The role of exonic lncRNAs in cancer and targets for cancer gene therapy

MO, junling

Japan Advanced Institute of Science and Technology

Doctoral Dissertation

The role of exonic lncRNAs in cancer and targets for cancer gene therapy

MO, junling

Supervisor: Toshifumi Tsukahara

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
Materials science
March 2023

Doctoral Dissertation

The role of exonic lncRNAs in cancer and targets for cancer gene therapy

Main theme Advisor	Professor Toshifumi Tsukahara
Sub-theme Advisor	Professor Takahiro Hohsaka

Judging Committee

Head of Committee	Professor Dr. Toshifumi Tsukahara
Internal Member	Professor Dr. Kenzo Fujimoto
Internal Member	Professor Dr. Takahiro Hohsaka
Internal Member	Associate Professor Dr. Takumi Yamaguchi
External Member	Professor Dr. Kunio Matsumoto

Mo, junling

s2020035

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
Materials science
November 2022

Abstract

In recent years, more and more attention has been paid to gene therapy. To provide targets and methods for gene therapy, we introduced the screening method of exonic lncRNA, the mechanism of lencRNA *PRKDC-210* and the feasibility of Double duplex invasion DNA for gene therapy in this paper.

In recent years, long noncoding RNAs (lncRNAs) have received increasing attention and have been reported to be associated with various genetic abnormalities. However, the functions of many lncRNAs, including those of long exonic noncoding RNAs (lencRNAs), have not yet been elucidated. Here, we used a novel tethering luciferase assay to analyze the transcriptional regulatory functions of five lencRNAs that are upregulated in cancer. We found that the lencRNA PRKDC-210 interacts with MED12, a component of the CDK8 complex, to regulate the transcription of several genes. The transcriptional activation ability of PRKDC-210 was abolished in siRNA-treated CDK8-depleted cells. We also confirmed the enrichment of PRKDC-210 on RNA polymerase II. RNA-seq analysis of cells in which PRKDC-210 or PRKDC mRNA was knocked down using antisense oligonucleotides revealed that PRKDC-210 can affect the expression levels of genes related to fatty acid metabolism. Finally, we used a ChIRP assay to examine PRKDC-210-enriched sites in the genome. Overall, our findings demonstrate that the lencRNA PRKDC-210 promotes transcription through the CDK8 complex pathway at the transcription initiation site. We propose that PRKDC-210 can affect the transcription of adjacent genes after its transcription and splicing.

Cancer is the deadliest disease now. Although there is a lot of research on cancer treatment, cancer is still difficult to cure. Here, we show a new anti gene that could be used for cancer treatment. Anti gene (double duplex invasion DNA) was developed by Fujimoto Lab in JAIST. This is a photo-cross-linking oligonucleotide toward the target gene by induction at 385 nm UV irradiation. To assess its potential in cancer treatment, we investigated the effect of this anti gene on cells. The EGFP gene was used as a target gene to evaluate the effect of the anti gene in cells. Because it's easier to observe. We used an EGFP stable expressing HeLa cell line in this study. The viability of cells after anti gene activation was confirmed by MTT assay. To assess apoptosis, the contents of caspase-3 and cleaved caspase-3 were detected by western blotting. We detected a large amount of caspase-3 and cleaved caspase-3 in the cells treated with anti gene. This proves that anti gene triggered apoptosis. Collectively, double duplex invasion DNA must be activated by 385nm UV radiation to photo-cross-linking with the target gene and trigger apoptosis. I think we can use this anti gene to target cancer-specific genes and use UV to control its activation. In this way, we can kill cancer cells and cause less damage to normal cells.

Keyword: lncRNA, transcriptional regulation, gene therapy, anti gene

Contents

List of Figures

List of Tables

Chapter 1
General introduction

Introduction

1.1 Gene expression of the central dogma

The process of gene expression was first summarized in 1958 when Francis Crick published the theory of the Central Dogma [1]. Gene expression is the process by which genetic information is transmitted, and it produces the functional molecules needed for life. In addition to well-known proteins, functional non-coding RNAs (such as rRNA and tRNA) are also produced by this process [2].

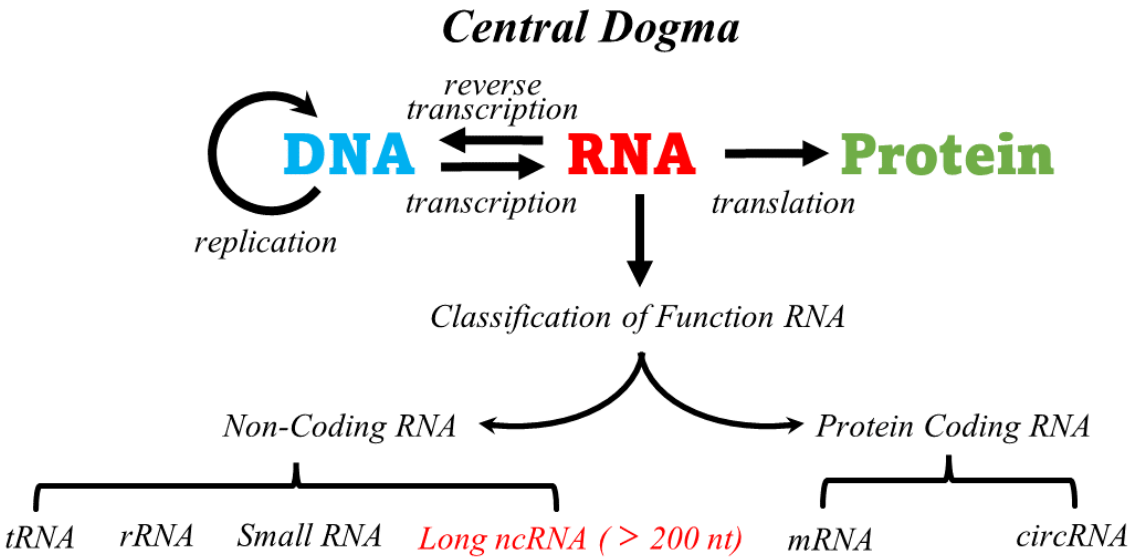

Figure 1: Central Dogma and classification of functional RNA. In cells, DNA replicates itself, and in some viruses, RNA can be reverse transcribed into DNA. DNA is used as a template for transcription to RNA, which is translated into proteins by codon. DNA is initially

transcribed as preRNA, but RNA is spliced to mature mRNAs or ncRNAs. Most of the known ncRNAs are directly synthesized by transcription. However, recently, many ncRNAs, especially lncRNAs, have been found to be generated by the splicing process[3].

Gene expression is affected by many factors. Under different morphologies of cells of the same genotype (i.e., cell differentiation), genes will be differentiated due to different environments during development and epigenetic changes at later stages. In the process of gene expression, there are corresponding regulatory mechanisms in different stages (Transcription and translation, and so on.), and these regulatory mechanisms are not fully understood[4–6]. However, the regulation between different stages may be interdependent, which needs more research[7].

1.2 Gene regulation

1.2.1 Transcriptional regulation

Transcriptional regulation refers to the regulation of the process of DNA transcription to produce RNA. Transcription is regulated in many ways, including the regulation of chromatin structure level (such as histone modifications) and the DNA level (such as modifications that recognize sequences or transcription factors)[8–11].

Transcriptional regulation is ubiquitous in prokaryotes and eukaryotes, although their mechanisms are quite different[12,13]. This regulatory mechanism allows cells to adapt to different environments. The lactose operon is one of the most common transcriptional regulatory mechanisms in prokaryotes. It can regulate the transcription of genes according to the concentration of lactose in the environment. This mechanism is also widely used in genetic engineering[14].

Transcriptional regulation involves all transcription processes, including elements such as RNA polymerases, transcription factors and enhancers. In this study, we focused on the transcription process in eukaryotes. The transcription system of eukaryotes has three RNA polymerases, RNA polymerase (I, II and III), and most of the RNA is transcribed by RNA polymerase II, including mRNA[15,16]. RNA polymerases are interacted with and regulated by transcription factors and attach to the promoter region of DNA. After transcription is completed, RNA polymerases are acted on by other transcription regulators and detach from DNA[17,18].

In eukaryotic cells, genomic DNA is highly compressed into chromosomes[19]. Double-stranded DNA is wrapped around histone octamers to form chromatin, the natural state of genes, in which RNA polymerases cannot bind DNA for

transcription. Through the regulation mechanism of histones, DNA can be transformed into transcribable. In addition to transcription, histones can affect epigenetics, DNA replication, and DNA repair functions and so on. There are many enzymes involved in histone regulation, such as histone acetyltransferases, deacetylases, methyltransferases, and kinases et al[20–22]. Modification of methylated of the histone octamer tail sequence can compress staining and lead to gene silencing, which is one of the important mechanisms leading to female x chromosome inactivation[23]. Different from the regulation of histones, the regulation of transcription factors is generally at the DNA level. Transcription factors are regulated in many ways, such as the methylation of C≡G base pairs will affect the binding of transcription factors to DNA[24]. Transcription factors directly affect the activity of RNA polymerase to synthesize RNA[25]. They have both activation and inhibition effects. Activators of transcription enhance the interaction between polymerase and DNA strands, while repressors of transcription do the opposite[26,27]. The joint action of these transcription factors can control the expression and silencing of genes. They will selectively act on DNA with the change of cell environment, and the change of their own concentration will also affect transcription, which is also one of the important

mechanisms of cell differentiation[28].

Almost all transcription factors need to interact with other transcription factors, and many transcription factors interact with homotypes or heterotypes of the same family to form polymers. Most of these transcription factor complexes need to bind to the corresponding DNA sequence to regulate transcription[29]. The binding between transcription factors and between transcription factors and DNA can only occur when transcription factors are in an activated. The function of cofactors is to regulate the protein of transcription factors[30]. Cofactor regulation is not specific to a single gene or transcription factor. For example, the cofactors of certain steroid receptors and the cofactors of the inflammatory factor NF-κB are universal, and such steroid receptors can affect the inflammatory response of certain tissues[31].

1.2.2 Post-transcriptional regulation

First, DNA is transcribed into preRNA, and then preRNA becomes functional RNA after various modifications[32]. These modifications can be used as regulatory checkpoints to detect or control the preRNA to mRNA process.

Splicing is the most widely known RNA modification. This process happens during transcription or shortly after transcription, and it completes in the

nucleus[33]. The splicing complex recognizes the RNA sequence and excise part of the sequence (usually introns), leaving the desired sequence (usually exons) to form mature mRNA. The exons of the spliced RNA are arranged in different sequences and combinations, resulting in different splice variants, a process known as alternative splicing[34]. These variants can code different proteins, which makes the diversity of genes. Moreover, the process of alternative splicing also produces intron-retaining RNAs, which are usually lncRNAs, and the formation mechanism of intron-retaining lncRNAs is still unclear.

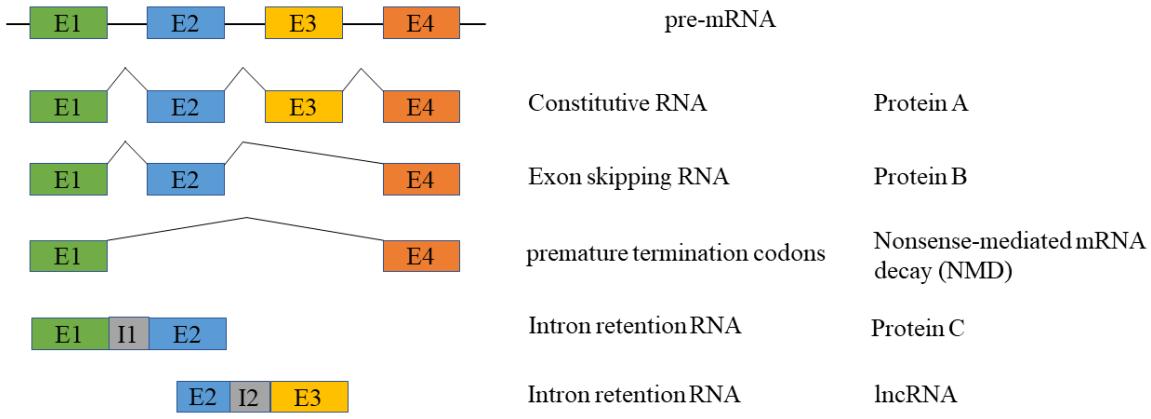

Figure 2: Types of RNA produced by alternative splicing. Alternative splicing of RNA removes introns and arranges exons in a different order to form different mature RNAs. Some of these different RNA variants are translated into different proteins, but some are also degraded by premature stop codons or via the nonsense-mediated decay (NMD)

pathway. In addition, intron retention may also occur, which can cause RNAs to lack open reading frames (ORFs), resulting in noncoding RNAs that cannot encode proteins. Reference: https://nanoporetech.com/applications/investigation/splice-variation.

After preRNA is transcribed, methyl guanosine is added to the 5' end of preRNA and the polyadenylation(Poly (A)) tail is added to the 3' end of preRNA. This modification is very important for RNA stability. The 5' -end cap modification removes the phosphodiester bond and protects the mature mRNA from degradation. And the 3' -end tail modification prevents exonuclease from entering the coding sequence, the length of the Poly(A) tail also determines the RNA life cycle[35]. These modifications all affect the level of RNA translation into protein. The 5' cap also guides RNA from the nucleus into the cytoplasm and localise zes to the translation initiation unit.

In addition, RNA editing also occurs at the post-transcriptional regulatory stage. RNA editing is able to alter the sequence of RNA so that the protein expressed by the RNA is changed, expression is silenced or reactivated. RNA editing is the modification of the original RNA sequence by insertion, deletion or replacement of single or multiple bases. Several enzymes are known to edit RNA in cells, such

as apolipoprotein B mRNA editing enzyme, catalytic polypeptide (APOBEC), which can cause Cytosine to Uracil editing of single-stranded RNA[36], and adenosine deaminase acting on RNA (ADAR), which causes double-stranded RNA Adenine to Inosine editing of RNA[37], Inosine can be recognized as a Guanine resulting in translation termination or amino acid change.

1.2.3 Translational regulation

The translation is the process of converting mRNA into protein, and translational regulation is the regulation of the amount of protein produced from mRNA. Generally speaking, translation regulation has two checkpoints, one is the stability of mRNA, and the other is the translation process itself. The stability of mRNA is determined by multiple factors. The RNA modification mentioned above regulates the stability of RNA in the nucleus. When mRNA enters the cytoplasm from the nucleus, it might be decomposed by microRNA or small RNA. This mechanism is called RNA interference[38,39]. The rate of protein production is affected by the activity of the ribosome. eIF-2 and eIF-4E are the main factors affecting the translation rate of ribosomes, and phosphorylation of translation initiation factors can activate them to increase the translation rate[40,41]. The 5' end of the mRNA is recognized by the ribosomal 30s complex and looks for the

start codon AUG. When the ribosomal 30s complex recognizes the start codon, methionyl tRNA (Met-tRNA) will be recruited to the ribosomal 30s subunit, and the ribosomal 50s subunit will also bind to the complex to form the final translation complex[42]. After the successful assembly of the translation complex, the peptide chain will be extended from Met under the regulation of various elongation factors. Finally, the complex is decomposed when it recognizes the stop codon (UAA, UAG, and UGA), and the polypeptide chain folds to form a higher-order structure(HOS). Multiple ribosomes can act on one mRNA chain at the same time, thus increasing the speed of translation. Ribosome activity determines the amount of protein produced[40]. Ribosomal skipping also occurs during translation, where the ribosomes skip the codon and recognize the codon that follows, which also leads to the diversity of proteins[43,44].

1.3 Long non-coding RNA

Noncoding RNAs (ncRNAs) are ubiquitous in eukaryotes, and those larger than 200 nt are classified as long noncoding RNAs (lncRNAs) [45,46]. lncRNAs are very abundant in *human* cells, but the expression levels of lncRNAs in different cells are not the same[47,48]. Compared with other cells, testis and neuronal cells contain many lncRNAs. Moreover, most lncRNAs are retained across species,

suggesting that lncRNAs are a common and important functional molecule[49].

Of course, the expression levels of lncRNAs vary among species.

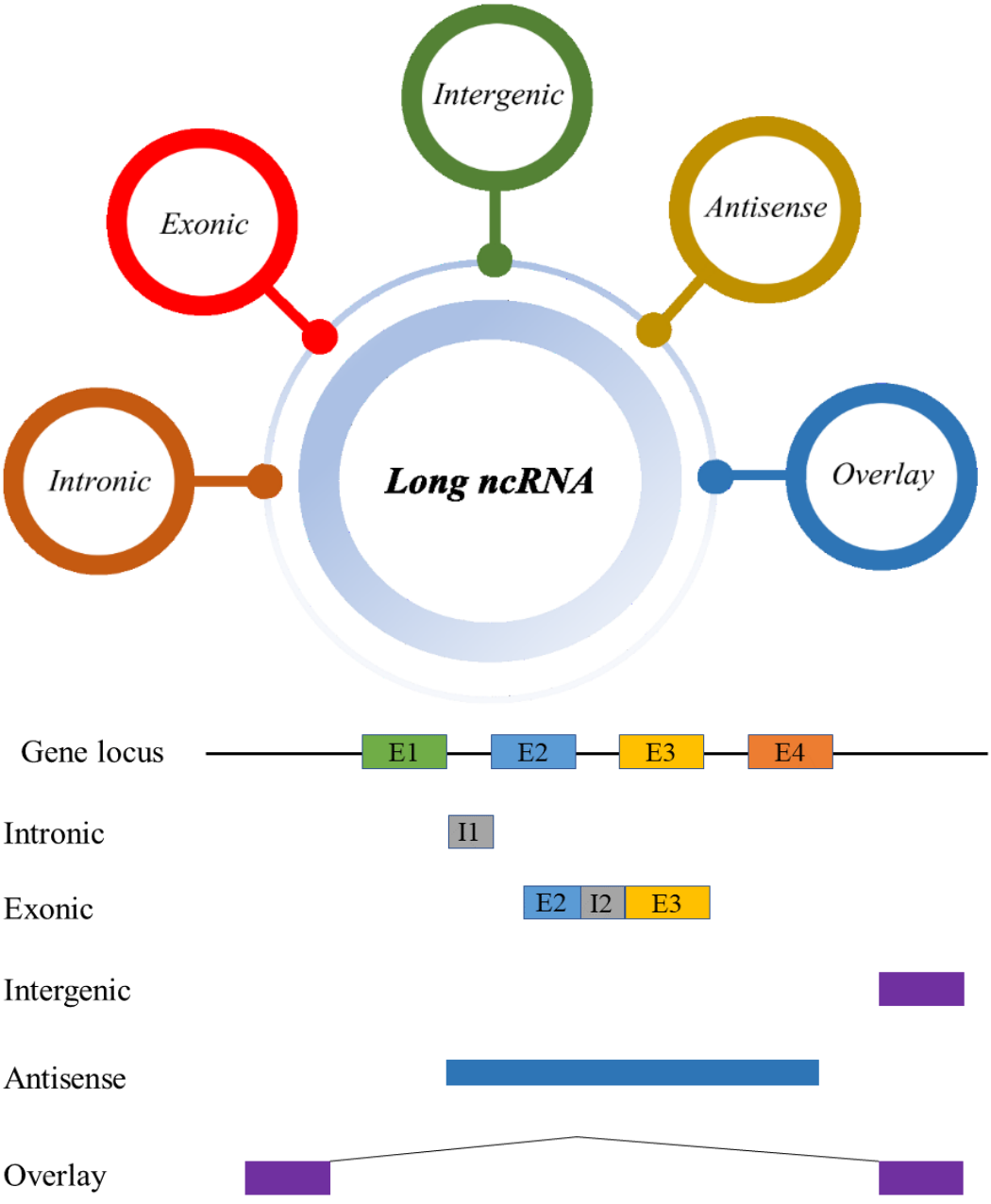

Figure 3: Types of lncRNA. Long non-coding RNAs can be classified into 5 types according to their position on the genome. Intronic is composed

only of introns of protein-coding genes. Exonic is composed of exons and introns of protein-coding genes. Intergenic consists of sequences between two protein-coding genes. Antisense is the transcriptional production of antisense strands of protein-coding genes. The overlay is composed of non-coding regions upstream and downstream of coding genes.

According to their genomic position, lncRNAs are further classified into five types, namely, intronic, exonic, intergenic, antisense, and overlay, although the vast majority of lncRNA research has focused on long intergenic noncoding RNAs (lincRNAs)[50–52]. benefited from the development of next-generation sequencing technology (NGS), many lncRNAs have been confirmed to be retained in mammals. More and more evidence has shown that these preserved lncRNAs have important functions, but there are very few lncRNAs whose functions have been clarified[53,54]. In the past, lncRNAs have received attention in epigenetic regulation. A series of lncRares is involved in the X chromosome inactivation (XCI) mechanism that controls the X chromosome imbalance between males and females, among which the 17 kbp X inactivation-specific transcript (Xist) is involved in the recruitment of silencing factors (e.g., PRC2)

renders the chromosome into transcriptional inactivation[55]. In recent years, lncRNAs have been reported[56,57] to perform various functions in cells through RNA–protein interactions and are also associated with genetic abnormalities in diseases such as cancer[58,59].

