โปรตีนอโกรนอท4 เป็นตัวกลางนำพาอาร์เอ็นเอสายสั้นในการเกิดเมทิลเลชั่นอีกครั้งในเซลล์มนุษย์

นายปิยพัทธ์ ปิ่นอ่อน

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาชีวเวชศาสตร์ (สหสาขาวิชา)
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2557

ARGONAUTE 4 PROTEINS MEDIATE SMALL-RNA-GUIDED DE NOVO METHYLATION IN HUMAN CELLS

Mr. Piyapat Pin-on

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Biomedical Sciences
(Interdisciplinary Program)
Graduate School
Chulalongkorn University
Academic Year 2014

Thesis Title	ARGONAUTE 4 PROTEINS MEDIATE SMALL-RNA-GUIDED DE NOVO METHYLATION IN HUMAN CELLS
By	Mr. Piyapat Pin-on
Field of Study	Biomedical Sciences
Thesis Advisor	Professor Apiwat Mutirangura, M.D., Ph.D.
Thesis Co-Advisor	Assistant Professor Chatchawit Aporntewan, Ph.D.

Accepted by the Graduate School, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

.....Dean of the Graduate School
(Associate Professor Sunait Chutintaranond, Ph.D.)

THESIS COMMITTEE

.....Chairman
(Professor Prasit Pavasant, D.D.S., Ph.D.)

.....Thesis Advisor
(Professor Apiwat Mutirangura, M.D., Ph.D.)

.....Thesis Co-Advisor
(Assistant Professor Chatchawit Aporntewan, Ph.D.)

.....Examiner
(Professor Nattiya Hirankarn, M.D., Ph.D.)

.....Examiner
(Assistant Professor Nipan Israsena, M.D., Ph.D.)

.....External Examiner
(Siwanon Jirawatnotai, Ph.D.)

ปิยพัทธ์ ปิ่นอ่อน : โปรตีนอโกรนอท4 เป็นตัวกลางนำพาอาร์เอ็นเอสายสั้นในการเกิดเมทิลเลชั่นอีกครั้งในเซลล์มนุษย์ (ARGONAUTE 4 PROTEINS MEDIATE SMALL-RNA-GUIDED DE NOVO METHYLATION IN HUMAN CELLS) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: อภิวัฒน์ มุทิรางกูร, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: ชัชวิทย์ อาภรณ์เทวัญ, 89 หน้า.

โปรตีนอโกรนอทเป็นโปรตีนที่ถูกอนุรักษ์ไว้ในเชิงวิวัฒนาการและเป็นโปรตีนที่พบว่ามีแสดงออกได้ในทุกๆส่วนของยูคาลีโอท์ชั้นสูง โปรตีนอโกรนอทมีหน้าที่สำคัญในหลายขบวนการทางชีววิทยาเช่นการเปลี่ยนแปลงของเซลล์ การหยุดแสดงออกของทรานส์โพซอน หน้าที่ทั่วไปของโปรตีนอโกรนอทคือเป็นโปรตีนที่สำคัญในขบวนการหยุดการแสดงออกของยีนโดยการเหนี่ยวนำอาร์เอ็นเอสายสั้น ในมนุษย์พบว่ามีความหลากหลายของโปรตีนอโกรนอท อย่างไรก็ตามบทบาทที่ชัดเจนของแต่ละโปรตีนอโกรนอทยังไม่ถูกศึกษามากนัก ดังนั้นข้าพเจ้าจึงมุ่งเน้นศึกษาว่าหลังจากลดการแสดงออกออกของโปรตีนอโกรนอท1, 2, 3 และ 4 ในเซลล์ HEK293แล้ว จะส่งผลต่อการแสดงออกของยีนที่มีไลน์-1แทรกอยู่อย่างไร โดยใช้แนวทางการศึกษาแบบชีวสารสนเทศน์ด้วยโปรแกรม Connection Up- or Down- Regulation Expression Analysis of Microarrays X (CU-DREAM-X) นอกจากนี้ข้าพเจ้ายังได้ศึกษาความสัมพันธ์ระหว่างโปรตีนอโกรนอท1, 2, 3 และ4 ต่อระดับเมทิลเลชั่นของไลน์-1 และความสัมพันธ์ระหว่างโปรตีนอโกรนอท4 ต่อระดับเมทิลเลชั่นของยีนที่มีความจำเพาะต่อการจับของโปรตีนอโกรนอท4 ด้วยวิธี Chromatin immunoprecipitation (ChIP), quantitative combined bisulfite restriction analysis of methylation pattern (Q-COBRA-MP) และ Pyrosequencing ผลการศึกษาพบว่าหลังจากลดการแสดงออกออกของโปรตีนอโกรนอท1, 2, 3 และ4 การแสดงออกของยีนที่มีไลน์-1แทรกเป็นป้องกันการแสดงออก, เพิ่มการแสดงออก, ไม่มีการเปลี่ยนแปลง และป้องกันการแสดงออกของยีนตามลำดับ และยังพบว่าโปรตีนอโกรนอทจับกับสภาวะไลน์-1เมทิลเลชั่นที่แตกต่างกันคือ โปรตีนอโกรนอท1และ4จะจับกับไลน์-1ที่มีระดับเมทิลเลชั่นไม่แตกต่างจากระดับเมทิลเลชั่นของจีโนม ขณะที่โปรตีนอโกรนอท2และ3จะจับกับไลน์-1ที่มีระดับเมทิลเลชั่นที่ต่ำกว่าระดับเมทิลเลชั่นของจีโนมอย่างมีนัยสำคัญ ผลการศึกษาระดับเมทิลเลชั่นของยีนที่จำเพาะต่อการจับของโปรตีนอโกรนอท4คือหลังจากที่ลดการแสดงออกของโปรตีนอโกรนอท4ในเซลล์ HEK293 โดยการเติม tetracycline เป็นเวลา 9 วัน ร่วมกับ 5-aza-deoxycytidine เป็นเวลา 2 วัน พบว่าระดับเมทิลเลชั่นของยีนลดลงอย่างมีนัยสำคัญ และเมื่อเลี้ยงเซลล์อย่างต่อเนื่องอีก 3 วัน โดยปราศจาก tetracycline ระดับเมทิลเลชั่นของยีนเพิ่มขึ้นอย่างมีนัยสำคัญ โดยสรุปแล้ววิทยานิพนธ์ฉบับนี้ได้พิสูจน์บทบาททางสภาวะเหนือพันธุกรรมของโปรตีนอโกรนอท1, 2, 3 และ4 ต่อระดับเมทิลเลชั่นของไลน์-1และการควบคุมการแสดงออกของยีนที่มีไลน์-1 แทรกอยู่ และพิสูจน์ว่าโปรตีนอโกรนอท4 เป็นตัวกลางนำพาอาร์เอ็นเอสายสั้นเข้าสู่ขบวนการเกิด de novo methylation ในเซลล์มนุษย์ การศึกษานี้มีความน่าสนใจเพื่อที่จะพิสูจน์ให้มีความเข้าใจมากขึ้นถึงความหลากหลายของหน้าที่ของโปรตีนอโกรนอท1, 2, 3 และ4

สาขาวิชา ชีวเวชศาสตร์ ลายมือชื่อนิสิต .....

ปีการศึกษา 2557 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

5287794620 : MAJOR BIOMEDICAL SCIENCES
KEYWORDS: ARGONAUTE PROTEINS / LINE-1 METHYLATION / DE NOVO METHYLATION

PIYAPAT PIN-ON: ARGONAUTE 4 PROTEINS MEDIATE SMALL-RNA-GUIDED DE NOVO METHYLATION IN HUMAN CELLS. ADVISOR: PROF. APIWAT MUTIRANGURA, M.D., Ph.D., CO-ADVISOR: ASST. PROF. CHATCHAWIT APORNTEWAN, Ph.D., 89 pp.

Argonaute proteins (AGOs) are evolutionarily conserved and ubiquitously expressed in all higher eukaryotes and have important function in several biological processes, such as cell differentiation and transposon silencing. The commonly known functions of AGOs are key in the gene-silencing pathways guided by small RNAs. There are many human AGOs; however, redundancy and the distinctive roles of different AGOs have not been well characterized. Here, I evaluated the expression of genes containing LINE-1 in AGO1-4-knocked down HEK293 cells using bioinformatics approach: Connection Up- or Down- Regulation Expression Analysis of Microarrays X (CU-DREAM-X). Furthermore, I measured the methylation levels of AGO1-4-bound LINE-1s and specific genes using Chromatin immunoprecipitation (ChIP) and quantitative combined bisulfite restriction analysis of methylation pattern (Q-COBRA-MP) and Pyrosequencing, respectively. The results showed that genes containing LINE-1s in AGO1, -2, -3 or -4-knocked down HEK293 cells prevented up-regulation, increased up-regulation, unchanged regulation and prevented up-regulation, respectively. Interestingly, AGO1-4 bound to LINE-1 differently in term of the methylation level. Although the methylation level of AGO1 and -4-bound LINE-1s was not different from the genome, AGO2- and AGO3-bound LINE-1s were hypomethylated. Finally, in case of specific genes, AGO4 bound to genes was reduced after Si-AGO4 HEK293 cell lines were treated with tetracycline for 7 days. Then, methylation levels of AGO4 binding genes were reduced after Si-AGO4 HEK293 cell lines were treated tetracycline together with 5-aza-deoxycytidine for 9 days and were recovered after Si-AGO4 HEK 293 cell lines were cultivated continuously for 3 days without tetracycline and 5-aza-deoxycytidine. Our experiments demonstrate the distinctive epigenetic roles of AGO1-4 in regulating genes containing LINE-1s and evaluate AGO4 mediate small-RNA-guided de novo methylation in human cells. It is interesting to further explore the underlining mechanism of these various functions.

Field of Study: Biomedical Sciences
Academic Year: 2014

Student's Signature .....
Advisor's Signature .....
Co-Advisor's Signature .....

ACKNOWLEDGEMENTS

This research has been supported by the Ratchadapiseksomphot Endowment Fund of Chulalongkorn University (CU-57-001-HR). I have been supported by Chula Dusadee Pipat Fund, Chulalongkorn University.

I would like to express my grateful thanks to my thesis advisor, Professor Apiwat Mutirangura, for his invaluable advice and constant encouragement throughout 6 years ago. His wisdom, knowledge and commitment motivated me to improve my research challenges. In addition, I am heartfelt thanks for my co-advisor, Assistant. Professor Chatchawit Aporntewan for continuous consultation and all of committees who contributed improve my research and led to my thesis's success. I am thankful Dr. Pornrutsami Jintaridh for pyrosequencing technique guidance. I would like to thank all members of Mutirangura lab, Faculty of Medicine, Chulalongkorn University for all the assistance.

My thesis would not be completed without all bottomless helps from my parents, my aunties and my girlfriend that I have always received throughout the period of Ph.D. student. They always encourage with theirs loving and supporting me even in the hardest time of my life. I would like to express my sincere appreciation to all of them who stand by me sides every situation.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (con't)

LIST OF FIGURES (con't)

LIST OF ABBREVIATIONS

Agos	Argonaute proteins
RITS	RNA-induced initiation of transcriptional silencing
PTGS	Post-transcriptional gene silencing PTGS
miRNAs	MicroRNAs
siRNAs	small interfering RNAs
RdDM	RNA-directed DNA methylation
ssRNAs	Single strand RNAs
POLIVa	DNA-Dependent RNA Polymerase IVa
RDR2	RNA-Dependent RNA Polymerase2
DCL3	Dicer-Like3
RNP	Ribonucleoprotein
DRM2	Domains rearranged methyltransferase2
LINE-1	Long interspersed nucleotide element1
CU-DREAM	Connection Up- and Down-Regulation Expression Analysis of Microarrays
ChIP	Chromatin immunoprecipitation
Q-COBRA-MP	Quantitative combined bisulfite restriction analysis of methylation pattern

LIST OF ABBREVIATIONS (con't)

CLIPZ	Cross-linking and immunoprecipitation
ChIP-qPCR	Chromatin immunoprecipitation and quantitative PCR
5-aza-CdR	5-aza-deoxycytidine
CH_3	Methyl group
SAM	S-adenosylmethionine
H3K9	Histone 3 Lysine-9 methylation
RNAPII	RNA polymerase II
OR	Odds ratio
CI	Confidence intervals
Tet+	Tetracycline treatment
Tet-	Without tetracycline treatment
SDS	Sodium dodecyl sulfate
DNMT	DNA methyltransferase
HEN1	Hua-enhancer 1
PolV	DNA dependent RNA polymerase V
DRD1	RNA-directed DNA methylation 1
DMS3	Defective in meristem silencing 3
DRM1	Required for DNA methylation 1

LIST OF ABBREVIATIONS (con't)

CTD	Carboxy-terminal domain
KTF1	Kow domain-containing transcription factor 1
RISC	RNA-induced silencing complex
C	Cytosine
G	Guanine
5-Cm	5-Methylcytosine
RC	Replication complex
MBDs	Methyl-CpG-binding proteins

CHAPTER I
INTRODUCTION

The mechanism of DNA methylation is a major epigenetic process found in the genome of various eukaryote and prokaryote organisms that is associated with regulating various cellular processes including embryonic development, transcription, chromatin structure, genomic stability, repetitive elements, and carcinogenesis. This mechanism is referring to the addition of a methyl group to the 5' position of cytosines which are present in the 5' regulatory regions closely promoter sequences of transcription start sites by DNA methyltransferase protein families. Four DNA methyltransferase, particularly DNMTl, DNMT2, DNMT3a, and DNMT3b have been identified in humans.

DNA methylation had been classified as either maintenance or de novo methylation. Maintenance methylation carried out by a DNA methyltransferase that reproduce DNA methylation patterns specifically on hemimethylated DNA between DNA replication. The mechanism is simplest to copies the parental-strand methylation pattern onto the daughter strand after each round of DNA replication. The major DNA methyltransferase is responsible for establishing and maintenance of methylation pattern as the DNA methyltransferase1 (Figure 1).

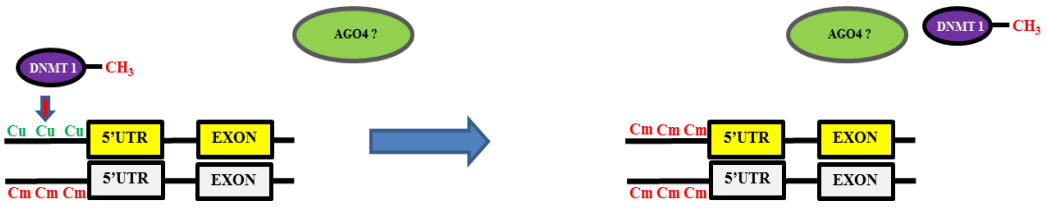


Figure 1. Mechanism of maintenance methylation in DNA replication.

Another mechanism is de novo methylation. In contrast, this mechanism is not necessary to recognizes methylation patterns of the parental DNA strand. There are several recent reports revealed that DNMT3a and3b were able to add methyl group to specific DNA that has never been methylated. It still unclear how to cells determine the locations of de novo DNA methylation in human cells. However, there are evidences suggest that Argonaute 4 has been implicated in RNA-directed DNA methylation (RdDM) in plant.

RdDM was first discovered in tobacco plants infected with pathogen circular RNA molecules. This phenomenon plays important roles as a cellular defence against RNA viruses, and associated with genome integrity, and transposon stability [1]. When the pathogen RNA molecules replicate, homologous DNA sequences becomes methylated [2]. In a following investigation also conducted, the constitutive expression of a dsRNA matching a genomic promoter sequence was shown to result in de novo DNA methylation and transcriptional silencing of the targeted promoter [3].

In brief, DNA dependent RNA polymerase IV (Pol IV) transcripts single-stranded RNA (ssRNA) (Figure 2-1) then copied into double-stranded RNA (dsRNA) by RNA dependent RNA polymerase 2 (RDR2) (Figure 2-2). Dicer-like 3 (DCL3) digests the dsRNA into 24-nucleotide

small interfering RNA (siRNA) duplexes (Figure 2-3) and methylated at 3′ overhang by hua-enhancer 1 (HEN1) (Figure 2-4). The single strand of siRNA binds to Argonaute 4 (AGO4) to form an RNA-induced silencing complex (RISC)–AGO4 complex (Figure 2-5). DNA dependent RNA polymerase V (PolV) transcription is supported by the DDR complex including defective in RNA-directed DNA methylation 1 (DRD1), defective in meristem silencing 3 (DMS3), required for DNA methylation 1 (DRM1) and DMS4 (Figure 2-6). AGO4 binds PolV transcripts by base-pair complementary with the siRNA sequences and is sustained by carboxy-terminal domain (CTD) and kow domain-containing transcription factor 1 (KTF1) (Figure 2-7). IDN2 may stabilize PolV transcript–siRNA pairing (Figure 2-8). The RDM1 protein binds AGO4 and the de novo cytosine methyltransferase, domains rearranged methylatransferase 2 (DRM2), adding methyl group into nearby DNA sequence (Figure 2-9) (Figure 2) [4].

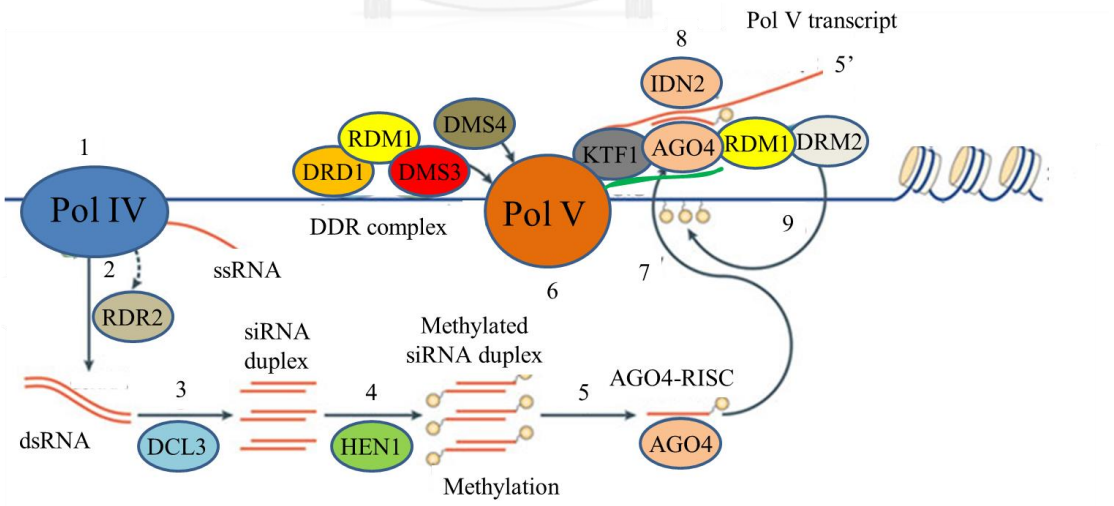


Figure 2. RNA-directed DNA methylation (RdDM) [4].

In plant, it is clear that AGO4 associated with miRNAs or siRNAs, and as an adaptor protein that binds to complementary target sequence

then recruits other proteins especially DNA methyltransferase such as DRM2 is responsible for cytosine methylation. However, it is not certain how the role of AGO4 in DNA methylation in human cell. Recently, we performed comparison between human promoter-methylation dataset (GSE20598) [5] and AGO binding sites, cross-linking and immunoprecipitation data (CLIPZ database) [6]. GSE20598 dataset provided the amount of promoter methylation in HEK-293 cells. CLIPZ database provided AGO binding sequences in the same cell. The result revealed that the numbers of AGO4 binding sites were significant correlation with promoter methylation. Moreover, the result showed that distance of AGO4 sites were associated with peak of methylation level at distance around 21 to 40 bp from the binding sites (Figure 3).

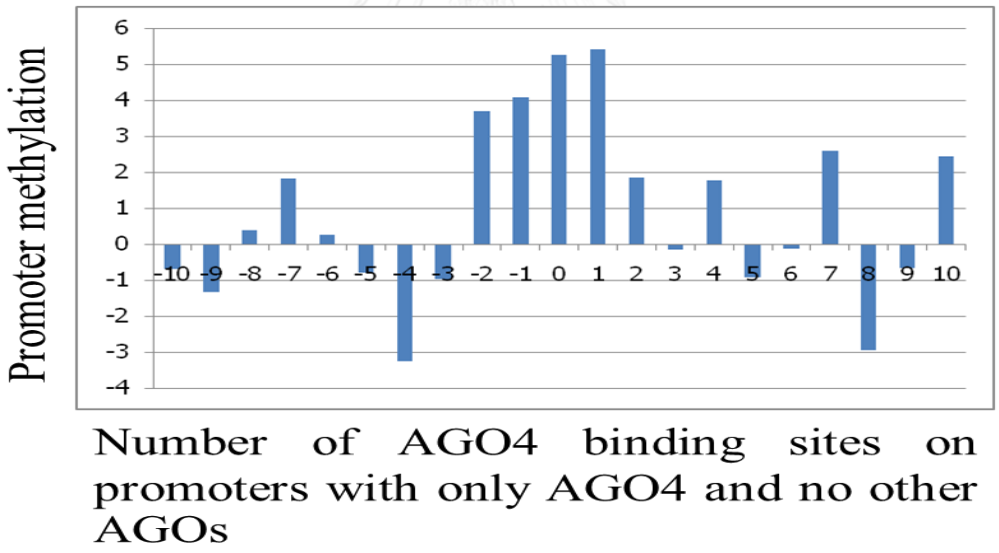


Figure 3. The correlation between the number of AGOs binding sites and promoter methylation (Chatchawit Aporntewan, personal communication).

The results from bioinformatics approach showed that AGO4 protein may play a role on the mechanism of specific promoter

methylation through a process of RdDM. So, I hypothesized that human AGO4 may be associated with promoter methylation similar to RdDM mechanism in plants due to as mediate to recruit DNA methylatransferase for promoter methylation (Figure 4).

