Doctor of Philosophy of Science

Investigation of BRCA1 target gene and role of Dicer mutation in breast cancer

유방암에서 BRCA1 의 타겟 유전자 발굴과 Dicer 돌연변이의 암 조절기작 규명

The Graduate School

of the University of Ulsan

Department of Medicine

Nguyen Dinh Duc

Doctor of Philosophy of Science

Investigation of BRCA1 target gene and role of Dicer mutation in breast cancer

유방암에서 BRCA1 의 타겟 유전자 발굴과 Dicer 돌연변이의 암 조절기작 규명

Supvervisor: Chang, Su-Hwan

The Graduate School

of the University of Ulsan

Department of Medicine

Nguyen Dinh Duc

Investigation of BRCA1 target gene and role of Dicer mutation in breast cancer

This certifies that the dissertation
of Nguyen Dinh Duc is approved

Choi, Sun-Cheol
Committee Chairman Dr.

Lee, Jung-Kul
Committee member Dr.

Kang, Min-Ji
Committee member Dr.

Kim, Yongsub
Committee member Dr.

Chang, Su-Hwan
Committee member Dr.

The Graduate School
of the University of Ulsan
August 2018

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BRCA1	Breast Cancer Type 1
ChIP	Chromatin imunoprecipation
CKS1B	Cyclin-dependent kinases regulatory subunit 1
COSMIC	Catalogue of Somatic Mutations in Cancer
DISC	death-inducing signaling complex
DMEM	Dulbecco's modified Eagle's medium
dsRBD	double-strands RNA binding domain
E.coli	*Escherichia coli*
ENCODE	ENCylopedia Of DNA Elements
FADD	Fas-associated protein with death domain
FBS	Fetal bovine serum
GEO	Gene Expression Omnibus
GFP	Green fluorescent protein
GST	Glutathione S-transferase
HDAC2	Histone deacetylase 2
IPTG	Isopropyl 1-thio-β-d-galactopyranoside
ITC	Isothermal titration calorimetry
MBP	Maltose –binding protein
MEIS2	Meis homeobox 2
miR/miRNA	micro-RNA
PCR	Polymerase Chain Reaction
PDI	Protein disulfide isomerase
pre-miR	micro-RNA precousor
RISC	RNA-induced silencing complex
RNaseIIIb	RNase IIIb domain of Dicer
RNaseIIIb_dsRBD	RNase IIIb and RNA binding domain of Dicer
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel

	electrophoresis
TCGA	The Cancer Genome Atlas
TEV	Tobacco etch virus
TNBC	Triple-Negative Breast Cancer
TSS	Transcription Start Site
WT/wt	wild-type

Abstract

Investigation of BRCA1 target gene and role of Dicer mutations in breast cancer

Nguyen Dinh Duc

The Graduate School

University of Ulsan

College of Medicine

BRCA1 is a multifunctional tumor suppressor involved in DNA damage repair, cell cycle regulation, metabolism and other essential cellular processes. Many of these functions are driven by or related with its transcriptional/epigenetic regulator activity. Despite of this, there has been no comprehensive study to reveal the transcriptional/epigenetic targets of the BRCA1. Here, report a comprehensive analysis of genomics / transcriptomics data to identify genome-wide, BRCA1 target genes. We first analyzed ENCODE data where BRCA1 ChIP-seq results from four different cells lines are available. To identify a set of genes whose promoter is occupied by BRCA1, we collected 3,085 loci with BRCA1 ChIP signal found in all of the 4 cell lines and calculated distance from the loci to the nearest gene transcription start site (TSS). The result revealed 66.5% of the BRCA1 bound loci fall into 2kb around TSS, suggesting the role of BRCA1 in transcriptional regulation.

Among these loci, we selected 20 genes based on the expression correlation data from two GEO datasets and TCGA containing human breast cancer expression profiles. The selected genes have significant correlation coefficient ranging from 0.25 to 0.54 (spearman correlation) to the BRCA1 Expression. According to the known function of selected genes, we further studied three genes and verified FADD as a novel direct target of BRCA1 by ChIP, real-time PCR and luciferase reporter assay. Altogether, our data demonstrate a genome-wide transcriptional regulation by BRCA1 and suggest target genes as biomarker candidates for BRCA1 associated breast cancer.

Dicer is an indispensable part of the microRNA biogenesis machinery. Mutations in this protein have been confirmed in recent literatures about many cases of cancers such as pleuropulmonary blastoma, breast, ovarian tumors, and the others. The level of expression and activity of the mature miRNAs, which are products of the dicing process, are often found to be dysregulated in cancer. In this research, we aimed to investigate the effect of mutations on the C-terminal domain of Dicer found from clinical data, in both *in-vitro* and *in-vivo* studies. Recently, a Dicer domain was expressed in *E.coli*, then purified and functionally tested by *in-vitro* interaction and kinetic assays. The three selected mutants R1851C, A1870S and A1914V showed their impact on miRNA synthesis in breast cancer cell via luciferase reporter assay. Future study may help us to understand the role of Dicer mutations in miRNA synthesis pathway,

which might lead to the discovery of regulatory factors amenable to pharmacological intervention in cancer.

Keywords: Breast cancer, transcriptional regulation, BRCA1, ENCODE, TCGA, GEO, Dicer, miRNA, pre-miRNA, tumor suppressor, mutations.

Chapter 1. Introduction

1.1. Breast cancer

Cancer has been the leading cause of death and a major health problem worldwide and in Korea for many years[1, 2]. Among that, breast cancer is the most common and lethal in women according to recent statistics (Figure 1.1). In United States, The average woman has about 12%, a risk of developing breast cancer in the lifetime (about 1 in 8 of the female population)[1]. Breast cancer was almost found entirely in women, but men can get breast cancer, too. Breast cancer is malignant disease; it results from out-of-control cell proliferation and apoptosis in breast tissue.

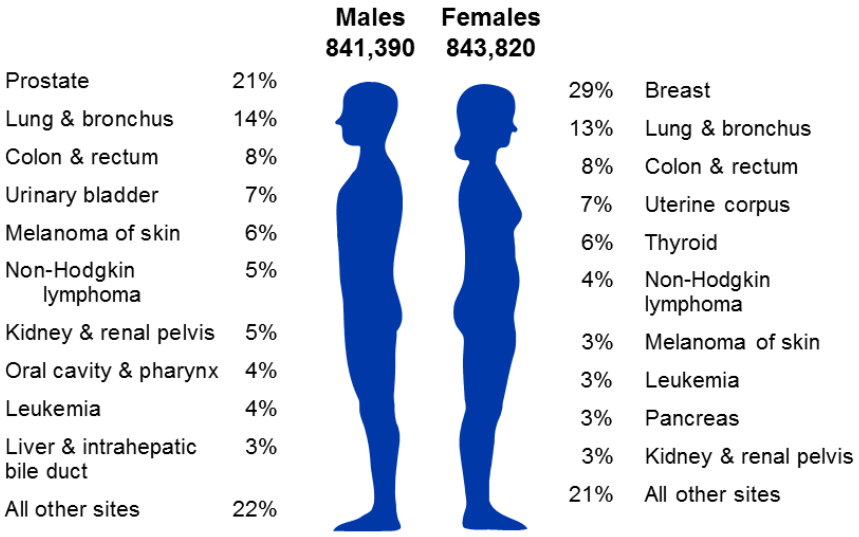

Figure 1.1. Breast cancer is a common and fatal disease in women (source from Breast cancer facts and figures 2017 – 2018[1])

Table 1.1. Cancer incidence, deaths, and prevalence by sex in Korea, 2014

Site/Type	New cases			Deaths			Prevalent cases[a)]		
	Both sexes	Men	Women	Both sexes	Men	Women	Both sexes	Men	Women
All sites	217,057	112,882	104,175	76,611	47,869	28,742	1,464,935	645,332	819,603
Lip, oral cavity, and pharynx	3,191	2,261	930	1,097	824	273	19,687	13,253	6,434
Esophagus	2,344	2,131	213	1,540	1,407	133	8,496	7,666	830
Stomach	29,854	20,087	9,767	8,917	5,767	3,150	235,172	156,264	78,908
Colon and rectum	26,978	16,182	10,796	8,338	4,760	3,578	202,295	121,057	81,238
Liver	16,178	12,058	4,120	11,566	8,616	2,950	57,691	43,192	14,499
Gallbladder[b)]	5,576	2,838	2,738	3,931	1,966	1,965	17,061	8,749	8,312
Pancreas	5,948	3,191	2,757	5,116	2,752	2,364	8,472	4,539	3,933
Larynx	1,111	1,048	63	410	371	39	9,262	8,714	548
Lung	24,027	16,750	7,277	17,440	12,785	4,655	63,460	40,098	23,362
Breast	18,381	77	18,304	2,271	17	2,254	158,916	622	158,294
Cervix uteri	3,500	-	3,500	960	-	960	45,189	-	45,189
Corpus uteri	2,214	-	2,214	264	-	264	18,381	-	18,381
Ovary	2,413	-	2,413	1,021	-	1,021	16,161	-	16,161
Prostate	9,785	9,785	-	1,667	1,667	-	62,256	62,256	-
Testis	259	259	-	14	14	-	2,570	2,570	-
Kidney	4,471	3,108	1,363	944	665	279	31,610	21,404	10,206
Bladder	3,949	3,182	767	1,354	1,016	338	28,559	23,293	5,266
Brain and CNS	1,725	917	808	1,285	716	569	9,500	4,898	4,602
Thyroid	30,806	6,174	24,632	346	84	262	328,072	54,696	273,376
Hodgkin lymphoma	278	167	111	70	38	32	2,318	1,471	847
Non-Hodgkin lymphoma	4,948	2,766	2,182	1,574	910	664	31,553	16,984	14,569
Multiple myeloma	1,396	758	638	864	473	391	4,811	2,513	2,298
Leukemia	3,080	1,771	1,309	1,671	921	750	17,151	9,534	7,617
Other and ill-defined	14,645	7,372	7,273	3,951	2,100	1,851	86,292	41,559	44,733

Source: Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2014[2]

Ordinarily, a gene responds to define a protein sequence and the protein works as an instruction manual for cell growth and function. Generally, several key proteins have been identified and found to affect signaling pathways regulating cell cycle progression, apoptosis, and gene transcription. Those genes are known as a tumor suppressor or oncogene in cancer. When changes or mistakes occur in the DNA, they may provide the wrong set of instructions, leading to faulty cell growth or function. Some genes are found to be associated with breast cancer with typographical errors; but still remain

other genetic mutations or genes are rare because these genetic mutations or protein have not been investigated.

In this study, we pay our attention to investigating the tumor suppressor BRCA1 as a transcriptional regulator; and novel mutants of Dicer relevant to microRNAs biogenesis in breast cancer cell. This might suggest a novel strategy or biomarker biomarker candidates for breast cancer therapeutics.

1.2. Breast Cancer Type 1 tumor suppressor – BRCA1

Breast Cancer Type 1 (BRCA1) was originally identified more than 20 years ago which normally expressed in breast and other tissue with multiple functions[3]. Mutations in BRCA1 protein was the first identified in familial breast cancer and thereafter, it is well established, strong breast and ovarian cancer susceptibility gene responsible for approximately half of all inherited breast cancer cases[4-7]. The BRCA1 gene located on chromosome 17q21 and its mutations are associated with a high risk of breast cancer in women with the lifetime risk of 50–85%[8].

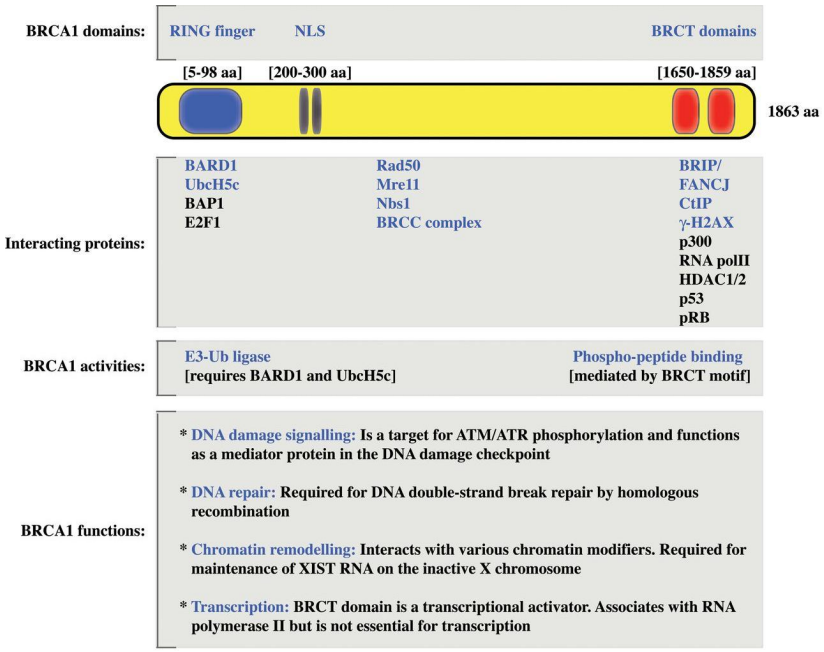

Figure 1.2. BRCA1 is a multifunctional protein Biochemical Society (source Transactions Oct 01, 2006,34(5)633-645[9])

Since the discovery of BRCA1, there have been numerous studies aimed to fully understand its biological function[10-14]. BRCA1, a large protein with several domains, known to be involved in multiple cellular processes and have multiple roles as a tumor suppressor or cell cycle regulator (Figure 1.2). These roles contribute to different vital functions such as DNA damage repair, cell cycle checkpoint control, and ubiquitination[15-18]. In addition to its mentioned roles, BRCA1 also acts on gene promoter as a transcriptional regulator since its C-terminal interaction to other transcription factors[19-25]. Thus, investigating the target genes regulated by BRCA1 would be an interesting approach for understanding the diverse function of BRCA1, especially, the transcriptional/epigenetic targets of BRCA1.

1.3. Dicer1 mutants and its roles in miRNA biogenesis

Dicer1 is a key enzyme of type III endoribonuclease, which plays a central role in non-coding RNA syntheses, such as microRNA[26-28]. The crystal structure of *Giardia* Dicer1 is got first detail looked by J. A. Doudna *et al.* since 2006[29] and other species were resolved then[30-34]. Recently, structural studies on the full-length of Dicer1 protein show that this protein owns an L shape overall[35]. The full structure of human Dicer1 has been solved and it is similar to several researches have been constructed of its domains separately (Figure 1.3). And its full architecture has been described base on those domains and the *Giardia* Dicer structures docking or Cryo-EM reconstruction[36, 37].

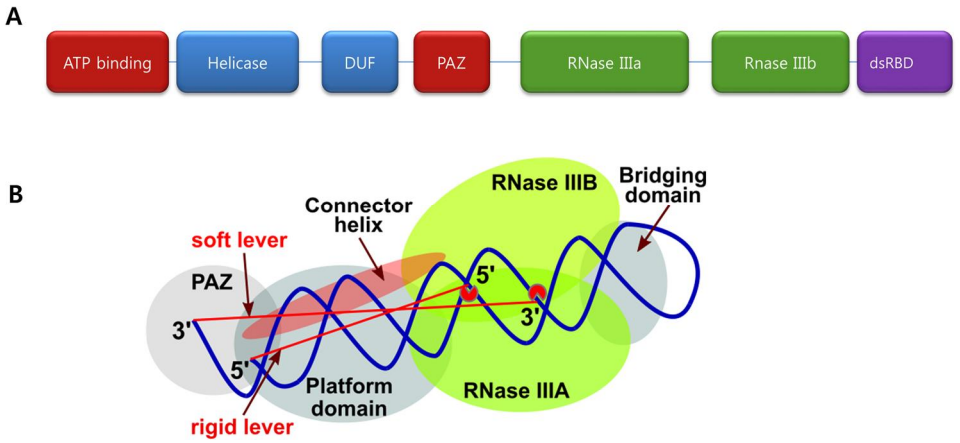

Figure 1.3. The scheme of Dicer constructs (A) and the proposed 2-lever model for Dicer protein structure and its substrate – miRNA precursor (B) (source *BMC Genomics* 2018 19(Suppl 3):114[38])

In the microRNA biogenesis pathway, Dicer carries out a central role in the process (Figure. 1.4). The miRNA precursor is transcribed from the genome in the nucleus. Subsequently, the long pri-miRNA is generated by the DGCR8–Drosha complex, to produce a 60- to 70-nucleotide precursor miRNA or pre-miRNA[39]. The pre-miRNA is exported to the cytoplasm via Exportin 5[40, 41] and further cleaved by the Dicer1 complex into the mature form of miRNA[42]. The mature miRNA is then loaded onto the Argonaute protein, forming a miRNA–protein complex known as the RNA-induced silencing complex (RISC)[43]. Afterward, it binds to mRNA and exerts its function of mRNA degradation or translational repression[44]. Therefore, knockout or mutants in Dicer1 are known to impact in cancer gene networks[45]. Deficiency of Dicer1 protein resulted in changes of the miRNA expression level, this leads the cell to an abnormal state and can cause rapid/adaptive changes in gene expression thus can be a cause of various diseases, especially in cancer[46-50].

