HP1BP3, a chromatin retention factor for co-transcriptional microRNA processing

APPROVED BY SUPERVISORY COMMITTEE

_____

Qinghua Liu, Ph.D.

_____

Benjamin Tu, Ph.D.

_____

Kim Orth, Ph.D.

_____

Michael Roth, Ph.D.

DEDICATION

To my parents and my grandparents for their continual support, guidance, and love.

HP1BP3, A CHROMATIN RETENTION FACTOR FOR CO-TRANSCRIPTIONAL PROCESSING OF MICRORNA

by

HAOMING LIU

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2016

Acknowledgement

First and foremost, I would like to thank my mentor, Dr. Qinghua Liu for all of his guidance and support throughout my dissertation research. His persistence and unwavering enthusiasm set an excellent example of what a scientist ought to be. From him, not only did I receive training as a biochemist, but I also learned how to face difficulties in life and tackle tough problems with determination and hard work. I am grateful to have had the opportunity to grow both as a person and as a scientist in Dr. Qinghua Liu's lab, and I will benefit from his wisdom and knowledge for the rest of my life.

I would like to thank my thesis committee members, Dr. Benjamin Tu, Dr. Kim Orth, and Dr. Michael Roth. They have always been both encouraging and resourceful in the past five years. Their advices have played a critical role in the completion of my thesis. I would also like to thank Sara and Frank Mcknight Foundation for supporting me financially through its pre-doctoral fellowship.

I would like to express my gratitude to all past and present members of Dr. Qinghua Liu's lab. I would like to thank Dr. Chunyang Liang, for both initially discovering the role of histone H1 in pri-miRNA processing, and his later collaboration and guidance to finish my current thesis project on HP1BP3. I would like to thank Dr. Xiong Ke for working with me in *Drosophila* H1 and initial research on HP1BP3. I would like to thank Dr. Zhiqiang Wang, Dr. Yibing Wang, Kejia Wu, and the rest of the lab for their assistance and advices, as well as making the lab a great environment to work. Lastly, I would like to thank Shimeng Liu, for the happiness we shared in the past three years, and hopefully for many years to come.

My thesis project would not be complete without the help of many excellent collaborators. I would like to thank Dr. David Corey and everyone in his lab, especially Dr. Masayuki Matsui, for teaching me the basics of ChIP and RIP assays which I relied heavily in my thesis research. I would like to thank Dr. Ralf Kittler and Rahul K. Kollipara for analyzing large sequencing data involved in this research. I would like to thank Dr. Yunsun Nam and her lab for their role in elucidating the interaction between HP1BP3 and Microprocessor complex. I was fortunate to have worked with all of you in the pursuit of science.

Finally, I would like to offer my sincere gratitude for all my friends who helped me during my time at UTSW. I am truly lucky to have the chance to meet and know some amazing people, and I wish all of them happiness and success in their future endeavors.

HP1BP3, A CHROMATIN RETENTION FACTOR FOR CO-TRANSCRIPTIONAL PROCESSING OF MICRORNA

Publication No. _____

Haoming Liu

The University of Texas Southwestern Medical Center at Dallas, 2016

Mentor: QINGHUA LIU, Ph.D.

RNA interference (RNAi) is a post-transcriptional gene silencing mechanism found in all eukaryotic organisms. It is characterized by a family of small non-coding RNAs, either endogenous (in the case of microRNAs) or exogenous (in the case of siRNAs), that inhibits gene expression post-transcriptionally. MicroRNAs (miRNAs) are a family of ~21-nt cellular RNAs that govern numerous pathological and physiological processes by mediating translational repression and deadenylation/decay of cognate mRNA. Dysregulation of miRNA expression have been associated with various types of cancer and developmental diseases. Typically, primary (pri-)miRNA transcripts are processed by Drosha complex into precursor (pre-)miRNAs, and then by cytoplasmic Dicer complex into mature miRNAs. The processing of pri-miRNAs is the most highly regulated step in the miRNA biogenesis

pathway. Therefore, understanding the molecular mechanisms of pri-miRNA processing and its regulation represents a very important objective in the miRNA field.

Recent studies suggest that the Drosha-DGCR8 complex can be recruited to chromatin to catalyze co-transcriptional processing of primary microRNAs (pri-miRNAs) in mammalian cells. However, the molecular mechanism of co-transcriptional miRNA processing is poorly understood. Here, we find that HP1BP3, a histone H1-like chromatin protein, specifically associates with the Microprocessor and promotes global miRNA biogenesis in HeLa cells. Accordingly, chromatin immunoprecipitation (ChIP) studies reveal genome-wide co-localization of HP1BP3 & Drosha and HP1BP3-dependent Drosha binding to actively transcribed miRNA loci. Moreover, HP1BP3 exhibits a novel pri-miRNA binding activity and promotes the Drosha-pri-miRNA association *in vivo*. Knockdown of HP1BP3 compromises pri-miRNA processing by resulting in premature release of pri-miRNA transcripts from the chromatin. Taken together, these studies suggest that HP1BP3 promotes co-transcriptional miRNA processing *via* chromatin retention of nascent pri-miRNA transcripts. This work expands the functional repertoire of the H1 family of proteins and suggests a new concept of chromatin retention factor for widespread co-transcriptional miRNA processing.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

H. Liu, C. Liang, R. K. Kollipara, M. Matsui, X. Ke, B. Jeong, Z. Wang, K. Yoo, G. Yadav, L. Kinch, N. V. Grishin, Y. Nam, D. Corey, R. Kittler, Q. Liu. HP1BP3, a chromatin retention factor for co-transcriptional miRNA processing. Under review at Molecular Cell.

K. Zhou, **H. Liu**, S. Zhang, X. Huang, Y. Wang, G. Huang. B.D. Sumer, J. Gao. (2012) Multicolored pH-Tunable and Activatable Fluorescence Nanoplatform Responsive to Physiologic pH Stimuli. J. Am. Chem. Soc. *134(18),* 7803-7811.

J. M. Womick, **H. Liu,** A. M. Moran, (2011) "Exciton Delocalization and Energy Transport Mechanisms in R-Phycoerythrin". J. Phys. Chem. A. *115*, 2471-2482.

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

LIST OF TABLES

LIST OF APPENDICES

LIST OF DEFINITIONS

HP1BP3 – Heterochromatin Protein 1, Binding Protein 3

miRNA – microRNA

RNAi – RNA interference

dsRNA – double-stranded RNA

CHS - chalcone synthase

nt - nucleotide

WT – wildtype

siRNA – small interfering RNA

S2 cell – Schneider 2 cell

RISC – RNA-induced silencing complex

Ago – Argonaute protein

RLC – RISC loading complex

piRNA- piwi-associated RNA

pri-miRNA – primary microRNA

pre-miRNA – precursor microRNA

mRNA – messenger RNA

rRNA – ribosomal RNA

lncRNA – long noncoding RNA

miRISC- miRNA-induced silencing complexes

DGCR8 – DiGeorge Syndrome chromosomal region 8

Pol II – RNA polymerase II

CPA – cleavage and polyadenylation

NTD – N-terminal Domain

GD – globular domain

CTD – C-terminal domain

U2OS – human osteosarcoma cells

PNK – T4 polynucleotide kindase

sRNA-seq – massively parallel sequencing of small RNAs

ChIP – chromatin immunoprecipitation

RIP – RNA immunoprecipitation

FISH - fluorescent *in situ* hybridization

"Look at this stone cutter hammering away at his rock, perhaps a hundred times without as much as a crack showing on it. Yet at the hundred-and-first blow it splits in two. And I know it was not the last blow that did it, but all that had gone before."

---- Jacob Riis

CHAPTER ONE

INTRODUCTION AND BACKGROUND INFORMATION

Discovery of RNA interference pathway

Overview of RNA interference pathway

RNA interference (RNAi) is one of the most important discoveries in biology. The phenomenon of RNAi was first observed in plants, but later found in many eukaryotic organisms. The molecular mechanism of RNAi remained a mystery to researchers, until in 1998, Andrew Fire and Craig Mello discovered that double stranded RNA (dsRNA) induces gene silencing through degradation of complementary mRNA (Fire et al., 1998). For their work on elucidating the mechanism of RNAi, Andrew Fire and Craig Mello were awarded Nobel Prize in Physiology or Medicine in 2006. Since the discovery of RNAi, not only has it been used as a powerful genetic tool in countless biological experiment, but also uncovered multiple classes of small regulatory RNAs (such as microRNA and piwi-associated RNA) that play important roles in many fundamental biological and disease processes **(Figure 1-1)**. The cells initially evolved the mechanism of RNAi as a form of innate viral defense, and now, this very mechanism holds great promises in future biological research, as well as developing novel treatment of human diseases.

Observation of RNAi in Petunia flowers

The first documented phenomenon of RNAi, initially called "co-suppression" or "transgene-induced gene silencing (TIGS)", was observed in *Petunia hybrid* **(Figure 1-2)**. It

was discovered accidentally by botanist Richard Jorgensen and coworkers, who attempted to overexpress chalcone synthase (CHS) in petunia flowers to enrich its pigmentation (Napoli et al., 1990). CHS is a key enzyme in petunia flowers responsible for the synthesis of pigments that gave the flower either pink or purple color. By overexpressing this enzyme, they expected to see a darker colored flower. To their surprise, overexpression of CHS resulted in partial or complete white flowers, indicating the function of CHS is reduced. Upon further examination, they discovered that CHS mRNA level was reduced by ~50 fold in the transgenic flowers as compared to wild-type (WT) flowers (Napoli et al., 1990). However, the molecular mechanism underlying this "co-suppression" phenomenon would remain a mystery for many years to come.

Discovery of dsRNA as the key of RNAi

In 1998, Andrew Fire and Craig Mello reported a potent sequence-specific gene silencing phenomenon in *C. elegans* after injection of double-stranded RNA (dsRNA) (Fire et al., 1998). The effect of dsRNA injection on gene silencing was more potent than the injection of purified sense or antisense single-stranded RNA alone. Remarkably, the effect of gene silencing was found to transmit from injected nematodes to their progenies, sometimes for up to forty generations. A few molecules per affected cell was enough for the suppression of its target mRNA, suggesting that dsRNA plays a catalytic role in the suppression of target genes. This dsRNA induced gene silencing was termed RNA interference (RNAi), and hence for the first time, the mechanism of dsRNA-induced gene silencing began to unfold. This breakthrough discovery laid the theoretical foundation for

uncovering multiple classes of small regulatory RNAs, such as small interfering RNA (siRNA), microRNA (miRNA), and piwi-associated RNA (piRNA), which play important roles in many fundamental biological and disease processes. For their work, Andrew Fire and Craig Mello were awarded the Nobel Prize in Physiology or Medicine in 2006.

Discovery of small interfering RNA

Understanding of the entire RNAi machinery will take many more years to complete. Tuschl and coworkers initiated the first step by reconstituting dsRNA-initiated gene silencing *in vitro* using *Drosophila* embryo lysate (Tuschl et al., 1999). Shortly after, using a modified version of Tuschl's system, Zamore and his group demonstrated that addition of dsRNA resulted in cleavage of target mRNA in a sequence-specific manner (Zamore et al., 2000). Further, they identified that 21-23 nucleotide (nt) RNAs, named small interfering RNA (siRNA), were a product of long dsRNA originally used to induce RNAi (Zamore et al., 2000). Later studies demonstrated that this siRNA was sufficient to induce cleavage of target RNA, and synthetic version of this siRNA can be used to repress the expression of endogenous genes in mammalian cells (Elbashir et al., 2001). From this point forward, siRNA became a powerful tool in reverse genetics, and opened a new path for the development of novel therapeutic treatment for human diseases.

Identification of key proteins involved in RNAi pathway

Classical biochemical analysis played a critical role in identifying proteins involved in the RNAi process. In 2001, Bernstein et al. used a candidate approach to express a

number of RNase-III enzyme in *Drosophila* Schneider 2 (S2) cells, and identified Dicer-1 could process long dsRNAs into siRNAs (Bernstein et al., 2003). Two years later, another RNase-III enzyme that also can generate siRNAs, named Dicer-2, was purified from S2 extract (Liu et al., 2003). Genetic and biochemical studies later revealed that Dicer-2 is required for processing dsRNA into siRNAs, and Dicer-1 is responsible for catalyze the processing of precursor microRNA (pre-miRNA) into mature miRNAs (Jiang et al., 2005; Lee et al., 2004b).

The RNA-induced silencing complex, or RISC, is the catalytic engine of RNAi. Its identification and characterization relied heavily on chromatographic fractionation. In 2000, Hammond et al. established a cell-free system to reconstitute the RNAi-induced mRNA cleavage by pre-loading S2 cells with dsRNA (Hammond et al., 2000). Then, using chromatographic fractionation, they identified a ribonuclease complex comprised of both proteins and small RNA, which catalyzed sequence-specific cleavage of target mRNA (Hammond et al., 2001). Further fractionation revealed that Argonaute 2 (Ago2) protein is the catalytic enzyme of RISC, and is responsible for the cleavage of target mRNA under the guide of single-stranded siRNA (Hammond et al., 2001; Song et al., 2004). The Argonaute family of proteins is evolutionarily conserved from fly to human, and is essential in RNAi pathway.

The mechanism of RNA-mediated gene silencing

After years of research, the molecular mechanism of RNAi began to emerge. Initially, a long dsRNA (either expressed endogenously or exogenously introduced into cells)

was processed by Dicer into either miRNA (endogenous) or siRNA duplexes. The siRNA or miRNA is then loaded onto RISC by the RISC loading complex (RLC) (Hutvagner and Zamore, 2002; Liu et al., 2003; Liu et al., 2006; Pham et al., 2004). After RISC binding, the siRNA duplex unwinds and separating into guide strand and passenger strand. Like its name's sake, the guide strand will stay with RISC and eventually catalyze the sequence-specific cleavage of target mRNA; while the other passenger strand is discarded. The active RISC, with the help of guide RNA, targets complementary sequence in mRNA and mediates its sequence-specific cleavage (siRNA), or mRNA decay and translational suppression (miRNA). RISC is a multiple-turnover enzyme, and its specificity and potency make it invaluable tool in post-transcriptional regulation of gene expression. Today, siRNA is being widely applied for reverse genetic, while numerous biological functions are linked to miRNAs and piwi-associated RNAs (piRNA).

For the past five years, our lab focused on the biogenesis of miRNA. More specifically, we aimed to identify novel regulators involved in the processing of pri-miRNA. Our study would not have been possible without the outstanding research done by countless scientists who came before, their discovery have established a foundation for all future studies in the field of RNAi.

Biogenesis of microRNA

Overview of miRNA and its biogenesis

Ever since the discovery of RNAi, studies have revealed an increasing role of small non-coding RNA in diverse cellular regulatory pathways, such as protein synthesis, metabolism, and gene specific silencing. MicroRNA, first discovered in *C. elegans* (Lee et al., 1993), is a family of 21-25-nucleotide non-coding small RNA that post-transcriptionally down-regulate gene expression (Bartel, 2004; He and Hannon, 2004). Mutations that cause disruption of miRNA biogenesis are linked to a variety of developmental defects in all classes of metazoans, such as germ-line defect (He and Hannon, 2004), abnormal embryogenesis (Knight and Bass, 2001), developmental arrests (Park et al., 2002), and depletion of stem cells (Wienholds et al., 2003). There is increasing evidence to suggest that miRNA acts as a critical regulator of cell fate determination and cell death (Ambros, 2004; Bernstein et al., 2003; Ma et al., 2010; Thum et al., 2008), and abnormal miRNA level was found in human cancer cells (Carthew and Sontheimer, 2009; Denli et al., 2004; Gregory et al., 2004; Kim et al., 2009; Liu and Paroo, 2010). Therefore, understanding the mechanisms of miRNA biogenesis and regulation represent an important research objective.

Typically, miRNA biogenesis is catalyzed sequentially by two RNase III enzymes, Drosha and Dicer (Kim et al., 2009) **(Figure 1-3)**. In the nucleus, primary miRNA (pri-miRNA) transcripts are processed by the Microprocessor (Drosha-DGCR8) complex into precursor miRNAs (pre-miRNAs) (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004; Lee et al., 2003). The pre-miRNA is then transported out of nucleus into the cytoplasm where it is further cleaved by the Dicer-TRBP/PACT complex into duplex

miRNAs (Chendrimada et al., 2005; Haase et al., 2005; Hutvagner et al., 2004; Lee et al., 2006; Paroo et al., 2009). The duplex miRNA is then loaded into the miRNA-induced silencing complexes (miRISC) (Hutvagner and Zamore, 2002; Liu et al., 2004; Maniataki and Mourelatos, 2005; Yoda et al., 2010), where miRNA serve as a guide for Argonaute (Ago) and associated proteins to affect target silencing through translational inhibition and mRNA decay (Carthew and Sontheimer, 2009; Djuranovic et al., 2012; Guo et al., 2010; Iwakawa and Tomari, 2015).

Drosha/DGCR8 complex and pri-miRNA processing

MicroRNAs are first transcribed as long primary transcripts (pri-miRNAs), and then processed by two RNase III class of enzyme in its maturation to ~22nt miRNAs. Dicer, previously identified as a critical component of the RNAi pathway, is responsible for catalyzing ~70-nt pre-miRNA into mature miRNA. The enzyme that processes pri-miRNA transcript remained a mystery until Dr. Narry Kim's group first identified another RNase III–Drosha – to be the core nuclease that catalyzes the first step of miRNA processing (Lee et al., 2003). Drosha, previously known to be involved in the processing of ribosomal RNA (rRNA), was an RNase III enzyme that is found exclusively in the nucleus, making it a strong candidate for pri-miRNA processing. Lee et al. demonstrated that recombinant and immuno-purified Drosha protein can efficiently cleave pri-miRNA into pre-miRNA *in vitro*, and depletion of Drosha by siRNA resulted in accumulation of pri-miRNAs in cell nucleus. The processing of pri-miRNA is performed by double-stranded RNA cleavage at the base of the pri-miRNA stem-loop structure (Lee et al., 2003). **(Figure 1-3)**. The DiGeorge Syndrome

chromosomal region 8, or DGCR8, was later identified by Han et al. to be another essential component of pri-miRNA processing complex (Han et al., 2004; Han et al., 2006; Landthaler et al., 2004). Based on the *in vitro* pri-miRNA processing assay developed earlier, combined with careful domain mapping of both Drosha and DGCR8, it was found that DGCR8 is the protein that directly binds to pri-miRNA, allowing its double strand cleavage by Drosha (Gregory et al., 2004; Han et al., 2004; Han et al., 2006; Landthaler et al., 2004). Taken together, Drosha and DGCR constitute the core of the microprocessor complex, which is responsible for the processing of pri-miRNA into pre-miRNA.

Co-transcriptional processing of pri-miRNA

There is increasing evidence from recent studies that suggest the majority of pri-miRNA processing occur co-transcriptionally. The majority of pri-miRNA transcripts are transcribed by RNA polymerase II (Pol II) and have 5' cap and 3' poly(A) tail as pre-mRNA transcripts. In 2007, the first evidence of co-transcriptional processing was discovered by Dr. Narry Kim's group, the research found that cleavage of intronic pri-miRNA occurred before splicing (Kim and Kim, 2007). Soon after, using chromatin immunoprecipitation (ChIP), Morlando et al. showed that Drosha is recruited to the site of transcription of both non-intronic and intronic pri-miRNAs. The chromatin localization of Drosha was Pol II dependent, suggesting that transcription and pri-miRNA processing is a coordinated event (Morlando et al., 2008). Processing of intergenic miRNA was also found to affect downstream transcription (Ballarino et al., 2009). Moreover, co-transcriptional pri-miRNA processing is also supported by native elongating transcript sequencing (mNET-seq) studies

for mammalian chromatin (Nojima et al., 2015). However, the molecular mechanisms of co-transcriptional miRNA processing are poorly understood.

Around the same time, it was found that artificial retention of pri-miRNA transcripts on the chromatin greatly enhanced production of mature miRNA *in vivo* (Pawlicki and Steitz, 2008). In their study, Pawlicki and Steitz found that by artificially retaining the pri-miRNA transcripts at the site of transcription, pri-miRNAs are processed more efficiently by Drosha-DGCR8 complex. Other methods of chromatin retention, such as flanking the pri-miRNA transcript with exons, also enhanced its processing efficiency (Pawlicki and Steitz, 2008). Excess pri-miRNAs that are not retained on chromatin were found to form nuclear foci, which are inaccessible by the microprocessor (Pawlicki and Steitz, 2008). This study indicates that processing of pri-miRNA is most efficient before its release from chromatin, but also proposed a new mechanism in which co-factors can regulate processing activity of Drosha-DGCR8 complex by modulating the chromatin retention time of pri-miRNA.

Histone H1 and HP1BP3

Histone H1 and its family of proteins

In eukaryotic cells, genomic DNA is packed into large nucleoprotein complex known as chromatin. The fundamental building block of chromatin, the nucleosome, consists of approximately 147 bp of DNA wrapped around a core histone particle, which is composed of four core histone proteins, H2A, H2B, H3, and H4 (Kornberg, 1974) **(Figure 1-4)**. Short stretches (~20-80 nt) of DNA, called linker DNA, connects the nucleosomes and gives the chromatin a structure resembling the "beads on a string"(Thoma et al., 1979). Histone H1 protein, also known as linker histone, binds to the linker DNA as it enters or exits the nucleosome. Thus, histone H1 plays an essential role in the establishment and maintenance of higher order chromatin structure. The packing of chromatin structure is a dynamic process, allowing the condensed DNA molecule to be accessible for essential cellular processes, such as DNA replication, transcription, and repair.

Histone H1 is a relatively small protein (~20kDa). It consists of three domains – a amino (N)-terminal tail domain (~20-35 aa), a central globular domain (GD) (~70 aa), and carboxyl (C)-terminal lysine rich domain (~100 aa). The secondary structure of the globular domain (a winged-helix motif) is highly conserved among the histone H1 subtypes (Hergeth and Schneider, 2015). The C-terminal domain (CTD) is largely unstructured, but studies found that CTD may adopt a folded pattern upon binding to DNA molecules (Roque et al., 2005).

The role of histone H1 was long thought to be purely structural. It is seen as a rigid component of chromatin that participates mainly in transcriptional suppression (Schlissel and

Brown, 1984). However, later studies revealed that Histone H1 is a highly diverse class of histones that dynamically control many important cellular process, such as modulating gene expression and DNA repair (Bustin et al., 2005; Hergeth and Schneider, 2015). Intriguingly, the H1 family is the fastest evolving among all histone families, having much higher sequence variability between different species compared to evolutionarily conserved core histones. It has expanded from single H1 in yeast or fly to eleven H1 subtypes in mammals (Hergeth and Schneider, 2015). In vertebrates, knockout of a single H1 subtype usually produces no pronounced phenotype. Human H1 family includes seven somatic variants (H1.1, H1.2, H1.3, H1.4, H1.5, H1.0, H1x), three testis variants (H1t, H1T2, HILS1), and one ovary variant (H1oo). These H1 subtypes exhibit different developmental and tissue-specific expression patterns and may have redundant as well as specific functions (Hergeth and Schneider, 2015).

Heterochromatin protein 1-binding protein 3

The Heterochromatin protein 1-binding protein 3 (HP1BP3), also known as HP1-BP74, was first identified as a binding protein of the heterochromatin protein HP1 (Hiragami and Festenstein, 2005; Le Douarin et al., 1996). HP1BP3 is found to be ubiquitously expressed in mice, and it binds to the nucleosomes and protects linker DNA from degradation by micrococcal nuclease (Garfinkel et al., 2015b; Hayashihara et al., 2010) **(Figure 1-5)**. Depletion of HP1BP3 in HeLa cells affected gene expression of roughly 40 genes (greater than 2-fold change), but has no effect on the global chromatin organization, consistent with previous finding that Histone H1 protein's structural role is often redundant

(Garfinkel et al., 2015b). HP1BP3-knockdout mice were found to have developmental deficiencies, characterized by general dwarfism and low bone mass (Garfinkel et al., 2015a).

HP1BP3 is 553 amino acid long, and contains three globular domains that share high sequence similarity to other members of histone H1 family **(Figure 1-6)**. The N-terminal domain (NTD) of HP1BP3 consists of 100 amino acids, it is rich in glutamine and predicted to be unstructured. The NTD is followed by three globular domains (GD), termed GD1, GD2, and GD3. Downstream of the three globular domains is the lysine rich and unstructured C-terminal domain (CTD). The domain organization, a globular domain flanked by unstructured NTD and CTD, highly resemble the typical histone H1 structure. However, unlike typical histones H1 proteins, HP1BP3 is much larger and contain not one but three globular domains. HP1BP3's resemblance to canonical H1 allows it to associate with chromatin and contribute to its structure and organization, while its triple-GD structure suggest that HP1BP3 may have unique biological functions compared to other H1 proteins.

In our study, we find that HP1BP3 promotes global miRNA biogenesis in human cells. Co-immunoprecipitation (IP) and Chromatin immunoprecipitation (ChIP) studies found that HP1BP3 specifically associates with Drosha/DGCR8 complex on chromatin, and facilitates the processing of pri-miRNA. Furthermore, we demonstrate that HP1BP3 exhibits a novel pri-miRNA binding activity and enhances co-transcriptional miRNA processing *via* chromatin retention of nascent pri-miRNA transcripts. Our study discovered a novel role of HP1BP3 in the regulation of global miRNA biogenesis, and suggests a new model in which co-transcriptional processing can be regulated by other histone H1 family of proteins through chromatin retention.

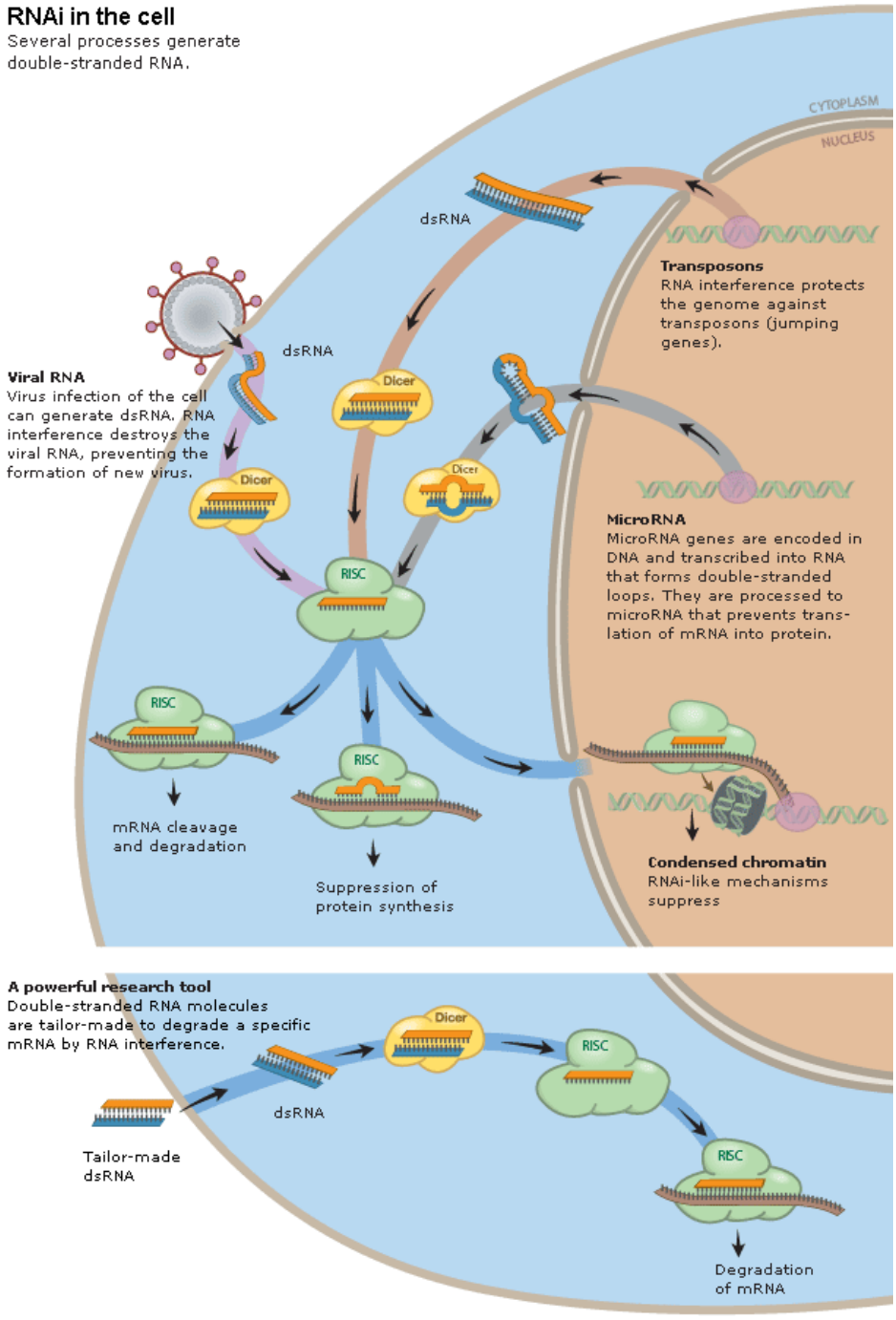

Figure 1-1 Overview of RNA Interference Pathway.
Image taken from Nobelprize.org. Copy right © The Nobel Committee for Physiology or Medicine at Karolinska Institute, 2006.

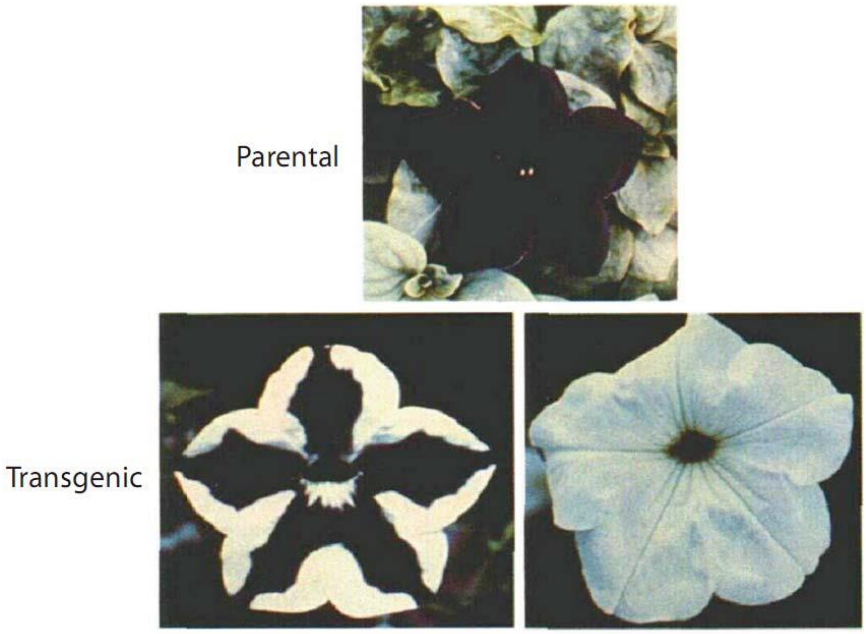

Figure 1-2. Observation of RNAi in Petunia flowers.
Partial or complete white flowers after ectopic expression of chalcone synthase, indicating the function of CHS is reduced in transgenic (bottom) petunia flowers compared to parental (top) flowers (Napoli et al., 1990).