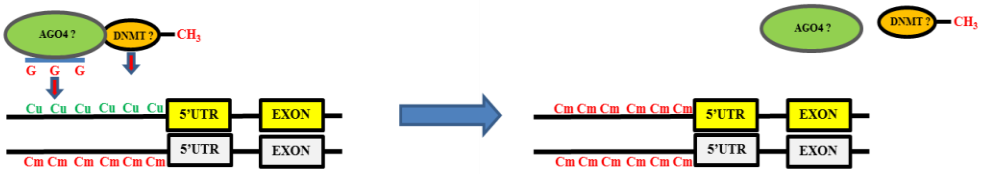


Figure 4. Schematic represents hypothesis of de novo methylation in human cell.

Addition, I evaluated the role of AGO4 in DNA methylation by establishing tetracycline-inducible sh-AGO4 HEK-293 cells. The methylation profiles were detected by pyrosequencing. The ability of AGO4 binding was investigated by chromatin immunoprecipitation and quantitative PCR (ChIP-qPCR). The results showed that down regulated AGO4 expression is strong associated with specific promoter hypomethylation after treated tetracycline and 5-aza-deoxycytidine (5-aza-CdR) on the 9th day.

The members of the Argonaute proteins are key players in gene silencing pathways in various processes such as cell differentiation, embryonic development and transposon silencing [7]. They are ubiquitously expressed in prokaryotic and eukaryotic organisms [8] and were first identified in plants [1]. They mediate guided by small RNAs that are anchored into binding pockets and lead to target mRNAs for gene silencing. In human, there are eight AGO family members, some of

which are investigated. However, even though AGO1-4 are capable of loading miRNA and endonuclease activity.

All AGOs share two main structural features, the PAZ and the PIWI domain. PAZ domain represents their containing the specific binding pocket that anchors the 3’ overhang of both siRNA and miRNA. Therefore, it is recognizes mRNA for mRNA degradation by base-pairing interaction. PIWI domain, it has the characteristic of RNase H thus it is essential for the target cleavage [9].

When referring to gene silencing pathway, the mechanism that is the most well-known and Ago has been also involved is called RNA interference. Ago2 acting as a trans-acting element to down-regulate gene expression by binding different classes of small noncoding RNAs, including microRNAs (miRNAs) and small interfering RNAs (siRNAs) 21–22 nt in length. These short double stranded fragments were digested by an RNase III-like enzyme (Dicer). Then, RNA-induced silencing complex (RISC) which cleaves the target mRNA strand complementary to their bound siRNAs [10-12].

Many researches on recent years have revealed that the role of Agos involving gene silencing pathways such as Ago1 is a component of the RNA-induced initiation of transcriptional silencing (RITS) complex in *S. pombe* [7]. Ago2 and Ago3 are confined to the siRNA directed post-transcriptional gene silencing (PTGS) and required for transposon silencing pathway respectively in *Drosophila* [13-17]. Ago4 has been found in the nucleus and was established well in plants that conducts siRNA-guided DNA methylation [18, 19].

Furthermore, our group reported that AGO2 down-regulates genes containing hypomethylated LINE-1 [20, 21]. In status of hypermethylated intragenic LINE-1, host genes are transcript into mRNAs whereas partial and hypomethylated LINE-1, LINE-1 RNAs will be transcribed and form with AGO2 then host genes reduce expression (Figure 5).

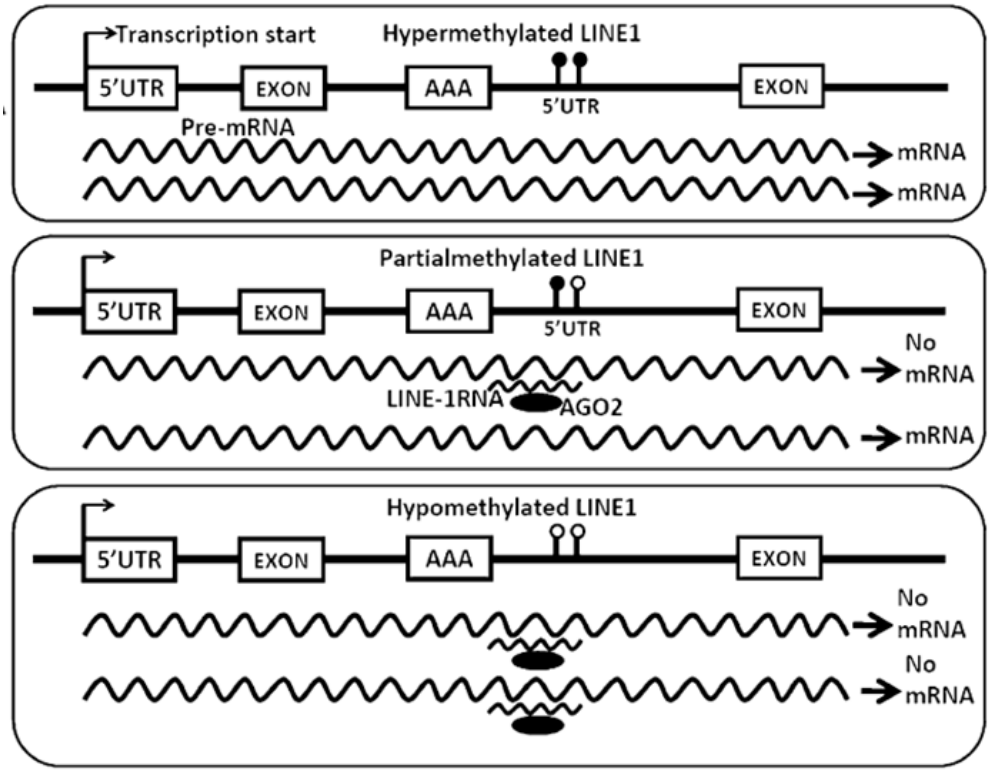


Figure 5. Intragenic hypomethylated LINE-1s repress host gene expression [21].

However, there were no reports how AGO1, 3 and 4 associated with LINE-1 methylation in human cell. Hence, in this study, I evaluated the role of AGO1-4 in affecting the expression of genes containing LINE-1 using a bioinformatics approach. The mRNA levels in the HEK293 cell line knocked down for AGO1-4 were compared with intragenic LINE-1 genes by The Connection Up- and Down-Regulation Expression Analysis

of Microarrays (CU-DREAM) software package (http://pioneer.netserv.chula.ac.th/~achatcha/cudream/) [22]. This software calculates statistical analyses, including Student's t-test and Pearson's chi-squared test, to analyze the gene regulatory functions of intragenic LINE-1. Moreover, to explore the association between AGOs and intragenic LINE-1 methylation, I performed chromatin immunoprecipitation (ChIP) and quantitative combined bisulfite restriction analysis of methylation pattern (Q-COBRA-MP) assays [23].

In conclusion, it is well known that AGOs play a role in controlling gene expression but there is limited information demonstrating whether the functions of human AGO1, 2, 3 and 4 are redundant or different. So, my exploration revealed whether the other AGOs serve as trans-acting elements that mediate gene expression through intragenic LINE-1 methylation. Interestingly, my research explored the novel function of AGO4 protein to regulate specific promoter methylation. This may be applied to use for cancer screening and diagnosis.

Objectives

1. To determine the association between AGO1, 2, 3 and 4 binding and intragenic LINE-1 methylation patterns.
2. To investigate the relationship between the role of AGO1, 2, 3 and 4 proteins and regulation of genes containing LINE-1.
3. To evaluate the role of AGO4 as mediate small-RNA-guided de novo methylation.

Research Questions

1. How do AGO proteins are associated with LINE-1 methylation pattern in human cell?
2. How do AGO proteins are associated with regulation of genes containing LINE-1 in human cell?
3. How does the role of AGO4 proteins as mediate small-RNA-guided de novo methylation in human cell?

Hypothesis

1. The binding of AGO1, 2, 3 and 4 proteins should be affect differently in methylation patterns of intragenic LINE-1.
2. The role of AGO1, 2, 3 and 4 proteins should be affect differently in regulation of genes containing LINE-1 expression.
3. AGO4 proteins play an important role as mediate small-RNA-guided de novo methylation in specific CpG promoters.

Keywords

Argonaute 4 protein (AGO4), de novo methylation, long interspersed element 1 sequences (LINE-1)

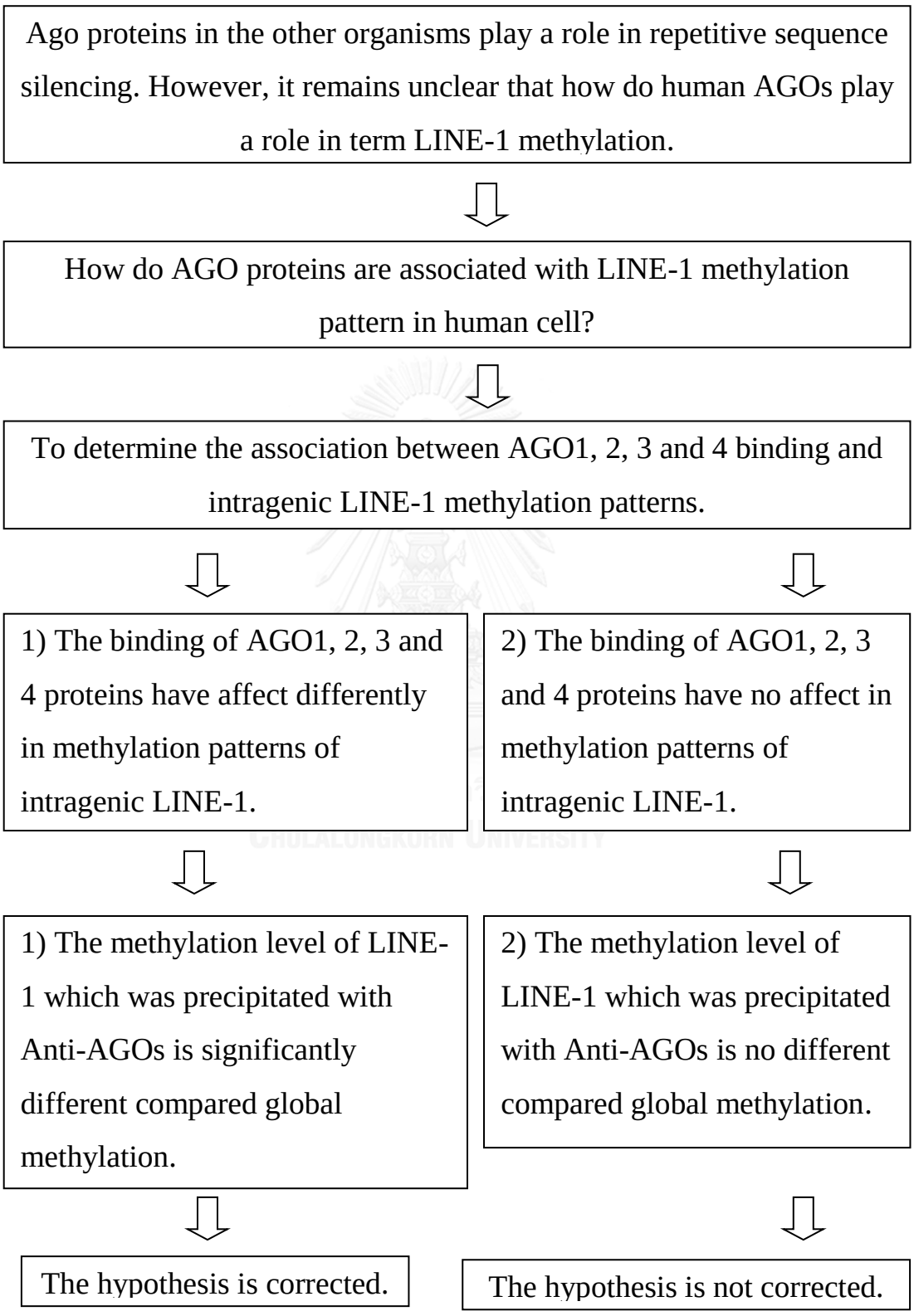
Expected Benefits and Application

It is well known that alteration of LINE-1methylation affect genetic instability and the regulation of gene expression. My research revealed new insights of the relationship between the role of AGO1, 2, 3 and 4 and LINE-1 methylation patterns and theirs genes containing

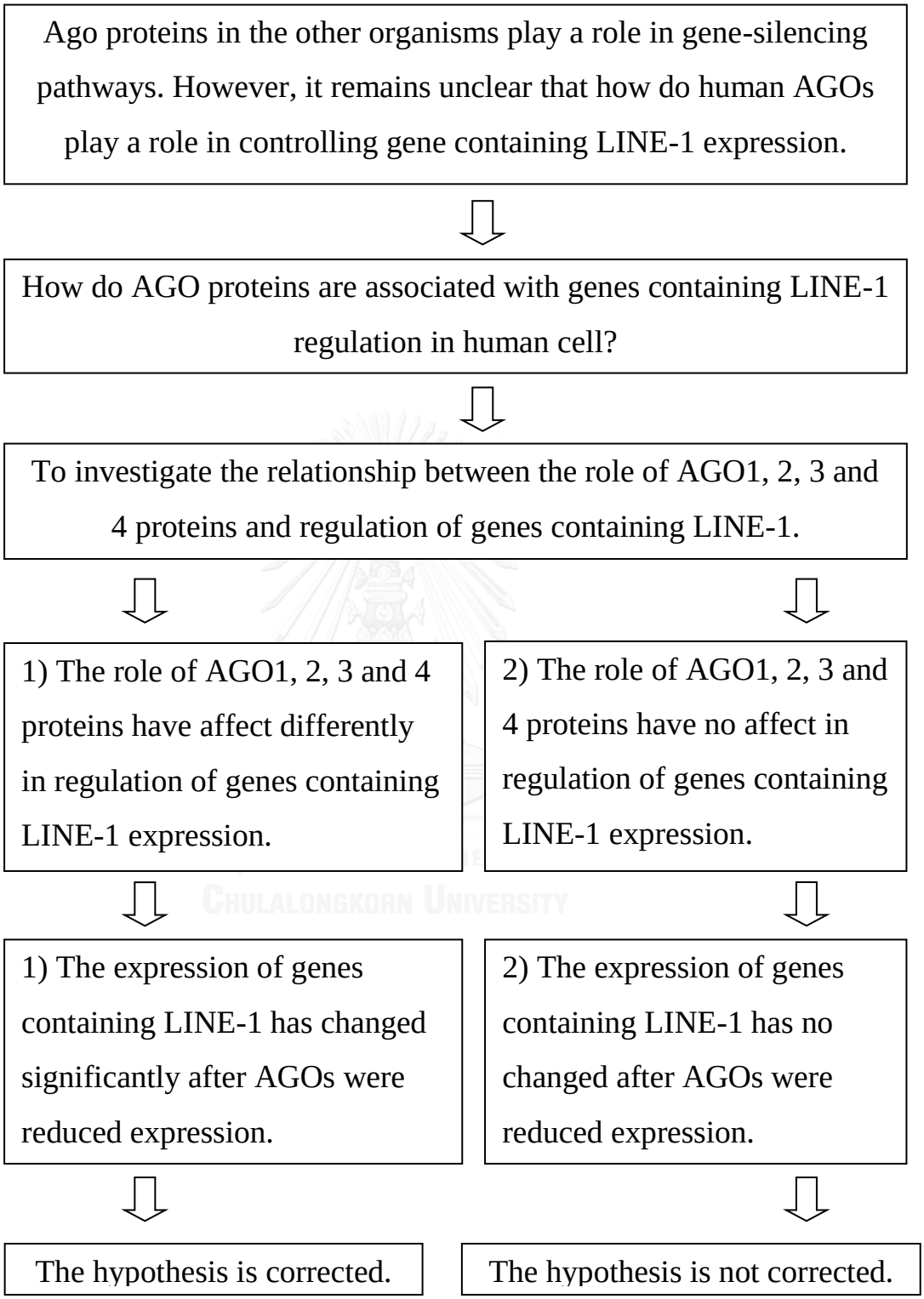
LINE-1 expression. Furthermore, my research also proved that AGO4 mediates small-RNA-guided de novo methylation in human cells. Understanding the distinctive role of AGOs effect on global methylation and specific genes will lead to new strategy for future cancer therapy approach as control methylation patterns of intragenic LINE-1 genes or AGO binding genes.

Conceptual Framework

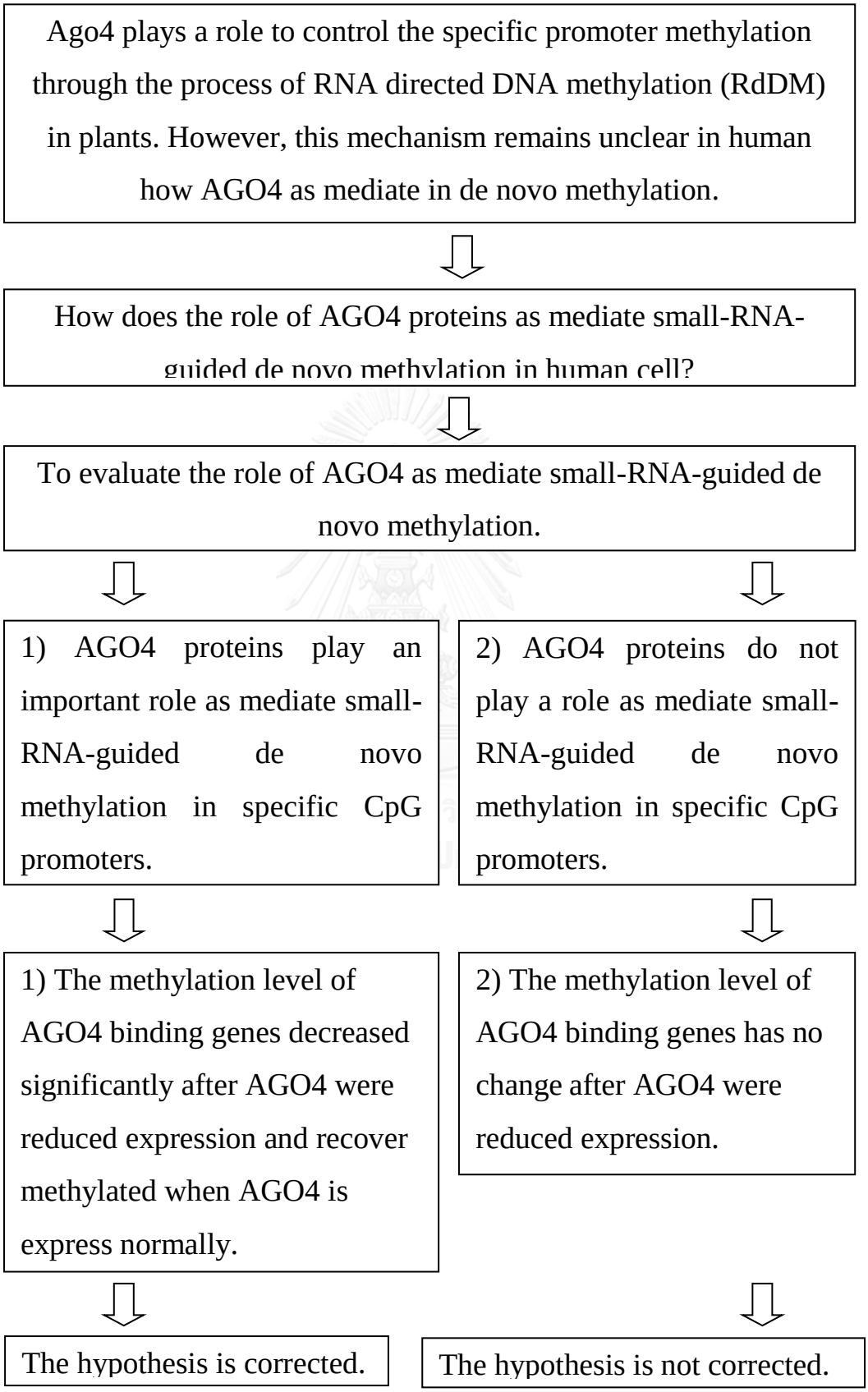
1.



2.



3.


Ago4 plays a role to control the specific promoter methylation through the process of RNA directed DNA methylation (RdDM) in plants. However, this mechanism remains unclear in human how AGO4 as mediate in de novo methylation.
How does the role of AGO4 proteins as mediate small-RNA-guided de novo methylation in human cell?
To evaluate the role of AGO4 as mediate small-RNA-guided de novo methylation.
1) AGO4 proteins play an important role as mediate small-RNA-guided de novo methylation in specific CpG promoters.
2) AGO4 proteins do not play a role as mediate small-RNA-guided de novo methylation in specific CpG promoters.
1) The methylation level of AGO4 binding genes decreased significantly after AGO4 were reduced expression and recover methylated when AGO4 is express normally.
2) The methylation level of AGO4 binding genes has no change after AGO4 were reduced expression.
The hypothesis is corrected.
The hypothesis is not corrected.

CHAPTER II
REVIEWS AND RELATED LITERATURES

DNA Methylation Mechanisms

DNA methylation is a biochemical modification that involves the transfer of a methyl group to the 5-carbon of the cytosine (C) ring that mainly upstream of guanine (G) in the DNA double-helix resulting in 5-methylcytosine (5-Cm). This process does not alter the information coded by the DNA but it participates to the regulation of gene expression. DNA methylation has been demonstrated in the mammalian genome that plays essential for normal development, differentiation, aging, X-chromosome inactivation and involved in control mechanisms of a variety of biological processes including genomic imprinting, suppression of repetitive elements and carcinogenesis[24-28].

The addition of methyl group is carried out by several different DNA methyltransferase enzymes (DNMTs). At least three independent DNA methyltransferases, DNMT1, DNMT3a and DNMT3b, that are involved in DNA methylation and are likely to be central to these processes. DNA methyltransferases divided into two categories based on their bearing to modify cytosine to 5-methylcytosine in unmethylated DNA strand, maintenance and de novo methyltransferases.

Maintenance methylation activity is necessary to preserve DNA methylation. DNMT1 is the major enzyme responsible for establishment and maintenance of DNA methylation patterns. After each round of DNA

replication cycle [29], the new synthesized DNA becomes is unmethylated at CpG sites in which the parent strand retains its methylation pattern. DNMT1, which is localized to the replication complex (RC), restores full methylation by methylating hemimethylated DNA with the help of UHRF1 protein. DNMT1 also recognizes hemimethylated sites and provides a mechanism for inheritance of a methylation profiles through generations, resulting in DNA methylation is stability. Furthermore, DNMT1 also maintains global methylation patterns to plays an essential role in the epigenetic modification that controls gene expression and genome stability.