1.4 Genetic disorders

Genetic disorders can be divided into two types (1) One is a congenital genetic disease, which is usually born with a certain disease due to genetic abnormalities of the parents. (2) One is the disease caused ban y acquired gene mutation. Most genetic disorders can be inherited or caused by mutations. Such as high blood pressure, diabetes, Alzheimer's, and cancer(https://www.genome.gov/For-Patients-and-Families/Genetic-Disorders, accessed on 1 September 2022). Genetic diseases can be caused by single or multiple genes. Recent studies have shown that alternative splicing of RNA is one of the main causes of genetic diseases[60]. Because the alternative splicing of pre-RNA leads to the diversity of proteins and the potential generation of ncRNA, it may directly lead to the loss of original functional proteins or the generation of irrelevant and harmful proteins[61]. Among them, lncRNA has been reported as a biomarker of potential genetic diseases. Especially in the field of cancer, there are more and more reports

on the involvement of lncRNA in the pathogenesis of cancer[62–64]. Therefore, research on the relationship between the mechanism of lncRNA itself and diseases is highly anticipated.

1.5 Cancer

Cancer is one of the deadliest diseases today. The number of deaths due to cancer is increasing year by year, and the number of deaths due to cancer in 2021 has approached 10 million. Although the death rate from cancer is decreasing, but the number of people diagnosed with cancer is still increasing (https://www.who.int/news-room/fact-sheets/detail/cancer, accessed on 20 October 2022). Cancer is a disease caused by damage or mutation in one's genes, not an infection. Although some cancers are caused by viral infections, but other factors are required to eventually develop into cancer. When normal cells become cancerous, the cells proliferate indefinitely, invade surrounding tissues, and rob the host of nutrients. These hyperproliferative cells can cause damage to blood vessels and internal organs, thereby jeopardizing the health of the host and causing great suffering (https://www.cancer.gov/about-cancer/understanding/what-is-cancer, accessed on 20 October 2022). There are many types of cancer treatments, such as chemotherapy, hormone therapy,

immunotherapy, radiation therapy, surgery, and targeted therapy(https://www.cancer.gov/about-cancer/treatment/types, accessed on 20 October 2022). Different cancers require different treatments or combinations of treatments. But almost all treatment methods will cause great damage to the patient itself, and it is difficult to cure. The pain caused by the process of cancer treatment may be similar to the pain caused by the er itself. Therefore, the research on cancer treatment still needs to be improved, and the optimization of current treatment methods and the development of new treatment methods are expected.

1.6 Gene therapy

Traditional gene therapy uses exogenous genes to supplement the mutated genes[65]. Generally, gene therapy uses vectors (e.g., viruses and plasmids) that carry therapeutic gene sequences. The treatment of adenosine deaminase deficiency by gene therapy was first reported in 1990, which opened the way for gene therapy[66,67]. After that, in 1999, Onodera's team improved the treatment of adenosine deaminase deficiency, making gene therapy more recognized[68]. In addition, miRNA and Antisense based gene therapy has been used in clinical therapy. Treatment based on miRNA and Antisense can regulate the

concentration of protein in cells by regulating the pathway of RNA to protein and can be used for the treatment of diseases with abnormal RNA transcription or abnormal protein expression[69–71]. Although still experimental, RNA editing techniques have been attempted for gene therapy, where ADAR-directed A to I editing has been found to occur in a variety of cancers[72,73]. The initiation of apoptosis by gene therapy has also been investigated as a potential approach for cancer therapy. Knockdown the of PTEN gene using shRNA to induce apoptosis has been reported[74].

1.7 Anti gene

Anti gene is a nucleic acid or nucleic acid-like sequence that specifically binds to target nucleic acid(DNA or RNA) and inhibits gene expression. Anti gene is complementary to nucleic acid through the specific sequence and uses special modification to affect the target nucleic acid sequence. Anti gene has the following types of action: (1) Combine with template DNA to form duplex or triplex to inhibit transcription or replication. (2) Combining with nascent RNA to form triple inhibit sequent modification and translation of RNA. (3) Combined with RNA to form DNA-RNA double-strand, which is recognized and degraded by RNase H[75–77]. Recently, a study by Flanga et al. reported the use of Peptide

Nucleic Acids (PNAs) to suppress cancer genes[78]. Previously, they also reported results using Locked Nucleic Acids (LNAs) on biological activity[79]. The use of Anti genes for gene therapy has gradually attracted the attention of researchers, and the development of Anti genes is imperative.

1.8 Photo cross link DNA

Photo cross linking is a covalent bonding structure between molecules due to chemical structural changes induced by light of specific wavelengths[80]. In particular, [2+2] photocyclization is a reversible reaction that enables controlled binding and cleavage of chemical structures using specific wavelengths of light irradiation. The covalent bond structure created by the binding of this light-responsive molecule to the DNA molecule cannot be separated by helicases in the cell. The use of oligodeoxyribonucleotides (ODNs) for photo cross linking with DNA has been reported as early as 2004[81]. In 2013, the photo cross linking of p-stilbazole and DNA was also reported[82]. These reports hint at the possibility of highly operable artificially interfering genes.

1.9 Aim of the study

Current lncRNA research focuses on intergenic lncRNAs (lincRNAs), while the mechanisms of intron-retained exonic lncRNAs (lencRNAs) have been

overlooked. Of course, the structure of lencRNA itself makes it difficult to identify. To this end, we developed a new analysis system—Tethering Luciferase Assay system to screen functional lencRNAs. Because lencRNAs are due to alternative splicing, and the likelihood that this occurs in genetically abnormal cells is very high, we focused our analysis on lencRNAs that are more abundant in cancer. In this study, the functions and mechanisms of lencRNAs screened in the Tethering Luciferase Assay system were analyzed. Furthermore, the effect of Double duplex invasion DNA on cells after photo cross linking with the genome in cells is also described in this article. This could provide a potential therapeutic approach for gene therapy.

The main purpose of this article is to

i) Development of a novel analytical system—Tethering Luciferase Assay system—for analyzing whether ncRNAs have transcriptional regulatory functions. Use this system to screen for functional lencRNAs.

ii) Analysis of the mechanism of lencRNA *PRKDC-210* in *Human* breast cancer cells MCF-7.

iii) To examine the possibility of Double duplex invasion of DNA in cells.

1.10 Expected impact of the study

This study developed a novel analysis method for lncRNA, provided a method for the study of lncRNA mechanism, and provided a basis for the study of lencRNA. At the same time, it also provides targets and methods for gene therapy for gene abnormal diseases.

1.11 Reference

1. CRICK, F. Central Dogma of Molecular Biology. Nature 1970, 227, 561–563, doi:10.1038/227561a0.

2. Gottlieb, G. Normally Occurring Environmental and Behavioral Influences on Gene Activity: From Central Dogma to Probabilistic Epigenesis. Psychol Rev 1998, 105, 792–802, doi:10.1037/0033-295X.105.4.792-802.

3. Srijyothi, L.; Ponne, S.; Prathama, T.; Ashok, C.; Baluchamy, S. Roles of Non-Coding RNAs in Transcriptional Regulation. In Transcriptional and Post-transcriptional Regulation; InTech, 2018.

4. Hausser, J.; Mayo, A.; Keren, L.; Alon, U. Central Dogma Rates and the Trade-off between Precision and Economy in Gene Expression. Nat Commun 2019, 10, 68, doi:10.1038/s41467-018-07391-8.

5. Boyer, L.A.; Lee, T.I.; Cole, M.F.; Johnstone, S.E.; Levine, S.S.; Zucker,

J.P.; Guenther, M.G.; Kumar, R.M.; Murray, H.L.; Jenner, R.G.; et al. Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. Cell 2005, 122, 947–956, doi:10.1016/j.cell.2005.08.020.

6. Davidson, E.; Levin, M. Gene Regulatory Networks. Proceedings of the National Academy of Sciences 2005, 102, 4935–4935, doi:10.1073/pnas.0502024102.

7. Takeuchi, N.; Kaneko, K. The Origin of the Central Dogma through Conflicting Multilevel Selection. Proceedings of the Royal Society B: Biological Sciences 2019, 286, 20191359, doi:10.1098/rspb.2019.1359.

8. Goodrich, J.A.; Tjian, R. Unexpected Roles for Core Promoter Recognition Factors in Cell-Type-Specific Transcription and Gene Regulation. Nat Rev Genet 2010, 11, 549–558, doi:10.1038/nrg2847.

9. Audia, J.E.; Campbell, R.M. Histone Modifications and Cancer. Cold Spring Harb Perspect Biol 2016, 8, a019521, doi:10.1101/cshperspect.a019521.

10. Merkenschlager, M.; Nora, E.P. CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. Annu Rev Genomics Hum Genet 2016, 17, 17–43, doi:10.1146/annurev-genom-083115-022339.

11. Zhang, Y.; Sun, Z.; Jia, J.; Du, T.; Zhang, N.; Tang, Y.; Fang, Y.; Fang, D.

Overview of Histone Modification. In; 2021; pp. 1–16.

12. Glisovic, T.; Bachorik, J.L.; Yong, J.; Dreyfuss, G. RNA-Binding Proteins and Post-Transcriptional Gene Regulation. FEBS Lett 2008, 582, 1977–1986, doi:10.1016/j.febslet.2008.03.004.

13. Matys, V. TRANSFAC(R) and Its Module TRANSCompel(R): Transcriptional Gene Regulation in Eukaryotes. Nucleic Acids Res 2006, 34, D108–D110, doi:10.1093/nar/gkj143.

14. Lewis, M. Allostery and the Lac Operon. J Mol Biol 2013, 425, 2309–2316, doi:10.1016/j.jmb.2013.03.003.

15. Woychik, N.A.; Young, R.A. RNA Polymerase II: Subunit Structure and Function. Trends Biochem Sci 1990, 15, 347–351, doi:10.1016/0968-0004(90)90074-L.

16. Mooney, R.A.; Landick, R. RNA Polymerase Unveiled. Cell 1999, 98, 687–690, doi:10.1016/S0092-8674(00)81483-X.

17. Gehring, N.H.; Wahle, E.; Fischer, U. Deciphering the MRNP Code: RNA-Bound Determinants of Post-Transcriptional Gene Regulation. Trends Biochem Sci 2017, 42, 369–382, doi:10.1016/j.tibs.2017.02.004.

18. Green, P.J.; Pines, O.; Inouye, M. THE ROLE OF ANTISENSE RNA IN

GENE REGULATION. Annu Rev Biochem 1986, 55, 569–597, doi:10.1146/annurev.bi.55.070186.003033.

19. Clapier, C.R.; Cairns, B.R. The Biology of Chromatin Remodeling Complexes. Annu Rev Biochem 2009, 78, 273–304, doi:10.1146/annurev.biochem.77.062706.153223.

20. Yan, L.; Wu, H.; Li, X.; Gao, N.; Chen, Z. Structures of the ISWI–Nucleosome Complex Reveal a Conserved Mechanism of Chromatin Remodeling. Nat Struct Mol Biol 2019, 26, 258–266, doi:10.1038/s41594-019-0199-9.

21. Yan, L.; Chen, Z. A Unifying Mechanism of DNA Translocation Underlying Chromatin Remodeling. Trends Biochem Sci 2020, 45, 217–227, doi:10.1016/j.tibs.2019.09.002.

22. Hargreaves, D.C.; Crabtree, G.R. ATP-Dependent Chromatin Remodeling: Genetics, Genomics and Mechanisms. Cell Res 2011, 21, 396–420, doi:10.1038/cr.2011.32.

23. Martinowich, K.; Hattori, D.; Wu, H.; Fouse, S.; He, F.; Hu, Y.; Fan, G.; Sun, Y.E. DNA Methylation-Related Chromatin Remodeling in Activity-Dependent Bdnf Gene Regulation. Science (1979) 2003, 302, 890–893, doi:10.1126/science.1090842.

24. Vinson, C.; Chatterjee, R. CG Methylation. Epigenomics 2012, 4, 655–663, doi:10.2217/epi.12.55.

25. Latchman, D.S. Transcription Factors: An Overview. Int J Biochem Cell Biol 1997, 29, 1305–1312, doi:10.1016/S1357-2725(97)00085-X.

26. Edmunds, J.W.; Mahadevan, L.C. MAP Kinases as Structural Adaptors and Enzymatic Activators in Transcription Complexes. J Cell Sci 2004, 117, 3715–3723, doi:10.1242/jcs.01346.

27. Shapiro, M.J.; Shapiro, V.S. Transcriptional Repressors, Corepressors and Chromatin Modifying Enzymes in T Cell Development. Cytokine 2011, 53, 271–281, doi:10.1016/j.cyto.2010.11.013.

28. Stadhouders, R.; Filion, G.J.; Graf, T. Transcription Factors and 3D Genome Conformation in Cell-Fate Decisions. Nature 2019, 569, 345–354, doi:10.1038/s41586-019-1182-7.

29. Fernandez Garcia, M.; Moore, C.D.; Schulz, K.N.; Alberto, O.; Donague, G.; Harrison, M.M.; Zhu, H.; Zaret, K.S. Structural Features of Transcription Factors Associating with Nucleosome Binding. Mol Cell 2019, 75, 921-932.e6, doi:10.1016/j.molcel.2019.06.009.

30. Lee, T.I.; Young, R.A. Transcriptional Regulation and Its Misregulation

in Disease. Cell 2013, 152, 1237–1251, doi:10.1016/j.cell.2013.02.014.

31. Nettles, K.W.; Gil, G.; Nowak, J.; Métivier, R.; Sharma, V.B.; Greene, G.L. CBP Is a Dosage-Dependent Regulator of Nuclear Factor-ΚB Suppression by the Estrogen Receptor. Molecular Endocrinology 2008, 22, 263–272, doi:10.1210/me.2007-0324.

32. Zhao, B.S.; Roundtree, I.A.; He, C. Post-Transcriptional Gene Regulation by mRNA Modifications. Nat Rev Mol Cell Biol 2017, 18, 31–42, doi:10.1038/nrm.2016.132.

33. Tilgner, H.; Knowles, D.G.; Johnson, R.; Davis, C.A.; Chakrabortty, S.; Djebali, S.; Curado, J.; Snyder, M.; Gingeras, T.R.; Guigó, R. Deep Sequencing of Subcellular RNA Fractions Shows Splicing to Be Predominantly Co-Transcriptional in the Human Genome but Inefficient for LncRNAs. Genome Res 2012, 22, 1616–1625, doi:10.1101/gr.134445.111.

34. Neil, C.R.; Fairbrother, W.G. Intronic RNA: Ad'Junk' Mediator of Post-Transcriptional Gene Regulation. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 2019, 1862, 194439, doi:10.1016/j.bbagrm.2019.194439.

35. Halbeisen, R.E.; Galgano, A.; Scherrer, T.; Gerber, A.P. Post-Transcriptional Gene Regulation: From Genome-Wide Studies to Principles.

Cellular and Molecular Life Sciences 2008, 65, 798, doi:10.1007/s00018-007-7447-6.

36. Wedekind, J.E.; Dance, G.S.C.; Sowden, Mark.P.; Smith, H.C. Messenger RNA Editing in Mammals: New Members of the APOBEC Family Seeking Roles in the Family Business. Trends in Genetics 2003, 19, 207–216, doi:10.1016/S0168-9525(03)00054-4.

37. Licht, K.; Jantsch, M.F. The Other Face of an Editor: ADAR1 Functions in Editing-Independent Ways. BioEssays 2017, 39, 1700129, doi:10.1002/bies.201700129.

38. Lyon, K.; Stasevich, T.J. Imaging Translational and Post-Translational Gene Regulatory Dynamics in Living Cells with Antibody-Based Probes. Trends in Genetics 2017, 33, 322–335, doi:10.1016/j.tig.2017.02.003.

39. Sajeev, N.; Bai, B.; Bentsink, L. Seeds: A Unique System to Study Translational Regulation. Trends Plant Sci 2019, 24, 487–495, doi:10.1016/j.tplants.2019.03.011.

40. Istomine, R.; Pavey, N.; Piccirillo, C.A. Posttranscriptional and Translational Control of Gene Regulation in CD4 + T Cell Subsets. The Journal of Immunology 2016, 196, 533–540, doi:10.4049/jimmunol.1501337.

41. Fox, T.D. Translational Control of Endogenous and Recoded Nuclear Genes in Yeast Mitochondria: Regulation and Membrane Targeting. Experientia 1996, 52, 1130–1135, doi:10.1007/BF01952112.

42. Springer, M.; Graffe, M.; Dondon, J.; Grunberg-Manago, M. TRNA-like Structures and Gene Regulation at the Translational Level: A Case of Molecular Mimicry in Escherichia Coli. EMBO J 1989, 8, 2417–2424, doi:10.1002/j.1460-2075.1989.tb08372.x.

43. Luke, G.A.; de Felipe, P.; Lukashev, A.; Kallioinen, S.E.; Bruno, E.A.; Ryan, M.D. Occurrence, Function and Evolutionary Origins of '2A-like' Sequences in Virus Genomes. Journal of General Virology 2008, 89, 1036–1042, doi:10.1099/vir.0.83428-0.

44. Doronina, V.A.; de Felipe, P.; Wu, C.; Sharma, P.; Sachs, M.S.; Ryan, M.D.; Brown, J.D. Dissection of a Co-Translational Nascent Chain Separation Event. Biochem Soc Trans 2008, 36, 712–716, doi:10.1042/BST0360712.

45. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; et al. Landscape of Transcription in Human Cells. Nature 2012, 489, 101–108, doi:10.1038/nature11233.

46. Montalbano, A.; Canver, M.C.; Sanjana, N.E. High-Throughput Approaches to Pinpoint Function within the Noncoding Genome. Mol Cell 2017, 68, 44–59.

47. Uszczynska-Ratajczak, B.; Lagarde, J.; Frankish, A.; Guigó, R.; Johnson, R. Towards a Complete Map of the Human Long Non-Coding RNA Transcriptome. Nat Rev Genet 2018, 19, 535–548, doi:10.1038/s41576-018-0017-y.

48. Jiang, S.; Cheng, S.-J.; Ren, L.-C.; Wang, Q.; Kang, Y.-J.; Ding, Y.; Hou, M.; Yang, X.-X.; Lin, Y.; Liang, N.; et al. An Expanded Landscape of Human Long Noncoding RNA. Nucleic Acids Res 2019, 47, 7842–7856, doi:10.1093/nar/gkz621.

49. Quinn, J.J.; Chang, H.Y. Unique Features of Long Non-Coding RNA Biogenesis and Function. Nat Rev Genet 2016, 17.

50. Chen, S.; Liang, H.; Yang, H.; Zhou, K.; Xu, L.; Liu, J.; Lai, B.; Song, L.; Luo, H.; Peng, J.; et al. LincRNA-P21: Function and Mechanism in Cancer. Medical Oncology 2017, 34.

51. Atianand, M.K.; Hu, W.; Satpathy, A.T.; Shen, Y.; Ricci, E.P.; Alvarez-Dominguez, J.R.; Bhatta, A.; Schattgen, S.A.; McGowan, J.D.; Blin, J.; et al. A

Long Noncoding RNA LincRNA-EPS Acts as a Transcriptional Brake to Restrain Inflammation. Cell 2016, 165, 1672–1685, doi:10.1016/j.cell.2016.05.075.

52. Kuwabara, Y.; Tsuji, S.; Nishiga, M.; Izuhara, M.; Ito, S.; Nagao, K.; Horie, T.; Watanabe, S.; Koyama, S.; Kiryu, H.; et al. Lionheart LincRNA Alleviates Cardiac Systolic Dysfunction under Pressure Overload. Commun Biol 2020, 3, doi:10.1038/s42003-020-01164-0.

53. Kim, E.-D.; Sung, S. Long Noncoding RNA: Unveiling Hidden Layer of Gene Regulatory Networks. Trends Plant Sci 2012, 17, 16–21, doi:10.1016/j.tplants.2011.10.008.

54. Dhanoa, J.K.; Sethi, R.S.; Verma, R.; Arora, J.S.; Mukhopadhyay, C.S. Long Non-Coding RNA: Its Evolutionary Relics and Biological Implications in Mammals: A Review. J Anim Sci Technol 2018, 60.

55. Tsai, M.-C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J.K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H.Y. Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. Science (1979) 2010, 329, 689–693, doi:10.1126/science.1192002.

56. Sun, Q.; Hao, Q.; Prasanth, K. v. Nuclear Long Noncoding RNAs: Key Regulators of Gene Expression. Trends in Genetics 2018, 34, 142–157.

57. Kung, J.T.Y.; Colognori, D.; Lee, J.T. Long Noncoding RNAs: Past, Present, and Future. Genetics 2013, 193, 651–669.

58. Huarte, M. The Emerging Role of LncRNAs in Cancer. Nat Med 2015, 21, 1253–1261.

59. Wapinski, O.; Chang, H.Y. Long Noncoding RNAs and Human Disease. Trends Cell Biol 2011, 21, 354–361.

60. Anna, A.; Monika, G. Splicing Mutations in Human Genetic Disorders: Examples, Detection, and Confirmation. J Appl Genet 2018, 59, 253–268, doi:10.1007/s13353-018-0444-7.

61. Montes, M.; Sanford, B.L.; Comiskey, D.F.; Chandler, D.S. RNA Splicing and Disease: Animal Models to Therapies. Trends in Genetics 2019, 35, 68–87, doi:10.1016/j.tig.2018.10.002.

62. Ghafouri-Fard, S.; Taheri, M. Long Non-Coding RNA Signature in Gastric Cancer. Exp Mol Pathol 2020, 113, 104365, doi:10.1016/j.yexmp.2019.104365.

