged in turn affecting the association with DNA (Oki, Aihara, and Ito 2007).

Lysine and arginine amino acids can go through methylation. In contrast to prior modifications, instead of changing the charge of amino acids, they introduce new functions to histones (Bedford and Clarke 2009).

1.1.3 Chromatin Regulation and Cancer

Chromatin architecture is controlled by a variety of molecules as its regulation has a key role in cellular activity. Disruption in chromatin regulation can lead to organizational failure in the chromatin which can have detrimental effects, including diseases (Tessarz and Kouzarides 2014). Interruption in structural maintenance and genome stability is one of the main causes of cancer (Hanahan and Weinberg 2011). Mutated chromatin modifiers alter

gene expression in favor of cancerous cells (Martín-Subero 2011). Histone modifications have a huge potential to become prognostic and diagnostic markers of cancer.

In cancer cells, mutations on the enzymatically active unit of the PRC complex are observed. These mutations disrupt methylation and alter chromatin compaction (Kraft et al. 2022). Alterations in chromatin compaction allow induction of oncogenes leading to cancer development. Genes encoding for cohesin proteins are another target of mutations in cancerous cells which leads to tumor initiation and development (Waldman 2020). Abnormal expression of lncRNAs and miRNAs are detected in cancer cells, making them prominent biomarkers for different cancer types (Arun, Diermeier, and Spector 2018).

1.2 Cervical Cancer

Cervical cancer is one of the most diagnosed cancers in women after breast cancer. Also, it is the leading cause of cancer death in 36 countries (Sung et al. 2021). It is described as cancerous tissue in the cervix, a region between the vagina and womb. The cause is almost all the time human papillomavirus, HPV. Thus, risk factors increase with HPV infection. Histological subtypes of cervical cancer are predominantly squamous cell carcinoma and adenocarcinoma (Cohen et al. 2019).

In cervical cancer, epigenetics has a crucial role in tumor progression. It makes epigenetic modification a great candidate for screening programs (Deldar et al. 2021). In the presence of HPV infection chromatin modification enzymes are upregulated (J. Fang, Zhang, and Jin 2014). Mutated chromatin modifiers alter gene expression in favor of cancerous cells (Martín-Subero 2011). Histone H3K27 causes downregulation of DNMT3A leading to overexpression of HAVCR2 and LGALS9, inducing immune tolerance to overexpression of cervical tumor cells (Paul, Pillai, and Kumar 2021). Likewise, non-histone proteins can play a role in tumor progression. PCR2 complex, downregulates the expression of E-cadherin in association with histone deacetylase 1 (C. Wang et al. 2013). Hypermethylation on the promoter region of Septin-9 acts as a biomarker for cervical cancer (Jiao et al. 2019).

Although the exact mechanism is not yet known, there are lncRNAs involved in oncogenesis. NEAT1 lncRNA is observed in P53 activation in cervical cancer which is enriched in the nucleus. MALAT1 is known to induce metastasis by regulating epithelial-mesenchymal transition (Sun et al. 2016).

1.3 lncRNAs

1.3.1 Definition

Most genes are encoded into transcripts which are not translated into functional units. These transcripts were accepted as "junk" for a long time. However, their function in gene regulation was discovered over time. They were named after their structural or functional features. Non-coding transcripts longer than 200 nucleotides are called long non-coding RNAs (S. Fang et al. 2018). This group includes lncRNAs transcribed by RNA polymerase II (RNA pol II), lncRNAs transcribed from intergenic regions or sense, and antisense transcripts overlapping genes (Yao et al., 2019).

1.3.2 Localization in the Nucleus

lncRNAs mostly contain a 5' end cap and 3' end poly(A) tails (Statello et al., 2021). However, in contrast to mRNAs, they tend to localize in the nucleus. New findings started to reveal the logic behind this distinct difference. lncRNAs have less exons which are longer in length (Uszczynska-Ratajczak et al. 2018). The nuclear localization, fate, of lncRNAs is determined by many factors involved in different levels of regulation. lncRNAs, transcribed by phosphorylated RNA pol II, have low co-transcriptional splicing activity and their termination occurs independently of polyadenylation signals (Nojima and Proudfoot, n.d.).

Thus, these lncRNAs tend to accumulate on chromatin leading to their degradation by RNA exosomes (Nojima and Proudfoot, n.d.).

lncRNAs avoiding degradation are hypothesized to be functional and they can also survive nuclear surveillance (Statello et al., 2021). Some lncRNAs might not need to go through nuclear surveillance as their localization on chromatin occurs via U1 small nuclear ribonucleoprotein (U1 snRNP) (Yin et al. 2020). Likewise, chromatin interaction via R-loops drives nuclear localization. This interaction causes the depletion of SPT6 (Vos et al. 2018). Its absence induces the transcription of lncRNAs as activation by methylation shifts from coding genes to non-coding ones (Nojima et al. 2018).

lncRNAs are weakly spliced in comparison to mRNAs because they have weaker internal splicing factors with a longer distance between the 3' splice site and branch point (Melé et al. 2017). Splicing factors are differentially expressed in lncRNAs contributing to the difference in splicing (Guo et al. 2020).

Another structural feature of lncRNAs causing nuclear retention is sequence motifs which are interacting with nuclear factors (Azam et al. 2019). Also, nuclear proteins can interact with repeat-derived sequences on lncRNAs such as Alu repeats or functional intergenic repeating RNA element (*FIRRE*) (Hacisuleyman et al. 2016).

1.3.3 Chromatin Regulation

Negatively charged lncRNAs can neutralize positively charged histone proteins and cause the unfolding of the chromatin structure (Dueva et al. 2019). Upon interaction, activation or repression of gene expression can be observed. Firstly, lncRNAs interact with proteins either preventing or inducing their binding to DNA. For example, lncRNA *ANRIL* interacts with PRC1 and PRC2 and directs them to neighbor genes or distal genes via *Alu* repeats (Holdt et al. 2013). Chromatin modifiers can also be mediated by lncRNAs to promote gene expression (K. C. Wang et al. 2011). lncRNA HOTTIP recruits modifiers to promoters of HOXA genes via chromatin looping (Fanucchi et al., 2019).

lncRNAs can directly interact with DNA to regulate chromatin. These interactions can be listed as R-loop or Triple helix, depending on the form of interaction. A triplex helix is formed between double-stranded DNA and RNA molecules (Li, Syed, and Sugiyama 2016). RNA forms a Hoogsteen hydrogen bond with purines on DNA strands in both directions, anti-parallel, and parallel (Kuo et al. 2019). Sphingosine kinase 1 *(SPHK1),* a proto-oncogene, is regulated by its enhancer region forming a triplex with antisense lncRNA *KHPS1* (Blank-Giwojna, Postepska-Igielska, and Grummt 2019).

R-loops are formed when there is base-pairing with a DNA strand and it displaces ssDNA (Niehrs and Luke 2020b). R-loops are another way of lncRNAs contributing to gene regulation. Also, they can initiate DNA repair mechanisms by the R loop processing factor CtIP (Ngo, Grimstead, and Baird 2021). R loops regulate chromatin, acting as an epigenetic marker. They bind to specific DNA sequences and they are recognized by R-loop-associated proteins (Niehrs and Luke 2020a). The recognition mechanism by proteins is unknown. lncRNA *TARID* forms an R loop with the promoter of *TCF21* in mouse embryonic stem cells. GADD45A recognizes the structure and recruits DNA demethylating factor TET1, resulting in the activation of *TCF21* (Arab et al. 2019)

1.4 iMARGI

1.4.1 Genome-wide Approaches

The starting point of approaches investigating lncRNA – chromatin interactions is protein-centric. Initially, lncRNAs interacting with chromatin modifiers were identified (Mishra and Kanduri 2019). However, in this perspective lncRNA chromatin association was not guaranteed. Then, one to all approaches came along where a specific RNA is featured. The association of this RNA in the genome is investigated. In these methods, biotinylated oligos are hybridized to targets (Simon et al. 2011a). An upgraded approach includes antisense RNA hybridization for greater specificity (Engreitz et al. 2013).

Non-RNA-centric methods target all chromatin-associated RNAs in the cell identifying global targets (Simon et al. 2011b). In these methods, cells are fixed by crosslinking in different conditions depending on the protocol. The interacting molecules are obtained by proximity ligation of linkers (Bonetti et al. 2020).

1.4.2 iMARGI

iMARGI (in situ mapping of RNA–genome interactome) is an all versus all method where the main goal is to identify all RNAs associated with the chromatin. The DNA-RNA interacting molecules are attained by pulling down oligomers ligated to DNA on one site and RNA on the other. In this protocol, mammalian cells can be used. The number of cells can vary but 10 million is recommended for the initial amount. Cells are treated with 16% formaldehyde, diluted to 1%, for crosslinking which aims to fix RNA associated with chromatin. Nuclei are isolated from cells and permeabilized in the second stage. Cells are homogenized with a mild non-ionic detergent to avoid breaking the nucleus membrane. RNA and chromatin fragmentation is achieved by RNase 1 and Alu 1 inside the nucleus.

Nuclei are isolated from cells and permeabilized in the second stage. Cells are homogenized with a mild non-ionic detergent to avoid breaking the nucleus membrane. RNA and chromatin fragmentation is achieved by RNase 1 and Alu 1 inside the nucleus.

The most critical step of iMARGI is proper ligation of oligomers, linkers to DNA and RNA inside an intact nucleus. There are two linkers named top and bottom which are complementary. These linkers are custom-made to include a BamHI restriction site.

Likewise, they are complementary to NEBNext Universal and Index primers for library generation steps (Wu et al. 2019). The top linker is also biotinylated allowing isolation via streptavidin-biotin association.

Before the annealing of linkers, the top linker is adenylated. RNA molecules' 3' phosphate groups are converted to 3' hydroxyl groups to be ligated to adenylated top strand. DNA is prepared for ligation by the addition of A base creating an exposed 3' end.

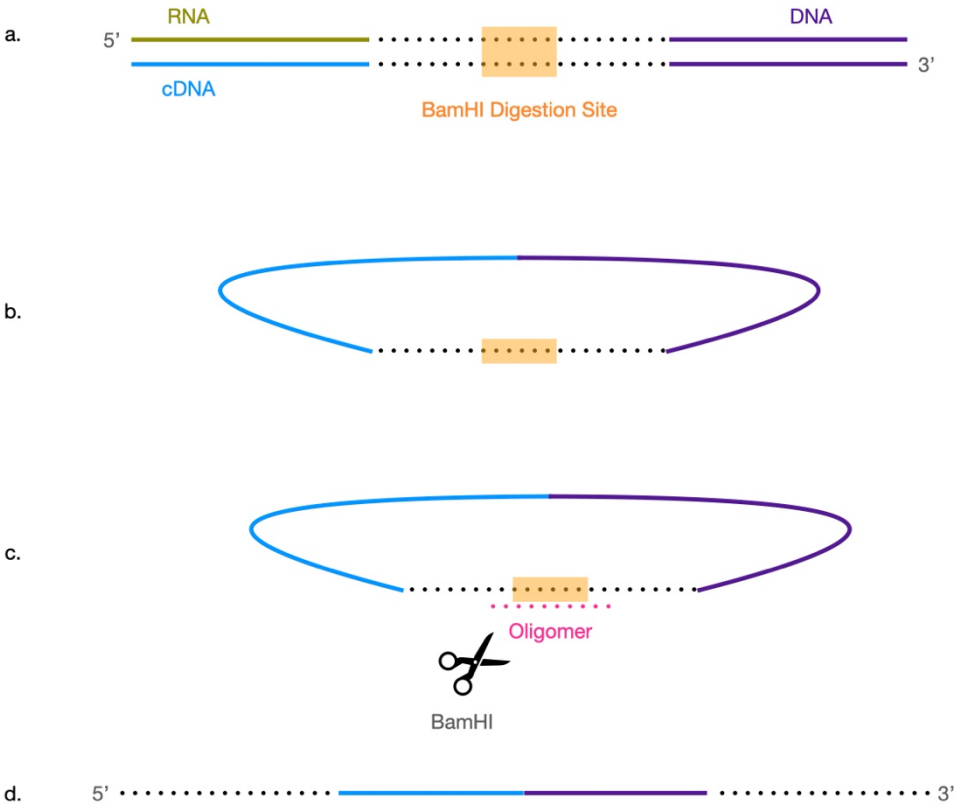

Figure 1.2. The workflow of iMARGI after isolation of linkers via streptavidin beads. a) Isolated chimeric sample with linkers in the middle. b) After circularization of the molecule. c) Custom-made oligomer is annealed on the BamHI site. The sample is re-linearized with the BamHI restriction enzyme. d) Linearized sample with linkers on each end and chimeric DNA - RNA sample in the middle.