While DNMT3a and DNMT3b are similar to mediate establishment of de novo DNA methylation that occurs essentially during embryonic development and also during the male and female gametogenesis [30, 31]. It is the process in which unmethylated sites in DNA are methylated resulting in formation of new methylation pattern. This process is no necessary methylated parent strand [32]. DNMT3b in particular is known to be required for de novo methylation in specific genomic regions, as mice or human patients with DNMT3b mutations are deficient in methylation of repetitive DNA sequences and CpG islands on the inactive X chromosome [32, 33].

However, overexpression of DNMT1 is leading to de novo methylation [34, 35]. Recent studies revealed that DNMT3a and DNMT3b are also required for maintenance methylation in repetitive sequences [36]. So the mechanisms of DNA methylation in mammalian cells may involve with another proteins beside DNMT.

Even though there is a little known about de novo methylation how specific DNA sequences are selected for DNA methylation. The observations indicate that de novo DNA methylation machinery is associated with inactive transcription.

In plants, it is clear that understanding how methylation patterns are regulated for de novo DNA methylation through RNA-directed DNA methylation mechanism (RdDM). RdDM is an epigenetic process in which siRNAs or miRNAs direct the cytosine methylation of DNA sequences that are complementary to the siRNAs or miRNAs. This process targets transposons and other types of repetitive sequences, leading to their transcriptional silencing and the maintenance of genome stability. In brief, double strand RNAs are transcribed from a genome sequences by RNA dependent polymerase 2 (RDR2). The dsRNAs are cleaved by dicer-like 3 (DCL3), and are methylated by HEN1. The ssRNAs bind to AGO4 and RNA-induced silencing complex (RISC). Then methyl group is transfer directly nearby the original genomic sequences. The sources of dsRNAs often obtain from RNA viruses or transposons. The target RNA sequence is transcribed by Pol V by assistance of the DDR complex (DRD1, DMS3, and RDM1) and DMS4. AGO4 binds via base pair complimentary with the target RNA sequence and is stabilized by KTF1, which also binds RNA. The de novo methyltransferase, DRM2 is loading methyl group to nearly genomic sequence. Interestingly, DRM2 protein share significant homology to the mammalian methyltransferases, DNMT3. DRM2 is the only enzyme that has been implicated as de novo DNA methyltransferase.

Interestingly, the recent research revealed that Piwi–piRNA complexes might guide DNMT3a/DNMT3b to transposon sequences,

although the mechanism induced DNA methylation remains unknown [37]. The complexes might recruit enzymes to establish epigenetically marks for promoting de novo DNA methylation through the activity of the DNMT3a–DNMT3b complex. It is possible that Argonaute family proteins may associate with de novo DNA methylation due to the piwi domain is a domain that found in a large number of related nucleic acid-binding proteins such as Argonaute proteins.

Interaction between DNMTs and Other Proteins

DNA methylation is not randomly distributed methylation, DNA methyltransferases demonstrate sequence specific in vitro than requiring the CpG dinucleotide recognition sequence. Therefore, recent identifying the protein interaction partners of each of the DNMTs how these interactions may mediate DNMT catalytic activity. For example, 1) HDACs**,** it has been well established that DNA methylation represses transcription partly by recruitment other co-repressor complexes, methyl-CpG-binding proteins (MBDs). Recent evidence, demonstrated that the DNMT1 and DNMT3 directly interact with one or more HDACs [38, 39]. 2) MBDs, MBD2 and MBD3 demonstrated co-localization with DNMT1 at late S-phase DNA replication. Furthermore, the MBD2/MBD3 complex represented binding affinity for both hemimethylated and fully methylated DNA [40, 41]. Furthermore, all MBD family members are involved in transcriptional repression by recruiting chromatin-remodeling complexes to methylated CpG sites [42]. 3) Myc, it interacts with DNMT3a and cooperation with other factors such as Miz-1, it is presumably play a role in targeting DNA methylation

to specific genomic regions. Proteins interact with methyltransferase are summarized in the table 1.

DNA Methyl-transferases	Interactive Proteins	Functions of Interactive Proteins	Putative Roles in Vivo
DNMT1	HDAC1 HDAC2	Histone deacetylase	Modification of chromatin by histone deacetylation, resulting in targeting DNA methylation.
	PML-RAR	Oncogenic transcription factor	DNA-binding and interaction with other transcriptional co-regulators, targeting methylation.
	DMAP1	Co-repressor	Recruiting other repressors, transcriptional repression
DNMT3a	RP58	Transcription factor	Sequence-specific DNA binding, targeting repression, possibly methylation as well.
	PML-RAR	Oncogenic transcription factor	DNA-binding and interaction with other transcriptional co-regulators, targeting methylation.
	HDAC1	Histone deacetylase	Modification of chromatin by histone deacetylation, resulting in targeting DNA methylation.
DNMT3b	HDAC1	Histone deacetylase	Modification of chromatin by histone deacetylation, resulting in targeting DNA methylation.

Table 1. Proteins interacting with methyltransferases [43].

The Role of Methylation in Gene Expression

The mechanism of DNA methylation in both plants and humans is referring to the addition of a methyl group to the 5' position of cytosine by methyltransferase protein families (Figure 6). This is a major epigenetic modification in multicellular organisms and essential for cell differentiation and embryonic development [44]. This mechanism associated with a number of key processes including repression of repetitive elements, and carcinogenesis. In mammals, this modification occurs at CpG sites, which are present in the 5' regulatory regions that can be in close to promoter regions of transcription start sites. The process plays an essential role in the regulation of gene expression by either interfering with the binding of transcription factors and leads to theirs reduced expression [45].

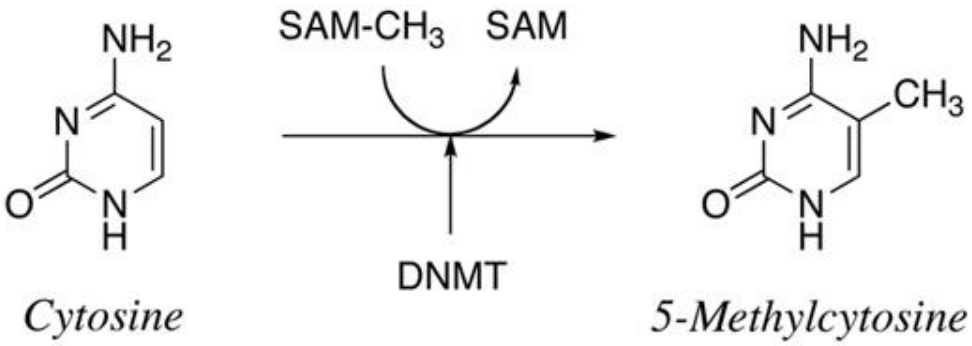


Figure 6. Conversion of cytosine to 5-methylctosine by methyltransferase (DNMT). DNMT catalyses the transfer of a methyl group (CH_3) from S-adenosylmethionine (SAM) to the 5-carbon position of cytosine [46].

There is evidence of this process has been revealed in studies that show that gene promoters methylation varies considerably depending on cell type, with more methylation of promoters association with low or no transcription [47]. In many disease processes, such as cancer, alterations of global and specific genes methylation have been recognized as an important component of cancer development (Table 2).

Genes	**Cancers-Associated Diseases**	**Correlations**
P16	NSCLC	1) Hypermethylation associated with years smoked and smoking duration. 2) In stage 1 adenocarcinoma, an independent risk factor predicting shorter post-surgery survival.
	ATL	Hypermethylation more frequently found in more malignant lymphoma and acute ATL types than the less malignant chronic & smoldering types.
HIC1	ALL,CML, NHL	*HCL1* gene methylation was found in 100% recurrent ALL and blast crisis, but not at initial diagnosis of CML.
P21	ALL	Methylation correlated significantly with decreased disease free and overall survival at 7 and 9 years
MGMT	Gliomas	Methylation correlated with tumor regression and prolonged overall 9 and disease-free survival in patients treated with alkylating agent.
Multiple genes	Cervical cancer	Methylation extent increased from non-dyplastic low grade CIN to high grade & finally to invasive carcinoma.

Table 2. The association between aberrant specific genes methylation and cancers-associated diseases [48].

Tumor suppressor genes are often silenced in cancer cells due to hypermethylation, whereas the genomes of cancer cells have been shown to be hypomethylated overall when compared to normal cells. Also, methylation patterns of specific promoters are related in individual humans.

The Stability of LINE-1 Methylation and Altered Gene Expression

LINE-1 are the largest element of abundant retrotransposon sequences and widely distributed approximately 15% in human genome [49, 50] so methylation patterns of LINE-1 are representative of global methylation patterns. Previously, LINE-1 transcripts were reported to be express only in germ line [51-54]. Full length LINE-1s are retrotransposon ability of causing insertion mutations and genome rearrangements. Recently, LINE-1 transcripts were detected in a variety of somatic tissues [55-57]. It was revealed that the LINE-1 expression play an important role in cell proliferation and differentiation, retaining gene expression and genomic stability. Some research explored that depletion of LINE-1 expression resulted in reduced of cancer cell proliferation [58]. Furthermore, the alteration of LINE-1 methylation was determined that associate with various human cancers (Table 3).

Alteration	Cancers	Remarks
Hypomethylation	Gastric cancer	Correlates with overall hypomethylation
Hypomethylation	Colon cancer	Begins in preneoplastic mucosa
Hypomethylation	Prostate cancer	Increases with stage and metastasis
Hypomethylation	Various	Cell lines
Hypomethylation and expression	Bladder cancer	Expression stronger in teratocarcinomas
Hypomethylation and expression	Liver carcinoma	Hypomethylation, but not expression cancer-specific
No hypomethylation	Renal carcinoma	Tissues only

Table 3. The association between altered LINE-1 methylation and various human cancers [59].

Moreover, It was reported that more than thousands genes containing intragenic LINE-1. These sequences are conserved and act as cis-regulatory elements and modulate gene expression within of their host genes. There are researches explored that intragenic LINE-1 methylation may be concerned with several biological processes such as DNA damage and repair, embryogenesis, cellular response to external stimuli and hormonal responses [60].

In normally, intragenic LINE-1sequences are methylated and silenced in human normal tissues [56, 57, 61]. While, hypomethylation of intragenic LINE-1 is thought to be linked to pathological processes including tumorigenesis, abnormal placental function, birth defects, aging and other chronic diseases [62, 63] and lead to their reactivation by

promoting genomic instability. It is interpreted that when intragenic LINE-1 methylation is decreased, the intragenic LINE-1 transcription is increased resulting LINE-1 mRNAs form RISC

The Roles of Argonaute Proteins in Regulation of Gene Expression

Argonaute proteins are a highly conserved and extensively expressed in many organisms, including plants, animals and humans (Table 4) [8]. The distinctive role of Argonaute protein families are key players in gene silencing pathways in various processes and species (Table 5). There are many reports that human AGOs are involved in organ growth and development [64], and recent studies have clarified its role as a potential factor related to oncogenesis in colonic cancer, breast cancer and prostate cancer [65-67].

Species	Number of genes
Homo sapiens	8
Rattus norvegicus	8
Mus musculus	8
Drosophila melanogaster	5
Caenorhabditis elegans	27
Arabidopsis thaliana	10
Schizosaccharomyces pombe	1
Neurospora crassa	2

Table 4. Argonaute proteins are highly conserved between species and many organisms encode multiple members of the family[68].

Argonaute proteins	Molecular functions
Schizosaccharomyces pombe	
Ago1	Heterochromatin silencing, transcriptional gene silencing, post-transcriptional gene silencing
Arabidopsis thaliana	
AGO1	miRNA-mediated gene silencing
AGO4	Heterochromatin silencing
AGO6	Heterochromatin silencing
AGO7	Leaf development
Drosophila melanogaster	
AGO1	miRNA-mediated gene silencing
AGO2	RNAi
AGO3	piRNA, transposon silencing
Murine/Human	
AGO1	Heterochromatin silencing
AGO2	RNAi, miRNA-mediated gene silencing, heterochromatin silencing

Table 5. Functions of Argonaute proteins in different species [69].

Furthermore, several works have revealed that Ago1 is associated with pre-transcriptional gene silencing. First, Ago1 in *S. pombe* is used to direct Histone 3 Lysine-9 methylation (H3K9) of local transposon sequences, resulting in nucleosome compaction and transposon silencing; RNA polymerase II (RNAPII) activity is thus restricted at the targeted locus, followed by heterochromatin formation [7]. Second, *Drosophila* Ago1 also has a role in heterochromatin formation: the Ago1 mutant disrupts pre-transcriptional gene silencing of transgene tandem repeats, with the loss of H3mK9 and silencing of heterochromatin [70]. Third, in *A. thaliana*, Ago1 may be the slicer component of RISC, acting as RNA

slicers that use the sequence of associated siRNAs to guide the cleavage of homologous RNAs in RNA-induced silencing complexes and in the maintenance of chromatin modifications at some loci. Furthermore, Ago1 knockdown resulted in developmental abnormalities [71] and implicated in the silencing activity of promoter-targeted small RNAs. In *Drosophila*, the function of Ago2 is confined to the siRNA directed post-transcriptional gene silencing (PTGS) pathway. One known mechanism is RNA interference, with Ago2 acting as a trans-acting element to down-regulate gene expression by binding different classes of small noncoding RNAs, including microRNAs (miRNAs) and small interfering RNAs (siRNAs), which guide proteins to their specific targets via sequence complementarity [10-12]. Another mechanism is found in plants and is termed RNA-directed DNA methylation (RdDM). Ago4 is associated with siRNA, and the ribonucleoprotein (RNP) complex is loaded onto chromatin via complementarity. In gene silencing, the complex guides methyltransferases to establish sequence-specific de novo DNA methylation [72].

Recently, we reported that AGO2 down-regulates genes containing hypomethylated LINE-1 [20]. Moreover, our follow-up report demonstrated that intragenic LINE-1 sequences are conserved and regulate several biological processes, including DNA damage and repair, inflammation, immune function and cell differentiation. We screened for genes that are involved in intragenic LINE-1 regulation networks using a bioinformatics approach and found that intragenic LINE-1 acts as a cis-regulatory element within genes to modulate host gene expression. Furthermore, AGOs are also trans-acting elements and coordinate with LINE-1 to regulate gene expression [60]. For example, genes containing

LINE-1 are up-regulated in sh-AGO2 bladder and gastric cancer cells. These results demonstrate that AGO2 plays a role in the control of gene expression through intragenic LINE-1.

The distinctive role of AGO1-4 in the control gene expression through upstream mononucleotide A-repeats has been reported [73]. We found that all members of the AGO family preferentially bind to A-repeats. However, AGO1-4 knockdown produced different patterns. AGO1 was up-regulated genes containing upstream A-repeats, whereas AGO2 and AGO3 were significantly associated with down-regulated genes containing upstream A-repeats. AGO4 did not significantly regulate gene expression. However, it is unclear whether AGOs are associated with intragenic LINE-1 methylation and whether they lead to gene up- and/or down-regulation.

The Association between Human Argonaute4 Proteins and Promoter Methylation

There have been revealed that the ability of Ago4 to control the specific promoter methylation through the process of RNA directed DNA methylation (RdDM) in plants [72]. This mechanism is a small-RNA-guided epigenetic process, adding a methyl group to precise DNA location. It is well known that Ago4 concerns siRNA-guided DNA methylation [18, 19]. It has been found that Ago4 localizes to distinct the nucleolus of *A. thaliana*. and supported that a various DNA methylation complex including Ago4 and siRNAs [74, 75]. In human, although siRNA or shRNA were proved to be able to generate DNA methylation,

this mechanism has not been proven in humans whether AGO4-associated siRNA sequence with directly DNA or transcripts.

Recently, we evaluated association between promoter-methylation dataset (GSE20598) [5] and AGO binding sites (CLIPZ database) [6]. GSE20598 dataset comprised the amount of global promoter methylation profiles in HEK-293 cells. CLIPZ database comprised AGO binding sequences in the same cell. Among all Ago proteins, only AGO4 shows the significant correlation with global promoter methylation summarized from multiple probes (Pearson correlation coefficients = 0.17 and p-value = 1.48E-03). The increasing number of AGO4 binding sites on a promoter is proportional to the amount of methylation on the promoter. Moreover, the binding position of AGO4 is correlated with the methylation of a single probe. The proximity between a binding site and a methylation site (a probe) determines the amount of methylation at the probe. However, the proximity is not correlated with methylation in a linear fashion. Our analysis of genome-wide data suggests that AGO4 may cause a peak of methylation level at distance around 21 to 40 bp from the binding sites (Figure 7).

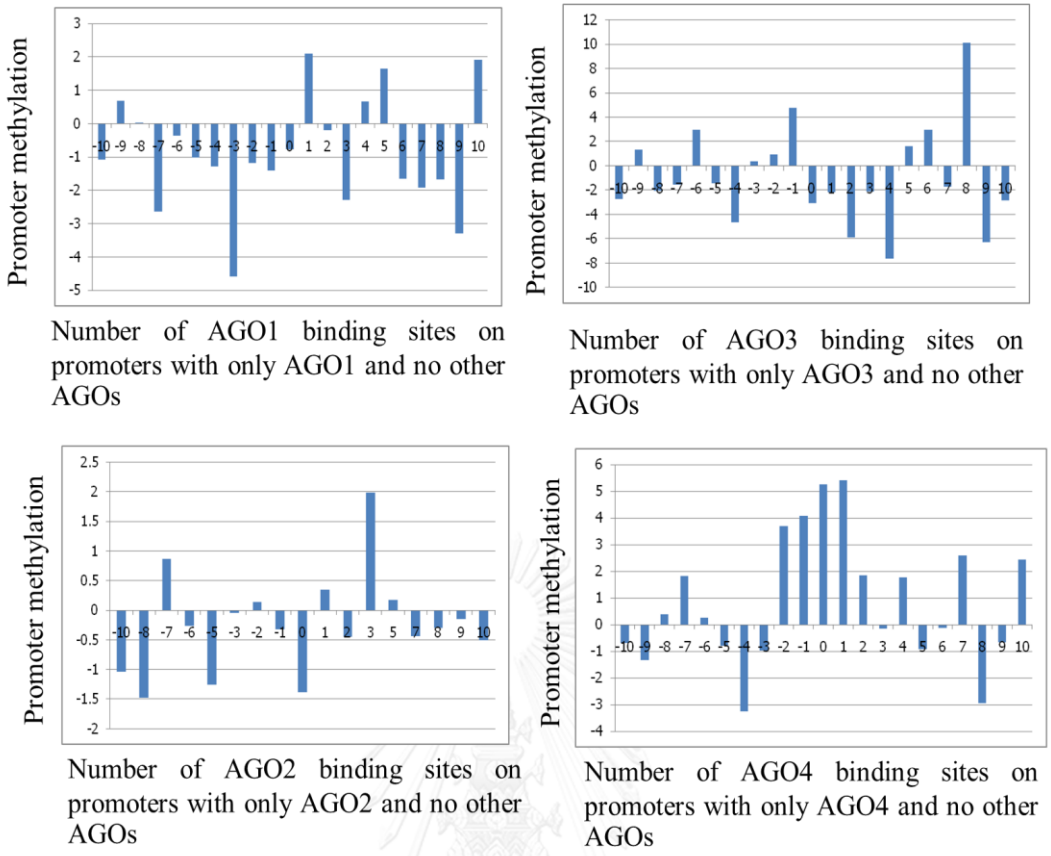


Figure 7. The correlation between the number of AGOs binding sites and promoter methylation (Chatchawit Aporntewan, personal communication).

.

CHAPTER III
MATERIALS AND METHODS

Data Mining for Promoter Selection and Primer Design

Only promoters bound with AGO4 proteins is considered in analysis in order to avoid any possible interactions between AGO proteins. The boundary of a promoter was defined as the region between the first and the fifteenth (the last) probes of each promoter. From previous study to determine the association between AGO protein site (CLIPZ database) and promoter methylation (GSE20598 dataset) in the whole genome of HEK293 cells by bioinformatics approach, we obtain probe ID, methylation level, position on the chromosome and AGO binding sequences.

To collect candidate promoter genes, probe ID are compared with genome_mapping database for providing bound by AGO4. Then, map ID are compared with mapped_sequences database for providing AGO4 binding sequences. These data are analyzed to know the candidate promoter sequences using Ensembl database. We downloaded the following database files from http://test.mirz.unibas.ch/smirnaWeb/geneBio/smiRNA/temp/10544043421949953483/samples (Figure 8).

We performed primer design for pyrosequencing using PyroMask Assay Design software. Following step1: Copy and paste the unconverted bisulfite sequence that we are interested to design primers with in the “Sequence Editor”. Step 2: Highlighting the targeted region in blue and right click to “Set Target Region”. Step 3: Click on the “Play button”.

The software will generate 100 primer set with a quality score given to indicate the quality of the design indirectly to determine the success rate for PCR. Choose the primer set with the highest score. Step 4: In order to view the report or summary of the primer set, click on the “report button” (Figure 9).