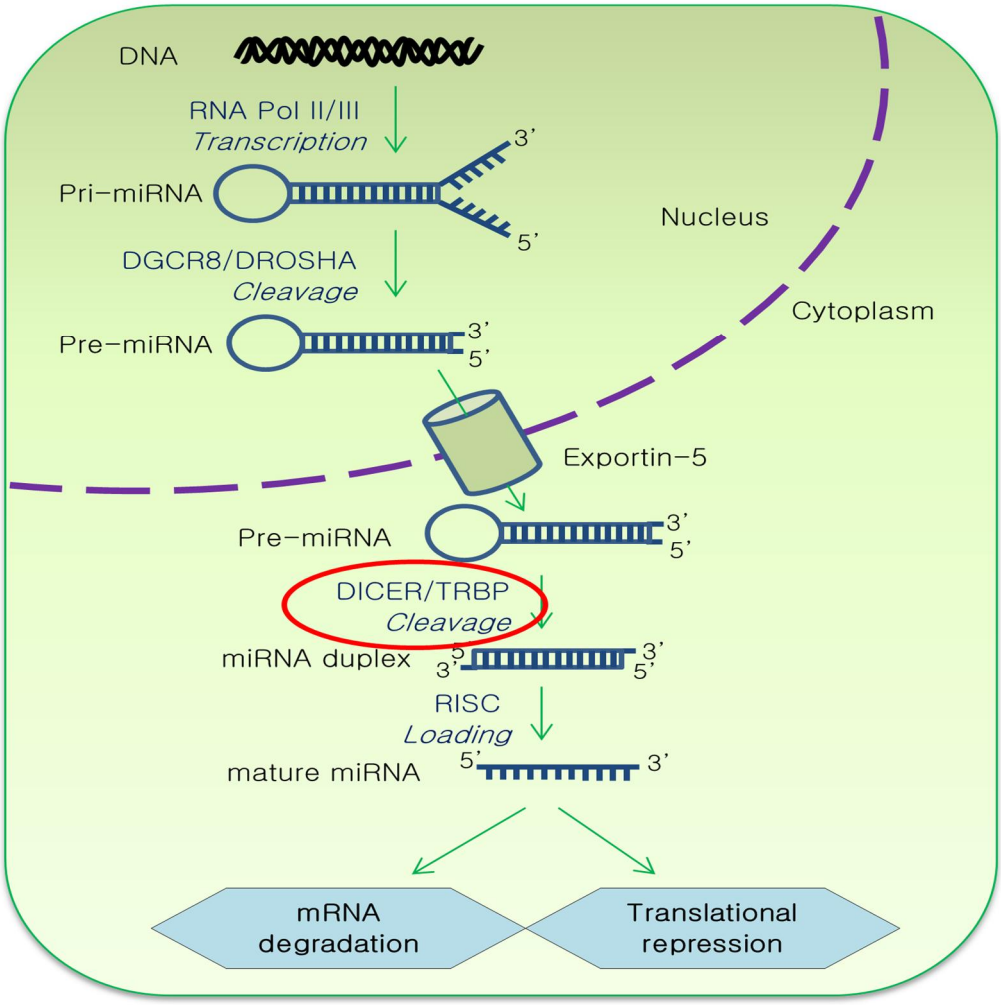

Figure 1.4. Dicer has a central role in miRNA biogenesis

Although Dicer mutants are found in the active site and cause failure in function, it remains unclear whether the Dicer protein can interaction and process pre-miRNAs loop. Here, we present a preliminary demonstrating that mutants of the C-terminal helicase domain of human Dicer indeed enhances or disrupts its ability to process endogenously pre-miRNAs into biologically active miRNAs *in-vitro* and *in-vivo*. At present, the C-terminal domains of Dicer have been expressed in bacteria and purified. Then the purified protein

has been used for ITC binding activity measurement and crystallization. The cellular based studies will be updated in future.

1.4. Bioinformatics analyses

Current bioinformatics developments have enabled the creation of a genome-wide database for multiple types of variations in cancer[51]. These kinds of data support scientist narrowing down times and targets. Single dataset mainly provides information about one type of variation, and singular analyses of this dataset have led to our understanding of a long list of single gene disorders. However, it is clear that isolated analyses of the different types of variations in the database may only provide a linear view of a multidimensional landscape because of systems disorders. Therefore, data integrated analyses is important to reach thoroughly understanding of common disorders and measure the possible interactions of risk factors or aberrant identified in genome-wide analyses[52]. Large-scale data-sharing projects such as TCGA[53] and ENCODE[54] include multiple types of data on a range of phenotypes. In the present study, we conduct those kinds of the database for 2 projects. Firstly, we make inquiries to BRCA1 target genes using ENCODE, GEO and TCGA integrated analyses. Secondly, we investigate the novel functional Dicer mutants from patient data of TCGA and COSMIC[55].

ENCODE (ENCyclopedia Of DNA Elements) is one of the data which provide massive information about elements involved in regulation of transcription including functional elements including enhancers, promoters, insulators, etc. The identification of their location in the genome and the interaction with transcriptional factors are keys for understanding location specific gene regulation. Particularly, the chromatin immunoprecipitation with massively parallel DNA sequencing (ChIP-seq) data generated in ENCODE consortium provided a substantial amount of regulatory information[54] (Figure 1.5). On the other hand, the GEO (Gene Expression Omnibus) collected huge data of gene expression and other functional genomics data sets[56]. The combinational analyses help us to narrow down BRCA1-targets and further confirmation by cellular studies.

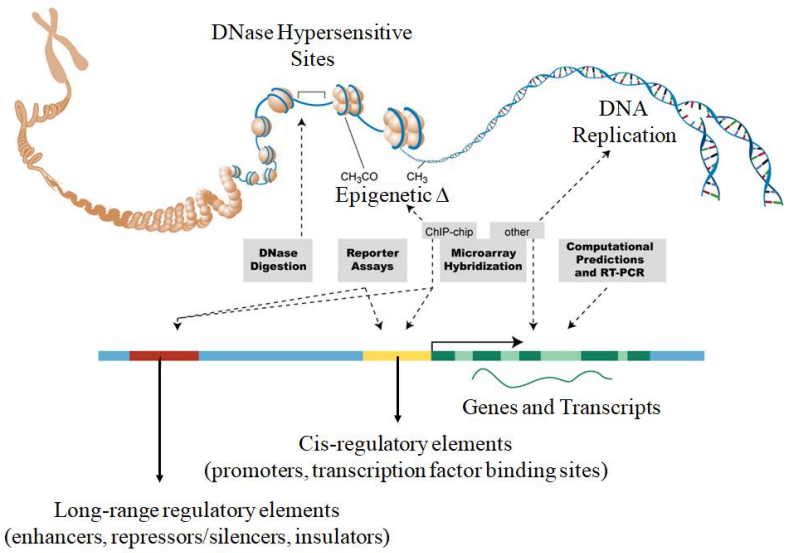

Figure 1.5. ENCODE data contains a huge number of gene regulatory elements (source Consortium, 2004[54])

Hotspot mutations of Dicer have been shown the key residues in the active site of this Ribonuclease III family protein. Beside of well-described mutations, there is much less of investigation near C-terminal dsRBD domain which believed to interact with a loop of miRNAs precursors. Using the TCGA / COSMIC clinical data and structure analysis, we conduct several mutants to understand this interaction and its impact on miRNA biogenesis in breast cancer cell. The interaction of miRNA precursor and Dicer may reveal an insight understanding on the microRNA biogenesis process, which might lead to the discovery of regulatory factors amenable to pharmacological intervention in cancer.

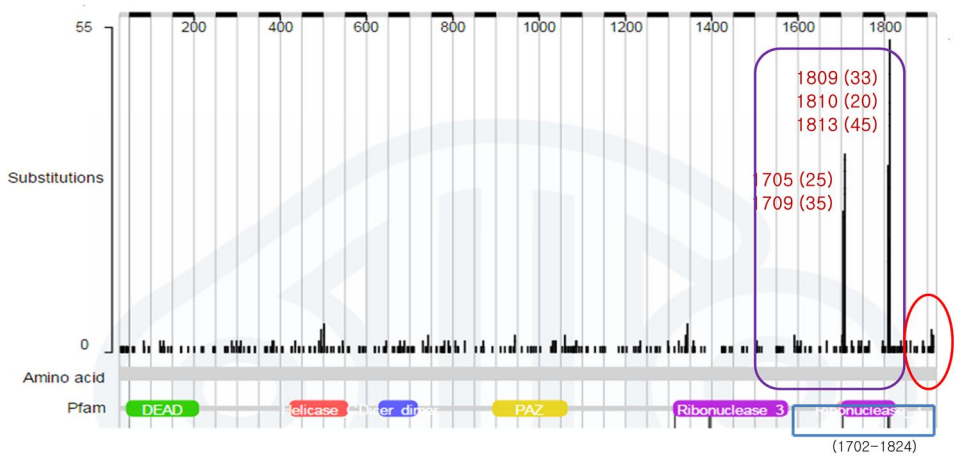

Figure 1.6. Dicer mutation distribution in cancer patients from COSMIC data. Hotpots of mutation locate in the RNase IIIa and b domains, number of residues are found with high frequency and well investigated. The mutation of dsRBD indicated in red circle is not well-known.

Chapter 2. Investigation of BRCA1 target gene in breast cancer

2.1. Aims of research

In this chapter, we want to verify novel direct target of BRCA1, which was known as a transcriptional regulator. The target genes will be suggested as biomarker candidates for BRCA1 associated breast cancer.

Here we report the results from the combined analysis of BRCA1-dependent expression profiles and BRCA1 ChIP-seq data using ENCODE, TCGA and GEO combination analyses, and verification of the results in breast cancer cell line. As example, we show the mRNA expression of three *FADD, MEIS2, CKS1B* genes have the significant correlation to the BRCA1 expression and the promoter of the three genes interact with BRCA1. Furthermore, we present a positive correlation between BRCA1 and the 3 targets in cancer cell lines and verify FADD as a novel direct target of BRCA1 by luciferase reporter assay, ChIP pull-down assay, and functional studies. Taken altogether, this study provides a comprehensive picture of the genome-wide transcriptional regulation by BRCA1 and suggests some of the target genes as biomarker candidates for BRCA1 associated breast cancer.

2.2. Material and method

2.2.1. Genome-wide investigation of BRCA1 target via GEO, ENCODE and TCGA data analysis

To verify the BRCA1 target genes, we identified BRCA1 binding sites using the BRCA1 ChIP-seq of 4 cell lines (hES, HeLa, HepG2, and GM12878) and calculated distance from the BRCA1 binding sites to the nearest gene transcription start site (TSS) from ENCODE data.

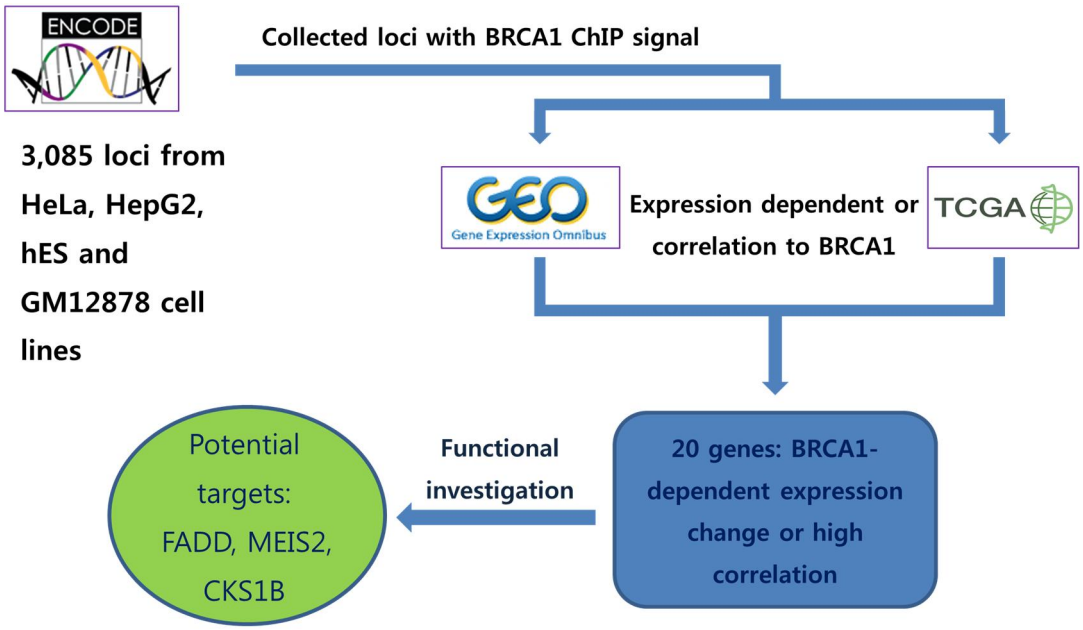

Figure 2.1. Schematic of investigating BRCA1 target via GEO, ENCODE and TCGA data integration analyses

To obtain the BRCA1-dependent expression profile, the two GEO dataset, GSE30822[57] and GSE22259 were used. TCGA containing human

breast cancer expression profiles were used to calculate the expression correlation between BRCA1 and other genes. Table 2.1 shows a list of top 20 genes from ENCODE, GEO and TCGA data analyses.

Table 2.1. Top 20 genes from integrative analyses

Gene Name	**ENCODE data (CHIP seq)**	**BRCA1-KD/control**	**BRCA1_overExp/ BRCA1_notExp _HCC1937**	**TCGA Correlation**
PSME3	unnamed (-265)	0.666859	1.279783	0.459748
CDC25C	unnamed (-6034)	0.656332	1.121402	0.431207
MLX	unnamed (-4998)	0.879046	1.275817	0.402135
NUSAP1	unnamed (-345)	0.798975	0.727273	0.393076
ASF1B	unnamed (-614)	1.090216	1.042378	0.38767
CENPQ	unnamed (-5)	0.693451	1.159113	0.386507
RAD51	unnamed (-196)	1.439286	1.078909	0.386464
BUB1B	unnamed (+176)	1.282035	0.882314	0.384827
KNTC1	unnamed (-477)	0.77683	0.527817	0.382264
NEIL3	unnamed (+107)	0.96816	0.826071	0.381176
OSR1	unnamed(+10168)	#N/A	0.92	-0.33
NFATC1	unnamed(+243883)	#N/A	1.43	-0.32
MEIS2	unnamed(+521171)	0.77	1.01	-0.32
ZBTB20	unnamed(-763032)	1.22	1.29	-0.31
FGF7	unnamed(+197663)	#N/A	1.28	-0.31
IFNGR1	unnamed (+434)	1.16	3.13	-0.30
SAV1	unnamed (-166)	1.15	0.96	-0.30
RPL13	unnamed (+5069)	0.83	1.10	-0.30
FBLN2	unnamed(+101496)	#N/A	1.653277	-0.28966
TRPM6	unnamed(-140879)	#N/A	1.062817	-0.28964

2.2.2. Cell culture

The MDA MB 436 cell line was used for overexpression of BRCA1, which is BRCA1-deficient. MCF7 cell line was used for the BRCA1 knockdown. MDA MB 436 and MCF7 cells were cultured in Dulbecco's Modified Eagle Medium supplemented with 10% (vol/vol) FBS and Penicillin (Gibco).

2.2.3. Overexpression and knockdown of genes

The BRCA1-siRNA and FADD-siRNA were chemically synthesized along with the scrambled siRNA as control (Genevolution, Korea). The sequences are followed as siBRCA1 - CAGCAGTTTATTACTCACTAAA; siFADD – GAACUCAAGCUGCGUUUAU. For siRNA treatments, subconfluent proliferating MCF7 cells were transfected with 50 nM of siRNA using the lipofectamine transfection reagent (Invitrogen, USA). After 24-h incubation, BRCA1 mRNA and protein level were measured. To obtain maximal knockdown, we carried out 48-h incubations with the siRNA. For BRCA1 overexpression, the MDA MB 436 cells were transfected with wtBRCA1 or C61G_BRCA1 mutant, and GFP as a control.

2.2.4. RNA extraction

RNA extraction was performed using TRizol (Invitrogen, Carlsbad, CA, USA) following the instructions from the manufacturer. One microgram of total RNA was used for cDNA synthesis (Superscript First-Strand Synthesis System; Invitrogen) by the manufacturer's protocol.

2.2.5. Real-time PCR analyses

The expression levels of BRCA1, FADD, MEIS2, UBE2S and CKS1B were measured by SYBR Green PCR Kit in LightCycler 480 II (Roche Applied Sciences, Indianapolis, IN, USA). Standard PCR conditions are: 10 minutes at 95 °C followed by 40 cycles of 15-second denaturation at 94 °C, 30-second annealing at 60 °C, and 30-second extension at 70 °C. The primer sequences for PCR are shown in Table 2.2. Human ribosomal protein gene RpL13a was used as an internal control gene. Relative quantification was calculated using the 2-(ΔΔCt) method.

Table 2.2. Sequences for real-time PCR Primers of BRCA1 targets

Gene name	Primer sequence
CKS1B	F: 5'-GGGGGG TATTCGGACAAATACGACGACG
	R: 5'-GGGGGG CGCCAAGATTCCTCCATTCAGA
FADD	F: 5'-GGGGGG ATGGACCCGTTCCTGGTGC
	R: 5'-GGGGGG TCAGGACGCTTCGGAGGTAG
MEIS2	F: 5'-GGGGGG GAGACCACGATGATGCAACCT
	R: 5'-GGGGGG ACTGTTGTCTAAACCATCCCCTT

2.2.6. Western-blot analyses

Western blot analysis was performed as previously described [18]. Briefly, cells were lysed in lysis buffer: 150 mM NaCl, 1 % Triton X-100, 1% Sodium deoxycholate, 50 mM Tris-HCl (pH 7.5), 2 mM EDTA (pH 8.0), and 0.1%

SDS. The 10~50 μg of protein were separated on SDS PAGE, transferred to Nitrocellulose membrane, probed with anti-BRCA1/FADD/MEIS2/CKS1B or FADD downstream antibodies (1:1000, Cell Signaling). Equal loading was confirmed with the anti β-actin antibody (1:1,000; Santa Cruz Biotechnology, CA, USA). Densitometric analysis was performed using the ImageJ 1.47v system.

2.2.7. Chromatin immunoprecipitation (ChIP) assay

ChIP was performed as previously desribed[18] with minor modifications. Briefly, MDA-MB-468 and MCF7 cells were cross-linked with 1% formaldehyde for 10 minutes at RT. Then cell pellets were lysed, re-suspended and subjected to sonication on ice. The 5 μl of sonicated chromatin was stored as a 5% input. The remaining samples were immunoprecipitated overnight at 4°C on a rocking platform using two kinds of BRCA1 antibody (#1 for Ab-1 (MS110, Merck, USA and #2 for home-made previously described[18]). Following overnight incubation, ChIP-grade Protein A/G Plus agarose beads were incubated for 2 h with the lysate at 4°C. Then the samples were washed with three different buffers and eluted in a buffer containing 5 M NaCl and 20 mg/mL Proteinase K. After that, the immunoprecipitated and input samples were reverse cross-linked at 65°C for 40 min. Finally, the elutes were used to detect ChIP signal by PCR (optimization), with the primers listed in Table 2.3.