DNA is ligated to the other side by sticky end ligation. Unligated linkers are removed by exonuclease after precipitation of nucleic acids. Desired DNA-linker-RNA molecules are harvested via streptavidin beads. The opposite strand of RNA is completed by reverse transcription. Newly-made strand containing cDNA is collected and circularized. Circularization is a crucial step to define the orientation of DNA and RNA. Another custom-made oligomer is annealed to the linker site, forming a BamHI digestion site with linkers (Figure 1.2a). The circular sample is then linearized by the BamHI restriction enzyme (Figure 1.2c). The linear sample has half of the linker on each side which allows the binding of primers for PCR amplification (Figure 1.2d). Lastly, paired-end sequencing is used for library formation using NEBNext Universal and Index primers.

1.4.3 Advantages & Disadvantages

iMARGI is an advanced procedure of MARGI, where "i" stands for in situ ligation in nuclei (Yan et al. 2019). This feature allows the iMARGI method to be applied to fewer mammalian cells. There are a couple of other methods that have a similar principle as iMARGI. ChAR-seq and GRID-seq are the most applied ones. The latter one is one step ahead which is applicable in mammalian cells. GRID seq (in situ capture of global RNA interactions with DNA by deep sequencing), uses a bivalent linker with adaptor ligation for the generation of RNA-chromatin library formation. The method also includes a bioinformatics pipeline (Zhou et al. 2019a).

One of the advantages of iMARGI is to be able to use the entire sequence in contrast to similar approaches, linkers are lost at the end of the experiment (Zhou et al. 2019b). For instance, in GRID-seq linkers should be read through to assign orientation which shortens the read length to 18 – 23 bp (Wu et al. 2019). iMARGI read length can increase up to 100 bp.

In Table 1.1., the three most known genome-wide applications are listed with their features. In the case of the number of cells, ChAR-seq is disadvantageous as it requires up to 100 million cells. Also, it is only suitable for fruit fly cells (Jukam et al. 2019). Ligation

reaction occurs in the nucleus for all methods which increases ligation efficiency. In that manner removing unligated samples are also important. Using an enzyme, Exo I is a much more convenient method as gel-based selection can cause material loss.

In contrast to other methods, the paired-end sequencing type is used in iMARGI allowing twice the number of reads. Linkers are not found in the result and the whole read pair can be used for mapping. This gives iMARGI another advantage in the read length category.

Table 1.1. The summary of different genome-wide methods.

	iMARGI	GRID-seq	ChAR-seq
Cell Number	2-5 million	2 million	100 million
Organism	Human	Human, mouse & fruit fly	Fruit fly
Sequencing Type	Paired-end	Single-end	Single-end
Usable sequencing length for mapping DNA/RNA	Up to 100 bp	18-23 bp	~65 bp
Ligation	In nucleus	In nucleus	In nucleus
Removal of unligated samples	Exonuclease and biotin selection	Gel-based size selection	No relevant experimental steps
Ligation of adaptors	Not applicable	Y-shaped adaptors	NEBNext Hairpin adaptors
Second-strand synthesis	Not applicable	Yes	Yes
Complementarity to adaptors	Yes	No	No

In iMARGI top linker is designed to have random bases to minimize bias in ligation (Wu et al. 2019). After ligation, the nucleic acid is isolated and the remaining reactions occur in the solution including cDNA synthesis, eliminating the possibility of inhibition in the nucleus (Yan et al. 2019). When compared with GRID-seq and ChAR-seq, iMARGI does not require complicated linker preparation steps only annealing is required. Also, unligated linkers are removed via an enzyme in contrast to gel-based removal in other methods (Wu et al. 2019). In sense of the final product, iMARGI does not demand second-strand DNA synthesis and adaptors (Wu et al. 2019).

iMARGI is a huge step forward for RNA – chromatin studies, however, it also has a few limitations. In this method at least 5 million cells should be used. The protocol takes two weeks to complete including many independent experiments. Additionally, the final library does not exactly show the active sites. iMARGI does not provide an overall statistical method, it is recommended to use statistics depending on the biological question and sample type.

The sequencing stage should be run to produce 300 million or more read pairs. It can be less, depending on the mode of RNA – chromatin interaction. The quality of reads can be checked with FastQC afterward (Andrews 2010). For the processing of data, a pipeline has been established in an iMARGI docker (Wu et al. 2019). The data is analyzed and finetuned according to the objective of the study. Data can be visualized with HiGlass as a heatmap or GIVE showing interactions on the chromosomal layout (X. Cao et al. 2018; Kerpedjiev et al. 2018).

1.5 The Aim

This thesis aims to form RNA – chromatin interaction libraries in HeLa cells with the iMARGI method. In this sense, the initial purpose is to optimize the protocol to be used in HeLa cells.

CHAPTER 2

MATERIALS & METHODS

2.1 Data Analysis

iMARGI method was designed on HEK cells and the data was published with the protocol article. The results comprised a file containing RNA – DNA interactions. The raw data was analyzed through the iMARGI pipeline in the iMARGI docker. The data matched the reference genome and the result had mapped RNA – DNA read pairs. The pairs were opened in R and data was scanned in terms of certain lncRNAs. Another file was created where DR5-AS interacting DNAs were listed. The list included 76 DNA molecules from 18 different chromosomes.

2.2 Cell Culture

HeLa ACC-57 cells were acquired from the German Collection of Microorganisms and Cell Cultures (DSMZ) GmbH. They were cultured with RPMI 1640 with L – Glutamine (Gibco) and 10% Fetal bovine serum (FBS) (Gibco). HeLa cells were cultured at 37°C with 5% CO_2 in a humidified incubator. Cell seeding density was 2 x 10^6 cells on each 15 cm dish. The next day, they were treated with DMSO at a final concentration of 0.1% (v/v) for 16 hours. The cells are harvested with 0.25% trypsin – EDTA (Gibco) and 1X PBS (Gibco). The suspended cells are centrifuged at 1000 rpm at room temperature for 5 min. The cell pellet was resuspended in 1x PBS. The cells were counted after staining with trypan – blue staining, 10 µL sample, and 90 µL of trypan– blue. Three replicates were combined to reach 10 x 10^6 cells.

2.3 Crosslinking

10 x 10^6 detached HeLa cells were centrifuged at 1500 rpm for 5 min at 4°C. The resulting cell pellet was resuspended with 1 mL 1X PBS and transferred into a 50-mL tube. 14 mL 1X PBS was added. 1 mL Formaldehyde (16% (wt/vol), Thermo Fischer Scientific) was added and cells were fixed at room temperature for 10 min with rotation. Fixation reaction was quenched with freshly prepared 4 mL 1M Glycine by incubating at room temperature for 10 min with rotation and then on ice for 10 min. Subsequently, the cells were spun down at 2,000 g for 5 min at 4°C, and the supernatant was discarded by pipetting. The cells were rinsed with 15 mL ice-cold 1x PBS without resuspending the cell pellet. Afterward, the cells were centrifuged at 2,000 g for 5 min at 4°C, and the supernatant was discarded. The tubes were left open for 2 minutes for aspirating purposes. The cell pellets were stored at -80°C.

2.4 Cell Lysis

The cells were thawed on ice angling the tube at 45°. The tube was inverted as the cells were thawed and the pellet was always facing the opposite direction of the ice. Afterward, the cells were resuspended in 2 mL of ice-cold lysis buffer and incubated on ice for 15 min. The cells were transferred into a 15-mL Dounce homogenizer (Merck). They were homogenized using pestle A. The pestle was moved slowly up and down 10 times, following incubation on ice for 1 min, and afterward, 10 more strokes were performed. The cells were incubated on ice further for 10 min. Two microcentrifuge tubes were weighed. The homogenized cell mixture was transferred into one of the tubes and cells were centrifuged at 2500 g for 5 min at 4°C. After nuclei isolation, control samples of 10 µL nuclei pellet was collected. The nucleus isolation was checked according to the DAPI staining protocol from 4.14.1. Afterward, the pellet was resuspended with 500 µL 1x CutSmart Buffer (NEB), and

250 µL of the mixture was transferred to the second tube, forming two technical replicates. The tubes were centrifuged again at 2500g at 4°C for 2 min, obtaining the nuclei pellet.

2.5 DNA Fragmentation

The supernatant was discarded from both tubes and the remaining pellets were weighed. From this step on, every step was applied for both tubes. The amount of 0.5% SDS for permeabilization was calculated according to the weight of the pellets. It was calculated by subtracting the tube weight from the total amount. Subsequently, the pellet was resuspended with SDS with a volume of three times the weight of the pellet for both tubes. The tubes were incubated at 62°C for 15 min in a thermomixer with 800 rpm shaking. The reaction was quenched with 1x CutSmart buffer and 10% Triton X-100 (Applichem). Final concentrations of SDS and Triton X-100 were adjusted as 0.1% (v/v) and 1% (v/v), respectively. The condition for quenching reaction was incubation at 37°C for 15 min in the thermomixer with 800 rpm shaking. Afterward, the cell pellet was acquired for both tubes. The pellet was washed by resuspending the nuclei with 500 µL of 1X CutSmart buffer (NEB) and subsequently discarding the supernatant. This step was repeated once. 10 µL of nuclei pellet was taken as control in-between the washes to check nuclear integrity with DAPI.

For DNA fragmentation by the Alu1 (NEB) enzyme the cell pellet was resuspended in the mixture shown in Table 2.1. The DNA fragmentation mix was incubated at 37°C on rotation, for 17h.

Table 2.1. The reagents used for DNA fragmentation with Alu1 restriction enzyme.

Reagent	Volume (µL)	Final Concentration
H_2O	198	
10x CutSmart Buffer	30	1x
Alu1 (10 U/µL)	70	2.3 U/µL

(cont. on the next page.)

Table 2.1. (cont.)

RNasin Plus (40 U/µL)	2	0.3 U/µL
Total	300	

2.6 RNA Fragmentation

RNase 1 (NEB) was diluted tenfold in 1x PBS. 1 µL of diluted RNase 1 (NEB) was added to samples incubated overnight. Afterward, the mixture was incubated at 37 °C for 3 min in the thermomixer (DLAB). Lastly, the reaction was terminated by incubation on ice for 5 min. The nuclei were pelleted by centrifuging at 2500g at 4°C for 2 min. The supernatant was collected as the control sample. The pellet was washed with 300 µL PNK Wash buffer and centrifuged at 2500g at 4°C for 2 min. The washing step was repeated. 5 µL of nuclei were taken for DAPI staining and 10 µL of nuclei was taken for Alu1 digestion control between the washing steps according to the protocol given in section 4.14.2.

2.7 Preparation of RNA and DNA end for ligation

The nuclei pellet was resuspended in the mixture with the following reagents listed in Table 2.2. The sample was incubated for 30 min at 37 °C with 800 rpm shaking in the thermomixer (DLAB).

Table 2.2. The volume of reagents given for RNA 3' end dephosphorylation

Reagent	Volume (µL)	Final Concentration
H_2O	148	
5x PNK phosphatase buffer pH 6.5 (NEB)	40	1x

(cont. on the next page.)

Table 2.2. (cont.)

T4 PNK (10 U/µL) (NEB)	10	0.5 U/µL
RNasin Plus (40 U/µL) (Promega)	2	0.4 U/µL
Total Volume	200	

After incubation, the nuclei were precipitated, and the supernatant was taken as a control sample. The pellet was washed again with 300 µL of PNK wash buffer on ice, twice. The DNA end was prepared for ligation by dA – tailing. The reagents of the procedure are given in Table 2.3.

Table 2.3. The reagents volume used in dA - tailing reaction.