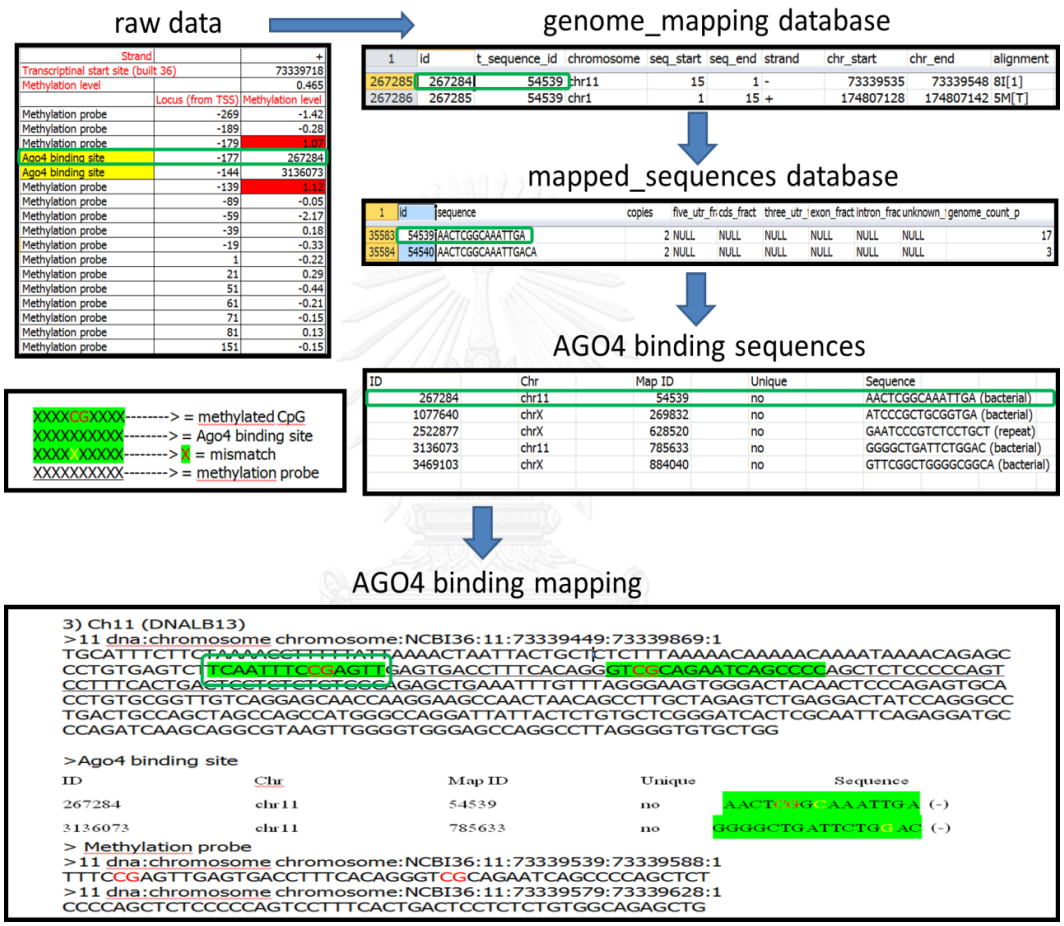


Figure 8. Strategy of AGO4 binding mapping.

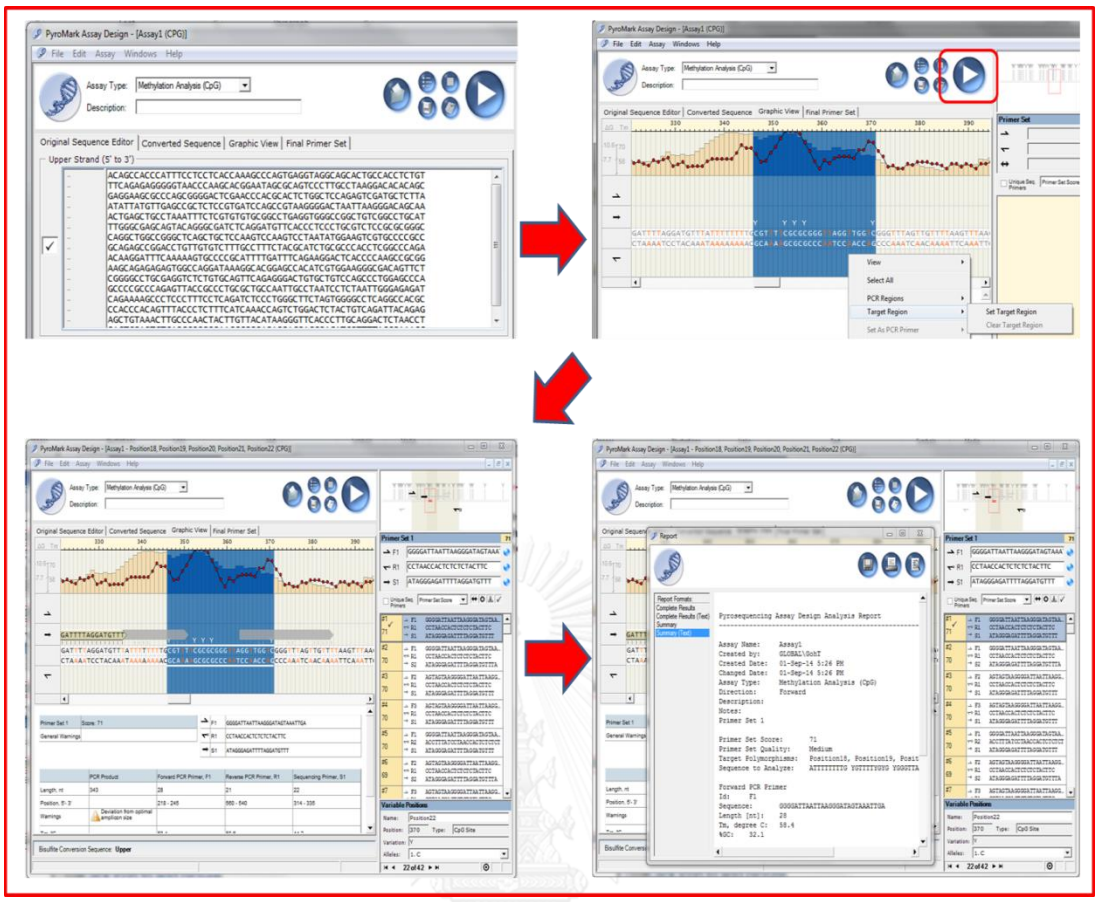


Figure 9. Illustration represents primer design for pyrosequencing using PyroMask Assay Design software.

Si-AGO4 HEK293 Construction

To construct Si-AGO4 HEK293 cell line, we perform lentivirus vector and TetR system according to manufacturer's instruction for coding double strand sh-AGO4. The target sequence 5'-GGCCAGAACTAATAGCAATT-3' were cloned into pENTR™/H1/TO (Invitrogen) and subcloned into pLenti4/BLOCK-iT™-DEST Gateway® Vector (Invitrogen) using LR Clonase® II (Invitrogen). Lentivirus expressed sh-AGO4 using *TetR* as activator for transcription. Briefly, HEK293FT cells were co-transfected with 5 µg of plenti4/H1/TO/shAgo4

and pLenti6/TR using FugeneHD transfection reagent (Roche, Indianapolis, IN). After transfection for 48 hours, the media were collected and filtered through a 0.45 μm pore-size filter. Virus-containing supernatants were centrifuged at 25,000 rpm for 90 minutes. Viral pellets were resuspended with Opti-MEM (Invitrogen).

HEK293 cells were transduced with lentivirus expressed sh-Ago4 and *TetR* supplemented with 6 μg/mL polybrene. After transduction for 48 hours, cells were dissociated and seeded in 10 cm^2 dish at 10% confluence and were cultured under 100 μg/mL Zeocin™ and 10 μg/mL Blasticidin selection for 3 weeks. Zeocin™- and Blasticidin-resistant clonal were isolated and tested for sh-AGO4 activity in a tetracycline-regulated manner using real-time PCR.

Cell Cultures and Treatments

Si-AGO4 HEK293 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 2mM L-glutamine, 10% (v/v) heat-inactivated fetal bovine serum, 10 mg/ml antibiotic/antimycotic, 500 μg/ml zeocin at 37^{o}C and 5% (v/v) CO_2. To knock down AGO4 protein expression, Si-AGO4 HEK293 cells were treated with 2 μg/ml tetracycline for 9 days. Si-AGO4 HEK293 cells (no treated tetracycline, Tet off) were also harvested on 7th, 9th and 12th day. Media and tetracycline were changed every 3 days later. All of them were added with 5-aza-CdR on 7th to 9th day and changed every day (Figure 10).

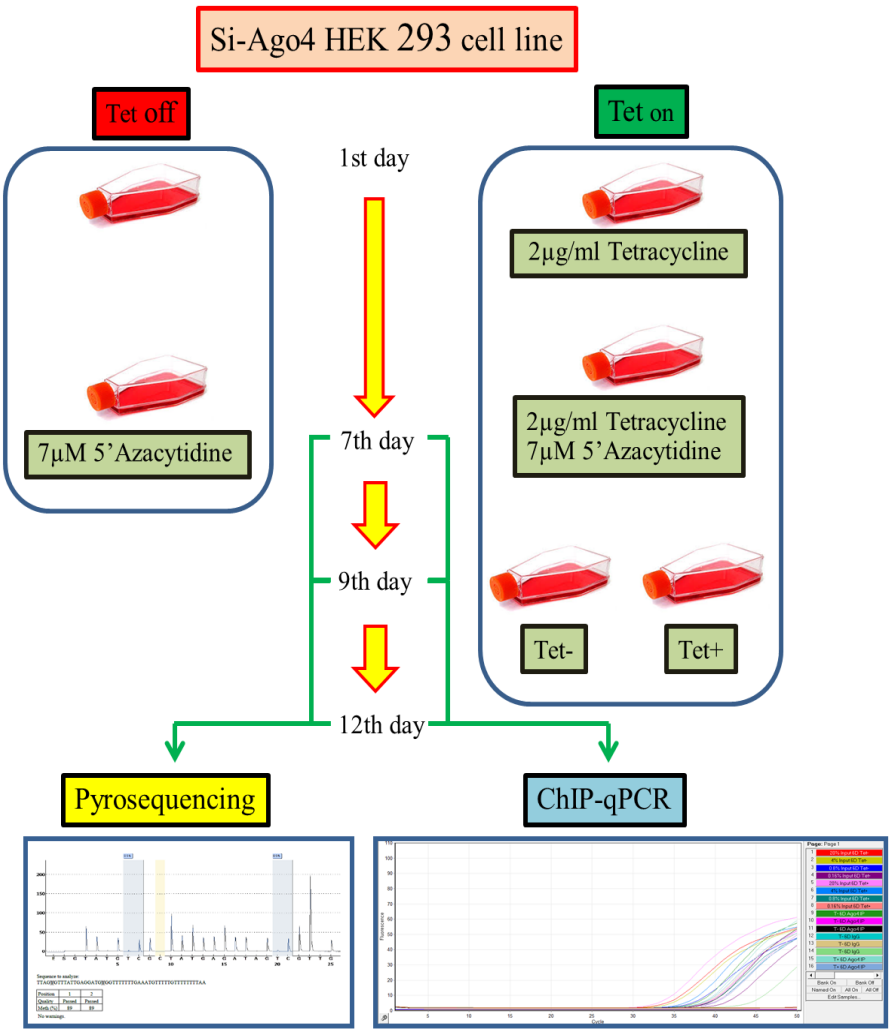


Figure 10. Methedology of cell cultures and treatments**.**

Quantitative Real-time PCR

AGO4 mRNA expression was determined whether the transfection of sh-RNA against the mRNA in Si-AGO4 HEK293 cells by quantitative real-time PCR. Total RNA was extracted from cell lines using the Trizol reagent (Invitrogen) according to the manufacturer's instructions. The RNA was treated with RNase-free DNaseI (Fermentas) to remove contaminating genomic DNA and with RiboLockTM Ribonuclease

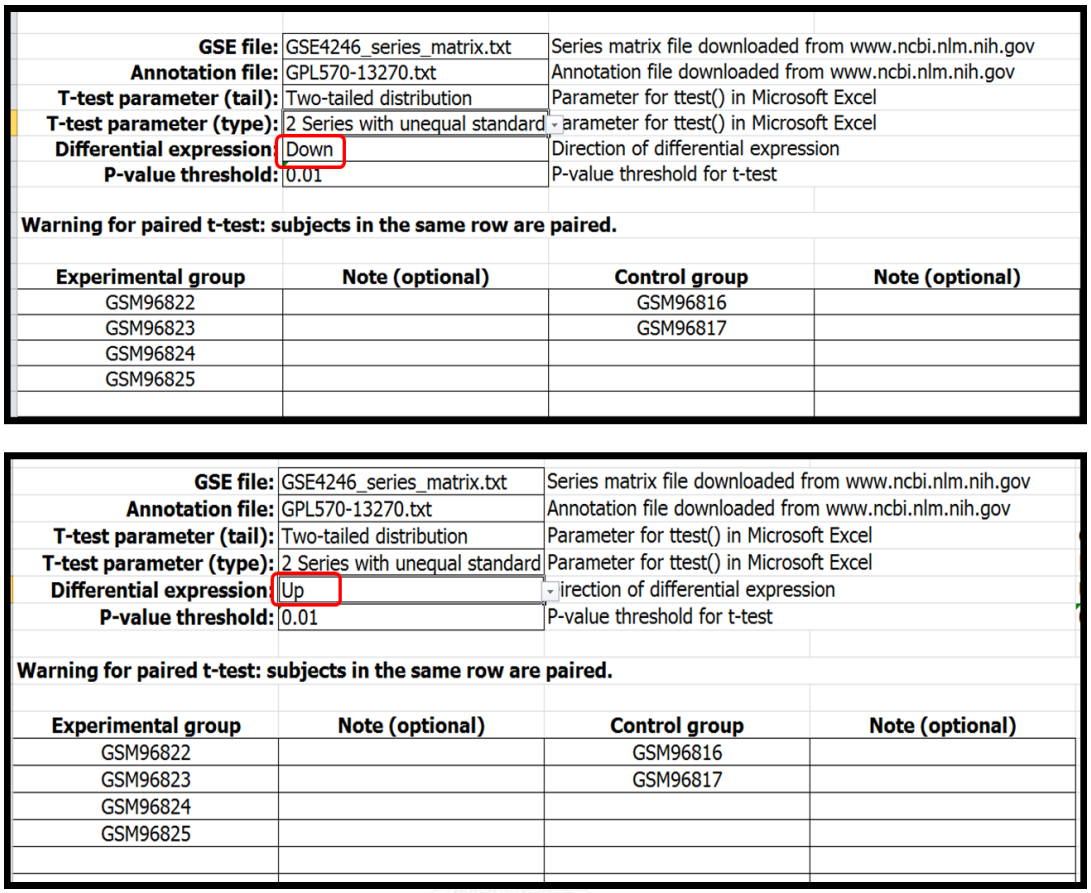
Inhibitor (Fermentas) to prevent degradation. To synthesize cDNA, 5 mg DNA-free RNA was dissolved in 12 ml of DEPC-treated water containing 0.5 mg oligo(dT)18 primer (Fermentas). The RNA was incubated for 5 min at 70^{o}C and chilled on ice for 5 min. Each sample was then incubated with 200U RevertAid^{TM} M-MuLV Reverse Transcriptase (Fermentas), 20 U Ribolock^{TM} Ribonuclease Inhibitor (Fermentas) and 20 mM dNTPs for 1 hr at 42^{o}C, followed by 10 min at 70^{o}C and subsequent chilling on ice.

The reaction mixtures contained 0.1 μM cDNA and were performed using Roter Gene machine (NumbleGEN, IN, USA) and QuantiTect SYBR Green I reagents (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The conditions used were as follows: 95^{o}C for 5 sec, followed by 40 cycles of 60^{o}C for 10 sec. All reactions were run in triplicate. The forward and reverse primers are: Ago4 exon16-17_F: CAG GAA TTC AGG GAA CCA GCC G and Ago4 exon 17-18_R: CTG CCT TCC GCA CTG TCA TGA TC. The relative expression of AGO4 was normalized with GAPDH.

Connection Up- or Down- Regulation Expression Analysis of Microarrays X (CU-DREAM-X)

To evaluate whether AGOs can control theirs host gene expression through intragenic LINE-1, we collected mRNAs of GSE 4246 from the GEO database [76] and prepared templates of the microarray, series matrix file and platform (Figure 11). Each gene was compared to the means of the experimental and control groups and classified as up- or down-regulated and not up- or not down-regulated, depending on the

statistical significance of Student's t-test. Subsequently, the distributions of up- or down-regulated genes were evaluated to whether the distributions were dependent on containing an intragenic LINE-1 using Pearson's chi-squared test. A list of genes containing LINE-1s was previously reported [60]. Genes were grouped into four groups, A through D. The significantly up- or down-regulated genes containing intragenic LINE-1 were included in group A. The significant not up- or not down-regulated genes containing intragenic LINE-1 were included in group B. The significant up- or down-regulated genes without intragenic LINE-1 were included in group C. The remaining genes (non-significant genes without LINE-1) were included in group D. The *P*-values of the odds ratio (ORs), *P*-values and lower and upper 95% confidence intervals (CI) of the genes in groups A through D were displayed in an MS Excel format. All of the statistical analyses were performed using extensions in the CU-DREAM software (http://pioneer.netserv.chula.ac.th/~achatcha/cu-dream/) (Table 7).



GSE file:	GSE4246_series_matrix.txt	Series matrix file downloaded from www.ncbi.nlm.nih.gov
Annotation file:	GPL570-13270.txt	Annotation file downloaded from www.ncbi.nlm.nih.gov
T-test parameter (tail):	Two-tailed distribution	Parameter for ttest() in Microsoft Excel
T-test parameter (type):	2 Series with unequal standard	arameter for ttest() in Microsoft Excel
Differential expression:	Down	Direction of differential expression
P-value threshold:	0.01	P-value threshold for t-test

Warning for paired t-test: subjects in the same row are paired.

Experimental group	Note (optional)	Control group	Note (optional)
GSM96822		GSM96816	
GSM96823		GSM96817	
GSM96824			
GSM96825			

GSE file:	GSE4246_series_matrix.txt	Series matrix file downloaded from www.ncbi.nlm.nih.gov
Annotation file:	GPL570-13270.txt	Annotation file downloaded from www.ncbi.nlm.nih.gov
T-test parameter (tail):	Two-tailed distribution	Parameter for ttest() in Microsoft Excel
T-test parameter (type):	2 Series with unequal standard	Parameter for ttest() in Microsoft Excel
Differential expression:	Up	irection of differential expression
P-value threshold:	0.01	P-value threshold for t-test

Warning for paired t-test: subjects in the same row are paired.

Experimental group	Note (optional)	Control group	Note (optional)
GSM96822		GSM96816	
GSM96823		GSM96817	
GSM96824			
GSM96825			

Figure 11. Template preparation for CU-DREAM-X analysis.

	Up- or down- regulated genes of experiment	Not Up- or not down-regulated genes of experiment
Genes containing LINE-1	Number of genes in 1st group (A)	Number of genes in 2nd group (B)
Genes without LINE-1	Number of genes in 3rd group (C)	Number of genes in 4th group (D)

Table 6. The principle of CU-DREAM-X. The table represents the intersection between the expression of each gene and genes containing intragenic LINE-1.

Chromatin Immunoprecipitation (ChIP)

A ChIP analysis was performed according to a previously published protocol [77] . Briefly, 1 x 10^6 HEK293 cells in a 75-cm^2 flask were grown to 80% confluence. The cells were harvested, and formaldehyde was added at a final concentration of 1% directly to the cell culture medium. Fixation was performed at 37°C for 10 min and stopped by the addition of glycine to a final concentration of 0.125 M. As much medium as possible was removed, and the cells were washed twice with ice-cold PBS containing protease inhibitors; the cells were scraped into a conical tube. The cells were collected by centrifugation and rinsed in cold phosphate-buffered saline. The cell pellets were resuspended in lysed buffer (0.5% NP40, 85 mM KCl, 5 mM PIPES, pH 8.0), incubated on ice for 20 min and homogenized. The nuclei were collected by microcentrifugation and then resuspended in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.1) and incubated on ice for 10 min. The samples were sonicated on ice with an Ultrasonics sonicator at setting 10 amplitude for 20 sec to achieve an average length of approximately 200-1000 bp; the samples were then microcentrifuged. The sonicated cell pellet was resuspended 10-fold in ChIP Dilution Buffer with protease inhibitors. The fragment solution was precleared with the addition of protein A-positive cells for 15 min at 4°C. The precleared chromatin was incubated with 1 µg of affinity-purified goat monoclonal antibody or no antibody and rotated at 4°C for approximately 12 to 16 h. The antibodies used included EIF2C1-4 (anti-AGO1-4) or a control non-immunized goat antibody. Protein A agarose (40 µl) was added for 2 h at 4ºC with rotation to collect the antibody/protein complex, and the complexes were washed and eluted. The pellet was gently

centrifugation (1084g at 4ºC, 1 min), and the protein A agarose/ antibody/ protein complex was washed for 3-5 minutes on a rotating platform with each of the following buffers in order: a) low-salt immune complex wash buffer (1 X 1 ml); b) high-salt immune complex wash buffer (1 X 1 ml) and c) LiCl immune complex wash buffer (1 X 1 ml).The complex was eluted with elution buffer and shaken for 15 min. The cross-links were reversed by the addition of NaCl to a final concentration of 200 mM, and RNA was removed by the addition of 10 µg of RNaseA per sample, followed by incubation at 65°C for 4 h. The samples were then precipitated at -20°C overnight by the addition of 2.5 volumes of ethanol and then pelleted by microcentrifugation. The samples were resuspended in 100 µl of Tris-EDTA (pH 7.5), 25 µl of proteinase K buffer (1.25% sodium dodecyl sulfate, 50 mM Tris (pH 7.5), 25 mM EDTA) and 1.5 µl of proteinase K (Sigma) and incubated at 45°C for 2 h. The samples were extracted with phenol-chloroform-isoamyl alcohol (25:24:1) followed by extraction with chloroform-isoamyl alcohol and then precipitated with 1/10 volume of 3 M NaOAc (pH 5.3), 5 µg of glycogen and 2.5 volumes of ethanol. The pellets were collected by microcentrifugation and resuspended in 30 µl of H_2O.

For ChIP and Quantitative-RT PCR, the pellets were collected by microcentrifugation, resuspended in 30 µl of H_2O. PCR mixtures contained 2µl of immunoprecipitate and were performed using Rotor-Gene SYBRGreen PCR Kit (Qiagen) on the Rotor-Gene Q system (Qiagen). The conditions used were as follows: 95^{o}C for 5 sec, followed by 40 cycles of 60^{o}C for 10 sec. All reactions were run in triplicate. The forward and reverse primers represent following table 8.