Table 2.3. Sequences for ChIP Primers of BRCA1 targets

Gene name (length of DNA sequence)	**Primer sequence**
CKS1B (292)	F: 5'-GGGGGG AGAAAAACTGGGCGACAGGG
	R: 5'-GGGGGG TGGAGCGGTAACCTAAGCTG
FADD (280)	F: 5'-GGGGGG CTTTGCAAACAGGTGGACTCG
	R: 5'-GGGGGG CTCCGGTGCCTGATTCACTA
MEIS2 (268)	F: 5'-GGGGGG AGACATGGTTTCTCGGCAGG
	R: 5'-GGGGGG CGCCAGAGTTTCAGTAGGCA

Promoters were amplified from genomic DNA that was isolated from total DNA of MDA-MB-468 and MCF7 human cell lines. Coordinates of promoters and primers used for amplification are presented in Table 2.4. Amplified promoters were double digested by appropriate restriction enzymes and inserted into pGL3 enhancer vector (Promega, WI). The selected clones, with pGL3 enhancer vector alone, were grown in LB medium containing 50μg ml ampicillin to extract plasmid DNA. All resulting clones showed sequences exactly matched with the corresponding NCBI GeneBank sequences.

Table 2.4. Primers sequences for promoter PCR and cloning the promoters of BRCA1 targets

Gene name	R.E.	Primer sequence
CKS1B	**SacI**	F: 5'-CCC**GAGCTC**GAAGTGAGGCTGGGAGTCTG
	KpnI	R: 5'-CCC**GGTACC**GAGGCGGGACACCTCTACTTTC
FADD	**XhoI**	F: 5'-CCC**CTCGAG**TTGAATGGTATGTGAATTATATC
	HindIII	R: 5'-CCC**AAGCTT**ACAGGGCTACTGCGAAAATTG
MEIS2	**NheI**	F: 5'-CCC**GCTAG**CAATGTCGACGAGCAACTCAGC
	XhoI	R: 5'-CCC**CTCGAG**AGGAAGTGAGGAGGCAGTCATC

2.2.8. DNA promoter constructs and luciferase assay

To construct the promoter, we integrated the signal from ENCODE analyses to UCSC genome browser and get the promoter regions with BRCA1 ChIP-seq. Each construct of the target was about 2000 bp in length. The MDA MB 436, a BRCA1 deficient cell line, was selected for promoter assay. Transient transfections were performed in 24-well plates using Lipofectamine 2000 (Invitrogen, USA), in triplicates. As a control, cells were transfected with 100 ng of pGL3 Basic Vector or pGL Promoter Vector (pGL3-PV) (Promega, WI). To examine BRCA1 dependency for each promoter, 100 ng of GFP/ WT BRCA1/BRCA1-C61G expression plasmid were co-transfected. 24/36 hours after transfection, cells were lysed and the activities of Renilla and firefly luciferases were measured using Dual-Luciferase Reporter Assay System (Bio-Rad, USA). The obtained values of the luminescence for the

firefly luciferase were normalized to the values for the Renilla luciferase, and the data were corrected for background luminescence, normalized by SV40 promoter activity and averaged. For each candidate under study, at least three independent transfections were performed.

2.2.9. alamarBlue® cell proliferation assay

We performed the assay according to previous study[58]. The cells in a 96 well plate were transfected with control, BRCA1 overexpression vector and treated with siFADD plus Fas ligand (a gift from Prof. Han Seok Choi, University of Ulsan College of Medicine, Korea). After 24 hours, a 1/10th volume of alamarBlue® reagent (Invitrogen) was added directly to culture media. The cells incubated for an additional 1 hour to measure viability, which was detected by a microplate fluorescence spectrophotometer (GenTeks Biosciences, Inc., San-Chong, Taipei).

2.2.10. Patients' survival analyses via PROGgeneV2 data

We conducted survival analysis for the FADD signature whose downregulation was consistent with BRCA1 inactivated using PROGgen. We accessed PROGgeneV2 portal to check the patient's survival. The FADD and BRCA1 survival analyses were inputted as the genes of interest and breast cancer type was selected to create plots. The bifurcation gene expression at median and overall survival measure by death were selected as parameters for analyses. The TNBC in the TCGA breast cancer dataset and ER-positive in GSE3494_U133A dataset were adjusted as the subtypes for the analyses. The

statistical method and prognostic plots were calculated by backend scripts as described in the previous report[59].

2.2.11. Data present and statistics

The resulting graph was visualized by Excel Microsoft office program. All data are presented in as means ±standard error mean (SEM). Statistical significance between two groups was calculated using t-test. The p-value less than 0.05 were considered to be significant.

2.3. Result

2.3.1. Identification of potential BRCA1 target genes via ChIP-seq data from ENCODE

To identify novel BRCA1 target genes, we analyzed ENCODE data by collecting the loci with BRCA1 ChIP signal. We collected 3,086 loci found in all of the 4 cell lines (hES, HeLa, HepG2 and GM12878) and calculated the distance from each of the loci to the nearest gene transcription start site (TSS). The results showed most ChIP signals fall within 2kb (for both directions) of its promoter. The graph in Figure 2.2 shows that there is a higher portion of ChIP signal come from upstream of TSS, it is suggesting the role of BRCA1 in transcriptional regulation.

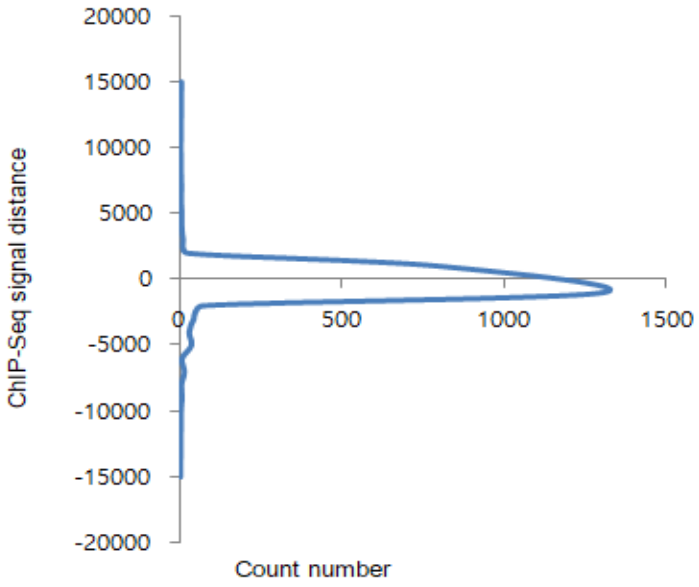

Figure 2.2. Schematic visualization of BRCA1 ChIP seq signal investigation from ENCODE data, showing the number of genes were according to the distance between the site of ChIP signal and TSS (transcription start site). Note the peak of the graph is bwlow of the x-axis, indicating higher number of ChIP signal comes from upstream of TSS compared to downstream.

2.3.2. GEO and TCGA integration analyses of BRCA1 and its targets correlation

From the integration data, the genes with BRCA1 binding signal were compared with the BRCA1-dependent expression profile data from GEO datasets. The GSE30822 contains expression profile of HCC1937, a BRCA1 deficient cell, along with HCC1937 reconstituted with wtBRCA1. On the other hand, the GSE22259 provides expression profile of HeLa cells with the BRCA1 knockdown. When we compare the two datasets, we found only weak

positive correlation (Figure 2.3.A). Therefore, we added breast cancer expression profiles from TCGA database to select the genes with high expression correlation with BRCA. The graph in Figure 2.3.B shows that there is a higher portion of genes positively correlated with BRCA1 in the three datasets.

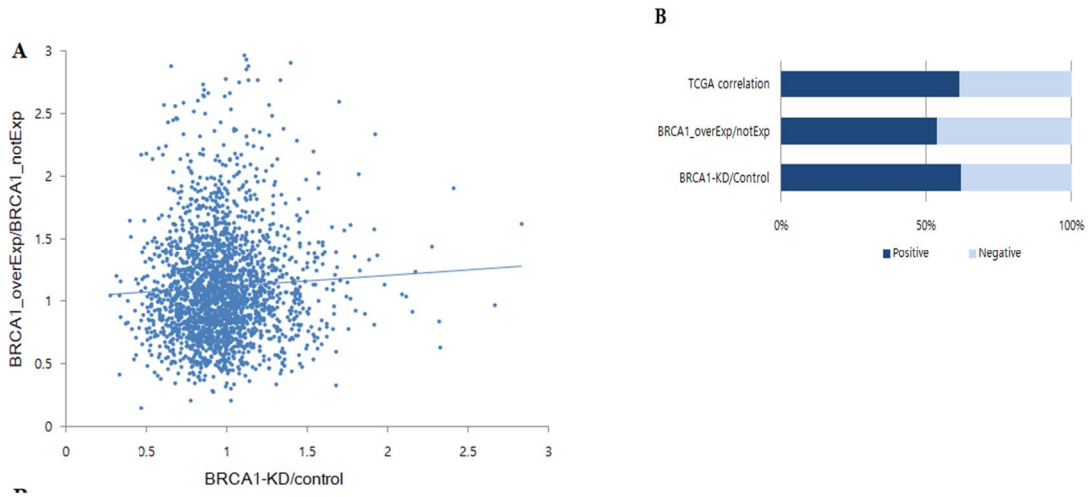

Figure 2.3. Integrated bioinformatics analysis of GEO and TCGA data to identify BRCA1 target genes

(A) Dot plot showing correlation between BRCA1 overexpression and knockdown. (B) Degree of expression correlation of BRCA1 with the other genes in overexpression, knockdown of BRCA1 and TCGA dataset. Positive correlation means the expression level of certain gene is increased by BRCA1 overexpression or decreased by BRCA1 knockdown.

2.3.3. BRCA1 target gene selection for cellular study

For further investigation, we selected 3 target genes FADD, MEIS2, CKS1B which have potential, functional relation to BRCA1 (Table 2.5).

Among them, MEIS2 was selected as a candidate for a long-range interacting target gene, as the BRCA1 ChIP signal was 521kb upstream of the TSS of MEIS2, which is the closest TSS annotated. We confirmed a significant correlation between these genes and the BRCA1 level, as determined by Spearman correlation coefficient (Table 2.5).

Table 2.5. Selected candidates of BRCA1 Target Genes from ENCODE / GEO / TCGA data analyses

Gene name	Function	ENCODE data (CHIP seq)[a]	BRCA1 KD/Ctr[b]	BRCA1 overExp/Ctr[c]	TCGA[d]
CKS1B	binding to the catalytic subunit of the cyclin dependent kinases and it is essential for their biological function	-319	0.66	1.90	0.29
FADD	an adaptor molecule that interacts with various cell surface receptors and mediates cell apoptotic signals	+152	1.06	2.10	0.24
MEIS2	binding to HOX or PBX proteins to form dimers, or to a DNA-bound dimer of these proteins and thought to have a role in stabilization of the homeoprotein-DNA complex for transcriptional regulation	+521171	0.77	1.01	-0.31

Note: [a] displays the site of ChIP-seq from TSS of targets via ENCODE data analyses; [b] shows the targets mRNA expression ratio of BRCA1 knockdown comparing to control from GEO datasets; [c] shows targets mRNA expression ratio of BRCA1 overexpressed comparing to control from GEO datasets; [d] shows correlation values from TCGA datasets analyses.

2.3.4. BRCA1 overexpression positively regulates FADD, MEIS2 and CKS1B expression

Following the bioinformatic analysis, we examined the BRCA1 dependent expression of selected 3 target genes in the BRCA1-deficient MDA

MB 436 cells, by the reconstitution of wild-type/C61G mutant BRCA1. The levels of BRCA1 RNA and proteins were accessed by western blotting and real-time PCR. The results in Figure 2.4A showed comparable expression of BRCA1 WT or C61G pathogenic mutant. Upon the BRCA1 expression, we observed significant upregulation of FADD, CKS1B, and MEIS2 by WT BRCA1, consistent with the bioinformatic analysis in Table 2.4 (Figure 2.3B). In contrast, the C61G pathogenic mutant BRCA1 failed to increase the expression of three genes, indicating the three genes are dependent on functional BRCA1. These results are further confirmed by western blot (Figure 2.4C and 2.4D for densitometry measurements).

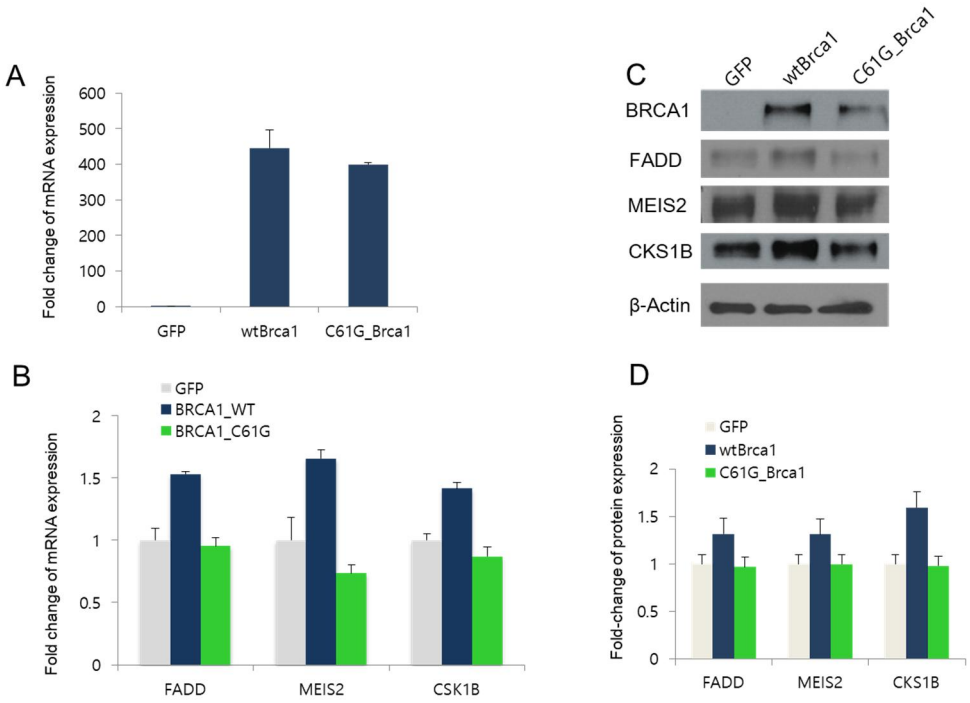

Figure 2.4. Identification of BRCA1 target genes by over-expression. (A). Real time PCR results of BRCA1 in BRCA1 deficient MDA-MB-436 cells or the cells

transfected with WT BRCA1 or C61G mutant BRCA1 construct. GFP expression plasmid was used as a control. (B) mRNA level of FADD, CKS1B and MEIS2 measured by real-time PCR after the expression of either WT or C61G BRCA1. (C) The same samples were analyzed by western blot to detect protein level. (D) Densitometry measurement of the data in (C).

2.3.5. BRCA1 knockdown negatively regulates FADD, MEIS2 and CKS1B expression

We next tested BRCA1 knockdown also can alter the expression of the three BRCA1 target candidate genes. Due to the poor efficiency of siRBA1 in MCF7, we introduced HEK393T cells (Fig. 2.5). As shown in the Figure 2.5A, the siRNA treatment of BRCA1 efficiently reduced BRCA1 expression. In the BRCA1 knockdown cells, we observed the expression of three candidate target genes are significantly reduced in mRNA level (Figure 2.5B), or in protein level (Figure 2.5C and 2.5D for quantization). Together, these data support BRCA1 transcriptionally unregulated FADD, MEIS2 and CKS1B candidates selected from the bioinformatics analysis.

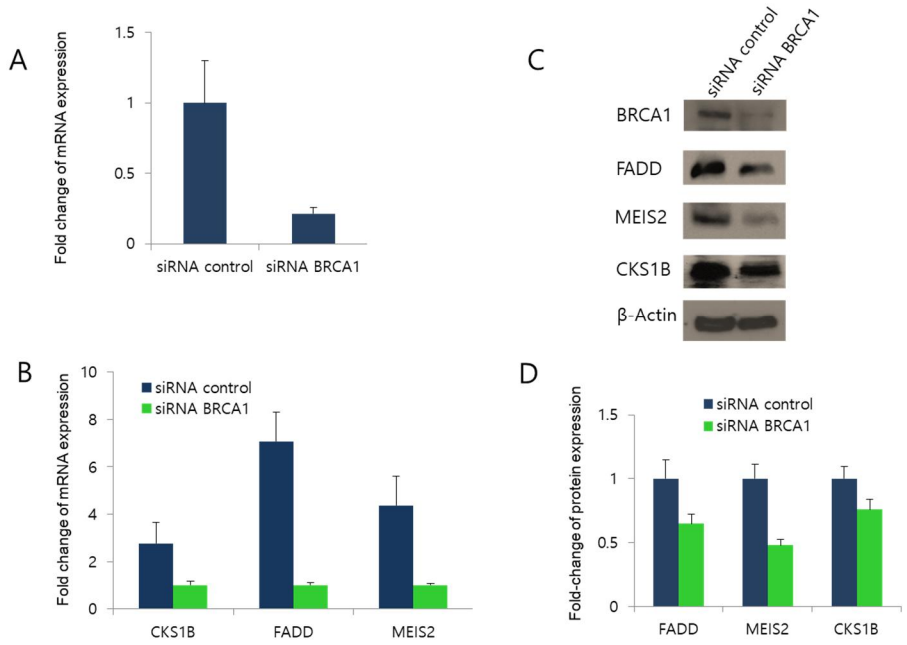

Figure 2.5. Knockdown of BRCA1 affect candidate target gene expression. (A). Real time PCR results of BRCA1 in HEK293 cells transfected with siRNA of BRCA1 or scrambled control. (B) mRNA level of FADD, CKS1B and MEIS2 measured by real-time PCR after the knockdown of BRCA1. (C) The same samples were analyzed by western blot to detect protein level. β-Actin was used as a loading control. (D) Densitometry measurement of the WB data in (C).

2.3.6. BRCA1 binds to FADD promoter among selected candidates

Further, clarify if the BRCA1 interacts with the FADD promoter, we performed ChIP analysis in MCF7 cells using two different BRCA1 antibodies and an antibody for HDAC2 that was shown to interact with BRCA1 in miR-155 regulation[18]. The results in Figure 2.6 show only #2 BRCA1 antibody detected BRCA1 interaction on the FADD promoter. Taken together, these

results indicate that induction of FADD mRNA by BRCA1, through the interaction on the FADD promoter and transcriptional activation.