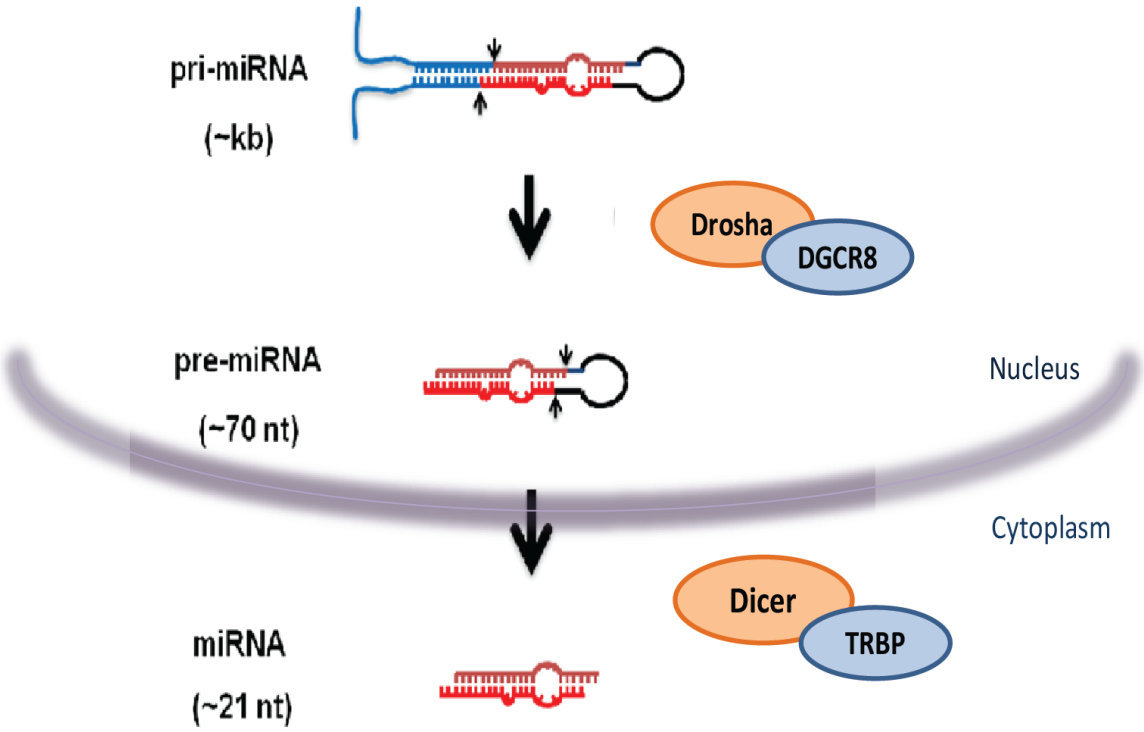

Figure 1-3. Overview of human miRNA biogenesis.
Pri-miRNA transcript is first processed by Drosha/DGCR8 complex into pre-miRNA. The pre-miRNA is then exported out of nucleus and into the cytoplasm where it is further processed by Dicer/TRBP into mature miRNAs.

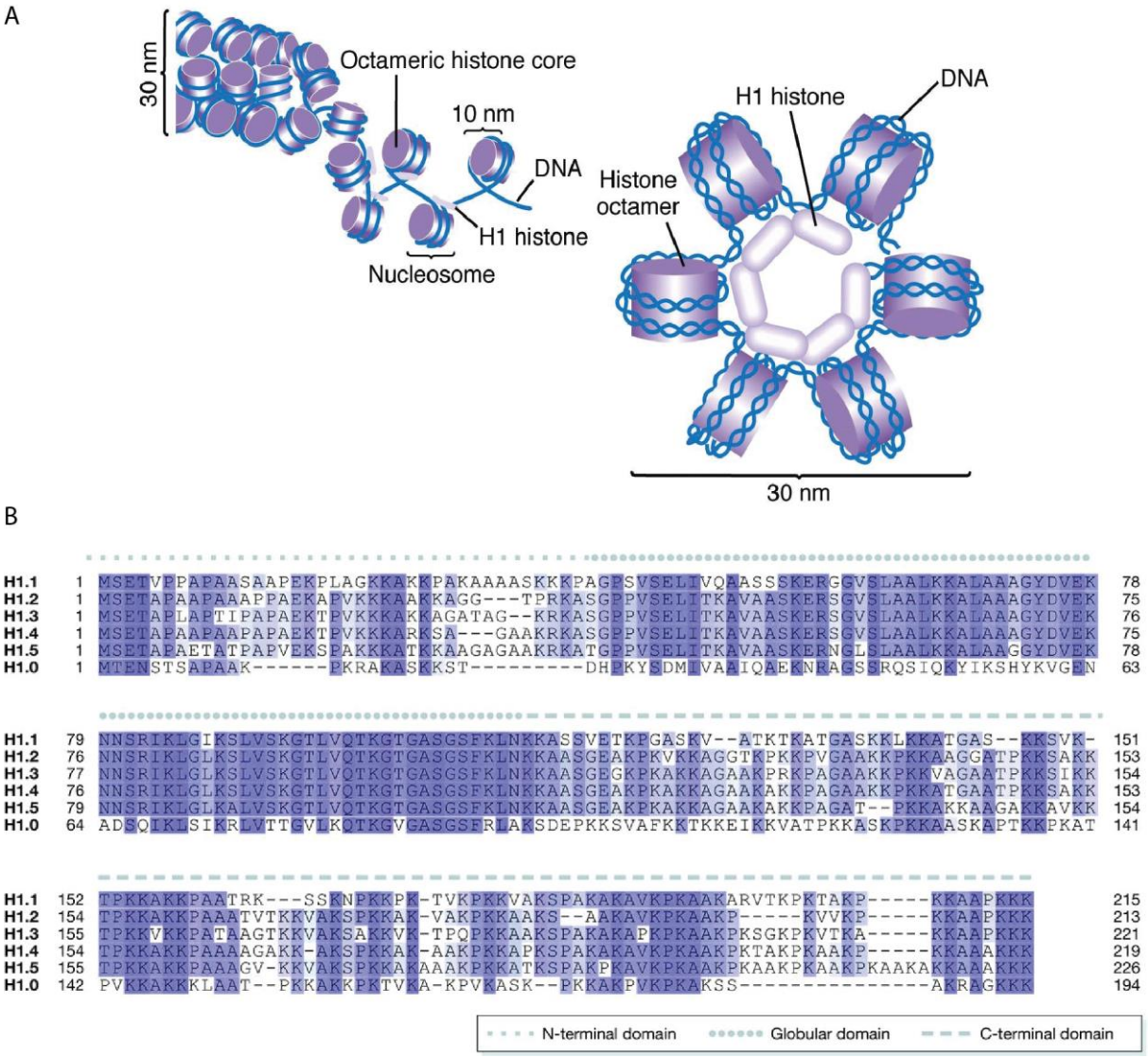

Figure 1-4. Histone H1 and its associated proteins.
(A) A graphic illustration of histones H1 and its role in chromatin compaction (Griffiths et al., 2004). **(B)** Sequence alignment of histone H1 isoforms in human, showing highly conserved globular domains. Conserved residue are highlight in blue, darker color indicates higher conservation (Hergeth and Schneider, 2015).

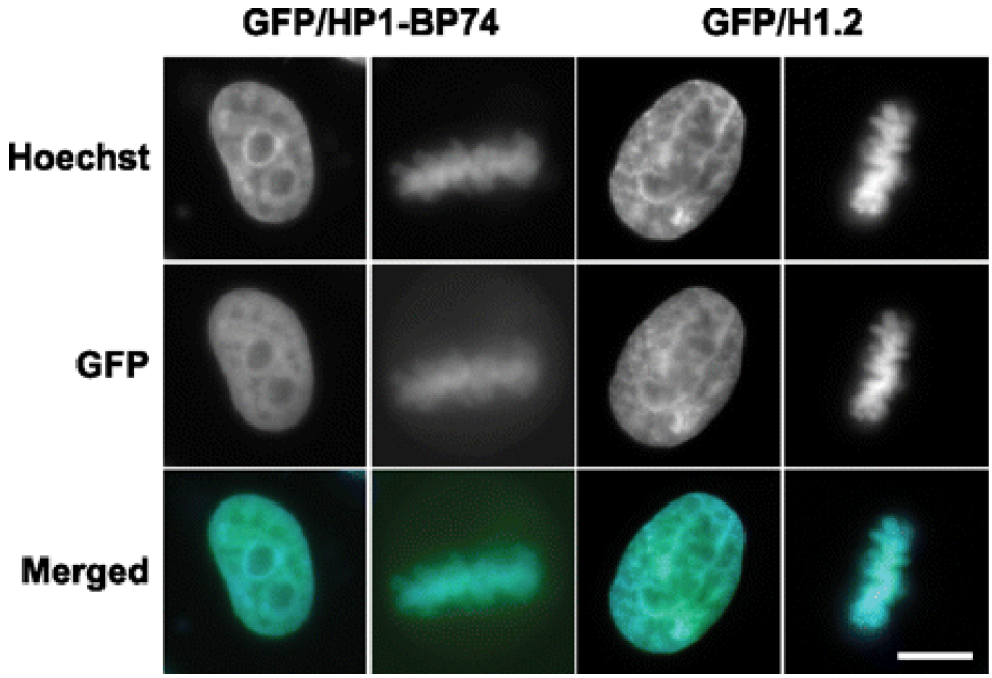

Figure 1-5. HP1BP3 is co-localized with chromatin throughout cell cycle.
HeLa cells were transfected with N-terminal GFP-fused HP1BP3 (HP1-BP74) and H1.2. GFP/HP1-BP74 was localized in nuclei in interphase and colocalized with chromosomes in mitosis(Hayashihara et al., 2010).

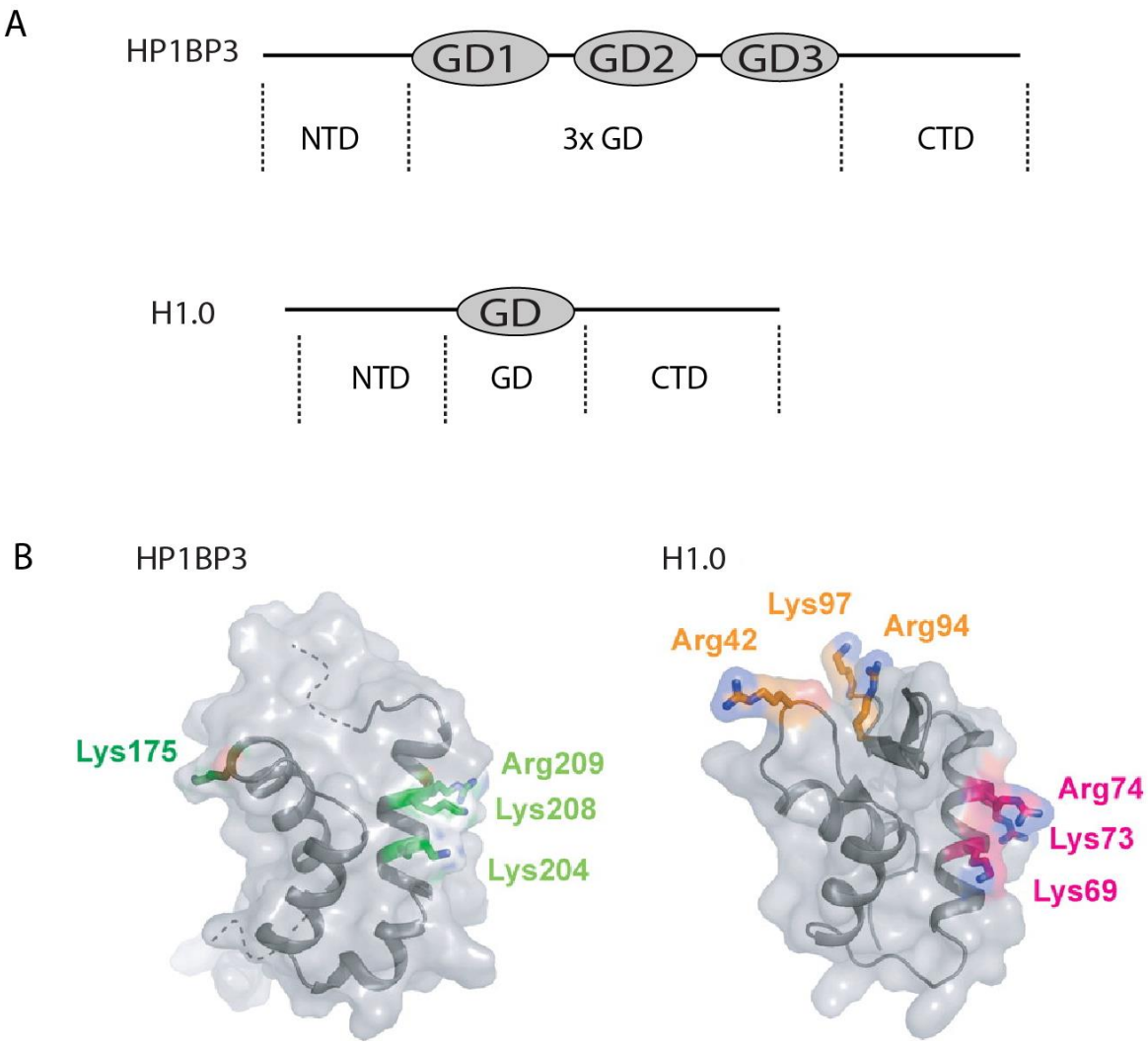

Figure 1-6. Structure of HP1BP3 protein
(A) Domain structure of HP1BP3, which is characterized by three globular domains flanked by unstructured N-terminal domain and lysine rich C-terminal domain. **(B)** 3-D structure of globular domain of HP1BP3 (left) and H1.0 (right), the predicted nucleosome-interacting residues are highlighted (Hayashihara et al., 2010).

CHAPTER TWO

HISTONE H1 PROMOTES MIRNA BIOGENESIS IN DROSOPHILA

*Note from author: The research content of this chapter was initiated by Dr. Chunyang Liang and Dr. Ke Xiong, and was completed as a collaboration between Dr. Chunyang Liang and myself. Although the research here was focused on fly H1, it laid the foundation for the identification of HP1BP3, a histone-H1 like protein, as a novel regulator of pri-miRNA processing in humans. The content of this chapter is included in the paper "**HP1BP3, a chromatin retention factor for co-transcriptional microRNA processing**"*

Abstract

MiRNAs are ~22-nt genome-encoded small RNAs that control target gene expression through translational repression and/or mRNA decay. Expression of miRNAs, which regulate diverse developmental and physiological processes, is highly regulated both transcriptionally and post-transcriptionally. Typically, the primary (pri-)miRNA transcripts are processed by nuclear Drosha complex into ~70-nt precursor (pre-)miRNAs, which are further cleaved by cytoplasmic Dicer complex into mature miRNAs. However, little is known about how the processing of miRNA is regulated. Our lab identified *Drosophila* histone H1 as a pri-miRNA-binding protein from the nuclear extract of S2 cells by chromatographic purification. In metazoans, histone H1 is known to bind to linker DNA between nucleosomes that contain core histones (H2A, H2B, H3, H4), and is involved in chromatin compaction and transcriptional repression. Unexpectly, *Drosophila* histone H1 can also function as a primary

miRNA (pri-miRNA) binding protein *in vitro and in vivo.* In *Drosophila* S2 cells, knockdown of histone H1 resulted in significant reduction of miRNAs and pre-miRNAs, while the levels of pri-miRNAs remained unaffected; suggesting that histone H1 regulates pri-miRNA processing.

Introduction

The majority of miRNAs are generated by sequential processing by two ribonuclease-III enzymes: Drosha and Dicer (Carthew and Sontheimer, 2009; Kim et al., 2009; Liu and Paroo, 2010). In the nucleus, the primary (pri)-miRNAs are transcribed by RNA polymerase II (Pol II) and processed by the Drosha/Pasha complex, also known as the microprocessor, into ~70-nt stem loop precursor (pre)-miRNA (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004; Lee et al., 2003). The pre-miRNAs are then exported to the cytoplasm with the help of Exportin 5-Ran-GTP (Lund et al., 2004; Yi et al., 2003). In the cytoplasm, pre-miRNAs are further cleaved into ~21-nt mature miRNAs by the Dicer-1 (Dcr-1)-Loqs-PB complex in *Drosophila melanogaster* (Forstemann et al., 2005; Jiang et al., 2005; Saito et al., 2005).

Global miRNA changes have been observed frequently during biological transitions, such as during embryonic development, stem cell differentiation, and tumorigenesis. As key regulators of genome activities, it is understandable that miRNA expression needs to be tightly controlled at transcriptional and post-transcriptional levels. Recent studies have just begun to uncover a series of post-transcriptional regulators of miRNA biogenesis, such as SMAD and Lin28 (Davis et al., 2008; Davis et al., 2010; Hagan et al., 2009; Heo et al., 2009;

Viswanathan et al., 2008). However, all of these miRNA regulators modulate the production of a small subset of miRNAs by recognizing specific sequences within pri-/pre-miRNAs (Davis et al., 2010; Heo et al., 2009; Trabucchi et al., 2009). It is likely that novel regulators of miRNA biogenesis, particularly those mediating global miRNA changes, remain to be identified.

We have recently identified *Drosophila* histone H1 as a pri-miRNA binding protein that promotes global miRNA biogenesis. Nucleosome is a fundamental unit of eukaryotic chromatin that is formed by wrapping DNA around the core histone octamer (2 copies each of H2A, H2B, H3, and H4). Histone H1 is known to bind the linker DNA region between nucleosomes and is involved in chromatin compaction and transcriptional repression. In our studies, we purified *Drosophila* histone H1 as a specific pri-miRNA binding protein. Knockdown of H1 expression in S2 cells caused significant reduction of many miRNAs, while the levels of pri-miRNAs were not affected. Based on these results, we conclude that histone H1 binds nascent pri-miRNA transcripts and promotes co-transcriptional processing of pri-miRNAs by the Drosha complex. Our study significantly expands the role of histone proteins, previously only known to have structural functions in chromatin compaction and transcriptional control. The newly discovered function of Drosophila H1 will lead not only to future studies of its role in pri-miRNA processing, but also opens new possibilities of histone proteins involvement in the regulation of other RNAs that are processed co-transcriptionally.

Result

Purification of Drosophila H1 as a specific pri-miRNA binding protein

Initially, our lab is interested in identification of novel factors involved in the biogenesis of miRNA. There are two hypotheses critical to our experimental design: 1) From an energy efficiency point of view, it is likely that key regulators of miRNA biogenesis are involved in the processing of pri-miRNA – the very first step of miRNA biogenesis after transcription. 2) If a regulator is involved in pri-miRNA processing, it must bind the pri-miRNA. Based on these two hypotheses, we employed an unbiased biochemical fractionation approach to identify novel pri-miRNA binding proteins in Drosophila S2 cells. S2 nuclear extract was fractionated using a Sp-Sepharose and Q-Sepharose columns **(Figure 2-1)**. Individual fractions were examined for the pri-miRNA-binding activity, and its protein content was resolved by SDS-PAGE followed by silver staining. After silver staining, the protein bands that correlated with pri-miRNA-binding activity were excised, digested by trypsin, and identified by Mass Spectrometry analysis. Using this method, we identified *Drosophila* histone H1 as a novel pri-miRNA binding protein from the nuclear extract of S2 cells **(Figure 2-1)**.

To address the question that any histone protein may bind to pri-miRNA in a non-specific manner, we further tested recombinant Drosophila H1's binding to pri-miRNA compared to other histone proteins using native gel shift assays. Surprisingly, Drosophila histone H1 binds preferentially to pri-miRNA compared to pre-miRNA and mature miRNA **(Figure 2-2A, B)**. Additionally, we also found that histone H1, but not core histone H2A, H3 or H4, efficiently bound pri-miRNA *in vitro*, further indicating that histone H1 binding to pri-miRNA is not due to non-specific interactions **(Figure 2-2C)**. Finally, RNA

immunoprecipitation (RIP) was used to show that Flag-H1 could specifically bind endogenous pri-miRNA transcripts *in vivo* **(Figure 2-3)**. In conclusion, we have identified that histone H1 is a novel and specific pri-miRNA binding protein, and a strong candidate for regulating the processing of pri-miRNA.

Drosophila H1 regulates pri-miRNA processing

To investigate the role of Drosophila H1 in the processing of pri-miRNA, we used siRNA-mediate knockdown to deplete Drosophila H1 in S2 cells. The level of pri-miRNA was measured using Taqman® pri-miRNA probes, and the level of pre-miRNA and mature miRNA is measured using Northern Blots. Initially, due to the strong binding exhibited between Drosophila H1 and pri-miRNA, we hypothesized that Drosophila H1 may be an inhibitor of pri-miRNA processing by competing with Drosha/Pasha complex for available pri-miRNA in the nucleus. However, in contrast to our expectations, northern blot analysis showed that knockdown of H1 expression decreased miRNA levels in S2 cells **(Figure 2-4A)**. Pri-miRNA level, on the other hand, either increased or remained unchanged after depletion of Drosophila histone H1 **(Figure 2-4B)**. The opposite trend between pri-miRNA and mature miRNA level after siH1 knockdown strongly suggests that Drosophila H1 plays a critical role in the processing of pri-miRNA. Removing Drosophila H1 severely disrupts the ability for Drosha/Pasha complex to catalyze the transition of pri-miRNA into pre-miRNA, leading to a decrease of product (pre- and mature miRNA) and an accumulation of reactant (pri-miRNA) in the miRNA biogenesis pathway.

Continuing work on fly Histone H1 in mammalian system

After the initial identification of Drosophila H1 as a novel regulator of pri-miRNA processing, attempts were made to further elucidate the molecular details of how histone H1 interacts with the microprocessor complex and facilitates the miRNA biogenesis. For example, we hypothesized that Drosophila H1 may recruit Drosha/Pasha complex to actively transcribed miRNA loci, or it may activate Drosha to catalyze pri-miRNA's processing. However, experiments designed to elucidate above hypothesis produced either negative or inconsistent results. Because there is only a single histone H1 isoform in Drosophila, it is difficult to separate its role in pri-miRNA processing from its other multitude of functions as an essential chromatin factor, which could explain the difficulties we encountered in our mechanistic studies. With the identification of HP1BP3 as the human homologue of fly H1 in pri-miRNA processing, efforts to elucidate molecularly how histone proteins are involved in miRNA biogenesis shifted to the human system. Unlike in Drosophila, there are eleven annotated H1 variants that are believed to have redundant as well as specific functions (Hergeth and Schneider, 2015). Our study in human system eventually lead to the discovery that HP1BP3, a histone H1-like protein, enhances pri-miRNA processing by enhancing the chromatin retention of nascent pri-miRNA transcripts. It is likely that Drosophila H1 may function in a similar manner to facilitate miRNA biogenesis in fly cells.

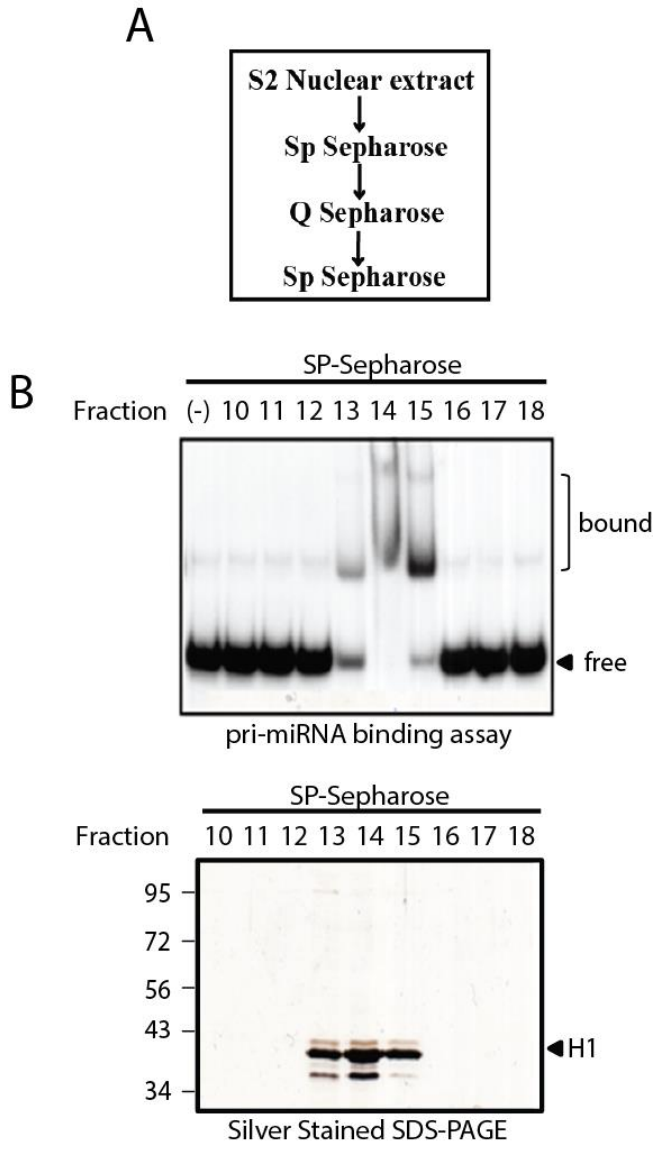

Figure 2-1. Purification of Drosophila H1 protein.
(A) A flow chart of purification procedure. **(B)** Following the final column, individual fractions were assayed in native gel shift assays for the pri-miRNA binding activity (top) and resolved by SDS-polyacrylamide gel (PAGE) followed by silver staining (bottom).

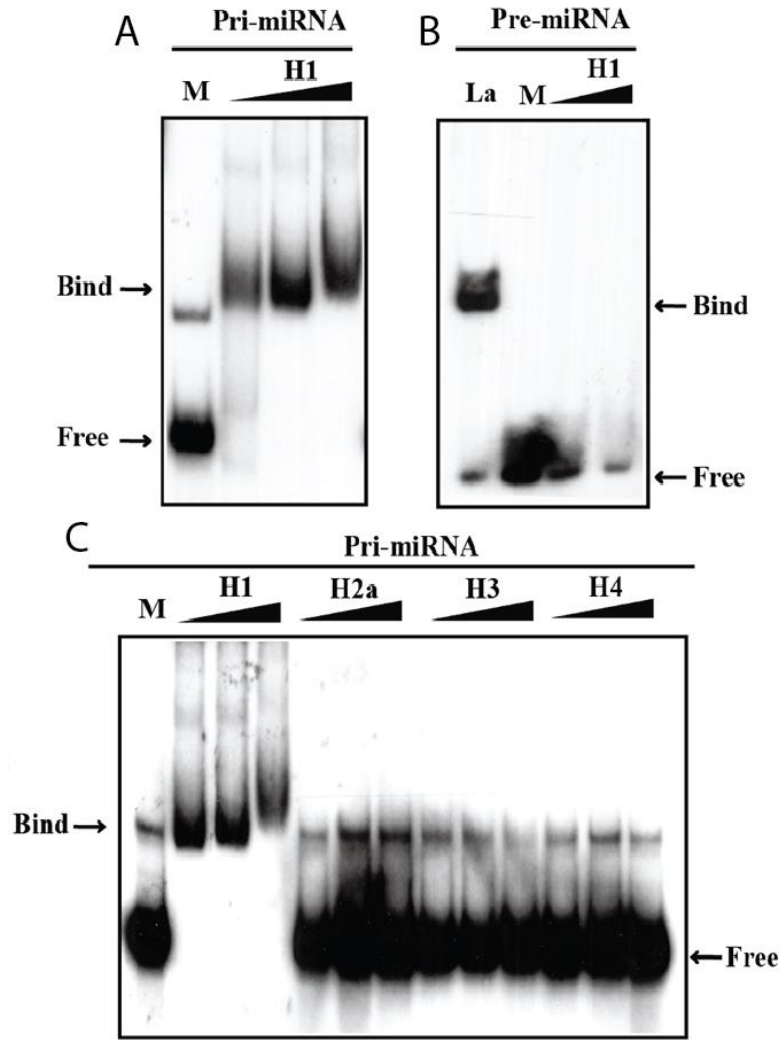

Figure 2-2. Drosophila H1 preferentially binds to pri-miRNA
(A) Native gel-shift assays were performed by incubating radiolabeled pri-miRNA in buffer alone or with 50, 100 and 200 nM of recombinant fly H1. **(B)** Native gel-shift assays were performed by incubating radiolabeled pre-miRNA with 100nM of protein La, in buffer alone, or with 100 and 200 nM of recombinant fly H1. **(C)** Native gel-shift assays were performed by incubating radiolabeled pri-miRNA in buffer alone or with 50, 100 and 200 nM of recombinant fly H1, H2a, H3 and H4. Arrows mark the positions of free and H1-bound pri-miRNA. (Courtesy of Dr. Chunyang Liang and Dr. Xiong Ke)

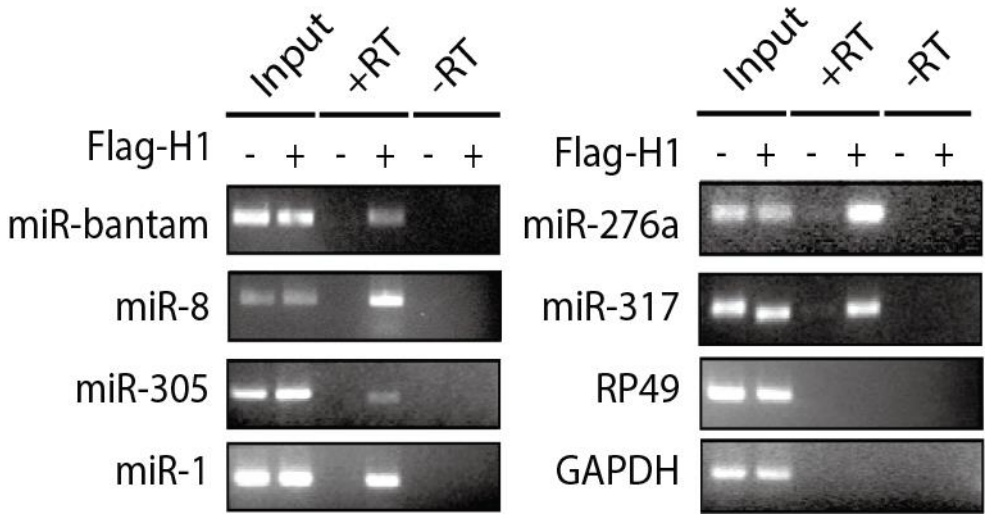

Figure 2-3 Drosophila H1 binds to pri-miRNA *in vivo*.
Two days after transfection of S2 cells with Flag- H1 construct, RNA immunoprecipitation (RIP) assays were performed using anti-Flag beads followed by RNA extraction and semi-quantitative RT-PCR to detect the H1-associated endogenous pri-miRNA, but not abundant RP49 or GAPDH mRNA. (Courtesy of Dr. Xiong Ke)

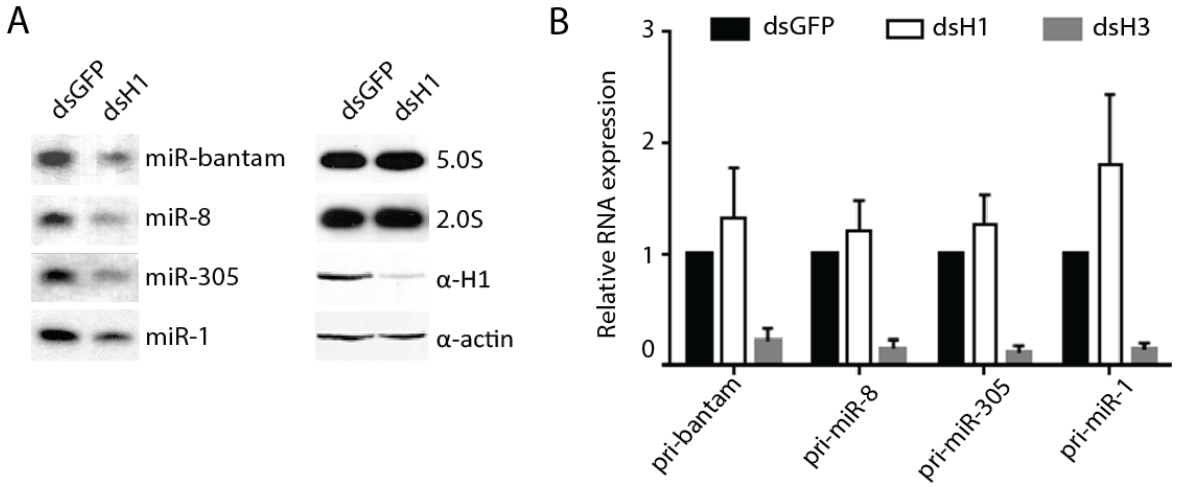

Figure 2-4 Drosophila H1 enhances miRNA biogenesis in Drosophila S2 cells.
(A) Following dsRNA-mediated knockdown of H1 in S2 cells, Northern blotting was performed to comparing the levels of mature miR-mature miR-bantam, miR-8, miR-305, and miR-1 with 5.0S and 2.0S rRNA as a loading controls. The efficiency of the knockdown was evaluated by western blot comparing the protein levels of histone H1 and Actin proteins.
(B) Quantitative analysis of pri-miRNA expression level by real time RT-qPCR in GFP, H1 and H3 dsRNA-treated S2 cells (triplicate experiments, data presented as Mean ± SD).