63. Zhang, X.; Hamblin, M.H.; Yin, K.-J. The Long Noncoding RNA Malat1: Its Physiological and Pathophysiological Functions. RNA Biol 2017, 14, 1705–1714, doi:10.1080/15476286.2017.1358347.

64. Finotti, A.; Fabbri, E.; Lampronti, I.; Gasparello, J.; Borgatti, M.; Gambari, R. MicroRNAs and Long Non-Coding RNAs in Genetic Diseases. Mol Diagn Ther 2019, 23, 155–171, doi:10.1007/s40291-018-0380-6.

65. Wirth, T.; Parker, N.; Ylä-Herttuala, S. History of Gene Therapy. Gene 2013, 525, 162–169, doi:10.1016/j.gene.2013.03.137.

66. Dubé, I.D.; Cournoyer, D. Gene Therapy: Here to Stay. CMAJ 1995, 152, 1605–1613.

67. Hirschhorn, R. Adenosine Deaminase Deficiency. Immunodefic Rev 1990, 2, 175–198.

68. Onodera, M.; Nelson, DavidM.; Sakiyama, Y.; Candotti, F.; Blaese, R.M. Gene Therapy for Severe Combined Immunodeficiency Caused by Adenosine Deaminase Deficiency: Improved Retroviral Vectors for Clinical Trials. Acta Haematol 1999, 101, 89–96, doi:10.1159/000040930.

69. Lu, T.X.; Rothenberg, M.E. MicroRNA. Journal of Allergy and Clinical Immunology 2018, 141, 1202–1207, doi:10.1016/j.jaci.2017.08.034.

70. Lee, Y.S.; Dutta, A. MicroRNAs in Cancer. Annual Review of Pathology: Mechanisms of Disease 2009, 4, 199–227, doi:10.1146/annurev.pathol.4.110807.092222.

71. Amado, D.A.; Davidson, B.L. Gene Therapy for ALS: A Review. Molecular Therapy 2021, 29, 3345–3358, doi:10.1016/j.ymthe.2021.04.008.

72. Romano, G.; Saviana, M.; Le, P.; Li, H.; Micalo, L.; Nigita, G.; Acunzo, M.; Nana-Sinkam, P. Non-Coding RNA Editing in Cancer Pathogenesis. Cancers (Basel) 2020, 12, 1845, doi:10.3390/cancers12071845.

73. Baysal, B.E.; Sharma, S.; Hashemikhabir, S.; Janga, S.C. RNA Editing in Pathogenesis of Cancer. Cancer Res 2017, 77, 3733–3739, doi:10.1158/0008-5472.CAN-17-0520.

74. Liu, L.; Huang, L.; He, J.; Cai, S.; Weng, Y.; Huang, S.; Ma, S. PTEN Inhibits Non small Cell Lung Cancer Cell Growth by Promoting G0/G1 Arrest and Cell Apoptosis. Oncol Lett 2018, doi:10.3892/ol.2018.9719.

75. Trojan, J.; Pan, Y.X.; Wei, M.X.; Ly, A.; Shevelev, A.; Bierwagen, M.; Ardourel, M.-Y.; Trojan, L.A.; Alvarez, A.; Andres, C.; et al. Methodology for Anti-Gene Anti-IGF-I Therapy of Malignant Tumours. Chemother Res Pract 2012, 2012, 1–12, doi:10.1155/2012/721873.

76. Hillion, J.; Roy, S.; Heydarian, M.; Cope, L.; Xian, L.; Koo, M.; Luo, L.Z.; Kellyn, K.; Ronnett, B.M.; Huso, T.; et al. The High Mobility Group A1 (HMGA1) Gene Is Highly Overexpressed in Human Uterine Serous Carcinomas

and Carcinosarcomas and Drives Matrix Metalloproteinase-2 (MMP-2) in a Subset of Tumors. Gynecol Oncol 2016, 141, 580–587, doi:10.1016/j.ygyno.2016.03.020.

77. Trojan, J.; Pan, Y.X.; Wei, M.X.; Ly, A.; Shevelev, A.; Bierwagen, M.; Ardourel, M.-Y.; Trojan, L.A.; Alvarez, A.; Andres, C.; et al. Methodology for Anti-Gene Anti-IGF-I Therapy of Malignant Tumours. Chemother Res Pract 2012, 2012, 1–12, doi:10.1155/2012/721873.

78. Falanga, A.P.; Cerullo, V.; Marzano, M.; Feola, S.; Oliviero, G.; Piccialli, G.; Borbone, N. Peptide Nucleic Acid-Functionalized Adenoviral Vectors Targeting G-Quadruplexes in the P1 Promoter of Bcl-2 Proto-Oncogene: A New Tool for Gene Modulation in Anticancer Therapy. Bioconjug Chem 2019, 30, 572–582, doi:10.1021/acs.bioconjchem.8b00674.

79. Lundin, K.E.; Højland, T.; Hansen, B.R.; Persson, R.; Bramsen, J.B.; Kjems, J.; Koch, T.; Wengel, J.; Smith, C.I.E. Biological Activity and Biotechnological Aspects of Locked Nucleic Acids. In; 2013; pp. 47–107.

80. Fujimoto, K.; Sasago, S.; Mihara, J.; Nakamura, S. DNA Photo-Cross-Linking Using Pyranocarbazole and Visible Light. Org Lett 2018, 20, 2802–2805, doi:10.1021/acs.orglett.8b00593.

81. Ihara, T.; Fujii, T.; Mukae, M.; Kitamura, Y.; Jyo, A. Photochemical Ligation of DNA Conjugates through Anthracene Cyclodimer Formation and Its Fidelity to the Template Sequences. J Am Chem Soc 2004, 126, 8880–8881, doi:10.1021/ja047784r.

82. Kashida, H.; Doi, T.; Sakakibara, T.; Hayashi, T.; Asanuma, H. P -Stilbazole Moieties As Artificial Base Pairs for Photo-Cross-Linking of DNA Duplex. J Am Chem Soc 2013, 135, 7960–7966, doi:10.1021/ja401835j.

Chapter 2
Identification of lencRNAs with transcriptional regulatory functions

2.1 Introduction

Long exonic noncoding RNAs (lencRNAs), which contain retained introns, have not received much attention to date. As they are splice variants of mRNAs, detection and recognition of lencRNAs are difficult; however, advancements in RNA detection technologies have revealed that lencRNAs are indeed retained in cells[83]. Consequently, we assumed that lencRNAs have specific cellular functions[84,85]. Recent studies have shown that several ncRNAs are associated with cancer[86,87]. In addition, several lncRNAs interact with transcriptional regulatory proteins to regulate gene transcription[56,57,84,88]. Therefore, we postulated that, after being spliced, lencRNAs may also affect the transcription of adjacent genes, resulting in abnormal gene expression in some genetic disorders. Here, we designed a tethering luciferase assay system to detect the effects of lencRNAs on gene transcription. This system was used to examine the abilities of five selected lencRNAs (*BAP1-206*, *PRKDC-210*, *TP53-215*, *PIK3R1-206*, and *PIK3R1-209*), all of which are expressed at high levels in cancer cell[83]. Mediator complex subunit 12 (MED12) has been reported as a potential biomarker of cancer[89] and is a constituent of the CDK8 complex, which is involved in RNA

pol II-mediated transcription. A past report demonstrated that MED12 interacts with ncRNA-A7 to upregulate gene transcription[90].

The results of the tethering luciferase assay and RNA immunoprecipitation (RIP)-qPCR assays revealed that the lencRNA PRKDC-210 interacts with MED12 to drive gene transcription, and this effect was abolished in cells lacking CDK8.

2.2 Materials and Methods

2.2.1 Construction of the MCP-GAL4DBD plasmid

2.2.1.1 MCP DNA fragment clone

MS2 coat protein (MCP) DNA fragment was cloned from the pJZC34 plasmid, components shown in table 1.

PCR components	Volume (μl)
pJZC34 plasmid (0.01 μg /μl)	1
10 × KOD neo plus Buffer	2.5
2 mM dNTPs	2.5
2.5 mM $MgSO_4$	1
10 μl primer mixer	1.5

KOD neo plus	0.5
H_2O	16
Total = 25 μl	

Table 1: PCR reaction components for cloning of MCP DNA fragment.

Perform PCR reactions using the following table 2.

Temperature (℃)	Time (sec)	cycle
94	120	1
98	10	38
55	15	
68	60	
4	∞	1

Table 2: PCR reaction conditions for MCP DNA fragment.

MCP DNA fragments were confirmed using agarose gel electrophoresis and recovered used NucleoSpin gel and PCR cleanup kit(Macherey-Nagel, Duren, Germany).

2.2.1.2 ligation of the MCP fragment and pM

Import the MCP fragment into pM (including GAL4-DBD), used EcoRI and SalI

to digest the MCP DNA fragment and pM, and the reaction components are as following table 3.

MCP fragment		pM	
Digestion Components	Volume (μl)	Digestion Components	Volume (μl)
MCP fragment	30	pM (0.5 μg)	5
10 × H buffer	3.5	10 × H buffer	3.5
EcoRI	0.5	EcoRI	0.5
SalI	0.5	SalI	0.5
H_2O	0.5	H_2O	25.5
Total = 35 μl			
37 ℃, overnight			

Table 3: Digestion reaction of MCP DNA and pM.

Digestion products were confirmed by agarose gel. Used NucleoSpin gel and PCR cleanup kit to purify the original digestion product. Diluted the product to 100 ng/μl with H_2O.

Use table 4 to perform ligation for MCP fragments and pM.

Ligation components	Volume (μl)
pM (100 ng/μl)	1
MCP fragment (100 ng/μl)	4
Ligation High buffer(Toyobo)	5
Total 10 μl	
16 ℃, 30 minutes	

Table 4: ligation reaction for pM-MCP.

Ligation products were used immediately for plasmid screening.

2.2.2 Construction of MS2-lncRNA plasmid

2.2.2.1 Construction of MS2 plasmid

The MS2 DNA fragment was cloned from the pJZC34 plasmid. The PCR reaction components are as follows in table 5.

Temperature (℃)	Time (sec)	cycle
94	120	1
98	10	35
60	15	

68	60	
4	∞	1

Table 5: PCR reaction for MS2 DNA fragment.

Following the construction method of the GAL4-MCP plasmid in 2.2.1, the MS2 fragment and pcDNA3 were digested with EcoRV and XbaI, the digested products were purified, ligated, and immediately used for transformation and screening.

2.2.2.2 clone of lncRNA fragments

Use the *human* testis RNA library to prepare a cDNA library, and the process of the RT-PCR is as following table 6.

RT-PCR components	Volume (μl)
Human testis RNA 1 μg	2
Random primer 50 pM	0.5
dNTPs 10mM	0.5
RNase free water	3.7
Incubated at 65 ℃, 5 minutes. After that keep on ice quickly.	
5 × FC buffer	2

DTT 0.1M	0.5
RNase inhibitor	0.3
SS III polymerase	0.5
Total = 10 μl Perform RT-PCR reaction	
Temperature (℃)	Time(minutes)
50	50
85	5
4	∞

Table 6: RT-PCR for *Human* testis RNA

Added 190 μl sterile water to dilute the RT-PCR product and stored at -20℃ until further use.

Used the cDNA of the *Human* testis as the template, the lncRNA was cloned by PCR used specific primers. The PCR process is as follows in Table 7 and Table 8.

PCR components	Volume (μl)
Testis cDNA	1
KOD one Mix(TAKARA)	12.5

10 μM primer mix	1.5
H_2O	10
Total = 25 μl	

Table 7: PCR reaction components for cloning of lncRNA fragments.

Temperature (°C)	Time (sec)	cycle
94	120	1
98	10	38
55	10	
68	20	
4	∞	1

Table 8: PCR reaction conditions for lncRNA fragments.

Used agarose gel electrophoresis to confirm the length of the PCR product, and if necessary, recovery form gel, otherwise directly used NucleoSpin Gel and PCR Clean-Up kit for purification.

For ligation, pMS2 was digested with EcoRV, and the digested product was purified using NucleoSpin Gel and PCR Clean-Up kit. DNA solution of pMS2 was adjusted to 50 ng/μl.

Used HiFi DNA Assembly to recombine pMS2 and lncRNA, the conditions are as follows in table 9.

Recombination components	Volume
HiFi DNA assembly master mix	10
pMS2 50 ng/μl	1
lncRNA solution 100 ng	X μl
H_2O	9-X μl
Total = 20 μl	
Incubated at 50 ℃ for 30 minutes	

Table 9: Recombination for pMS2-lncRNA.

Recombination products were used directly for plasmid transformation and screening.

2.2.3 Recombinant plasmid transformation, screening, and purification

Taken 2 μl ligation or recombination products added to 30 μl DH5 α competent cells, mixed gently and placed on ice for 5 minutes. heat shock 42 ℃ for 45 seconds, quickly removed and placed on ice for 5 minutes. All the mixed solution

was added to 900 μl soc medium, shaking culture, 180 rpm, 37 ℃, 3 hours. Taken 100 μl culture medium and evenly spread on Amp LB solid medium, inverted culture, 37 ℃, overnight. Picked several single colonies and performed colony PCR using plasmid-specific primers. Bacteria were picked from positive colonies and added to Amp LB liquid medium for shaking culture, 200 rpm, 37 ℃, overnight.

NucleoSpin Plasmid easyPure kit (Macherey-Nagel, Duren, Germany) was used for plasmid extraction from the medium, and the Nucleobond Xtra Mini kit (Macherey-Nagel, Duren, Germany) was used for plasmid extraction when the plasmid was used for mammalian cell transfection.

Digested the plasmid and used agarose gel electrophoresis to confirm the correct size of the plasmid, then use the DNA sequencing service of Eurofins Genomics (Tokyo, Japan) to check the correctness of the plasmid sequence.

2.2.4 Cell Lines

Two cell lines, HEK293T and MCF-7, were used in this study. HEK293T was used for all tethering luciferase assays for detecting exogenous lencRNA. MCF-7 cells were used to detect endogenous lencRNA. It was used for the knockdown assay,

RNA immunoprecipitation assay, RNA-Seq and chromatin isolation by RNA purification assay, respectively.

2.2.5 Cell Culture

HEK293T cells and MCF-7 cells were cultured in high-glucose DMEM (Nacalai Tesque, Kyoto, Japan) supplemented with 10% FBS (BioWest, Funakoshi, Japan). The cells were maintained in a 5% CO_2 incubator at 37 ℃[91].

2.2.6 Transfection

Twenty-four hours before transfection, cells were seeded into 12-well cell culture plates (TrueLine, New York, NY, USA) at a density of 2–3 × 10^5 cells/well. The culture medium was replaced with 1 mL of the fresh medium before transfection. For the tethering luciferase assay in HEK293T cells, a mixture containing 400 ng of the Firefly Luc plasmid, 10 ng of the Renilla Luc plasmid, 50 ng of the MCP-GAL4 plasmid, 50 ng of the MS2-lencRNA plasmid, 2.5 μL of polyethylenimine (1.0 μg/μL; Polysciences, Warrington, PA, USA), and 100 μL of OPTI-MEM (Thermo Fisher Scientific, Waltham, MA, USA) was incubated at room temperature for 20 min. Subsequently, the transfection master mix was added to the cell culture medium, and cells were incubated for 24 hours[92,93].

For siRNA transfections, HEK293T cells were seeded into 12-well plates at a density of 1 × 10^5 cells/well. Following the manufacturer's protocol, 2 μL of Lipofectamine™ RNAiMAX (Thermo Fisher Scientific) and 24 pmol of siRNA were added to 100 μL of OPTI-MEM, and the mixture was incubated at room temperature for 20 min. The transfection master mix was then added to the cell culture medium, and the tethering luciferase assay was performed 24 hours later[94]. The Mission™ CDK8 siRNA (SASI_WI_00000016) and control siRNA (Universal Negative Control #1) were obtained from Sigma Aldrich (Burlington, MA, USA).

2.2.7 Dual-Luciferase Reporter Assay

The Dual-Luciferase® Reporter Assay System (Promega, Madison, WI, USA) was used to detect luciferase activity in the tethering luciferase assay. After transfection for 24 h, the culture medium was removed, and the cells were washed once with PBS(-) (Nacalai Tesque). Subsequently, 250 μL of 1× passive lysis buffer was added, and the cells were shaken for 15 min at room temperature. To detect the firefly luciferase signal, a 10 μL aliquot of the lysed cells was added to a luminometer tube along with Luciferase Assay Reagent II. Subsequently, Stop &

Glo® Reagent was added to the tube to enable the detection of the Renilla luciferase signal. The ratio of firefly to Renilla luciferase was then determined [95]. The signals were detected using a Gene Light 55 luminometer (WAKENYAKU, Kyoto, Japan) with the following detection conditions: delay time 3 s, count time 6 s, repeat time 1.

2.2.7 RNA Extraction

RNA extraction was performed using TRIzol® reagent (Thermo Fisher Scientific), according to the manufacturer's instructions. Briefly, the culture medium was removed, and the cells were washed once with PBS(-), followed by the addition of 400 μL of TRIzol® reagent and RNA extraction using the basal protocol [96].

2.2.8 RT-qPCR

Total RNA (500 ng) was reverse transcribed into cDNA using ReverTra Ace® reagent (TOYOBO, Osaka, Japan) and random primers (5′ -NNNNNNNNN-3′). In the RIP assays, the RNA pellet was dissolved in nuclease-free water, and the entire sample was used for reverse transcription.

The qPCR assays were performed using TB Green® Premix Ex Taq™ II (Tli RNaseH Plus) (TaKaRa, Kusatsu, Japan), according to the manufacturer's

protocol, and primers were obtained from Eurofins Genomics. Analyses were performed using an Mx3000p™ qPCR system (Agilent, Santa Clara, CA, USA)[97].

2.2.9 RNA Immunoprecipitation

Cells were grown to 80–90% confluency in 10 cm culture dishes. The medium was then changed, and the cells were incubated for 30 min at 37 °C. Subsequently, the medium was removed, and the cells were washed three times with PBS(-); then, 5 mL of 1 × nuclear isolation buffer (3 mL of Milli-Q water, 1 mL of PBS, 1 mL of 5 × NIB (1.28 M sucrose, 40 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 4% Triton X-100)) was added. The cells were then scraped from the dish, transferred to a 15 mL centrifuge tube, placed on ice for 20 min, and centrifuged for 15 min at 2000 RCF. Next, the supernatant was removed, and the pellet was resuspended in 500 µL of RNA isolation buffer (150 mM KCl, 25 mM Tris-HCl pH 7.5, 5 mM EDTA, 0.5% NP-40, 0.5 mM DTT, protease inhibitor cocktail, RNase inhibitor, DNase I) using a 25 G syringe tip. After centrifuging at 12,500 RCF for 10 min, the supernatant was collected, and a 50 µL sample (used as input) was stored at 4 °C. Another 400 µL sample was mixed with 2 µg of the specific antibody (Abcam,

Cambridge, UK) and 10 µL of Dynabeads™ Protein G (Thermo Fisher Scientific), and then incubated overnight at 4° C with gentle rotation. The next morning, the supernatant was separated from the beads using a magnetic device, and the pelleted beads were washed three times with RNA isolation buffer and once with PBS(-)[98]. Finally, 400 µL of TRIzol® reagent was added to the input sample and beads, and RNA extraction was performed as described in Section 2.2.8.

2.2.10 Primer

List of primers used in this section.

No.	Name	Sequence 5'-3'
1	MCP clone fw	gggGGATCCATGCCCAAAAAGAAAAGAAA
2	MCP clone rv	gggGATATCATCGGCTCGTCCGGAGCCGT
3	pMCP seq 1	TCATCGGAAGAGAGTAG
4	pMCP seq 2	AGCAAGTAAAACCTCTAC
5	MS2 clone fw	gggGATATCGTTTAAGAGCTATGCTGGAA
6	MS2 clone rv	gggTCTAGAGGAAGACTCCCCAGTGACTG
7	BAP1-206 clone fw	gggAAGCTtTTGCTGAGCGTGCTCCTGAA
8	BAP1-206 clone	cccGGATCcCACCAGCGGGGACTTGTTGC

	rv	
9	PRKDC-210 clone fw	gggGAATTcCCAAAGAGTCTGAAACACTC
10	PRKDC-210 clone rv	cccGATATcACTGTTGAATTTTCTGAAGT
11	TP53-215 clone fw	gggAAGCTtGCAGAAGGCAAGCCCGGAGG
12	TP53-215 clone rv	cccGGATCcTGTTTATAGCAGCTTTATTT
13	PIK3RI-206 clone fw	gggAAGCTtAAAGAAAGCCGGACTCTTTT
14	PIK3RI-206 clone rv	cccGGATCcTCAACCACAGAACTGAAGGT
15	PIK3RI-209 clone fw	gggGAATTcTAAGAAAACAATGCCATTTT
16	PIK3RI-209 clone rv	cccGATATcTCAGATGTTCAGGGACATCA
17	pcDNA3 seq 1	AGGTCTATATAAGCAGAGCT
18	pcDNA3 seq 2	TAGAAGGCACAGTCGAGG
19	Firefly luc-adaptor rv	CTGATCTAGAGGTACCGGATCCcacggcgatctttccgccct
20	Renilla luc-adaptor rv	CTGATCTAGAGGTACCGGATCCCTGCTCGTTCTTCAGCACGC
21	Adaptor XKB rv	CTGATCTAGAGGTACCGGATCC

22	Firefly luc fw	gcgcggaggagttgtgtttg
23	Renilla luc fw	TTGTCGAGGGAGCTAAGAAG
24	hBAP1-206 fw4	CTTATTTCCTGCCTGGGCTGCCTG
25	hBAP1-206 rv4	CCTTCCAACAAGCTGTATGAGGGGCC
26	hPRKDC-210 fw3	GGCAAAGAGGTGGCAGTTAAAATG
27	hPRKDC-210 rv3	CAGGCCCAGTTTGAAAGCCATC
28	hTP53-215 fw3	GATCGTGCCACTACACTCCAGCCT
29	hTP53-215 rv3	CAGTACATTGAAGGACAGAAGCCCG
30	hPIK3R1-206 fw1	GATACAGCAGACGGGACCTT
31	hPIK3R1-206 rv1	CAGCCCAAAACCCTAATTTC
32	hPIK3R1-209 fw3	CTGAGTTTGGAACATCTTGTAGG
33	hPIK3R1-209 rv3	GGGACATCATTATGGACACAGAG
34	hPRKDC-201 fw1	CCACGACTCTGCTAAACACCTC
35	hPRkDC-201 rv1	GTTGAAACCTATGCTTGCGGGC
36	hATP5B fw	GGCAGGGTCAGTCAAGTCAT
37	hATP5B rv	CCCTTCTGCTGTGGGCTAT

Table 10: Primer used in Chapter 2.