Reagent	Volume (µL)	Final Concentration
H_2O	164	
10× NEBuffer 2 (NEB)	20	1x
Klenow fragment (3′–5′ exo-; 5 U/µL) (NEB)	12	0.3 U/µL
10 mM dATP (NEB	2	0.1 mM
10% (vol/vol) Triton X-100	2	0.1% (vol/vol)
Total Volume	200	

The tube was incubated at 37°C for 30 min with 800 rpm shaking in the thermomixer. Afterward, nuclei were precipitated, and the supernatant was collected as a control sample. The pellet was washed again twice with 300 µL PNK wash buffer. The pellet was kept on ice while linkers were prepared for ligation.

2.8 Linker Preparation

The top linker was adenylated before the linker annealing. The adenylation was performed according to the conditions given in Table 2.4.

The reagents were mixed in a 0.2 mL PCR tube. The reaction was incubated in a thermocycler (Blue–Raytech) at 65°C for 1 hour and then at 85°C for 5 min. Afterward, 2 µL of the bottom linker was added to the PCR tube containing adenylated top linker. For annealing, the tube was placed into the thermocycler with the following program: 95°C for 2 min; then 71 cycles of 20 s, starting from 95°C and decreasing the temperature by 1 °C at each cycle, down to 25°C, and hold at 25°C.

Table 2.4. The reagents for top linker adenylation.

Reagent	Volume (µL)	Final Concentration
H_2O	3.5	
10× 5’ DNA Adenylation Reaction Buffer (NEB)	3	1x
1 mM ATP (NEB)	3	0.1 mM
Top Linker	2.5	30 µM
50 µM Mth RNA Ligase (NEB)	18	30 µM
Total Volume	30	

The annealed linker complex was purified with 200 µL Dynabeads MyOne Silane beads (Invitrogen). Silane beads were vortexed only for a second to obtain a homogenous mixture. 200 µL of beads are transferred into a fresh 1.5 mL microcentrifuge tube. The tube was placed on a magnetic stand where beads were attracted to the magnet leaving a clear supernatant inside the tube. This solution was discarded. The beads were resuspended in 300 µL RLT buffer (Qiagen) and placed on the magnetic stand. The supernatant was discarded again. After removing the tube from the magnetic stand, 3.5-fold the sample volume of RLT

buffer and 34.2 µL annealed linker sample was added into the tube. Finally, 4.5x-fold the sample volume of isopropanol (Sigma) was added. The tubes were inverted 5 times to mix the solution. They were incubated for 10 minutes at room temperature on rotation. Afterward, the tubes were placed on the magnetic stand, and the supernatant was removed. The beads were washed with 70% Ethanol (Merck) twice, in between washes the ethanol was removed after placing the tube on the magnetic stand. After the second wash, tubes were placed lid open on the magnetic stand for air-dry purposes for 10 min. The beads should lose their shiny appearance when dried properly. The beads were incubated in 50 µL Ultra-pure H_2O for elution, to allow the transfer of the linker sample to the solution. Afterward, the tubes were placed again on a magnetic stand, and the supernatant was collected. For the control of annealing efficiency, 1 µL of annealed linker replicates were transferred to a new tube. The efficiency was checked as stated in 4.14.3.

The linker – RNA ligation was achieved by mixing the following reagents (Table 2.5.). in a 2 mL microcentrifuge tube to facilitate mixing. The reagents and volumes were shown in Table 2.5.

Table 2.5. The reagents used for linker - RNA ligation reaction.

Reagent	Volume (µL)	Final Concentration
H_2O	40	
Linker	46	
10x T4 RNA Ligase Reaction Buffer (NEB)	20	1x
50% (wt/vol) PEG 8000 (NEB)	80	20% (vol/vol)
10% (vol/vol) Triton X – 100	2	0.1% (vol/vol)
RNasin Plus (40 U/µL)	2	0.4 U/µL
T4 RNA Ligase 2, truncated KQ (200 U/µL) (NEB)	10	10 U/µL
Total	200	

The RNA – linker ligation samples were incubated at 22°C for 6 h and then at 16 °C for 17 h with rotation in a cold incubator. The detailed sequence information of used primers is given in Table 2.6. The list consists of linkers and cut_oligo primer detailed sequence. *stands for biotinylation and Invd T is written for inverted thymine base.

Table 2.6. The list of primers used in the iMARGI protocol.

Primer Name	Sequence
Top Linker	5'/Phos/NNAGATCGGAAGAGCGTCGTGTAGGGAGGATCCGTTCAG ACG TGTGCTCTTCC*GA/iBiodT/CT3'
Bottom Linker	5'/Phos/GATCGGAAGAGCACACGTCTGAACGGATCCTCCCTACACG A CGCTCT3'
Cut_oligo	5'TCGTGTAGGGAGGATCCGTTCAGACGTGTGCTCT/3InvdT/3'

2.9 Proximity Ligation

RNA – linker reaction was terminated by adding 20 µL of 0.5 M EDTA (Sigma) was put in both tubes and incubation at 16°C. The nuclei were pelleted, and supernatants of both tubes were collected as control samples. The pellets were washed with 500 µL PNK Wash Buffer 5 times. The supernatant was discarded. This step is crucial for removing free linkers.

Table 2.7. The reagents used for proximity ligation.

Reagent	Volume (µL)	Final Concentration
H_2O	1660	
10x DNA Ligase Reaction Buffer (NEB)	200	1x

(cont. on the next page.)

Table 2.7 (cont.)

BSA (20 mg/mL) (NEB)	100	1 mg/mL
T4 DNA Ligase (2000 U/µL) (NEB)	4	4 U/µL
10% (vol/vol) Triton X – 100	20	0.1% (vol/vol)
RNasin Plus (40 U/µL)	16	0.5 U/µL
Total	2000	

Reagents shown in Table 2.7., except T4 DNA Ligase, were mixed in a 5 mL centrifuge tubes. Approximately 1 mL of the reaction mixture was used for resuspending the pellets and transferred back to 5 mL tube including the pellet. Finally, T4 DNA ligase was added to the tubes. The tubes were incubated in the incubator at 16°C for 17 h on rotation.

2.10 Reverse Crosslinking and DNA/RNA Extraction

The proximity ligation reactions were terminated by the addition of 200 µL of 0.5M EDTA and incubation for 15 min at 16°C. The tubes were centrifuged to obtain nuclei pellets and supernatants were collected. The pellets were washed with 1x PBS. The mixture was transferred to a 1.5 mL microcentrifuge tube after the second wash.

The pellets were resuspended with 250 µL extraction buffer to extract nucleic and reverse the crosslinking. The tubes were incubated at 65°C for 3 hours with 800 rpm shaking. Afterwards, 250 µL phenol/chloroform/isoamyl alcohol (PCI) (25:24:1) (pH 8) (BioShop) was added into the tubes. The tubes were vortexed until the solution turns white. Phase-lock gel tubes (VWR) were centrifuged at 1500 xg for 1 minute at room temperature (RT) empty to collect chemicals on the walls of the tubes. The samples were added on top of the phase lock gel tubes and centrifuged at 12,000 xg for 5 minutes at RT. The aqueous phases on top of the tubes are collected. PCI of 250 µL was added to the collected samples and vortexed again until the mixtures turned white. Afterward, the samples were transferred into new, pre-spun, phase lock-gel tubes (VWR) and centrifuged 12,000 xg for 5 minutes at RT. The supernatants were again collected, 1/10th volume of 3M sodium acetate (Applichem) (pH 5.2)

and 3-fold volume of ice-cold 100% Ethanol were added into the tubes. The tubes were mixed well 3 times before incubating the tubes at -80°C for 30 minutes. After incubation, the tubes were centrifuged at 16,000 xg for 30 minutes at 4°C. The supernatants of both tubes were discarded, and the pellets are air-dried. The pellets were eluted in 50 µL UP H_2O. The concentrations of the samples were measured. They can be stored at -20°C for up to 6 months.

2.11 Removal of biotin from unligated linkers

The samples were divided into 0.2 mL microcentrifuge tubes, each containing a maximum of 8 µg of the sample. Table 2.8. shows the conditions for the removal of biotin from unligated linkers.

Table 2.8. The following reagents are used for the first step of removing unligated linkers.

Reagent	Volume (µL)	Final Concentration
H_2O	To 144	
10x NEBuffer 2	15	1x
BSA (20 mg/mL) (NEB)	1	0.1 mg/mL
Exo1 (20 U/µL) (NEB)	5	0.7 U/µL
RNasin Plus (40 U/µL)	2	0.6 U/µL
Nucleic Acids	Variable	
Total	144	

The following reagents were added to each PCR tube with a suitable volume of nucleic acids to complete the total volume. The tubes were incubated at 37°C for 30 minutes in the thermocycler. After the incubation, reagents shown in Table 2.9. were added to the tubes making a final volume of 150 µL.

Table 2.9. The reagents used for the second step of removing unligated linkers.

Reagent	Volume (µL)	Final Concentration
10 mM dATP (NEB)	1.5	0.1 mM
10 mM dGTP (NEB)	1.5	0.1 mM
T4 DNA Polymerase (3 U/µL) (NEB)	3	0.06 U/µL
Total	6	

The tubes were incubated again in a thermocycler at 12°C for 2h. In the end, all samples from the same replicate were combined in individual 1.5 mL microcentrifuge tubes. Adequate volume of EDTA was added in both tubes to have 0.5 M of EDTA as the final concentration. The tubes were stored at -20°C overnight.

2.12 Biotin pull-down of RNA-DNA chimeric sequences, reverse transcription, and ssDNA circularization

200 µL of Dynabeads MyOne streptavidin C1 (Invitrogen) beads were transferred to a 15 mL centrifuge tube for each sample. The beads were vortexed briefly to obtain a homogenous mixture followed by a wash step with 300 µL 1x B&W buffer. The supernatant was discarded in between washes by placing the tube on the magnet for a minute. This step was repeated three times. The beads were resuspended in the 2x B&W buffer equal to the volume of each sample. The samples were added on top of the bead–buffer mixture. The tubes were incubated on rotation at room temperature for 30 minutes. Afterward, the tubes were placed on the magnet and washed seven times with 7 mL of high salt biotin wash buffer. Tubes were incubated with the buffer for 2 minutes for each wash until the solution became clear before discarding the supernatant. Subsequently, the beads were resuspended with 1 mL of PNK wash buffer and the mixture was transferred to a new 2 mL microcentrifuge tube. The buffer was also removed after the tubes were placed onto a magnet. Washing with PNK

wash buffer is crucial to remove any trace of high-salt biotin wash buffer which can inhibit enzymatic reactions in the next steps. For the reverse transcription step, the mixture was prepared with the following reagents given in Table 2.10. for each sample.

Table 2.10. The reagents required for reverse transcription are listed in the table.

Reagent	Volume (µL)	Final Concentration
H2O	22	
5x First-strand buffer (ThermoFisher)	8	1x
10 mM dNTP (ThermoFisher)	2	0.5 mM
100 mM DTT (ThermoFisher)	2	5 mM
RNasin Plus (40 U/µL)	2	2 U/µL
SuperScript RT (200 U/µL) (ThermoFisher)	4	20 U/µL
Total	40	

The beads were resuspended in the mixture and incubated at 50°C for 1 hour with shaking at 800 rpm in the thermomixer. After the incubation, the tubes were placed on the magnet and the supernatant was discarded. The beads were washed with 300 µL 1x B&W buffer twice and with 300 µL PNK wash buffer once.

Table 2.11. The reagents required for DNA 5' end phosphorylation.

Reagents	Volume (µL)	Final Concentration
H2O	74	
10x T4 PNK reaction buffer, pH 7.6 (NEB)	10	1x
T4 PNK (10 U/ µL)	5	0.5 U/µL
10% (vol/vol) Triton X-100	1	0.1% (vol/vol)
10 mM ATP	10	1 mM
Total	100	

The following reagents stated in Table 2.11. were mixed and used for resuspension of the bead mixture of each sample. The samples were incubated at 37°C for 1 hour. After incubation, the beads are washed with 300 µL of 1x B&W buffer. Subsequently, the beads were resuspended in a 100 µL denaturing buffer to obtain the single-stranded cDNA-DNA chimeric sequence. The opposite strand, RNA – DNA, remained bound to the beads. The mixture was incubated at room temperature for 15 minutes, thus single-stranded cDNA – DNA molecules were released into the solution. Afterward, the tubes were placed on a magnet and the clear solution, supernatant, was collected. To fully hydrolyze the complementary RNA strand the solution was incubated at 98°C for 20 min. Then, it was neutralized with 10 µL of 1M HCl and 10 µL of 1M Tris-HCl, pH 7.5. The chimeric sequence was isolated from the solution with 100 µL of silane beads, following the same protocol mentioned earlier. The beads were eluted with 20 µL H_2O.