CHAPTER THREE

HP1BP3 IS A NOVEL GLOBAL REGULATOR OF PRI-MIRNA PROCESSING

*Note from author: The content of this chapter is included in the paper "**HP1BP3, a chromatin retention factor for co-transcriptional microRNA processing**"*

Abstract

The heterochromatin protein 1 binding protein 3 (HP1BP3), also known as HP1-BP74 (hereafter HP1BP3 for simplicity), has recently been identified as a histone H1-related protein that also binds to the nucleosomes and protects linker DNA from degradation by micrococcal nuclease (Garfinkel et al., 2015b; Hayashihara et al., 2010). Here, we find that HP1BP3, but not other histone H1 variants (H1.2, H1.5, H1.0, H1.X), specifically associates with the Drosha/DGCR8 complex and promotes global miRNA biogenesis in human cells. Additionally, ChIP and ChIP-seq analysis shows that HP1BP3 and Drosha co-localized at actively transcribed miRNA loci, indicating that HP1BP3 is an important co-factor to Drosha/DGCR8 and facilitates the processing of pri-miRNA.

Introduction

In the previous chapter, we have identified histone H1 as a novel regulator of pri-miRNA processing in Drosophila. In this chapter, we will continue to examine the role of human histone H1 and its associated proteins in the biogenesis of miRNA. In human, miRNA biogenesis is catalyzed sequentially by two RNase III enzymes, Drosha and Dicer (Kim et

al., 2009). In the nucleus, primary miRNA (pri-miRNA) transcripts are processed by the Microprocessor (Drosha-DGCR8) complex into ~70-nt stem-loop precursor miRNAs (pre-miRNAs) (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004; Lee et al., 2003). DGCR8 is a double-stranded RNA (dsRNA)-binding partner for Drosha that is essential for efficient and accurate processing of pri-miRNA (Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004). In the cytoplasm, pre-miRNAs are further cleaved by the Dicer-TRBP/PACT complex into mature miRNAs (Chendrimada et al., 2005; Haase et al., 2005; Hutvagner et al., 2004; Lee et al., 2006; Paroo et al., 2009), which are then assembled into the miRNA-induced silencing complexes (miRISC) (Hutvagner and Zamore, 2002; Liu et al., 2004; Maniataki and Mourelatos, 2005; Yoda et al., 2010). In mammalian miRISC, miRNA guides Argonaute (Ago) and associated proteins to effect target silencing through translational inhibition and mRNA decay (Carthew and Sontheimer, 2009; Djuranovic et al., 2012; Guo et al., 2010; Iwakawa and Tomari, 2015). Although key catalytic enzymes of the miRNA biogenesis pathway have been identified, little is known of how this process is regulated post-transcriptionally. Based on our previous finding in S2 cells, histone H1 family of proteins represent a strong candidate in pri-miRNA processing regulation. Using bioinformatics analysis, genetic, and large parallel sequencing techniques, we aim to identify the human histone that participates in the evolutionarily conserved role of facilitating the biogenesis of miRNA.

Histone H1 plays an essential role in the establishment and maintenance of higher order chromatin structure. The nucleosome -the basic unit of chromatin- consists of ~147 base pairs of DNA wrapping around an octamer comprised of two copies of core histones

H2A, H2B, H3 and H4 (Kornberg, 1974). Linker histone H1 binds to the nucleosomes as well as the ~20 to 80-nt DNA linking the nucleosomes, which facilitates the compaction of chromatin from the "beads on a string" configuration into a 30 nm chromatin fiber (Hergeth and Schneider, 2015). Intriguingly, the H1 family is the fastest evolving among all histone families, which has expanded from single H1 in yeast or fly to eleven H1 subtypes in mammals. Human H1 family includes seven somatic variants (H1.1, H1.2, H1.3, H1.4, H1.5, H1.0, H1x), three testis variants (H1t, H1T2, HILS1), and one ovary variant (H1oo). These H1 subtypes exhibit different developmental and tissue-specific expression patterns and may have redundant as well as specific functions (Hergeth and Schneider, 2015).

The heterochromatin protein 1 binding protein 3 (HP1BP3), also known as HP1-BP74 (hereafter HP1BP3 for simplicity), has recently been identified as a histone H1-related protein that also binds to the nucleosomes and protect linker DNA from degradation by micrococcal nuclease (Garfinkel et al., 2015b; Hayashihara et al., 2010). Here, we find that HP1BP3, but not canonical H1 variants, specifically associates with the Microprocessor and promotes global miRNA biogenesis in human cells. ChIP analysis reveals chromatin co-localization of HP1BP3 and Drosha across the genome. Based on all the above evidence, we conclude that HP1BP3 is a novel chromatin factor that facilitates global pri-miRNA processing. These studies will provide new insights into the regulatory mechanisms of the miRNA with significant impact on multiple fields in biology.

Result

HP1BP3, but not H1 variants, is involved in miRNA biogenesis in human cells

Unlike in Drosophila, where there is only one histone variant, there are as many as eleven annotated H1 variants in human, each believed to have redundant as well as specific functions (Hergeth and Schneider, 2015). This presents a challenge to identify the human histone H1 that regulates pri-miRNA processing in a similar manner as *Drosophila* H1, however, it also presents a unique opportunity to identify a specific histone H1 variants that specialize in enhancing pri-miRNA processing. HP1BP3 can be viewed as the twelfth human H1 variant because of its high sequence and functional similarity to classical H1 variants (Garfinkel et al., 2015b; Hayashihara et al., 2010). The H1 family of linker histones consists of a central winged helix globular domain (GD) surrounded by a variable amino (N)-terminal tail and a carboxyl (C)-terminal lysine rich tail (Hergeth and Schneider, 2015). By contrast, HP1BP3 contains three tandem GD repeats flanked by lysine rich sequences that could function similarly as the H1 tails. The three GD repeats of HP1BP3 resemble the GD of H1 variants, with the first GD showing 37% identity (53.33% similarity) to H1.0, 42% identity (64% similarity) to H1oo and 33% identity (56% similarity) to fly H1 **(Figure 3-1)**. Phylogenetic tree analysis suggests that the GDs of HP1BP3 are more closely related to the GD of fly H1 or H1oo than other human H1 variants **(Figure 3-2)**.

To identify the human orthologue of fly H1 in miRNA biogenesis, we performed siRNA-mediated knockdown of HP1BP3, H1.0, H1.2, H1.5, or H1x in HeLa cells. The efficiency of target mRNA knockdown was measured by quantitative RT-PCR **(Figure 3-3A)**. Consistent with phylogenetic tree analysis, knockdown of HP1BP3, but not knockdown of other H1 variants, significantly reduced the ratio of mature let-7a and miR-21 relative to

pri-let-7a and pri-miR-21, respectively **(Figure 3-3B, C)**. A similar result was obtained when HP1BP3 was depleted in human osteosarcoma (U2OS) cells **(Figure 3-4)**. These studies suggest that HP1BP3 is a histone H1-like chromatin protein that is specifically involved in miRNA processing in human cells.

HP1BP3, but not H1 variants, specifically associates with Microprocessor

Next, we examined the association between the Microprocessor and HP1BP3 or H1 variants by co-transfecting HeLa cells with the constructs expressing Flag-tagged Drosha-DGCR8 complex and GFP-tagged HP1BP3, H1.0, H1.2 or H1.5. Co-immunoprecipitation (IP) studies showed that Drosha-DGCR8 complex specifically associated with HP1BP3, but not other H1 variants **(Figure 3-5)**. Furthermore, we constructed bacterial artificial chromosome (BAC) transgenic HeLa cell lines that stably expressed a C-terminal GFP-tagged Drosha, HP1BP3 or H1.5 protein **(Figure 3-6A)**. The use of BACs that harbor large (>150 kb) genomic regions encompassing all exons, introns and regulatory regions allows for near physiological expression of transgenes as previously demonstrated (Kittler et al., 2005; Kittler et al., 2013; Poser et al., 2008). All three GFP-tagged proteins were localized in the nucleus and, as expected for chromatin factors, both H1.5-GFP and HP1BP3-GFP lighted up the mitotic chromosomes **(Figure 3-6B)**, suggesting that these fusion proteins were indeed functional. Reciprocal co-IP experiments indicated that HP1BP3-GFP, but not H1.5-GFP, specifically associated with Flag-tagged and endogenous Drosha-DGCR8 complex in the BAC transgenic cells **(Figure 3-7A, B)**. Due to the relative low salt condition (250mM) that is required to observe this association, we suspect that HP1BP3 may indirectly bind to

Drosha through RNA. We further tested this hypothesis by adding RNase in the co-IP procedure and found significant decrease of association between HP1BP3 and Drosha **(Figure 3-7C)**. This result was also consistent with later mechanistic studies presented in Chapter 3. Taken together, our co-IP data demonstrates that HP1BP3 is the unique human histone H1 isoform that interact with Drosha/DGCR8 complex in a RNA-dependent manner. We generated a series of truncated HP1BP3-GFP constructs to investigate whether the three GD domains of HP1BP3 are important for its association with the Microprocessor in vivo. Co-IP experiments showed that Flag-tagged Drosha-DGCR8 complex exhibited weak, medium and strong association, respectively, with the GFP-tagged HP1BP3-1GD (1-250), 2GD (1-330) and 3GD (1-420) fusion proteins **(Figure 3-8)**. These results suggest that the three GDs of HP1BP3 function cooperatively to mediate its interaction with the Microprocessor in vivo.

HP1BP3 is involved in the processing, not the transcription, of pri-miRNAs

To investigate whether HP1BP3 affected pri-miRNA processing, we compared the levels of pri-, pre-, and mature miRNAs for let-7a, miR-16, miR-21 and miR-23a between the control siRNA (siCtrl) and HP1BP3 siRNA (siBP3) treated HeLa cells **(Figure 3-9)**. Consistently, both pre-miRNAs and mature miRNAs were significantly reduced, but the pri-miRNA levels remained unchanged or increased upon depletion of HP1BP3 (3-9B, C). Since a rescue experiment is the gold-standard of any siRNA study, we generated a GFP-HP1BP3 construct that is resistant to our siRNA knockdown **(Figure 3-10A)**. Direct base-pairing between siRNA-guide and mRNA is disrupted with a silent mutation one in every three

bases. Based on our western blot **(Figure 3-10B)**, this mutant construct is resistant to siRNA mediated knockdown in Hela cells. The miRNA expression defect observed previously could also be rescued in HP1BP3-depleted cells by transfection of the siRNA-resistant HP1BP3 expression construct **(Figure 3-10)**. This confirms that reduction of pri-miRNA processing efficiency after HP1BP3 depletion was not due to an off-target effect. Additionally, HP1BP3 depletion did not affect Drosha or DGCR8 expression, suggesting that the mature miRNA reduction was not due to decreased expression level of the Microprocessor **(Figure 3-9)**. Since HP1BP3 is a histone-H1 like protein, and histone H1 is known to be involved in chromatin compaction and transcriptional control, great care were taken to confirm that the phenotype observed after HP1BP3 knockdown did not affect the transcriptional level to pri-miRNA. Using qRT-PCR to measure pri-miRNA level was insufficient because pri-miRNA level is affect by both pri-miRNA transcription and decay. To directly measure pri-miRNA transcription, we employed Pol II immunoprecipitation assay, which we find that HP1BP3 depletion had no effect on Pol II occupancy at the promoters of these miRNA genes **(Figure 3-11)**. Taken together, these observations suggest that HP1BP3 is specifically involved in the processing, not the transcription, of pri-miRNAs.

HP1BP3 promotes global miRNA biogenesis

We performed massively parallel sequencing of 15 to 40-nt small RNAs (sRNA-Seq) to compare global miRNA expression profile between the control and HP1BP3-depleted HeLa cells **(Figure 3-12)**. We analyzed two independent libraries for the siCtrl and siBP3 samples, respectively. Sequencing reads were mapped to unique sites in the human genome

(hg19) using Bowtie (v.2.2.5) (Langmead et al., 2009). Whereas the majority of small RNAs represent degradation fragments of rRNA, tRNA and snoRNA, ~0.2% of the sequencing reads correspond to *bona fide* miRNAs according to a 523 re-annotated human miRNA list (Fromm et al., 2015). Although most of small RNAs remained unchanged, 148 of 164 of expressed miRNAs were reduced, of which 68 miRNAs showed >2 fold reduction in the HP1BP3-depleted cells **(Figure 3-12 and Table 3-1)**. The remaining 16 miRNAs showed similar trend, but the results were inconclusive due to low counts. Thus, HP1BP3 promotes global miRNA biogenesis in human cells.

Genome-wide co-localization of HP1BP3 & Drosha at active miRNA

If HP1BP3 were involved in co-transcriptional miRNA processing, we expected that HP1BP3 should co-localize with the Microprocessor at the chromatin loci of actively transcribed miRNA genes. To test this hypothesis, we took advantage of our BAC transgenic cell lines to compare the chromatin binding maps of HP1BP3-GFP, H1.5-GFP and Drosha-GFP by ChIP and massively parallel sequencing (ChIP-Seq). We used SICER algorithm (v1.1) (Zhang et al., 2008) to define genome-wide occupied islands for these proteins from ChIP-Seq data. As reported for somatic H1 variants (Millan-Arino et al., 2014), H1.5-GFP showed a basal level of mostly uniform binding throughout the genome **(Figure 3-13)**. By contrast, HP1BP3-GFP was enriched at 9,149 islands across the genome, suggesting a different binding mode from that of H1.5-GFP. Similarly, Drosha-GFP was enriched at 5,814 islands, including many non-miRNA loci **(Figure 3-13)**. This result was consistent with previous reports that Drosha could also cleave mRNA with long hairpins and regulate other

biological processes (Gromak et al., 2013; Han et al., 2009; Kadener et al., 2009; Knuckles et al., 2012).

For annotation of Drosha, HP1BP3, or H1.5 islands, we used annotatePeaks function in Homer tools (Heinz et al., 2010) to assign islands relative to their specific positions in the genome. Furthermore, we used 5,000 base pair (bp) regions flanking the center of islands to generate density plots showing Drosha, HP1BP3, and H1.5 ChIP signals across all Drosha islands, HP1BP3 islands, or H1.5 islands **(Figure 3-14)**. These genomic analyses revealed a significant overlap in the overall binding sites between HP1BP3 & Drosha, but not between H1.5 & Drosha or between H1.5 & HP1BP3 (odds ratio=1.32, *p* value=2.97e-5). A similar result was obtained when we repeated this ChIP-Seq experiment to compare the chromatin binding maps of HP1BP3-GFP, H1.5-GFP and endogenous Drosha proteins using a ChIP-grade anti-Drosha antibody. The genome-wide chromatin co-localization of HP1BP3 and Drosha suggests an important functional link between these two proteins.

Previous studies have examined the chromatin binding of Drosha or DGCR8 proteins at a limited number of exogenous and endogenous miRNA genes (Gromak et al., 2013; Morlando et al., 2008; Pawlicki and Steitz, 2008). Thus, it is uncertain whether the majority of pri-miRNA transcripts undergo co-transcriptional miRNA processing. To address this question, we examined the chromatin binding of HP1BP3-GFP and Drosha-GFP at many expressed miRNA genes in HeLa cells. In our ChIP-Seq data, we found that HP1BP3-GFP bound to 42 expressed miRNA loci, and Drosha/Drosha-GFP bound to 18 expressed miRNA loci **(Table 3-2)**, which was consistent with previous finding that Drosha preferentially bound to actively transcribed miRNA loci (Morlando et al., 2008). Importantly, HP1BP3

also bound to thirteen (~72%) of eighteen Drosha-bound miRNA loci **(Table 3-2)**, indicative of a significant overlap of Drosha/HP1BP3 binding at active miRNA loci.

We suspected that our ChIP-Seq analysis underestimated the number of Drosha or HP1BP3-bound miRNA loci because individual ChIP assays could identify Drosha/HP1BP3-bound miRNA loci that were not detected by ChIP-Seq. Thus, we expanded individual ChIP analysis to examine the binding of HP1BP3-GFP and endogenous Drosha at the stem-loop regions of seventeen miRNA loci that were expressed at high or moderate levels (Figure 3-15). Notably, we chose one miRNA to represent every polycistronic miRNA gene that contains multiple expressed miRNAs in HeLa cells **(Figure 3-16)**. Thus, this series of ChIP analysis in effect covered a total of 39 active miRNA loci. Without an exception, both HP1BP3 and Drosha showed binding to all of these seventeen miRNA loci **(Figure 3-15)**. Therefore, our combined ChIP-Seq and ChIP analyses suggest genome-wide co-localization of HP1BP3 and Drosha at a total of 42 active miRNA loci **(Table 3-2)**. These results, together with our finding that HP1BP3 promotes global miRNA biogenesis **(Figure 3-12)**, suggest that HP1BP3 likely plays a key role in co-transcriptional processing of many if not all pri-miRNA transcripts.

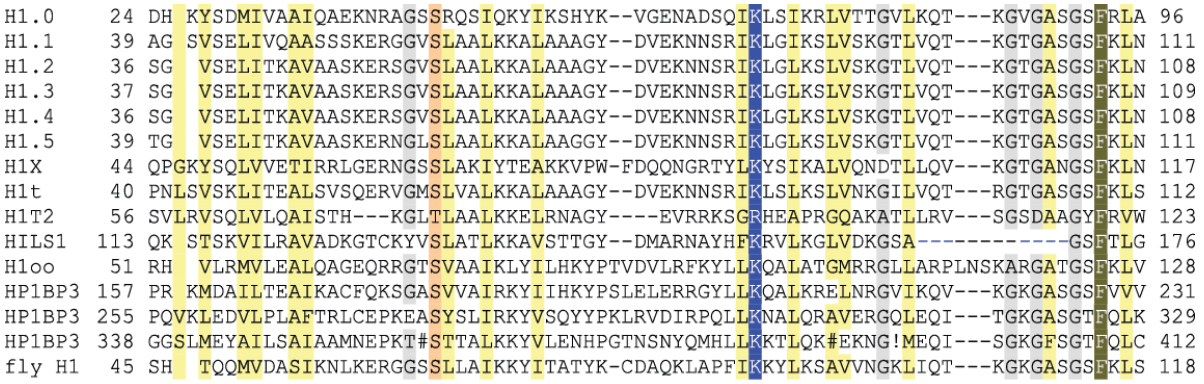

Figure 3-1. A multi-sequence alignment of human histone isoforms and Drosophila H1. Amino acid sequence of human histone isoforms and Drosophila H1 is aligned based on homology of globular domain (first globular domain is used in the case of HP1BP3). Residue positions are highlighted according to conservation: hydrophobic positions in yellow, charged residues blue, polar residues in orange, and glycine in gray. (Courtesy of Dr. Lisa Kinch)

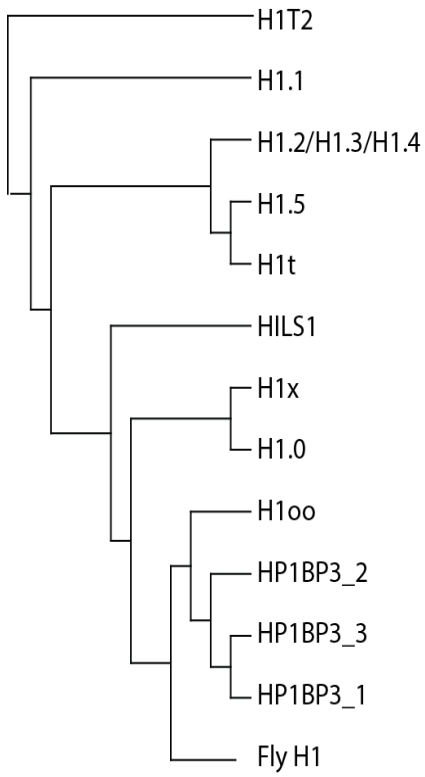

Figure 3-2. A phylogenetic tree of human histone variants and fly H1.
Phylogenetic tree showing that the relative evolutionary distance among HP1BP3, eleven human H1 variants and fly H1 based on the sequence homology of the central globular domains. (Courtesy of Dr. Lisa Kinch)

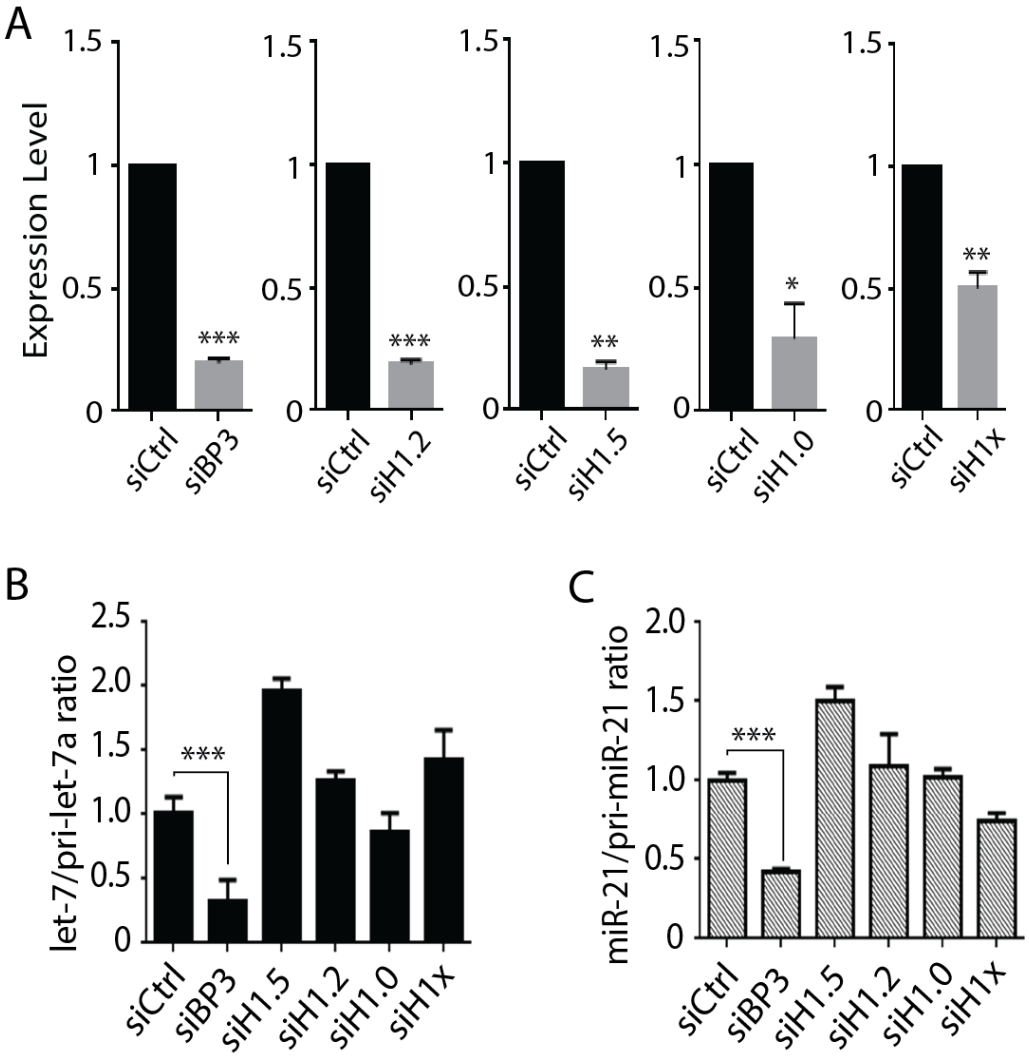

Figure 3-3 HP1BP3 is the specific human isoform that is involved in pri-miRNA processing.

(A) Following siRNA-mediated knockdown in HeLa cells, the mRNA level of HP1BP3, H1.0, H1.2, H1.5 or H1x was respectively measured by real time RT-qPCR. Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$). **(B, C)** Following siRNA-mediated knockdown of HP1BP3, H1.0, H1.2, H1.5, or H1x in HeLa cells, the ratio of let-7a/pri-let-7a **(B)** or miR-21/pri-miR-21 **(C)** was measured by Taqman qPCR, respectively (triplicate samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$).

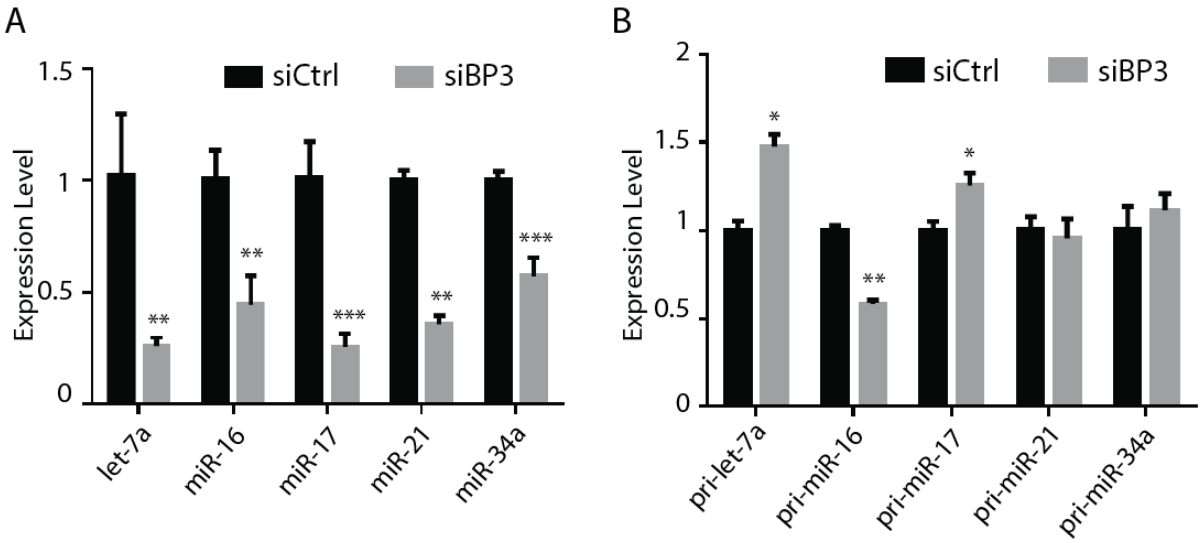

Figure 3-4 HP1BP3 regulates pri-miRNA processing in U2OS cells.
(A) Following siRNA-mediated knockdown of HP1BP3 in U2OS cells, the expression of let-7a, miR-16, miR-17, miR-21, and miR-34a were measured by Taqman qPCR, (triplicate samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** p< 0.001, ** p <0.01, * p<0.05). **(B)** Following siRNA-mediated knockdown of HP1BP3 in U2OS cells, the expression level of pri-let-7a, pri-miR-16, pri-miR-17, pri-miR-21, and pri-miR-34a were measured by Taqman qPCR, (triplicate samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** p< 0.001, ** p <0.01, * p<0.05).