Primer Name	Sequence (5'-----> 3")
ChIP_F_C8ORF42	AGA AGG CAC TGC TGG TGC
ChIP_R_ C8ORF42	CCT CAT TTT TTT TGC TCT CCC TTA
ChIP_F_C16ORF89	GGG TGG CCT GGG CAA A
ChIP_R_ C16ORF89	GCA GCC CTC TCC CCA
ChIP_F_C6ORF134	GGG TGT GTA GGG ATT GTG TA
ChIP_R_ C6ORF134	AGG CAT GCC ACA AGG AAA AG
ChIP_F_FEZ2	GGG AAG AAA TGT TTC CAG GGT C
ChIP_R_FEZ2	CCC TCT CTT TCT CCT TTT TGT TGT
ChIP_F_MSN	CTG TCT AGA GAA GGA AAA GAT AGG TAG
ChIP_R_MSN	CTC CTC TCC CAC CCT GGT CA

Table 7. List of primer sequences for ChIP and Quantitative-RT PCR.

Quantitative Combined Bisulfite Restriction Analysis of Methylation pattern (Q-COBRA-MP)

The immunoprecipitates were used for bisulfite modification, as previously described [23]. The bisulfite-treated DNA was amplified using COBRA LINE-1 primers. The 20 μl of PCR reaction includes: 25 U HotStarTaq DNA Polymerase (Qiagen, Germany) with 1× PCR buffer (Qiagen, Germany), 2.5 mM of Magnesium chloride (Qiagen, Germany), 200 μM deoxynucleotide triphosphate (dNTP) (Promega, USA), 1 μM of each primer, and 2 μl sample of bisulfited DNA. The reactions were incubated at 95°C for 10 min, followed by 35 cycles of 95°C for 30 s, 54°C for 30 s and 72°C for 30 s, with a final extension at 72°C for 7 min.

The amplicons were digested in 10 ml reaction volumes with 2U of TaqI in 1xTaqI buffer (MBI Fermentas) at 65°C overnight and then electrophoresed through 8% non-denaturing polyacrylamide gels. The intensities of DNA fragments were measured with a PhosphorImager and analyzed using Excel (Microsoft). The methylated amplicons (TaqI positive) and unmethylated amplicons (uncut amplicon) yielded were calculated as a percentage (the intensity of the methylated candidate gene digested by TaqI divided by the sum of the unmethylated amplicon and the TaqI-positive amplicons) following this formula: A=92/92, B=60/56, C=50/48, D=42/40, E=32/28 and F=((D+E)-(B+C))/2. The LINE-1 methylation level was calculated from [(A+2C+F)x100]/(2A+2B+2C+2F) and statistical analysis by paired T-test.

Pyrosequencing

Si-AGO4 HEK293 with and without tetracycline treatment (Tet+ and Tet-) cells were harvested by trypsinization and extracted for DNA. Brief, cells were lysised with 10% sodium dodecyl sulfate (SDS) (Sigma-Aldrich, USA), lysis buffer II (0.75 M NaCl, 0.024 M EDTA, pH 8) and 20 mg/ml proteinase K (USB, USA) at 50^{o}C for 7 days. Phenol/chloroform extraction and ethanol precipitation were carried out as previously described. The amount 750 ng of each genomic DNA sample was used for bisulfite modification using EZ DNA methylation-Gold™kit (Zymo research, USA) according to the manufacturer's instruction. After resuspend with Tris-EDTA buffer (pH 8.0). The eluted DNA was amplified consequently by PCR. The PCR mixture volume is 30 μl includes: 1 unit Taq Pol with 1× buffer, 2.5 mM $MgCl_2$, 200 μM

dNTP, 1 μM primers, and 3 μl bisulfited DNA. The reactions were incubated at 95°C for 10 min, followed by 35 cycles of 95°C for 30 s, 54-60°C for 30 s, and 72°C for 30 s, and a final extension at 72°C for 7 min. PCR assays were designed to amplify a part of the specific CpG in the gene promoters. Confirmation of PCR product quality was on 8% acrylamide gels with ethidium bromide staining. The forward and reverse primers represent following table 9.

Primer Name	Sequence (5'-----> 3")
C8ORF42-F	AGA AGG TAT TGT TGG TGT TAT AGG
C8ORF42-RBiotin	CCT CAT TTT TTT TAC TCT CCC TTA CT
C16ORF89-F	GGG TGG TTT GGG TAA AGG A
C16ORF89-RBiotin	ACA ACC CTC TCC CCA ACA
C6ORF134-F	GGG TGT GTA GGG ATT GTG TAT
C6ORF134-RBiotin	AAA CAT ACC ACA AAA AAA AAC TCT CTA
FEZ2-F	GGG AAG AAA TGT TTT TAG GGT TAT ATG A
FEZ2-RBiotin	CCC TCT CTT TCT CCT TTT TAT TAT ATC T
MSN-F	TTG TTT AGA GAA GGA AAA GAT AGG TAG TGA
MSN-RBiotin	CTC CTC TCC CAC CCT AAT CAC AA

Table 8. List of primer sequences for pyrosequencing.

Pyrosequencing was carried out using the PyroMask® Gold Q24 machine (Qiagen) according to manufacturer's protocol. Pyrosequencing primers were subsequently designed to focus on 3-5 targets CpG dinucleotides of C8ORF42, C16ORF89, C6ORF134, FEZ2 and MSN. In

brief, pyrosequencing was applied to each sample in triplicate. Twenty microliters of PCR product was immobilized on Streptavidin Sepharose HP beads (GE Healthcare Life Sciences) and purified using the PyroMark Q24 vacuum work station (Qiagen), according to the manufacturer's instructions. After denaturation at 80°C for 2 min, PCR products were annealed to a sequencing primer, and subjected to pyrosequencing using the PyroMark Q24 instrument (Qiagen) and the PyroMark Gold Q24 reagents (Qiagen). The nucleotide dispensation order and the sequence to analyze were determined with the PyroMark Q24 software 2.0.6. Negative-control nucleotides were automatically incorporated, and a cytosine was added at position to check the efficiency of bisulfite conversion. Pyrograms were analyzed in CpG assay mode in order to quantify the methylation percentage of the CpGs.

The principle of pyrosequencing, based on PCR, DNA will be polymerized by polymerase enzyme and release pyrophosphate. Adenosine 5' phosphosulphate and pyrophosphate will be catalyzed by sulfulylase and release ATP. This ATP drives luciferase for convert luciferin to oxyluciferin that generate visible light. Apyrase degrades ATP and unincoperate dNTP. Methylation level will be calculated from the formular C peak height divide sum of C and T peak height. The sequence primers represent following table 10.

Primer Name	Sequence (5'-----> 3")
C8ORF42	GGT TGG TTT TTG GAT TAT AA
C16ORF89	AAG GGT GGT TGT AGG
C6ORF134	GTG AAG TTT TTG GAG TAA T
FEZ2	GTG AAG TTT TTG GAG TAA T
MSN	GAG TGG GGG TGG GGT

Table 9. List of sequencing primers for pyrosequencing.

CHAPTER IV
RESULTS

Connection Up- or Down- Regulation Expression Analysis of Microarrays X (CU-DREAM-X)

In this study, we determined whether the roles of human AGO1-4 in the epigenetic regulation of genes containing LINE-1 are different. The mRNA profiles of HEK293 cells knocked down for AGOs from GSE4246 was categorized as up- or down-regulated and not up- or not down-regulated; the statistical significance of 1454 genes containing or not intragenic LINE-1 was determined using Student's t-test and the chi-square test. The results showed that the reduction of AGO2 expression in HEK293 cells significantly promoted the up-regulation of genes containing intragenic LINE-1, OR = 1.88, *P*-value = 6.19E-04 and 95% CI = 1.30-2.70. This result supports that the biological function of AGO2 is to down-regulate genes containing intragenic LINE-1. In contrast, HEK293 cells knocked down for AGO3 did not show regulated gene expression. HEK293 cells knocked down for AGO1 and -4 showed hampered up-regulation, with significant ORs = 0.55 and 0.33, *P*-values = 4.67E-04 and 1.34E-08, 95% CI = 0.39-0.77 and 0.22-0.50 respectively. Conversely, AGO1-4 siRNA did not significantly altered down-regulating genes containing LINE-1 elements. These results support that AGO1 and AGO4 play a role that is opposite to AGO2 in terms of the expression of genes containing intragenic LINE-1 (Table 11).

		Si-AGO1			Si-AGO2			Si-AGO3			Si-AGO4		
		OR	*P*-value	95% CI	OR	*P*-value	95% CI	OR	*P*-value	95% CI	OR	*P*-value	95% CI
Up	LINE-1	0.55	4.67E-04	0.39-0.77	1.88	6.19E-04	1.30-2.70	0.78	4.45E-01	0.41-1.48	0.33	1.34E-08	0.22-0.50
Down	LINE-1	1.29	5.21E-02	1.00-1.67	0.82	4.57E-01	0.48-1.39	1.29	3.03E-01	0.79-2.11	1.27	7.99E-02	0.97-1.66

Table 10. The result of CU-DREAM-X. The ORs, P-values and lower and upper 95% confidence intervals (95% CI) of Si-AGO1-4 are grouped in terms of up- and down-regulated gene expression.

Chromatin immunoprecipitation (ChIP) and quantitative combined bisulfite restriction analysis of methylation pattern (Q-COBRA-MP)

To evaluate whether the epigenetic modification of AGOs bound to LINE-1s are different, we investigated the methylation status of LINE-1-bound AGO1-4 proteins by ChIP and Q-COBRA-MP. The results revealed that AGO1 and AGO4 preferentially bind to normally methylated LINE-1s, whereas AGO2 and AGO3 preferentially bind to hypomethylated LINE-1s (Figure 12).

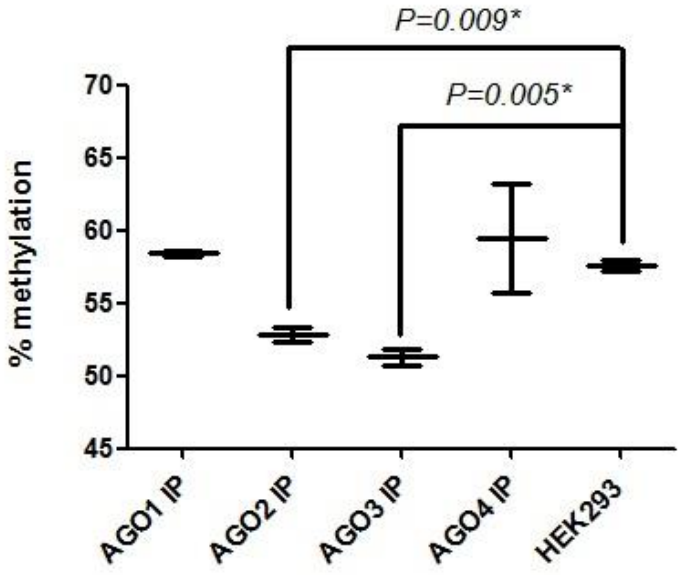


Figure 12. ChIP and Q-COBRA-MP. Methylation percentage of AGO1 and AGO4 shows normality compared with the HEK293 cell line, whereas the methylation percentage of AGO2 and AGO3 shows hypomethylation compared with the HEK293 cell line.

AGO4 Binding Site Mapping

From the analysis to explore association between promoter methylation and AGO4 binding sites by bioinformatics approach, we categorized the genes into two groups, AGO4 binding genes (AGO+) that are hypermathylated comprise C8ORF42, C16ORF89, C6ORF134 and FEZ2. The other is non-AGO binding gene (AGO-), MSN. Obtaining of probe numbers and AGO4 binding sequences makes us derive genomics DNA sequences. Then AGO4 binding site are located into the sequence, up-and downstream CpGs nearby are examined methylation patterns. The forward and reverse pyrosequencing primers of each gene were designed

by PyroMask Assay Design (Qiagen). The amplicon sizes are approximately 90-250 base pairs (Figure 13).

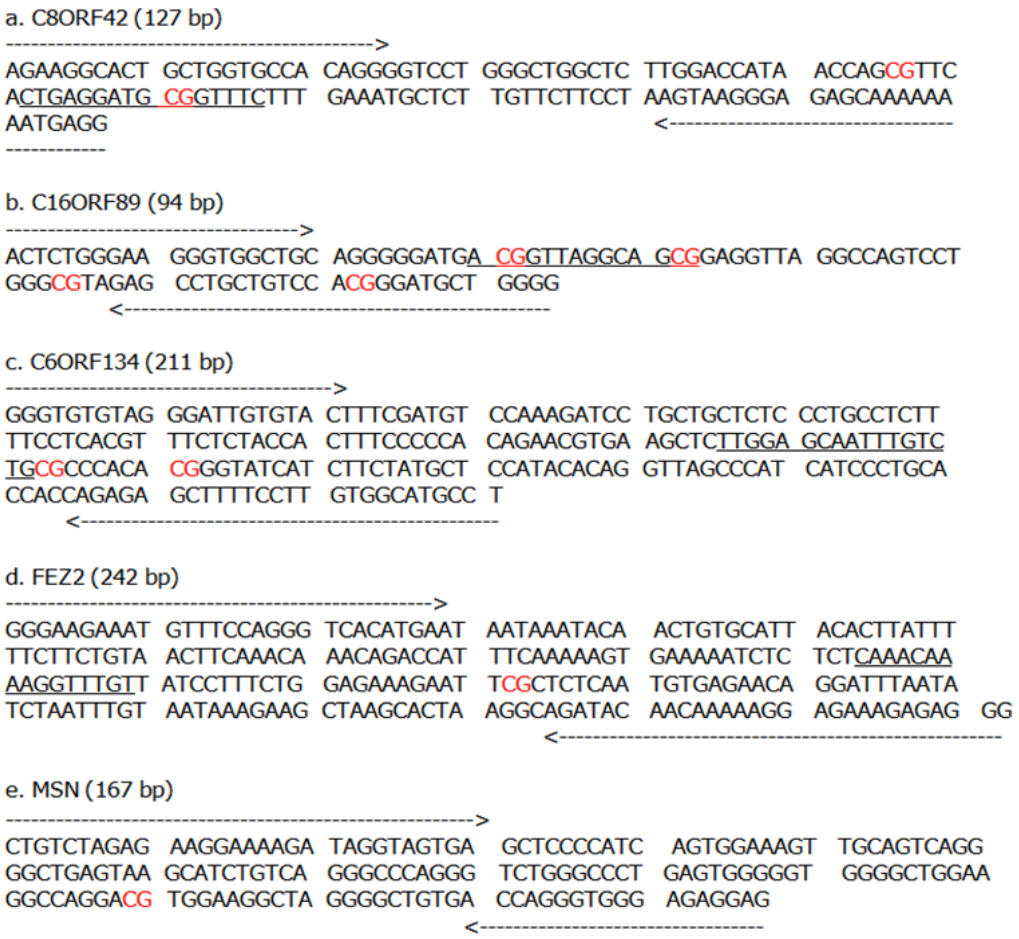


Figure 13. Amplicon sequences of C8ORF42, C16ORF89, C6ORF134, FEZ2 and MSN. Red alphabets are methylated CpGs, underlines are AGO4 binding sites, forward and reverse arrows are forward and reverse primers.

Argonaute 4 Proteins Mediate Small RNA Guided De Novo Methylation in Human Cells

After Si-AGO4 HEK 293 cell lines were treated with tetracycline for 7 and 9 days. We investigated AGO4 expression of Tet on Si-AGO4 HEK 293 cell compare with Tet off Si-AGO4 HEK 293 cell. The results revealed that AGO4 expression was decreased significantly (Figure 14).

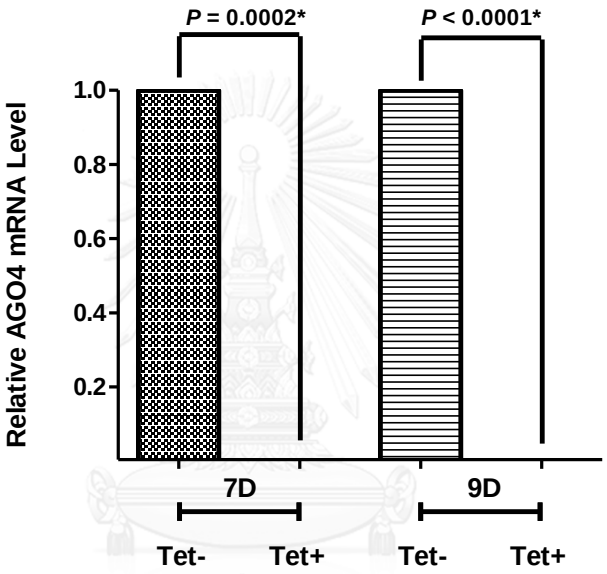


Figure 14. AGO4 expression of tetracycline-inducible sh-AGO4 HEK-293 cells was cultured without tetracycline treatment (Tet-) and with tetracycline (Tet+) for 7 and 9 days.

We evaluated the role of AGO4 in DNA methylation by establishing tetracycline-inducible sh-AGO4 HEK-293 cells. Using bioinformatic information from promoter-methylation dataset (GSE20598) and AGO binding sites (CLIPZ database), 9 loci of four known AGO4 binding methylated genes (AGO+), C8ORF42, FEZ2, C16ORF89 and C6ORF134, were evaluated whereas AGO4 lacking gene (AGO-), MSN as a control.

First, to determine AGO4 binds with AGO4 binding gene. Sh-AGO4 HEK293 cells were cultured and treated with tetracycline to repress AGO4 expression for 9 days. Then Sh-AGO4 HEK293 cells were harvested and were determined DNA recovery using ChIP technique. The results showed that the DNA recovery of all of AGO4 binding genes reduced significantly when compared with cells without tetracycline treatment (Tet-), while DNA recovery of AGO4 lacking gene, MSN was not changed (Figure 15).

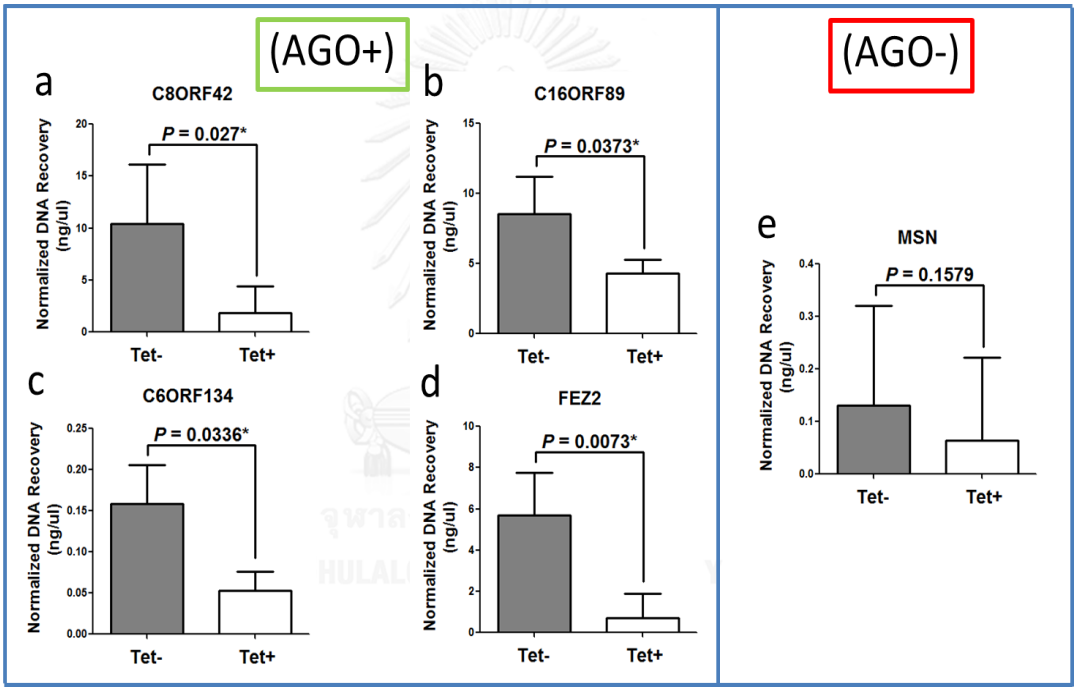


Figure 15. Normalized DNA recovery of AGO+ and AGO- genes in Sh-AGO4 HEK293 cells were treated with tetracycline for 9 days.

Then, to determine AGO4 associated with DNA methylation. Sh-AGO4 HEK293 cells were cultured and treated with tetracycline to repress AGO4 expression for 7 days. Sh-AGO4 HEK293 cells were harvested and determined DNA methylation using pyrosequencing technique. The results showed that the percentage of methylation levels

both AGO4 binding genes and AGO4 lacking gene not changed when compared to cells without tetracycline treatment (Tet-) (Figure 16).

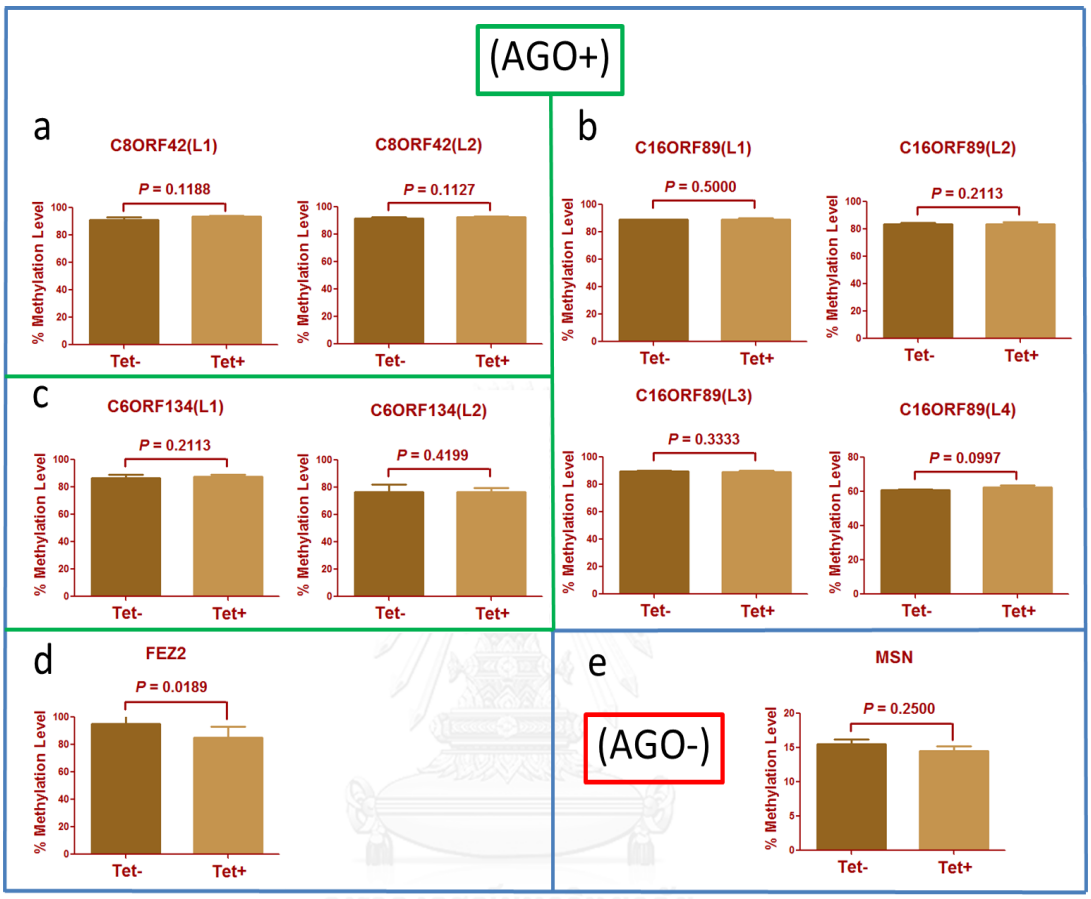


Figure 16. The percentage of methylation levels of AGO+ and AGO- genes in Sh-AGO4 HEK293 cells were treated with tetracycline for 7 days.