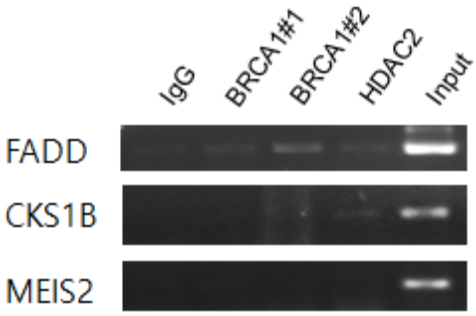

Figure 2.6. A ChIP assay results indicating BRCA1 is physically associated with the FADD promoter. IgG was used as a negative control. Two different BRCA1 antibodies marked as #1 and #2 were used for ChIP.

2.3.7. BRCA1 activates FADD promoter via luciferase expression

Next, we explore the mechanism(s) of how BRCA1 transcriptionally up-regulates the three candidates' gene expression. As the ChIP signal from ENCODE database suggested putative BRCA1 interaction on the promoter region of each gene (except for MEIS2), we tested the effect of BRCA1 expression on the activation of the candidate gene promoters. Each candidate promoter reporter construct was co-transfected with either wild-type or C61G mutant BRCA1expression vector into human breast carcinoma MDA MB 436 cells, followed by luciferase reporter assay. As a control, PLG3 vector was included. Among those 3 candidates, FADD promoter reporter constructs (Figure 2.7A) showed significant activation following the expression of WT

BRCA1, but not by the C61G mutant, in MDA MB 436 cell lines (Figure 2.7B for FADD and Figure 2.7C for other candidates).

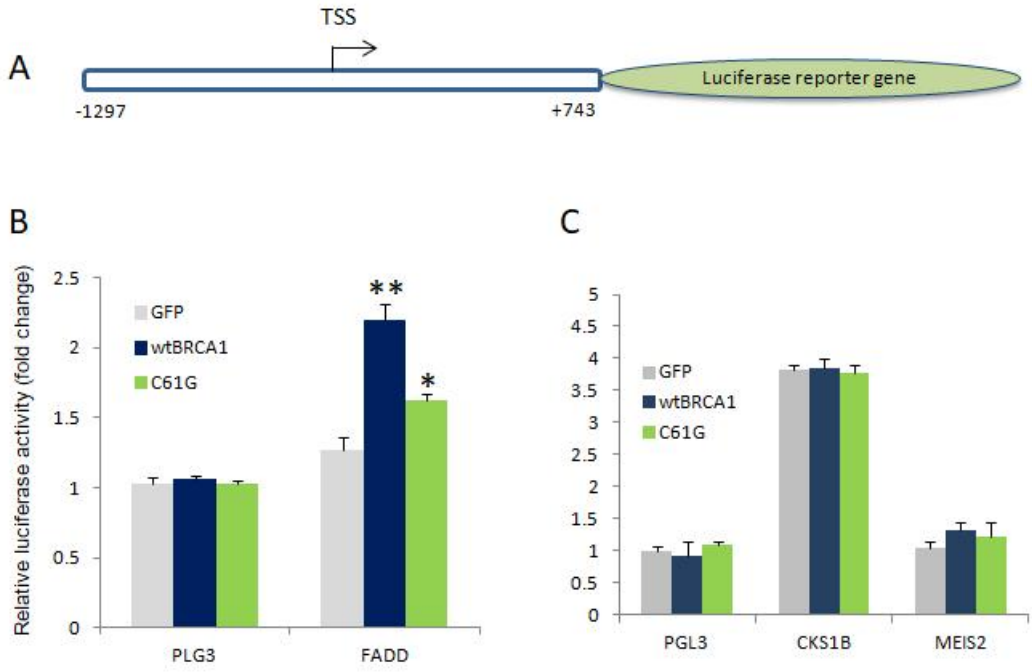

Figure 2.7. Promoter constructs and luciferase activity assay of BRCA1 candicates (A) Schematic presentation of the promoter construct of FADD used for luciferase assay. (B, C) Graphs showing the results of promoter assay. (B) FADD gene promoter was activated by WT BRCA1 (in deep blue) which is markedly reduced by the C61G mutant (in green). PGL3 vector only was used as a negative control while (C) CKS1B and MEIS2 gene promoters showed no changing luciferase activity in the same conditions.

2.3.8. FasL signaling and BRCA1 expression reducing proliferation

To understand how BRCA1 exerts its tumor suppressive effects by regulating FADD in breast cancer cells, we next checked the effect of BRCA1 overexpression on FADD related signaling. As shown in the Table 2.5, FADD is an adaptor molecule involved in Fas ligand-receptor mediated apoptosis. Therefore, we examined if the BRCA1 mediated FADD upregulation can sensitize cells to Fas ligand-mediated apoptosis. For this, BRCA1 is over-expressed in MDA MB 436 cells, in combination with Fas ligand and siRNA of FADD. Cell proliferation assay showed the Fas ligand treatment inhibited cell growth only with the expression of BRCA1, which is reversed by siRNA of FADD (Figure 2.8A). Western blot analysis confirmed increased FADD by BRCA1 expression and efficient knockdown of the FADD by siRNA (Figure 2.8B). These data indicate that BRCA1 expression in MDA-MB-436 cells stimulates Fas ligand induced cell death signaling via increased FADD level.

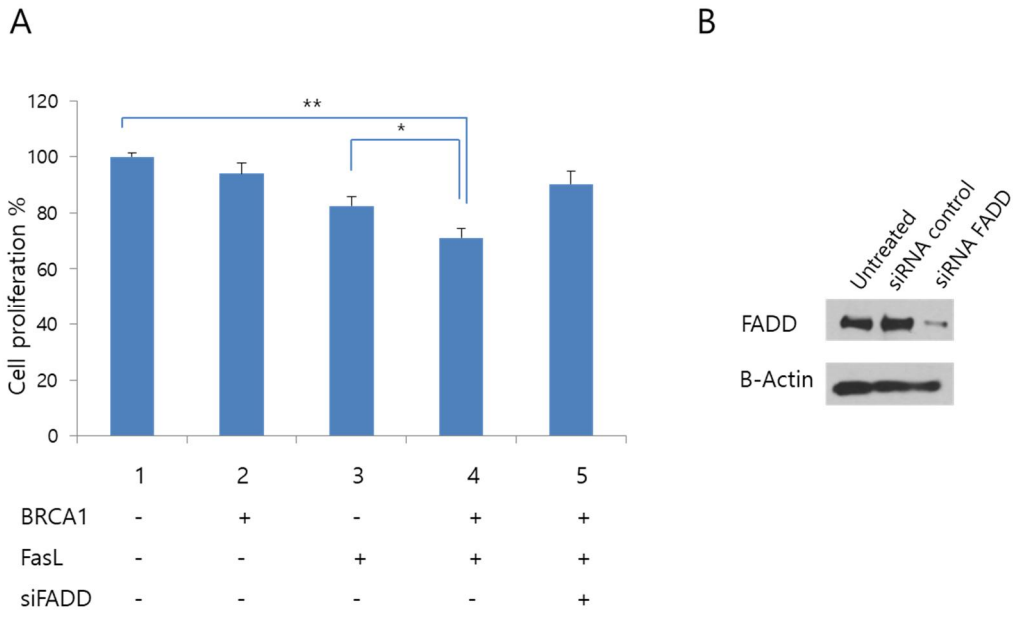

Figure 2.8. BRCA1 expression sensitizes cancer cell to FasL via FADD.

(A) Proliferation assay of MDA MB 436 cells treated with FasL in the combination with BRCA1 and siRNA of FADD.

(B) Western blot analysis to confirm knockdown siRNA of FADD. Beta-Actin was used as a loading control.

2.3.9. Low expression of FADD is correlated with poor survival of TNBC patients, meanwhile the opposite occurs in estrogen receptor (ER)-positive cases

Based on our results showing *FADD* as a BRCA1 target, we further examined whether this finding is valid in breast cancer patients. To determine whether the expression level of FADD could be a biomarker for patient survival, we used PROGgeneV2[59].

As BRCA1 is frequently inactivated in the TNBC subtype, we initially focused on TNBC cases. Figure 2.9A shows that low FADD mRNA expression (low FADD, n = 14; high FADD, n = 7; HR = 0.37, p = 0.011) was significantly correlated with low BRCA1 expression (Figure 2.9B) (low BRCA1, n = 11; high BRCA1, n = 9; HR = 0.87, p = 0.60) and poor overall survival of TNBC patients, supporting our results that FADD is a mediator of BRCA1 function in suppressing tumor cell proliferation (Figure 2.8A). In contrast, the survival curve for ER-positive type breast cancers (Figure 2.9C) showed that low expression of FADD (low FADD, n = 18; high FADD, n = 31; HR = 2.24, $p < 0.001$) and BRCA1 (Figure 2.9D) (low BRCA1, n = 20; high BRCA1, n = 29; HR = 3.13, p = 0.021) are beneficial for survival, suggesting that the effect of FADD on the patient's survival differs in a subtype-specific manner.

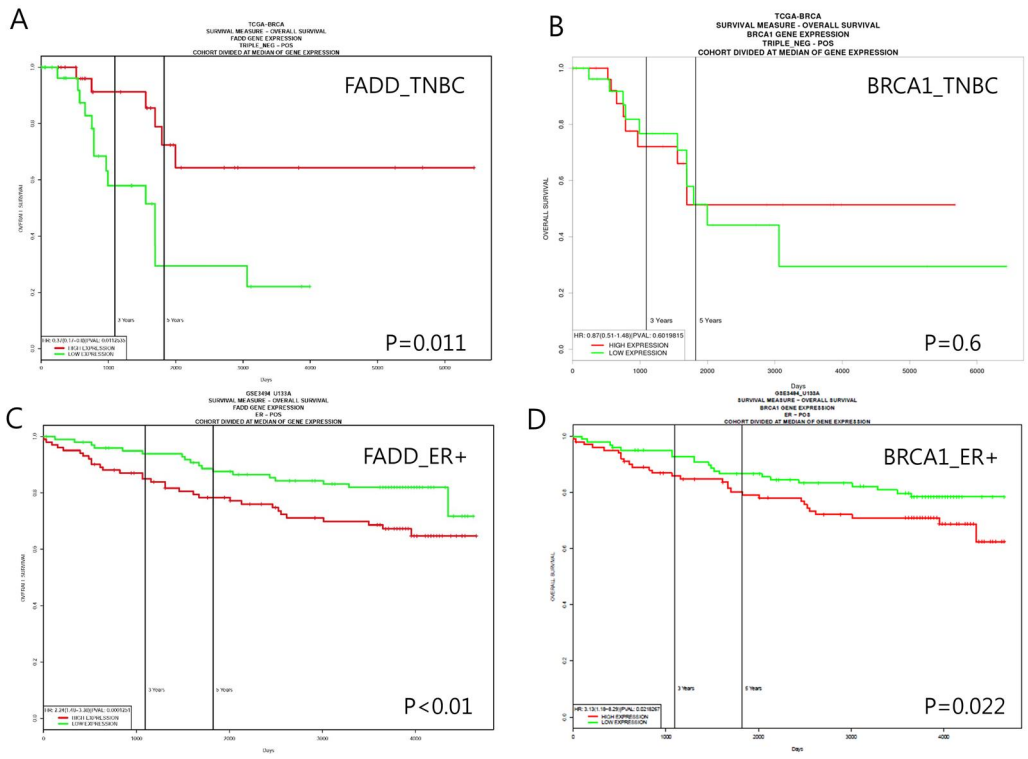

Figure 2.9. PROGeneV2 analysis showing that low *FADD* mRNA expression is significantly correlated with poor overall survival of TNBC patients. Overall survival curve of TNBC patients grouped according to the (**A**) *FADD* and (**B**) *BRCA1* expression level. Note that a low level (green line) of *FADD* and *BRCA1* expression was correlated with poorer survival, compared to higher level of FADD (red line). Overall survival curve of ER-positive breast cancer patients grouped according to (**C**) *FADD* and (**D**) *BRCA1* expression level. Note that a low level (green line) of *FADD* and *BRCA1* expression was significantly correlated with better survival for ER-positive cancer patients.

2.4. Discussion

Many studies have shown the important roles of BRCA1 tumor suppressor in DNA repair, cell cycle progression and transcriptional regulation. People have been put much effort into investigating the mechanism of how BRCA1 is involved in these important biological events. Previous reports have shown a number of genes targeted by BRCA1, but it was unclear whether these expression changes are due to the direct interaction of BRCA1 on the regulatory elements. In this regard, the integration of several genome-wide data of ENCODE, TCGA and GEO can allow researchers to identify the BRCA1 associated genetic elements and develop it as putative prognostic markers for breast cancer. The present study by the analysis of ENCODE ChIP-seq data combined with GEO and TCGA data suggests that the BRCA1 expression may directly regulate a large number of gene expression.

Known as a multifunctional protein, BRCA1 suppressor tumor formation via a transcriptional co-activator or co-repressor function[60]. Consistent with this, BRCA1 was shown to display a genome-wide impact on gene expression by forming BRCA1-containing complexes. For example, BRCA1 was known to associate with ERα protein and represses its signaling, ERα-dependent transcription[61, 62]. Moreover, BRCA1 was shown to bind to HDAC2, a member of chromatin remodeling factor, suggesting it has a function in

epigenetic control. Indeed, BRCA1 was reported to modulate the activity of various co-factors within promoter-bound complexes[63].

The three targets FADD, CKS1B and MEIS2 were selected among the candidate genes from raw data, with related literature review. Those three genes have roles in different cell functions such as cell proliferation, regulation of cell cycle, and signal transduction (Table 2.5). The low-expression of these target genes were also reported in different tumorigenesis events and considered as important markers. Fas-associated death domain (FADD) is an adapter protein that is recruited to the death-inducing signaling complex (DISC) during signaling via death receptors. Hence, FADD plays an important role in apoptosis in aged lymphocytes and reported as a co-tumor suppressor in cancer cell lines[64-66]. On the other hand, CKS1B is a protein with a well-documented role in the cell-cycle regulation. The CKS1B protein binds to the catalytic subunit of the cyclin-dependent kinases and is essential for their biological function. Also, CKS1B controls MEK/ERK and JAK/STAT3 pathway[67]. MEIS2 was identified in an array screening as one of the top 50 downregulated genes in prostate cancer[68]. The study of MEIS2 in low-grade prostate tumors suggests that it plays a critical function related to poor prognosis and might be a useful biomarker or therapeutic target. In our study, BRCA1 binds 521kb upstream of the TSS of MEIS2 but surprisingly, the BRCA1 expression could increase MEIS2 RNA about two-fold (Figure 2.4B). As we could not observe promoter activation of CKS1B and MEIS2 or

positive ChIP signal of BRCA1 (Figure 2.6), we suggest altered post-transcriptional regulation by microRNA or other non-coding RNA might cause the increase of CKS1B and MEIS2 expression by BRCA1. Indeed, there are a set of miRNAs reported to be regulated by BRCA1[60], which needs to be further examined for the possible regulation of the two genes.

The reporter assay showed activation of FADD following the expression of BRCA1 in MDA MB 436 cell line. The over-expression of BRCA1 has been shown to increase Fas and Fas-ligand level in MCF7 cells[69]. In the report, they also show that dominant negative form of FADD can inhibit BRCA1 induced apoptosis. In our study, however, we could not observe apoptosis by the BRCA1 expression alone (Figure 2.8A), but by the co-treatment of Fas-ligand. We consider this is because we used MDA-MB-436 which is different from MCF7 in BRCA1 status. A number of studies have suggested that BRCA1 be involved in the processes of apoptosis[70-73] but the role of FADD is, to our knowledge, not reported so far. Therefore, our results add novel information for the apoptosis via BRCA1.

Lastly, the survival curve in Figure 2.9 showed poor prognosis of TNBC patient with the low expression level of FADD. Related to this, a recent meta-analysis in head and neck cancer showed FADD as a prognostic marker, consistent with our results[74]. Considering that the knockdown of BRCA1 caused reduction of the FADD level, we suggest that the loss of BRCA1 (or its inactivation) in a tumor can cause reduced levels of FADD that in turn

desensitizes cells to a FasL-mediated anti-proliferative effect. Interestingly, the correlation was reversed in ER-positive or luminal-type breast tumors. At present, it is not clear why this difference occurs, but we speculate that the presence of BRCA1 in ER-positive or luminal tumors disables the ability to overcome FADD-induced cell death. Further functional studies of putative BRCA1 target genes may help to fully understand the roles of this multifunctional tumor suppressor.

Chapter 3. Role of Dicer mutations in breast cancer via miRNA biogenesis pathway

3.1. Aims of research

In this chapter, we want to investigate the effect of Dicer mutations via clinical data. Then we examine the effective mutations on the C-terminal domain of Dicer protein – RNaseIIIb and dsRBD (double strand RNA binding) domains – in miRNA biogenesis and understanding of the microRNA precursor-Dicer interaction specifically for miR-inhibition.

The mutations are the key to investigating dsRBD domain as a well-defined specificity and provide valuable tools to the current understanding of the miR155-Dicer interaction specifically. The human Dicer domains and mutations were expressed in E.coli and purified by His-tag affinity and FPLC chromatography. The purified proteins were used for binding assay and crystallization. The structural interaction of Dicer/miRNA precursor is performed by auto-made docking. To investigation impact of Dicer mutants, we are going to introduce those mutants into Dicer deficient cells. The initial results show the reduction miR-155, miR-21 in breast cancer cell via luciferase activity assay.

3.2. Material and method

3.2.1. Clinical data observation of Dicer mutants

We collect the clinical data from TCGA, COSMIC and Asan Medical Center mutant of Dicer in cancer patients. To test whether mutants in the dsRBD domain can effectively generate mature miRNAs in human cancer cells and in-vitro, we select the C-terminal mutants of human Dicer.