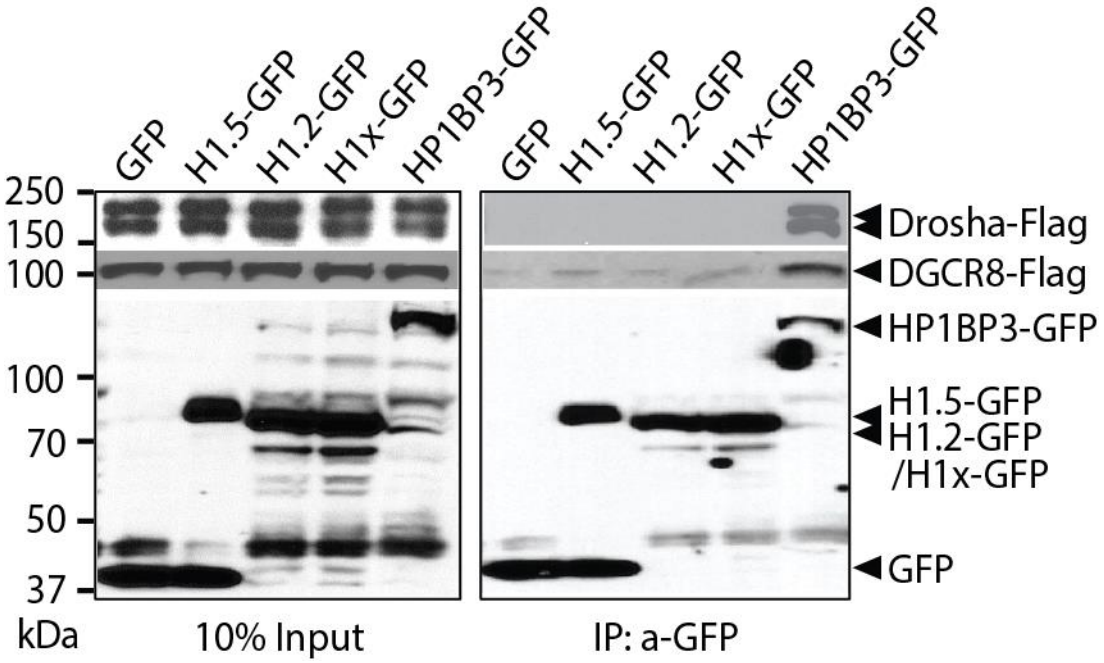

Figure 3-5 HP1BP3 associates with Drosha/DGCR8
HeLa cells were co-transfected with constructs expressing GFP-tagged HP1BP3, H1.2, H1.5 or H1x and Flag-tagged Drosha-DGCR8. Co-immunoprecipitation (co-IP) was performed using anti-GFP antibody. Western blotting (WB) was performed with anti-Flag and anti-GFP antibodies to visualize interaction between histone H1 isoforms and Drosha/DGCR8. Left panel shows 10% input sample, right panel shows co-IP result.

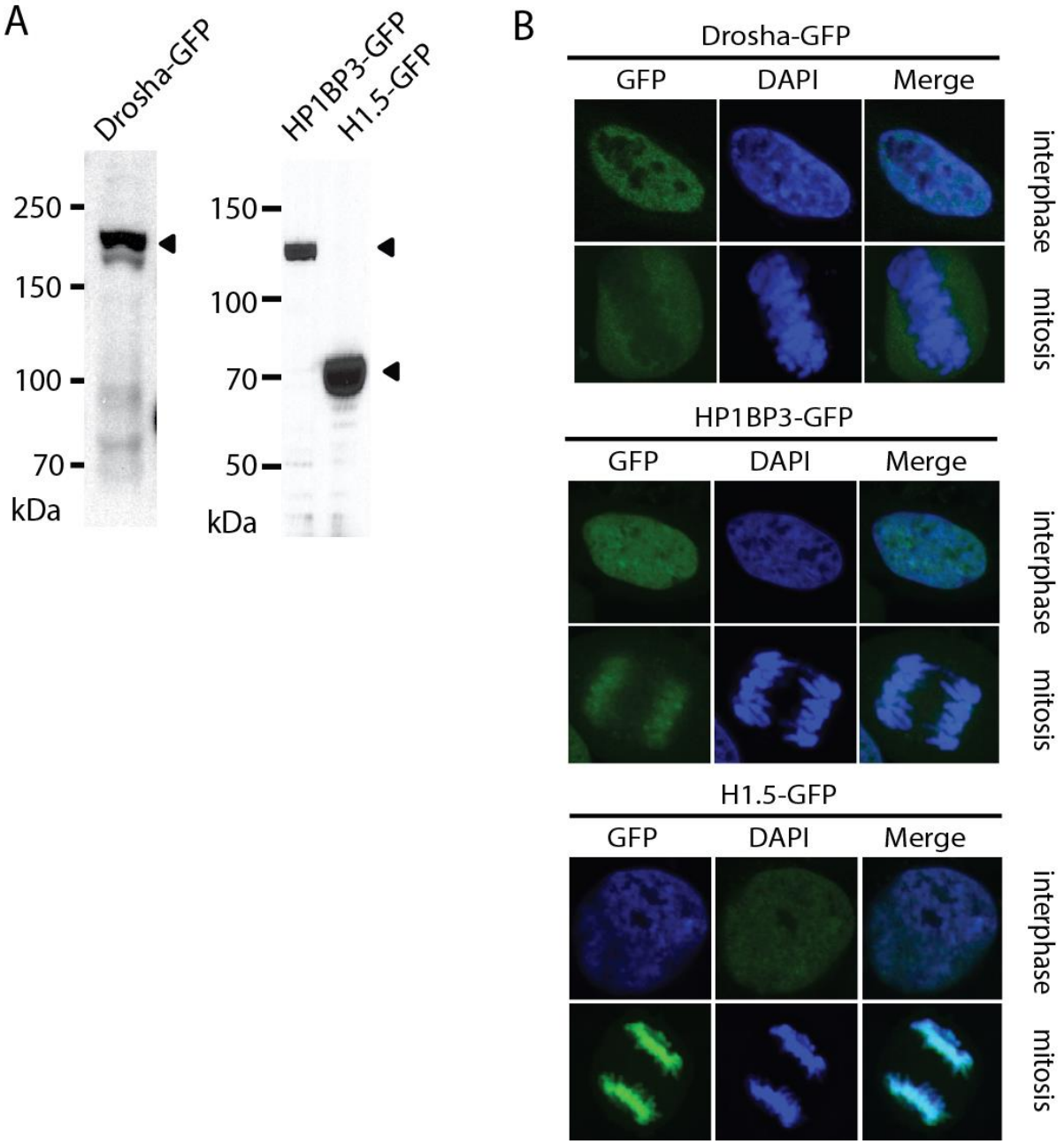

Figure 3-6 Generation of Bac-transgenic cell line.
(A) Western blotting was performed using anti-GFP antibodies to detect the expression of Drosha-GFP, HP1BP3-GFP and H1.5-GFP in different BAC-transgenic HeLa cells.
(B) Confocal microscopy images showing the localization of Drosha-GFP, HP1BP3-GFP and H1.5–GFP (green) in the interphase or mitosis phase of BAC transgenic HeLa cells. DNA was stained with DAPI (blue). (Courtesy of Dr. Zhiqiang Wang)

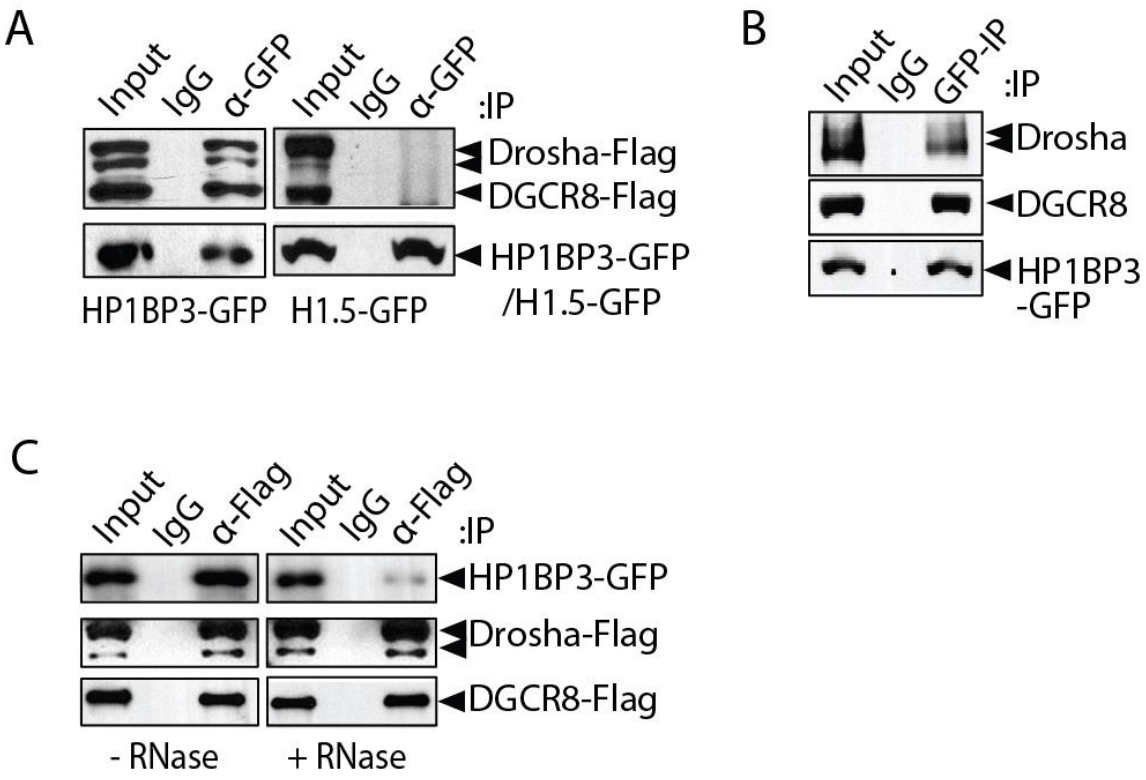

Figure 3-7 Co-immunoprecipitation of GFP-HP1BP3 and Drosha/DGCCR8 complex
(A) H1.5-GFP (left) or HP1BP3-GFP (right) BAC transgenic HeLa cells were transfected with Flag-tagged Drosha and DGCR8 constructs followed by co-IP using anti-GFP antibody. Western blotting was performed with anti-Flag and anti-GFP antibodies.
(B) From HP1BP3-GFP BAC transgenic HeLa cells, co-IP of HP1BP3-GFP was performed with anti-GFP antibody followed by Western blotting to detect HP1BP3-GFP and endogenous Drosha or DGCR8 proteins with the corresponding antibodies.
(C) HP1BP3-GFP BAC transgenic HeLa cells were transfected with Flag-tagged Drosha and DGCR8 constructs. Cell lysates were untreated or treated with RNase A before co-IP using anti-Flag antibody. Western blotting was performed with anti-GFP and anti-Flag antibodies.

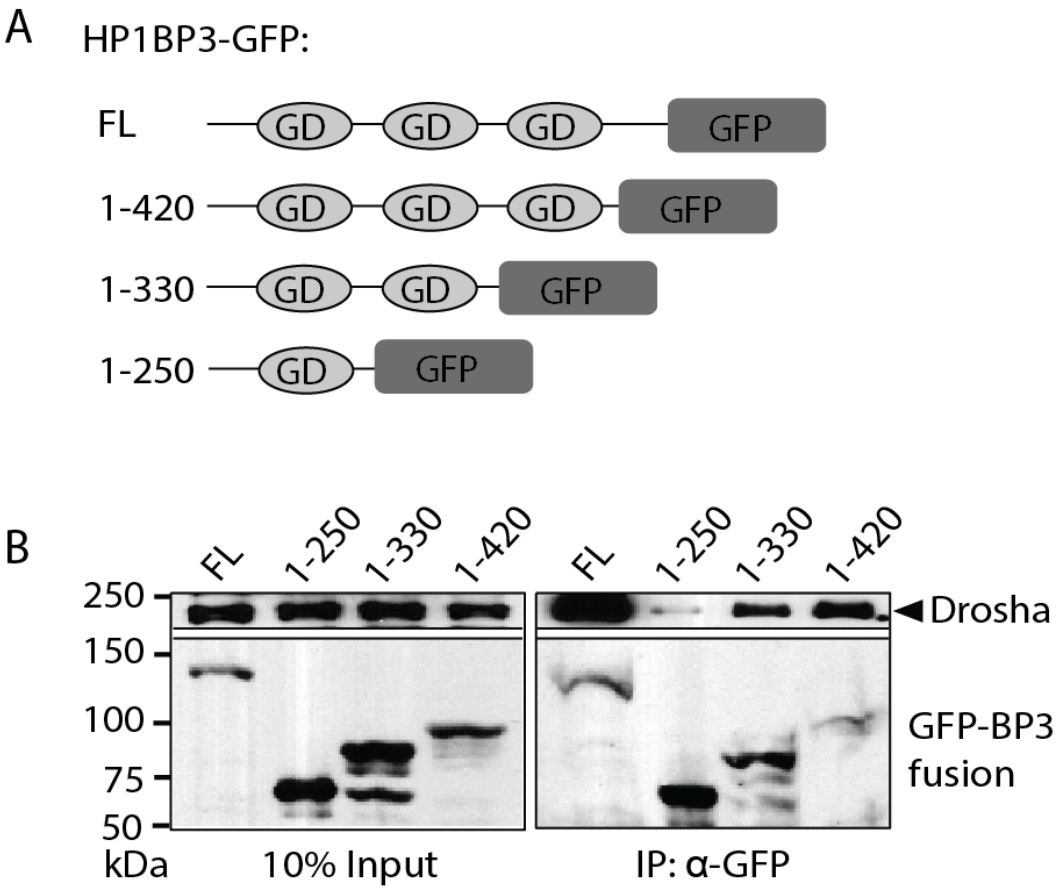

Figure 3-8 Co-immunoprecipitation of truncated HP1BP3 and Drosha
(A) A schematic diagram of various truncated HP1BP3-GFP constructs, retaining either one, two, or all three of the globular domain of HP1BP3. **(B)** After co-transfection of HeLa cells with Drosha-Fla and GFP-tagged HP1BP3 constructs, co-IP was performed with anti-GFP antibody followed by Western blotting with anti-Flag and anti-GFP antibodies, respectively. Western blotting (WB) was performed with anti-Flag and anti-GFP antibodies to visualize interaction between different truncated construct and Drosha protein. Left panel shows 10% input sample, right panel shows co-IP result.

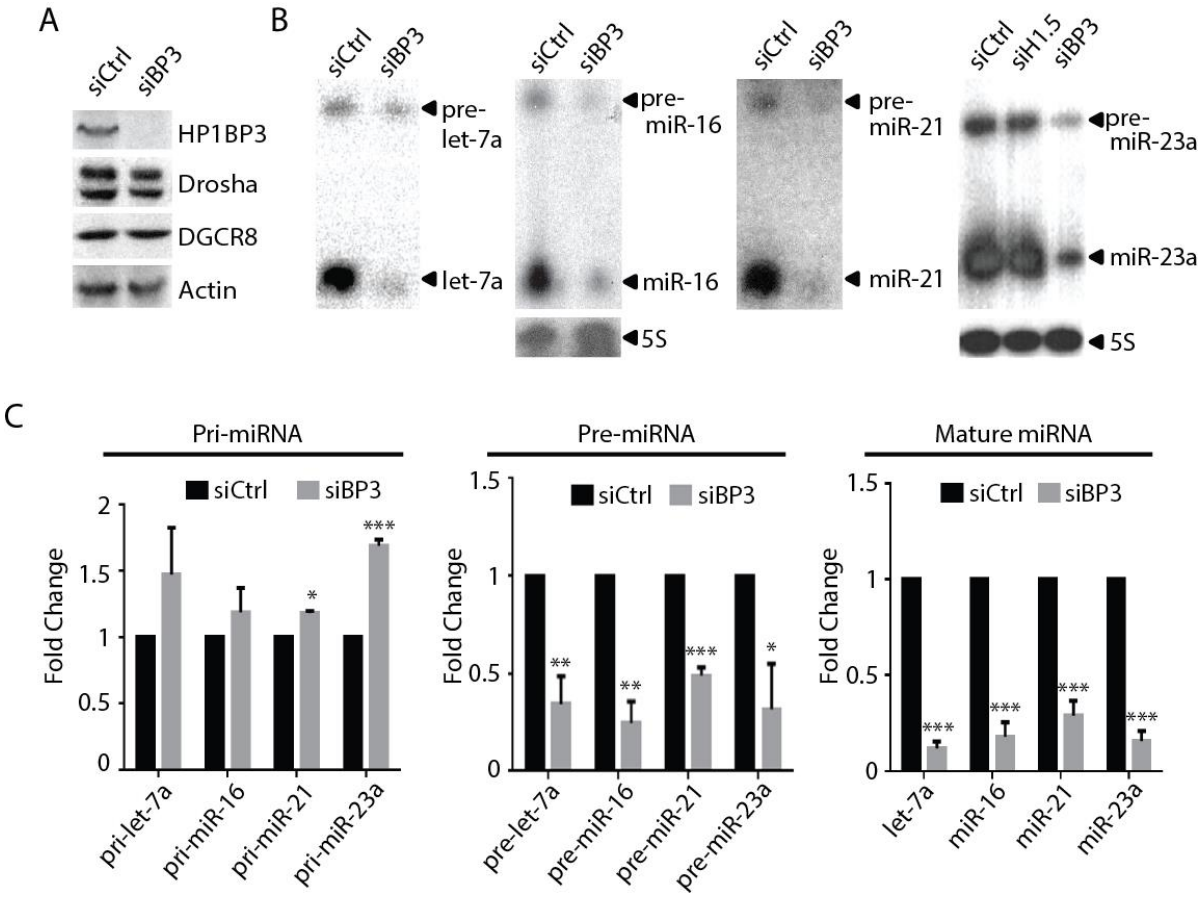

Figure 3-9 HP1BP3 enhances processing of pri-miRNA into pre- and mature miRNA
(A) Western blotting was performed using the corresponding antibodies to detect HP1BP3, Drosha, DGCR8, and Actin proteins in the control and HP1BP3-depleted HeLa cell extracts. **(B)** Northern blotting was performed to compare the levels of let-7a, miR-16, miR-21, miR-23a between the control and HP1BP3-depleted HeLa cells. 5S RNA was used as a loading control. **(C)** Quantitative analysis of the expression fold change of pri-miRNA (by Taqman PCR), pre-miRNA and miRNA (by Northern blotting in B) between the control and HP1BP3-depleted HeLa cells (triplicate samples, data shown as mean ± SD). Unpaired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$).

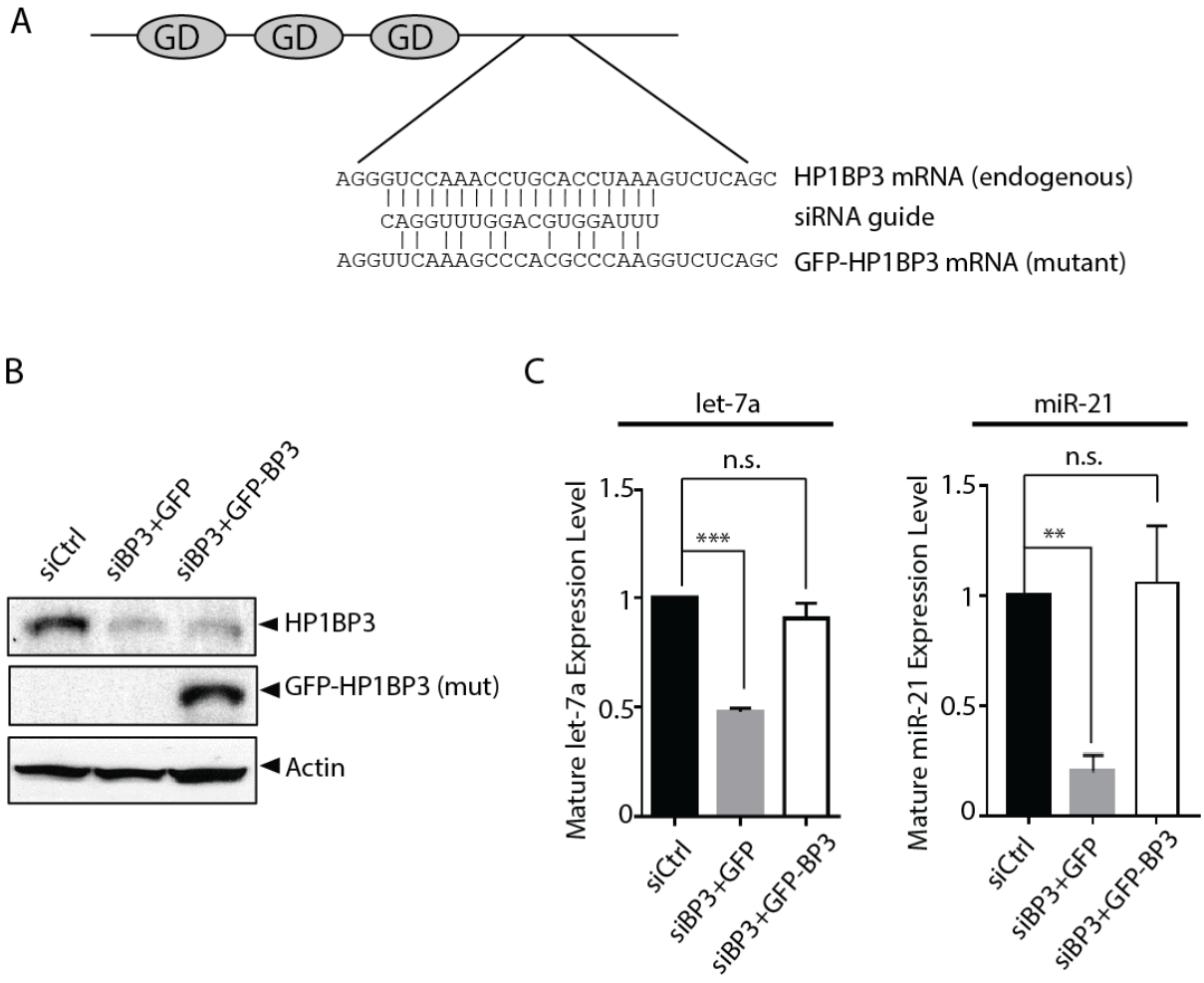

Figure 3-10 HP1BP3 enhances miRNA biogenesis

(A) A schematic diagram of base pairing between the siRNA guide RNA with wild-type or mutant (mut) HP1BP3 cDNA sequence. **(B)** HeLa cells were co-transfected with control (siCtrl) or HP1BP3 siRNA (siBP3) together with GFP or GFP-HP1BP3 (mut) construct in Hela cells. Western blot was used to show effective knockdown of endogenous HP1BP3 and expression of mutant GFP-HP1BP3 after seventy-two hours post transfection. **(C)** The levels of mature miRNAs (let-7a and miR-21) were measured by Taqman qPCR seventy-two ours after co-transfection of control (siCtrl) or HP1BP3 siRNA (siBP3) together with GFP or GFP-HP1BP3 (mut) construct in Hela cells. (triplicate samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$).

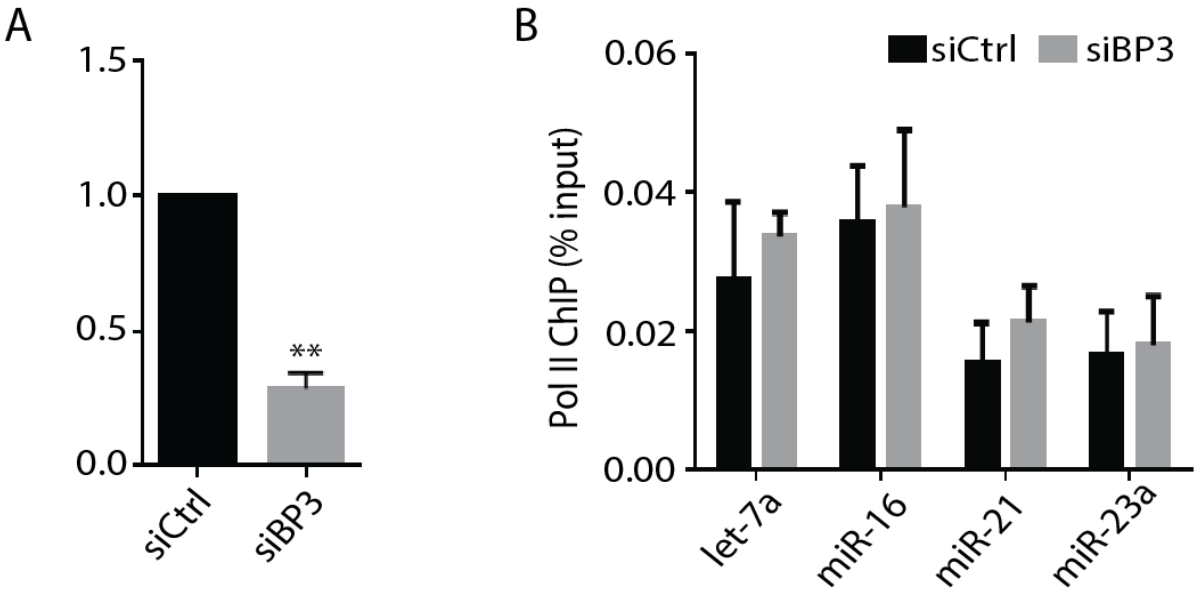

Figure 3-11 HP1BP3 does not affect transcription level of pri-miRNA
(A) Following siRNA-mediated knockdown in HeLa cells, the mRNA level of HP1BP3, was measured by real time RT-qPCR. Paired t-test was used for statistical analysis (** $p < 0.01$).
(B) ChIP analysis comparing the Pol II occupancy at the promoter regions of let-7a, miR-16, miR-21, and miR-23a genes between the control and HP1BP3-depleted HeLa cells (triplicate samples, data shown as mean ± SD).

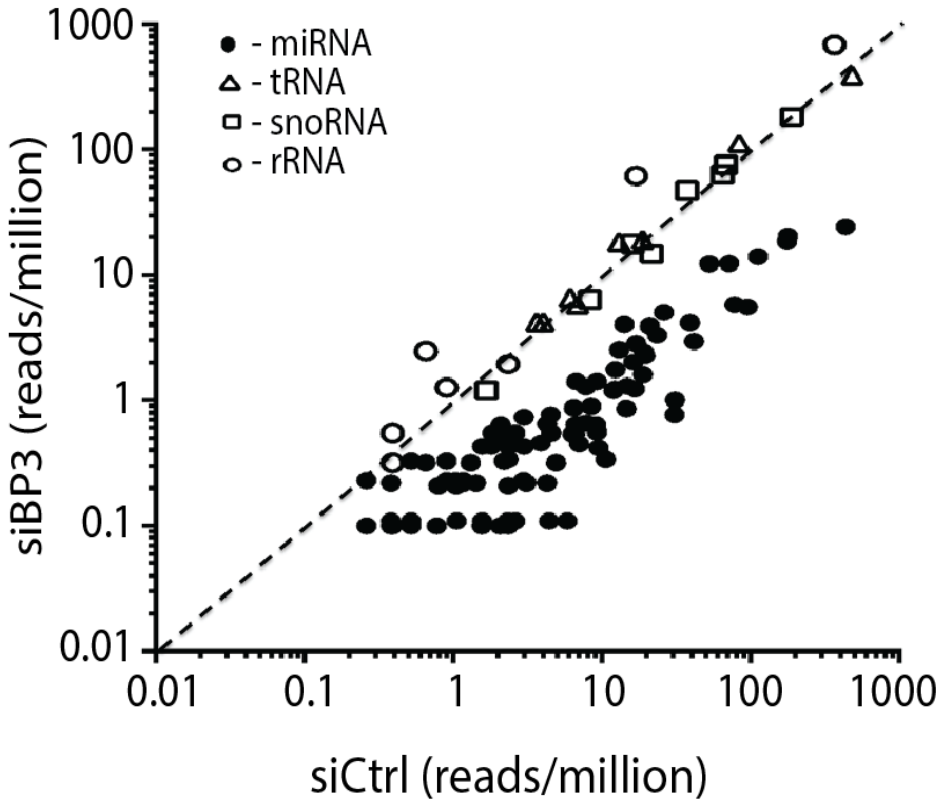

Figure 3-12 HP1BP3 regulates global miRNA biogenesis
Small RNA sequencing was performed to compare global miRNA expression profile between the control and HP1BP3-depleted HeLa cells. Each black circle represents an annotated miRNA expressed in HeLa cells. Each white triangle, square, and circle represents a tRNA, snoRNA, and rRNA, respectively. Diagonal line marks the 1:1 ratio of siCtrl and siBP3 samples.