After harvest Sh-AGO4 HEK293 cells were cultured with and without tetracycline for 7 days (Tet + and Tet-). Cell were treated with 5-aza-CdR and cultured for 2 days. Then, cells were harvested to determine DNA methylation. Results of experiment shows that the percentage of methylation level all of AGO4 binding genes in Tet +, Aza + cells decreased significantly while AGO4 lacking genes unchanged compared with Tet-, Aza- cells (Figure 17).

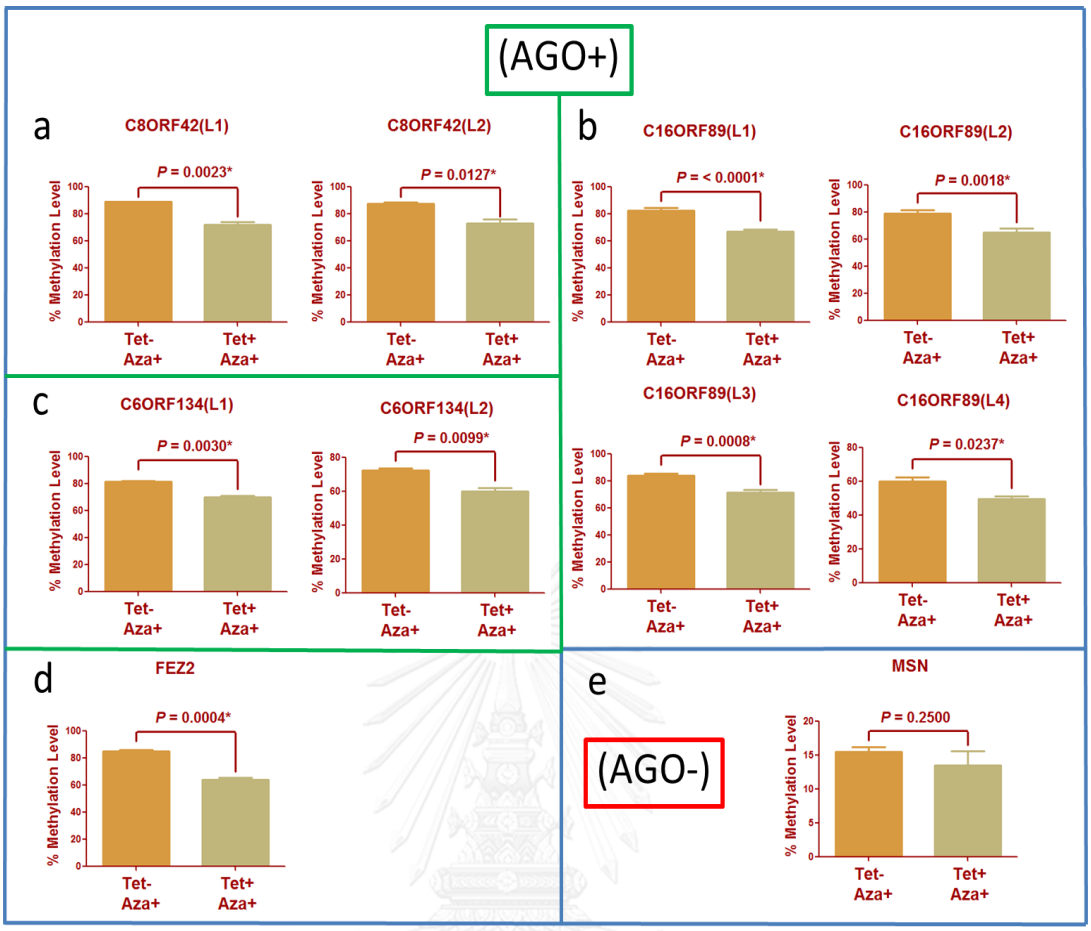


Figure 17. The percentage of methylation levels of AGO+ and AGO- genes in Sh-AGO4 HEK293 cells were treated with 5-aza-CdR for 2 days.

After harvest Sh-AGO4 HEK293 cells were treated tetracycline together 5-aza-CdR for 9 days (Tet + and Tet). Cells were splited and treated with and without tetracycline treatment for 3 days. Then, cells were harvested to determine DNA methylation. Results of the experiments showed that the percentage of methylation level almost all of AGO4 binding genes and AGO4 lacking genes in Tet- cells recovered significantly compared to the Tet + cells (Figure 18).

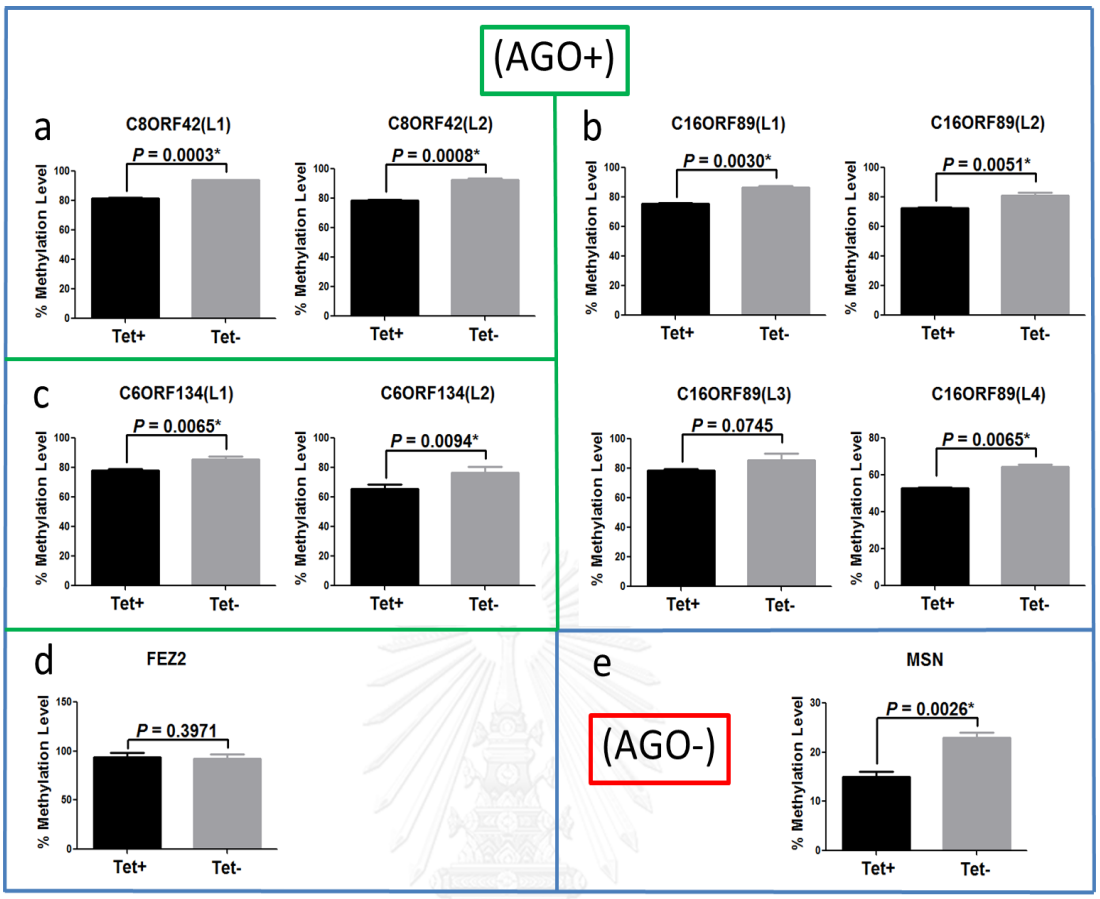


Figure 18. The percentage of methylation levels of AGO+ and AGO- genes in Sh-AGO4 HEK293 cells were treated with tetracycline for 9 days and were cultured continuously for 3 days with and without tetracycline treatment.

CHAPTER V
DISCUSSION

Distinctive Roles of Human Argonaute1, 2, 3 and 4 Proteins on LINE-1 Methylation and Regulation of Intragenic LINE-1-Containing Genes

Previously, we demonstrated that AGO2 represses genes containing hypomethylated LINE-1s. Here, we demonstrate that AGO1 and AGO4 regulate genes containing LINE-1s in an opposite manner as that of AGO2. Furthermore, AGO1 and AGO4 bind to different LINE-1 loci from AGO2. Whereas AGO2 and AGO3 bind to hypomethylated LINE1s, AGO1 and AGO4 bind to LINE-1s that have methylation levels similar to the entire genome.

The AGO2-mediated repression of genes containing hypomethylated LINE-1s is mediated by LINE-1 RNA. When intragenic LINE-1s are hypomethylated, LINE-1 RNA is transcribed; the LINE-1 RNA then forms a complex with pre mRNA and AGO2, resulting in mRNA depletion. On the other hand, when AGO2 were knocked down, LINE-1 RNA is still transcribed however LINE-1 RNA cannot form a complex with AGO2, resulting in increased up regulated genes. AGO3 binds to hypomethylated LINE-1s, similar to AGO2. Nevertheless, the depletion of AGO3 did not alter gene expression. The similar roles between AGO2 and AGO3 and the opposite roles between AGO1 and AGO2 were demonstrated previously with regard to the regulatory roles of mononucleotide A repeats. We speculate that the AGO2 and AGO3

functions are redundant; thus, the depletion of AGO3 was compensated for by AGO2 (Figure 19).

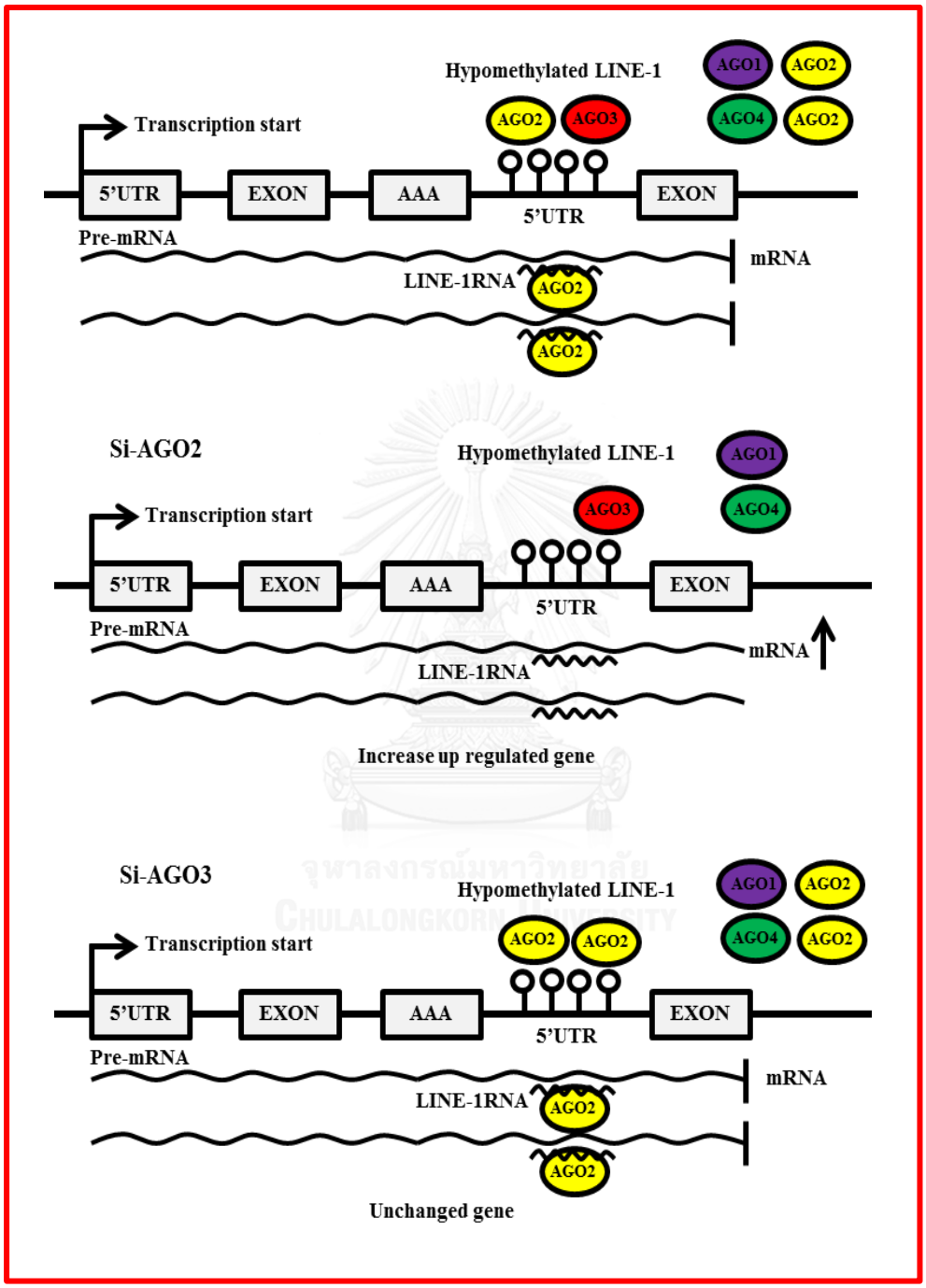


Figure 19. Illustration represents hypothesis mechanisms how do AGO2 and AGO3 proteins are associated with regulation of genes containing LINE-1 in human cells.

The down-regulation of AGO1 and AGO4 prevents the up-regulation of gene expression. It is possible that AGO1 and AGO4 prevent LINE-1 transcription. Consequently, the presence of AGO1 and AGO4 at normally methylated LINE-1s prevents intragenic LINE-1 RNA production, and AGO1 and AGO4 consequently limit LINE-1 RNA from interfering with gene transcription. In normal cell, intragenic LINE-1 is hypomethylated. The host gene is transcribes. When either AGO1 or AGO4 is knocked down, the other AGOs remain bind to hypermethylated LINE-1. So, LINE-1 RNA cannot transcribe, resulting in prevented up regulated genes (Figure 20).

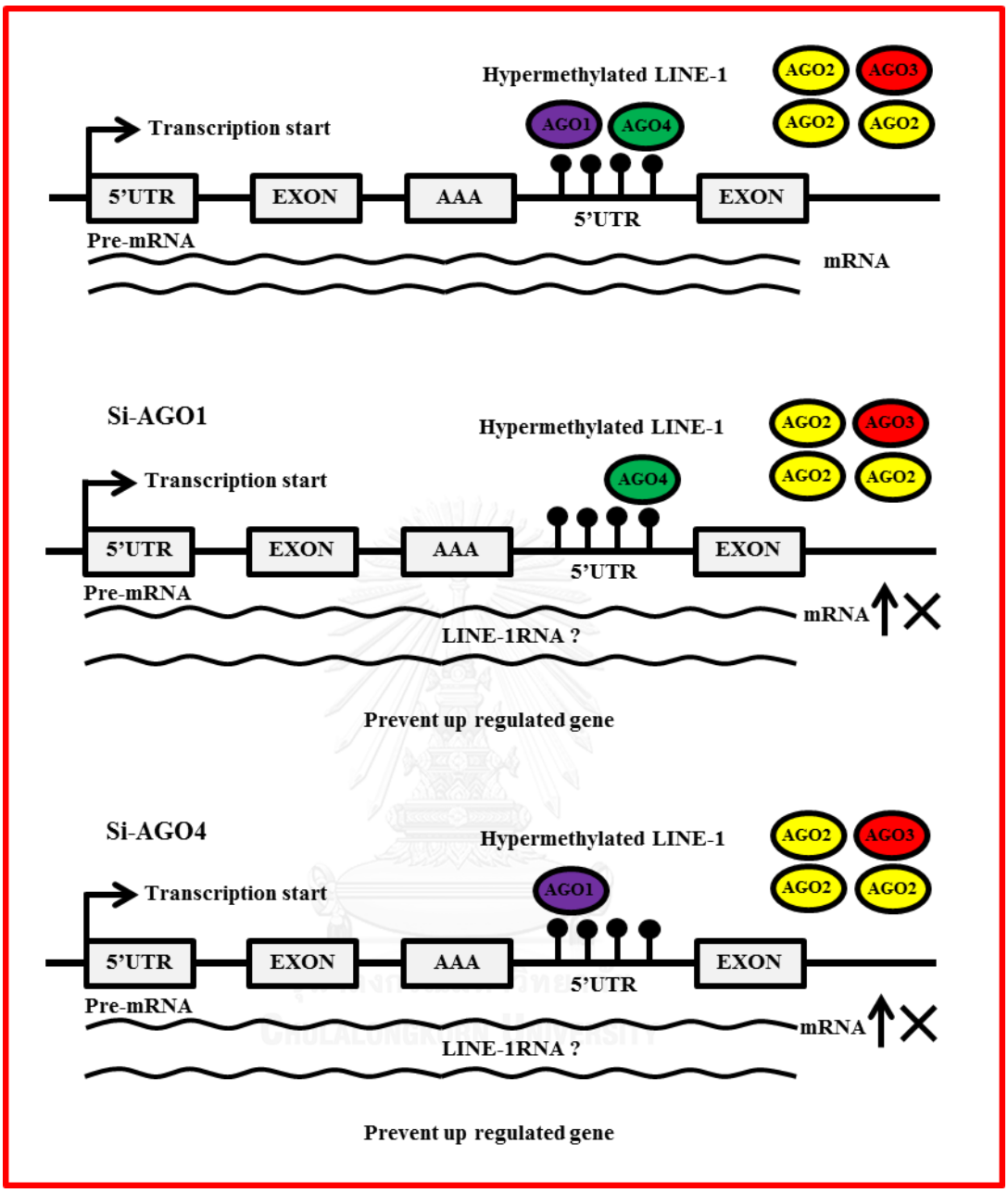


Figure 20. Illustration represents hypothesis mechanisms how do AGO1 and AGO4 proteins are associated with regulation of genes containing LINE-1 in human cells.

In a variety of living organisms, AGO1 has the function of promoting heterochromatin; in plants, AGO4 promotes DNA methylation. It is possible that AGO1 and AGO4 in humans may be

associated with heterochromatin and methylation-associated LINE-1s, with AGO1 and AGO4 binding to transcription-inactiveLINE-1s.

In conclusion, we explored whether AGO proteins play a different role in controlling gene expression. AGO2 repress genes containing hypomethylated LINE-1. AGO3 may be redundant to AGO2, while AGO2 plays a major role. The AGO1 and AGO4 roles are distinctive from AGO2 and have opposite effects in regulating gene expression. It is interesting to further explore the mechanisms underlying these differences.

Argonaute 4 Proteins Mediate SmallRNA Guided De Novo Methylation in Human Cells

From literature reviews are reported that the ability of Ago4 to control the specific promoter methylation through the process of RNA directed DNA methylation (RdDM) in plants. Furthermore, we demonstrated that AGO4 shows the significant correlation with global promoter methylation. So it is interesting to define which DNA methylation related biological processes are AGO4 dependent.

Then, we evaluated association between AGO4 binding and mehtylation status. AGO4 binding genes in tetracycline treatment (Tet+) were reduced when compared with cells without tetracycline treatment (Tet-). This result confirms AGO4 protein was interfered by tetracycline treatment. When Si-AGO4 HEK 293 cells were treated tetracycline on 7th days, methylation level of AGO4 binding genes was not changed suggesting AGO4 is not a prerequisite to maintain DNA methylation. It is possible that DNMT1 still load methyl group to specific loci due to

DNMT1 acts as maintain DNA methyltransferase. It will add the methyl group in daughter strand after every DNA replication. (Figure 21).