Table 2.12. The reagents required for circularization reaction.

Reagent	Volume (µL)	Final Concentration
ssDNA	15	
CircLigase 10x reaction buffer (Epicentre)	2	1x
1mM ATP (Epicentre)	1	0.05 mM
50 mM $MnCl_2$ (Epicentre)	1	2.5 mM
CircLigase ssDNA Ligase (100 U/µL) (Epicentre)	1	5 U/µL
Total	20	

The reagents listed in Table 2.12. were mixed with the 15 µL ssDNA sample as elution is not achieved with full efficiency. The circularization reaction was incubated at 60°C for 4 h in a thermocycler and at 80°C for 10 minutes. The samples were held at 4°C in the thermocycler overnight.

2.13 Cut_oligo annealing, BamHI digestion, and library amplification

The annealing mix was prepared with the reagents stated in Table 2.13. The mixtures were added to 20 µL of circularized ssDNA of both tubes and incubated in the thermocycler with the same annealing program used for the annealing of linkers. At the end of the program, 3 µL of BamHI-HF (NEB) was added and incubated for 1 hour at 37°C in the thermocycler to linearize the DNA.

Table 2.13. The reagents used for circularization.

Reagent	Volume (µL)	Final Concentration
H_2O	23	
10x CutSmart buffer	3	1x
10 µM Cut_oligo	1	0.2 µM
Total	37	

Another bead isolation was performed with 50 µL of silane beads to isolate linearized DNA and the sample was eluted with 25 µL H_2O. This sample concludes all DNA – RNA interactions as a linearized DNA complex. To form our library, the sample should be amplified via PCR. The optimal PCR cycle number was determined through the optimization step. In a 0.2 mL PCR tube, 5 µL of linearized DNA, 25 µL of 2x NEBNext High-Fidelity PCR Master Mix, 1 µL of Universal primer (10 µM), 1 µL Index Primer (10 µM), and 18 µL of H_2O were mixed. The mixture was divided into 5 different tubes forming 10 µL aliquots. Five different PCR cycles were tested: 10, 13, 15, 18, and 22.

Table 2.14. The reagents used for library formation PCR.

Reagent	Volume (µL)	Final Concentration
H_2O	40	
Linearized DNA	5	
2x NEBNext High-Fidelity PCR MM (NEB)	50	1x
Universal Primer (10 µM)	2.5	0.25 µM
Index Primer (10 µM)	2.5	0.25 µM
Total	100	

The linearized DNA sample was amplified to form our library by mixing the reagents stated in Table 2.14. and the PCR program in Table 2.15. The library samples were kept at 4°C overnight.

Table 2.15. The PCR program for library amplification.

Cycle No	Denaturation	Annealing	Extension
1	98°C, 30 s		
2 – 17	98°C, 10 s	65°C, 30 s	72°C, 30 s
18			72°C, 5 min

2.14 Library Size Selection

The purification of the final product was performed by MN NucleoSpin Gel and PCR Clean-up kit. For 100 µL of PCR product, 200 µL NTI (binding buffer) buffer was added to the sample. The mixture was loaded into a column. The tube was centrifuged for 30 seconds at 11,000 xg and the flow-through was discarded. Subsequently, 700 µL of NT3 (wash

buffer) buffer was added to the column and centrifuged for 30 seconds at 11,000 xg. The flow-through was discarded. This step was repeated. The column was transferred into a new microcentrifuge tube and centrifuged for 3 minutes at 11,000 xg to remove the excess buffer. Afterward, the column was heated for 3 minutes at 70°C. 12 µl NE Buffer was added for elution and incubated for 1 min at room temperature. The sample was centrifuged at 11,000 xg for a minute. This step was repeated with 12 µL NE Buffer and the flow-through was collected in the microcentrifuge tube.

The sample was loaded on 1.8% agarose gel with a 1 kb ladder (NEB) and run for 50 min at 100V. After the illumination of the gel under UV – light, the sample was cut out from the gel between the sizes 250 to 1000 bp. The selected library was extracted from the gel by NucleoSpin Gel and PCR clean-up (Macherey-Nagel). The gel pieces were transferred into a 5 mL centrifuge tube. To obtain the correct weight of the gel piece, the weight of the empty tube was subtracted from the total weight. NTI Buffer at a volume of twice the gel weight was added into the tube. The gel was incubated at 50°C and it was vortexed every 2 min until complete dissolution. The dissolved mixture was loaded into a column and for the remaining isolation steps, the PCR clean-up method was followed. The final 24 µL sample was stored at -20°C and sent for sequencing in suitable conditions.

2.15 Control Experiments

2.15.1 DAPI Staining

It is crucial to isolate nuclei and maintain their integrity in the iMARGI method. After the initial stages nuclei sample was taken as a control. The control samples were obtained after nuclei isolation, nuclear membrane permeabilization with SDS and Alu1 digestion steps. 10 µL sample was mixed with 90 µL 100% Ethanol for fixation. The fixed samples were stained with DAPI (4′,6-diamidino-2-phenylindole), a blue fluorescent dye, which binds to DNA. For 9 µL of the sample, 1 µL of 1 µg/µL DAPI was used. The mixture was

transferred to a microscope slide. The observation was performed for 4x, 10x and 20x magnifications.

2.15.2 Alu1 Digestion

At first, the chromatin digestion protocol required optimization as the digestion of DNA was not efficient enough. Also, the thermomixer explained in the iMARGI protocol had different functions than the one found in our lab, such as intermittent mixing. Hence, two different devices were tested for chromatin digestion. One was our thermomixer and the other one was an orbital shaker which was placed in the incubator. The purpose of intermittent shaking is to avoid nuclei precipitation. The orbital shaker accomplished the aim by rotating the tube upside down. The optimal digestion conditions were defined by changing SDS/Nucleus ratio and incubation time.

Table 2.16. Optimization conditions for chromatin digestion. The ratio is referencing to the ratio between SDS and nuclei pellet weight.

Condition #	Time (min)	Ratio
1	10	3
2	15	3
3	18	3
4	10	5
5	15	5

The conditions given in Table 2.16, were tested for chromatin digestion protocol. Alu1 amount could be adjusted as well; however, a cost-effective approach was chosen. Isolated nuclei samples from the same sample were treated in these conditions.

The digestion level of chromatin is a key criterium in iMARGI. The digestion was checked on the agarose gel. After Alu1 digestion, 10 µL of the sample was taken and resuspended in 100 µL extraction buffer. The mixture was incubated at 65 °C for 2 hours to extract nucleic acids and reverse crosslinking. After incubation, nucleic acid was isolated with 50 µL silane beads and eluted in 50 µL UP H_2O. The concentration of the sample was measured, and 500 ng DNA was subjected to gel electrophoresis on 1% agarose gel. The gel ran for 45 min at 120V with NEB 1kb ladder.

2.15.3 Annealing of Linkers

The custom top linker was observed as two bands on the agarose gel which suggested either contamination or annealing on itself. An alkaline gel was prepared to see if the second band was caused by a secondary structure. The alkaline gel prevents the hydrogen bonding between bases as they lose protons at high pH (Green and Sambrook 2021). Thus, DNA is observed in denatured from as ssDNA.

The PCR conditions include an initial heating stage to avoid primer dimers. Thus, the presence of contamination was tested by boiling the sample at 98°C for 10 min. Also, the top linker was incubated with the initial stage of the annealing PCR cycle, at 95°C for 2 min. PCR setup was sufficient to remove dimers in case of self-annealing.

The annealing of linkers was checked on 2% agarose gel due to its small size. For the sample preparation, 0.5 µL of each linker and annealed linker were mixed with 9.5 µL H_2O.

2.15.4 Nucleic Acid Ratio

The experiments of the first week were performed in the nuclei. Thus, after each enzymatic reaction supernatant was collected to compare concentrations with DNA and RNA extracted from the nuclei at the end of the first week. Nucleic acid from the supernatant was

isolated via streptavidin beads after treatment with 10 µL of Proteinase K (NEB), to reverse crosslinking. The samples were incubated for 2h at 65°C. After elution with 15 µL UP H_2O step, the sample was divided into two (DNA and RNA) without removing the beads. The DNA sample was subjected to 1 µL RNase A (Thermo Fisher Scientific) while the RNA sample was treated with 1µL Turbo DNase and 2.5 µL 10x TURBO Reaction Buffer (Invitrogen). The samples were incubated at 37°C for 30 min. Nucleic acid samples were isolated with the beads remaining in the tubes. The concentrations of samples were measured and multiplied by 2 as the samples were divided in the beginning. Afterward, the concentration was divided to the nuclear nucleic acid concentration.

CHAPTER 3

RESULTS

3.1 Data Analysis

iMARGI method was developed on HEK cells and the data was published. The data was analyzed, and it showed lncRNAs including DR5 – AS found by our lab which was shown to have a role in cell proliferation (Gurer et al. 2021). 77 DNA interactions were found for our candidate. Four of the interacting DNA were coding for proteins functioning inside the nucleus. CTCF was also found in the data suggesting a role of DR5 – AS in chromatin folding. iMARGI was also used previously on endothelial cells and shown lncRNA – chromatin interactions under diabetic hyperglycemic and inflammatory conditions (Sriram et al. 2020).

3.2 Nuclear Integrity

3.2.1 DAPI Staining

DAPI stained nuclei images from different stages are shown in Figure 3.1. Nuclei are stained by DAPI because of the presence of DNA (Ferro et al. 2017). Figure 3.1a shows the nuclei after the initial isolation stage. The nuclei can be observed as blue indication of DAPI binding to DNA. The proper circular shape without scattered particles indicates that the nucleus was not leaked. After Alu1 digestion and SDS permeabilization (Figure 3.1b, c, d, e) the same pattern was observed

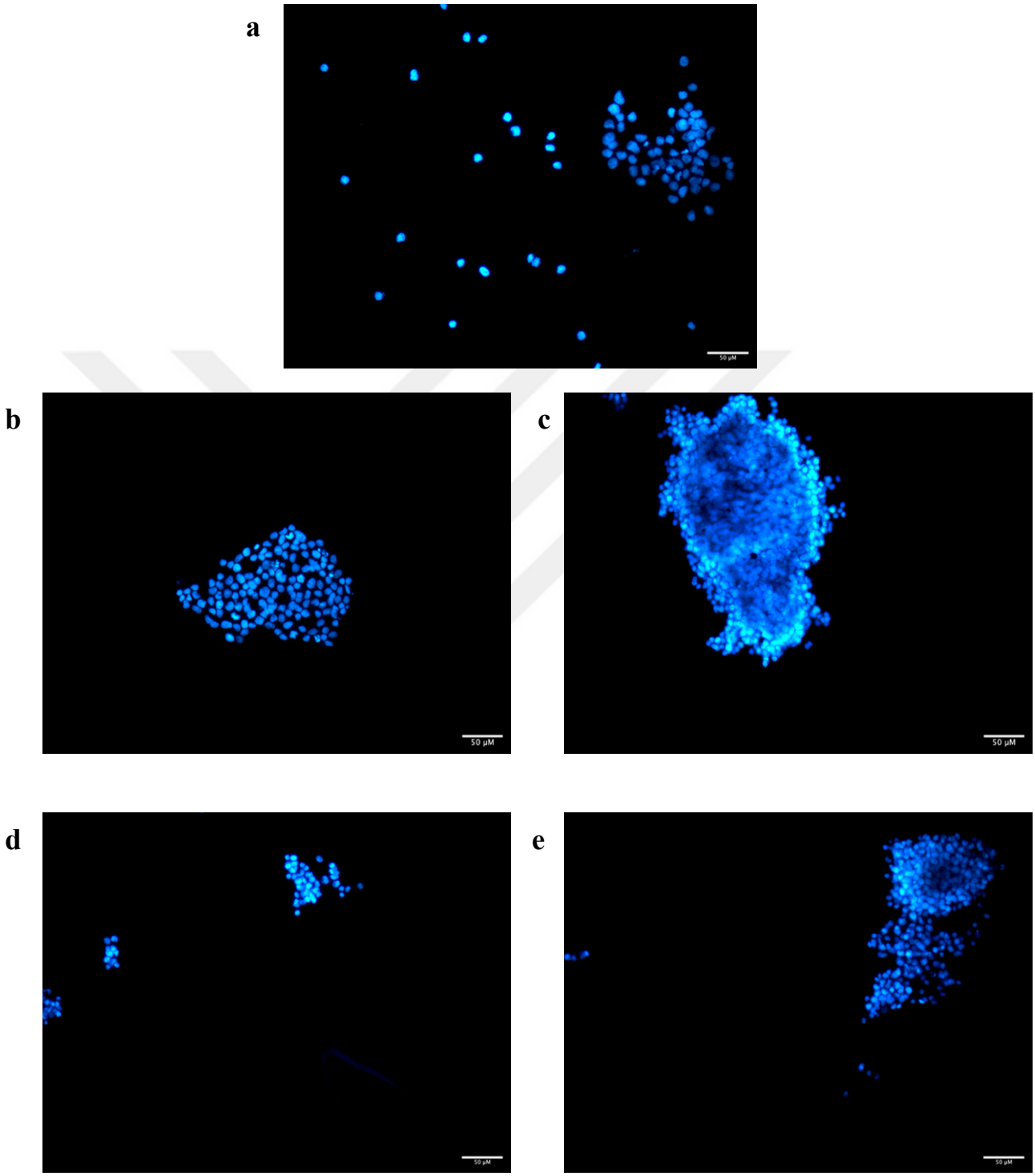

Figure 3.1. The DAPI stained nuclei images with 20x magnification and 50 µM scale bar. a) After nuclei isolation, the sample was not divided yet. b) iMARGI-1 sample after SDS permeabilization. c) iMARGI-2 sample after SDS permeabilization. d) iMARGI-1 after Alu1 digestion. e) iMARGI-2 sample after Alu1 digestion.