2.3 Result

2.3.1 Tethering Luciferase Assay System

To examine whether lencRNAs can affect gene transcription, we designed a novel screening technique called the tethering luciferase assay system (Figure 4). This system is divided into three parts: a lencRNA bound to MS2 RNA (MS2-lencRNA), a fusion protein comprising the MS2 coat protein (MCP) and the GAL4 DNA binding domain (MCP-GAL4DBD), and a luciferase reporter gene containing a GAL binding site.

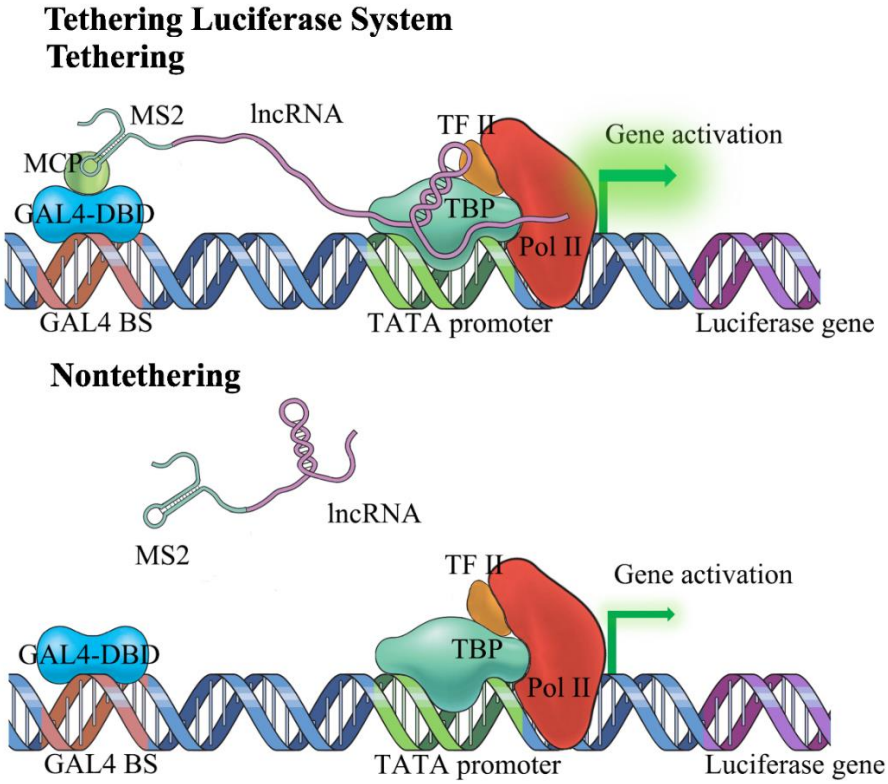

Figure 4: Schematic illustration of Tethering Luciferase Assay System.

We cloned the MCP fragment and MS2 fragment separately from the pJZC34 plasmid. We inserted the MCP fragment into the pM plasmid to prepare the pMCP plasmid(Figure 5). The pM plasmid contains the expression region of the GAL4-DNA binding domain (GAL4-DBD) protein. After inserted the MCP, the pMCP can express the GAL4DBD-MCP fusion protein. We inserted the MS2 fragment into the pcDNA3 plasmid to prepare the pcDNA3-MS2 plasmid. Meanwhile, we cloned 5 lncRNAs (*BAP1-206*, *PRKDC-210*, *TP53-215*, *PIK3R1-206*, and *PIK3R1-209*) from the *human* testis cDNA library. Next, the lncRNA fragment was inserted into pcDNA3-MS2 to obtain the pcDNA3-MS2-lncRNA plasmid(Figure 6). This plasmid is used to express MS2-lncRNA, which cannot be translated.

MCP sequence:

GTAGTATGGCTTCTAACTTTACTCAGTTCGTTCTCGTCGACAATGGCGGAACT
GGCGACGTGACTGTCGCCCCAAGCAACTTCGCTAACGGGATCGCTGAATGG
ATCAGCTCTAACTCGCGTTCACAGGCTTACAAAGTAACCTGTAGCGTTCGTC
AGAGCTCTGCGCAGAATCGCAAATACACCATAAAGTCGAGGTGCCTAAAGGC
GCCTGGCGTTCGTACTTAAATATGGAACTAACCATTCCAATTTTCGCCACGAA
TTCCGACTGCGAGCTTATTGTTAAGGCAATGCAAGGTCTCCTAAAAGATGGA
AACCCGATTCCCTCAGCAATCGCAGCAAACTCCGGCATCTACGGCTCCGGAC
GAGCC

MS2 sequence:

GGGAGCACATGAGGATCACCCATGTGCGACTCCCACAGTCACTGGGGAGTCT
TCC

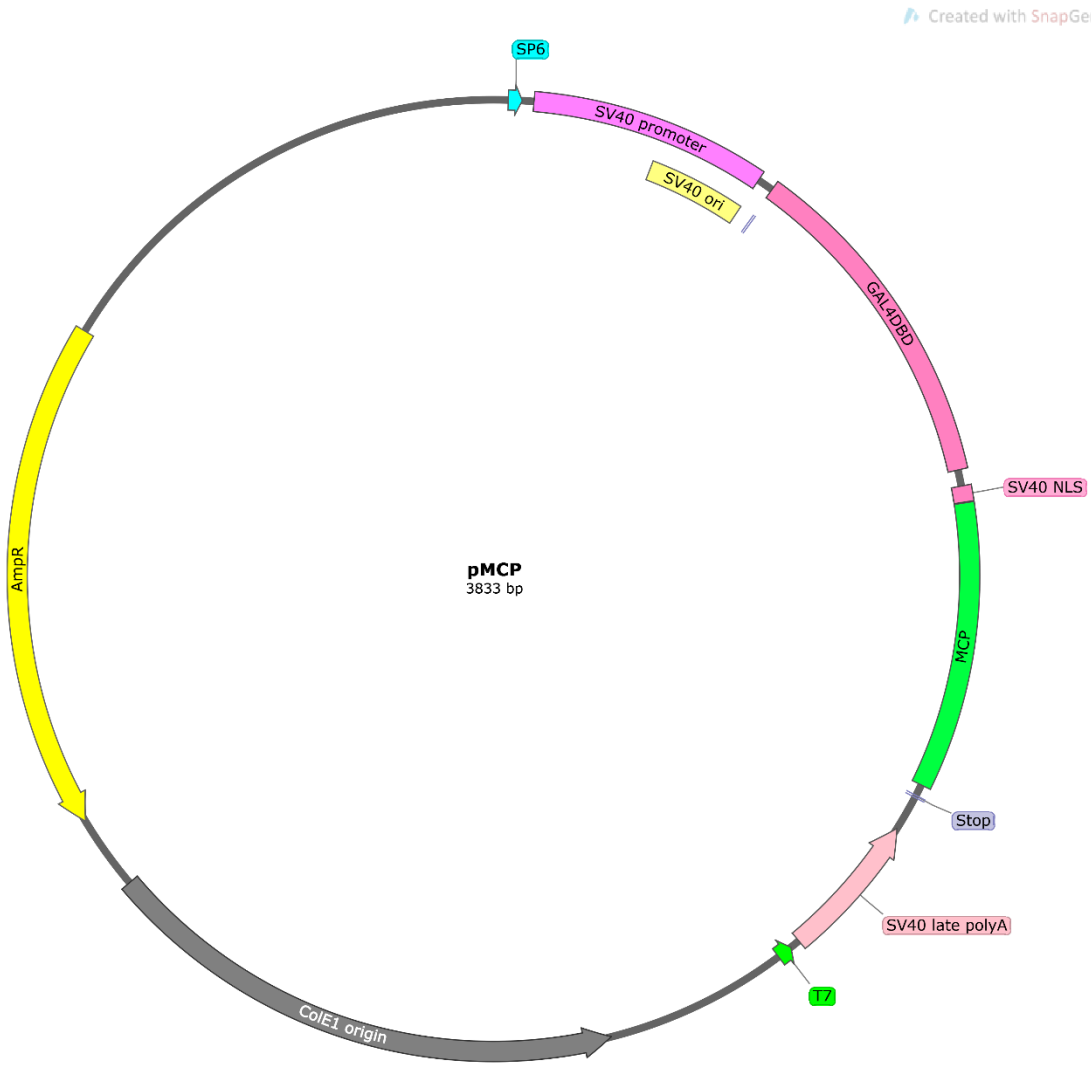

Figure 5: Gene map of pMCP.

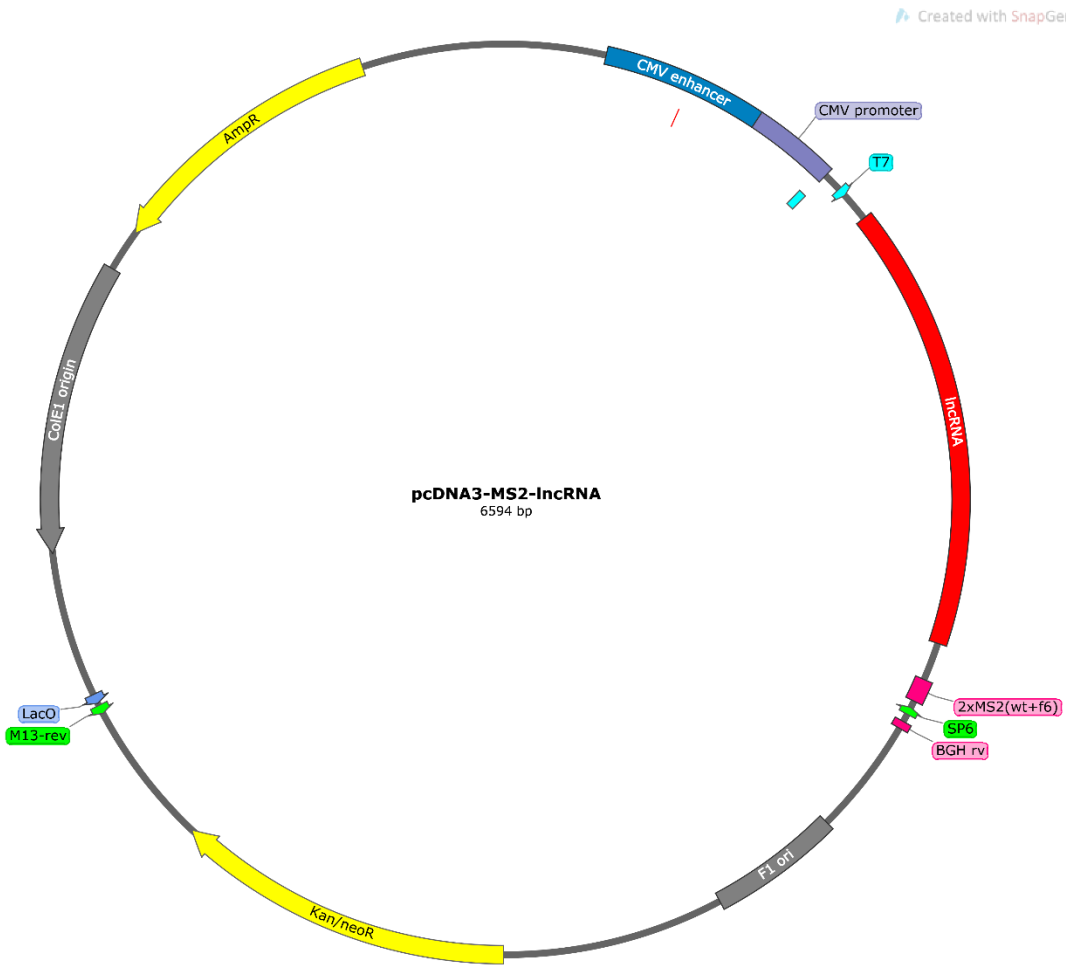

Figure 6: gene map of pcDNA3-MS2-lncRNA.

The lencRNA is tethered upstream of the luciferase gene by the interaction between MS2 and MCP-GAL4DBD, to construct an enhancer-like structure or cofactor recruiting structure. A TATA box is used as the promoter for the reporter gene because only a few factors bind to this type of element [99,100], which makes it easier to evaluate the role of the lencRNA in transcriptional regulation. The GAL4DBD protein (without MCP), which was unable to tether the lencRNA

upstream of the luciferase gene, was used as a nontethering control. Using this system, we anticipated that the tethered lencRNA would be enriched upstream of the luciferase gene, thereby enabling it to regulate the transcription of the reporter gene. By contrast, in the nontethering control, the lencRNA would be expected to float in cells and have little effect on the transcription of the luciferase gene. We used HEK293T cells for all tethering luciferase assays because HEK293T is highly transfected and easy to culture.

2.3.2 Identification of Long Exonic Non-Coding RNA with Transcriptional Regulatory Functions

We selected five genes with high intron retention in cancer and selected their lencRNA variant from the gene bank (http://asia.ensembl.org/index.html, accessed on 2 October 2019). The *BAP1-206*, *PRKDC-210*, *TP53-215*, *PIK3R1-206*, and *PIK3R1-209* lencRNAs were inserted into plasmids containing MS2 RNA sequences, and the empty MS2 plasmid was used as a negative control. Plasmids harboring the three components of the tethering luciferase assay system were then co-transfected into HEK293T cells, and relative luciferase activities were examined in cells expressing the tethered or nontethered lencRNAs. The

relative luciferase activity in cells expressing tethered *PRKDC-210* or *PIK3R1-209* was significantly higher than that in cells expressing the nontethered versions of these lencRNAs (Figure 7a). Moreover, compared with that in negative control cells, the luciferase mRNA level was higher in cells expressing *BAP1-206*, *PRKDC-210*, *PIK3R1-206*, or *PIK3R1-209* (Figure 7b). The enhancement level of mRNA detection was generally higher than that of the luciferase assay. We guessed that the difference should be caused by different detection objects and methods, as well as uncontrollable factors in the process of transcription to translation. Moreover, the higher transcription levels confirmed what we suspected, namely, that the regulatory function of these lencRNAs occurs at the transcription level.

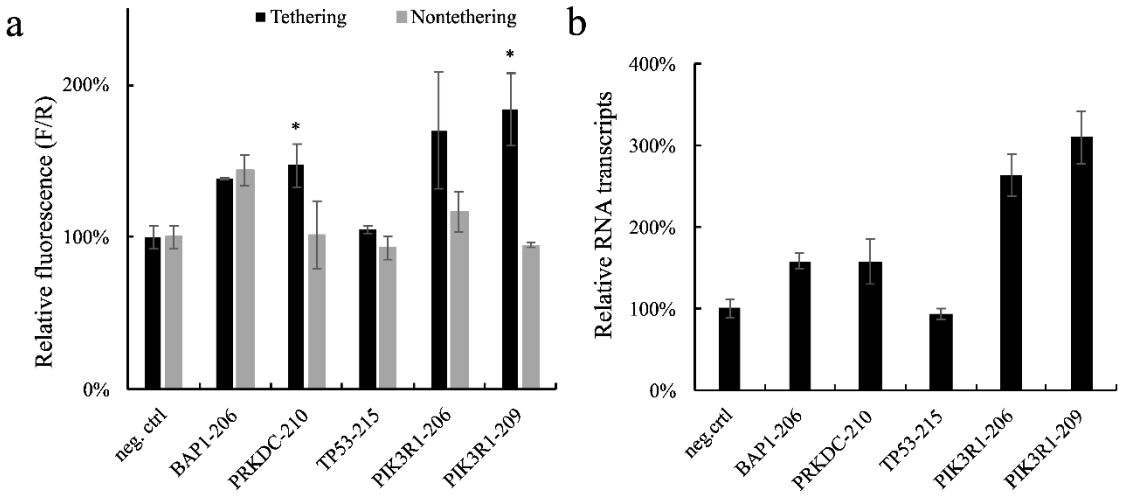

Figure 7: Tethering Luciferase Assay for lencRNA. Tethering luciferase

assay using HEK293T cells, relative fluorescence and mRNA levels were detected. We used the empty MS2 plasmid instead of the MS2 lencRNA plasmid as negative control (neg. ctrl). Value = sample/neg.ctrl. (a) We used the Dual-Luciferase® Reporter assay to detect the results of the tethering assay; the luciferase gene expression levels of different lencRNAs in the tethering or nontethering were compared. F/R, Firefly luciferase/Renilla luciferase. (b) Total RNA was extracted from cells transfected with tethering luciferase assay plasmid, and the differences in luciferase mRNA transcript between different lencRNAs were compared. Total RNA was reverse transcribed into cDNA using random primers and quantified by qPCR using a specific primer for the luciferase gene. value = sample/input. Data in all panels are represented as the mean ± s.e.m. of three independent experiments. * $p < 0.05$ by a two-tailed Student's t-test (compared with nontethering(a)).

Next, we performed a RIP-qPCR assay of *human* breast cancer MCF7 cells using an anti-MED12 antibody, because these lencRNAs are present in relatively high levels in human breast cancer cells[83]. The results indicated that PRKDC-210

can interact with MED12 or CDK8 complex (Figure 8). Because the TATA promoter is not regulated by MED12 or its complexes, we hypothesized that PRKDC-210 promotes the enrichment of MED12 or its complex, thereby enabling it to affect RNA pol II activity. Therefore, we focused our study on the lencRNA PRKDC-210.

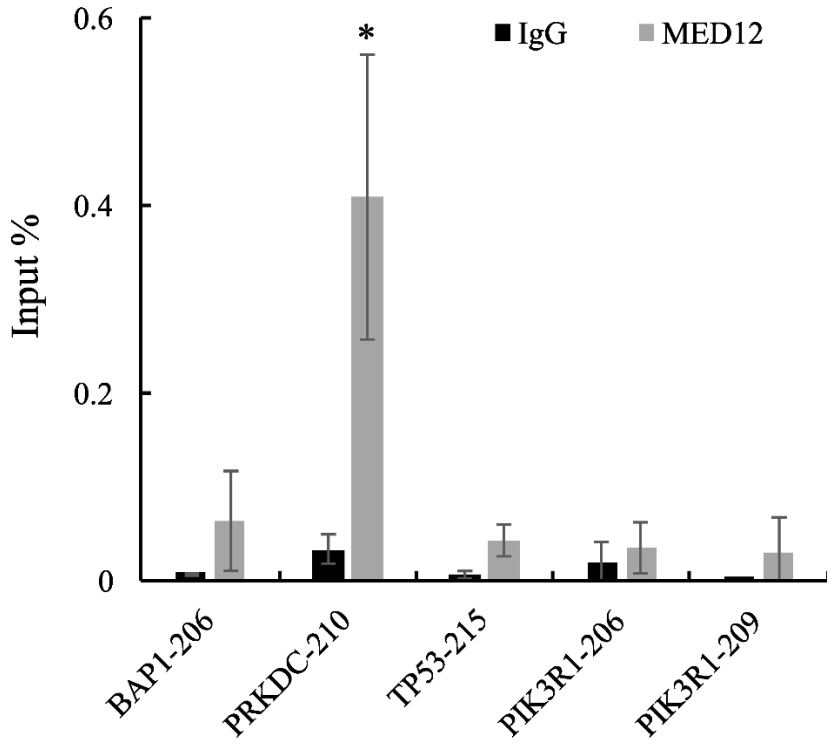

Figure 8: RIP-qPCR for med12. The binding of different lencRNAs to MED12 in MCF7 cells was compared by RIP-qPCR assay. Value = sample/input. Data in all panels are represented as the mean ± s.e.m. of three independent experiments. * $p < 0.05$ by a two-tailed Student's

t-test (compared with IgG (d)).

2.3.3 The long exonic non-coding RNA PRKDC-10 Promotes Activity of the CDK8 Complex to Drive Transcription

As a constituent of the CDK8 complex, MED12 plays a role in activating the phosphorylation of RNA pol II at Ser 5 to promote transcription[101]. Given that the RIP-qPCR analysis identified an interaction between *PRKDC-210* and MED12, we hypothesized that *PRKDC-210* may also affect the activity of the CDK8 complex. To test this hypothesis, we knocked down intracellular CDK8 expression in HEK293T using a predesigned siRNA and performed a tethering luciferase assay in CDK8-depleted HEK293T cells. The ability of *PRKDC-210* to upregulate luciferase activity was abolished in CDK8-depleted cells (Figure 9).