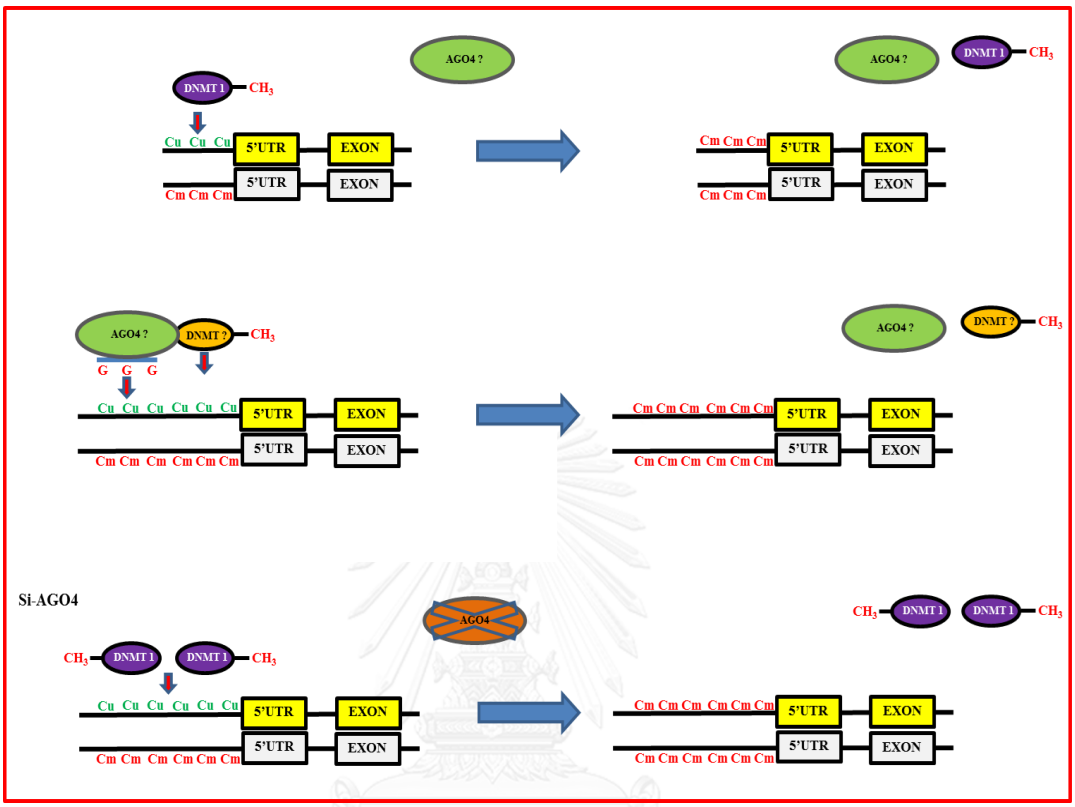


Figure 21. Illustration represents hypothesis mechanisms of DNA methylation in Si-AGO4 HEK293 cells were treated tetracycline for 7 days.

Whereas, Si-AGO4 HEK 293 cell lines were treated tetracycline together with 5'azacytidine on 9th days, the results of combination between reduction of AGO4 binding to target loci and a DNA demethylating agent, 5-aza-CdR, treatment reduced significantly DNA methylation levels at the loci than 5-aza-CdR treatment alone. It is possible that AGO4 acts as de novo methylation by recruits the other DNA methylatransferase to add methyl group such as DNMT3a DNMT3b. These DNA methylatransferase were reported act as de novo

methylation that play an important role to add methyl group at specific loci without methylated parent strand.

To evaluate AGO4 acts as de novo methylation molecules, the 9th days treated tetracycline and 5'azacytidine Si-AGO4 HEK 293 cell lines were cultivated continuously for 3 days with tetracycline but without 5-aza-CdR. Re-methylation of C8ORF42, C16ORF89 and C6ORF134 in Si-AGO4 HEK293 cell lines can be observed at higher levels by re-introducing AGO4. This finding was comprehended that AGO4 depletion and 5-aza-CdR treatment decreased DNA methylation by different mechanism. 5-aza-CdR is chemical reagent to interfere maintain DNA methyltransferase during DNA replication, preventing adding methyl group to a daughter DNA strand. Therefore, it is possible that AGO4 dependent RdDM contributes to large number of de novo methylation loci in human cells (Figure 22).

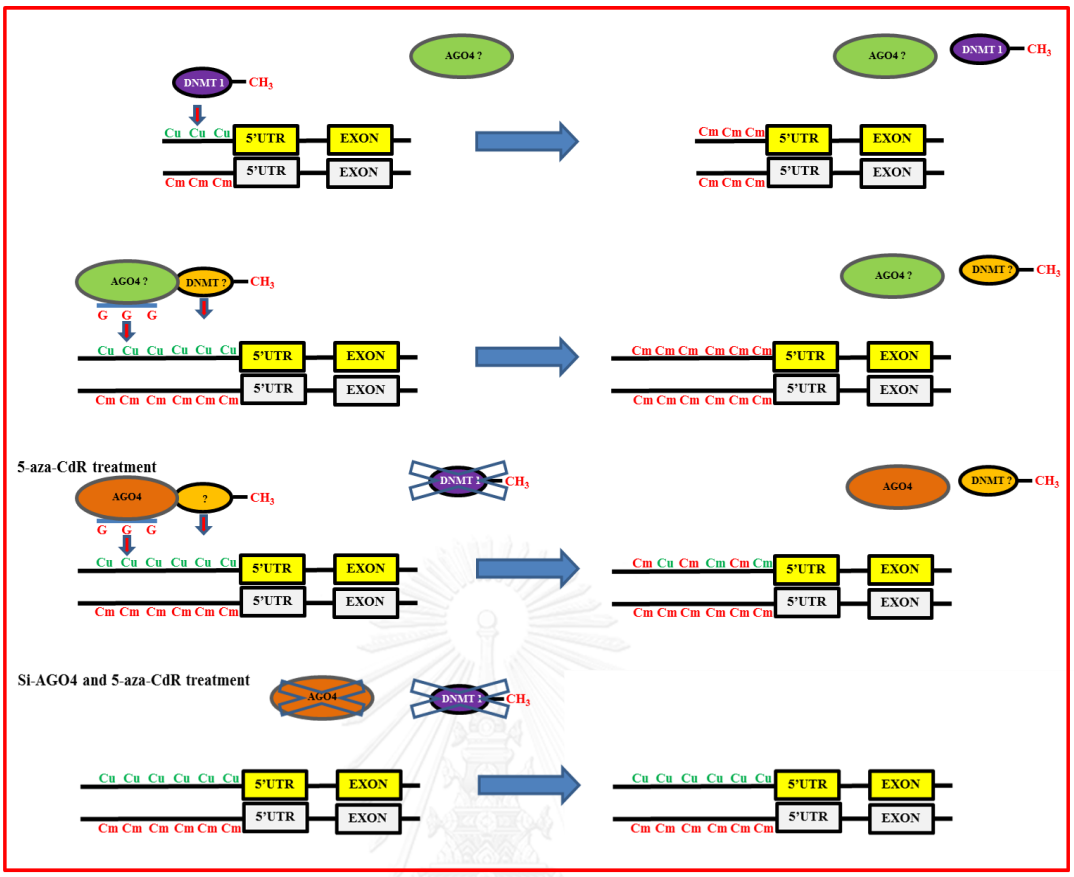


Figure 22. Illustration represents hypothesis mechanisms how does the role of AGO4 proteins as mediate small-RNA-guided de novo methylation in human cell.

REFERENCES

1. Matzke, M., et al., *RNA-mediated chromatin-based silencing in plants.* Current opinion in cell biology, 2009. **21**(3): p. 367-376.

2. Wassenegger, M., et al., *RNA-directed de novo methylation of genomic sequences in plants.* Cell, 1994. **76**(3): p. 567-576.

3. Mette, M., et al., *Transcriptional silencing and promoter methylation triggered by double-stranded RNA.* The EMBO Journal, 2000. **19**(19): p. 5194-5201.

4. Haag, J.R. and C.S. Pikaard, *Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing.* Nature reviews Molecular cell biology, 2011. **12**(8): p. 483-492.

5. Komashko, V.M. and P.J. Farnham, *5-azacytidine treatment reorganizes genomic histone modification patterns.* Epigenetics, 2010. **5**(3): p. 229-240.

6. Khorshid, M., C. Rodak, and M. Zavolan, *CLIPZ: a database and analysis environment for experimentally determined binding sites of RNA-binding proteins.* Nucleic acids research, 2010: p. gkq940.

7. Volpe, T.A., et al., *Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi.* science, 2002. **297**(5588): p. 1833-1837.

8. Wei, K.F., et al., *Structural evolution and functional diversification analyses of argonaute protein.* Journal of cellular biochemistry, 2012. **113**(8): p. 2576-2585.

9. Rashid, U.J., et al., *Structure of Aquifex aeolicus argonaute highlights conformational flexibility of the PAZ domain as a potential regulator of RNA-induced silencing complex function.* Journal of Biological Chemistry, 2007. **282**(18): p. 13824-13832.

10. Braun, J.E., E. Huntzinger, and E. Izaurralde, *The role of GW182 proteins in miRNA-mediated gene silencing*, in *Ten Years of Progress in GW/P Body Research*. 2013, Springer. p. 147-163.

11. Chen, S., et al., *Ago-2-mediated slicer activity is essential for anti-flaviviral efficacy of RNAi.* PloS one, 2011. **6**(11): p. e27551.

12. Shekar, P.C., et al., *Argonaute-2-null embryonic stem cells are retarded in self-renewal and differentiation.* Journal of biosciences, 2011. **36**(4): p. 649-657.

13. Okamura, K., et al., *Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways.* Genes & development, 2004. **18**(14): p. 1655-1666.

14. Aravin, A.A., et al., *Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line.* Molecular and cellular biology, 2004. **24**(15): p. 6742-6750.

15. Brennecke, J., et al., *Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in< i> Drosophila</i>*. Cell, 2007. **128**(6): p. 1089-1103.

16. Kalmykova, A.I., M.S. Klenov, and V.A. Gvozdev, *Argonaute protein PIWI controls mobilization of retrotransposons in the*

Drosophila male germline. Nucleic acids research, 2005. **33**(6): p. 2052-2059.

17. Vagin, V.V., et al., *A distinct small RNA pathway silences selfish genetic elements in the germline.* Science, 2006. **313**(5785): p. 320-324.

18. Qi, Y., et al., *Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation.* Nature, 2006. **443**(7114): p. 1008-1012.

19. Zilberman, D., X. Cao, and S.E. Jacobsen, *ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation.* Science, 2003. **299**(5607): p. 716-719.

20. Aporntewan, C., et al., *Hypomethylation of intragenic LINE-1 represses transcription in cancer cells through AGO2.* PloS one, 2011. **6**(3): p. e17934.

21. Kitkumthorn, N. and A. Mutirangura, *Long interspersed nuclear element-1 hypomethylation in cancer: biology and clinical applications.* Clinical epigenetics, 2011. **2**(2): p. 315-330.

22. Aporntewan, C. and A. Mutirangura, *Connection up-and down-regulation expression analysis of microarrays (CU-DREAM): a physiogenomic discovery tool*, in *Asian Biomed*. 2011. p. 257.

23. Chalitchagorn, K., et al., *Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis.* Oncogene, 2004. **23**(54): p. 8841-8846.

24. Leonhardt, H. and M.C. Cardoso, *DNA methylation, nuclear structure, gene expression and cancer.* Journal of cellular biochemistry, 2000. **79**(S35): p. 78-83.

25. Leonhardt, H., et al., *Dynamics of DNA replication factories in living cells.* The Journal of cell biology, 2000. **149**(2): p. 271-280.

26. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes & development, 2002. **16**(1): p. 6-21.

27. Ehrlich, M., *DNA methylation in cancer: too much, but also too little.* Oncogene, 2002. **21**(35): p. 5400-5413.

28. Yoder, J.A., C.P. Walsh, and T.H. Bestor, *Cytosine methylation and the ecology of intragenomic parasites.* Trends in genetics, 1997. **13**(8): p. 335-340.

29. Leonhardt, H., et al., *A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei.* Cell, 1992. **71**(5): p. 865-873.

30. Kafri, T., et al., *Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line.* Genes & development, 1992. **6**(5): p. 705-714.

31. Okano, M., et al., *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.* Cell, 1999. **99**(3): p. 247-257.

32. Okano, M., S. Xie, and E. Li, *Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases.* Nature genetics, 1998. **19**(3): p. 219-220.

33. Miniou, P., et al., *Abnormal methylation pattern in constitutive and facultative (X inactive chromosome) heterochromatin of ICF patients.* Human molecular genetics, 1994. **3**(12): p. 2093-2102.

34. Vertino, P.M., et al., *De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5-)-methyltransferase.* Molecular and cellular biology, 1996. **16**(8): p. 4555-4565.

35. Mikovits, J.A., et al., *Infection with human immunodeficiency virus type 1 upregulates DNA methyltransferase, resulting in de novo methylation of the gamma interferon (IFN-γ) promoter and subsequent downregulation of IFN-γ production.* Molecular and cellular biology, 1998. **18**(9): p. 5166-5177.

36. Liang, G., et al., *Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements.* Molecular and cellular biology, 2002. **22**(2): p. 480-491.

37. Aravin, A.A., et al., *A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice.* Molecular cell, 2008. **31**(6): p. 785-799.

38. Fuks, F., et al., *DNA methyltransferase Dnmt1 associates with histone deacetylase activity.* Nature genetics, 2000. **24**(1): p. 88-91.

39. Fuks, F., et al., *Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription.* The EMBO journal, 2001. **20**(10): p. 2536-2544.

40. Hendrich, B. and A. Bird, *Identification and characterization of a family of mammalian methyl-CpG binding proteins.* Molecular and cellular biology, 1998. **18**(11): p. 6538-6547.

41. Meehan, R.R., et al., *Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs.* Cell, 1989. **58**(3): p. 499-507.

42. Bird, A.P. and A.P. Wolffe, *Methylation-induced repression-belts, braces, and chromatin.* Cell, 1999. **99**(5): p. 451-454.

43. Robertson, K.D., *Epigenetic mechanisms of gene regulation*, in *DNA Methylation and Cancer Therapy*. 2005, Springer. p. 13-30.

44. Jaenisch, R. and A. Bird, *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals.* Nature genetics, 2003. **33**: p. 245-254.

45. Kobayashi, H., J. Ngernprasirtsiri, and T. Akazawa, *Transcriptional regulation and DNA methylation in plastids during transitional conversion of chloroplasts to chromoplasts.* The EMBO journal, 1990. **9**(2): p. 307.

46. Wu, H. and Y. Zhang, *Reversing DNA methylation: mechanisms, genomics, and biological functions.* Cell, 2014. **156**(1): p. 45-68.

47. Suzuki, M.M. and A. Bird, *DNA methylation landscapes: provocative insights from epigenomics.* Nature Reviews Genetics, 2008. **9**(6): p. 465-476.

48. Szyf, M., *DNA methylation and cancer therapy.* Drug resistance updates, 2003. **6**(6): p. 341-353.

49. Ostertag, E.M. and H.H. Kazazian Jr, *Biology of mammalian L1 retrotransposons.* Annual review of genetics, 2001. **35**(1): p. 501-538.

50. Babushok, D.V. and H.H. Kazazian, *Progress in understanding the biology of the human mutagen LINE-1.* Human mutation, 2007. **28**(6): p. 527-539.

51. Kazazian, H.H., *Mobile elements: drivers of genome evolution.* Science, 2004. **303**(5664): p. 1626-1632.

52. Ostertag, E.M., et al., *SVA elements are nonautonomous retrotransposons that cause disease in humans.* The American Journal of Human Genetics, 2003. **73**(6): p. 1444-1451.

53. Schulz, W.A., *L1 retrotransposons in human cancers.* BioMed Research International, 2006. **2006**.

54. Kano, H., et al., *L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism.* Genes & development, 2009. **23**(11): p. 1303-1312.

55. Belancio, V.P., et al., *Somatic expression of LINE-1 elements in human tissues.* Nucleic acids research, 2010. **38**(12): p. 3909-3922.

56. Rangwala, S.H., L. Zhang, and H.H. Kazazian Jr, *Many LINE1 elements contribute to the transcriptome of human somatic cells.* Genome Biol, 2009. **10**(9): p. R100.

57. Ergün, S., et al., *Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues.* Journal of Biological Chemistry, 2004. **279**(26): p. 27753-27763.

58. Sciamanna, I., et al., *Inhibition of endogenous reverse transcriptase antagonizes human tumor growth.* Oncogene, 2005. **24**(24): p. 3923-3931.

59. Schulz, W., C. Steinhoff, and A. Florl, *Methylation of endogenous human retroelements in health and disease*, in *DNA Methylation: Development, Genetic Disease and Cancer*. 2006, Springer. p. 211-250.

60. Wanichnopparat, W., et al., *Genes associated with the cis-regulatory functions of intragenic LINE-1 elements.* BMC genomics, 2013. **14**(1): p. 205.

61. Steinhoff, C. and W. Schulz, *Transcriptional regulation of the human LINE-1 retrotransposon L1. 2B.* Molecular Genetics and Genomics, 2004. **270**(5): p. 394-402.

62. Wilson, A.S., B.E. Power, and P.L. Molloy, *DNA hypomethylation and human diseases.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2007. **1775**(1): p. 138-162.

63. Wilhelm-Benartzi, C.S., et al., *Research| Children's Health.* Environmental health perspectives, 2012. **120**(2): p. 297.

64. Li, W., et al., *Evaluation of Argonaute protein as a predictive marker for human clear cell renal cell carcinoma.* International journal of clinical and experimental pathology, 2013. **6**(6): p. 1086.

65. Chen, Z., et al., *Hypoxia-responsive miRNAs target argonaute 1 to promote angiogenesis.* The Journal of clinical investigation, 2013. **123**(3): p. 1057-1067.

66. Li, L., et al., *Argonaute proteins: potential biomarkers for human colon cancer.* BMC cancer, 2010. **10**(1): p. 38.

67. Yoo, N.J., et al., *Immunohistochemical analysis of RNA-induced silencing complex-related proteins AGO2 and TNRC6A in prostate and esophageal cancers.* Apmis, 2010. **118**(4): p. 271-276.

68. Phillips, T., *The role of methylation in gene expression.* Nature Education, 2008. **1**(1): p. 116.

69. Hutvagner, G. and M.J. Simard, *Argonaute proteins: key players in RNA silencing.* Nature reviews Molecular cell biology, 2008. **9**(1): p. 22-32.

70. Pal-Bhadra, M., et al., *Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery.* Science, 2004. **303**(5658): p. 669-672.

71. Jones, L., et al., *Virus-induced gene silencing of argonaute genes in Nicotiana benthamiana demonstrates that extensive systemic silencing requires Argonaute1-like and Argonaute4-like genes.* Plant physiology, 2006. **141**(2): p. 598-606.

72. Wassenegger, M., *RNA-directed DNA methylation*, in *Plant Gene Silencing*. 2000, Springer. p. 83-100.

73. Aporntewan, C., et al., *Upstream mononucleotide A-repeats play a cis-regulatory role in mammals through the DICER1 and Ago proteins.* Nucleic acids research, 2013. **41**(19): p. 8872-8885.

74. Li, C.F., et al., *Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in Arabidopsis thaliana.* PLoS genetics, 2008. **4**(2): p. e27.

75. Pontes, O., et al., *The< i> Arabidopsis</i> Chromatin-Modifying Nuclear siRNA Pathway Involves a Nucleolar RNA Processing Center*. Cell, 2006. **126**(1): p. 79-92.

76. Schmitter, D., et al., *Effects of Dicer and Argonaute down-regulation on mRNA levels in human HEK293 cells.* Nucleic acids research, 2006. **34**(17): p. 4801-4815.

77. Kuo, M.-H. and C.D. Allis, *< i> In Vivo</i> Cross-Linking and Immunoprecipitation for Studying Dynamic Protein: DNA Associations in a Chromatin Environment.* Methods, 1999. **19**(3): p. 425-433.

APPENDIX

SUBMITTED ARTICLE

Distinctive roles of human Argonaute1, 2, 3 and 4 proteins on LINE-1 methylation and regulation of intragenic LINE-1-containing genes

Piyapat Pin-on[a], Apiwat Mutirangura[b,*]

[a]Inter-Department Program of Biomedical Sciences, Faculty of Graduate School, Chulalongkorn University, Bangkok 10330, Thailand

[b]Center of Excellence in Molecular Genetics of Cancer and Human Diseases, Department of Anatomy, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand

[*]Corresponding author, e-mail: mapiwat@chula.ac.th

KEYWORDS: Argonaute proteins, LINE-1 methylation, chromatin immunoprecipitation**,** Connection Up- or Down-Regulation Expression Analysis of Microarrays X

ABSTRACT: Argonaute proteins (AGOs) are evolutionarily conserved and ubiquitously expressed in all higher eukaryotes and have important function in several processes, such as cell differentiation and transposon silencing. The commonly known functions of AGOs are key in the gene-silencing pathways guided by small RNAs. There are many human AGOs; however, redundancy and the distinctive roles of different AGOs have not been well characterized. Previously, we demonstrated that AGO2 down-regulates the expression of genes containing hypomethylated LINE-1s. Here, we evaluate the expression of genes containing LINE-1 in AGO1-4-knocked down HEK293 cells. Furthermore, we measured the methylation levels of AGO1-4-bound LINE-1s. Genes containing LINE-1s in AGO1, -2, -3 or -4-knocked down HEK293 cells showed prevented up-regulation, increased up-regulation, unchanged regulation and prevented up-regulation, respectively. Interestingly, AGO1-4 bound to LINE-1 differently in term of the methylation level. Although the methylation level of AGO1 and -4-bound LINE-1s was not different from the genome, AGO2- and AGO3-bound LINE-1s were hypomethylated. Our experiments demonstrate the distinctive epigenetic roles of AGO1-4 in regulating genes containing LINE-1s. It is interesting to further explore the underlining mechanism of these differences

INTRODUCTION

It is well known that Argonaute proteins (AGOs) play a role in controlling gene expression. There are 8 AGOs in humans[1, 2], but there is limited information demonstrating whether the functions of each AGO are redundant or different[3]. AGOs are a highly conserved and extensively

expressed in many organisms, including plants, animals and humans[4]. There are many reports that human AGOs are involved in organ growth and development[5], and recent studies have clarified its role as a potential factor related to oncogenesis in colonic cancer, breast cancer and prostate cancer[6-8]. Furthermore, several works have revealed that AGO1 is associated with pre-transcriptional gene silencing. First, Ago1 in *S. pombe* is used to direct Histone 3 Lysine-9 methylation (H3K9) of local transposon sequences, resulting in nucleosome compaction and transposon silencing; RNA polymerase II (RNAPII) activity is thus restricted at the targeted locus, followed by heterochromatin formation[9]. Second, *Drosophila* Ago1 also has a role in heterochromatin formation: the Ago1 mutant disrupts pre-transcriptional gene silencing of transgene tandem repeats, with the loss of H3mK9 and silencing of heterochromatin[10]. Third, in *A. thaliana*, Ago1 may be the slicer component of RISC, acting as an RNA slicers that uses the sequence of associated siRNAs to guide the cleavage of homologous RNAs in RNA-induced silencing complexes and in the maintenance of chromatin modifications at some loci. Furthermore, Ago1 knockdown resulted in developmental abnormalities[11]. In addition, Ago1 has been implicated in the silencing activity of promoter-targeted small RNAs. DNA methyltransferase (DNMT3a) has been shown to co-immunoprecipitate with small RNAs at the promoters of some targeted genes. One known mechanism is RNA interference, with Ago2 acting as a trans-acting element to down-regulate gene expression by binding different classes of small noncoding RNAs, including microRNAs (miRNAs) and small interfering RNAs (siRNAs), which guide proteins to their specific targets via sequence complementarity[12-14]. Another mechanism is found in plants

and is termed RNA-directed DNA methylation (RdDM). Ago4 is associated with siRNA, and the ribonucleoprotein (RNP) complex is loaded onto chromatin via complementarity. In gene silencing, the complex guides methyltransferases to establish sequence-specific de novo DNA methylation[15].