3.2.2. Structural prediction of Dicer and miRNA precursors

To evaluate the selected mutants of Dicer in the previous step, we analyzed the interaction of dsRBD of Dicer and the two pre-miRNA loops of miR-155 and miR-4502. Because dsRBD domain structure of human Dicer has been solved, the dsRBD domain of mouse Dicer was used for the docking. The miRNA loops of miR-155 and miR-4502 were building using HADDOCK programs[75]. Then these loops were used for NPDock web-based automated-docking to optimize the interaction[76]. PyMOL program was used for structure visualization[77].

3.2.3. DNA constructs and cloning

The human *Dicer1* gene in encodes a total of 1922 amino-acid residues. Such a big protein, we could not express full length in bacteria. A homodimer of RNaseIIIB and dsRBD domains of mouse Dicer1 is reported that it could be expressed in bacteria and process the biogenesis for miRNA maturation in-

vitro[78]. Though we decided to construct this partial protein for bacterial expression of human Dicer and named as RNaseIIIb / RNaseIIIb_dsRBD, their length are 170 and 243 amino acids. The sequence of DNA and protein in this study are listed in the Table 3.1.

Table 3.1. Amino acid sequences of truncated human Dicer constructs

Constructs	**Sequences**
RNaseIIIBD	YLLQAFTHASYHYNTITDCYQRLEFLGDAILDYLITKHLYEDPR QHSPGVLTDLRSALVNNTIFASLAVKYDYHKYFKAVSPELFHVI DDFVQFQLEKNEMQGMDSELRRSEEDEEKEEDIEVPKAMGDIFE SLAGAIYMDSGMSLETVWQVYYPMMRPLIEKFSANVP
RNaseIIIBD&dsRBD	YLLQAFTHASYHYNTITDCYQRLEFLGDAILDYLITKHLYEDPR QHSPGVLTDLRSALVNNTIFASLAVKYDYHKYFKAVSPELFHVI DDFVQFQLEKNEMQGMDSELRRSEEDEEKEEDIEVPKAMGDIFE SLAGAIYMDSGMSLETVWQVYYPMMRPLIEKFSANVPR**S**PVREL LEMEPETAKFSPAE**R**TYDGKVRVTVEVVGKGKFKGVGRSYRIAK SAAARRALRSLK**A**NQPQVPNS

Note: The red-highlight residues are selected for mutant study.

The RNaseIIIb / RNaseIIIb_dsRBD were amplified *via* PCR using cDNA from HEK 293 cell total RNA extract as a template. The DNA sequences were cloned using Gateway Cloning systems (Invitrogen, Carlsbad,

CA). The sequences of the forward and reverse oligonucleotide primers designed from the NCBI genome sequence were in Table 3.2, respectively.

Table 3.2. Primer for PCR and cloning of truncated human Dicer constructs

Primer names	**Sequences**
F:attb_ RNaseIIIBD&d:	ggggg**acaagtttgtacaaaaaagcaggcttcgaaaacctgtattt**
	tcagggcTACCTTATTCTTCAGGCTTTGAC
R:attb_ RNaseIIIBD	ggggg**accactttgtacaagaaagctgggtc**ttaGGGTACATTTGC
	AGAAAAC
R:attb_Dicer	ggggg**accactttgtacaagaaagctgggtc**ttaTCAGCTATTGGG
	AACCTGAG

Note: The bases in bold designate the *atta&attb* sites[79, 80].

The amplicon was ligated into the pDONOR207 entry vector by BP recombination reaction (Figure 3.1). Then the DNA coding sequence in entry vector was shifted into destination expression vectors by LR recombination reaction (Figure 3.1). We used four vector pDEST-HGWA, pDEST-HMGWA, pDEST-HXGWA and pDEST-PDIb'a' which encodes four different tags 6xHis, 6xHis_MBP, 6xHis_TRX and 6xHis_PDI tag, respectively. Those four tags were inserted upstream of a TEV cleavage site in the vectors. The correct construction sequences that resulted from the BP and LR recombination

reactions were confirmed using a DNA sequencing service (Macrogen, Daejon, Korea). All constructs contained the tags for ease of purification.

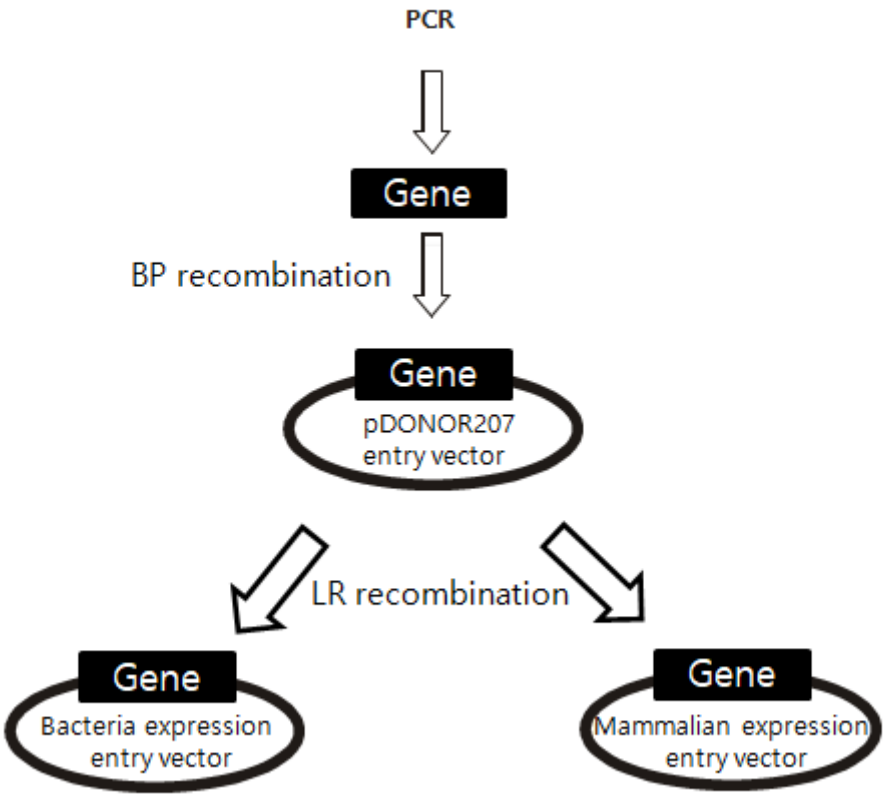

Figure 3.1. Gate-way cloning systems was used for RNaseIIIb and RNaseIIIb_dsRBD cloning

3.2.4. Expression of hDicer in E.coli

Overexpression of the constructs containing RNaseIIIb and RNaseIIIb_dsRBD were tested in bacteria. Each expression vector was transformed into competent E.coli BL21 (DE3) host cells and cultured at 37^0C

in Luria Bertani medium (containing 100mg/mL ampicillin) until the A600 optical density value reached 0.4~0.6. The expression system was induced using 1mM 1-thio-b-d-galactopyranoside (IPTG) for overnight at low-temperature 18^0C and for 6h at high-temperature 37^0C. The harvested cells from 5 ml of the cultured cell were sonicated with 500µl lysis buffer (Table 3.3), and then the soluble and insoluble parts were separated using centrifugation at 13,000rpm for 15min at 4^0C. The expression and solubility of protein were assessed using SDS-PAGE and 10% tricine gels.

Table 3.3. Buffer conditions were used for protein purification

Lysis/binding buffer	**Wash buffer I**	**Wash buffer II**	**Elution buffer**
20mM Tris pH 8.0 300mM NaCl 20mM Imidazole 5% glycerol	20mM Tris pH 8.0 600mM NaCl 20mM Imidazole 5% glycerol	20mM Tris pH 8.0 300mM NaCl 30mM Imidazole 5% glycerol	20mM Tris pH 8.0 100mM NaCl 250mM Imidazole 5% glycerol

Then the optimized condition of RNaseIIIb and RNaseIIIb_dsRBD proteins were cultured at a condition of 18^0C for 16h and collected by centrifugation at 3600rpm for 30min at 4^0C. Then the pellets were stored at -20^0C until further study.

3.2.5. Purification of hDicer

The frozen cell pellets were homogenized in lysis buffer (Table 3.3) at a 10mL/g ratio, and the lysate was sonicated 30 times (on/off for 10/10s) to disrupt the cells. The soluble protein in the supernatant was collected and obtained by centrifugation at 15,000rpm for 12min at 4^0C, filtered using a 0.45mm pore size filter before purification. The filtered supernatant was applied to a 2mL Ni-NTA HisTrap affinity column (GE Healthcare, Piscataway, NJ) that was equilibrated in the lysis buffer. The column was washed with 2 volumes of Wash buffer I and II (Table 3.3) to remove any nonspecific binding, and the elution step was performed using Elution buffer containing high imidazole concentration. A 10kDa cut-off membrane (Viskase Corporation, Darien, IL) was used to reduce the salt concentration of the purified MBP-RNaseIIIb and MBP-RNaseIIIb_dsRBD in a buffer A (20mM Tris-HCl, 5% glycerol [v/v], 100mM NaCl, 2mM DTT; pH 8.0) for TEV cleavage. Purified TEV was added to the sample at a ratio of 1:10 (TEV:fusion protein [w/w]) and incubated for 24 h at 25^0C with shaking at 100rpm, but the TEV could not remove MBP tags efficiently. The MBP-RNaseIIIb and MBP-RNaseIIIb_dsRBD sample were concentrated to 5mg/ml and then dialyzed in buffer D (25mM Tris-HCl pH 7.4, 5% glycerol [v/v], 2mM DTT) using a dialysis membrane with a 10kDa cut-off membrane (Viskase, Darien, IL) for 5h at 4^0C. The purification steps were checked using SDS-PAGE with 14%

tricine gels. Protein concentrations were measured using the BCA protein assay (Pierce Biotechnology, Rockford, IL) using BSA as the standard.

3.2.6. Site directed-mutagenesis

Primers for site directed-mutagenesis were designed base on pDEST-HMGWA and Plex307 vectors containing RN3BD_dsRBD constructs via PrimerX web-based program[81]. The RN3BD_dsRBD mutants were prepared by using KOD-Plus-Mutagenesis Kit (Toyobo, Osaka, Japan) in accordance with the manufacturer's protocol with modification (Table 3.4). Then the sequences of the mutant clones were confirmed by automated DNA sequencing service (Macrogen, Daejon, Korea).

Table 3.4. Site-directed mutagenesis primers and protocols for Dicer mutants

Mutants	**Primer sequences**	**Annealing**
R1850C	F: GCAAATGTACCC**T**GTTCCCCTGTGCG	58^0C – 30secs
	R:CGCACAGGGGAAC**A**GGGTACATTTGC	
A1870S	F:CTGCCAAATTTAGCCCG**T**CTGAGAGAACTTACGAC	60^0C – 30secs
	R:GTCGTAAGTTCTCTCAG**A**CGGGCTAAATTTGGCAG	
A1914V	F:CCCTCCGAAGCCTCAAAG**T**TAATCAACCTCAGGTTC	58^0C – 30secs
	R:GAACCTGAGGTTGATTA**A**CTTTGAGGCTTCGGAGGG	

Note: The mutation sites are indicated in red color.

3.2.7. Isothermal titration calorimetry (ITC) assay

The binding of pre-miR155 and miR4502 to wild-type MBP-RNaseIIIb and MBP-RNaseIIIb_dsRBD was monitored by ITC. Titrations were performed on a Nano ITC low-volume titration calorimeter (TA Instruments). In a typical ITC experiment, pre-miR-155 and miR-4502 (10mM) was titrated with wild-type MBP-RNaseIIIb or MBP RNaseIIIb_dsRBD (0.1mM). The instrument was equilibrated at 25°C until the baseline was flat and stable. All experiments were performed in 25mM Tris pH 8.0 buffers with stirring at 300rpm.

3.2.8. Cell culture and transfection

The MDA MB 436 cell line was used for overexpression of Dicer and mutants that were cultured in Dulbecco's Modified Eagle Medium supplemented with 10% (vol/vol) FBS and Penicillin (Gibco). Transient transfections were performed in 24-well plates using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) in triplicates.

3.2.9. Western blot

Preparation of cell lysates and Western blot analysis were performed as previously described in chapter 2. Membranes were probed with anti-Dicer (1:500, Santa Cruz Biotechnology, CA, USA), anti-6HIS (1:500, (Invitrogen, Carlsbad, CA, USA). As loading control, anti-beta-Actin (1:1, 000; Santa Cruz Biotechnology) antibody was used.

3.2.10. Luciferase activity assay

Cells were seeded in 24-well plate and transfected with 150 ng of sponge or reporter plasmid DNA, which has one copy of miRNA binding site for four target miRNAs. As an internal control, 10 ng of SV40-Luc (Renilla luciferase) per well was used. Forty-eight hours after transfection, cells were lysed and luciferase activity was measured using Dual-Luciferase Reporter Assay System (Promega) and GloMax Luminometer (Promega), according to the manufacturer's protocol.

3.2.11. Crystal structure of protein for modeling and docking

The full length Dicer from mouse was retrieved from Protein Data Bank (PDB ID: 3C4B[78]) because the human Dicer structure is lack of dsRBD at the time of investigation. The pre-miR-155 and pre-miR-4502 loops structure were build base on 23S RNA (PDB 3CC2) template structure[82] and the complex-docking predicted via NPDOCK[76]. The dsRBD/miR-loop complexes were visualized using PyMOL software[83].

3.2.12. sgRNA design for Dicer K/O

To construct the sgRNA-cloning vector, the sgRNAs of Dicer were constructed by CHOPCHOP v2 web server[84]. We selected 3 constructs base on the results of the analyses then order for oligo-DNA synthesis service (Macrogen, Daejon, Korea). Then those 3 sets were annealed and inserted behind the promoter sequences of PLKO-cloning vector. The target sequences

of each sgRNA are listed in Table 3.5. We used lentiCas9-Blast plasmid (Addgene plasmids #52962) for cas9 gene expression.

Table 3.5. sgRNA target sequences for Dicer knockout in breast cancer cells

sgRNA_Dicer	Nucleic acid sequences
sgRNA_1	F: CACCG CCTTCGTTTCGTGGAACCTGGTC
	R: AAAC GACCAGGTTCCACGAAACGAAGG
sgRNA_2	F: CACCG TTCAACGTGGAGCTTACCAGGGG
	R: AAAC CCCCTGGTAAGCTCCACGTTGAA
sgRNA_3	F: CACCG CCCCTGGTAAGCTCCACGTTGAA
	R: AAAC TTCAACGTGGAGCTTACCAGGGG

3.3. Result

3.3.1. Mutants selection and structural prediction

To investigate whether mutants can effectively generate mature miRNAs in the database, we collected and created a list of human Dicer mutants. Furthermore, we pay attention to the other literatures have been studied on those mutants and find out that the interaction of dsRBD domain and miRNA precursors were not well-known up to date. Moreover, expression and purification of RNaseIIIb and dsRBD domain have not been reported

previously. We decided to clone the RNaseIIIb and RNaseIII_dsRBD domain and express in bacteria, respectively (Table 3.1).

Considering among 20 amino acids, the polar amino acid owns a higher affinity to the DNA. Following this path, there is hypothesis that the mutations of polar residue to the non-polar following affinity reduction and vice versa. Indeed, this was also confirmed by previous studies[85, 86] (Figure 3.2). As a final, we selected three typical mutants in the dsRBD domain from data: R1851C as an example of polar to non-polar mutation; A1870S for non-polar to polar mutant; and A1914V for the neutral mutant.

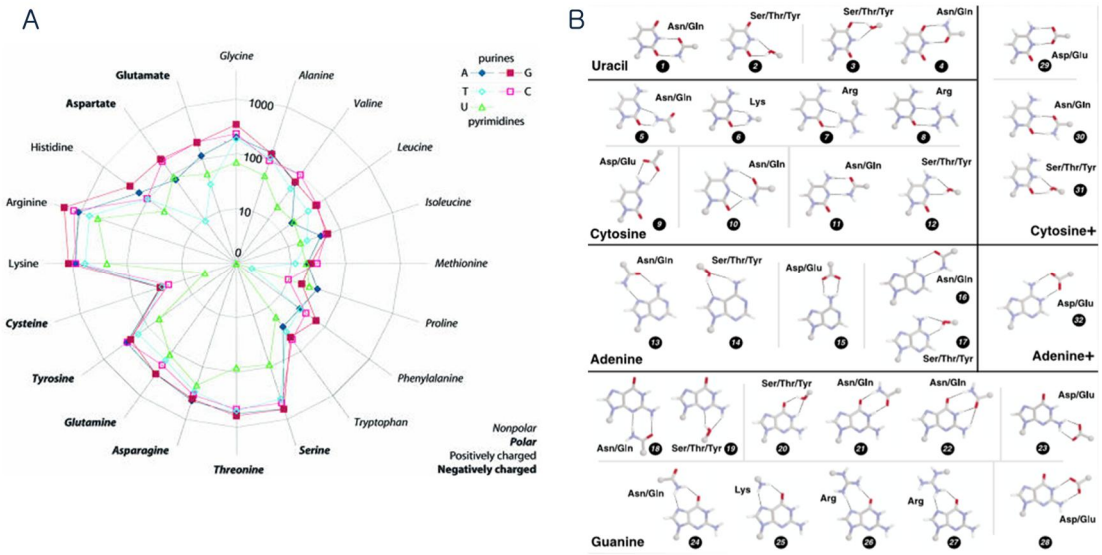

Figure 3.2. Polar amino acids have a high potential attractive to ribonucleotides RNA by forming the hydrogen bonds.