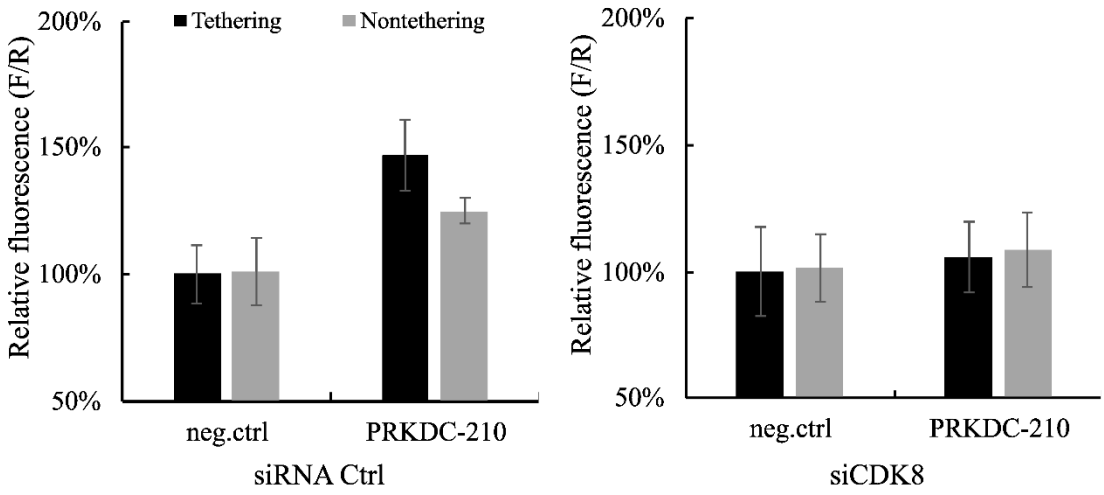

Figure 9: Tethering Luciferase Assay of PRKDC-210 in CKD8 depleted HEK239T cell. This figure shows the expression of luciferase after tethering of PRKDC-210 in control siRNA-treated HEK293T cells and siCDK8-treated HEK293T cells. F/R, firefly luciferase/Renilla luciferase. Data in all panels are represented as the mean ± s.e.m. of three independent experiments. *** $p < 0.001$ by a two-tailed Student's t-test.

Next, we performed RIP-qPCR assays of MCF7 cells using antibodies targeting total RNA pol II and RNA pol II phosphorylated at Ser 5 (S5P) or Ser 2 (S2P). After immunoprecipitation, RT-qPCR was used to analyze various regions of the gene encoding *PRKDC*, including the 5 kbp upstream flanking region (used as a negative control), the transcription start site (TSS; used as a positive control), *PRKDC-210*, and the stop site. The results indicated that *PRKDC-210* was enriched on RNA pol II S5P (Figure 10). At the initiation of transcription, the CDK8 complex activates the phosphorylation of RNA pol II at Ser 5[101]; therefore, taken together with our previous findings, these results suggest that *PRKDC-210* may act as a linker between the MED12–CDK8 complex and/or the CDK8–RNA pol II S5P complex to promote transcription.

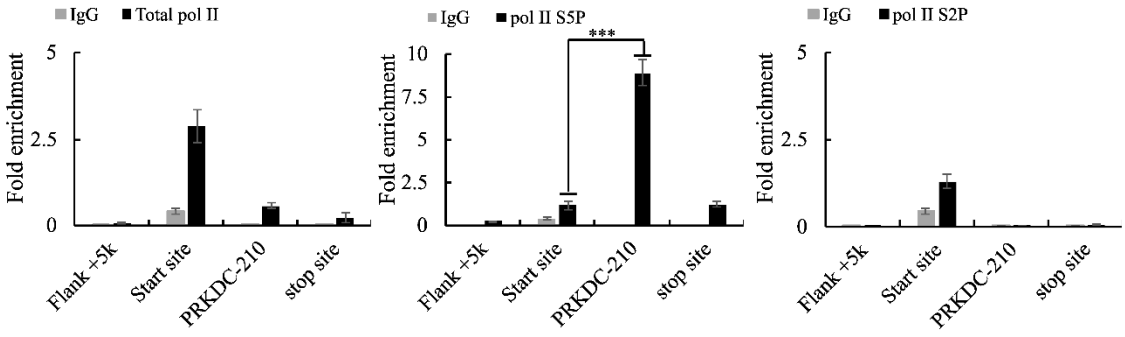

Figure 10: RIP-qPCR for RNA polymerase II. RIP-qPCR assays of MCF7 cells using antibodies targeting total RNA pol II, RNA pol II S5P, and RNA pol II S2P. The enrichment of PRKDC mRNA at different sites and the enrichment of PRKDC-210 in different phosphorylation of RNA Pol II are shown. Data in all panels are represented as the mean ± s.e.m. of three independent experiments. *** $p < 0.001$ by a two-tailed Student's t-test.

2.4 Discussion

In the past, limitations in detection technologies have meant that the roles of lencRNAs have been overlooked. In particular, as lencRNAs are splice variants of mRNAs, they are often confused with mRNAs during analyses. Here, we designed a novel tethering luciferase assay to detect the regulatory functions of exogenous lencRNAs in transcription and found that the lencRNA *PRKDC-210* interacts

with MED-12 and/or the CDK8 complex to upregulate gene expression. Since the tethering luciferase assay utilizes a TATA box, which is not associated with MED12 enrichment, we assumed that *PRKDC-210* promoted the interaction of MED12 and its complex with RNA polymerase II S5P to drive transcriptional upregulation. In addition to *PRKDC-210*, the lencRNAs *BAP1-206*, *PIK3R1-206* and *PIK3R1-209* also increased transcription of the luciferase gene in our initial screen (Figure 7,8). We did not analyze the functions of these lencRNAs further because they were unable to interact with MED12 (Figure 8); however, their potential regulatory roles and interaction with other transcriptional complexes will be examined in the future.

Based on the results of the tethering luciferase assay and RIP-qPCR analyses (Figure 7-10), we speculate that *PRKDC-210* acts directly on the CDK8 complex in the nucleus. Notably, lencRNAs such as *PRKDC-210* are not modified by poly(A) tailing; hence, they may act rapidly on nearby proteins and promote gene expression immediately after their generation. The regulatory action of lencRNAs may be a cause of abnormal gene expression in cancer. The lencRNAs examined in our current study are all expressed at high levels in cancer. As a product of

alternative splicing, the increased generation of lencRNAs is likely attributable to genetic abnormalities. Indeed, there have been numerous reports of cancers producing abnormal mRNA splicing[83,102,103]. Although these reports have focused mainly on protein-coding RNAs, there are also reports confirming that lncRNAs are spliced from coding genes[104]. More in-depth studies are required to analyze the mechanisms underlying the production and functions of lencRNAs. Such studies may expand our current understanding of diseases caused by splicing abnormalities and lay the foundation for future therapeutic developments.

2.5 Reference

83. Shiraishi, Y.; Kataoka, K.; Chiba, K.; Okada, A.; Kogure, Y.; Tanaka, H.; Ogawa, S.; Miyano, S. A Comprehensive Characterization of Cis-Acting Splicing-Associated Variants in Human Cancer. Genome Res 2018, 28, 1111–1125, doi:10.1101/gr.231951.117.

84. Wang, K.C.; Chang, H.Y. Molecular Mechanisms of Long Noncoding RNAs. Mol Cell 2011, 43, 904–914.

85. Kopp, F.; Mendell, J.T. Functional Classification and Experimental Dissection of Long Noncoding RNAs. Cell 2018, 172, 393–407.

86. Bhan, A.; Soleimani, M.; Mandal, S.S. Long Noncoding RNA and Cancer: A New Paradigm. Cancer Res 2017, 77, 3965–3981.

87. Wang, L.; Cho, K.B.; Li, Y.; Tao, G.; Xie, Z.; Guo, B. Long Noncoding RNA (LncRNA)-Mediated Competing Endogenous RNA Networks Provide Novel Potential Biomarkers and Therapeutic Targets for Colorectal Cancer. Int J Mol Sci 2019, 20.

88. Engreitz, J.M.; Haines, J.E.; Perez, E.M.; Munson, G.; Chen, J.; Kane, M.; McDonel, P.E.; Guttman, M.; Lander, E.S. Local Regulation of Gene Expression by LncRNA Promoters, Transcription and Splicing. Nature 2016, 539, 452–455, doi:10.1038/nature20149.

89. Broude, E. v; Győrffy, B.; Chumanevich, A.A.; Chen, M.; Mcdermott, M.S.J.; Shtutman, M.; Catroppo, J.F.; Roninson, I.B. Expression of CDK8 and CDK8-Interacting Genes as Potential Biomarkers in Breast Cancer HHS Public Access; 2015; Vol. 15;.

90. Lai, F.; Orom, U.A.; Cesaroni, M.; Beringer, M.; Taatjes, D.J.; Blobel, G.A.; Shiekhattar, R. Activating RNAs Associate with Mediator to Enhance Chromatin Architecture and Transcription. Nature 2013, 494, 497–501,

doi:10.1038/nature11884.

91. Freshney, R.I. Culture of Animal Cells; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2010; ISBN 9780470649367.

92. González-Domínguez, I.; Grimaldi, N.; Cervera, L.; Ventosa, N.; Gòdia, F. Impact of Physicochemical Properties of DNA/PEI Complexes on Transient Transfection of Mammalian Cells. N Biotechnol 2019, 49, 88–97, doi:10.1016/j.nbt.2018.09.005.

93. Yue, Y.; Wu, C. Progress and Perspectives in Developing Polymeric Vectors for in Vitro Gene Delivery. Biomater Sci 2013, 1, 152–170.

94. Technologies, L. Lipofectamine ® RNAiMAX Reagent Lipofectamine ® RNAiMAX Transfection Protocol Scaling Up or Down Transfections Limited Product Warranty and Disclaimer Details Package Contents;

95. Corporation, P. Dual-Luciferase ® Reporter Assay System Instructions for Use of Products E1910 and E1960;

96. Fisher Scientific, T. Procedural Guidelines;

97. Mx3000P & Mx3005P Real-Time QPCR System;

98. Gagliardi, M.; Matarazzo, M.R. RIP: RNA Immunoprecipitation. In;

2016; pp. 73–86.

99. Burley, S.K.; Roeder, R.G. TATA Box Mimicry by TFIID. Cell 1998, 94, 551–553, doi:10.1016/S0092-8674(00)81596-2.

100. Savinkova, L.K.; Ponomarenko, M.P.; Ponomarenko, P.M.; Drachkova, I.A.; Lysova, M. v.; Arshinova, T. v.; Kolchanov, N.A. TATA Box Polymorphisms in Human Gene Promoters and Associated Hereditary Pathologies. Biochemistry (Moscow) 2009, 74, 117–129, doi:10.1134/S0006297909020011.

101. Knuesel, M.T.; Meyer, K.D.; Donner, A.J.; Espinosa, J.M.; Taatjes, D.J. The Human CDK8 Subcomplex Is a Histone Kinase That Requires Med12 for Activity and Can Function Independently of the Mediator. Mol Cell Biol 2009, 29, 650–661, doi:10.1128/mcb.00993-08.

102. Yoshida, K.; Ogawa, S. Splicing Factor Mutations and Cancer. Wiley Interdiscip Rev RNA 2014, 5, 445–459, doi:10.1002/wrna.1222.

103. Srebrow, A.; Kornblihtt, A.R. The Connection between Splicing and Cancer. J Cell Sci 2006, 119, 2635–2641.

104. Huang, M.S.; Zhu, T.; Li, L.; Xie, P.; Li, X.; Zhou, H.H.; Liu, Z.Q. LncRNAs and CircRNAs from the Same Gene: Masterpieces of RNA Splicing.

Cancer Lett 2018, 415, 49–57.

Chapter 3
Functions of endogenous lencRNA PRKDC-210 in MCF7 cell line.

3.1 Introduction

The results of the tethering luciferase assay and RNA immunoprecipitation (RIP)-qPCR assays revealed that the lencRNA *PRKDC-210* interacts with MED12 to drive gene transcription, and this effect was abolished in cells lacking CDK8. Further examination of the function of the endogenous lencRNA *PRKDC-210* is very important, because the mechanism of transcription regulation in different cells may be different[105]. The effects of endogenous *PRKDC-210* on gene transcription were examined using RNA knockdown, RNA-sequencing (RNA-seq), and chromatin isolation by RNA purification (ChIRP) assays. We compared the effects of the knockdown of *PRKDC* mRNA and lencRNA *PRKDC-210* on the overall transcription of cells. According to the results of RNA sequencing, the knockdown of *PRKDC-210* will affect the transcription of a large number of genes, and this process may directly affect the transcription of target genes. We identified the start sites of *PRKDC-210* acting on target DNA through RNA-DNA interaction, and these target genes are relatively close to *PRKDC* genes, which is consistent with our original speculation on the mechanism of action of lncRNA.

3.2 Material and method

3.2.1 RNA knockdown

Antisense Oligonucleotide (ASOs) is a commonly used RNA regulatory, which can control the action of ASOs in the nucleus or the nucleus through different modifications[106]. We used Phosphorothioate-modified ASOs to bind the target RNA in the nucleus and induce the breakdown of the target RNA, thus reducing the function of the mRNA or lncRNA(figure 11).

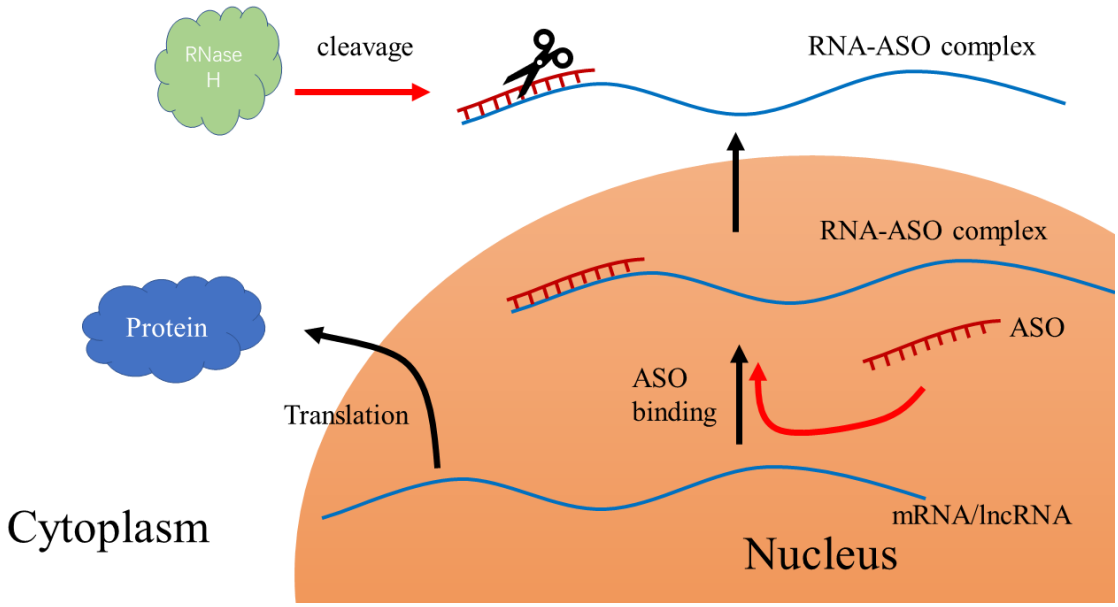

Figure 11: Mechanism of ASOs. Reference: https://www.ataxia.org/scasourceposts/snapshot-what-is-an-antisense-oligonucleotide-aso-aon/

3.2.2 Transfection for ASOs

For knockdown of PRKDC mRNA and PRKDC-210 in MCF-7 cells, following the manufacturer's protocol, 5 μL of Lipofectamine™ 2000 (Thermo Fisher Scientific) was mixed with 100 pmol of ASOs in 100 μL of OPTI-MEM and incubated at room temperature for 5 min. The transfection master mix was then added to the cell culture medium, and the cells were incubated for 24 hours [107]. Phosphorothioate-modified ASOs were obtained from Eurofins Genomics (Tokyo, Japan).

3.2.3 RNA Sequencing

Total RNA was extracted from cells transfected with ASOs as described in Section 4.4. Subsequently, the samples were analyzed using the RNA-seq service of Azenta Life Sciences (Tokyo, Japan). The level of gene expression was indicated by the read density and calculated as shown below (where FPKM represents the fragments per kilo bases per million reads) based on read counts from HT-seq (V. 0.6.1) [108].

$$\mathrm{FPKM} = \frac{\text{total exon fragments}}{\text{mapped reads(millions)} \times \text{exon length (Kbp)}}$$

The ratio of total exon fragments to mapped reads is normalized to the gene length

(exon length), such that the expression levels of genes with different sequencing depths and lengths are comparable. RNAs with reads > 10 and FPKM > 2 were screened as valid samples.

3.2.4 Chromatin Isolation by RNA Purification

The experimental method followed the protocol of Ci Chu et al.[109]. To isolate *PRKDC-210*, a probe set targeting *PRKDC-210* needed to be designed. Probes were designed on the *PRKDC-210* at about 100 bp intervals, with a total of 6 probes. The ChIRP-probe set was modified with 3' BiotinTEG and ordered from Eurofins Genomics.

Cells were grown to 80–90% confluency in 10 cm cell culture dishes. The medium was then removed, and the cells were washed three times with PBS(-) before being resuspended in PBS(-), transferred to a 50 mL tube, and centrifuged. Next, the supernatant was removed, and the pellet was resuspended in 10 mL of 1% glutaraldehyde (prepared immediately before use) and incubated for 10 min at room temperature, with gentle shaking to crosslink the cells. To stop the crosslinking, 1 mL of 1.25 M glycine was added, and the sample was incubated for 5 min at room temperature. Subsequently, the tube was centrifuged, the

supernatant was removed, the pellet was washed once with PBS(-), and the sample was then centrifuged again. Next, the pellet in PBS(-) was transferred to a 1.5 mL centrifuge tube, the PBS(-) was removed, and 500 µL of cell lysis buffer (50 mM Tris-HCl pH 7.0, 10 mM EDTA, 1% SDS, protease inhibitor cocktail, RNase inhibitor) was added. The sample was then sonicated until it became clear. A 10 µL aliquot of the lysate was reserved as the DNA/RNA input sample, and then 1 mL of hybridization buffer (750 mM NaCl, 1% SDS, 50 mM Tris-HCl pH 7.0, 1 mM EDTA, 15% formamide, protease inhibitor cocktail, RNase inhibitor) and 100 pmol of ChIRP probes (Eurofins Genomics) were added to the remaining sample, followed by incubation for 4 h at 37 ° C with rotation. Next, Dynabeads Biotin Binder beads (100 µL; Thermo Fisher Scientific) were washed with cell lysis buffer, added to the sample, and incubated for 30 min at 37 ° C with shaking. Subsequently, the beads were washed five times with 1 mL of wash buffer (2× SSC, 0.5% SDS) and then resuspended in 1 mL of wash buffer. A 100 µL aliquot was reserved for RNA isolation. The remaining 900 µL of the bead suspension was removed from the wash buffer and added to 300 µL of elution buffer to elute the DNA from the beads.

For the RNA component, the wash buffer was removed from the beads for treatment with proteinase K. After the addition of TRIzol® reagent (as described in Section 2.2.7), RT-qPCR analyses were performed using specific primers for RNA products.

For the DNA component, the sample was treated with proteinase K. Subsequently, 300 µL of PCI (phenol: chloroform: isoamyl alcohol; 25:24:1) was added, the sample was shaken vigorously for 10 min, and then the supernatant was removed following centrifugation. Next, 30 µL of NaOAc and 900 µL of 100% EtOH were added, and the sample was incubated overnight at −20 ℃. The next morning, the supernatant was removed by centrifugation, and the DNA pellet was washed with 70% EtOH, centrifuged again to remove the supernatant, air-dried, and dissolved in Milli-Q water. qPCR analyses were performed using specific primers.

3.2.5 Primer

List of primers used in this section.