3.2.2 Nucleic Acid Ratio

Table 3.1. The supernatant and nuclear DNA ratio levels in percentage after each step of samples iMARGI-1 and iMARGI-2.

	iMARGI - 1	iMARGI – 2
Step 23	0.28%	0.15%
Step 28	0.14%	0.09%
Step 32	0.23%	0.32%
Step 50	0.83%	1.31%
Step 56	0.20%	0.36%

In Table 3.1, the supernatant and nuclear DNA ratio for each step of both iMARGI replicates are given. The nuclear integrities of nuclei were checked by comparing DNA and RNA concentrations from supernatant and nucleus. The ratio values are given as percentages. All values except step 50 for iMARGI-2 are below 1%. Other values were much lower.

Table 3.2. The supernatant and nuclear DNA & RNA ratio of iMARGI-2 after each step.

	iMARGI - 1	iMARGI - 2
Step 23	11.62%	6.56%
Step 28	0.21%	0.26%
Step 32	0.42%	0.33%
Step 50	0.05%	0.54%
Step 56	0.53%	0.38%

The ratio values between supernatant and nuclear RNA are given for both replicates in Table 3.2. Only the ratio of iMARGI – 1 for step 23 exceeded 10%. The 23rd step result for iMARGI was also high but did not exceed the 10% limit.

3.3 Linker Annealing

The first result of linkers showed the presence of multiple bands on the gel, indicating there might be a contamination or folding. Top and bottom linkers were loaded on 1% alkaline gel to check for contamination. Both linkers were observed on the gel as a single band. Their sizes were close the 50 bp. The top linker was checked with two different conditions on 2% agarose gel (Figure 3.2). The conditions were boiling and the initial denaturation stage.

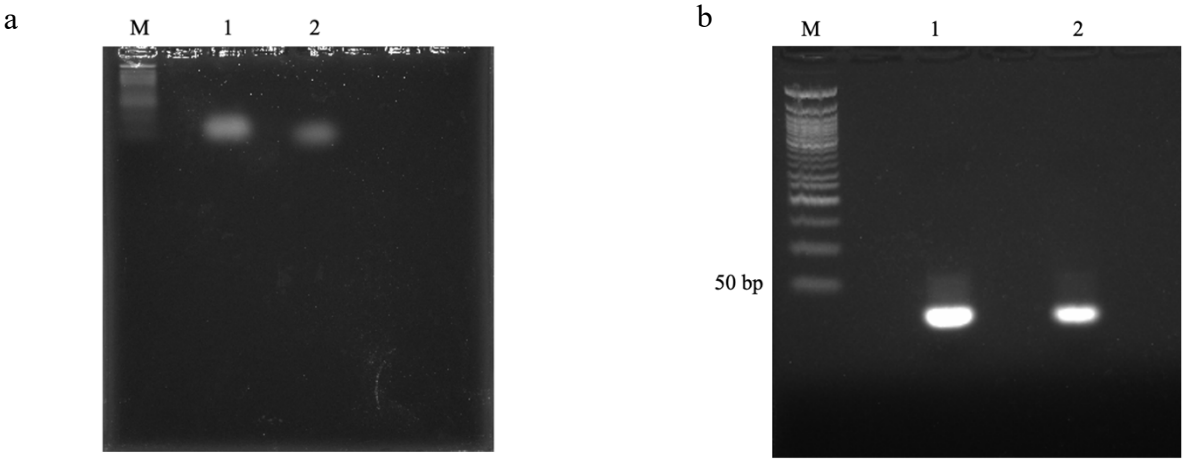

Figure 3.2. (a) The top, 1 and bottom, 2 linkers on %1 alkaline gel. (b) The top linker load on 2% AGE after incubation in two different conditions: 1, boiling and 2, initial denaturation step.

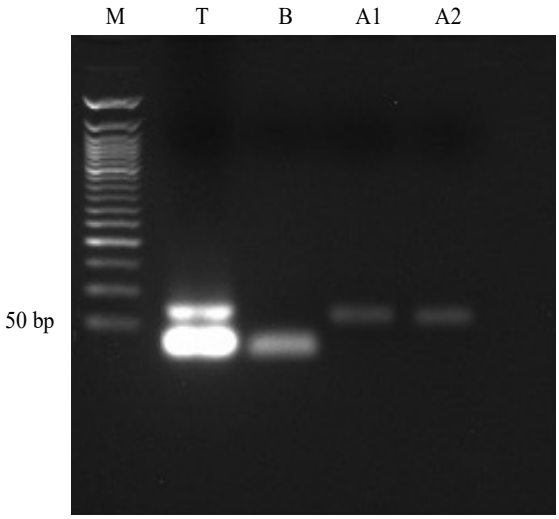

Figure 3.3. Top (T), bottom (B), and annealed linkers (A1, A2) on 2% Agarose gel. 50 bp marker was used for scaling.

The samples are observed under the 50 bp size level. Top and bottom linkers were loaded on 2% Agarose gel with annealed linker samples (Figure 3.3). Annealed linkers were from both replicates, A1 and A2. The top linker was observed as two bands. The denser band was located under the 50 bp scale. The bottom linker was also located in a closer range. Annealed linkers samples were observed at the same region, around 50 bp.

3.4 Nucleic Acid Extraction

After isolation of nucleic acids, the concentration values were measured. The results are given in Table 3.4. The amount of DNA in the sample for iMARG-1 was calculated as 46 µg and for iMARG-2 it was 64.45 µg. DNA concentration values for each replicate samples were in the range given in the protocol, 10 – 15 µg. Also, the absorbance values were stated to comment on purity.

Table 3.4. The nanodrop results of isolated nucleic acids.

Sample	Concentration (ng/µL)	260/230	260/280
iMARGI-1 DNA	958.43	2.33	1.93
iMARGI-1 RNA	769.19	2.33	1.93
iMARGI-2 DNA	1371.43	2.24	2
iMARGI-2 RNA	1091.81	2.24	2.01

3.5 Alu1 Digestion

Initially, the device used for chromatin digestion was checked. In Figure 3.4., both samples were seen on 1% agarose gel. Lane 1, showed a bigger range of DNA fragments compared to Lane 2. Also, the band stuck in the upper part of the gel was brighter in Lane 2.

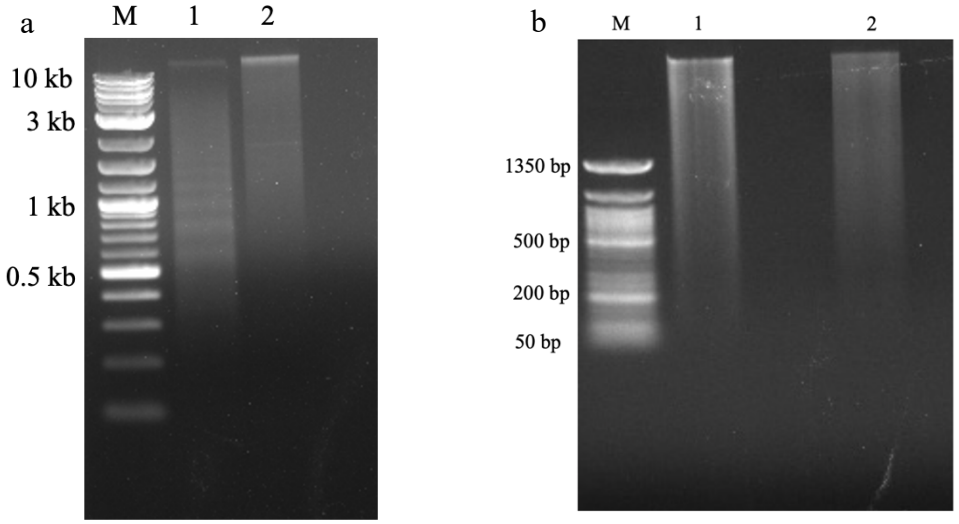

Figure 3.4. (Left) The digested DNA samples on 1% Agarose gel with 1 kb plus NEB marker. The sample incubated in the orbital shaker is shown with 1 and the thermomixer with 2. (Right) The DNA sample on 1% Agarose gel from both replicates was labeled as 1 for iMARGI-1 and 2 for iMARGI-2.

As seen in Figure 3.4. (right), digested chromatin was seen on the agarose gel. In both samples, the digestion was observed starting from 200 bp. The digestion range was covering a big area, including big pieces of fragments.

In Figure 3.5., the samples incubated in different conditions as stated in Table 2.16, were shown on 1% agarose gel. Smear formation was seen for all samples. However, in the first two lanes, there were bright band formations on top of the gel. In the third lane, the sample was brighter on the middle part of the gel whereas for the fourth one it was located more on the upper part. The fifth sample had a similar result to the third one

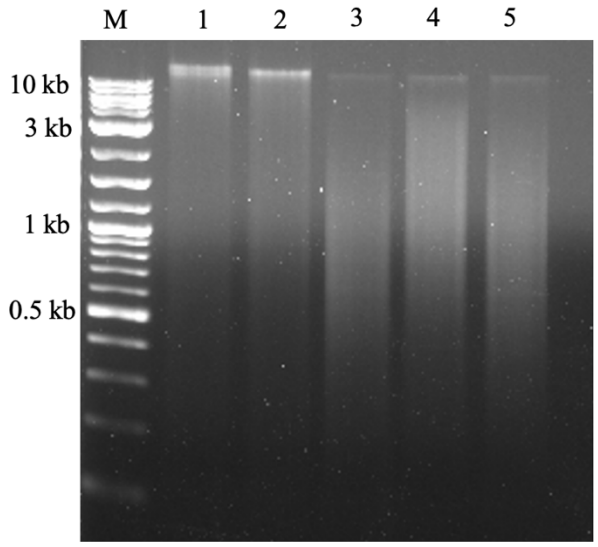

Figure 3.5. The digested DNA samples on agarose gel with different Alu1 digestion conditions, each lane represents another condition.

3.6 PCR Cycle Optimization

The end sample was incubated in a thermocycler with different PCR programs. The results were loaded on the 1% agarose gel. Smear formation was observed in 15, 18, and 22 cycle samples for both replicates, Figure 3.6. The fragments were located between 200 and 750 bp. For the initial lower cycle numbers of PCR, smear formation was not observed in both samples. Instead, a band was seen at 100 bp level. This band vanished when the cycle number was increased.