(A) Polar residues have higher affinity to the nucleic acid (source[85])

(B) Amino acid–base interactions with two or more hydrogen bonds. Interactions are grouped by base type, and are subgrouped by the interacting face (source[86])

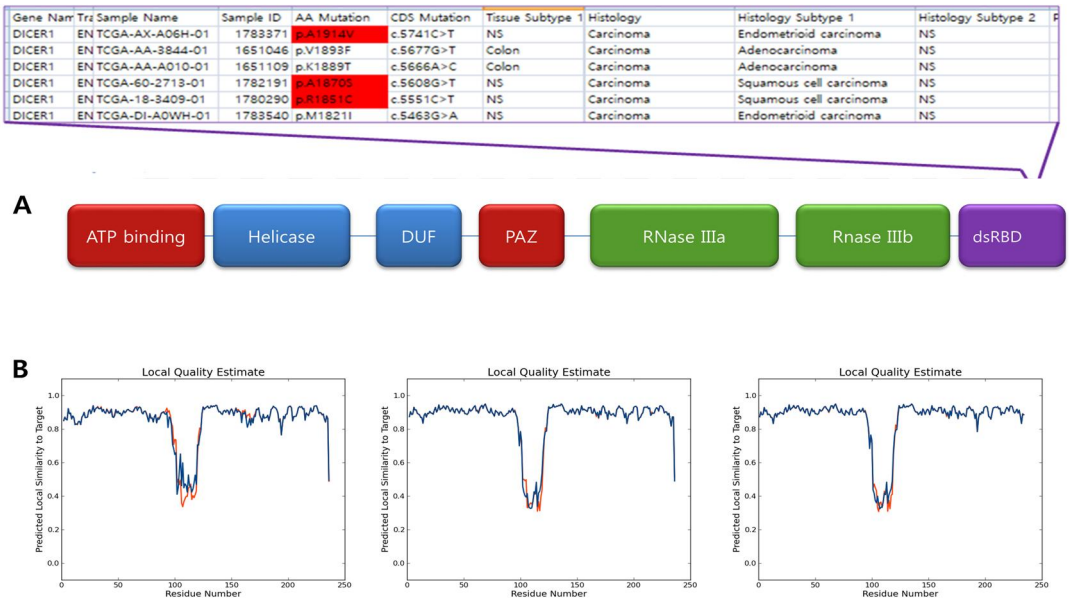

Figure 3.3. Selected mutants for further study

(A) R1851C, A1870S and A1914V mutants in dsRBD domain are selected from clinical data.

(B) Structural of mutant dsRBD are compare to the wild-type dsRBD, the local quality estimate peak shows no change in structure affected by the mutants, from left to right R1851C, A1870S and A1914V. The computation was analyses using SWISS-MODEL homology-modeling server[87].

3.3.2. Structure modeling of miRNA - precursor loops and miRNA-dsRBD domain complexes

The pre-mi155 and pre-miR4502 loops were building by ModeRNA automated modeling server[88]. The best results base on calculation of pre-miR155 and pre-miR4502 was used for dsRBD-loop complex docking (Figure 3.4). The results of automate docking was showed in different manners of

interaction between miR155 and miR4502 loops, even though they are high similarity of sequences. However, the directions of interaction are similar in both complexes. The R1851 and A1870 residues are near to the loop, while the A1914 is far-away of interaction.

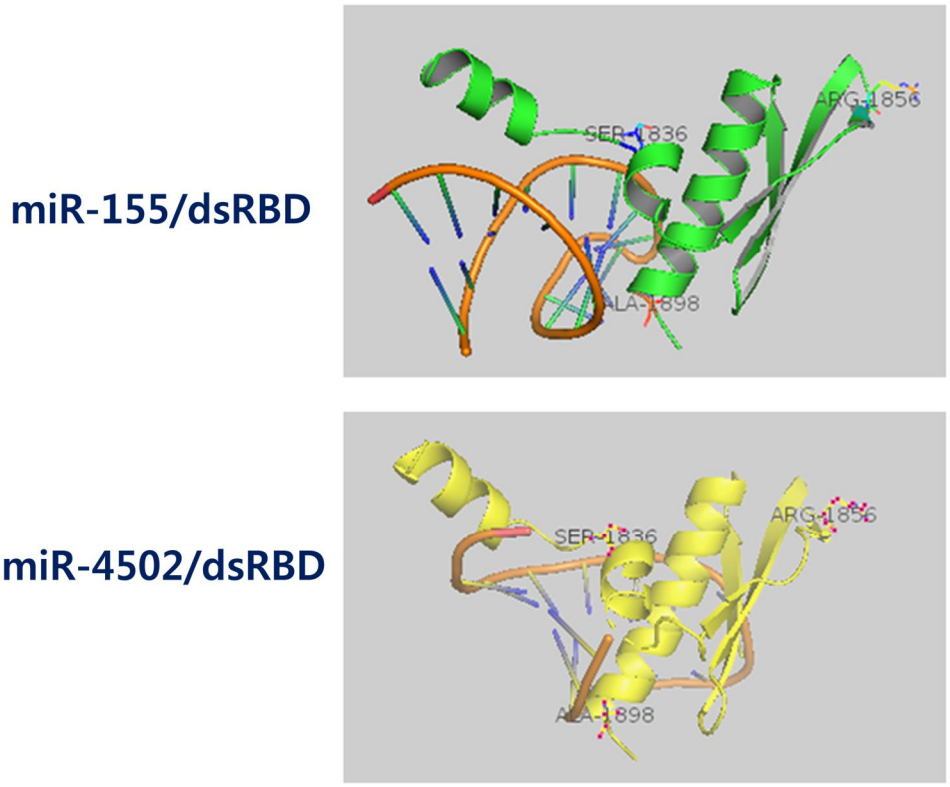

Figure 3.4. Structural docking of mouse dsRBD to pri-miRNAs. The pri-miR-155 (A) and mi-R4502 (B) loop were modeled and used for docking. Labeled residues are selected for mutation.

3.3.3. Cloning of RNaseIIIb and RNaseIIIb_dsRBD

Previous studies reported the expression of human Dicer in the eukaryotic expression systems, but not in bacteria. Firstly, the construct of two

RNaseIIIb and RNaseIIIb_dsRBD DNA sequences were inserted into pDONOR207 entry vector and confirmed by sequencing (Figure 3.5).

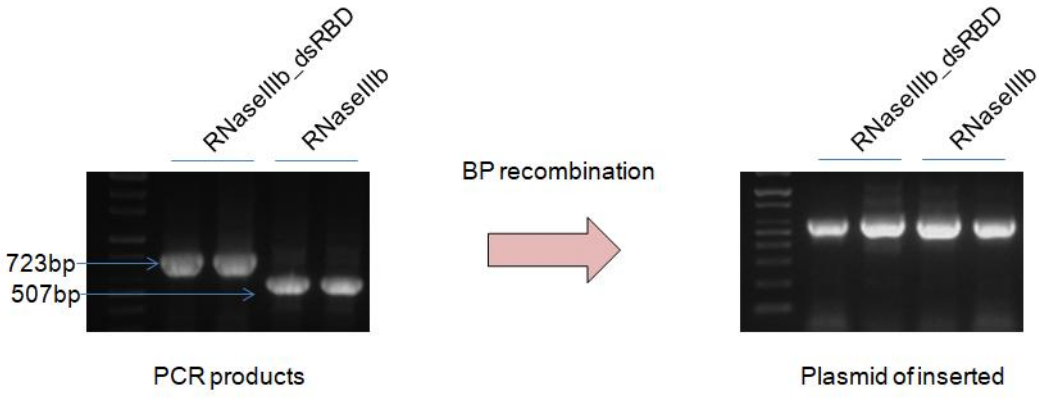

Figure 3.5. DNA constructs of RNaseIIIb / RNaseIIIb_dsRBD are amplified from HEK293T cell cDNA and inserted into pDONOR207 plasmid

To determine whether RNaseIIIb and RNaseIIIb_dsRBD are able to be expressed as soluble forms in *E. coli*, eight plasmids expressing N-terminal His6, His6_GST, His6_MBP, and His6_PDI–tagged variants were successfully shifted into expression vectors containing the four tags (Figure 3.6), respectively. For mammalian expression vector, the RNaseIIIb and RNaseIIIb_dsRBD were inserted into Plex307 by LR recombination reaction. To obtain tag-free RNaseIIIb and RNaseIIIb_dsRBD, a TEV restriction site was placed between each tag and RNaseIIIb or RNaseIIIb_dsRBD sequences. All bacteria constructs also contained His6 tags to facilitate the affinity

purification (Fig. 3.6). The constructs were then transformed into BL21 (DE3) to allow the expression of the proteins.

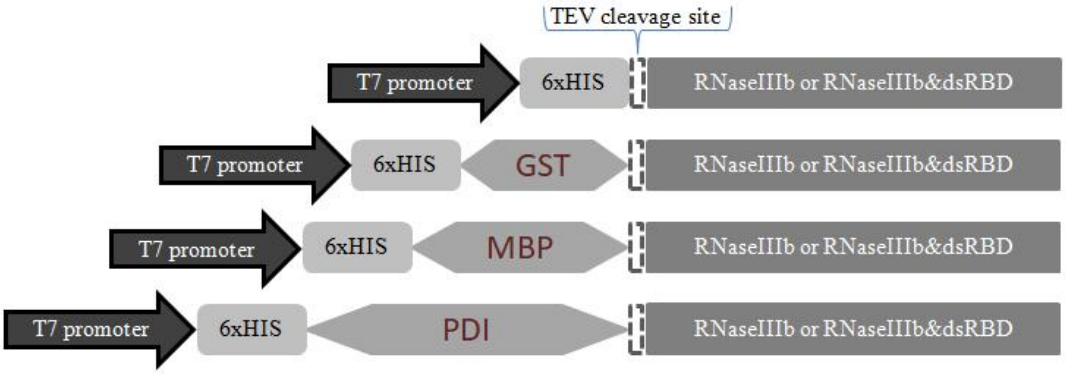

Figure 3.6. Schematic visualization of the domain structure of the RNaseIIIb or RNaseIIIb&dsRBD fusion variants in bacteria expression plasmid. The eight fusion variants of RNaseIIIb and RNaseIIIb&dsRBD were shifted into expression plasmid using the LR recombination Gateway cloning method: His6, hexa(poly)histidine; GST, glutathione-S-transferase; MBP, maltose-binding protein; PDI, full-length human PDI. The expression of the fusion proteins is driven by the IPTG-inducible T7 promoter with ampicillin as the selection marker.

3.3.4. Expression evaluation of RNaseIIIb and RNaseIIIb_dsRBD

The RNaseIIIb or RNaseIIIb_dsRBD fusion protein expression vectors were designed to be driven under the control of the IPTG-inducible T7 promoter. The expression level and solubility varied depending on the tag and temperature. Among four constructs we could not get an expression with the single His6-tag proteins (data not showed). Overall evaluation, the GST and

MBP-tagged protein showed the highest expression levels in insoluble forms and PDI in soluble form at 37^0C (Figure 3.7). Lower expression temperature at 15°C significantly increases the expression level of any of MBP tagged protein in the soluble form. Moreover, the low temperature was known to improve the solubility of the protein significantly. As a result, we chose the His6MBP-RNaseIIIb or His6MBP–RNaseIIIb_dsRBD construct for further study. More conventionally, the MBP tag allowed for straightforward downstream purification.

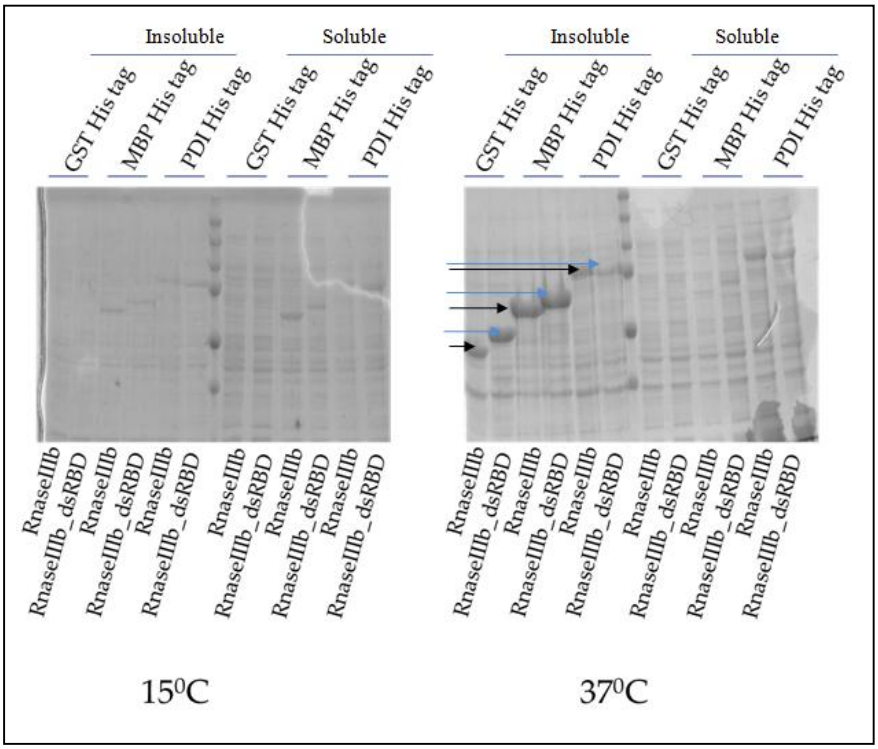

Figure 3.7. Bacteria expression was screening of RNaseIIIb and RNaseIIIb&dsRBD with different tags at low and high temperatures. The blue arrows show the band of RNaseIIIb&dsRBD; and the black arrows show the band of RNaseIIIb.

3.3.5. Purification of MBP-RNaseIIIb and MBP-RNaseIIIb_dsRBD

A total amount of the 5ml purified proteins was obtained in an overall yield of 1.5mg/ml. The His6MBP-RNaseIIIb and His6MBP-RNaseIIIb_dsRBD proteins were further purified using HiTrap Q ion exchange column (GE Healthcare) and were analyzed via SDS-PAGE (Fig. 3.8). The molecular weight of purified 6xHisMBP-RNase IIIB / -RNase IIIB& dsRBD are 58kDa / 70kDa.

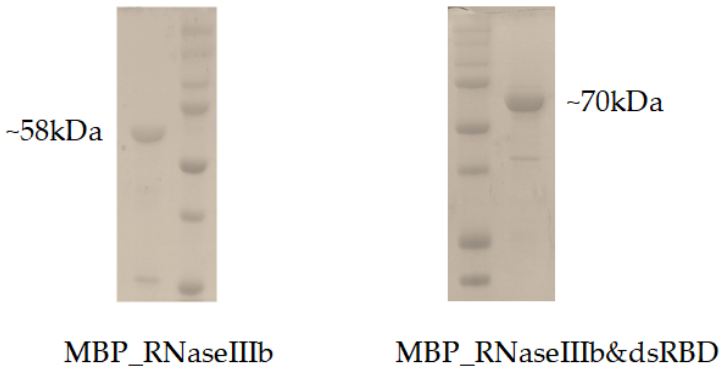

Figure 3.8. Purification of 6xHisMBP-RNase IIIB / -RNase IIIB& dsRBD proteins

For crystallization, the protein solution was concentrated using Centri-Prep (Millipore) to a final concentration of 3 mg/ml in a buffer consisting of 25 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl, 5% glycerol (v/v) but we fail to get the crystal growth after several months of incubation.

3.3.6. Binding affinity of MBP-RNaseIIIB and MBP-RNaseIIIB&dsRBD to miRNA precursors

To examine the role of the dsRBD domain, we check the binding activity of the recombinant proteins with/without carrying the dsRBD domain. When the MBP-RNaseIIIb_dsRBD activity was analyzed and compared to the RNase IIIb by ITC assay. The preliminary results showed that dsRBD domain was significantly altered binding activity to pre-miRNAs, both miR-155 and miR-4502 precursors (Figure 3.9). Thus, the dsRBD domain was considered as a crucial determinant of binding activity to pre-miRNAs and was further investigated by site-directed mutagenesis with selected mutants.

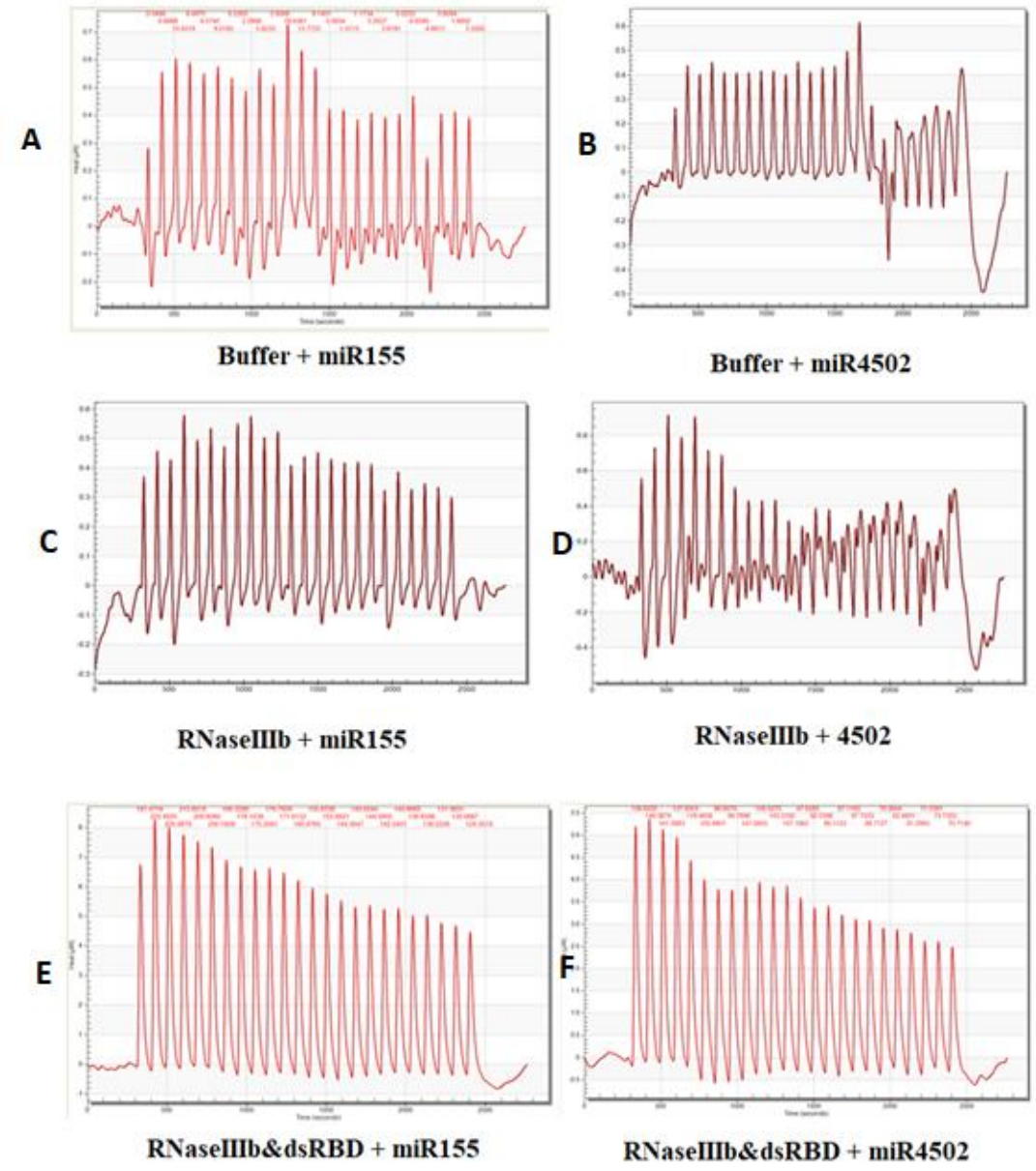

Figure 3.9. Preliminarily ITC assay analyses of humans partial Dicer and pre-miRNAs

(A, B) Control of buffer and pre-miRNAs

(C, D) Biding affinity of RNaseIIIb to pre-miRNAs

(E, F) Significant binding of RNaseIIIb_dsRBD to pre-miRNAs

3.3.7. Dicer mutants expression effecting to level of miRNAs in breast cancer cell

To over-express Dicer and its mutants, we check the level of Dicer expression in cancer cell lines. The expression of Dicer was shown very low both of mRNA and protein level in the MDA-MB-436 breast cancer cell line (Figure 3.11). We used the MDA-MB-436 for study of Dicer and its mutants.