No.	**Name**	**Sequence 5'-3'**
1	hPRKDC-201 (5`flank -14116) fw	cttcattaacaacccccaagac
2	hPRKDC-201 (5`flank -14116) rv	ggtgctttgatgagctttgc
3	hPRKDC-201 (5`flank -4044) fw	catgatcttggctcactgcaac

4	hPRKDC-201 (5`flank -4044) rv	caaaattagtcaggcgtggtgg
5	hPRKDC-201 (upstream -1394) fw	catcctcactccgagcttgt
6	hPRKDC-201 (upstem -1394) rv	ggttcagtggtgagtcataatg
7	hPRKDC-201 (promoter) fw	cgcttgttttatctgcctctgg
8	hPRKDC-201 (promoter) rv	ccagtaagcgcgcctctttg
9	hPRKDC-201 (start) fw	CTGCAGGAGACCTTGTCCGCT
10	hPRKDC-201 (start) rv	CTGCTGCTCAGGACGCATTCCTG
11	hPRKDC-201 (E6-E7) fw	CTGGGTGAACTTAAGACCCAGg
12	hPRKDC-201 (E6-E7) rv	GAACAGGTAGTTTGGGCTCTC
13	hPRKDC-201 (E41) fw	CTAAGTTTCACACGCCAGTCC
14	hPRKDC-201 (E41) rv	CATTGATGTGTTGAAGTCCAGG
15	hPRKDC-201 (E64-I64) fw	GCTAAGCAGTATGATGAGgt
16	hPRKDC-201 (E64-I64) rv	cattgaaccaacaaggatctgc
17	hPRKDC-201 (stop) fw	GATGAGCTACTCCTGGGTCATGA
18	hPRKDC-201 (stop) rv	CTTCTGAAAGCCCACTCTCTGG
19	hPRKDC-201 (3`UTR) fw	GGTCATAGCATTCACTTTTGC
20	hPRKDC-201 (3`UTR) rv	CTATCCCTAACAACAACCATGGG

21	CEBPD fw	GTGACAGCCTCGCTTGGAC
22	CEBPD rv	CGGTTCGTAGAAGGGCGCAG
23	FNTA fw	CACAATGGAAGTTATGGTGATGC
24	FNTA rv	GTTGCCCAAGCTGGAGTGCAG
25	UBE2V2 fw	CTGCAGGAGAAGATGGCGGTC
26	UBE2V2 rv	CTTCTCCCGCACGGTCAG
27	MCM4 fw	CTGCTGCTCAGGACGCATTC
28	MCM4 rv	GTGTGCGTTGCTCCCTGCTG
29	TCEA1 fw	GAAGCCACGCCTGCCCACTAG
30	TCEA1 rv	CAGGCAGCGACAATCGAACAC
31	TDRD7 fw	CTTGAGGTGATTCCCAAGCCGC
32	TDRD7 rv	GGACCTGCACTCACCTCGTTC
33	ARMC10 fw	GAGAAGCTGCGATAGTACCCTC
34	ARMC10 rv	GGGCGTGAAGCAACCATTCCTC
35	ETV3 fw	CTCTCTGCTGAGGATGGAAGTG
36	ETV3 rv	GTCCGTCTTCGGTCCGGTCTC
37	LYPLA1 fw	tttGACTCGCCGCTCGCAC
38	LYPLA1 rv	GCGGGGTTGACATGTTATTG
39	AASDH fw	CTCTGGGATAGCTTTACGGC
40	AASDH rv	CTCGCGGATTAGGTTTGATC
41	BAAT fw	GAAATACACCCTGGTCTCCTG
42	BAAT rv	CAGAAAACAGGGTTCAAGAGC
43	ACSM2B fw	GAGGACTGCTGTTACCCTTTG
44	ACSM2B rv	GAGCACAGTAACTCAAGCAG
45	SLC27A1 fw	CAAGGAGAGCAGAAAACAAGC
46	SLC27A1 rv	TACTGTGCGGGCTTCTGCTTTG
47	ACSM2A fw	GACTCCAAAGTTATCCACTGTC
48	ACSM2A rv	GGAAGAGAGCACAGTAACTC

Table 11: Primer used in Chapter 3.

List of ASOs used in this section.

No.	**Name**	**Sequence 5'-3'**	**Modification**
1	**Ctrl ASO**	ATTTTCTCCTCGCAATTTA	Phosphorothioate-

			modified DNA
2	**PRKDC ASOs Exon-intron**	CCATGGAAGAAGgtattgcttgg	Phosphorothioate-modified DNA
3	**PRKDC ASOs Random intron**	gcgagttttgtgtaattctgtc	Phosphorothioate-modified DNA
4	**PRKDC-210 Intron only**	CCCTAATTTATCTGACTCTAATTG	Phosphorothioate-modified DNA
5	**PRKDC-210 Exon-Intron**	CTTCAGCCTTGTCAGGTAATG	Phosphorothioate-modified DNA

Table 12: ASOs for the knockdown assay used in Chapter 3.

List of ChIRP probes used in this section.

No.	**Name**	**Sequence 5'-3'**	**Modification**
1	prPRKDC1	GAGTCTGAAACACTCTCCTG	BiotinTEG 3' DNA
2	prPRKDC2	GGCCTCTTGTTTGACCTTTC	BiotinTEG 3' DNA
3	prPRKDC3	CAAACTGGGCCTGAGCTATACC	BiotinTEG 3' DNA
4	prPRKDC4	CTTCAGCCTTGTCAGGTAATG	BiotinTEG 3' DNA
5	prPRKDC5	CTGAAATTTAAGCCCAGTGTG	BiotinTEG 3' DNA
6	prPRKDC6	GACTCCTAATACACTGCTTTGC	BiotinTEG 3' DNA

Table 13: ChIRP probes for ChIRP assay in Chapter 3.

3.3 Result

3.3.1 Functions of PRKDC-210 in MCF7 Cells

To further analyze lencRNA *PRKDC-210* cellular function, we used an antisense

oligo (ASO) to knock down endogenous *PRKDC-210* in MCF7 cells. In addition, *PRKDC* mRNA expression was also knocked down. RT-qPCR analyses performed 24 h after transfection with each ASO confirmed successful knockdown of both *PRKDC-210* and the *PRKDC* mRNA (Figure 12a). Notably, the expression level of the *PRKDC* mRNA was lower in cells expressing the *PRKDC-210* ASO than in those expressing the *PRKDC* mRNA ASO (Figure 12b).

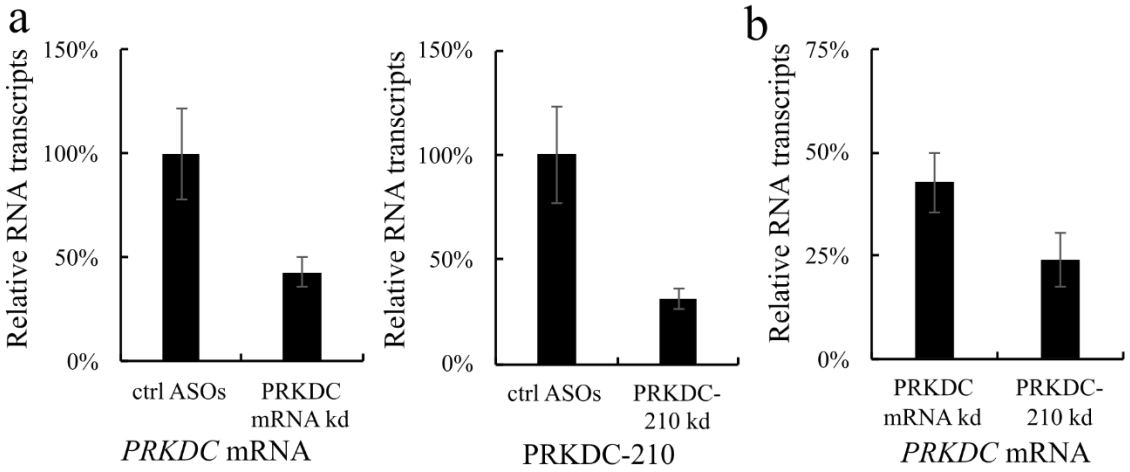

Figure 12: RNA knockdown for PRKDC. Total RNA was extracted from MCF7 cells expressing control (ctrl) ASOs or ASOs targeting the PRKDC mRNA or PRKDC-210, and the RNA transcripts levels were detected via RT-qPCR. Kd means knockdown. Data are represented as the mean ± s.e.m. of three independent experiments.

Next, we performed RNA-seq analysis of total RNA extracts from *PRKDC-210*

knockdown and *PRKDC* mRNA knockdown cells. This analysis confirmed the reduction in *PRKDC* mRNA expression following the knockdown of *PRKDC-210*. There was no significant difference between overall RNA transcription in *PRKDC-210* knockdown and *PRKDC* mRNA knockdown cells, and only a few RNAs were affected by *PRKDC-210* knockdown (Figure 12a,b). Compared with those in *PRKDC* mRNA knockdown cells, the expression levels of 614 and 511 RNAs were downregulated and upregulated, respectively, by knockdown of *PRKDC-210* (Figure 12c,d).

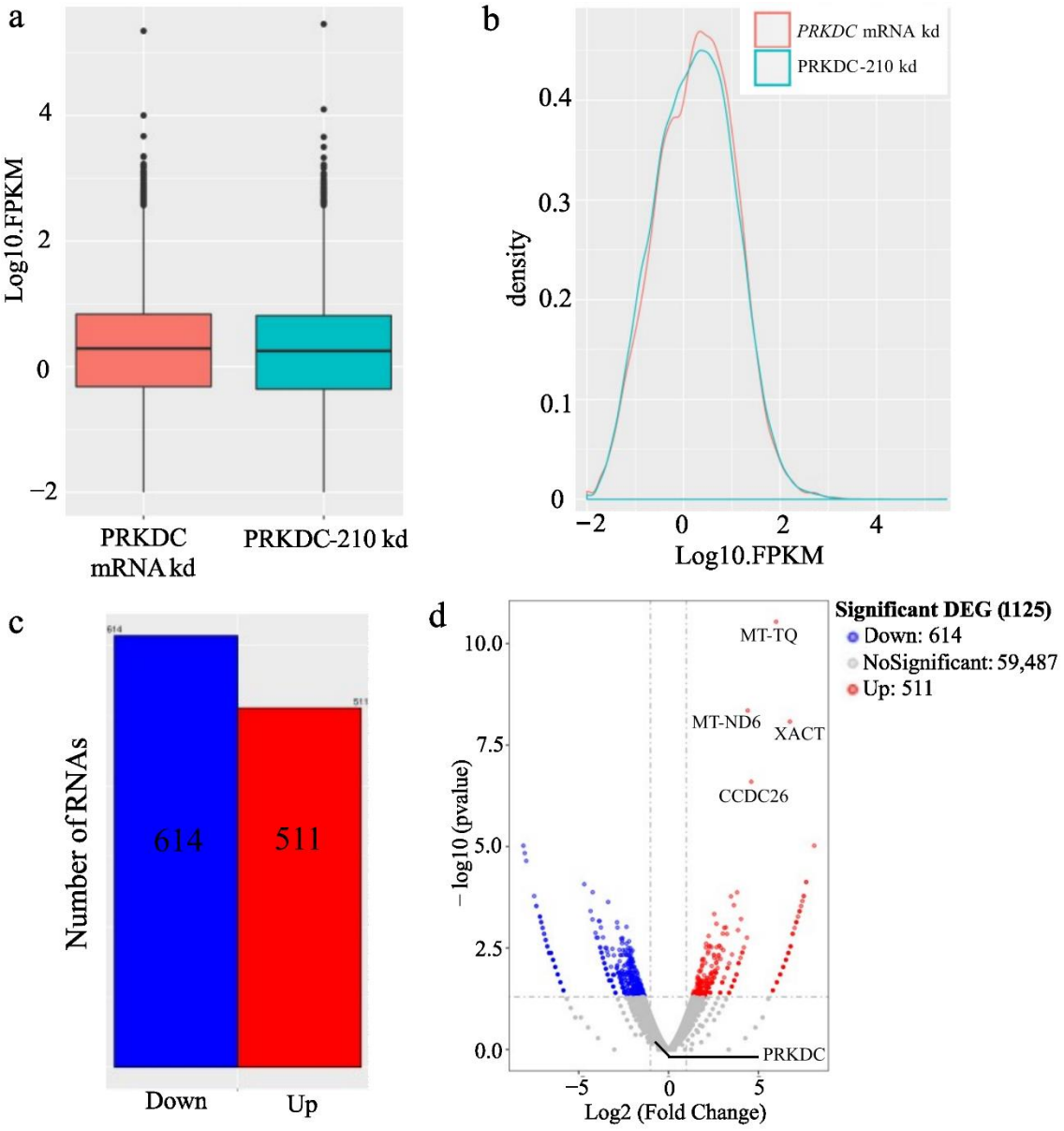

Figure 13: RNA-seq of *PRKDC* knockdown. RNA-seq analyses of PRKDC mRNA knockdown and PRKDC-210 knockdown cells. (a,b) FPKM profiles showing the overall gene expression levels. (c) The numbers of RNAs with altered transcript levels in *PRKDC-210*

knockdown cells versus control cells. (d) Genes that were significantly differentially expressed (fold change >2 and $p < 0.05$) in *PRKDC-210* knockdown cells versus PRKDC mRNA knockdown cells.

A gene ontology (GO) analysis of the affected RNAs indicated that *PRKDC-210* may affect the expression of fatty acid metabolism-related genes (Figure 14). Notably, fatty acid metabolism is an additional energy supply pathway in cancer cells[110].

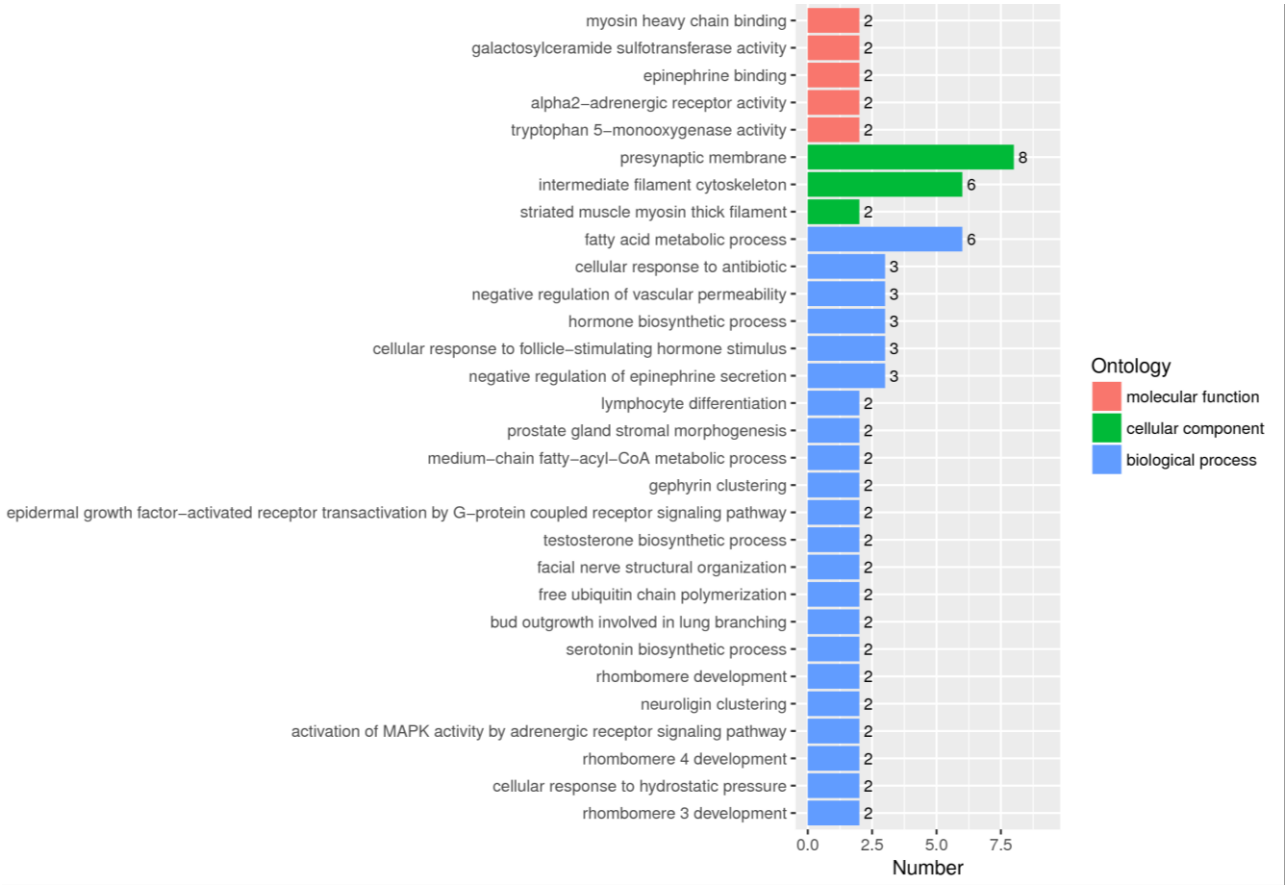

Figure 14: Gene ontology analysis of RNA-seq. Gene ontology

enrichment analysis focuses on biological processes.

3.3.2 PRKDC-210 Acts on the Endogenous Genome

Given its interaction with the CKD8 complex and RNA pol II S5P, we hypothesized that *PRKDC-210* would be enriched at the promoter or transcription initiation regions of genes. To examine this possibility, we used a ChIRP assay to assess RNA–DNA binding[109]. The RT-qPCR analysis confirmed that *PRKDC-210* was precipitated by the ChIRP probe set (Figure 15a). We suspected that *PRKDC-210* would act on adjacent genes, so we examined the enrichment of various regions of the *PRKDC* locus on the immunoprecipitated *PRKDC-210* RNA. We examined the 15 kbp upstream flanking region (used as a negative control), 1 kbp upstream site, promoter region, transcription start site (TSS), Exon region (6, 41 and 64), transcription stop site and the 3' UTR, respectively. As expected, *PRKDC-210* was enriched at the TSS of the *PRKDC* gene (Figure 15b).

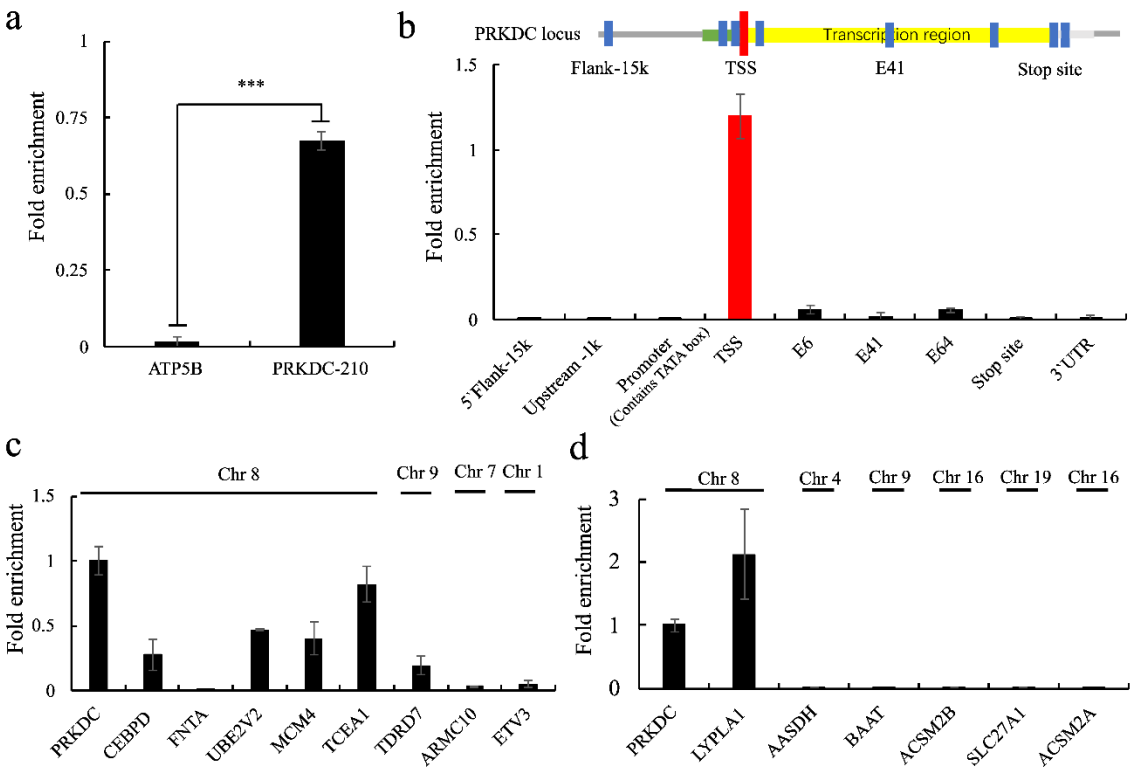

Figure 15: ChIRP assay for *PRKDC-210*. MFC7 was used for all ChIRP assays. (a) RT-qPCR analyses of *PRKDC-210* and *ATP5B* (negative control) levels in the RNA precipitated by ChIRP. (b–d) The DNAs associated with the immunoprecipitated RNAs in the ChIRP assay were analyzed by qPCR using specific primers. (b) The distribution of *PRKDC-210* at the *PRKDC* locus. (c) Enrichment of *PRKDC-210* at the TSSs of various genes. (d) Enrichment of *PRKDC-210* at the TSSs of genes related to fatty acid metabolism. Data in all panels are represented as the mean ± s.e.m. of three independent experiments. *** $p < 0.001$

by a two-tailed Student's t-test.

Next, to determine whether *PRKDC-210* can act directly on genes other than *PRKDC*, we performed ChIRP-qPCR analyses of eight randomly selected genes that were detected as differentially expressed in the RNA-seq analyses following the knockdown of *PRKDC-210*. Among them, *CEBPD*, *FANTA*, *UBE2V2*, *MCM4* and *TCEA1* are located on the same chromosome as *PRKDC* (chromosome 8), whereas *TDRD7*, *ARMC10*, and *ETV3* are located on chromosomes 9, 7 and 1, respectively. *PRKDC-210* was enriched at the TSSs of the *CEBPD*, *UBE2V2*, *MCM4*, *TCEA1* and *TDRD7* genes (Figure 15c). We also performed a ChIRP-qPCR analysis of the TSSs of genes included in the GO fatty acid metabolic process, including *LYPLA1* (chromosome 8), *AASDH* (chromosome 4), *BAAT* (chromosome 9), *ACSM2B* (chromosome 16), *SLC27A1* (chromosome 19) and *ACSM2A* (chromosome 16). *PRKDC-210* was enriched at the *LYPLA1* gene only (Figure 15d). Overall, these results suggest that genes on chromosome 8 are more likely to be targeted by *PRKDC-210*.