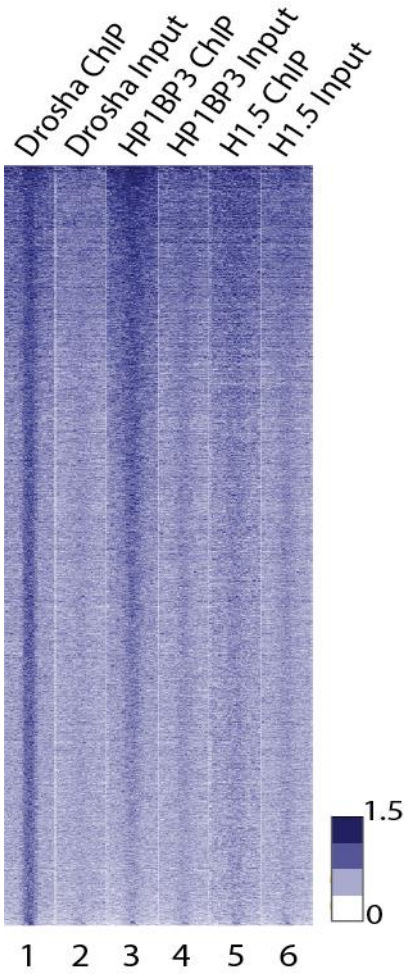

Figure 3-13 HP1BP3 co-localized with Drosha across the genome
A global view of the ChIP-Seq read density heatmap of Drosha-GFP islands throughout the genome. *X*-axis represents the distance to island center. In *Y*-axis, each row in the heatmap represents one island. Color (from dark blue to red) indicates the read density of each island. (Courtesy of Dr. Ralf Kittler)

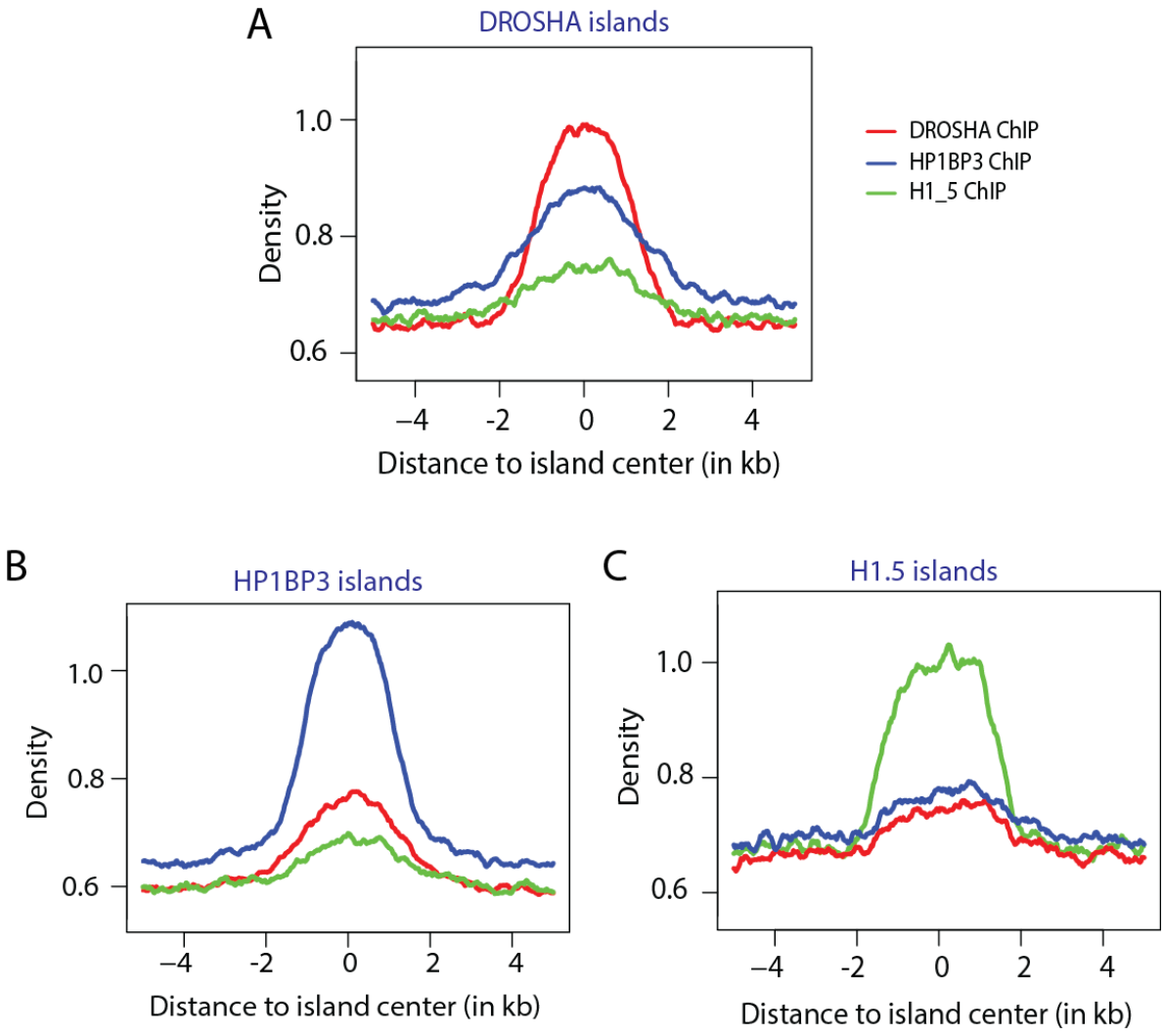

Figure 3-14 HP1BP3, not H1.5 co-localizes with Drosha
(A) Density map showing the enrichment of Drosha, HP1BP3, and H1.5 cumulative ChIP signal across all Drosha islands. *X*-axis represents the distance to island center. *Y*-axis represents per base read coverage. **(B, C)** Density map showing the enrichment of Drosha, HP1BP3 and H1.5 cumulative ChIP signal across all HP1BP3 **(B)** or H1.5 **(C)** islands. *X*-axis represents the distance to island center. *Y*-axis represents per base read coverage. (Courtesy of Dr. Ralf Kittler)

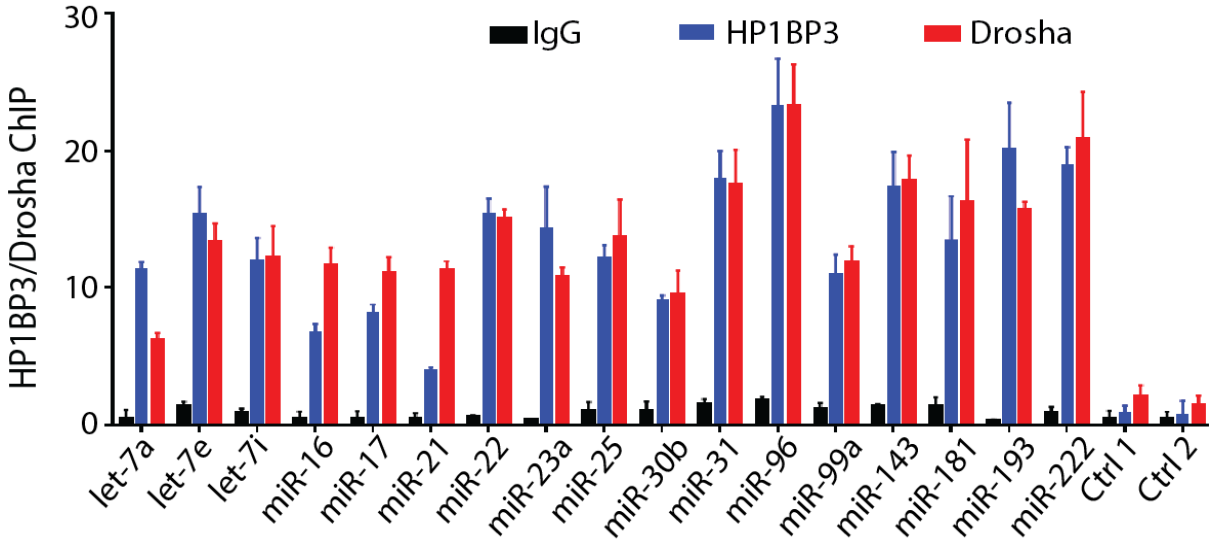

Figure 3-15 HP1BP3 co-localizes with Drohs/DGCR8 at miRNA transcription loci. Individual ChIP analysis comparing chromatin co-localization of HP1BP3-GFP and Drosha at the stem-loop regions of seventeen active miRNA loci in HeLa cells (triplicate samples, data shown as mean ± SD).

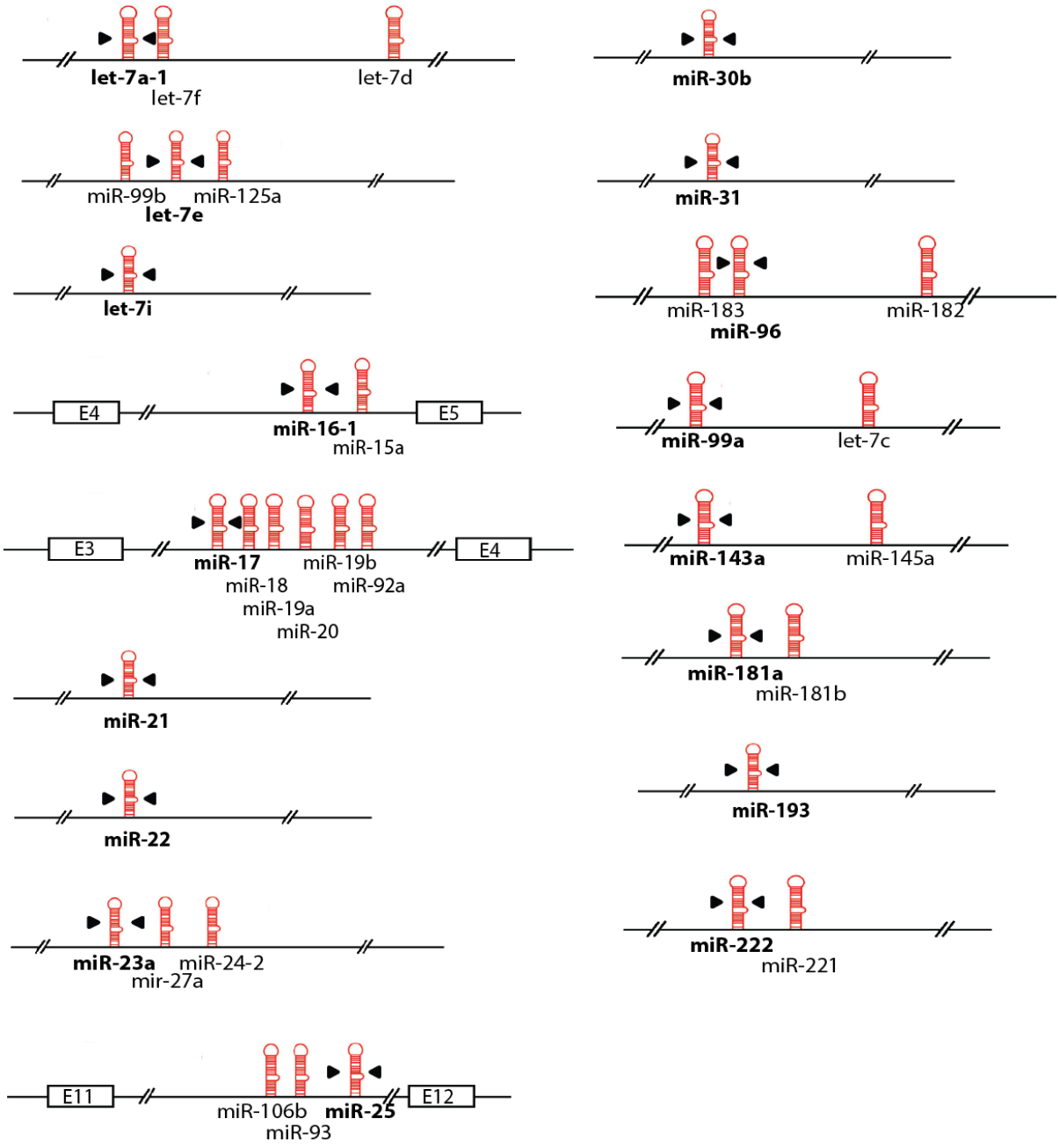

Figure 3-16 Schematic figure of primers used for ChIP analysis
A schematic figure of the seventeen expressed miRNA loci selected for ChIP analysis in Figure 3-15, including six individual miRNA genes and eleven miRNA loci that represent all of the major clusters of expressed polycistronic miRNA genes.

Table 3-1 miRNAs downregulated after BP3 knockdown (numbers shown in counts per million)

Gene ID	siCTRL_1	siCTRL_2	siBP3_1	siBP3_2
mir-31	481.5	384.8	27.35	21.29
let-7a-3	161.7	188.8	16.86	20.37
mir-21	160.2	194.5	15.83	24.72
mir-99b	103.6	87.56	4.935	6.18
mir-30a	98.25	125.8	16.24	11.67
mir-182	82.13	72.88	6.168	5.493
mir-23a	73.94	68.69	11.72	13.05
let-7i	50.92	54.53	11.1	13.5
let-7a-2	36.08	46.66	2.467	3.433
let-7b	32.75	45.09	2.879	5.493
mir-143	25.33	36.7	0.411	1.602
mir-92a-1	24.82	27.26	4.318	5.722
mir-20a	23.28	15.73	2.056	2.518
let-7f-1	23.28	23.59	2.262	4.349
mir-92a-2	23.03	18.88	3.495	4.349
let-7c	22	39.32	0.617	0.915
mir-30d	19.7	17.83	1.85	1.373
mir-181a-2	19.45	18.88	1.85	2.975
mir-33a	18.68	10.49	1.028	0.687
let-7a-1	16.12	16.25	2.467	1.602
mir-224	14.07	4.195	0.822	0.458
mir-93	13.82	19.4	0.411	2.06
mir-222	11.77	22.02	2.262	3.433
mir-30c-2	10.75	13.11	1.028	1.373
mir-181b-1	9.978	8.913	0.617	0.229
mir-16-1	9.211	18.88	3.29	4.806
mir-17	9.211	4.195	1.234	1.602
mir-16-2	8.955	15.73	1.234	2.289
mir-30c-1	8.699	17.3	1.645	3.433
mir-125a	7.932	21.5	1.234	1.373
mir-98	7.932	5.243	0.617	0.458
mir-186	7.42	13.63	0	0.687
let-7e	7.42	7.34	0.617	0.687
mir-181a-1	7.164	6.816	0.206	0.687

mir-221	7.164	8.389	1.439	1.144
let-7f-2	6.908	11.53	1.234	1.602
mir-19b-1	6.652	6.292	0.822	0.915
mir-22	6.652	17.83	0.411	2.06
mir-100	6.396	10.49	0.206	1.602
mir-24-2	6.396	2.622	0.617	0.915
mir-25	6.396	6.292	0.617	0.458
mir-374a	6.141	8.389	0	0
mir-28	5.629	3.146	0	0.229
let-7d	5.373	6.292	0	0.229
mir-181b-2	5.117	10.49	0.411	0.915
let-7g	4.861	13.63	0.206	0.915
mir-106b	4.861	8.389	0.822	0.458
mir-196a-2	4.35	1.573	0	0.458
mir-26b	4.35	4.195	0.617	0.687
mir-155	4.35	4.195	0.206	0.229
mir-671	4.35	4.719	0.411	0.687
mir-151a	4.094	3.67	0	0.915
mir-324	3.582	2.622	0.206	0.229
mir-101-2	3.582	0.524	0.206	0
mir-425	3.326	1.049	0.206	0.458
mir-455	3.326	2.622	0.411	0.458
mir-196a-1	3.07	1.049	0	0
mir-574	3.07	0.524	0.411	0.458
mir-766	3.07	1.573	0	0.229
mir-193a	2.814	3.146	1.028	0.458
mir-15a	2.559	2.097	0.411	0.458
mir-183	2.559	2.097	0.206	0
mir-744	2.303	0.524	0	0
mir-30e	2.047	3.146	0.411	0.687
mir-26a-2	2.047	3.146	0.411	0.687
mir-10a	2.047	1.049	0.206	0
mir-18a	1.791	1.049	0.206	0.229
mir-191	1.791	2.097	0.822	0.229
mir-501	1.791	2.622	0	0
mir-660	1.791	1.049	0	0
mir-1307	1.535	3.146	0	0.687
mir-193b	1.535	2.622	0.822	0.458

mir-454	1.535	3.67	0	0.229
mir-27a	1.535	0.524	0	0.458
mir-181d	1.535	1.049	0	0
mir-23b	1.535	3.146	0.411	0
mir-101-1	1.279	0.524	0	0.458
mir-92b	1.279	1.049	0	0.458
mir-139	1.279	0	0	0
mir-145	1.279	1.049	0.206	0.229
mir-27b	1.279	2.622	0	0
mir-452	1.279	0.524	0.206	0.458
mir-210	1.023	0	0	0
mir-103a-2	1.023	0.524	0.206	0
mir-301b	1.023	0.524	0.206	0
mir-26a-1	1.023	1.573	0.411	0.229
mir-138-1	1.023	2.097	0	0.229
mir-551b	1.023	0	0	0
mir-29a	1.023	2.097	0.411	0.458
mir-2278	1.023	0	0	0
mir-32	1.023	0.524	0	0
mir-503	1.023	1.049	0	0
mir-34a	0.768	0	0.206	0.229
mir-146b	0.768	1.049	0	0
mir-1908	0.768	0.524	0	0
mir-125b-1	0.768	0	0.206	0
mir-19a	0.768	0.524	0	0
mir-365a	0.768	0	0	0.229
mir-296	0.768	0.524	0.411	0.229
mir-125b-2	0.768	2.097	0	0
mir-339	0.768	0.524	0	0
mir-196b	0.768	0	0	0
mir-335	0.768	0	0.206	0
mir-532	0.768	0.524	0	0
mir-361	0.768	0	0	0
mir-345	0.512	3.146	0	0
mir-138-2	0.512	0.524	0	0
mir-195	0.512	0	0	0
mir-301a	0.512	0.524	0.206	0.458
mir-181c	0.512	1.573	0	0

mir-149	0.512	0	0.206	0
mir-378a	0.512	3.146	0.411	0.687
mir-340	0.512	0	0	0
mir-219a-1	0.512	0.524	0.206	0
mir-30b	0.512	1.573	0.411	0
mir-7-1	0.512	0	0	0
mir-362	0.512	0	0	0
mir-148b	0.256	1.049	0	0
mir-331	0.256	1.049	0	0
mir-190a	0.256	1.573	0	0
mir-152	0.256	1.573	0	0
mir-7-3	0.256	0.524	0	0
mir-330	0.256	1.049	0	0
mir-2355	0.256	0.524	0	0
mir-99a	0.256	1.049	0	0
mir-15b	0.256	0.524	0.206	0
mir-103a-1	0.256	1.573	0	0
mir-126	0.256	1.049	0	0
mir-652	0.256	0.524	0.206	0
mir-19b-2	0.256	0.524	0	0
mir-505	0.256	0.524	0	0
mir-137	0	1.573	0	0
mir-197	0	0.524	0	0
mir-130a	0	2.097	0	0.229
mir-200c	0	0.524	0	0
mir-615	0	1.049	0	0
mir-132	0	1.573	0.411	0
mir-769	0	1.049	0	0
mir-128-1	0	0.524	0	0
mir-130b	0	0.524	0	0
mir-3619	0	0.524	0	0
mir-582	0	1.049	0	0.229
mir-589	0	0.524	0.206	0
mir-148a	0	0.524	0	0
mir-96	0	1.049	0	0
mir-29b-1	0	0.524	0	0.458
mir-500a	0	1.049	0	0.229

Table 3-2. List of expressed miRNA loci co-localized with Drosha and HP1BP3.

Gene ID	Drosha	HP1BP3	Fold Change	Method
mir-92b	+	-	0.19663	CS
mir-26b	+	-	0.15256	CS
mir-191	+	+	0.27038	CS
mir-425	+	+	0.15163	CS
mir-219a-1	+	+	0.19846	CS
mir-151a	+	-	0.11792	CS
mir-4664	+	+	0.80361	CS
mir-365a	+	-	0.29817	CS
mir-23a	+	+	0.17364	both
mir-24-2	+	+	0.16992	both
mir-27a	+	+	0.22226	both
mir-181c	+	+	*	CS
mir-181d	+	+	*	CS
let-7e	+	+	0.08831	both
mir-99b	+	+	0.05814	CS
mir-125a	+	+	0.08858	CS
mir-1249	+	+	*	CS
mir-3619	+	+	*	CS
let-7a	+	+	0.12571	IC
let-7i	+	+	0.23336	IC
mir-16-1	+	+	0.28825	IC
mir-17	+	+	0.21154	IC
mir-21	+	+	0.11433	IC
mir-22	+	+	0.10095	IC
mir-25	+	+	0.08469	IC
mir-30b	+	+	0.19726	IC
mir-31	+	+	0.05613	IC
mir-96	+	+	*	IC
mir-99a	+	+	*	IC
mir-143	+	+	0.03245	IC
mir-181a	+	+	0.12953	IC
mir-193	+	+	0.24928	IC
mir-222	+	+	0.16853	IC
mir-574	-	+	0.24174	CS

mir-378a	-	+	0.30015	CS
mir-589	-	+	0.39215	CS
mir-196b	-	+	*	CS
mir-106b	+	+	0.96615	both
mir-93	+	+	0.07439	both
mir-126	-	+	*	CS
mir-1908	-	+	*	CS
mir-615	-	+	*	CS
mir-18a	+	+	0.15301	both
mir-20a	+	+	0.11724	both
mir-19a	+	+	*	both
mir-19b-1	+	+	0.13426	both
mir-92a-1	+	+	0.19276	both
mir-342	-	+	*	CS
mir-324	-	+	0.07004	CS
mir-769	-	+	*	CS
let-7f	+	+	0.14101	both
mir-15	+	+	0.18664	both
mir-183	+	+	0.04416	both
mir-182	+	+	0.07523	both
let-7c	+	+	0.02499	both
mir-181b	+	+	0.04477	both
mir-221	+	+	0.16612	both

a. * - No miRNA was detected in either siBP3_1 or siBP3_2 samples of RNA-seq
b. CS – ChiP-seq analysis. IC – individual ChIP analysis. Both – CS and IC analysis

CHAPTER FOUR

HP1BP3 IS A NOVEL CHROMATIN RETENTION FACTOR THAT ENHANCES PRI-MIRNA PROCESSING

*Note from author: The content of this chapter is included in the paper "**HP1BP3, a chromatin retention factor for co-transcriptional microRNA processing**"*

Abstract

HP1BP3, a histone H1-like protein, was identified to facilitate global co-transcriptional processing of pri-miRNA by Drosha/DGCR8 complex. Through detailed ChIP analysis, we found that HP1BP3 is enriched at genomic loci that correspond to the stem-loop region of actively transcribed miRNA. Further, depletion of HP1BP3 resulted in reduction of Drosha's localization of miRNA transcription site. Further, we found that HP1BP3 is a unique histone protein that can bind specifically to pri-miRNA as well as genomic DNA simultaneously, and its presence significantly enhanced Drosha's association with nascent transcribed miRNA. Knockdown of HP1BP3 compromises pri-miRNA processing by resulting in premature release of pri-miRNA transcripts from the chromatin. Taken together, these studies suggest that HP1BP3 promotes co-transcriptional miRNA processing *via* chromatin retention of nascent pri-miRNA transcripts.

Introduction

In eukaryotic cells, the processing (e.g. capping, splicing, and 3' end formation) of precursor messenger RNA (pre-mRNA) is closely coupled to its transcription by RNA

polymerase II (Pol II) (Bentley, 2014). Nascent pre-mRNA transcripts can be capped shortly after their 5' ends emerge from Pol II complex (Rasmussen and Lis, 1993). The majority of active spliceosomes are chromatin bound and probably function co-transcriptionally in mammalian cells (Brugiolo et al., 2013). Accordingly, sequencing of chromatin–associated RNA detected an enrichment of exon over neighboring intron sequences, suggesting that the bulk of splicing occurs on chromatin (Brugiolo et al., 2013). Moreover, termination of transcription requires the recognition and cleavage of the 3' "AAUAAA" cleavage and polyadenylation (CPA) signal sequence (Whitelaw and Proudfoot, 1986; Zaret and Sherman, 1982). Cleavage of CPA provides an entry site for exonuclease XRN2/Rat1 to promote the eviction of Pol II from the DNA template (West et al., 2004). Thus, the 3' end formation of pre-mRNA must occur co-transcriptionally prior to transcription termination. It is commonly believed that co-transcriptional processing enhances the efficiency and accuracy of pre-mRNA maturation and allows for temporally coordinated regulation of various processing events on chromatin.

Accumulating studies have suggested that the processing of pri-miRNA can occur co-transcriptionally. The majority of pri-miRNA transcripts are transcribed by Pol II and have 5' cap and 3' poly(A) tail as pre-mRNA transcripts (Lee et al., 2004a). In mammals, pri-miRNAs frequently reside in the introns of other protein coding transcripts. An intronic pri-miRNA is processed before splicing that occurs co-transcriptionally (Morlando et al., 2008). Accordingly, chromatin immunoprecipitation (ChIP) analysis reveals that the Drosha-DGCR8 complex associates with several actively transcribing miRNA loci (Ballarino et al., 2009; Morlando et al., 2008). On the other hand, artificial retention of pri-miRNA transcripts

on the chromatin, e.g. by blocking 3' end formation, greatly enhances miRNA production *in vivo* (Pawlicki and Steitz, 2008). Both the kinetics and efficiency of pri-miRNA processing are enhanced when coupled to pri-miRNA transcription by Pol II in the nuclear extract of HeLa cells (Yin et al., 2015). Moreover, co-transcriptional pri-miRNA processing is also supported by native elongating transcript sequencing (mNET-seq) studies for mammalian chromatin (Nojima et al., 2015). However, the molecular mechanisms of co-transcriptional miRNA processing are poorly understood. It remains uncertain whether all or a subset of pri-miRNAs undergo co-transcriptional processing. It is also unclear whether specific chromatin factors regulate co-transcriptional miRNA processing.

In the previous chapter, we have established HP1BP3 as a novel regulator of co-transcriptional processing of pri-miRNA. It specifically associates with the Drosha/DGCR8 complex, co-localized globally with Drosha/DGCR8 at miRNA loci, and promotes global miRNA biogenesis in human cells. However, mechanistic insight of the interaction between HP1BP3, Drosha/DGCR8, and pri-miRNA remains unknown. In this chapter, we will be focused on answering the question: "How does HP1BP3 facilitate co-transcriptional processing of pri-miRNA?" In the process, we revealed that HP1BP3 enhances Drosha's localization to pri-miRNA transcription loci, as well as Drosha's association with pri-miRNA itself. Further, we found that HP1BP3 is a unique histone H1-like protein whose domain structure allows it to simultaneously bind to both pri-miRNA and DNA. Due to this special property, HP1BP3 can function like a "molecular hook" at site of pri-miRNA transcription, retaining the nascent pri-miRNAs on chromatin. This effectively increases the local concentration of pri-miRNA at site of transcription, which allows Drosha to more efficiently

co-transcriptionally catalyze its transformation into pre-miRNA. Our study established a working model of how a chromatin protein can facilitate co-transcriptional processing of pri-miRNA, and proposed a novel mechanism for which may be applied to link chromatin protein with the processing of other RNAs.

Results

Binding of Drosha to actively transcribed miRNA loci is dependent on HP1BP3

In previous studies, it was shown that Drosha is recruited to the gene region corresponding to the stem-loop structure of pri-miRNA (Morlando et al., 2008). Based on our finding that HP1BP3 interact with Drosha/DGCR8 complex, we hypothesize that HP1BP3 must also localize to the same genomic region to facilitate the processing of pri-miRNA. Therefore, we performed extensive detailed ChIP analysis to compare the chromatin binding patterns of HP1BP3-GFP and Drosha at miRNA transcription locus by designing PCR primers for every 100-200 bp within a 4,000 bp region surrounding the let-7a-1 and miR-16 miRNA. Both HP1BP3-GFP and Drosha showed a single chromatin-binding peak at the stem-loop region of let-7a-1 and miR-16 locus **(Figure 4-1)**. Studies have also shown that Drosha is localized to miRNA transcription site in a Pol II-dependent manner, suggesting that Drosha selectively binds to actively transcribed miRNA loci. To test if HP1BP3 also exhibit this selectivity, we performed ChIP experiment following qRT-PCR with primers designed for miRNA with high expression level in Hela (e.g. let-7a-1, miR-16, miR-17, miR-21) with miRNA with low to no expression (e.g. miR-1, miR-9). The miRNA expression profile of Hela cells was taken from control samples in previous small RNA sequencing result (Chapter

2). We found that HP1BP3, unlike H1.5, is specifically enriched at actively transcribed miRNA loci **(Figure 4-2)**. Furthermore, knockdown of HP1BP3 expression resulted in a significant reduction in the Drosha ChIP signal at all six miRNA loci that we examined in HeLa cells **(Figure 4-3)**. Thus, it is possible that HP1BP3 plays a critical role in the recruitment or retention of the Microprocessor at active miRNA loci to catalyze co-transcriptional pri-miRNA processing. Since previous co-IP result indicated that HP1BP3 interact with pri-miRNA in a RNA dependent manner, understanding the molecular detail of how HP1BP3 interact with pri-miRNA will be critical for elucidating how HP1BP3 recruits Drosha to miRNA transcription loci.

HP1BP3 exhibits specific pri-miRNA binding activity in vitro and in vivo

As previously reported (Hayashihara et al., 2010), poly-histidine-tagged HP1BP3 and H1.5 recombinant proteins exhibited double-stranded DNA (dsDNA) binding activity **(Figure 4-4)**. Unexpectedly, we found that recombinant HP1BP3, but not H1.5, also exhibited a specific pri-miRNA-binding activity *in vitro* **(Figure 4-4)**. Whereas pre-miRNA typically consists of a terminal loop and a ~22 bp imperfect double-stranded stem with 2-nt 3' overhang, the corresponding pri-miRNA carries a longer (e.g. ~33 bp) stem as well as two single-stranded tails. Recombinant HP1BP3 efficiently bound to pri-miRNA and, to a lesser extent, pre-miRNA, but not duplex miRNA **(Figure 4-5A)**, suggesting that HP1BP3 probably recognize the shared stem-loop structure of pri-/pre-miRNA. Neither pre-miRNA nor miRNA duplex could compete off binding of radiolabeled pri-miRNA by HP1BP3 **(Figure 4-5B)**. Thus, HP1BP3 displayed a higher affinity for pri-miRNA than for pre-

miRNA, suggesting that the single-stranded tails of pri-miRNA (absent in pre-miRNA) could also contribute critically to HP1BP3 binding. This notion that HP1BP3 recognizes the common structure features of pri-miRNAs is also consistent with two of our findings: 1) HP1BP3 promotes global miRNA biogenesis; 2) HP1BP3 ChIP signals peak at the stem-loop regions of many active miRNA loci.

To investigate whether HP1BP3 bound dsDNA and pri-miRNA *via* the same mechanism, we performed *in vitro* competition experiments by using non-radiolabeled pri-miRNA or dsDNA of the same sequence to compete with the binding of radiolabeled pri-miRNA by recombinant HP1BP3. Although dsDNA could not compete off binding of radiolabeled pri-miRNA **(Figure 4-6)**, addition of dsDNA upshifted the HP1BP3-pri-miRNA complex, suggesting that HP1BP3 may bind dsDNA and pri-miRNA simultaneously *via* different mechanisms.

We performed RNA immunoprecipitation (RIP) to examine whether HP1BP3 associated with endogenous pri-miRNA transcripts *in vivo*. Following *in vivo* formaldehyde crosslinking of the RNA/protein (RNP) complexes, we immunoprecipitated HP1BP3-GFP or H1.5-GFP from the BAC transgenic HeLa cells and measured the abundance of HP1BP3 or H1.5-associated pri-miRNAs by Taqman qPCR. This experiment showed that HP1BP3-GFP associated with all five pri-miRNAs that we examined, whereas H1.5-GFP moderately associated with pri-miR-16, but not four other pri-miRNAs **(Figure 4-7A)**. Moreover, the GFP-tagged HP1BP3 1-250, 1-330 and 1-420 truncated proteins showed increasing affinity for endogenous pri-miRNAs, suggesting that the three GDs of HP1BP3 function cooperatively to bind pri-miRNAs *in vivo* **(Figure 4-7B)**. This result was consistent with

previous co-immunoprecipitation data, as inclusion of more GD also enhanced HP1BP3's binding to Drosha. Our result further indicates that HP1BP3 interact with Drosha/DGCR8 in an RNA dependent manner. Taken together, these results suggest that HP1BP3 exhibits a specific pri-miRNA binding, as well as DNA binding, both *in vitro* and *in vivo*.