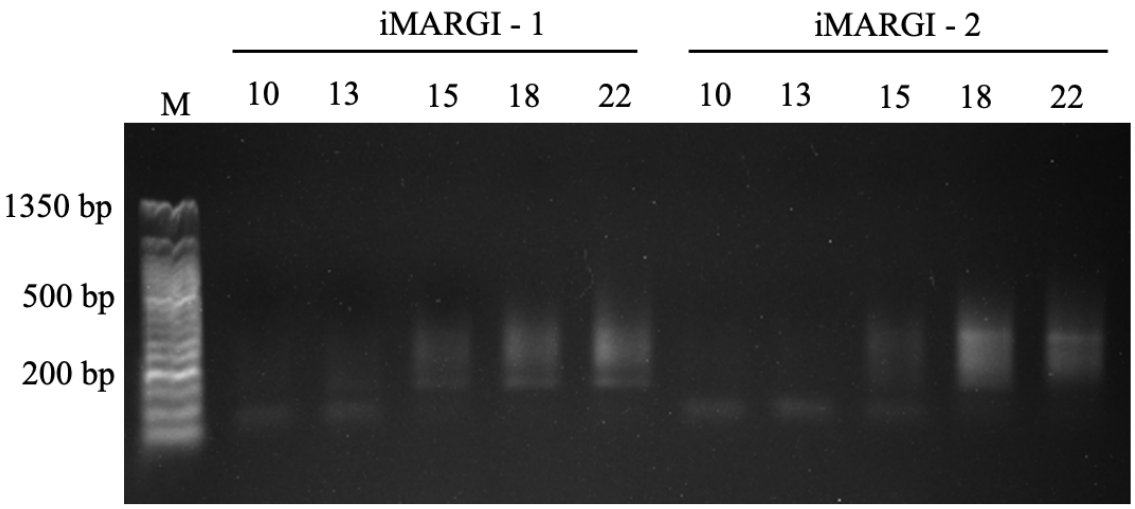

Figure 3.6. Different PCR cycle trials for each iMARGI replicate with a 50 bp ladder.

CHAPTER 4

DISCUSSION

The 3D structure of chromatin is regulated with lncRNAs in cancer cells. New studies are developed to investigate all lncRNAs interacting with the chromatin. iMARGI is selected for this purpose as it can be applied to mammalian cells. Also, it gives the longest read length compared to similar methods. The purpose of this thesis was to generate lncRNA – chromatin interaction libraries from HeLa cells. Thus, another aim was to optimize the iMARGI method to fit our lab and HeLa cells in different conditions. The results gathered from the published iMARGI data served as the basis for our hypothesis. It was proposed that there are lncRNAs interacting with the chromatin in HeLa cells regulating chromatin architecture.

lncRNA – chromatin interactions were pulled down as ssDNA fragments of different sizes. After the first week, the required amount of nucleic acid was obtained. The final library results containing ssDNA were observed for HeLa cells.

In iMARGI, it was crucial to apply the initial steps of ligation inside the nucleus. The first step is to isolate the nucleus and permeabilize it without compromising nuclear integrity. As seen in Figure 3.1a, the isolation of nuclei was successful for both replicates. All observed circular structures were stained with DAPI and emitted blue under the fluorescence microscope. After the nuclear membrane is permeabilized with SDS, the structures are successfully stained again, and nuclei structures are observed in both samples (Figure 3.1b&c). The same pattern is preserved in the samples gathered after chromatin digestion (Figure 3.1d&e).

Two conditions were compared in terms of DNA isolation and agarose gel electrophoresis results determined the method. Most of the fragments should be located between 200 bp and 1500 bp. The sample incubated in the orbital shaker (Figure 3.4a) showed a variety of fragments concentrated between 600 bp and 2500 bp, whereas the Thermomixer sample was comprised of fragments bigger than 3 kb and a bright band with a size larger than 10 kb. Thus, the orbital shaker was chosen for Alu1 digestion.

In the first two samples of different SDS conditions, the ratio was held at 3 whereas the duration was increased for the second sample. The digestion was not sufficient for both samples as there was a large bright band located at the upper part from 10 kb. The third sample had the ratio of 3 with 18 min of incubation. It had a clear smear formation, and the brighter part was located between 150 bp and 2.5 kb. For the fourth sample, the ratio was increased to 5 with 10 min of incubation, in which most of the fragments were located on the upper side of the gel, starting from 1 kb. The fifth sample had a ratio of 5 with 15 min incubation. Its digestion was nearly as efficient as the third sample but, the third sample's fragment size range was bigger. The result on agarose gel indicated that 18 min of incubation with a ratio of 3 gives the best digestion.

The Alu1 digestion results of replicates were similar to each other during the application of the iMARGI protocol. The size distribution started from 200 bp. Most fragments were between 200 bp and 2.5 kb. The digestion was efficient for the protocol.

The ratio of supernatant and nuclear nucleic acid shows if there is nuclear leakage or not. The threshold ratio is 0.01 (1%) for every step of DNA and RNA except the initial value of the RNA ratio. This value can be as high as 0.1 (10%).

The ratio values of replicates were shown in Tables 3.1 & 3.2. The ratio for DNA of iMARGI-1 did not exceed the threshold value in any of the steps. However, in step 50 the ratio of iMARGI – 2 was 0.3% higher than 1%. The value was not drastic, so the experiment was continued. In RNA ratio levels iMARGI – 1 there was an unexpected value for step 23 with 11.62%. Nevertheless, the process was not terminated. iMARGI – 2 had all acceptable ratio values for RNA.

The linker annealing was checked on 2% agarose gel. A shift in size compared to the top or bottom linker indicated that annealing was successful. At the initial trials of the methods, the top linker was observed as multiple bands on the gel. There was a suspect that the issue could be caused by either contamination or the formation of secondary structures of the top linker. First, the linkers were run on an alkaline gel to check the presence of contamination. In Figure 3.2a, both top and bottom linkers are observed as a single band. This result indicates the second band observed was a secondary structure caused by the folding of the linker on itself. Subsequently, the PCR conditions were checked with a regular denaturing boiling protocol to see if the initial stage was sufficient to denature the second band. The top linker

was run on the gel after the initial stage of PCR and boiling. The top linker was observed as a single band at the right size, indicating during the annealing protocol the top linker is found as a single band.

The results of replicates were amplified with PCR using different cycle numbers. Most of the library should be located between 150 bp and 1000 bp. The libraries of both replicates were located between 150 bp and 500 bp after 18 cycles of PCR.

CHAPTER 5

CONCLUSION

In this study, the RNA – DNA interaction molecules were ligated to a linker sequence. The formed RNA – linker – DNA molecule was pulled down with streptavidin-biotin interaction. The pulled-down molecule was converted to cDNA – linker – DNA and circularized afterward. The circularized molecule was re-linearized and amplified with PCR, forming lncRNA – chromatin interaction libraries. The optimal PCR cycle number was determined, and the library was observed on the gel for non-treated HeLa cells. The optimal conditions for the iMARGI procedure were determined including Alu1 digestion, linker annealing, and library amplification. The library sequencing results will provide a basis for lncRNA – chromatin interaction in HeLa cells. Ideally, these findings can be repeated for different cancer cells with or without treatment.

Future research will focus on HeLa cells in apoptotic conditions. For this purpose, HeLa cells will be treated with different pro-apoptotic drugs, TNF-alpha mAb, and Cisplatin. This would provide a good starting point for investigating chromatin architecture regulation by lncRNA in apoptotic conditions. lncRNAs selected from these libraries will be selected as candidates for future applications.

REFERENCES

Andrews, S. 2010. “FastQC: A Quality Control Tool for High Throughput Sequence Data. .” Https://Www.Bioinformatics.Ba Braham.Ac.Uk/Projects/Fastqc/ . 2010.

Arab, Khelifa, Emil Karaulanov, Michael Musheev, Philipp Trnka, Andrea Schäfer, Ingrid Grummt, and Christof Niehrs. 2019. “GADD45A Binds R-Loops and Recruits TET1 to CpG Island Promoters.” *Nature Genetics* 51 (February): 217–23. https://doi.org/10.1038/s41588-018-0306-6.

Arun, Gayatri, Sarah D Diermeier, and David L Spector. 2018. “Therapeutic Targeting of Long Non-Coding RNAs in Cancer.” https://doi.org/10.1016/j.molmed.2018.01.001.

Azam, Sikandar, Shuai Hou, Baohui Zhu, Weijie Wang, Tian Hao, Xiangxue Bu, Misbah Khan, and Haixin Lei. 2019. “Nuclear Retention Element Recruits U1 SnRNP Components to Restrain Spliced LncRNAs in the Nucleus.” *RNA Biology* 16 (8): 1001–9. https://doi.org/10.1080/15476286.2019.1620061/SUPPL_FILE/KRNB_A_1620061_SM1917.ZIP.

Bannister, Andrew J, and Tony Kouzarides. 2011. “Regulation of Chromatin by Histone Modifications.” *Cell Research* 21: 381–95. https://doi.org/10.1038/cr.2011.22.

Bedford, Mark T., and Steven G. Clarke. 2009. “Protein Arginine Methylation in Mammals: Who, What, and Why.” *Molecular Cell* 33 (1): 1–13. https://doi.org/10.1016/J.MOLCEL.2008.12.013.

Blank-Giwojna, Alena, Anna Postepska-Igielska, and Ingrid Grummt. 2019. “LncRNA KHPS1 Activates a Poised Enhancer by Triplex-Dependent Recruitment of Epigenomic Regulators.” *Cell Reports* 26 (11): 2904-2915.e4. https://doi.org/10.1016/J.CELREP.2019.02.059.

Bonetti, Alessandro, Federico Agostini, Ana Maria Suzuki, Kosuke Hashimoto, Giovanni Pascarella, Juliette Gimenez, Leonie Roos, et al. 2020. “RADICL-Seq Identifies General and Celltype–Specific Principles of Genome-WideRNA-Chromatin Interactions.” *Nature Communications*, 18. https://doi.org/10.1038/s41467-020-14337-6.

Cao, Ru, Liangjun Wang, Hengbin Wang, Li Xia, Hediye Erdjument-Bromage, Paul Tempst, Richard S. Jones, and Yi Zhang. 2002. “Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing.” *Science* 298 (5595): 1039–43. https://doi.org/10.1126/SCIENCE.1076997/SUPPL_FILE/CAOR1076997.SUP.PDF.

Cao, Xiaoyi, Zhangming Yan, Qiuyang Wu, Alvin Zheng, and Sheng Zhong. 2018. “GIVE: Portable Genome Browsers for Personal Websites.” *Genome Biology*, 19–92. https://doi.org/10.1186/s13059-018-1465-6.

Chu, Ci, Qiangfeng Cliff Zhang, Simão Teixeira da Rocha, Ryan A. Flynn, Maheetha Bharadwaj, J. Mauro Calabrese, Terry Magnuson, Edith Heard, and Howard Y. Chang. 2015. “Systematic Discovery of Xist RNA Binding Proteins.” *Cell* 161 (2): 404–16. https://doi.org/10.1016/J.CELL.2015.03.025.

Cohen, Paul A, Anjua Jhingran, Ana Oaknin, and Lynette Denny. 2019. “Cervical Cancer.” *The Lancet* 393 (10167): 169–82. http://clinicaltrials.gov,.

Deldar, Mahshid, Abad Paskeh, Sepideh Mirzaei, Mohammad Hossein Gholami, Ali Zarrabi, Amirhossein Zabolian, Mehrdad Hashemi, et al. 2021. “NC-ND License Cervical Cancer Progression Is Regulated by SOX Transcription Factors: Revealing Signaling Networks and Therapeutic Strategies.” *Biomedicine & Pharmacotherapy* 144: 112335. https://doi.org/10.1016/j.biopha.2021.112335.

Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. 2012. “Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions.” *Nature 2012 485:7398* 485 (7398): 376–80. https://doi.org/10.1038/nature11082.

Dueva, Rositsa, Karen Akopyan, Chiara Pederiva, Davide Trevisan, Soniya Dhanjal, Arne Lindqvist, and Marianne Farnebo. 2019. “Neutralization of the Positive Charges on Histone Tails by RNA Promotes an Open Chromatin Structure.” *Cell Chemical Biology* 26 (10): 1436-1449.e5. https://doi.org/10.1016/J.CHEMBIOL.2019.08.002.