3.4 Discussion

An interesting phenomenon was observed in the ChIRP assay. Most of the genes

affected by *PRKDC-210* are located on chromosome 8, which is where the *PRKDC* locus is situated (Figure 14c,d). Among the 14 genes examined in the ChIRP assay (eight randomly selected genes and six fatty acid metabolic process-related genes), *CEBPD*, *FNTA*, *UBE2V2*, *MCM4*, and *LYPLA1* are located on chromosome 8, which has a total length of 146 Mbp. The maximum distance between these genes and the *PRKDC* locus is about 6.5 Mbp. Notably, the *CEBPD* and *MCM4* loci are located <50 Kbp and 7 bp from the *PRKDC* locus, respectively. The pairwise genetic distances were calculated according to Rutgers Map v.3, a combined linkage-physical map of the human genome (http://compgen.rutgers.edu/map_interpolator.shtml, accessed on 27 September 2022). The minimal genetic distances of intergenic recombinations were obtained for *PRKDC*, *CEBPD*, *MCM4*, and *UBE2V2* (<0.05 cM). The genetic distance between *FNTA* and *PRKDC* was about 0.5 cM. In general, linkage disequilibrium is supposedly noticeable at such short distances. *PRKDC* is located far from *LYPLA1*, with a genetic distance of about 5.0 cM. Unlike *CEBPD*, *MCM4*, and *UBE2V2*, *LYPLA1* does not appear to display linkage disequilibrium with *PRKDC*. Nonetheless, *LYPLA1* was also affected by *PRKDC-210*. In addition,

because chromosomes move randomly in the nucleus, the relative positions of two genes on different chromosomes are uncertain, while the relative positions of two genes on the same chromosome do not change significantly. Of course, this could be due to chromosome folding or another configuration that causes *PRKDC-210* to enhance transcription within a certain range of chromatin 8[111]. However, heterochromatin structure should not be associated with the transcriptional enhancement of *PRKDC-210*, as heterochromatin structure is thought to be unable to express genes[112]. We propose that genes on the same chromosome as the *PRKDC* locus may be more susceptible to *PRKDC-210* than those on other chromosomes, as after transcription and splicing, *PRKDC-210* may be bound rapidly by nearby CDK8 complexes or RNA pol II, which then act on neighbouring genes.

3.5 Conclusion

In conclusion, lencRNA is a type of lncRNA produced by abnormal splicing caused by genetic deviations such as cancer. Here, we found that the lencRNA *PRKDC-210* interacts with the CDK8 complex, resulting in transcriptional upregulation, which may be a major cause of abnormal gene expression in cancer. Additional

studies are needed to understand the roles of lencRNAs in disease states.

3.6 Reference

105. Panni, S.; Lovering, R.C.; Porras, P.; Orchard, S. Non-Coding RNA Regulatory Networks. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **2020**, *1863*, 194417, doi:10.1016/j.bbagrm.2019.194417.

106. Chi, X.; Gatti, P.; Papoian, T. Safety of Antisense Oligonucleotide and SiRNA-Based Therapeutics. *Drug Discov Today* **2017**, *22*, 823–833, doi:10.1016/j.drudis.2017.01.013.

107. Technologies, L. *Lipofectamine ® 2000 Reagent Protocol Outline*; 2000;

108. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and Quantifying Mammalian Transcriptomes by RNA-Seq. *Nat Methods* **2008**, *5*, 621–628, doi:10.1038/nmeth.1226.

109. Chu, C.; Quinn, J.; Chang, H.Y. Chromatin Isolation by RNA Purification (ChIRP). *J Vis Exp* **2012**, doi:10.3791/3912.

110. Currie, E.; Schulze, A.; Zechner, R.; Walther, T.C.; Farese, R. v. Cellular Fatty Acid Metabolism and Cancer. *Cell Metab* 2013, *18*, 153–161.

111. Dixon, J.R.; Gorkin, D.U.; Ren, B. Chromatin Domains: The Unit of

Chromosome Organization. *Mol Cell* 2016, *62*, 668–680.

112. Liu, J.; Ali, M.; Zhou, Q. Establishment and Evolution of Heterochromatin. *Ann N Y Acad Sci* 2020, *1476*, 59–77.

Chapter 4
Sequence-specific targeted DNA therapy using Double duplex DNA

4.1 Introduction

In the previous chapters, we introduced the mechanism of the function lencRNA act in the cell. Genetic mutations result in changes to the DNA sequence, which may cause abnormal splicing and modification after its transcription into pre-mRNA. This may result in the production of abnormal proteins or lncRNAs[113,114]. We considered that these abnormal genes may be specific to genetically abnormal diseases such as cancer[115]. And so, they may have the potential to be used as targets for gene therapy[116].

Anti gene therapy targeting specific nucleic acid sequences is gaining attention. Anti gene therapies are generally aimed at inhibiting gene expression, either by inhibiting the transcription or translation of specific DNA or RNA or by degradation of the target to block the disease-causing gene. Among other things, antibody therapies targeting dsDNA have been used for Systemic lupus erythematosus (SLE)[117]. In addition, RNA-specific therapies using siRNA, miRNA, dsRNA and antisense strands have been used for diseases such as spinal muscular atrophy (SMA) and cancer[118–120]. Unlike these traditional Anti gene therapies, we have discovered a highly manipulative, nucleic acid antibody against

dsDNA, which we have attempted to use in gene therapy.

The laboratory of Pro.f Kenzo Fujimoto reported that the Anti gene designed using 3-cyanovinylcarbazole nucleoside (^{CNV}K) was successfully photo cross linked with DNA[121]. This Anti gene is called Double duplex invasion DNA. Using a ^{CNV}K-designed gene probe, the ^{CNV}K group formed a covalent bond with the complementary chain pair to the latter thymine in the 3' direction of the base under the induction of 385 nm UV (Figure 16). Inhibition of in vitro transcription after double duplex invasion DNA forms a covalent bond with target DNA has been validated. Also, it was observed that the expression of the target gene continued to decrease several days after the introduction of Double duplex invasion DNA into cells and induction using 385 nm UV. We considered this function holds promise for gene therapy.

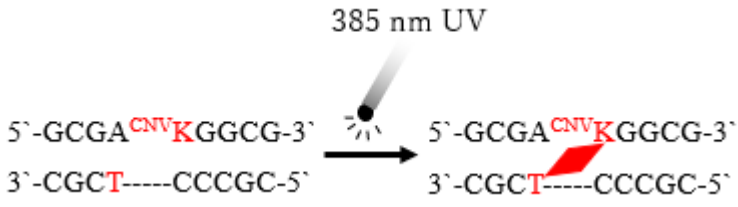

Figure 16: Schematic illustration of Double duplex invasion DNA.

In this chapter, we want to explore the possibility of Double duplex invasion DNA in mammalian cells. Different from the previous Anti gene, Double duplex

invasion DNA needs to be controlled by UV, which is very beneficial for local treatment. At present, the function and mechanism of Double duplex invasion DNA in cells are still unknown. We want to clarify these and use them for the development of gene therapy.

In this chapter, a series of intracellular functional studies were performed using Double duplex invasion DNA developed by Pro. Fujimoto's laboratory. We screened a HeLa cell line with stable expression of EGFP. The EGFP gene was introduced into the genome and expressed stably, and was used as the target gene of Double duplex invasion DNA. We performed cell fluorescence, cell proliferation, and cell apoptosis assays on EGFP-HeLa cells treated with Double duplex invasion DNA. The results of these assays were evaluated, and the potential use of Double duplex invasion DNA for gene therapy was discussed.

4.2 Material and Method

4.2.1 Cell line

(1). **HeLa cell line:** *Human* cervical cancer cells, the first cancer cells to be isolated from *homo sapiens*. HeLa cells proliferate quickly and is very adherent.

(2). **EGFP-HeLa cell line:** HeLa cell line containing the enhanced green

fluorescent protein (EGFP) gene. The EGFP gene was introduced into the genome of the HeLa cell line, and the cell lines with stable expression were screened out. We purchased EGFP-HeLa cell line (Cell Biolabs; code: AKR-213) and prepared a 100% stable expression EGFP-HeLa cell line using a single-cell screening method.

4.2.2.1 Screening of EGFP-HeLa cell line

The purchased EGFP-HeLa cell line was cultured in 200 μM G418 DMEM for one week. Wash 3 times with PBS (-), and added trypsin, incubated cells at 37 ℃ for 5 min. Collect the cells by centrifugation at 900 x g for 2 min, resuspended the cells in DMEM without addition, and counted. Used 200 μM G418 DMEM to dilute the cells to 1 cell/100 μl, and seeded 100 μl of cell dilution to each well of a 96-well plate. Incubated for 30 days at 37 ℃, 5% CO_2. Media changed every 7 days used 200 μM G418 DMEM. After 30 days, the EGFP 100% expression cell line was selected for expanded culture and cryopreserved.

4.2.2 Anti gene

The double duplex invasion DNA developed in Pro.f Fujimoto's laboratory was used, and the double duplex invasion DNA used in this chapter has the same

sequence as that in the previous in vitro study. Double duplex invasion DNA was synthesized and provided by Pro. Fujimoto Laboratory.

The double duplex invasion DNA is designed as two ssDNA anti gene probes(Figure 17).

```
5`-ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGAC
3`-TACCACTCGTTCCCGCTCCTCGACAAGTGGCCCCACCACGGGTAGGACCAGCTCGACCTG

   GGCGACGTAAACGGCCACAAGTTCAGCGT---GTCCGGCG--AGGGCGAGGGCGATGCCACCTAC
           Probe1 3`-TTCAAGTCGCA[CNV]KAGGCCGC[CN]UC-5`-NLS
                Probe2   NLS-5`-G[CN]UCCGGCGA[CNV]KGGCGAGGGCGA-3`
   CCGCTGCATTTGCCGGTGTTCAAGTCGCAC__AGGCCGCT___CCCGCTCCCGCTACGGTGGATG

   GGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACC
   CCGTTCGACTGGGACTTCAAGTAGACGTGGTGGCCGTTCGACGGGCACGGGACCGGGTGG

   CTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAG
   GAGCACTGGTGGGACTGGATGCCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTACTTC

   CAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTC
   GTCGTGCTGAAGAAGTTCAGGCGGTACGGGCTTCCGATGCAGGTCCTCGCGTGGTAGAAG

   TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTG
   AAGTTCCTGCTGCCGTTGATGTTCTGGGCGCGGCTCCACTTCAAGCTCCCGCTGTGGGAC

   GTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCAC
   CACTTGGCGTAGCTCGACTTCCCGTAGCTGAAGTTCCTCCTGCCGTTGTAGGACCCCGTG

   AAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAAC
   TTCGACCTCATGTTGATGTTGTCGGTGTTGCAGATATAGTACCGGCTGTTCGTCTTCTTG

   GGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCC
   CCGTAGTTCCACTTGAAGTTCTAGGCGGTGTTGTAGCTCCTGCCGTCGCACGTCGAGCGG

   GACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCAC
   CTGGTGATGGTCGTCTTGTGGGGGTAGCCGCTGCCGGGGCACGACGACGGGCTGTTGGTG

   TACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC
   ATGGACTCGTGGGTCAGGCGGGACTCGTTTCTGGGGTTGCTCTTCGCGCTAGTGTACCAG

   CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA-3`
   GACGACCTCAAGCACTGGCGGCGGCCCTAGTGAGAGCCGTACCTGCTCGACATGTTCATT-5`
```

Figure 17: Schematic illustration of double duplex invasion DNA probe for EGFP gene. ^{CNV}K: photoresponsive group. ^{CN}U: A modified base used to prevent self-link of probe1 and probe2. NLS: nuclear localization signal, a short peptide chain: probe-5`-CVKRKKKP.

4.2.3 Transfection for Anti gene

Twenty-four hours before transfection, cells were seeded into 48-well cell culture plates (TrueLine, New York, NY, USA) at a density of 1–4 × 10^4 cells/well. Following the manufacturer's protocol, for each well, 0.2 μl of Lipofectamine™ 3000 (Thermo Fisher Scientific) and 0.2 μl of p3000 solution was mixed with 30 pmol of Anti gene in 100 μl of OPTI-MEM and incubated at room temperature for 15 min. The transfection master mix was then added to the cell culture medium, and the cells were incubated for 24 hours[122].

4.2.4 Photo Cross Link

In the case of intracellular induction, prolonged irradiation is harmful to the cells. According to previous research reported by Fujimoto Lab, in this experiment, the cells were irradiated with 385 nm UV, 10 s, 0.149 J, and various detection experiments were performed after incubation at 5% CO_2, 37 °C for a certain

period.

4.2.5 MTT assay

The MTT assay is a colorimetric method that measures the enzymatic activity that reduces MTT and similar dyes (XTT, MTS, WST) to formazan dyes. When it reacts with dehydrogenase in the mitochondria of living cells, it produces water-soluble formazan, whose absorbance can be detected with a microplate reader. In this study, we evaluated the viability and proliferation rate of cultured cells by this method. [Live cell counting reagent SF] (Nacalai tesque, Kyoto, Japan) utilizes the MTT-like dye WST-8 and is a reagent for MTT assay[123].

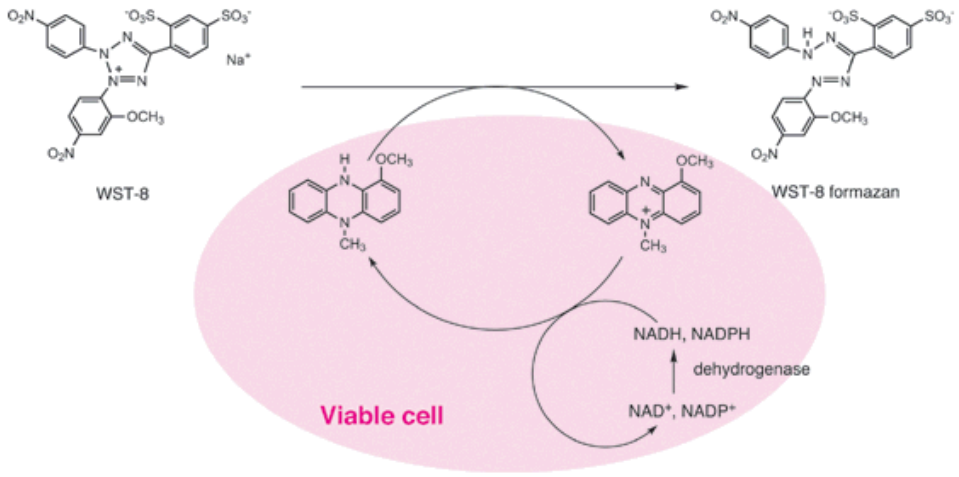

Figure 18: The mechanism of WST-8 intracellular action. From: https://www.dojindo.co.jp/products/CK04/

For each sample in 48-well plate, added 10 μl/well cell counting reagent SF,

incubated at 5% CO_2, 37 ℃ for 3 hours. Used Microplate reader to detect the absorbance of each sample at 450 nm, use the added counting reagent SF DMEM without cells was used as blank.

4.2.6 Western Blotting

Western blotting (WB) is a method in which proteins are transferred to a hydrophobic membrane after SDS-PAGE, and specific proteins are detected using antibodies. Polyacrylamide electrophoresis (SDS-PAGE) is a protein analysis method that uses differences in protein molecular weight to separate proteins.

After recovering the protein from the sample with RIPA Buffer, the protein was separated by SDS-PAGE and transferred to a membrane (PVDF membrane). The membrane was blocked with TBST blocking buffer of 5% skim milk and washed with TBST. Anti-caspase-3 and anti-β-actin antibodies were used as primary antibodies for antibody reaction, and the membrane was washed with TBST. An anti-rabbit IgG HRP antibody was used as a secondary antibody for antibody reaction, and the membrane was washed with TBST. The membrane was reacted with ECL solution and photographed with LAS3000. The resulting photographs

were subjected to grayscale analysis with the image analysis software Image J. Open the Western blotting photo with Image J and draw the background. Homogeneous protein bands are selected and measured in rectangular areas of the same area. The measurement results are expressed in IntDen (Integrated Density). The IntDen values were compared to give the relative protein content of the samples.

4.3 Result

Anti gene functions were evaluated with EGFP gene as the target of Anti gene. Since EGFP-HeLa cells stably express EGFP, intracellular fluorescence loss can be observed with a fluorescence microscope (JuLI™; Digital Bio) if the Anti gene functions. Anti gene was imported into EGFP-HeLa cells, and cell shape, presence or absence of fluorescence, cell viability, and apoptosis were detected to verify the effectiveness of the Anti gene.

4.3.1 Confirmation of transcription inhibition to target gene by Double duplex invasion DNA

To verify whether the Anti gene functions in cells, 0.03 nmol of the Anti gene was introduced into 1-4 × 10^4 cells using the "lipofectamine 3000 kit", incubated

overnight. cells were irradiated with 0.149 J, 385 nm UV for 10 s to induce photo cross link. After 24 and 48 hours of photo cross link, cell shape and fluorescence were observed with a fluorescence microscope JuLI™.

We observed a significant decrease in the total number of cells after 24 hours of UV treatment compared to before UV treatment(Figure 19). Due to the introduction of the Anti gene, many cells with loss of green fluorescence were observed. Before the UV treatment, the introduction of Anti gene has almost no effect on the cells. However, after UV treatment, the green fluorescence in the cells was significantly reduced, which may be the result of Anti gene inhibiting the expression of target genes. Also, UV treatment itself causes cell death, although we use very small doses of UV treatment. Although these results suggest that the Anti gene may inhibit the expression of target genes after photo cross linking, but the loss of green fluorescence may be due to cell death. To this end, we further detected the cell viability after UV and Anti gene treatment to confirm the function of Anti gene.

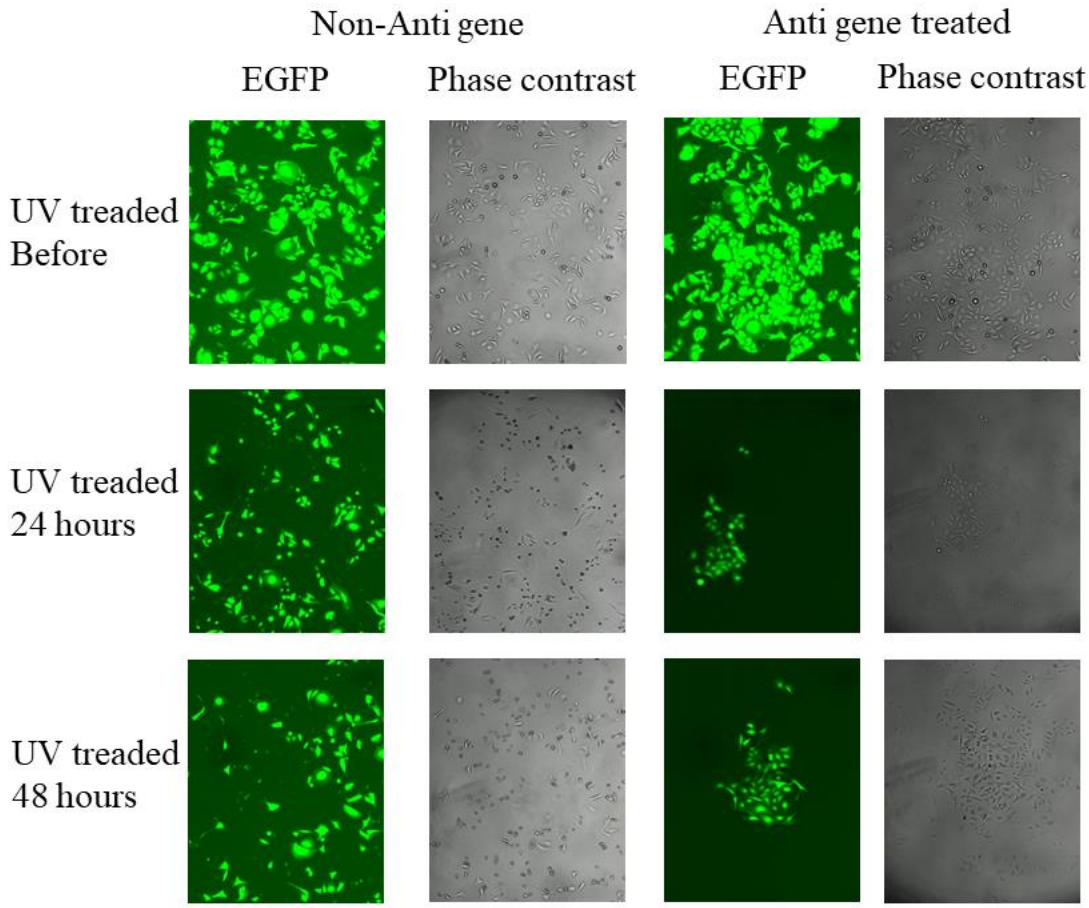

Figure 19: Fluorescence observation of Anti gene treatment. Anti gene was imported into EGFP HeLa cell line, and EGFP HeLa cells without Anti gene were used as control. The condition of the cells and the fluorescence of EGFP were compared before UV treatment, after 24 hours of UV treatment and after 48 hours of UV treatment. Magnification 40X.