HP1BP3 promotes Drosha-pri-miRNA association in vivo

Furthermore, we compared by RIP the *in vivo* association of endogenous Drosha and pri-miRNA transcripts between the control and HP1BP3 or H1.5-depleted HeLa cells. Consistently, depletion of HP1BP3, but not depletion of H1.5, resulted in a significant reduction in Drosha binding to all four endogenous pri-miRNAs that we examined **(Figure 4-8)**. These results indicate that HP1BP3 promotes the Drosha/pri-miRNA association *in vivo*. Next, we performed native gel-shift assay to examine whether HP1BP3 directly enhance the binding of Drosha-DGCR8 complex to pri-miRNA *in vitro* **(Figure 4-9)**. Although HP1BP3 did not enhance the pri-miRNA binding, it resulted in a supershift of the Drosha-DGCR8/pri-miRNA complex (Figure 4-9), suggesting that HP1BP3 and Drosha-DGCR8 could co-occupy the same pri-miRNA to form a higher order complex. HP1BP3 also slightly inhibited the *in vitro* pri-miRNA processing activity of recombinant Microprocessor **(Figure 4-10)**. In our reconstitution system, however, the lack of chromatin component and/or use of truncated Drosha/DGCR8 proteins could potentially explain the inability to fully recapitulate the physiological activity of HP1BP3 *in vitro*.

HP1BP3 promotes chromatin retention of nascent pri-miRNA transcripts

At this point, we have established that 1) HP1BP3 co-localize with Drosha to actively transcribed miRNA loci; 2) HP1BP3 can specifically bind to pri-miRNA, while retaining its association with both DNA and Drosha/DGCR8; 3) Depletion of HP1BP3 reduces Drosha's localization to miRNA loci, as well as its binding to pri-miRNA. From the above observations, a hypothetical model emerged that HP1BP3 may function as a "molecular hook" that helps pri-miRNA retain on chromatin at the site of its transcription, thus allowing its more efficient co-transcriptional processing by Drosha/DGR8. Knockdown of HP1BP3 could cause the premature release of pri-miRNA transcripts from the chromatin and thus, resulting in less chromatin binding of Drosha to miRNA loci and diminished Drosha-pri-miRNA association. In fact, Steitz and colleagues previously showed that pri-miRNA transcripts retained on chromatin, e.g. by inhibition of 3' polyadenylation, were processed by the Microprocessor more efficiently in HeLa cells, leading to an increase of mature miRNA production (Pawlicki and Steitz, 2008). Inspired by this study, we used the same strategy to test our hypothesis by comparing exogenous miRNA expression from transfected Pri-lin-4 or Pri-lin-4ΔpA constructs that contained or lacked the "AAUAAA" cleavage and polyadenylation (CPA) signal **(Figure 4-11)**. We expect that HP1BP3 will promote lin-4 production similarly to endogenous miRNA in Hela cells transfected with regular Pri-lin-4 construct. However, for Pri-lin-4ΔpA transfected Hela cells, depletion of HP1BP3 would have little to no effect on lin-4 production, because the nascent pri-miRNA transcript is retained on chromatin in a mechanism that is independent of HP1BP3.

First, we must establish that HP1BP3 was required for expression of lin-4 from the Pri-lin-4 construct in a similar fashion as endogenous miRNAs. Using Taqman PCR, we first

established that the processing of pri-lin-4 is dependent on HP1BP3 **(Figure 4-12A)**. Then, ChIP analysis revealed that HP1BP3 associated with transfected Pri-lin-4 plasmid **(Figure 4-12B)**, similar to how HP1BP3 is localized to endogenous miRNA transcription loci. Next, we compared the expression of pri-lin-4 and mature lin-4 from the Pri-lin-4 and Pri-lin-4ΔpA constructs, respectively, in control and HP1BP3-depleted HeLa cells **(Figure 4-13)**. Similar to endogenous miRNA, knockdown of HP1BP3 did not affect pri-lin-4 transcription, but reduced mature lin-4 production from the Pri-lin-4 construct. In contrast, neither pri-lin-4 nor mature lin-4 expression from the Pri-lin-4ΔpA construct was affected by HP1BP3 depletion. These results are consistent with our hypothesis that HP1BP3 promotes co-transcriptional pri-lin-4 processing. In the case of Pri-lin-4ΔpA, artificial retention of pri-lin-4 transcript at the transcription site, due to deletion of CPA, allowed for efficient co-transcriptional pri-lin-4 processing independent of HP1BP3.

To further confirm if HP1BP3 promoted chromatin retention of pri-lin-4 transcript, we isolated the chromatin and nucleoplasm fractions from the control and HP1BP3-depleted HeLa cells as described (Gagnon et al., 2014; Morlando et al., 2008; Pawlicki and Steitz, 2008). A schematic diagram of this procedure is shown in **Figure 4-14**. We then quantified the relative abundance of pri-lin-4 in both chromatin and nucleoplasm fraction by Taqman qPCR **(Figure 4-15)**. For the Pri-lin-4 construct, knockdown of HP1BP3 resulted in a two to four-fold reduction of pri-lin-4 from the chromatin, accompanied by a two-fold increase in pri-lin-4 in nucleoplasm. Since the total amount of pri-lin-4 was unchanged, depletion of HP1BP3 resulted in the premature release of pri-lin-4 transcript from the chromatin into the nucleoplasm. As expected for the Pri-lin-4ΔpA construct, depletion of HP1BP3 did not

change the abundance of pri-lin-4 transcript in the chromatin or nucleoplasm fraction **(Figure 4-15)**. Taken together, these results suggest that HP1BP3 promotes retention of pri-lin-4 transcript at the transcription site to enhance co-transcriptional pri-lin-4 processing.

Furthermore, we compared the relative abundance of seven endogenous pri-miRNA transcripts in the chromatin and nucleoplasm fractions isolated from the control and HP1BP3-depleted HeLa cells. Consistently, depletion of HP1BP3 resulted in a significant reduction of all five pri-miRNAs from the chromatin and a concomitant increase of these pri-miRNAs in the nucleoplasm **(Figure 4-16)**. Moreover, we verified that nascent pri-miRNA transcripts were prematurely released from the chromatin by using PCR primers spanning the introns of pri-miR-25 and pri-miR-140 **(Figure 4-16)**. By contrast, the abundance of histone H4 transcript remained the same in the chromatin and nucleoplasm fractions in HP1BP3-depleted cells **(Figure 4-16)**. However, knockdown of H1.2 or H1.5 did not significantly affect the levels of pri-miRNA in either the chromatin or nucleoplasm fractions **(Figure 4-17)**. Taken together, these series of experiments strongly suggest that HP1BP3 promotes co-transcriptional miRNA processing through chromatin retention of nascent pri-miRNA transcripts.

In conclusion, we have identified HP1BP3 is a unique histone H1 protein that can binds simultaneously to both RNA and DNA through distinct mechanism. Depletion of HP1BP3 results in reduced microprocessor's association with chromatin and with pri-miRNA. Further, we demonstrated that HP1BP3 increase pri-miRNA processing efficiency by retaining nascent pri-miRNA transcript on chromatin. The mechanism of how HP1BP3

enhances miRNA biogenesis has not been previously identified, and our research established a novel link between histone protein and co-transcriptional processing of RNAs.

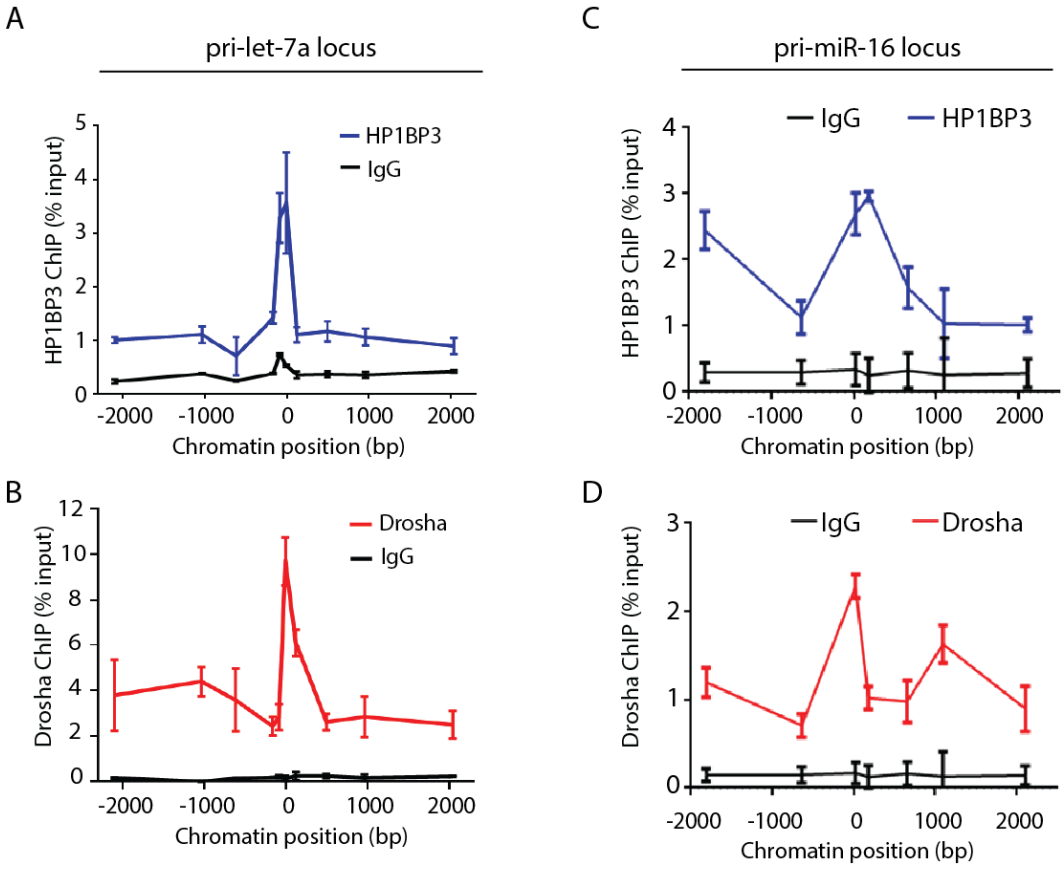

Figure 4-1 HP1BP3 and Drosha co-localize to transcription loci corresponding to the stem-loop region of pri-miRNA.
(A, B) Comprehensive ChIP analysis comparing chromatin co-localization of HP1BP3-GFP **(A)** and Drosha-GFP **(B)** in a 4 kb region surrounding the stem-loop region of the let-7a-1 locus. **(C, D)** Comprehensive ChIP analysis comparing chromatin co-localization of HP1BP3-GFP **(C)** and Drosha-GFP **(D)** in a 4 kb region surrounding the stem-loop region of the miR-21 locus.

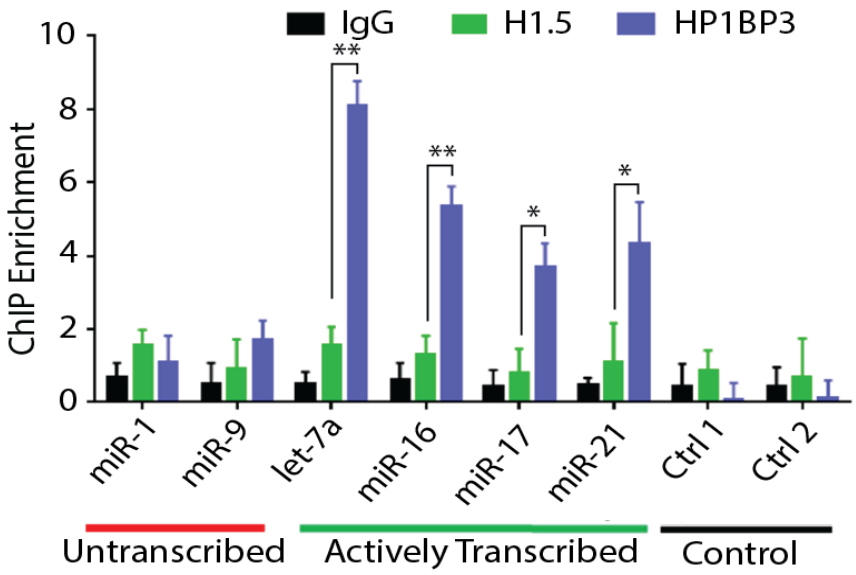

Figure 4-2 HP1BP3 is enriched at actively transcribed miRNA loci.
ChIP analysis comparing chromatin binding of H1.5-GFP and HP1BP3-GFP at the stem-loop regions of several inactive (miR-1, miR-9) and actively transcribed miRNA loci (let-7a, miR-16, miR-17, miR-21). Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$).

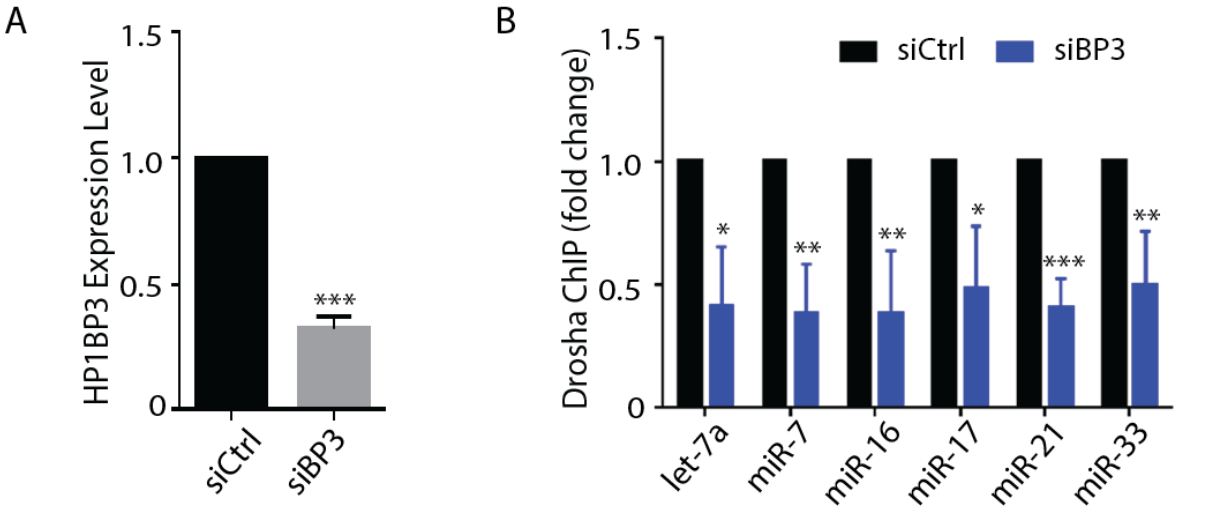

Figure 4-3 HP1BP3 enhances Drosha's localization to miRNA loci.
(A) Following siRNA-mediated knockdown in HeLa cells, the mRNA level of HP1BP3, H1.0, H1.2, H1.5 or H1x was respectively measured by real time RT-qPCR (triplet samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$). **(B)** Individual ChIP analysis comparing chromatin binding of Drosha at six active miRNA loci between the control and HP1BP3-depleted HeLa cells (quadruplet samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$).

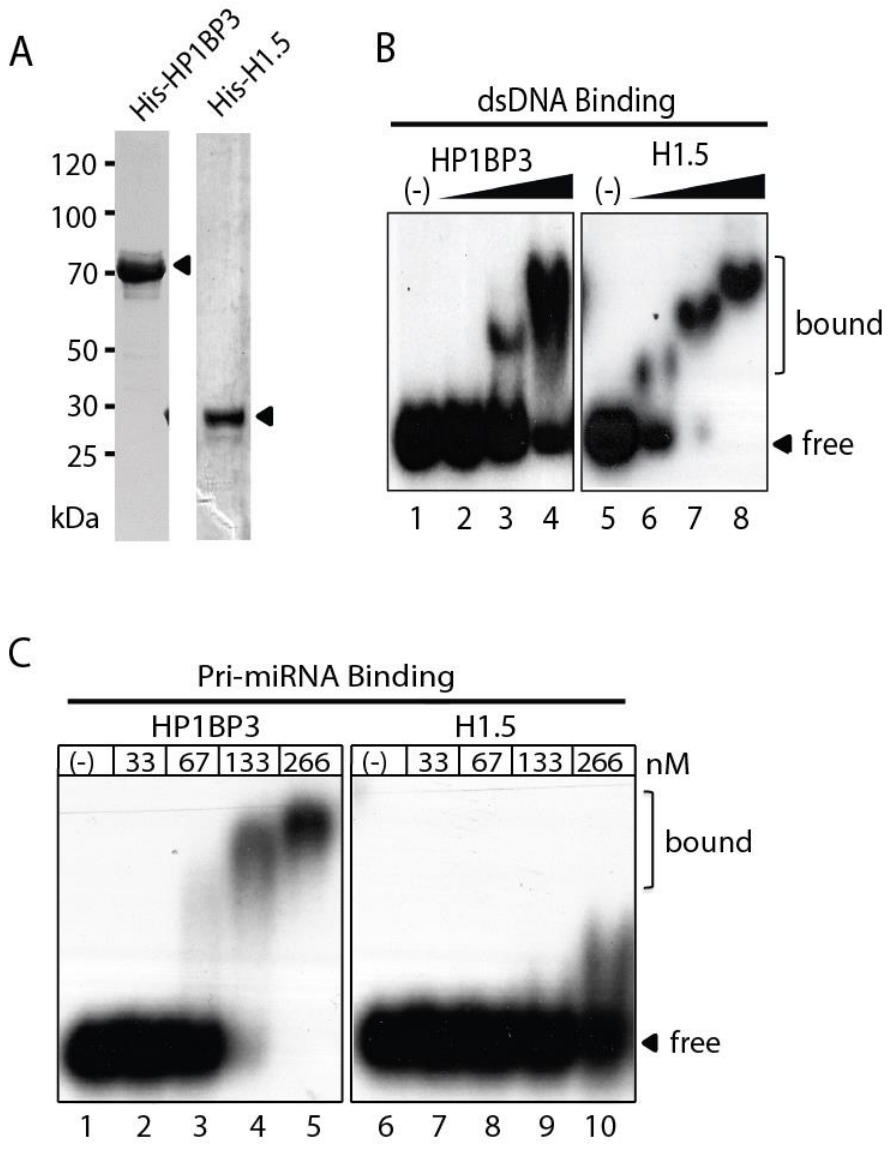

Figure 4-4 HP1BP3, not H1.5, specicially associates with pri-miRNA *in vitro*.
(A) A Coomassie-stained SDS-PAGE showing purified His_6-HP1BP3 and His_6-H1.5 recombinant proteins. **(B)** Native gel-shift assays were performed by incubating 5' radiolabeled pri-let-7a dsDNA in buffer alone (lane 1, 5) or with increasing concentration of recombinant HP1BP3 (lanes 2-4) and H1.5 (lane 6-8) proteins. **(C)** Native gel-shift assays were performed by incubating 5' radiolabeled pri-let-7a RNA in buffer alone or with increasing concentration (33, 67, 133, and 266nM) of recombinant HP1BP3 (lanes 1-5) and H1.5 (lanes 6-10). Arrowhead marks free pri-let-7a RNA substrate and bracket corresponds to the HP1BP3-bound pri-let-7a RNA.

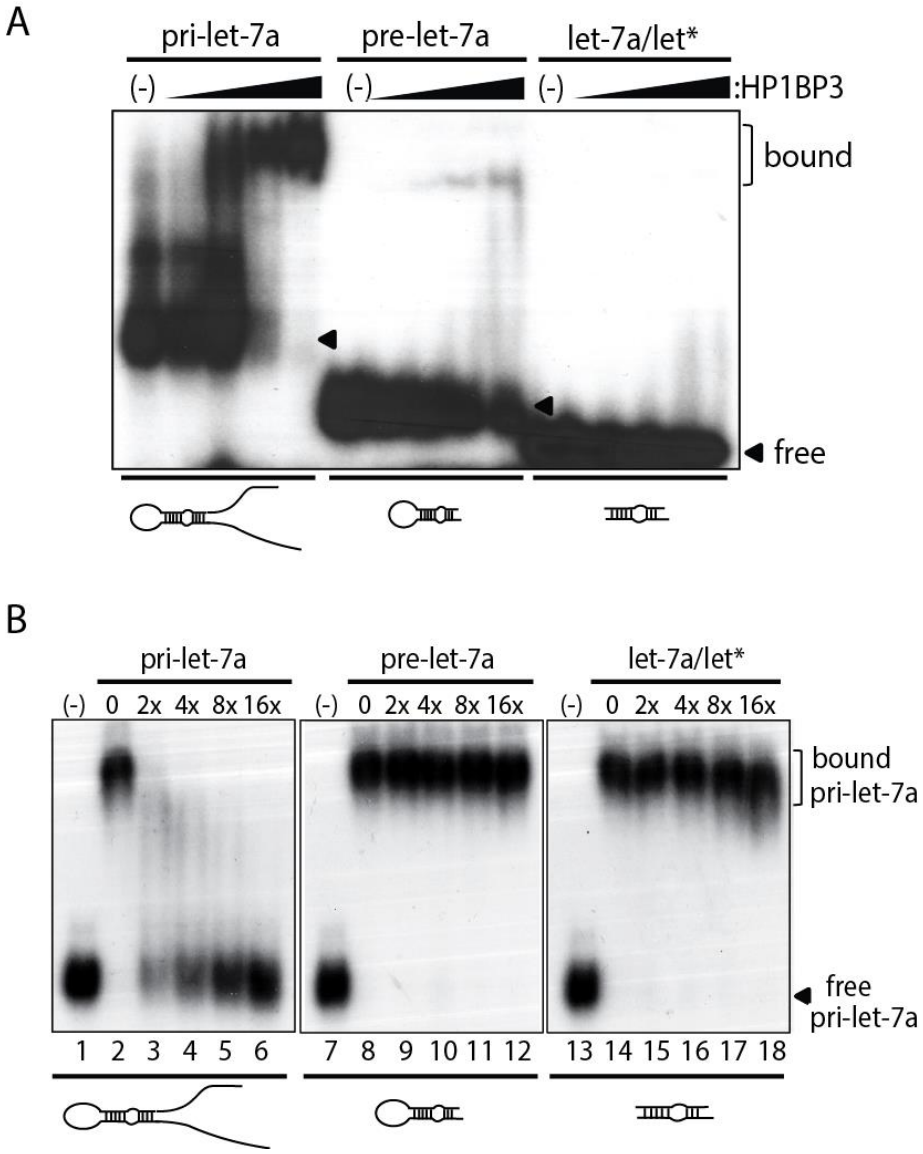

Figure 4-5 HP1BP3 specifically associates with pri-miRNA in vitro.
(A) Native gel-shift assays were performed by incubating 5' radiolabeled pri-let-7a (lanes 1-5), pre-let-7a (lanes 6-10), and let7a/let* duplex (lanes 11-15) RNA in buffer alone (lanes 1, 6, 11), or with increasing concentration (38nM, 75nM, 150nM, and 300nM) of recombinant HP1BP3. Arrowhead marks free RNA substrate, whereas bracket corresponds to HP1BP3-bound pri-miRNA or pre-miRNA. **(B)** Native gel-shift assays were performed by using excess non-radiolabeled pri-let-7a, pre-let-7a, or duplex let-7a RNA to compete with the binding of recombinant HP1BP3 (133nM) to radiolabeled pri-let-7a RNA. (Courtesy of Dr. Chunyang Liang)

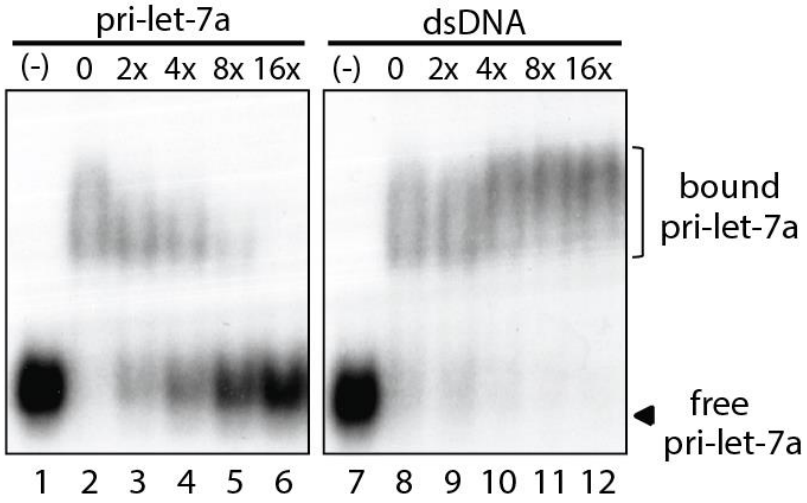

Figure 4-6 HP1BP3 binds to pri-miRNA and DNA via different mechanism.
Native gel-shift assays were performed by using excess non-radiolabeled pri-let-7a RNA (left panel) or dsDNA (right panel) to compete with the binding of recombinant HP1BP3 (133nM) to radiolabeled pri-let-7a RNA. (Courtesy of Dr. Chunyang Liang)

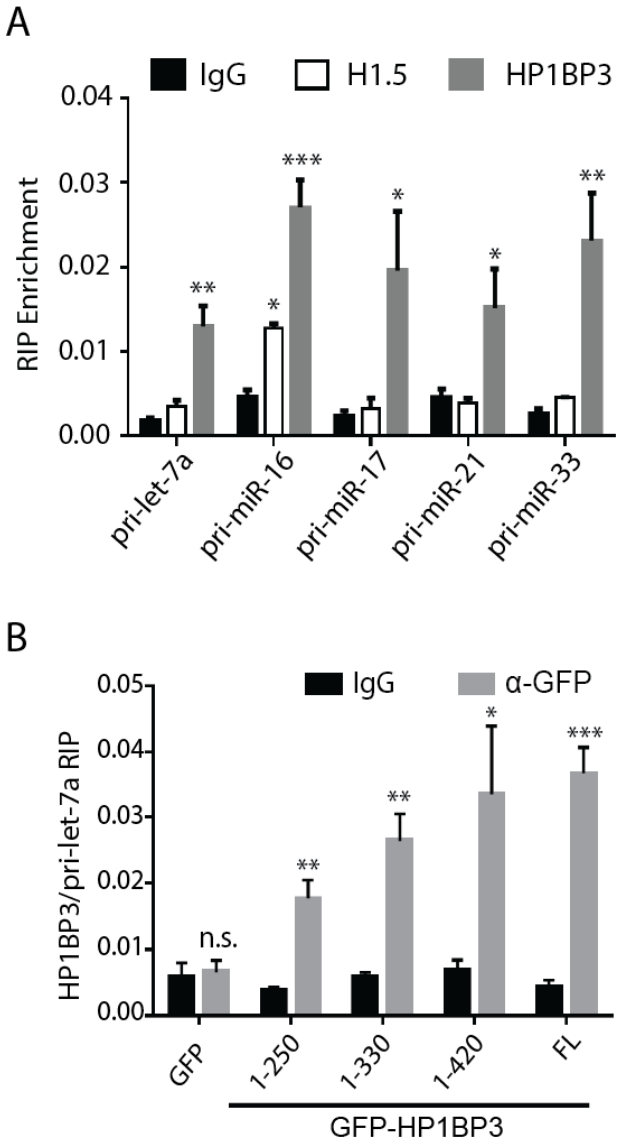

Figure 4-7 HP1BP3 specifically associates with pri-miRNA *in vivo*.
(A) RNA immuneprecipitation (RIP) assays were performed to measure *in vivo* association of H1.5-GFP and HP1BP3-GFP with endogenous pri-miRNA transcripts by (triplicate experiments, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** p< 0.001, ** p <0.01, * p<0.05). **(B)** RIP assays were performed to measure *in vivo* association of GFP-tagged full-length and various truncated constructs with endogenous pri-let-7a in HeLa cells (triplicate samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** p< 0.001, ** p <0.01, * p<0.05).