Engreitz, Jesse M., Amy Pandya-Jones, Patrick McDonel, Alexander Shishkin, Klara Sirokman, Christine Surka, Sabah Kadri, et al. 2013. “The Xist LncRNA Exploits Three-Dimensional Genome Architecture to Spread across the X Chromosome.” *Science* 341 (6147). https://doi.org/10.1126/SCIENCE.1237973/SUPPL_FILE/ENGREITZ.SM.PDF.

Fang, Jinchuan, Hai Zhang, and Sufang Jin. 2014. “Epigenetics and Cervical Cancer: From Pathogenesis to Therapy.” *Tumour Biology : The Journal of the International Society for Oncodevelopmental Biology and Medicine* 35 (6): 5083–93. https://doi.org/10.1007/S13277-014-1737-Z.

Fang, Shuangsang, Lili Zhang, Jincheng Guo, Yiwei Niu, Yang Wu, Hui Li, Lianhe Zhao, et al. 2018. “NONCODEV5: A Comprehensive Annotation Database for Long Non-Coding RNAs.” *Nucleic Acids Research* 46 (D1): D308–14. https://doi.org/10.1093/NAR/GKX1107.

Fanucchi, Stephanie, Ezio T Fok, Emiliano Dalla, Youtaro Shibayama, Kathleen Bamp, Erin Y Chang, Stoyan Stoychev, et al. n.d. “Immune Genes Are Primed for Robust Transcription by Proximal Long Noncoding RNAs Located in Nuclear Compartments.” *Nature Genetics*. https://doi.org/10.1038/s41588-018-0298-2.

Ferro, Anabela, Tânia Mestre, Patrícia Carneiro, Ivan Sahumbaiev, Raquel Seruca, and João M Sanches. 2017. “Blue Intensity Matters for Cell Cycle Profiling in Fluorescence DAPI-Stained Images.” *Laboratory Investigation* 97: 615–25. https://doi.org/10.1038/labinvest.2017.13.

Furlan-Magaril, Mayra, Nataliya Soshnikova, Giovanni Messina, Viviana Valadez-Graham, and Mauro Magaña-Acosta. 2009. “Chromatin Remodelers in the 3D Nuclear Compartment” 11: 600615. https://doi.org/10.3389/fgene.2020.600615.

Ghirlando, Rodolfo, and Gary Felsenfeld. 2016. “CTCF: Making the Right Connections.” *Genes & Development* 30 (8): 881–91. https://doi.org/10.1101/GAD.277863.116.

Green, Michael R., and Joseph Sambrook. 2021. “Alkaline Agarose Gel Electrophoresis.” *Cold Spring Harbor Protocols* 2021 (11): pdb.prot100438. https://doi.org/10.1101/PDB.PROT100438.

Guo, Chun Jie, Xu Kai Ma, Yu Hang Xing, Chuan Chuan Zheng, Yi Feng Xu, Lin Shan, Jun Zhang, et al. 2020. “Distinct Processing of LncRNAs Contributes to Non-Conserved Functions in Stem Cells.” *Cell* 181 (3): 621-636.e22. https://doi.org/10.1016/J.CELL.2020.03.006.

Gurer, Dilek Cansu, İpek Erdogan, Ulvi Ahmadov, Merve Basol, Osama Sweef, Gulcin Cakan-Akdogan, and Bünyamin Akgül. 2021. “Transcriptomics Profiling Identifies Cisplatin-Inducible Death Receptor 5 Antisense Long Non-Coding RNA as a Modulator of Proliferation and Metastasis in HeLa Cells.” *Frontiers in Cell and*

Developmental Biology 9 (August): 2248. https://doi.org/10.3389/FCELL.2021.688855/BIBTEX.

Hacisuleyman, Ezgi, Chinmay J. Shukla, Catherine L. Weiner, and John L. Rinn. 2016. "Function and Evolution of Local Repeats in the Firre Locus." *Nature Communications 2016 7:1* 7 (1): 1–12. https://doi.org/10.1038/ncomms11021.

Hanahan, Douglas, and Robert A. Weinberg. 2011. "Hallmarks of Cancer: The Next Generation." *Cell* 144 (5): 646–74. https://doi.org/10.1016/J.CELL.2011.02.013.

Hansen, Anders S, Claudia Cattoglio, Xavier Darzacq, and Robert Tjian. 2018. "Recent Evidence That TADs and Chromatin Loops Are Dynamic Structures." *NUCLEUS* 9 (1): 20–32. https://doi.org/10.7554/eLife.25776.

Hirano, Tatsuya. 2015. "Chromosome Dynamics during Mitosis." *Cold Spring Harbor Perspectives in Biology* 7 (6): 1–15. https://doi.org/10.1101/CSHPERSPECT.A015792.

Holdt, Lesca M., Steve Hoffmann, Kristina Sass, David Langenberger, Markus Scholz, Knut Krohn, Knut Finstermeier, et al. 2013. "Alu Elements in ANRIL Non-Coding RNA at Chromosome 9p21 Modulate Atherogenic Cell Functions through Trans-Regulation of Gene Networks." *PLOS Genetics* 9 (7): e1003588. https://doi.org/10.1371/JOURNAL.PGEN.1003588.

Jiao, Xinlin, Siying Zhang, Jun Jiao, Teng Zhang, Wenjie Qu, Guy Mutangala Muloye, Beihua Kong, Qing Zhang, and Baoxia Cui. 2019. "Promoter Methylation of SEPT9 as a Potential Biomarker for Early Detection of Cervical Cancer and Its Overexpression Predicts Radioresistance." Clinical Epigenetics . 2019. https://doi.org/10.1186/s13148-019-0719-9.

Jukam, David, Charles Limouse, Owen K. Smith, Viviana I. Risca, Jason C. Bell, and Aaron F. Straight. 2019. "Chromatin-Associated RNA Sequencing (ChAR-Seq)." *Current Protocols in Molecular Biology* 126 (1): e87. https://doi.org/10.1002/CPMB.87.

Kerpedjiev, Peter, Nezar Abdennur, Fritz Lekschas, Chuck Mccallum, Kasper Dinkla, Hendrik Strobelt, Jacob M Luber, et al. 2018. "HiGlass: Web-Based Visual Exploration and Analysis of Genome Interaction Maps." *Genome Biology*, 19–125. https://doi.org/10.1186/s13059-018-1486-1.

Kim, Somi, Nam Kyung Yu, and Bong Kiun Kaang. 2015. “CTCF as a Multifunctional Protein in Genome Regulation and Gene Expression.” *Experimental & Molecular Medicine 2015 47:6* 47 (6): e166–e166. https://doi.org/10.1038/emm.2015.33.

Kraft, Katerina, Kathryn E. Yost, Sedona E. Murphy, Andreas Magg, Yicheng Long, M. Ryan Corces, Jeffrey M. Granja, et al. 2022. “Polycomb-Mediated Genome Architecture Enables Long-Range Spreading of H3K27 Methylation.” *Proceedings of the National Academy of Sciences of the United States of America* 119 (22). https://doi.org/10.1073/PNAS.2201883119.

Kuo, Chao-Chung, Sonja H ¨ Anzelmann, Nevcin Sentürk, Sent¨ Sentürk Cetin, Stefan Frank, Barna Zajzon, Jens-Peter Derks, et al. 2019. “Detection of RNA-DNA Binding Sites in Long Noncoding RNAs.” *Nucleic Acids Research* 47 (6): 32. https://doi.org/10.1093/nar/gkz037.

Lee, Jong Hyuk, Edward W. Kim, Deborah L. Croteau, and Vilhelm A. Bohr. 2020. “Heterochromatin: An Epigenetic Point of View in Aging.” *Experimental & Molecular Medicine 2020 52:9* 52 (9): 1466–74. https://doi.org/10.1038/s12276-020-00497-4.

Li, Yue, Junetha Syed, and Hiroshi Sugiyama. 2016. “RNA-DNA Triplex Formation by Long Noncoding RNAs.” *Cell Chemical Biology* 23 (11): 1325–33. https://doi.org/10.1016/J.CHEMBIOL.2016.09.011.

Mariño-Ramírez, Leonardo, Maricel G. Kann, Benjamin A. Shoemaker, and David Landsman. 2005. “Histone Structure and Nucleosome Stability.” *Expert Review of Proteomics* 2 (5): 719. https://doi.org/10.1586/14789450.2.5.719.

Martín-Subero, Jose Ignacio. 2011. “How Epigenomics Brings Phenotype into Being.” *Pediatric Endocrinology Reviews : PER* 9 Suppl 1 (SUPPL. 1): 506–10. https://europepmc.org/article/med/22423506.

Melé, Marta, Kaia Mattioli, William Mallard, David M. Shechner, Chiara Gerhardinger, and John L. Rinn. 2017. “Chromatin Environment, Transcriptional Regulation, and Splicing Distinguish LincRNAs and MRNAs.” *Genome Research* 27 (1): 27–37. https://doi.org/10.1101/GR.214205.116.

Mishra, Kankadeb, and Chandrasekhar Kanduri. 2019. “Understanding Long Noncoding RNA and Chromatin Interactions: What We Know So Far.” *Non-Coding RNA 2019, Vol. 5, Page 54* 5 (4): 54. https://doi.org/10.3390/NCRNA5040054.

Morrison, Olivia, and Jitendra Thakur. 2021. “Molecular Sciences Molecular Complexes at Euchromatin, Heterochromatin and Centromeric Chromatin.” https://doi.org/10.3390/ijms22136922.

Ngo, Greg H.P., Julia W. Grimstead, and Duncan M. Baird. 2021. “UPF1 Promotes the Formation of R Loops to Stimulate DNA Double-Strand Break Repair.” *Nature Communications 2021 12:1* 12 (1): 1–15. https://doi.org/10.1038/s41467-021-24201-w.

Niehrs, Christof, and Brian Luke. 2020a. “Regulatory R-Loops as Facilitators of Gene Expression and Genome Stability.” *Nature Reviews Molecular Cell Biology* 21. https://doi.org/10.1038/s41580-019-0206-3.

———. 2020b. “Regulatory R-Loops as Facilitators of Gene Expression and Genome Stability.” *Nature Reviews Molecular Cell Biology 2020 21:3* 21 (3): 167–78. https://doi.org/10.1038/s41580-019-0206-3.

Nojima, Takayuki, and Nick J Proudfoot. n.d. “Mechanisms of LncRNA Biogenesis as Revealed by Nascent Transcriptomics.” https://doi.org/10.1038/s41580-021-00447-6.

Nojima, Takayuki, Michael Tellier, Jonathan Foxwell, Claudia Ribeiro de Almeida, Sue Mei Tan-Wong, Somdutta Dhir, Gwendal Dujardin, Ashish Dhir, Shona Murphy, and Nick J. Proudfoot. 2018. “Deregulated Expression of Mammalian LncRNA through Loss of SPT6 Induces R-Loop Formation, Replication Stress, and Cellular Senescence.” *Molecular Cell* 72 (6): 970-984.e7. https://doi.org/10.1016/J.MOLCEL.2018.10.011.

Nora, Elphège P., Bryan R. Lajoie, Edda G. Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Piolot, et al. 2012. “Spatial Partitioning of the Regulatory Landscape of the X-Inactivation Centre.” *Nature 2012 485:7398* 485 (7398): 381–85. https://doi.org/10.1038/nature11049.

Oki, Masaya, Hitoshi Aihara, and Takashi Ito. 2007. “Role Of Histone Phosphorylation In Chromatin Dynamics And Its Implications in Diseases.” *Sub-Cellular Biochemistry* 41: 323–40. https://doi.org/10.1007/1-4020-5466-1_14.

Olins, Donald E., and Ada L. Olins. 2003. “Chromatin History: Our View from the Bridge.” *Nature Reviews Molecular Cell Biology 2003 4:10* 4 (10): 809–14. https://doi.org/10.1038/nrm1225.

Parthun, M R. 2007. “Hat1: The Emerging Cellular Roles of a Type B Histone Acetyltransferase.” *Oncogene* 26: 5319–28. https://doi.org/10.1038/sj.onc.1210602.

Paul, Aswathy Mary, Madhavan Radhakrishna Pillai, and Rakesh Kumar. 2021. “Prognostic Significance of Dysregulated Epigenomic and Chromatin Modifiers in Cervical Cancer.” https://doi.org/10.3390/cells10102665.