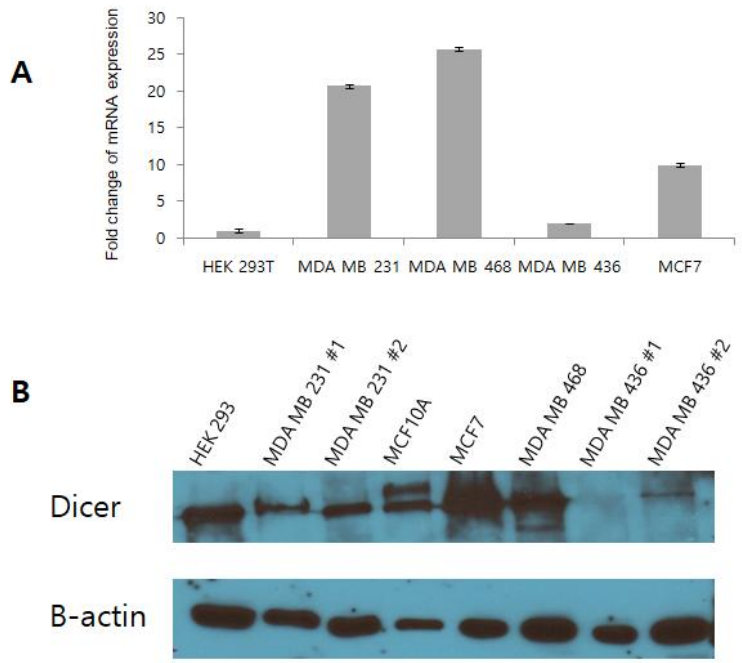

Figure 3.10. mRNA expression level of Dicer in breast cancer cell lines. (A) mRNA level of Dicer expression; (B) Protein level of Dicer expression

Using the MDA-MB-346 breast cancer cell line, we hypothesized whether the expression of Dicer mutants could reduce or increase the miRNAs in the transected cell line. The reporter plasmid was designed for miR-155 and

miR-21 inhibit in our lab as previously describe[89]. To test the impact of mutants and wtRNaseIIIb_dsRBD, the level of each target miRNA was measured by luciferase activity after the transfection 48h. As shown in Figure 3.12, we observed dramatic reduction of all of the 2 miRNAs upon the expression of the partial Dicer mutants compared to the wild-type. Those showing the expression of Dicer mutants functionally decrease the mature miR-155 and miR-21 thus these cause increasing of luciferase activity.

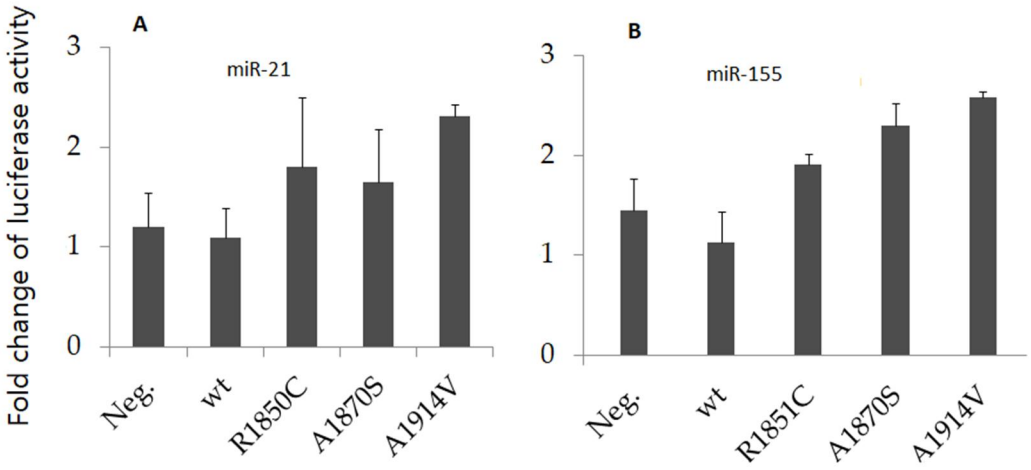

Figure 3.11. Expression of Dicer mutants decreasing the miR-21 (A) and miR-155 (B) in MDA-MB-436 breast cancer cell line via luciferase reporter assay.

3.4. Discussion and future plan

Dicer was known as a key protein in the miRNA biogenesis pathway. *DICER1* mutation carriers have an increased predisposition to cancer, such as pleuropulmonary blastoma[90] and Sertoli-Leydig cell tumor[91]. Although breast

cancer has been reported with *DICER1* mutation carriers that was not well established[92, 93]. The structure of human Dicer also solved but not fully understanding on interaction mechanism to miRNA precursors, especially on the dsRBD domain[30, 78, 94]. It is valuable to fulfill this lacking of mechanism about Dicer mutants and its structural interaction. The Dicer has an important role in miRNA biogenesis, which cleavage directly pre-miRNA into mature miRNA in the cell[95-98]. The miRNA has been known as a therapeutic target, as an oncogene or tumor suppressor or even a dual-function, in cancer[99-101]. Many mutations of Dicer have been shown in numerous studies and were associated with the global reduction of miRNAs in cancer[45, 90, 91, 102]. Those investigated mutations are in the RNaseIIIa&b domains, which have role in cleavage activity of the enzyme. However, there is limitation in understanding the interaction of dsRBD to pre-miRNA. The three selected mutations R1851C, A1870S and A1914V in this study may support a novel understanding of the interaction could be effected via structural analyses and functional studies. Furthermore, the investigation could answer the question about the mutation of Dicer may affect in different manner to the miRNAs.

In the present study, the partial Dicers, the RNaseIIIb and RNaseIIIb_dsRBD domains, were expressed and purified in E.coli systems successfully. An in-vitro binding assay of pre-miR155 and miR4502 to wild-type MBP-RNaseIIIb and MBP-RNaseIIIb_dsRBD was introduced using ITC assay. And the preliminary result of this experiment assists our hypothesis

about dsRBD domain and pre-miRNAs affinity. The truncated-dsRBD protein showed the lower affinity to both pre-miRNAs, suggesting the binding function of this domain in Dicer to the pre-miRNA loop. Thus, the position of this domain was considered as a crucial determinant of pre-miRNA loop binding activity, and was further investigated by site□directed mutagenesis of this domain.

In the next step, we checked for the production of miRNAs appears to be less or more efficient by the selected mutants in cellular studies. The preliminary results shown decreasing defective of mutations R1851C, A1870S and A1914V via luciferase activity assay. Dicer has an important role in microRNA pathway, which has been previously reported with numerous significant mutations[91, 93]. The mutations in the hostpot of RNaseIIIb were almost well-known, but still unclearly understanding of the rest of C-terminal domain – dsRBD. The three mutants are selected in the dsRBD from cancer clinical patients, which might be considered as missense mutations because of having distinct effects on miRNA processing in the breast cancer. This effect could be explained according to the role of Dicer-interacting protein. The mutated residues may maintain the stabilization of the complex Dicer/pre-miRNA/cofactor thus the change in these residues could affect on the interaction and defects the miRNAs biogenesis. This is an aspect that has not been explored to date and might have a significant influence on the miRNA biological processes regulated by the Dicer and co-factor proteins.

For further understanding of this, we will express and purify the RNaseIIIb_dsRBD with the mutation and check for the changing in affinity to miRNAs by ITC assay and immuno-precipitation pull down. The K/O Dicer will be generated and introduce the wild-type or mutants Dicer to compare the level of miRNAs in the cells. Although, we will examine the cancer cell characteristics changing by those mutant, such as cell proliferation, migration or apoptosis. For the interaction impact, the yeast two-hybrid assay could be used to understand the interactions between Dicer and its co-factor, such as TRBP[103], the system is to allow for detecting moderate to subtle changes in protein interaction affinities. Whether downstream of miRNAs can ever be an effective response in mammalian cells, including cancer cells. Moreover, currently remains unclear of those mutations are specifically effect to certain miRNA or not, we did not know up to now and that question needs to be answered in the additional investigation. The results could add novel information for understanding the microRNA synthesis mechanism, which provide valuable knowledge to the pharmaceutical research for miRNA targeting in breast cancer.

Chapter 4. Conclusion

Firstly, our study provided a novel the role of BRCA1 as a transcriptional factor on transcriptional regulation, on a genome-wide scale, by combining ChIP signal from ENCODE and expression data from TCGA and GEO integration analyses. Further functional study of putative BRCA1 target genes aids us to better understand this multifunctional tumor suppressor. Among genome-wide candidates, we further studied three genes and verified FADD as a novel direct target of BRCA1 by ChIP, real-time PCR and luciferase reporter assay and suggested as biomarker candidates for BRCA1 associated breast cancer. The functional assay showed activation of FADD following the expression of BRCA1 in our investigation. The over-expression of BRCA1 has been shown to decrease cell proliferation via Fas and Fas-ligand. Moreover, the high expression of BRCA1 and FADD was shown beneficiation in breast cancer patient survival via prognosis data inquiry. Therefore, our results add novel information for understanding the apoptosis mechanism via BRCA1.

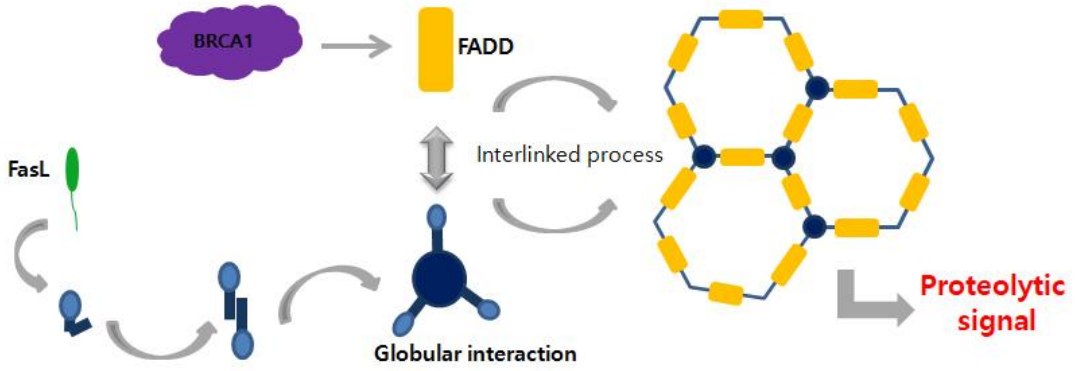

Figure 4.1. BRCA1 activates expression of FADD. The consequence promotes interlinked DISC formation and clustering, which open Fas molecules interacting via the globular interaction are linked by a multitude of dimer-Fas/FADD bridges leading to overall stable DISC clusters. This results in activation of caspase-8, and induces of the proteolytic signal for the cell apoptosis.

In the second study, we investigated of Dicer mutants via clinical data and selected three R1851C, A1870S and A1914V mutants for further characterization. We also visualization the interaction of dsRBD domain to miR155 and miR4502 precursor loops. We also fused MBP-RNaseIIIb and MBP-RNaseIIIb_dsRBD proteins with four special tags and compared their expression levels and solubility in different *E. coli* strains. An efficient MBP-RNaseIIIb and MBP-RNaseIIIb_dsRBD expression and purification protocol as a soluble form in *E. coli* was introduced. An in-vitro biding affinity showed the potential of the dsRBD domain to pre-miRNAs are confirmed. Preliminary examination of miR155 and miR21 levels affect by the three mutants R1851C, A1870S and A1914V showed their defective on miRNA synthesis in breast

cancer cell via luciferase reporter assay. We also suggest three mutants in the dsRBD domain of Dicer and further studies to characterize the individual novel missense mutations will need to include *in vitro*, cell and/or animal models. The understanding of Dicer mutations in miRNA synthesis pathway might lead to the discovery of regulatory factors amenable to pharmacological intervention in breast cancer.

References

1. AmericanCancerSociety. Breast Cancer Facts & Figures 2017-2018. *Atlanta: American Cancer Society, Inc* 2017.
2. KoreanCancerAssociation. Cancer Statistics in Korea: Incidence, Mortality, Survival, and Prevalence in 2014. *Cancer Research and Treatment* 2017; **49**(2): 292-305.
3. Hall JM, Lee MK, Newman B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 1990; **250**(4988): 1684-9.
4. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 1994; **266**(5182): 66-71.
5. Rosen EM, Fan S, Pestell RG, Goldberg ID. BRCA1 gene in breast cancer. *Journal of cellular physiology* 2003; **196**(1): 19-41.
6. Narod SA, Salmena L. BRCA1 and BRCA2 mutations and breast cancer. *Discovery medicine* 2011; **12**(66): 445-53.
7. Hilakivi-Clarke L. Estrogens, BRCA1, and breast cancer. *Cancer research* 2000; **60**(18): 4993-5001.
8. James CR, Quinn JE, Mullan PB, Johnston PG, Harkin DP. BRCA1, a potential predictive biomarker in the treatment of breast cancer. *The oncologist* 2007; **12**(2): 142-50.
9. Boulton SJ. Cellular functions of the BRCA tumour-suppressor proteins. *Biochemical Society transactions* 2006; **34**(Pt 5): 633-45.
10. Brodie KM, Henderson BR. Characterization of BRCA1 protein targeting, dynamics, and function at the centrosome: a role for the nuclear export signal, CRM1, and Aurora A kinase. *The Journal of biological chemistry* 2012; **287**(10): 7701-16.

11. Coyne RS, McDonald HB, Edgemon K, Brody LC. Functional characterization of BRCA1 sequence variants using a yeast small colony phenotype assay. *Cancer biology & therapy* 2004; **3**(5): 453-7.

12. Vallon-Christersson J, Cayanan C, Haraldsson K, et al. Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families. *Human molecular genetics* 2001; **10**(4): 353-60.

13. Ruffner H, Verma IM. BRCA1 is a cell cycle-regulated nuclear phosphoprotein. *Proceedings of the National Academy of Sciences of the United States of America* 1997; **94**(14): 7138-43.

14. Anantha RW, Simhadri S, Foo TK, et al. Functional and mutational landscapes of BRCA1 for homology-directed repair and therapy resistance. *eLife* 2017; **6**.

15. Scully R, Chen J, Ochs RL, et al. Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell* 1997; **90**(3): 425-35.

16. Cortez D, Wang Y, Qin J, Elledge SJ. Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. *Science* 1999; **286**(5442): 1162-6.

17. Scully R, Anderson SF, Chao DM, et al. BRCA1 is a component of the RNA polymerase II holoenzyme. *Proceedings of the National Academy of Sciences of the United States of America* 1997; **94**(11): 5605-10.

18. Chang S, Wang RH, Akagi K, et al. Tumor suppressor BRCA1 epigenetically controls oncogenic microRNA-155. *Nat Med* 2011; **17**(10): 1275-82.

19. Monteiro AN, August A, Hanafusa H. Evidence for a transcriptional activation function of BRCA1 C-terminal region. *Proceedings of the National Academy of Sciences of the United States of America* 1996; **93**(24): 13595-9.

20. Chapman MS, Verma IM. Transcriptional activation by BRCA1. *Nature* 1996; **382**(6593): 678-9.

21. Zheng L, Pan H, Li S, et al. Sequence-specific transcriptional corepressor function for BRCA1 through a novel zinc finger protein, ZBRK1. *Molecular cell* 2000; **6**(4): 757-68.

22. Zhang H, Somasundaram K, Peng Y, et al. BRCA1 physically associates with p53 and stimulates its transcriptional activity. *Oncogene* 1998; **16**(13): 1713-21.

23. Jin S, Zhao H, Fan F, et al. BRCA1 activation of the GADD45 promoter. *Oncogene* 2000; **19**(35): 4050-7.

24. Fan W, Jin S, Tong T, et al. BRCA1 regulates GADD45 through its interactions with the OCT-1 and CAAT motifs. *The Journal of biological chemistry* 2002; **277**(10): 8061-7.

25. Chang S, Sharan SK. The role of epigenetic transcriptional regulation in BRCA1-mediated tumor suppression. *Transcription* 2013; **4**(1): 24-8.

26. Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 2001; **409**(6818): 363-6.

27. Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes & development* 2001; **15**(20): 2654-9.

28. Knight SW, Bass BL. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* 2001; **293**(5538): 2269-71.

29. MacRae IJ, Zhou KH, Li F, et al. Structural basis for double-stranded RNA processing by dicer. *Science* 2006; **311**(5758): 195-8.

30. Lau PW, Potter CS, Carragher B, MacRae IJ. Structure of the human Dicer-TRBP complex by electron microscopy. *Structure* 2009; **17**(10): 1326-32.

31. Tsutsumi A, Kawamata T, Izumi N, Seitz H, Tomari Y. Recognition of the pre-miRNA structure by Drosophila Dicer-1. *Nature structural & molecular biology* 2011; **18**(10): 1153-8.