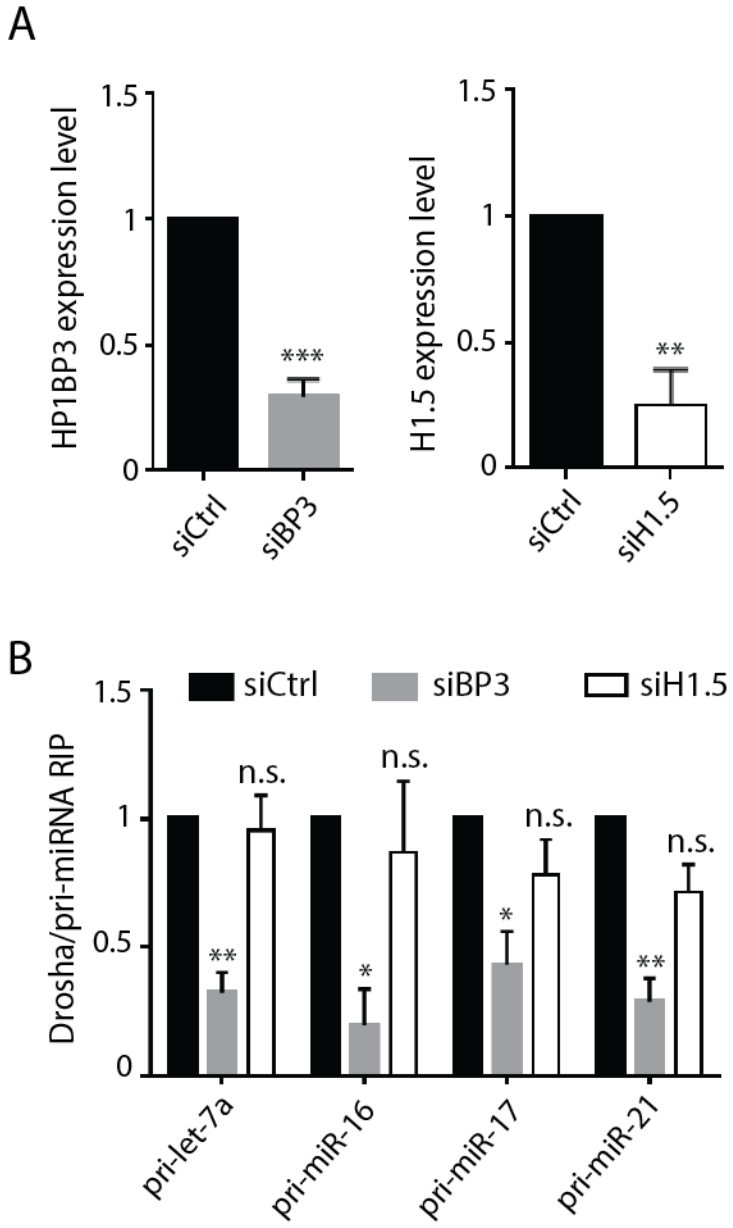

Figure 4-8 HP1BP3 enhances Drosha's association with pri-miRNA *in vivo*.
(A) Quantitative analysis of HP1BP3 and H1.5 expression level after siRNA-mediated knockdown of HP1BP3 (left) and H1.5 (right) by RT-qPCR (triplicate samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$).
(B) RIP assays were performed to measure *in vivo* association of endogenous Drosha with pri-miRNAs between the control, HP1BP3-depleted, and H1.5-depleted HeLa cells (triplicate samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$). (Courtesy of Dr. Masayuki Matsui and Dr. David R. Corey)

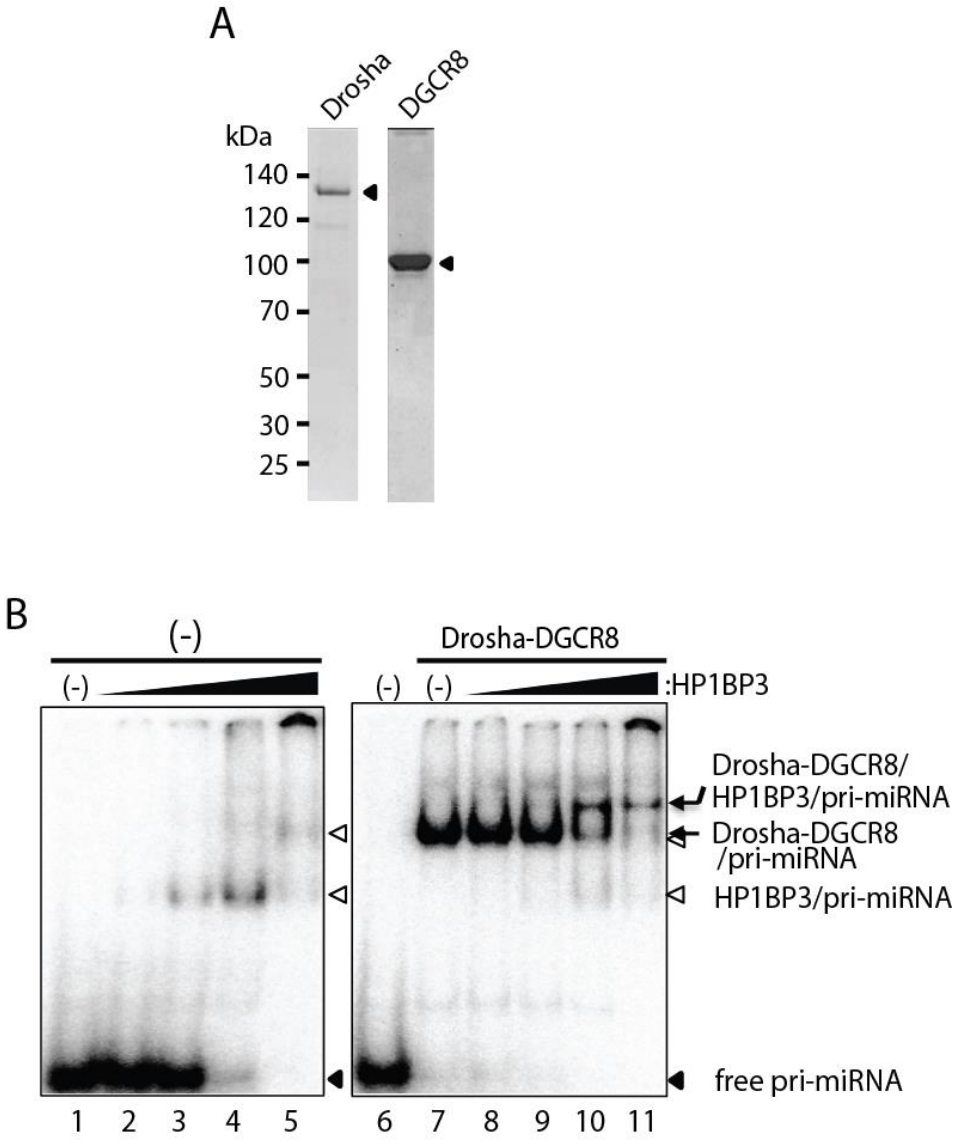

Figure 4-9 HP1BP3 co-binds with Drosha/DGCR8 on pri-miRNA
(A) Coomassie-stained SDS-PAGE showing purified truncated Drosha (aa 398 – 1374) and DGCR8 (aa 276 – 751) recombinant proteins. **(B)** Native gel-shift assay was performed by incubating radiolabeled pri-let-7a in buffer alone (lanes 1 and 6) or increasing concentration (65nM, 125nM, 250nM, 500nM) of recombinant HP1BP3 in the absence (lanes 2-5) or presence (lanes 7-11) of recombinant Drosha-DGCR8 complex (25nM). (Courtesy of Dr. Byung-Cheon Jeong,, Kyoung Shin Yoo, and Dr. Yunsun Nam)

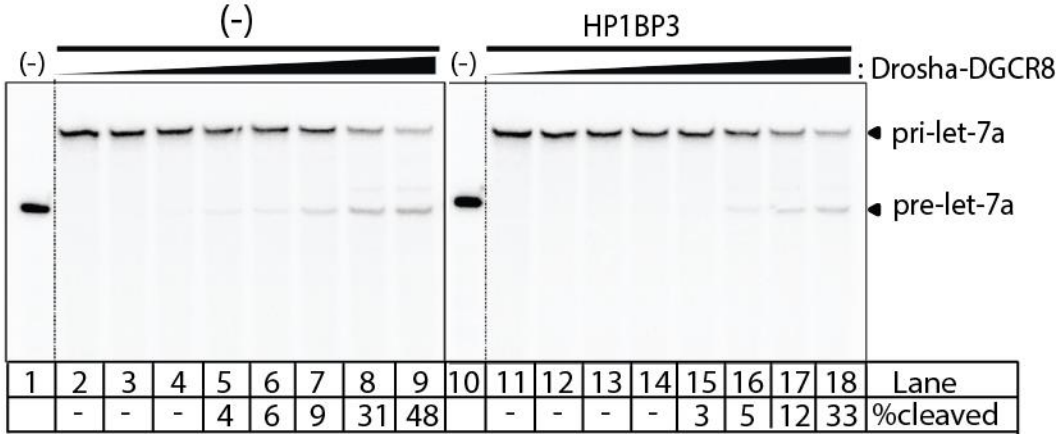

Figure 4-10 HP1BP3 does not enhance processing of pri-miRNA by *in vitro*.
The pri-miRNA processing assay was performed by incubating radiolabeled pri-let-7a with increasing concentration of recombinant Drosha-DGCR8 complex in the absence (lanes 1-9) and presence (lanes 10-18) of recombinant HP1BP3. Synthetic pre-let-7a was used as a marker (lanes 1 and 10). The pri-miRNA processing activity is measured by the ratio of pre-let-7a/(pri-let-7a+ pre-let-7a) as listed below each lane. (Courtesy of Dr. Byung-Cheon Jeong,, Kyoung Shin Yoo, and Dr. Yunsun Nam)

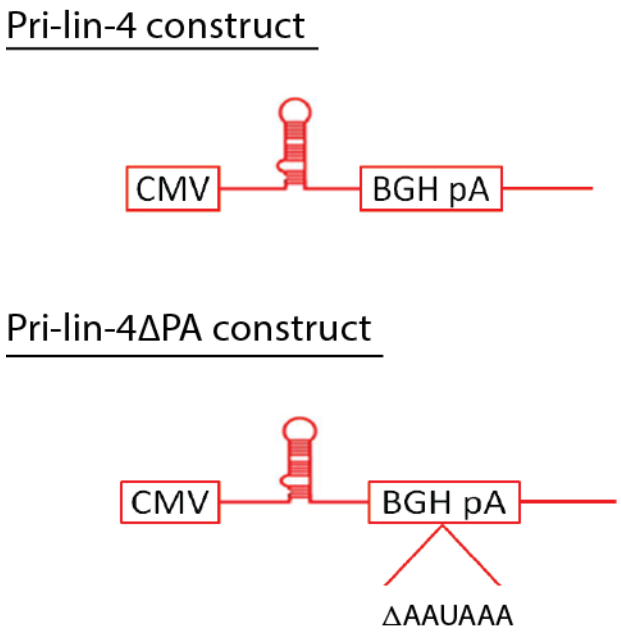

Figure 4-11 Schematic diagram of Pri-lin-4 and Pri-lin-4ΔPA construct (Pawlicki et al., 2008).

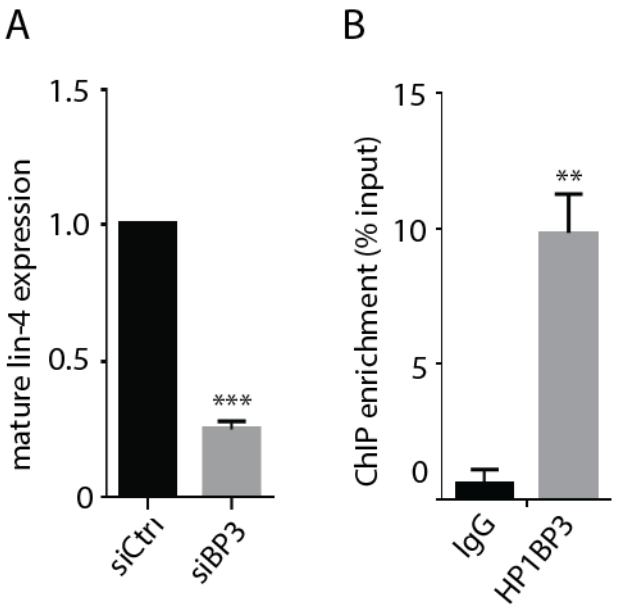

Figure 4-12 Processing of pri-lin-4 can be regulated by HP1BP3.
(A) Quantitative analysis of mature lin-4 expression from transfected Pri-lin-4 plasmid between the control and HP1BP3-depleted Hela cells by Taqman qPCR (triplicate samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$). **(B)** ChIP analysis of HP1BP3-GFP binding to the transfected pri-lin4 plasmid DNA in the HP1BP3-GFP BAC transgenic HeLa cells (triplicate samples, data shown as mean ± SD). Paired t-test was used for statistical analysis (** $p < 0.01$).

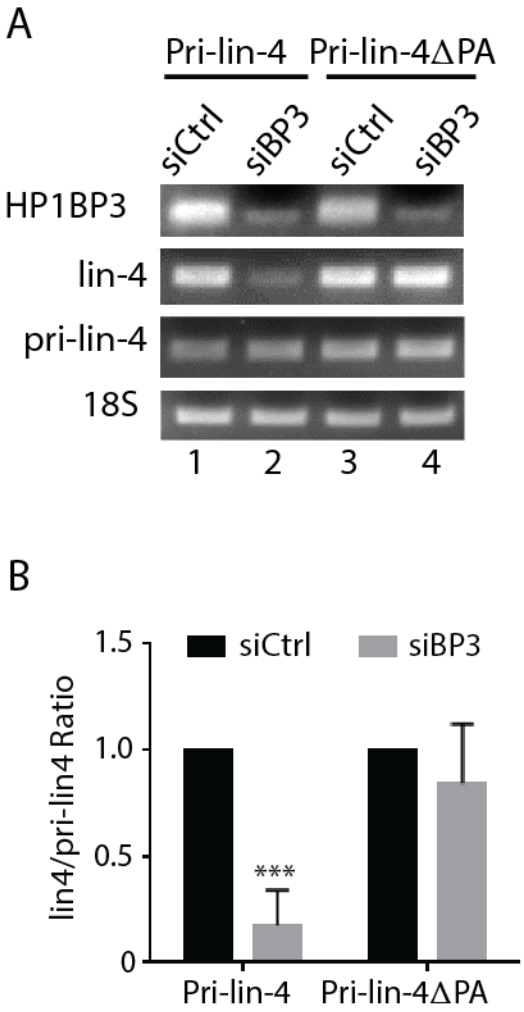

Figure 4-13 Processing of pri-lin-4ΔPA is independent of HP1BP3.
(A) Agarose gel electrophoresis images showing semi-quantitative RT-PCR results that compare the pri-lin-4 and mature lin-4 expression from transfected Pri-lin-4 and Pri-lin-4ΔpA constructs, respectively, between the control and HP1BP3-depleted HeLa cells.
(B) Quantitative analysis (by Taqman qPCR) of the pri-lin-4 and mature lin-4 expression from transfected Pri-lin-4 and Pri-lin-4ΔpA constructs between the control and HP1BP3-depleted HeLa cells (triplicate experiments, data shown as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$).

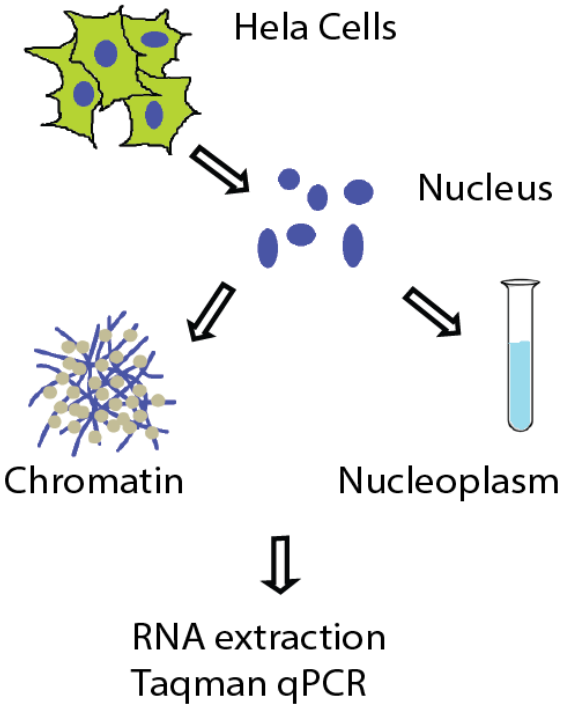

Figure 4-14 A schematic diagram of the procedure for chromatin and nucleoplasm fractionation.

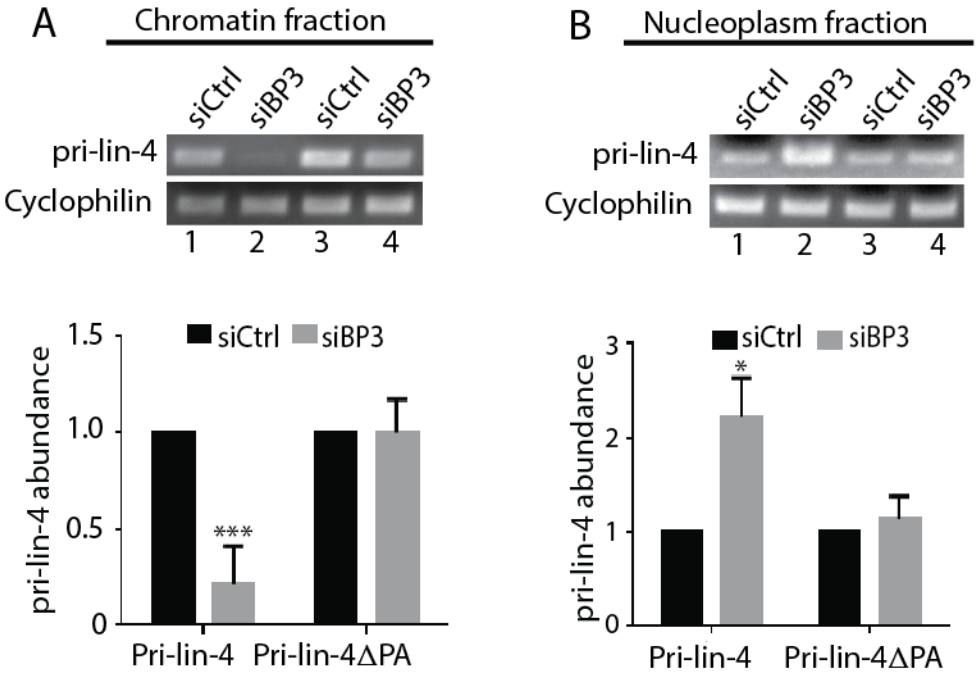

Figure 4-15 HP1BP3 enhances pri-lin-4 processing by retaining its nascent transcript on chromatin.

(A) Semi-quantitative RT-PCR (top) and Taqman qPCR (bottom) were performed to measure the abundance of chromatin-associated pri-lin-4 transcripts expressed from transfected Pri-lin-4 and Pri-lin-4ΔpA constructs between the control and HP1BP3-depleted HeLa cells (triplicate samples, data present as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$). **(B)** Semi-quantitative RT-PCR (top) and Taqman qPCR (bottom) were performed to measure the abundance of nucleoplasmic pri-lin-4 transcripts expressed from transfected Pri-lin-4 and Pri-lin-4ΔpA constructs between the control and HP1BP3-depleted HeLa cells (triplicate samples, data present as mean ± SD). Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$).

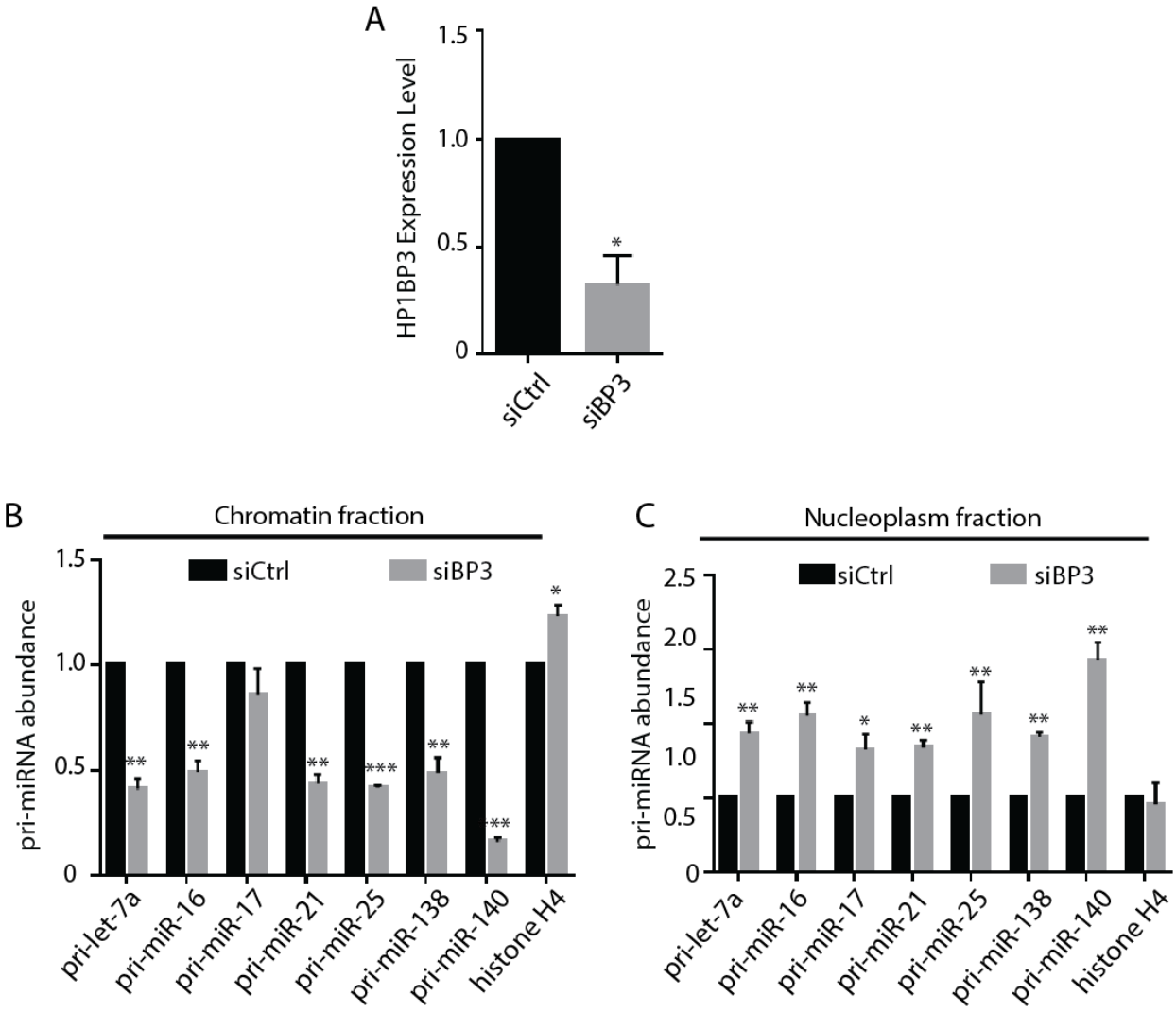

Figure 4-16 HP1BP3 retains nascent pri-miRNA transcript on chromatin.
(A) Following siRNA-mediated knockdown in HeLa cells, the mRNA level of HP1BP3, H1.2 or H1.5 was respectively measured by real time RT-qPCR. Paired t-test was used for statistical analysis (* $p < 0.05$). **(B, C)** Taqman qPCR were performed to measure of the relative abundance of endogenous pri-miRNA and histone H4 transcripts in the chromatin (G) and nucleoplasm (H) fractions between the control and HP1BP3-depleted HeLa cells (quadruplet samples, data present as mean ± SD). Notably, primers for amplifying the introns of nascent pri-miR-25 and pri-miR-140 transcripts were used. 18S rRNA was used as a loading control. Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$).

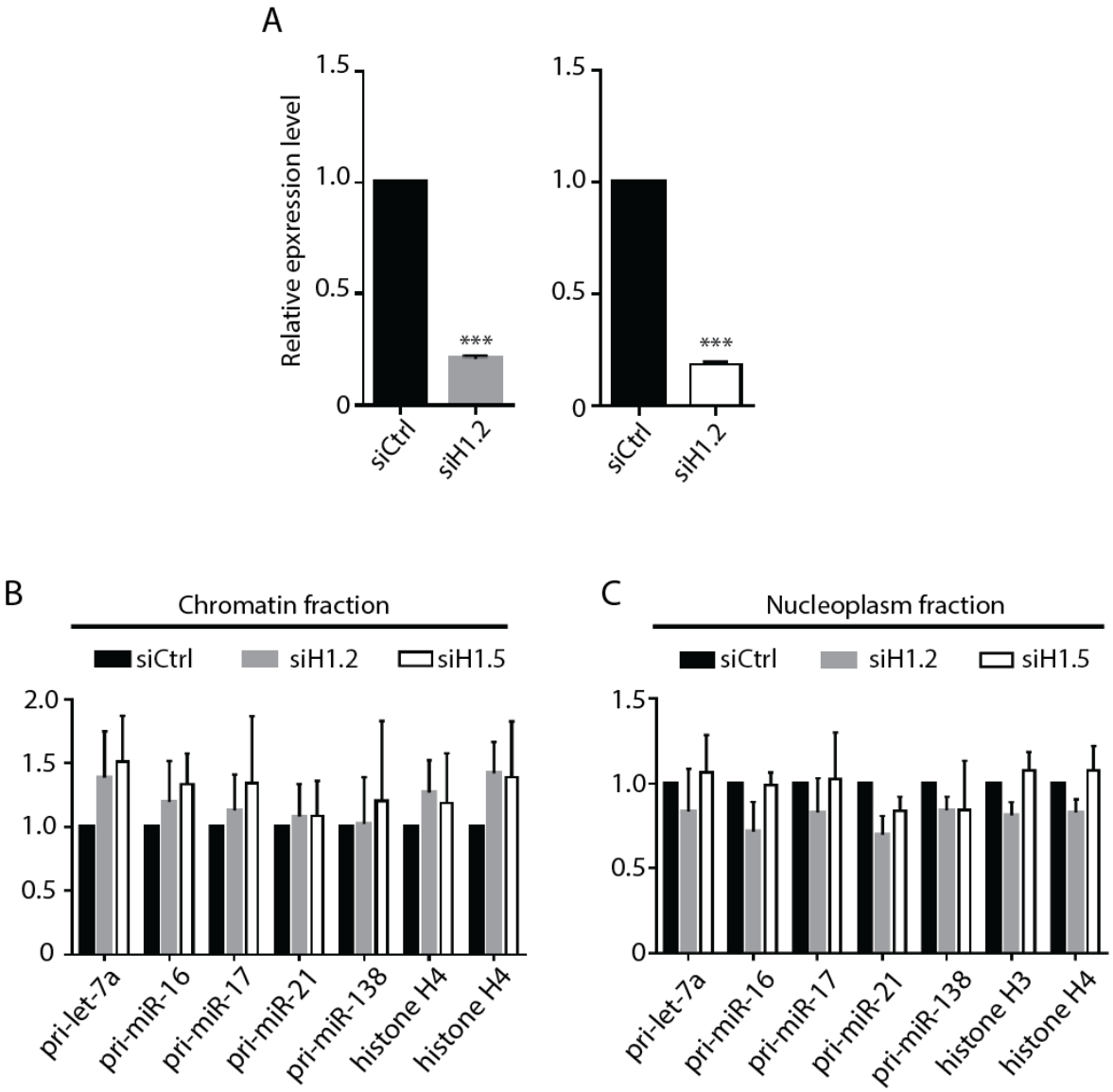

Figure 4-17 Neither H1.2 nor H1.5 plays a significant role in the retention of pri-miRNA transcript on chromatin.

(A) Following siRNA-mediated knockdown in HeLa cells, the mRNA level of HP1BP3, H1.2 or H1.5 was respectively measured by real time RT-qPCR. Paired t-test was used for statistical analysis (*** $p < 0.001$). **(B, C)** Taqman qPCR was performed to measure the relative abundance of endogenous pri-miRNA and histone H4 transcripts in the chromatin (D) and nucleoplasm (E) fractions of the control, H1.2-depleted and H1.5-depleted HeLa cells (triplet samples, data present as mean ± SD). 18S rRNA was used as a loading control. Paired t-test was used for statistical analysis (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$).

CHAPTER FIVE

CONCLUSION AND FUTURE DIRECTIONS

A new concept: chromatin retention factor for co-transcriptional miRNA processing

In this study, we find that HP1BP3, a histone H1-like chromatin protein, possesses a novel pri-miRNA binding activity and promotes co-transcriptional miRNA processing *via* chromatin retention of nascent pri-miRNA transcripts. Based on our findings, we propose a working model for co-transcriptional miRNA processing in mammalian cells **(Figure 5-1)**: First, the pri-miRNA transcript is transcribed by RNA polymerase II (Pol II) from the miRNA gene. Second, HP1BP3 binds both chromatin DNA and nascent pri-miRNA transcript to help retain pri-miRNA on chromatin. Third, retention of pri-miRNA at the site of transcription provides more opportunity for Drosha-DGCR8 complex to bind pri-miRNA on chromatin. Moreover, since HP1BP3 associates with both Drosha/DGCR8 and pri-miRNA, it may actively facilitate the Microprocessor to bind and/or process pri-miRNA co-transcriptionally *in vivo*. On the other hand, in the absence of HP1BP3, nascent pri-miRNA transcript prematurely dissociates with chromatin, and may translocate to nuclear foci where it is inaccessible to Drosha/DGCR8 complex **(Figure 5-2)**.

In support of this model, we showed that HP1BP3 was involved in global miRNA biogenesis, and specifically, at the step of pri-miRNA processing in human cells. ChIP analysis revealed genome-wide chromatin co-localization of HP1BP3 & Drosha and HP1BP3-dependent Drosha binding to actively transcribed miRNA loci. Importantly, knockdown of HP1BP3 compromises co-transcriptional miRNA processing by resulting in

premature release of pri-miRNAs from the chromatin and reduction in the Drosha-pri-miRNA association. In the nucleoplasm, the released pri-miRNAs may become vulnerable for nuclease-mediated degradation or localize in regions, such as the SC35 bodies, that prevent efficient binding or processing by the Drosha-DGCR8 complex (Pawlicki and Steitz, 2009). Furthermore, we found that *Drosophila* H1 exhibited a similar pri-miRNA binding activity and promoted miRNA processing in S2 cells. Thus, it will be interesting for future studies to investigate whether histone H1 and related proteins play a conserved role in co-transcriptional miRNA processing in flies, worms, and other organisms.

Regulation of co-transcriptional pri-miRNA processing

As the initial step of miRNA biogenesis, it is understandable that pri-miRNA processing is one of the most highly regulated steps in the miRNA pathway. Recent studies have identified a number of proteins that regulate pri-miRNA processing. For example, TAR DNA-binding protein 43 (TDP-43) controls the stability of Drosha protein and thus, globally affecting the production of miRNAs during neurodifferentiation (Di Carlo et al., 2013). MECP2, a methyl CpG DNA binding protein linked to neurodevelopmental disorders such as RETT syndrome and autism, disrupts global pri-miRNA processing by interfering with the assembly of Drosha-DGCR8 complex (Cheng et al., 2014). Moreover, transcription activators, such as SMADs and p53, promote the processing of a specific set of pri-miRNAs through interaction with DDX5/p68 helicase that interacts with the Microprocessor (Davis et al., 2008; Suzuki et al., 2009). RNA-binding proteins, such as LIN28, hnRNPA1 and KSRP, recognize the terminal loop of a subset of pri-miRNAs and inhibit or enhance their

processing by the Microprocessor (Guil and Caceres, 2007; Newman et al., 2008; Trabucchi et al., 2009). In most of these studies, however, it is unclear whether these regulatory events occur on the chromatin or in the nucleoplasm.

Our comprehensive ChIP analyses suggest that most if not all pri-miRNA transcripts undergo co-transcriptional miRNA processing. Therefore, a logical prediction of this result would be that many known regulators of pri-miRNA processing should in effect regulate co-transcriptional miRNA processing on chromatin. It is entirely plausible for transcription activators (e.g. SMADs and p53) to bind specific chromatin DNA elements and recruit the Microprocessor to adjacent pri-miRNA transcripts to facilitate co-transcriptional miRNA processing. LIN28 has recently been shown to bind nascent pri-let-7 transcript co-transcriptionally and inhibits Drosha-mediated processing of pri-let-7 in *C. elegans* and human ES cells (Van Wynsberghe et al., 2011). Moreover, FUS/TLS (fused in sarcoma/translocated in liposaroma) protein, an RNA binding protein that is linked to Amyotrophic Lateral Sclerosis (ALS), stimulates co-transcriptional miRNA processing by facilitating Drosha recruitment to specific miRNA loci (Morlando et al., 2012). It is possible that some of these RNA binding proteins may also possess DNA binding activity or interacts with chromatin factors to promote co-transcriptional miRNA processing. Finally, our current study suggests that HP1BP3 promotes co-transcriptional miRNA processing *via* chromatin retention of nascent pri-miRNA transcripts. Therefore, mounting evidence suggest an emerging theme that the chromatin regulation of co-transcriptional miRNA processing is a widespread and conserved mechanism in eukaryotes.