Peters, Antoine H.F.M., Dónal O’Carroll, Harry Scherthan, Karl Mechtler, Stephan Sauer, Christian Schöfer, Klara Weipoltshammer, et al. 2001. “Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability.” *Cell* 107 (3): 323–37. https://doi.org/10.1016/S0092-8674(01)00542-6.

Radman-Livaja, Marta, and Oliver J. Rando. 2010a. “Nucleosome Positioning: How Is It Established, and Why Does It Matter?” *Developmental Biology* 339 (2): 258–66. https://doi.org/10.1016/J.YDBIO.2009.06.012.

———. 2010b. “Nucleosome Positioning: How Is It Established, and Why Does It Matter?” *Developmental Biology* 339 (2): 258–66. https://doi.org/10.1016/J.YDBIO.2009.06.012.

Rea, Stephen, Frank Eisenhaber, Do O Â nal, Brian D Strahl, Zu-Wen Sun, Manfred Schmid, Susanne Opravil, et al. 2000. “Regulation of Chromatin Structure by Site-Speci®c Histone H3 Methyltransferases.” *NATURE*. Vol. 406. www.nature.com593.

Sikorska, Natalia, and Tom Sexton. 2020. “Defining Functionally Relevant Spatial Chromatin Domains: It Is a TAD Complicated.” *Journal of Molecular Biology* 432 (3): 653–64. https://doi.org/10.1016/J.JMB.2019.12.006.

Simon, Matthew D., Charlotte I. Wang, Peter v. Kharchenko, Jason A. West, Brad A. Chapman, Artyom A. Alekseyenko, Mark L. Borowsky, Mitzi I. Kuroda, and Robert E. Kingston. 2011a. “The Genomic Binding Sites of a Noncoding RNA.” *Proceedings of the National Academy of Sciences of the United States of America* 108 (51): 20497–502. https://doi.org/10.1073/PNAS.1113536108/-/DCSUPPLEMENTAL/PNAS.1113536108_SI.PDF.

———. 2011b. “The Genomic Binding Sites of a Noncoding RNA.” *Proceedings of the National Academy of Sciences of the United States of America* 108 (51): 20497–502. https://doi.org/10.1073/PNAS.1113536108/-/DCSUPPLEMENTAL/PNAS.1113536108_SI.PDF.

Sriram, Kiran, Zhijie Qi, Sheng Zhong, and Zhen Chen. 2020. “Mitochondrial RNA-Chromatin Interactome Regulates Endothelial-Mesenchymal Transition.” *The FASEB Journal* 34 (S1): 1–1. https://doi.org/10.1096/FASEBJ.2020.34.S1.02664.

Statello, Luisa, Chun-Jie Guo, Ling-Ling Chen, and Maite Huarte. n.d. “Gene Regulation by Long Non-Coding RNAs and Its Biological Functions.” *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/s41580-020-00315-9.

Sun, Ruili, abcd Changfei Qin, abc Binyuan Jiang, abc Shujuan Fang, abc Xi Pan, abc Li Peng, abc Zhaoyang Liu, abc Wenling Li, abc Yuehui Li abc, and Guancheng Li. 2016. “Down-Regulation of MALAT1 Inhibits Cervical Cancer Cell Invasion and Metastasis by Inhibition of Epithelial-Mesenchymal Transition †.” *Mol. BioSyst* 12: 952. https://doi.org/10.1039/c5mb00685f.

Sung, Hyuna, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. 2021. “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.” *CA: A Cancer Journal for Clinicians* 71 (3): 209–49. https://doi.org/10.3322/CAAC.21660.

Tessarz, Peter, and Tony Kouzarides. 2014. “Histone Core Modifications Regulating Nucleosome Structure and Dynamics.” https://doi.org/10.1038/nrm3890.

Tremethick, David J. 2007. “Higher-Order Structures of Chromatin: The Elusive 30 Nm Fiber.” *Cell* 128 (4): 651–54. https://doi.org/10.1016/J.CELL.2007.02.008.

Uszczynska-Ratajczak, Barbara, Julien Lagarde, Adam Frankish, Roderic Guigó, and Rory Johnson. 2018. “Towards a Complete Map of the Human Long Non-Coding RNA Transcriptome.” *Nature Reviews Genetics 2018 19:9* 19 (9): 535–48. https://doi.org/10.1038/s41576-018-0017-y.

Vos, Seychelle M, Lucas Farnung, Marc Boehning, Christoph Wigge, Andreas Linden, Henning Urlaub, and Patrick Cramer. 2018. “Structure of Activated Transcription Complex Pol II-DSIF-PAF-SPT6.” *Nature*. https://doi.org/10.1038/s41586-018-0440-4.

Waldman, Todd. 2020. “Emerging Themes in Cohesin Cancer Biology.” *Nature Reviews Cancer 2020 20:9* 20 (9): 504–15. https://doi.org/10.1038/s41568-020-0270-1.

Wang, Cheng, Xiqiang Liu, Zujian Chen, Hongzhang Huang, Yi Jin, Antonia Kolokythas, Anxun Wang, Yang Dai, David T.W. Wong, and Xiaofeng Zhou. 2013. "Polycomb Group Protein EZH2-Mediated E-Cadherin Repression Promotes Metastasis of Oral Tongue Squamous Cell Carcinoma." *Molecular Carcinogenesis* 52 (3): 229–36. https://doi.org/10.1002/MC.21848.

Wang, Kevin C, Yul W Yang, Bo Liu, Amartya Sanyal, Ryan Corces-Zimmerman, Yong Chen, Bryan R Lajoie, et al. 2011. "A Long Noncoding RNA Maintains Active Chromatin to Coordinate Homeotic Gene Expression." *Nature*. https://doi.org/10.1038/nature09819.

Wu, Weixin, Zhangming Yan, Tri C Nguyen, Zhen Bouman Chen, Shu Chien, and Sheng Zhong. 2019. "Mapping RNA-Chromatin Interactions by Sequencing with IMARGI." *Nature Protocols* 14 (11): 3243–72. https://doi.org/10.1038/s41596-019-0229-4.

Yan, Zhangming, Norman Huang, Weixin Wu, Weizhong Chen, Yiqun Jiang, Jingyao Chen, Xuerui Huang, et al. 2019. "Genome-Wide Colocalization of RNA–DNA Interactions and Fusion RNA Pairs." *Proceedings of the National Academy of Sciences of the United States of America* 116 (8): 3328–37. https://doi.org/10.1073/PNAS.1819788116/SUPPL_FILE/PNAS.1819788116.SAPP.PDF.

Yao, Wen, Yang Wang, Run-Wen Yao, and Ling-Ling Chen. n.d. "Cellular Functions of Long Noncoding RNAs." *Nature Cell Biology*. https://doi.org/10.1038/s41556-019-0311-8.

Yin, Yafei, J Yuyang Lu, Xuechun Zhang, Wen Shao, Yanhui Xu, Pan Li, Yantao Hong, et al. 2020. "U1 SnRNP Regulates Chromatin Retention of Noncoding RNAs." *Nature* 580: 147. https://doi.org/10.1038/s41586-020-2105-3.

Zhou, Bing, Xiao Li, Daji Luo, Do-Hwan Lim, Yu Zhou, and Xiang-Dong Fu. 2019a. "GRID-Seq for Comprehensive Analysis of Global RNA-Chromatin Interactions." *Nature Protocols* 14: 2036–68. https://doi.org/10.1038/s41596-019-0172-4.

———. 2019b. "GRID-Seq for Comprehensive Analysis of Global RNA-Chromatin Interactions." *Nature Protocols* 14 (11): 2036–68. https://doi.org/10.1038/s41596-019-0172-4.

APPENDIX A

LIST OF REAGENTS

1 M Glycine
3.74 g Glycine is dissolved in 50 mL Ultra Pure distilled H_2O.

0.5 M Glycine
25 mL of 1M glycine is mixed with 25 mL Ultra Pure distilled H_2O.

50x Protease Inhibitor Cocktail
One protease inhibitor cocktail tablet (Roche) is dissolved in 1 mL of Ultra Pure distilled H_2O and aliquoted into 20 centrifuge tubes.

0.1 M DTT
5 mL of 1M DTT (Sigma Aldrich) was diluted with 4.5 mL Ultra Pure distilled H_2O.

10 mM dATP
50 µL of 100 mM dATP (NEB) was diluted with 450 µL Ultra Pure distilled H_2O.

1 M NaOH
4 g NaOH pellet (Merck) is mixed with 100 mL Ultra Pure distilled H_2O.

0.1 M NaOH
5 mL of 1M NaOH was diluted with 45 mL Ultra Pure distilled H_2O.

1M HCl
4.93 mL of 37% HCl (Merck) was mixed with Ultra Pure distilled H2O to have a 50 mL total solution.

10% (vol/vol) Tween 20
10 µL of 100% Tween 20 (Applichem) was mixed with 90 µL Ultra Pure distilled H_2O.

70% Ethanol
35 mL of 100% Ethanol (Merck) was mixed with 15 mL Ultra Pure distilled H_2O.

20% Nonidet P 40
20 µL of 100% Nonidet P 40 (Applichem) was mixed with 80 µL of Ultra Pure distilled H_2O.

5 M NaCl
292.2 g of NaCl (Applichem) was dissolved with Ultra Pure distilled H_2O to have 1 L of a total solution.

1 M NaCl
50 mL of 5 M NaCl was dissolved with 50 mL Ultra Pure distilled H_2O.

1M Tris-HCl (pH 7.6)
12.11g of Tris Base (Applichem) was mixed with 80 mL Ultra Pure distilled H_2O. The pH was adjusted with 37% HCl solution. The volume was filled with Ultra Pure distilled H_2O to have a 50 mL solution.

1M Tris-HCl (pH 6.5)
12.11g of Tris Base (Applichem) was mixed with 80 mL Ultra Pure distilled H_2O. The pH was adjusted with 37% HCl solution. The volume was filled with Ultra Pure distilled H_2O to have a 50 mL solution.

1x Cell Lysis Buffer
10 µL of 1M Tris-HCl (pH 7.5), 10 µL of 1 M NaCl, 10 µL of 20% Nonidet P 40, 20 µL of 50X protease inhibitor cocktail and 950 µL Ultra Pure distilled H_2O were mixed.

1x CutSmart buffer
100 µL 10x CutSmart buffer (NEB) was diluted with 950 µL Ultra Pure distilled H_2O.

20% (wt/vol) SDS
20 gr SDS (Applichem) is dissolved in 100 mL of Ultra Pure distilled H_2O.

0.5% (wt/vol) SDS
5 µL of 20% SDS was diluted with 100 µL 1x CutSmart buffer, and 95 µL Ultra Pure distilled H_2O.

10% (vol/vol) Triton X-100
20 µL of 100% Triton X-100 (Applichem) is mixed with 100 µL 1x CutSmart buffer, and 80 µL of Ultra Pure distilled H_2O.

5x PNK phosphatase buffer
350 µL of 1M Tris-HCl, (pH 6.5), 0.5 mL of 1 M $MgCl_2$, 100 µL of 0.1 M DTT, and 500 µL of Ultra Pure distilled H_2O was mixed.

PNK wash buffer
1 mL of 1 M Tris-HCl (pH 7.5), 0.5 mL of 1 M $MgCl_2$, 0.1 mL of 100% Tween 20, and 48.4 mL Ultra Pure distilled H_2O were mixed.

Extraction buffer
25 µL of 1 M Tris-HCl (pH 7.5), 25 µL of 20% SDS, 1 µL of 0.5 M EDTA, 10 µL of 5 M NaCl, 25 µL of proteinase K, and 414 µL of Ultra Pure distilled H_2O were mixed.

2x B&W buffer
300 µL 1 M Tris-HCl (pH 7.5), 60 µL 0.5 M EDTA, 12 mL of 5 M NaCl, 60 µL of 10% Tween 20, and 55.8 mL of Ultra Pure distilled H_2O.

1x B&W buffer
10 mL of 2x B&W buffer was diluted with 10 mL of Ultra Pure distilled H_2O.

Denaturing buffer
100 µL of 1 M NaOH, 1 µL of 100 mM EDTA, and 899 µL of Ultra Pure distilled H_2O were mixed.

High-salt biotin wash buffer
3 mL of 1 M Tris-HCl (pH 7.5), 240 mL of 5 M NaCl, 600 µL of Tween 20, 600 µL of 0.5 M EDTA, and 55.8 mL Ultra Pure distilled H_2O were mixed.