32. Lau PW, Guiley KZ, De N, Potter CS, Carragher B, MacRae IJ. The molecular architecture of human Dicer. *Nature structural & molecular biology* 2012; **19**(4): 436-40.

33. Macrae IJ, Zhou K, Li F, et al. Structural basis for double-stranded RNA processing by Dicer. *Science* 2006; **311**(5758): 195-8.

34. Takeshita D, Zenno S, Lee WC, Nagata K, Saigo K, Tanokura M. Homodimeric structure and double-stranded RNA cleavage activity of the C-terminal RNase III domain of human dicer. *Journal of molecular biology* 2007; **374**(1): 106-20.

35. Liu Z, Wang J, Cheng H, et al. Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. *Cell* 2018; **173**(6): 1549-50.

36. Lau PW, Guiley KZ, De N, Potter CS, Carragher B, MacRae IJ. The molecular architecture of human Dicer. *Nat Struct Mol Biol* 2012; **19**(4): 436-40.

37. Taylor DW, Ma EB, Shigematsu H, et al. Substrate-specific structural rearrangements of human Dicer. *Nature structural & molecular biology* 2013; **20**(6): 662-+.

38. Titov, II, Vorozheykin PS. Comparing miRNA structure of mirtrons and non-mirtrons. *BMC genomics* 2018; **19**(Suppl 3): 114.

39. Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. *Genes & development* 2004; **18**(24): 3016-27.

40. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & development* 2003; **17**(24): 3011-6.

41. Bohnsack MT, Czaplinski K, Gorlich D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *Rna* 2004; **10**(2): 185-91.

42. Saito K, Ishizuka A, Siomi H, Siomi MC. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. *PLoS biology* 2005; **3**(7): e235.

43. MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA. In vitro reconstitution of the human RISC-loading complex. *Proceedings of the National Academy of Sciences of the United States of America* 2008; **105**(2): 512-7.

44. Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & development* 2006; **20**(5): 515-24.

45. Foulkes WD, Priest JR, Duchaine TF. DICER1: mutations, microRNAs and mechanisms. *Nature reviews Cancer* 2014; **14**(10): 662-72.

46. Swahari V, Nakamura A, Deshmukh M. The paradox of dicer in cancer. *Molecular & cellular oncology* 2016; **3**(3): e1155006.

47. Shan W, Sun C, Zhou B, et al. Role of Dicer as a prognostic predictor for survival in cancer patients: a systematic review with a meta-analysis. *Oncotarget* 2016; **7**(45): 72672-84.

48. Merritt WM, Lin YG, Han LY, et al. Dicer, Drosha, and outcomes in patients with ovarian cancer. *The New England journal of medicine* 2008; **359**(25): 2641-50.

49. Bian X, Shen Y, Zhang G, et al. Expression of dicer and its related miRNAs in the progression of prostate cancer. *PloS one* 2015; **10**(3): e0120159.

50. Davalos V, Esteller M. Rolling the dice to discover the role of DICER in tumorigenesis. *Cancer cell* 2012; **21**(6): 717-9.

51. Visscher PM, Wray NR, Zhang Q, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. *American journal of human genetics* 2017; **101**(1): 5-22.

52. Thingholm LB, Andersen L, Makalic E, Southey MC, Thomassen M, Hansen LL. Strategies for Integrated Analysis of Genetic, Epigenetic, and Gene Expression Variation in Cancer: Addressing the Challenges. *Frontiers in genetics* 2016; **7**: 2.

53. Wang Z, Jensen MA, Zenklusen JC. A Practical Guide to The Cancer Genome Atlas (TCGA). *Methods in molecular biology* 2016; **1418**: 111-41.

54. Consortium EP. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 2004; **306**(5696): 636-40.

55. Forbes SA, Bhamra G, Bamford S, et al. The Catalogue of Somatic Mutations in Cancer (COSMIC). *Current protocols in human genetics* 2008; **Chapter 10**: Unit 10 1.

56. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* 2013; **41**(Database issue): D991-5.

57. Tanic M, Zajac M, Gomez-Lopez G, Benitez J, Martinez-Delgado B. Integration of BRCA1-mediated miRNA and mRNA profiles reveals microRNA regulation of TRAF2 and NFkappaB pathway. *Breast cancer research and treatment* 2012; **134**(1): 41-51.

58. Kim S, Rhee JK, Yoo HJ, et al. Bioinformatic and metabolomic analysis reveals miR-155 regulates thiamine level in breast cancer. *Cancer letters* 2015; **357**(2): 488-97.

59. Goswami CP, Nakshatri H. PROGgeneV2: enhancements on the existing database. *BMC Cancer* 2014; **14**: 970.

60. Chang S, Sharan SK. Epigenetic control of an oncogenic microRNA, miR-155, by BRCA1. *Oncotarget* 2012; **3**(1): 5-6.

61. Zheng L, Annab LA, Afshari CA, Lee WH, Boyer TG. BRCA1 mediates ligand-independent transcriptional repression of the estrogen receptor. *Proceedings of the National Academy of Sciences of the United States of America* 2001; **98**(17): 9587-92.

62. Fan S, Ma YX, Wang C, et al. Role of direct interaction in BRCA1 inhibition of estrogen receptor activity. *Oncogene* 2001; **20**(1): 77-87.

63. Yarden RI, Brody LC. BRCA1 interacts with components of the histone deacetylase complex. *Proceedings of the National Academy of Sciences of the United States of America* 1999; **96**(9): 4983-8.

64. Gupta S, Kim C, Yel L, Gollapudi S. A role of fas-associated death domain (FADD) in increased apoptosis in aged humans. *Journal of clinical immunology* 2004; **24**(1): 24-9.

65. Chinnaiyan AM, O'Rourke K, Tewari M, Dixit VM. FADD, a novel death domain-containing protein, interacts with the death domain of Fas and initiates apoptosis. *Cell* 1995; **81**(4): 505-12.

66. Boldin MP, Varfolomeev EE, Pancer Z, Mett IL, Camonis JH, Wallach D. A novel protein that interacts with the death domain of Fas/APO1 contains a sequence motif related to the death domain. *The Journal of biological chemistry* 1995; **270**(14): 7795-8.

67. Shi L, Wang S, Zangari M, et al. Over-expression of CKS1B activates both MEK/ERK and JAK/STAT3 signaling pathways and promotes myeloma cell drug-resistance. *Oncotarget* 2010; **1**(1): 22-33.

68. Chen JL, Li J, Kiriluk KJ, et al. Deregulation of a Hox protein regulatory network spanning prostate cancer initiation and progression. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2012; **18**(16): 4291-302.

69. Yan Y, Haas JP, Kim M, Sgagias MK, Cowan KH. BRCA1-induced apoptosis involves inactivation of ERK1/2 activities. *The Journal of biological chemistry* 2002; **277**(36): 33422-30.

70. Benezra M, Chevallier N, Morrison DJ, MacLachlan TK, El-Deiry WS, Licht JD. BRCA1 augments transcription by the NF-kappaB transcription factor by binding to the Rel domain of the p65/RelA subunit. *The Journal of biological chemistry* 2003; **278**(29): 26333-41.

71. Quinn JE, Kennedy RD, Mullan PB, et al. BRCA1 functions as a differential modulator of chemotherapy-induced apoptosis. *Cancer research* 2003; **63**(19): 6221-8.

72. Thangaraju M, Kaufmann SH, Couch FJ. BRCA1 facilitates stress-induced apoptosis in breast and ovarian cancer cell lines. *The Journal of biological chemistry* 2000; **275**(43): 33487-96.

73. Cohen M, Pierredon S, Wuillemin C, Delie F, Petignat P. Acellular fraction of ovarian cancer ascites induce apoptosis by activating JNK and inducing BRCA1, Fas and FasL expression in ovarian cancer cells. *Oncoscience* 2014; **1**(4): 262-71.

74. Reddy RB, Bhat AR, James BL, et al. Meta-Analyses of Microarray Datasets Identifies ANO1 and FADD as Prognostic Markers of Head and Neck Cancer. *PloS one* 2016; **11**(1): e0147409.

75. van Zundert GCP, Rodrigues JPGLM, Trellet M, et al. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *Journal of molecular biology* 2016; **428**(4): 720-5.

76. Tuszynska I, Magnus M, Jonak K, Dawson W, Bujnicki JM. NPDock: a web server for protein-nucleic acid docking. *Nucleic Acids Research* 2015; **43**(W1): W425-W30.

77. DeLano WL. PyMOL molecular viewer: Updates and refinements. *Abstr Pap Am Chem S* 2009; **238**.

78. Du Z, Lee JK, Tjhen R, Stroud RM, James TL. Structural and biochemical insights into the dicing mechanism of mouse Dicer: a conserved lysine is critical for dsRNA cleavage. *Proceedings of the National Academy of Sciences of the United States of America* 2008; **105**(7): 2391-6.

79. Magnani E, Bartling L, Hake S. From Gateway to MultiSite Gateway in one recombination event. *BMC molecular biology* 2006; **7**: 46.

80. Fu C, Wehr DR, Edwards J, Hauge B. Rapid one-step recombinational cloning. *Nucleic Acids Res* 2008; **36**(9): e54.

81. Lapid C. Primer X. 2003. www.bioinformatics.org/primerx/cgi-bin/DNA_1.cgi.

82. Blaha G, Gurel G, Schroeder SJ, Moore PB, Steitz TA. Mutations outside the anisomycin-binding site can make ribosomes drug-resistant. *Journal of molecular biology* 2008; **379**(3): 505-19.

83. DeLano WL. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography* 2002; **4**: 82-92.

84. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res* 2016; **44**(W1): W272-6.

85. Hoffman MM, Khrapov MA, Cox JC, Yao J, Tong L, Ellington AD. AANT: the Amino Acid-Nucleotide Interaction Database. *Nucleic Acids Res* 2004; **32**(Database issue): D174-81.

86. Cheng AC, Chen WW, Fuhrmann CN, Frankel AD. Recognition of nucleic acid bases and base-pairs by hydrogen bonding to amino acid side-chains. *Journal of molecular biology* 2003; **327**(4): 781-96.

87. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 2003; **31**(13): 3381-5.

88. Rother M, Rother K, Puton T, Bujnicki JM. ModeRNA: a tool for comparative modeling of RNA 3D structure. *Nucleic Acids Res* 2011; **39**(10): 4007-22.

89. Jung J, Yeom C, Choi YS, et al. Simultaneous inhibition of multiple oncogenic miRNAs by a multi-potent microRNA sponge. *Oncotarget* 2015; **6**(24): 20370-87.

90. Fernandez-Martinez L, Villegas JA, Santamaria I, et al. Identification of somatic and germ-line DICER1 mutations in pleuropulmonary blastoma, cystic nephroma and rhabdomyosarcoma tumors within a DICER1 syndrome pedigree. *BMC Cancer* 2017; **17**(1): 146.

91. Kato N, Kusumi T, Kamataki A, Tsunoda R, Fukase M, Kurose A. DICER1 hotspot mutations in ovarian Sertoli-Leydig cell tumors: a potential association with androgenic effects. *Human pathology* 2017; **59**: 41-7.

92. Caffrey E, Ingoldsby H, Wall D, et al. Prognostic significance of deregulated dicer expression in breast cancer. *PloS one* 2013; **8**(12): e83724.

93. Cao WM, Gao Y, Yang HJ, et al. Germline mutations of DICER1 in Chinese women with BRCA1/BRCA2-negative familial breast cancer. *Genetics and molecular research : GMR* 2014; **13**(4): 10754-60.

94. Tian Y, Simanshu DK, Ma JB, et al. A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Molecular cell* 2014; **53**(4): 606-16.

95. Koscianska E, Starega-Roslan J, Krzyzosiak WJ. The role of Dicer protein partners in the processing of microRNA precursors. *PloS one* 2011; **6**(12): e28548.

96. Kim YK, Kim B, Kim VN. Re-evaluation of the roles of DROSHA, Export in 5, and DICER in microRNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 2016; **113**(13): E1881-9.

97. Starega-Roslan J, Galka-Marciniak P, Krzyzosiak WJ. Nucleotide sequence of miRNA precursor contributes to cleavage site selection by Dicer. *Nucleic Acids Res* 2015; **43**(22): 10939-51.

98. Nguyen DD, Chang S. Development of Novel Therapeutic Agents by Inhibition of Oncogenic MicroRNAs. *International journal of molecular sciences* 2017; **19**(1).

99. Yamamoto H, Mori M. MicroRNAs as Therapeutic Targets and Colorectal Cancer Therapeutics. *Advances in experimental medicine and biology* 2016; **937**: 239-47.

100. Nana-Sinkam SP, Croce CM. MicroRNAs as therapeutic targets in cancer. *Translational research : the journal of laboratory and clinical medicine* 2011; **157**(4): 216-25.

101. Shah MY, Calin GA. MicroRNAs as therapeutic targets in human cancers. *Wiley interdisciplinary reviews RNA* 2014; **5**(4): 537-48.

102. Wu MK, de Kock L, Conwell LS, et al. Functional characterization of multiple DICER1 mutations in an adolescent. *Endocrine-related cancer* 2016; **23**(2): L1-5.

103. Fareh M, Yeom KH, Haagsma AC, Chauhan S, Heo I, Joo C. TRBP ensures efficient Dicer processing of precursor microRNA in RNA-crowded environments. *Nature communications* 2016; **7**: 13694.

국문초록

유방암에서 BRCA1 의 타겟 유전자 발굴과 Dicer 돌연변이의 암 조절기작 규명

Nguyen Dinh Duc

울산대학교

의과대학

BRCA1 은 DNA 손상 복구, 세포 주기 조절, 신진대사 및 기타 필수 세포 과정에 관여하는 다기능 종양 억제 유전자이다. 이러한 기능 중 많은 부분은 전사 및 후성학적 조절에 의해 유도되거나 연관성이 있다. 그러나 인간 유전체에 대한 BRCA1 의 전사/후성학적 타겟에 대한 연구가 부족한 상황이다. 그러므로 본 연구에서는 유전체/전사체 데이터 분석을 통하여 BRCA1 타겟 유전자들을 발굴 하였다. 먼저, ENCODE 데이터에서 4 종의 세포에서 BRCA1 의 ChIP-seq 을 수행한 결과를 얻어 분석하였다. 프로모터에 BRCA1 이 결합한 유전자들을 확인하기 위하여, 4 종의 모든 세포에서 발견된 BRCA1 ChIP 신호를 가지는 3085 지역을 확인하였으며 유전자 전사 시작 사이트(TSS)로부터의 거리를 계산하였다. 그 결과 BRCA1 결합된 지역의 66.5%는 전사시작 사이트로부터 2Kb 내에 떨어져있었으며, 이것은 유전자의 전사조절에 BRCA1 이 작용 할 수 있음을 시사한다. 두 종류의 GEO 데이터 세트와 유방암 환자의 유전자 발현 프로파일을 포함하는 TCGA 를 이용한 유전자 발현 상관 관계 분석을 통하여 20 개의 유전자를 선택하였다. 선택된

유전자는 0.25~0.54(Spearman 상관 관계) 수준의 유의한 상관 계수를 가졌다. 이 유전자들 중에 3 개의 유전자를 선택하여 ChIP, 실시간 PCR 및 Luciferase 리포터 실험을 한 결과, FADD 를 BRCA1 의 새로운 타겟으로 확인하였다. 본 연구결과를 통해 BRCA1 에 의한 유전체 범위 전사조절을 확인하였으며, BRCA 의 타겟 유전자들은 BRCA1 관련 유방암의 바이오 마커 후보로 활용될 수 있음을 보여준다.

Dicer 는 세포 내 microRNA 의 생합성 시스템에 필수적인 단백질이다. 이 단백질의 돌연변이는 흉막과 폐의 모세포종, 유방암 난소암 등 많은 암에서 발견된다는 것이 최근 많은 논문에 보고되었다. Dicing 공정의 생성물인 mature miRNA 의 발현과 활성의 수준은 암에서 비정상적으로 조절되어 있다. 본 연구에서는 miRNA 생합성과정에 영향을 미칠 수 있는 Dicer C-터미널 지역의 돌연변이를 분석하고자 하였다. Dicer C-터미널 지역의 도메인을 E.coli 에서 발현 및 정제하였으며, 상호작용과 상호동역학 분석을 수행하였다. Luciferase reporter 분석을 통하여 R1851C, A1870S and A1914V 세 종류의 돌연변이가 miRNA 합성에 영향을 미칠 수 있음을 확인하였다. 본 연구결과는 miRNA 합성과정에 Dicer 돌연변이의 활성 억제 효과를 이해하고, 약리학적 조절 인자를 개발하는데 도움을 줄 수 있으리라 기대한다.

Acknowledgment

Foremost, I would like to express my sincere gratitude to my Ph.D. thesis advisor, Prof. Chang Suhwan, for his patience, motivation, enthusiasm, immense knowledge and supporting me during these past 4 years living here. His kindly help and guidance help me can go further and I believe I can work independently my research field in the future. Moreover, this work would not have been possible without the financial support from grants of the LIO laboratory and the AF-1 Scholarship from University of Ulsan, The Graduate School.

Besides my adviser, I am grateful to the members of my Ph.D. committee, Prof. Choi Sun-Cheol, Prof. Lee Jung-Kul, Prof. Kang Min-Ji, and Prof. Kim Yongsub for their helpful advices and constructive suggestions to me during writing this thesis. Also, thanks to Dr. Kang from Dankuk University for his contribute in the bioinformatics analyses.

My great thanks to my lab mates for their helps and for all of fun we have had in the last four years. Many thanks for what I have learn from Dr. Kim, Dr. Choi, Ms. Eun-Ji, Jae-Yun, Min-Ji to make everything run smoothly in my research and our lab works. Thanks to the other students Eun-Song, Euna, Jeongmin, Bruce, Myongji and Dr. Hu for working together and helps.

With a grateful mention to all my friends, AsanersVN group (Trang, Tan, Hang, anh Duong, Mai), nhom Theresa of Vietnam catholic in Seoul community, nhom loi ich (anh Hoang, anh Phong, Thong), KUFC, 07CS in Korea, Tony and other who support, help and spend time together during my Ph.D. student life in Korea.

My deep and sincere gratitude sends to my family for their continuous and unparalleled love, help and support. This journey would not have been possible without them.

My special thank to my beloved Nana, for her love and support.