4.3.2 Effect of Double duplex invasion DNA on cell viability

Because UV treatment will cause the number of cells to decrease, and Anti gene may also cause cell death. To verify whether Anti gene can induce cell death, we performed an MTT assay on the cells that introduced Anti gene and UV treatment. 0.03 nmol Anti gene was introduced into $1\sim4\times10^4$ EGFP-HeLa cells using the lipofectamine 3000 kit, cultured overnight. cells were treated with 385 nm UV, 10 s, 0.149 J for photo cross link. After 24 hours of UV irradiation, the live cell counting reagent SF was added to the cell culture plate, cultured at 37 ℃ for 3 hours, and absorbance was measured at 450 nm using a Microplate reader.

We first confirmed whether the Anti gene influences normal HeLa cells without the target gene. After UV treatment, imported of Anti gene will not affect the viability of normal HeLa cells(Figure 20).

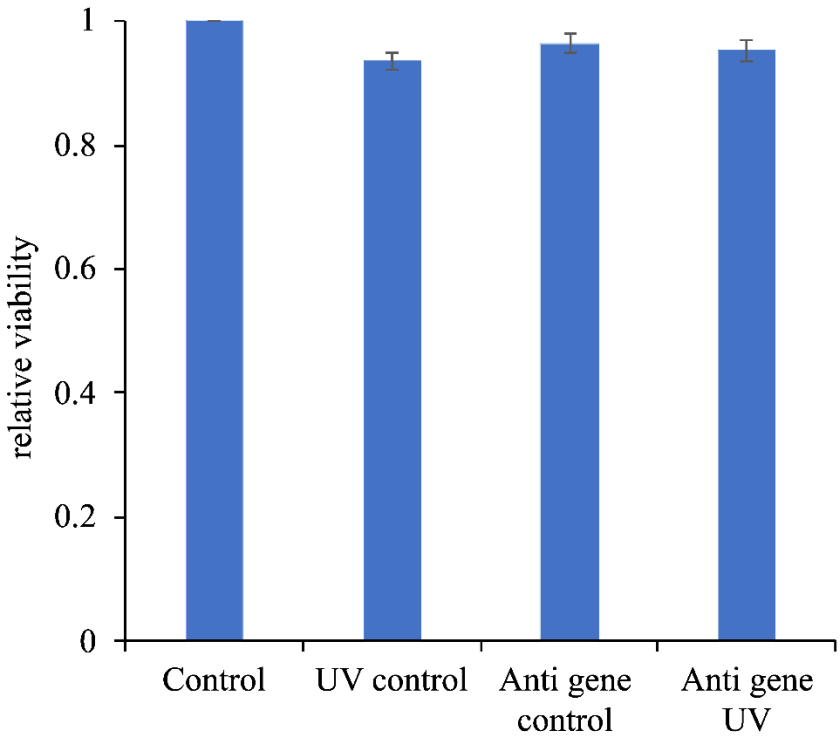

Figure 20: Relative viability of Anti gene treated on HeLa cell. The viability of HeLa cells after UV treatment for 24 hours was detected. Control: cells without any treatment. UV control: no Anti gene, only UV treatment; Anti gene control: imported Anti gene, but no UV treatment. Anti gene UV: imported of Anti gene and perform UV treatment. Data in all panels are represented as the mean ± s.e.m. of three independent experiments.

Next, we detected the effect of the Anti gene on the viability of EGFP HeLa cells containing the target gene. According to the results we obtained, the single

treatment of UV or Anti gene does not have a great impact on cell viability. However, after Anti gene and UV treatment, cell viability was greatly reduced (about 30% less viability compared to UV-treated cells) (Figure 21a). This means that Anti gene will cause cell death or reduce activity after UV induced photo cross linking. Of course, this could be due to a temporary reduction in the viability of the cells due to the complex treated. Therefore, we observed the cell viability three days after UV treatment to prove that the decrease in cell viability was caused by the Anti gene. The results proved that our considered was correct, Anti gene would lead to cell death rather than temporary activity reduction after UV induced photo cross linking(Figure 21b). This makes the possibility of Anti gene used in cancer therapy more anticipated. Because the single imported of Anti gene will not affect the viability of the cell, and the function of Anti gene requires the existence of target gene and is controlled by UV, which greatly improves the safety of treatment.

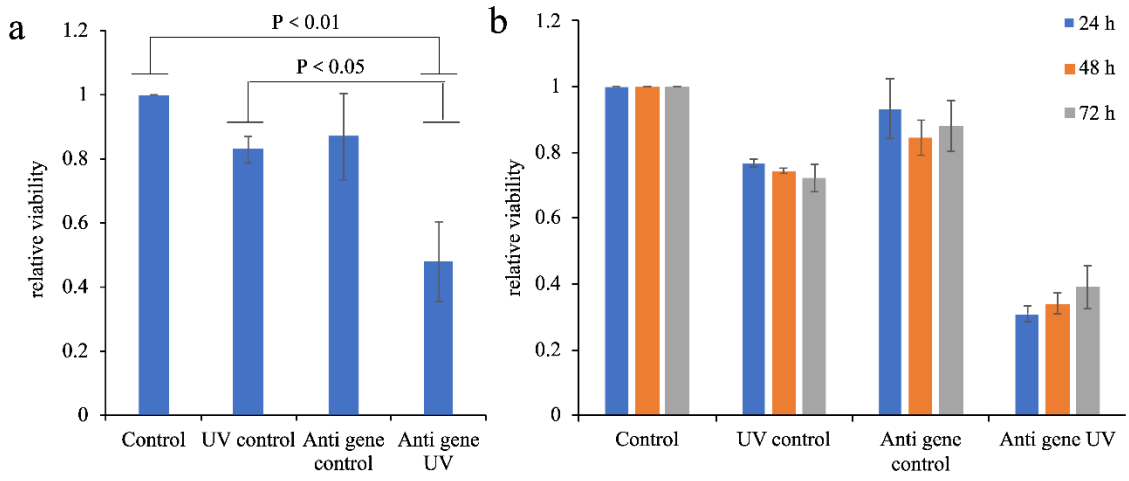

Figure 21: Relative viability of Anti gene treated on EGFP HeLa cell. The viability of EGFP HeLa cells after UV treatment was detected. Based on the detected absorbance of Control, calculate the relative absorbance of the sample. Value=$A_{sample}/A_{Control}$. Control: cells without any treatment. UV control: no Anti gene, only UV treatment; Anti gene control: imported Anti gene, but no UV treatment. Anti gene UV: imported of Anti gene and perform UV treatment. (a) Relative viability after UV treated 24 hours. (b)Relative viability after UV treated 24, 48, and 72 hours. Data in all panels are represented as the mean ± s.e.m. of three independent experiments. P value calculated by a two-tailed Student's t-test.

4.3.3 Double duplex invasion DNA induces apoptosis.

Cell death can generally be divided into two types, necrosis, and apoptosis. Necrosis is the disorderly destruction of cells caused by abnormal physical and chemical or environmental factors. Due to the incomplete degradation of cells, necrosis may lead to local inflammation. Apoptosis is different, apoptosis is the programmed death of cells. During the process of apoptosis, most of the cell components will be completely degraded, and the remaining small part will be absorbed by the surrounding cells, and this process will not cause damage to the surrounding cells. Anti gene inhibited DNA transcription or replication, and this process may trigger apoptosis mechanism leading to cell death. If the Anti gene causes apoptosis instead of necrosis, it is very beneficial for treatment. Therefore, we examined an important protein (Caspase 3) in the apoptosis process to confirm whether apoptosis occurred.

Caspase 3 will increase in the prophase and middle stage of apoptosis and be activated into cleaved caspase 3 to degrade DNA. We extracted total protein from EGFP HeLa cells 24 hours after UV treatment and performed Western blotting

detection of caspase 3. Through Western blot detection, we observed that cells treated with UV and anti gene produced a large amount of Caspase 3 and Cleaved Caspase 3(Figure 22a). We performed grayscale analysis on the results of western blot and compared the relative contents of Caspase 3 and Cleaved Caspase 3 in UV, Anti gene treatment and UV treatment only. We confirmed that the relative content of Caspase 3 increased by 50% and the relative content of Cleaved Caspase 3 increased by 20% when the Anti gene was imported compared to UV treatment only(Figure 22b). This shows that Anti gene may induce apoptosis mechanism after UV-induced photo cross link. Combined with the results of the cell viability analysis, although only UV treatment can also cause damage to cells and induce apoptosis, the effect is obviously small. After the introduction of the Anti gene, cell apoptosis is obviously enhanced, which may be very beneficial in actual treatment.

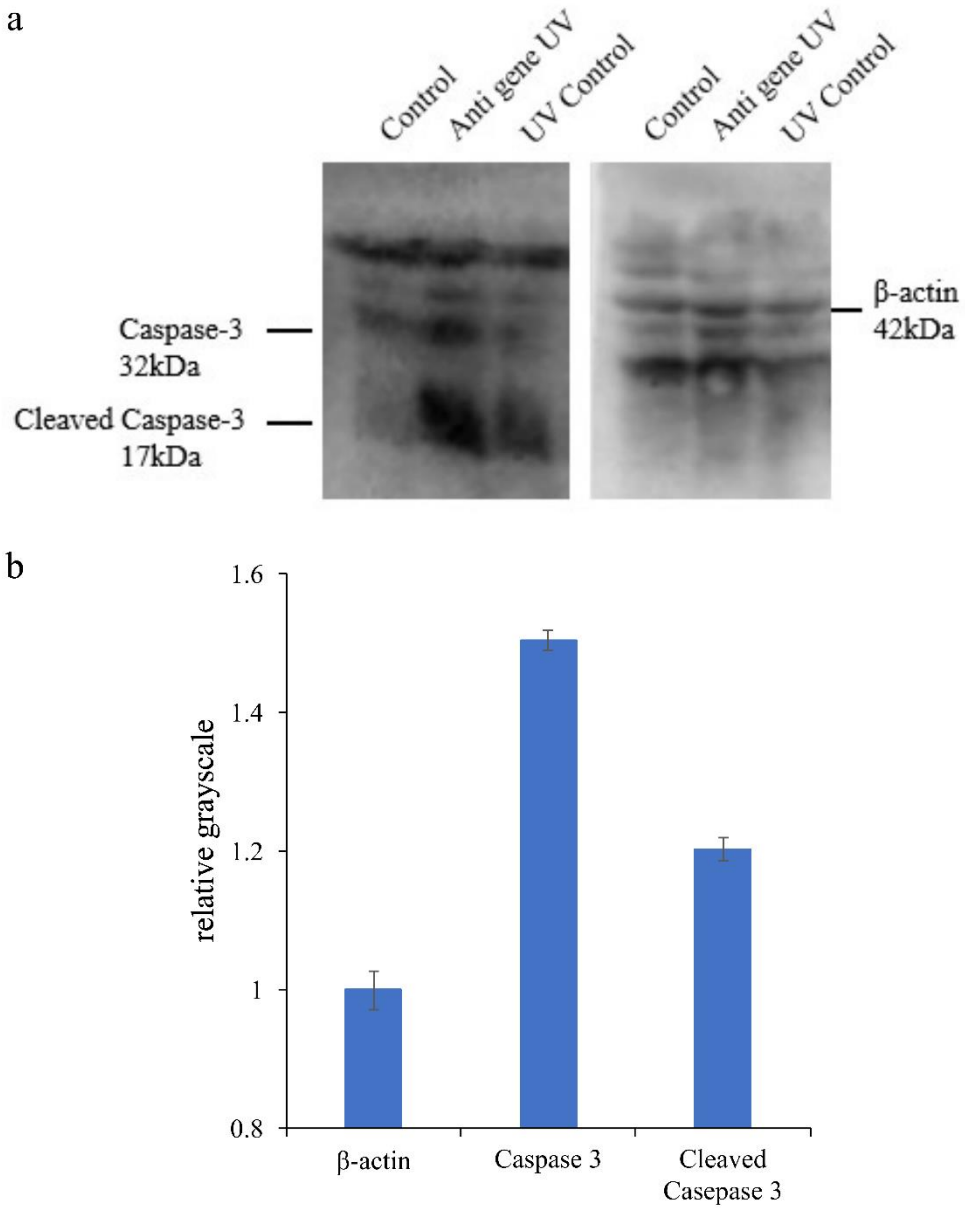

Figure 22: Western blotting for caspase 3. Anti gene transfection and UV treatment were carried out in EGFP cell line, and protein was extracted for Western blot assay after 24 hours of UV treatment. Control: cells without anti gene introduction and no irradiation, negative control. UV control: cells with only

UV irradiation without introduction of Anti gene. Anti gene UV: cells with UV irradiation after introduction of Anti gene. (a) Western blot, Caspase 3 antibody on the left, Caspase 3 (32 kDa) and Cleaved Caspase 3 (17 kDa) can be observed. On the right side, β-actin antibody was used as an internal reference. (b) grayscale analysis, Compared the gray scale of the western blot results of UV treatment only and Anti gene imported UV treatment. Relative Grayscale = IntDen(Anti gene UV)/IntDen(UV control). Data in all panels are represented as the mean ± s.e.m. of three independent experiments.

4.4 Discussion

This study sheds light on some of the effects of double duplex invasion DNA on cells. After introduction of double duplex invasion DNA into cells, if no photo cross link is induced or the target gene is not present, double duplex invasion DNA does not affect the cells. It was confirmed that photo cross link induces cell death when the target gene is present.

It was confirmed that the apoptosis characteristic proteins caspase-3 and cleaved caspase-3 increased in the samples photo cross link with double duplex invasion DNA, and apoptosis occurred.

Previous experiments strongly suggested that double duplex invasion DNA induces cell apoptosis after inducing photo cross link with 385 nm UV. We believe that the properties of double duplex invasion DNA are advantageous for cancer therapy. By controlling the activity of Anti gene with light, the effect on cells other than the affected area is small, and by targeting cancer-specific genes, it is thought that only cancer cells can be suppressed. Double duplex invasion DNA is expected as a cancer therapy.

In the future, it is necessary to study the effectiveness of double duplex invasion DNA in animal bodies in animal experiments. Unpredictable reactions may occur in the animal body. In addition, UV transmittance and introduction efficiency of Anti gene are also difficult points for practical use of double duplex invasion DNA.

4.5 Reference

113. Conboy, J.G. Unannotated Splicing Regulatory Elements in Deep Intron Space. *WIREs RNA* **2021**, *12*, doi:10.1002/wrna.1656.

114. Baralle, D. Splicing in Action: Assessing Disease Causing Sequence Changes. *J Med Genet* **2005**, *42*, 737–748, doi:10.1136/jmg.2004.029538.

115. Li, Y.; Yuan, Y. Alternative RNA Splicing and Gastric Cancer. *Mutation*

Research/Reviews in Mutation Research **2017**, *773*, 263–273, doi:10.1016/j.mrrev.2016.07.011.

116. Monteuuis, G.; Schmitz, U.; Petrova, V.; Kearney, P.S.; Rasko, J.E.J. Holding on to Junk Bonds: Intron Retention in Cancer and Therapy. *Cancer Res* **2021**, *81*, 779–789, doi:10.1158/0008-5472.CAN-20-1943.
117. Wang, X.; Xia, Y. Anti-Double Stranded DNA Antibodies: Origin, Pathogenicity, and Targeted Therapies. *Front Immunol* **2019**, *10*, doi:10.3389/fimmu.2019.01667.
118. FRIEDRICH, I.; SHIR, A.; KLEIN, S.; LEVITZKI, A. RNA Molecules as Anti-Cancer Agents. *Semin Cancer Biol* **2004**, *14*, 223–230, doi:10.1016/j.semcancer.2004.04.001.
119. Khan, P.; Siddiqui, J.A.; Lakshmanan, I.; Ganti, A.K.; Salgia, R.; Jain, M.; Batra, S.K.; Nasser, M.W. RNA-Based Therapies: A Cog in the Wheel of Lung Cancer Defense. *Mol Cancer* **2021**, *20*, 54, doi:10.1186/s12943-021-01338-2.
120. Crooke, S.T.; Witztum, J.L.; Bennett, C.F.; Baker, B.F. RNA-Targeted Therapeutics. *Cell Metab* **2018**, *27*, 714–739, doi:10.1016/j.cmet.2018.03.004.
121. Nakamura, S.; Kawabata, H.; Fujimoto, K. Double Duplex Invasion of DNA

Induced by Ultrafast Photo-Cross-Linking Using 3-Cyanovinylcarbazole for Antigene Methods. *Chemical Communications* **2017**, *53*, 7616–7619, doi:10.1039/c7cc01746d.

122. *Lipofectamine ™ 3000 Reagent Protocol Protocol Outline*; 2016;

123. Tesque, N. *Cell Cultivation Handbook*;

Chapter 5
General Discussion

In this paper, we discuss some basic functions and mechanisms of exonic lncRNAs. This could be a potential biomarker and therapeutic target for genetically abnormal diseases (e.g., cancer). At the same time, we also introduced a highly operable DNA sequence-specific Anti gene, Double duplex invasion DNA. We hope that these studies on lencRNA and gene therapy strategies can provide help for future gene regulation research and corresponding treatments.

In Chapter 2, we designed and developed a novel transcription function analysis method, Tethering Luciferase Assay system. Although we describe the method in this article for detecting the function of lncRNA at the transcriptional level, not only lncRNA but any ncRNA can be detected using this system. The Tethering Luciferase Assay system was developed because the function of lencRNA is difficult to detect or identify by other methods. LencRNA is an intron-retained ncRNA. In most detection methods, lencRNA is an easily identifiable biomarker, and it is easy to be confused with preRNA, resulting in interference or even failure of detection. Therefore, we used an exogenous gene transcription system to detect lencRNA alone. Not only that, by changing the type of the reporter gene or the sequence of the promoter and enhancer in the system, the influence of specific

transcription factors and functional ncRNAs can be studied in a targeted manner. In Chapter 2 and Chapter 3, we studied 5 lencRNAs (*BAP1-206*, *PRKDC-210*, *TP53-215*, *PIK3R1-206*, and *PIK3R1-209*) that are highly abundant in cancer. We found that most lencRNAs can enhance transcription, but their targets may be different. This paper studies the interaction between MED12 and lencRNA, in which *PRKDC-210* can interact with MED12 and its complex, and acts on RNA Pol II S5P at the initiation of transcription. Indeed, the transcriptional upregulation function of *PRKDC-210* was lost after depletion of CDK8, a core protein of the MED12 acting complex. We therefore determined that *PRKDC-210* acts on the CDK8 complex and may act as a linker. Next, we detected *PRKDC-210* in the cells, by comparing *PRKDC* mRNA knockdown and *PRKDC-210* knockdown, we confirmed that the knockdown of *PRKDC-210* would lead to a decrease in the transcription of many genes. This shows that *PRKDC-210* not only acts on the *PRKDC* gene, but may also act on other genes. To this end, we examined the interaction of *PRKDC-210* with the genome. We first checked that the site of action of *PRKDC-210* on the *PRKDC* locus is located at the transcription initiation site, which is consistent with our previous detection of the

site of action of the CDK8 complex as well as RNA Pol II S5P. Then we found that *PRKDC-210* was more likely to target genes located on chromosome 8. This confirmed our initial guess that the lencRNA would act on nearby genes.

In Chapter 4, we introduced an Anti gene, Double duplex invasion DNA. It has potential for use in gene therapy. We confirmed that Double duplex invasion DNA can be photo cross link by 385 nm UV manipulation and induce apoptosis, and it also has sequence specificity. This operability and targeting will greatly increase the safety and effectiveness if it is used in gene therapy.

In conclusion, our analysis of the mechanism of lencRNA contributes to the understanding of the basic transcriptional regulation mechanism, and the Tethering Luciferase Assay system we developed can be used for the functional detection of difficult-to-identify ncRNA. For lencRNA, it may be produced by unconventional alternative splicing caused by gene abnormalities, so it is also expected to become a biomarker or target for gene therapy. This can provide a basis for future gene therapy. We conclude by introducing a DNA-targeted gene therapy tool, Double duplex invasion DNA, which we believe is helpful for current gene therapy. It has the advantages of sequence specificity and light manipulation,

which is different from the traditional Anti gene which can only recognize special sequences. Although the actual possible application of tissues and organs is limited because of the need for UV control, it greatly improves the safety of treatment. The application of Double duplex invasion DNA remains to be studied. We consider that lencRNA may be caused by a mutation in the DNA sequence and may therefore be different from normal cellular DNA, making it a possible target for Anti gene therapy. However, so many questions remain unclear, and we need more research on the origin of lencRNA to confirm its potential. Finally, we hope this article can be helpful for future gene therapy.

List of publications

Mo, J.; Fan, G.; Tsukahara, T.; Sakari, M. The Role of the Exonic LncRNA PRKDC-210 in Transcription Regulation. Int J Mol Sci 2022, 23, 13783, doi:10.3390/ijms232213783.

Acknowledgements

Thanks to Professor Toshifumi Tsukahara for his funding and guidance on this research. At the same time, I would like to thank the members of the Tsukahara Laboratory for their help and support during the five years. Thank you for all your help in my research and life. Especially the guidance and help of postdoctoral Matomo Sakari for lncRNA.
Thanks to Professor Kenzo Fujimoto for his help and guidance on Double duplex invasion DNA research. also, I would like to thank the members of Fujimoto's laboratory for providing a lot of help in the study of Double duplex invasion DNA.
Thanks to Professor Takahiro Hohsaka for his help and guidance on my subtheme research. Also, I would like to thank Assistant Professor Watanabe for his guidance on the sub-project research.