Recently, we reported that AGO2 down-regulates genes containing hypomethylated intragenic Long Interspersed Elements-1 (LINE-1)[16]. LINE-1s comprise a group of abundant retrotransposon sequences that are found in large numbers in the human genome. Moreover, our follow-up report demonstrated that intragenic LINE-1 sequences are conserved and regulate several biological processes, including DNA damage and repair, inflammation, immune function and cell differentiation. We screened for genes that are involved in intragenic LINE-1 regulation networks using a bioinformatics approach and found that intragenic LINE-1 acts as a cis-regulatory element within genes to modulate host gene expression. Furthermore, AGOs are also trans-acting elements and coordinate with LINE-1 to regulate gene expression[17]. For example, genes containing LINE-1 are up-regulated in si-AGO2 embryonic kidney cell lines. These results demonstrate that AGO2 plays a role in the control of gene expression for thousands of genes containing LINE-1s. Herein, we demonstrate whether the other AGOs serve as trans-acting elements that mediate gene expression through intragenic LINE-1 elements.

The distinctive role of AGO1-4 in the control gene expression through upstream mononucleotide A-repeats has been reported[3]. We found that all members of the AGO family preferentially bind to A-repeats. However, AGO1-4 knockdown produced different patterns.

AGO1 was up-regulated genes containing upstream A-repeats, whereas AGO2 and AGO3 were significantly associated with down-regulated genes containing upstream A-repeats. AGO4 did not significantly regulate gene expression.

However, it is unclear whether AGOs are associated with intragenic LINE-1 methylation and whether they lead to gene up- and/or down-regulation. Thus, in this study, we evaluated the role of AGO 1-4 in affecting the expression of genes containing LINE-1 using a bioinformatics approach. The mRNA levels in the HEK293 cell line knocked down for AGO1-4 were compared with intragenic LINE-1 genes by The Connection Up- and Down-Regulation Expression Analysis of Microarrays (CU-DREAM) software package (http://pioneer.netserv.chula.ac.th/~achatcha/cudream/)[18]. This software calculates various statistical analyses, including Student's t-test and Pearson's chi-squared test, to analyze the gene regulatory functions of intragenic LINE-1. To explore the association between AGOs and intragenic LINE-1 methylation, we performed chromatin immunoprecipitation (ChIP) and quantitative combined bisulfite restriction analysis of methylation pattern (Q-COBRA-MP) assays [19].

MATERIALS AND METHODS

Cell culture

HEK293 cells were maintained in Dulbecco's modified Eagle's medium (Gibco BRL, Life Technologies) supplemented with 2 mM L-glutamine,

10% (v/v) heat-inactivated fetal bovine serum (Sigma), 10 mg/ml antibiotic/antimycotic (Gibco BRL, Life Technologies) at 37°C and 5% (v/v) CO_2. The medium was changed 3 days later. The cells were harvested for the ChIP assay on the 7th day.

Chromatin immunoprecipitation (ChIP)

A ChIP analysis was performed according to a previously published protocol[20]. Briefly, 1 x 10^6 HEK293 cells in a 75-cm^2 flask were grown to 80% confluence. The cells were harvested, and formaldehyde was added at a final concentration of 1% directly to the cell culture medium. Fixation was performed at 37°C for 10 min and stopped by the addition of glycine to a final concentration of 0.125 M. As much medium as possible was removed, and the cells were washed twice with ice-cold PBS containing protease inhibitors; the cells were scraped into a conical tube. The cells were collected by centrifugation and rinsed in cold phosphate-buffered saline. The cell pellets were resuspended in lysed buffer (0.5% NP40, 85 mM KCl, 5 mM PIPES, pH 8.0), incubated on ice for 20 min and homogenized. The nuclei were collected by microcentrifugation and then resuspended in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.1) and incubated on ice for 10 min. The samples were sonicated on ice with an Ultrasonics sonicator at setting 10 amplitude for 20 sec to achieve an average length of approximately 200-1000 bp; the samples were then microcentrifuged. The sonicated cell pellet was resuspended 10-fold in ChIP Dilution Buffer with protease inhibitors. The fragment solution was precleared with the addition of protein A-positive cells for 15 min at 4°C. The precleared chromatin was incubated with 1 µg of affinity-purified goat

monoclonal antibody or no antibody and rotated at 4°C for approximately 12 to 16 h. The antibodies used included EIF2C1-4 (anti-AGO1-4) or a control non-immunized goat antibody. Protein A agarose (40 µl) was added for 2 h at 4ºC with rotation to collect the antibody/protein complex, and the complexes were washed and eluted. The pellet was gently centrifugation (1084g at 4ºC, 1 min), and the protein A agarose/ antibody/ protein complex was washed for 3-5 minutes on a rotating platform with each of the following buffers in order: a) low-salt immune complex wash buffer (1 X 1 ml); b) high-salt immune complex wash buffer (1 X 1 ml) and c) LiCl immune complex wash buffer (1 X 1 ml).The complex was eluted with elution buffer and shaken for 15 min. The cross-links were reversed by the addition of NaCl to a final concentration of 200 mM, and RNA was removed by the addition of 10 µg of RNaseA per sample, followed by incubation at 65°C for 4 h. The samples were then precipitated at -20°C overnight by the addition of 2.5 volumes of ethanol and then pelleted by microcentrifugation. The samples were resuspended in 100 µl of Tris-EDTA (pH 7.5), 25 µl of proteinase K buffer (1.25% sodium dodecyl sulfate, 50 mM Tris (pH 7.5), 25 mM EDTA) and 1.5 µl of proteinase K (Sigma) and incubated at 45°C for 2 h. The samples were extracted with phenol-chloroform-isoamyl alcohol (25:24:1) followed by extraction with chloroform-isoamyl alcohol and then precipitated with 1/10 volume of 3 M NaOAc (pH 5.3), 5 µg of glycogen and 2.5 volumes of ethanol. The pellets were collected by microcentrifugation and resuspended in 30 µl of H_2O.

Q-COBRA-MP

The immunoprecipitates were used for bisulfite modification, as previously described[19]. The bisulfite-treated DNA (2 µl) was amplified using COBRA LINE-1 primers. The reactions were incubated at 95°C for 10 min, followed by 35 cycles of 95°C for 30 s, 54°C for 30 s and 72°C for 30 s, with a final extension at 72°C for 7 min. The amplicons were digested in 10-ml reaction volumes with 2U of TaqI in 1xTaqI buffer (MBI Fermentas) at 65°C overnight and then electrophoresed through 8% non-denaturing polyacrylamide gels. The intensities of DNA fragments were measured with a PhosphorImager and analyzed using Excel (Microsoft). The methylated amplicons (TaqI positive) and unmethylated amplicons (uncut amplicon) yielded were calculated as a percentage (the intensity of the methylated candidate gene digested by TaqI divided by the sum of the unmethylated amplicon and the TaqI-positive amplicons).

Connection Up- or Down- Regulation Expression Analysis of Microarrays X (CU-DREAM-X)

To evaluate whether intragenic LINE-1 can control host gene expression through AGOs, we collected mRNAs of GSE 4246 from the GEO database[21] and prepared templates of the microarray, series matrix file and platform. Each gene was compared to the means of the experimental and control groups and classified as up- or down-regulated and not up- or not down-regulated, depending on the statistical significance of Student's t-test. Subsequently, the distributions of up- or down-regulated genes were evaluated a to whether the distributions were dependent on containing an intragenic LINE-1 using Pearson's chi-squared test. A list of genes containing LINE-1s was previously reported[17]. Genes were

grouped into four groups, A through D. The significantly up- or down-regulated genes containing intragenic LINE-1 were included in group A. The significant not up- or not down-regulated genes containing intragenic LINE-1 were included in group B. The significant up- or down-regulated genes without intragenic LINE-1 were included in group C. The remaining genes (non-significant genes without LINE-1) were included in group D. The *P*-values of the odds ratio (ORs), *P*-values and lower and upper 95% confidence intervals (CI) of the genes in groups A through D were displayed in an MS Excel format. All of the statistical analyses were performed using extensions in the CU-DREAM software (http://pioneer.netserv.chula.ac.th/~achatcha/cu-dream/) (Table 1).

RESULTS

In this study, we determined whether the roles of human AGO1-4 in the epigenetic regulation of genes containing LINE-1 are different. The mRNA profiles of HEK293 cells knocked down for AGOs from GSE4246 was categorized as up- or down-regulated and not up- or not down-regulated; the statistical significance of 1454 genes containing or not intragenic LINE-1 was determined using Student's t-test and the chi-square test. The results showed that the reduction of AGO2 expression in HEK293 cells significantly promoted the up-regulation of genes containing intragenic LINE-1, OR = 1.88, *P*-value = 6.19E-04 and 95% CI = 1.30-2.70. This result supports that the biological function of AGO2 is to down-regulate genes containing intragenic LINE-1. In contrast, HEK293 cells knocked down for AGO3 did not show regulated gene expression. HEK293 cells knocked down for AGO1 and -4 showed

hampered up-regulation, with significant ORs = 0.55 and 0.33, *P*-values = 4.67E-04 and 1.34E-08, 95% CI = 0.39-0.77 and 0.22-0.50 respectively. Conversely, AGO1-4 siRNA did not significantly altered down-regulating genes containing LINE-1 elements. These results support that AGO1 and AGO4 play a role that is opposite to AGO2 in terms of the expression of genes containing intragenic LINE-1 (Table 2).

To evaluate whether the epigenetic modification of AGOs bound to LINE-1s are different, we investigated the methylation status of LINE-1-bound AGO1-4 proteins by ChIP and Q-COBRA-MP. The results revealed that AGO1 and AGO4 preferentially bind to normally methylated LINE-1s, whereas AGO2 and AGO3 preferentially bind to hypomethylated LINE-1s (Figure 1).

DISCUSSION

Previously, we demonstrated that AGO2 represses genes containing hypomethylated LINE-1s. Here, we demonstrate that AGO1 and AGO4 regulate genes containing LINE-1s in an opposite manner as that of AGO2. Furthermore, AGO1 and AGO4 bind to different LINE-1 loci from AGO2. Whereas AGO2 and AGO3 bind to hypomethylated LINE1s, AGO1 and AGO4 bind to LINE-1s that have methylation levels similar to the entire genome.

The AGO2-mediated repression of genes containing hypomethylated LINE-1s is mediated by LINE-1 RNA. When intragenic LINE-1s are hypomethylated, LINE-1 RNA is transcribed; the LINE-1 RNA then forms a complex with pre mRNA and AGO2, resulting in

mRNA depletion. The down-regulation of AGO1 and AGO4 prevents the up-regulation of expression. It is possible that AGO1 and AGO4 prevent LINE-1 transcription. Consequently, the presence of AGO1 and AGO4 at normally methylated LINE-1s prevents intragenic LINE-1 RNA production, and AGO1 and AGO4 consequently limit LINE-1 RNA from interfering with gene transcription. AGO3 binds to hypomethylated LINE-1s, similar to AGO2. Nevertheless, the depletion of AGO3 did not alter gene expression. The similar roles between AGO2 and AGO3 and the opposite roles between AGO1 and AGO2 were demonstrated previously with regard to the regulatory roles of mononucleotide A repeats[3]. We speculate that the AGO2 and AGO3 functions are redundant; thus, the depletion of AGO3 was compensated for by AGO2.

In a variety of living organisms, AGO1 has the function of promoting heterochromatin; in plants, AGO4 promotes DNA methylation. It is possible that AGO1 and AGO4 in humans may be associated with heterochromatin and methylation-associated LINE-1s, with AGO1 and AGO4 binding to transcription-inactiveLINE-1s.

In conclusion, we explored whether AGO proteins play a different role in controlling gene expression. AGO2 repress genes containing hypomethylated LINE-1. AGO3 may be redundant to AGO2, while AGO2 plays a major role. The AGO1 and AGO4 roles are distinctive from AGO2 and have opposite effects in regulating gene expression. It is interesting to further explore the mechanisms underlying these differences.

*Acknowledgement*s: This work was supported by "CU Research Cluster: 2014 Ratchadapisek Sompoch Endowment Fund, Chulalongkorn University". Piyapat Pin-on is supported by Chula Dusadee Pipat Fund, Chulalongkorn University.

REFERENCES

1. Zhou X, Guo H, Chen K, Cheng H, Zhou R (2010) Identification, chromosomal mapping and conserved synteny of porcine Argonaute family of genes. *Genetica* **138**, 805-12.

2. Sasaki T, Shiohama A, Minoshima S, Shimizu N (2003) Identification of eight members of the Argonaute family in the human genome. *Genomics* **82**, 323-30.

3. Aporntewan C, Pin-on P, Chaiyaratana N, Pongpanich M, Boonyaratanakornkit V, Mutirangura A (2013) Upstream mononucleotide A-repeats play a cis-regulatory role in mammals through the DICER1 and Ago proteins. *Nucleic acids research* **41**, 8872-85.

4. Wei KF, Wu LJ, Chen J, Chen Yf, Xie Dx (2012) Structural evolution and functional diversification analyses of argonaute protein. *Journal of cellular biochemistry* **113**, 2576-85.

5. Li W, Liu M, Feng Y, Xu Y-F, Che J-P, Wang G-C, Zheng J-H, Gao H-J (2013) Evaluation of Argonaute protein as a predictive marker for human clear cell renal cell carcinoma. *International journal of clinical and experimental pathology* **6**, 1086.

6. Chen Z, Lai T-C, Jan Y-H, Lin F-M, Wang W-C, Xiao H, Wang Y-T, Sun W, et al. (2013) Hypoxia-responsive miRNAs target argonaute 1 to promote angiogenesis. *The Journal of clinical investigation* **123**, 1057-67.

7. Li L, Yu C, Gao H, Li Y (2010) Argonaute proteins: potential biomarkers for human colon cancer. *BMC cancer* **10**, 38.

8. Yoo NJ, Hur SY, Kim MS, Lee JY, Lee SH (2010) Immunohistochemical analysis of RNA-induced silencing complex-related proteins AGO2 and TNRC6A in prostate and esophageal cancers. *Apmis* **118**, 271-6.

9. Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**, 1833-7.

10. Pal-Bhadra M, Leibovitch BA, Gandhi SG, Rao M, Bhadra U, Birchler JA, Elgin SC (2004) Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science* **303**, 669-72.

11. Jones L, Keining T, Eamens A, Vaistij FE (2006) Virus-induced gene silencing of argonaute genes in Nicotiana benthamiana demonstrates that extensive systemic silencing requires Argonaute1-like and Argonaute4-like genes. *Plant physiology* **141**, 598-606.

12. Braun JE, Huntzinger E, Izaurralde E. The role of GW182 proteins in miRNA-mediated gene silencing. Ten Years of Progress in GW/P Body Research: Springer; 2013. p. 147-63.

13. Chen S, Chahar HS, Abraham S, Wu H, Pierson TC, Wang XA, Manjunath N (2011) Ago-2-mediated slicer activity is essential for anti-flaviviral efficacy of RNAi. *PloS one* **6**, e27551.

14. Shekar PC, Naim A, Sarathi DP, Kumar S (2011) Argonaute-2-null embryonic stem cells are retarded in self-renewal and differentiation. *Journal of biosciences* **36**, 649-57.

15. Wassenegger M. RNA-directed DNA methylation. Plant Gene Silencing: Springer; 2000. p. 83-100.

16. Aporntewan C, Phokaew C, Piriyapongsa J, Ngamphiw C, Ittiwut C, Tongsima S, Mutirangura A (2011) Hypomethylation of intragenic LINE-1 represses transcription in cancer cells through AGO2. *PLoS One* **6**, e17934.

17. Wanichnopparat W, Suwanwongse K, Pin-On P, Aporntewan C, Mutirangura A (2013) Genes associated with the cis-regulatory functions of intragenic LINE-1 elements. *BMC genomics* **14**, 205.

18. Aporntewan C, Mutirangura A (2011) Connection up-and down-regulation expression analysis of microarrays (CU-DREAM): a physiogenomic discovery tool. *Asian Biomed* **5**, 257.

19. Chalitchagorn K, Shuangshoti S, Hourpai N, Kongruttanachok N, Tangkijvanich P, Thong-ngam D, Voravud N, Sriuranpong V, et al. (2004) Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. *Oncogene* **23**, 8841-6.

20. Kuo MH, Allis CD (1999) Cross-Linking and Immunoprecipitation for Studying Dynamic Protein: DNA Associations in a Chromatin Environment. *Methods* **19**, 425-33.

21. Schmitter D, Filkowski J, Sewer A, Pillai RS, Oakeley EJ, Zavolan M, Svoboda P, Filipowicz W (2006) Effects of Dicer and Argonaute down-regulation on mRNA levels in human HEK293 cells. *Nucleic acids research* **34**, 4801-15.

TABLE LEGENDS

Table 1 CU-DREAM-X procedure. The table represents the intersection between the expression of each gene and genes containing intragenic LINE-1. These intersections are referred to as groups A through D. Group A consists of genes that are up- or down-regulated and contain intragenic LINE-1.Group B consists of genes that are not up- or not down-regulated and contain intragenic LINE-1. Group C consists of genes that are up- or down regulated and without intragenic LINE-1.Group D consists of genes that are not up- or down regulated and without intragenic LINE-1.

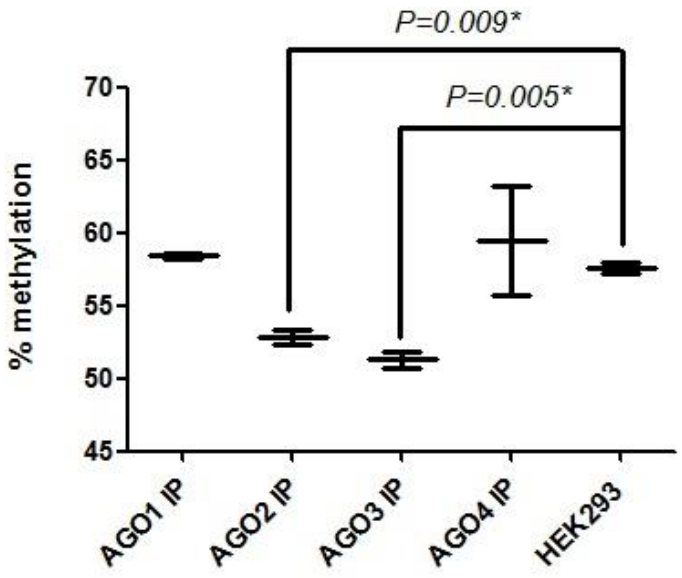
	Up- or down-regulated genes of experiment	Not Up- or not down-regulated genes of experiment
Genes containing LINE-1	Number of genes in 1st group (A)	Number of genes in 2nd group (B)
Genes without LINE-1	Number of genes in 3rd group (C)	Number of genes in 4th group (D)

Table 2 The results of CU-DREAM-X. The ORs, *P*-values and lower and upper 95% confidence intervals (95% CI) of Si-AGO1-4 are grouped in terms of up- and down-regulated gene expression.

		Si-AGO1			Si-AGO2			Si-AGO3			Si-AGO4		
		OR	*P*-value	95% CI	OR	*P*-value	95% CI	OR	*P*-value	95% CI	OR	*P*-value	95% CI
Up	LINE-1	0.55	4.67E-04	0.39-0.77	1.88	6.19E-04	1.30-2.70	0.78	4.45E-01	0.41-1.48	0.33	1.34E-08	0.22-0.50
Down	LINE-1	1.29	5.21E-02	1.00-1.67	0.82	4.57E-01	0.48-1.39	1.29	3.03E-01	0.79-2.11	1.27	7.99E-02	0.97-1.66

FIGURE LEGENDS

Fig. 1 ChIP and Q-COBRA-MP. Methylation percentage of AGO1 and AGO4 shows normality compared with the HEK293 cell line, whereas the methylation percentage of AGO2 and AGO3 shows hypomethylation compared with the HEK293 cell line.



VITA

Piyapat Pin-on is now candidate Ph.D. student in interdisciplinary program in Biomedical Sciences, Chulalongkorn University. He was born on April 27, 1982. After finished high school education at Taweethapisek School, Bangkok, in 1999, he has studied in Bachelor degree of Science (Genetics) at Chulalongkorn University until in 2003 after that he studied in Master of Science (Genetic Engineering) at Kasetsart University and graduated in 2007.

Over a period of six years of being a doctoral student in the Center of Excellence in Molecular Genetics of Cancer and Human Diseases under Professor Apiwat Mutirangura, head and his adviser, he has received a great opportunity and good advice for research from his adviser.

He has been supported by Chula Dusadee Pipat Fund, Chulalongkorn University for research on the topic of Argonaute 4 proteins mediate small-RNA-guided de novo methylation in human cells. In the first part of his thesis was became 1 international publication in Science Asia journal on the topic Distinctive roles of human Argonaute1, 2, 3 and 4 proteins on LINE-1 methylation and regulation of intragenic LINE-1-containing genes. Piyapat Pin-on, Apiwat Mutirangura.

On cooperation between Professor Apiwat Mutirangura and Professor J. Silvio Gutkind from, National Institutes of Health, USA. He had a great opportunity to go aboard for doing short term research project for six months at National Institute of Dental and Craniofacial Research, National Institutes of Health, Marryland, USA, under supervision of Professor J. Silvio Gutkind and supported by Chula Dusadee Pipat Fund, Chulalongkorn University.

After finishing doctoral qualifying examination (February 9, 2011) and thesis proposal defense (April 25, 2012). He finished his thesis defense on January and will be graduated by the end of in the same year.