Essential roles of HP1BP3 in miRNA biogenesis and during development

HP1BP3 is ubiquitously expressed in all somatic tissues and is highly enriched in the brain (Garfinkel et al., 2015b). Knockout of HP1BP3 in mice causes early post-natal lethality (~60% pups die within 24 hours after birth) (Garfinkel et al., 2015a; Garfinkel et al., 2015b). Homozygous Hp1bp3$^{-/-}$ mice that survive to weaning are fertile and have a normal life span, but are significantly smaller than their littermates since birth (Garfinkel et al., 2015a; Garfinkel et al., 2015b). These mutant mice show proportional reduction in body weight, body length and organ weight as well as severe impairment in bone development. However, wild-type and Hp1bp3$^{-/-}$ mouse embryonic fibroblast (MEF) exhibit a similar sensitivity of chromatin to micrococcal nuclease digestion (Garfinkel et al., 2015b). Moreover, the lack of HP1BP3 only impacts the expression of a limited set of genes (Garfinkel et al., 2015b). These observations suggest that the phenotypes of Hp1bp3$^{-/-}$ mice are probably not due to a global defect in chromatin organization. Rather, that HP1BP3 promotes global miRNA processing provides a likely explanation for the lethality and growth defects of Hp1bp3$^{-/-}$ mice.

Conclusions

The H1 family of linker histones is often assumed to be non-specific DNA binding proteins with prosaic and well-understood functions. However, it is still a mystery why there is a major expansion from single H1 in flies to eleven H1 variants plus HP1BP3 in mammals. Because of their proximity to chromatin, histone H1 and related proteins are well positioned to play important roles in various transcriptional and co-transcriptional processes. The

possibility of unanticipated functions of various linker histones is underscored by our surprising discovery that HP1BP3 possesses a novel pri-miRNA binding activity and promotes co-transcriptional miRNA processing in human cells. The current work significantly expands the functional repertoire of the H1 family of proteins, and suggests a new mechanism that chromatin factors retain nascent transcripts at the site of transcription to enhance co-transcriptional RNA processing.

Future Directions

Our study establishes that HP1BP3 can facilitate miRNA biogenesis through enhancing retention of nascent pri-miRNA on chromatin, allowing more efficient processing by Drosha/DGCR8 complex. However, there are still outstanding questions that remain unanswered in our current model. First, it remains to be seen if the mechanism if miRNA biogenesis regulation is conserved in Drosophila. Co-IP experiments will elucidate interaction between *Drosophila* H1 and Drosha/Pasha Microprocessor complex in fly. ChIP experiments will determine if *Drosophila* H1 co-localizes with Drohsa/Pasha to pri-miRNA transcription loci and its involvement in recruiting Drohsa to nascent pri-miRNA. Most importantly, it is still unclear if *Drosophila* H1 participates in enhancing chromatin retention of nascent pri-miRNAs. Attempts were made to knockdown *Drosophila* H1 in S2 cells and measure chromatin retention of endogenous fly pri-miRNAs. However, the stability of H1 knockdown was difficult to control due to toxicity involved in depleting an essential chromatin protein. The results obtained were promising but inconsistent. Because unlike in human where there are as many as 13 different H1 isoform, there is only one H1 isoform in

the *Drosophila*; therefore, mechanistic study of *Drosophila* H1 faces the ultimate challenge of distinguishing its role in pri-miRNA processing from its other multitude of functions as an essential chromatin factor.

For HP1BP3, more detailed mechanism remains undiscovered. For example, HP1BP3 is characterized by three GDs flanked by NTD and highly basic CTD. Our preliminary study suggests that CTD is largely responsible for DNA binding, while three GDs cooperatively bind to pri-miRNA. Therefore, our model proposes that HP1BP3 functions as a molecular link connecting chromatin to nascent pri-miRNA transcript. Removal of one or two GDs partially retains HP1BP3's binding to both pri-miRNA and Drosha/DGCR8 complex. However, whether or not the truncated HP1BP3 is biologically functional remains to be seen. To elucidate the functional significance of each of HP1BP3's domain, I propose to use the rescue experiment that we established in Chapter Three. Instead of full length HP1BP3, different truncated version of HP1BP3-GFP construct will be transfected into Hela cells depleted of endogenous HP1BP3. Both processing efficiency (by measuring pri-miRNA to mature miRNA ratio) and chromatin retention time (by chromatin fractionation) of pri-miRNA will be measured. I hypothesize that removing either CTD or all three GDs will abolish HP1BP3's involvement in pri-miRNA processing, but removing one or two GDs will still allow HP1BP3 to function partially. We also could not rule out the possibility that HP1BP3 may enhance pri-miRNA processing through an unknown pathway independent of its role in pri-miRNA chromatin retention, of which we hope to reveal through our functional study of each domain of HP1BP3.

In our study, we have found that HP1BP3 not only can regulate the processing of endogenous pri-miRNAs, but also exogenous pri-miRNAs (pri-lin-4) from a transfected construct (Figure 4-12). With an inducible Lin-4 system already established, the role of HP1BP3 in pri-miRNA processing can be studied more dynamically. For example, using ChIP experiment, we can elucidate whether HP1BP3 or the microprocessor is recruited first to the site of transcription. Lastly, Joan Steitz's group claimed that unprocessed pri-lin-4 RNA falls off chromatin and translocate to nuclear foci inaccessible to the microprocessor (Pawlicki and Steitz, 2008). It remains to be seen if this is also true for endogenous pri-miRNA as well. Using fluorescent *in situ* hybridization (FISH), we can visualize the localization of endogenous pri-miRNA in WT HeLa cells as well as HP1BP3-depleted HeLa cells. Additionally, using the GFP-Drosha BAC transgenic HeLa cell line established in our study, Drosha's accessibility to these RNA foci can also be evaluated.

In eukaryotic cells, the processing (e.g. capping, splicing, and 3' end formation) of precursor pre-mRNA also happens co-transcriptionally (Bentley, 2014). It is likely that HP1BP3, or some other protein in the histone H1 family, may also facilitate the co-transcriptional processing of pre-miRNA. As demonstrated in Chapter Four, HP1BP3 recognize pri-miRNAs through their secondary structure. This observation raises the possibility that HP1BP3 may also interact with other RNAs that contain stem-loop structure. Additionally, preliminary data indicates that HP1BP3 reduces chromatin retention of actin pre-mRNA, suggesting that HP1BP3 may play an important role in pre-mRNA splicing. To investigate this hypothesis, we can design primers specifically measure pre- and post-spliced mRNA and measure their relative level in WT and HP1BP3 depleted cells. Co-IP assay can

be employed to detect association between HP1BP3 and known spliceosomes. Chromatin fractionation assay will reveal if HP1BP3 can enhance pre-mRNA chromatin retention, similarly to pri-miRNA.

Finally, our ChIP-seq results show that HP1BP3 and Drosha co-localize to genomic regions with no pri-miRNA transcription. Recent studies have also shown that Drosha participate in functions independent of its pri-miRNA processing, such as modulating gene expression (Gromak et al., 2013) and cleavage of many classes of RNAs, such as messenger RNA (mRNA) (Han et al., 2009; Kadener et al., 2009), ribosomal RNA (rRNA) (Wu et al., 2000), and long non-coding RNA (lncRNA) (Ganesan and Rao, 2008). It is still unclear if HP1BP3 also functions as a co-factor with Drosha in these processes.

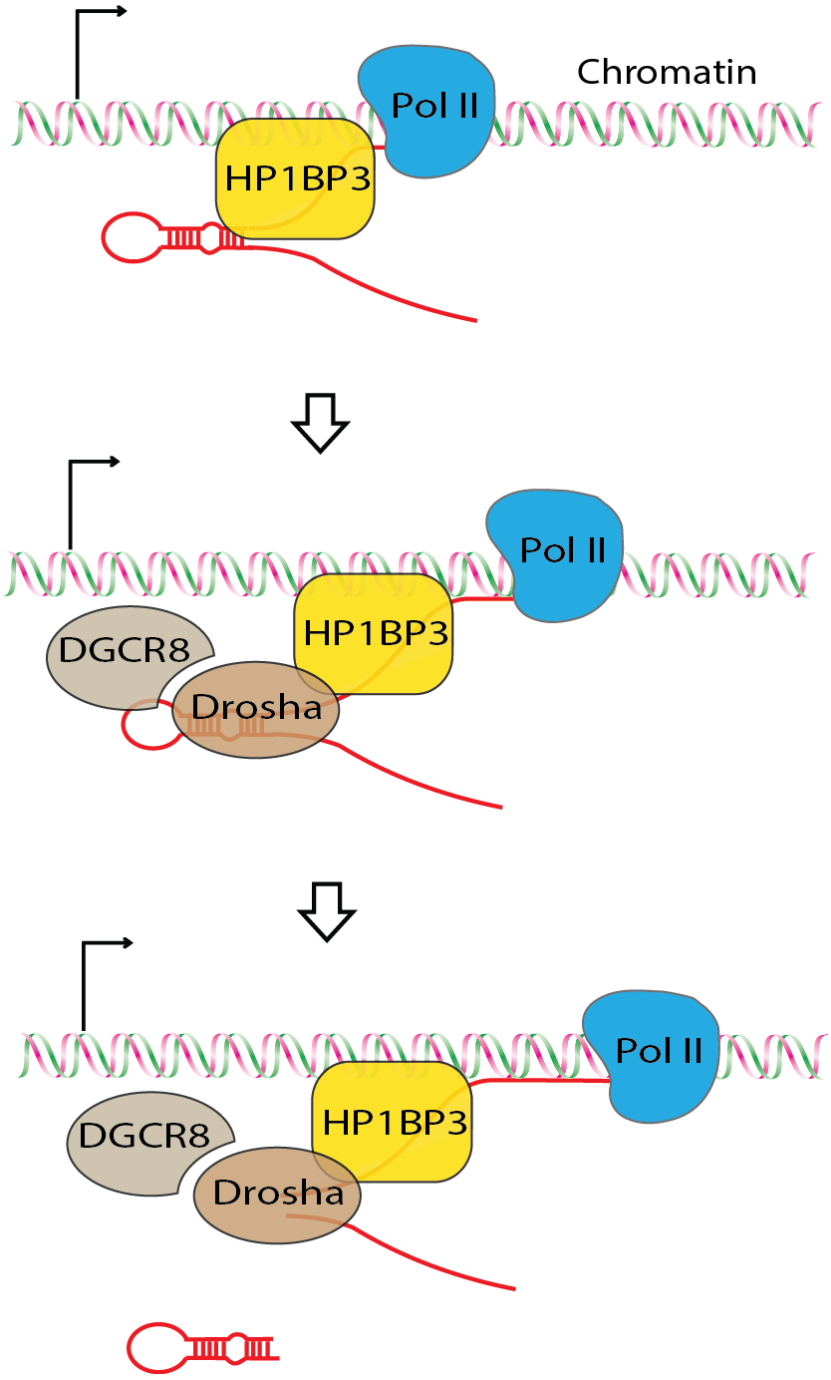

Figure 5-1 A model of co-transcriptional pri-miRNA processing in the presence of HP1BP3.
A schematic diagram showing that HP1BP3 promotes co-transcriptional pri-miRNA processing through chromatin retention of nascent pri-miRNA transcript.

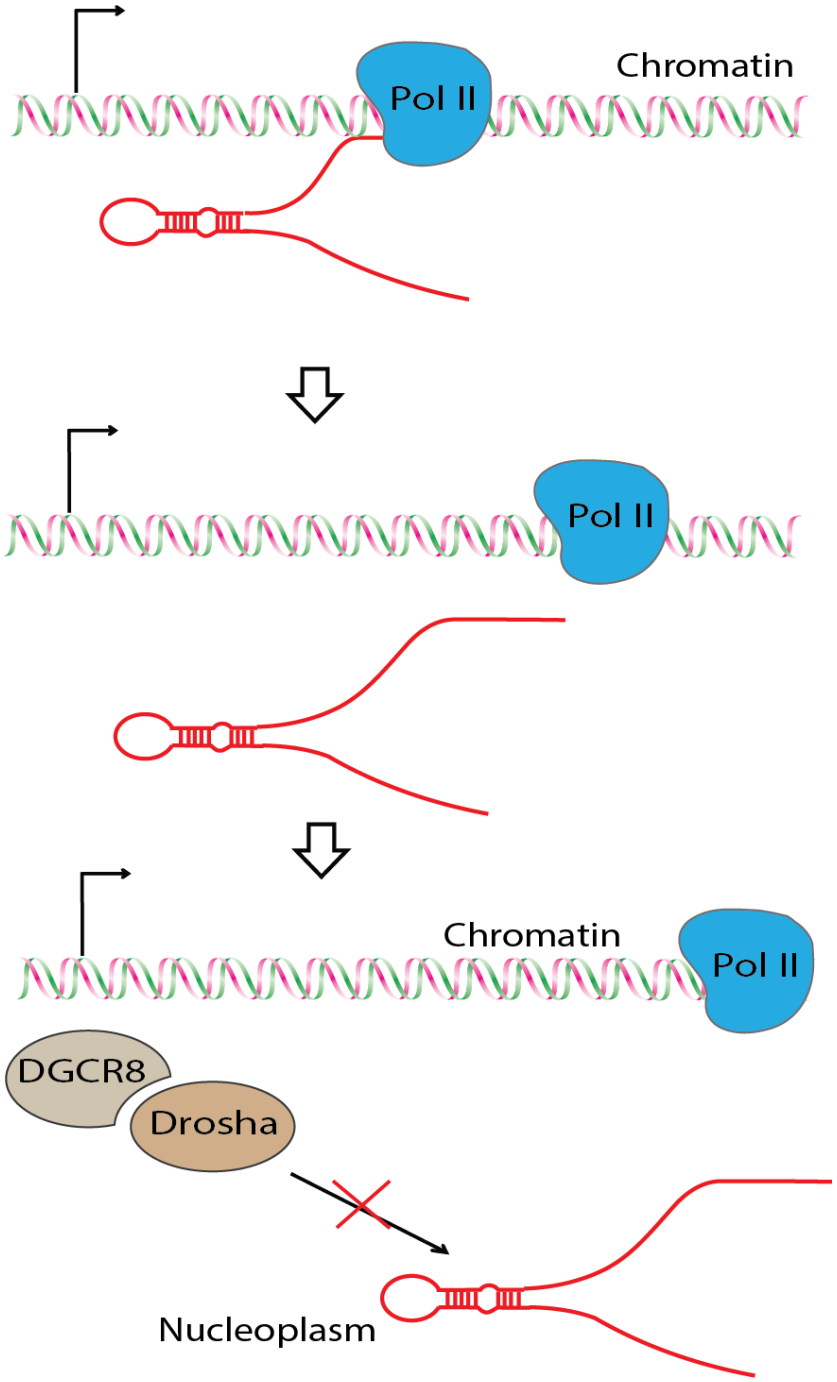

Figure 5-2 A model of co-transcriptional pri-miRNA processing in the absence of HP1BP3.
A schematic diagram showing that in the absence of HP1BP3, pri-miRNA is prematurely released from chromatin, leading to inefficient processing by Drosha/DGCR8 complex.

CHAPTER SIX

MATERIAL AND METHODS

Purification of fly histone H1 as a pri-miRNA-binding protein

Drosophila S2 cells were grown in suspension in SFX media (Invitrogen), harvested at ~2 x 10^7 cells/ml, and washed twice by PBS. Cells were incubated in five volumes of hypotonic buffer A (10 mM KOAc, 10 mM HEPES, pH7.4, 2 mM $Mg(OAc)_2$, 5mM DTT) supplemented with fresh proteinase inhibitorz on ice for 20 minutes. Then, cells were broken with a Douncer by 40 strokes and centrifuged at 2,000g at 4 ºC for 30 minutes. The nuclei pellet was washed twice in buffer A followed by centrifugation at 2,000 *g* for 10 min, and was resuspended in five volumes of nuclear extraction buffer (20 mM HEPES, pH 7.4, 10% glycerol, 350 mM NaCl, 0.1% Triton X-100, 1 mM DTT) supplemented with fresh proteinase inhibitors at 4°C for 1 hour. After centrifugation at 20,000 *g* for 30 min at 4°C, the S10 supernatant was further centrifuged at 100,000g for 1 hour at 4ºC to obtain nuclear S100 extract. The purification of histone H1 was carried out using a three-step chromatographic purification procedure (Nuclear S100 → SP-Sepharose → Q-Sepharose → SP-Sepharose) on the ÄKTA FPLC system (GE Healthcare Inc) at 4 ºC. First, the nuclear S100 extract was fractionated by SP-Sepharose column. The pri-miRNA binding activity was measured by native gel-shift assay by incubating individual fractions with radiolabeled pri-miRNA. The fractions with peak activity were dialyzed overnight and fractionated on Q-Sepharose column. The peak activity fractions were dialyzed overnight and fractionated again by SP-Sepharose column. After the final step, individual fractions were examined for the priRNA-

binding activity and resolved by SDS-PAGE followed by silver staining. After silver staining, the protein bands that correlated with pri-miRNA-binding activity were excised, digested by trypsin, and identified by Mass Spectrometry analysis.

Generation of recombinant histone proteins

Polyhistidine-tagged *Drosophila* histone H1, H2A, H3 and H4 and human HP1BP3 recombinant proteins were expressed in the pET-28a vector in *E. coli* BL21(DE3) strain as previously described (Liang et al., 2015). The various His-tagged recombinant proteins were purified using Ni-NTA agarose (Qiagen) at 4ºC followed by Q-Sepharose chromatography.

Native gel-shift assays

Native gel-shift was carried as described (Liang et al., 2013), pri-miRNA (more description), pre-miRNA, duplex miRNA and dsDNA (180-nt pri-let-7a DNA sequence) were resepectively 5' radiolabeled with [γ-^{32}P] ATP by T4 polynucleotide kinase (PNK) followed by purification by G25 column (Ambion). Typically, recombinant proteins and 5×10^4 cpm radiolabeled RNA were incubated at 37 °C for 30 min in a 10-µl reaction (100 mM KOAc, 15 mM HEPES pH7.4, 2.5 mM EDTA, 2.5 mM DTT, pH7.4). The reaction mixture was resolved by a 4.5% native PAGE and exposed to X-ray film.

RNAi knockdown in S2 cells

Different dsRNAs were prepared using Megascript T7 *in vitro* transcription kit (Ambion). Briefly, 1 mL 2 x 10^6 S2 cells were soaked with 30 ug GFP and histone H1 or H3

dsRNA for 6 hours in serum-free SFX media in 6-well dish, then add 2 ml medium to the cells. After 48 hours, the media was removed and dsRNA treatment was repeated. After another 48 hours, the dsRNA treated S2 cells were harvested for Western blotting, Northern Blotting, and real-time RT-qPCR.

Multi-sequence alignment and phylogenetic tree analysis

Linker histone H1 sequences were collected with PSI-BLAST (Altschul et al., 1997) against the UniProtKB/Swiss-Prot (Magrane and Consortium, 2011) database (default cutoffs, 5 iterations to convergence) initiated from a query sequence (gi|6016184, residue range 32-129). The query sequence range includes the H1 globular domain (GD) defined by the conserved domain database (CDD)(Marchler-Bauer et al., 2015) with some additional anchoring sequence on either side. The resulting 145 sequences producing significant alignments were from animals (112), green plants (24), fungi (6), Olisthodiscus (2), and cellular slime mold (1), with human possessing 14 variants. One reported human H1 sequence (PMID: 26474902) encoded by the H1T2 gene was not identified by this method, so we also performed more sensitive sequence searches against the human genome with the same query using the HHPRED server (Hildebrand et al., 2009). HHPRED confidently identified human sequences with probability > 95% that corresponded to the same genes identified using PSI-BLAST, as well as the H1T2 sequence with slightly lower confidence (probability 92.26%). Taxonomic distribution of HP1BP3 sequences was defined with BLAST using the full-length human sequence as a query, restricting identified sequences to those containing tandem G15 domains defined by CDD. Selected sequences corresponding to

the GD were aligned using MAFFT(Katoh and Standley, 2014), and the resulting multiple sequence alignment was used to estimate phylogenetic distances (JTT with frequencies option) and build a maximum likelihood tree using local rearrangement of an initial NJDIST tree with the Molphy package(Adachi and Hasegawa, 1992). Because the H1T2 sequence is more distantly related to the other H1 GDs, we displayed the tree using H1T2 as an outgroup.

Mammalian cell culture and RNAi knockdown

HeLa and U2OS cells were cultured at 37 °C in 5% CO_2 in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin /streptomycin. At 70% confluency, HeLa/U2OS cells were transfected with 20nM of siRNAs (Table S4) by Lipofectamine RNAiMax (Thermofisher Scientic) according to manufacturer's protocol. A second siRNA transfection was repeated 48 hours later. Transfected cells were harvested after another 48 hours for further analysis, e.g. RT-qPCR, Northern blotting, Western blotting, ChIP, RIP, and chromatin fractionation. Total RNA was extracted using Trizol® reagent (Thermofisher Scientic). The knockdown efficiency of HP1BP3 and H1 variants in Figure 1B was evaluated by RT-qPCR using primers listed in Table S5. The knockdown efficiency of HP1BP3 and H1.5 is further examined by Western blotting.

Northern blotting

Northern blotting was performed to measure expression of pre- and mature miRNAs in the control and H1.5 or HP1BP3-depleted HeLa cells as previously described (Liang et al.,

2013). In brief, 30 µg total RNA was resolved by 12% Urea-PAGE, transferred to GT membrane (Bio-Rad), and crosslinked by ultra violet (UV) light. The sequences of antisense RNA probes for detection of miRNAs and 5S RNA are listed in Table S7. The RNA probes were 5′ radiolabeled with γ-^{32}P ATP by T4 polynucleotide kinase (NEB). Hybridization was carried in Ultrasensitive Hybridization solution (Ambion) at 40°C overnight. The membrane was washed 3 times at 40°C for 15 minutes with 2 ×SSC, 0.5% SDS followed by autoradiography.

Small RNA sequencing

Sequence libraries were filtered for adapter contamination using cutadapt (v1.8.3) (Martin, 2011) software tool. This software cut the adapter sequence from the sequencing reads and filtered the reads whose length is greater than or equal to 15bp for further analysis. Filtered reads were aligned to the human reference genome (hg19) using Bowtie (v.2.2.5) (Langmead et al., 2009). Reads that mapped with ≤2 mismatches to the reference sequence were retained for further analysis. For siBP3 replicates one and two we obtained 4,863,552 reads and 4,369,227 reads respectively, while for control replicates one and two we obtained 3,908,404 reads and 1,907,245 reads respectively. We used the 523 re-annotated miRNA list (Fromm et al., 2015), 624 tRNA regions (GenCode) and 1,769 rRNA regions (UCSC genome browser) to count the mapped reads using featurecounts (v1.4.6) module from subread package (Liao et al., 2014). These counts were normalized to library size and performed differential expression analysis using edgeR bioconductor package (v.3.8.6)

(Robinson et al., 2010). Limma (v.3.22.7) (Ritchie et al., 2015) package was used to calculate differential expression change.

Immunostaining and co-immunoprecipitation (IP)

The Drosha-GFP, HP1BP3-GFP and H1.5-GFP BAC transgenic HeLa cells grown on coverslips were fixed with 4% paraformaldehyde in PBS (pH 7.4) for 15 min at 37 °C, and then permeabilized with 0.2% Triton X-100 in PBS for 10 min at room temperature. After washing, the cells were stained with 1 μg/ml 4′,6-diamidino-2-phenylindole (DAPI) and imaged under the LSM 700 Laser Scanning Microscope (Carl Zeiss). For co-IP studies, cell lysate (containing 500μg of total protein) was pre-cleared using 40μL of protein A/G beads for 2 hours at 4 ºC. When necessary, RNaseA (1μg) was added to cell extract for additional 30 minutes of incubation at room temperature. Pre-cleared HeLa extract was first mixed with 2μg of antibody at 4 ºC overnight, then mixed with 40μL protein A/G plus Argarose (Santa Cruz Biotechnology) resin at 4 ºC for 2 hours. Resin was washed 5 times with IP wash buffer (10mM Tris pH 7.5, 50mM NaCl, 2mM $MgCl_2$, 0.05% NP-40, 0.025% SDS). Protein associated with resin was eluted by SDS-PAGE loading buffer at 100 ºC for 2min, and then analyzed by western blot. For IP of Flag-tagged proteins, Anti-Flag (M2) Affinity Gel (Sigma) was used directly in place of Protein A/G resin and antibody.

ChIP-Seq experiments

BAC transgenic cells (primers used to generate BAC transgenic cells listed in table S3) at 80% confluency (~$1x10^7$) were cross-linked with 1% formaldehyde for 10 minutes at

37°C, and quenched with 125 mM glycine at room temperature for 5 minutes. The fixed cells were washed twice with cold PBS, scraped, and transferred into 1 ml PBS containing protease inhibitors (Roche). After centrifugation at 700 g for 4 minutes at 4°C, the cell pellets were resuspended in 100 µl ChIP lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl [pH 8.1] with protease inhibitors) and sonicated at 4°C with a Bioruptor (Diagenode) (30 seconds ON and 30 seconds OFF at highest power for 15 minutes). The sheared chromatin with a fragment length of ~200 – 600 bp) was centrifuged at 20,000 g for 15 minutes at 4°C. 100 µl of the supernatant was used for ChIP or as input. A 1:10 dilution of the solubilized chromatin in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 167 mM NaCl 16.7 mM Tris-HCl [pH 8.1]) was incubated at 4°C overnight with 6 µg/ml of a goat anti-GFP (raised against His-tagged full-length eGFP and affinity-purified with GST-tagged full-length eGFP). Immunoprecipitation was carried out by incubating with 40 µl pre-cleared Protein G Sepharose beads (Amersham Bioscience) for 1 hour at 4°C, followed by five washes for 10 minutes with 1ml of the following buffers: Buffer I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl [pH 8.1], 150 mM NaCl), Buffer II (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl [pH 8.1], 500 mM NaCl), Buffer III (0.25 M LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl [pH 8.1]); twice with TE buffer [pH 8.0]. Elution from the beads was performed twice with 100 µl ChIP elution buffer (1% SDS, 0.1 M NaHCO3) at room temperature for 15 minutes. Protein-DNA complexes were de-crosslinked by heating at 65°C in 192 mM NaCl for 16 hours. DNA fragments were purified using QiaQuick PCR Purification kit (QIAGEN) and eluted into 30 µl H_2O according to the manufacturer's protocol after treatment with RNase A and Proteinase K. For ChIP-Seq,

barcoded libraries of Drosha-GFP, HP1BP3-GFP and H1.5-GFP ChIP and input DNA were generated with the TruSeq® ChIP Sample Preparation Kit (Illumina), and 50-nt single-end reads were generated with the HiSeq2000 system (Illumina). Sequence reads were aligned to the human reference genome (hg19) using Bowtie (v.0.12.7) (Langmead et al., 2009). Uniquely mapped reads with ≤2 mismatches to the reference sequence were retained for further analysis; for Drosha-GFP we obtained 28,343,999 reads from 28,580,957 input reads, for HP1BP3-GFP-GFP we obtained 26,606,724 reads from 30,200,875 input reads, for H1.5-GFP-GFP we obtained 21,719,813 reads from 31,055,062 input reads. We used SICER algorithm (v1.1) (Zhang et al., 2008) to define genome wide occupied islands for these proteins from ChIP-Seq data. We used a window size of 200bp, a fragment size of 200bp, and a gap size of 600bp parameters to quantify the Drosha, HP1BP3 or H1.5 ChIP signal fold enrichment over input sample in each island. A false discovery rate (FDR) cutoff of 1% was used to select islands for further analysis. For annotation of Drosha or HP1BP3 or H1.5 islands, we used annotate Peaks function in Homer tools (Heinz et al., 2010) to assign Drosha islands relative to their specific positions in the genome. This function takes the Drosha islands coordinates and tag directories as input, and extends each tag by their estimated ChIP-fragment length. We normalized each library to 10 million reads and calculate ChIP fragment coverage represented in per base pair per peak. We used 5000 bp regions flanking the islands center to generate density plots showing Drosha, HP1BP3 and H1.5 ChIP signal across all Drosha islands.

Chromatin immunoprecipitation (ChIP) and RNA immunoprecipitation (RIP)

ChIP and RIP were performed as described previously (Sakurai et al., 2010). Cells were crosslinked with 1% formaldehyde and then nuclear fractions were isolated. The nuclear lysates were incubated with anti-Drosha antibody (2 μg; Abcam) or rabbit IgG (2 μg; Millipore) overnight and antibody-protein-DNA/RNA complexes were recovered using protein G plus/protein A agarose beads (Millipore). After reverse crosslinking and proteinase K treatment, immunoprecipitated DNA/RNA were purified by phenol-chloroform extraction and ethanol precipitation. For RIP, the samples were treated with DNase I to remove genomic DNA and then reverse transcribed to generate cDNA. Real time RT-PCR was performed using Taqman pri-miRNA kit in RIP or designed primers for miRNA gene locus (listed in table S6).

Chromatin and nucleoplasm fractionation

Chromatin fractionation was performed as previously described (Gagnon et al., 2014). HeLa cells were lysed in hypotonic buffer (HLB) (10mM Tris pH 7.5, 10mM NaCl, 3mM $MgCl_2$, 0.3% (vol/vol) NP-40, 1% protease inhibitor cocktail). Cell nuclei were isolated by centrifugation at 1000g for 4min and washed three times using 1mL of HLB. The nuclei were suspended in Modified Wuarin-Schibler buffer (MWS) (10mM Tris pH 7.0, 4mM EDTA, 0.3M NaCl, 1M urea, and 1% (vol/vol) NP-40) and spun at 1000g for 4min at 4 ºC. While the supernatant was collected as the nucleoplasm fraction, the pellet was washed three times with 1mL MWS buffer and then collected as the chromatin fraction. RNA from nucleoplasm was precipitated by ethanol and extracted by Trizol. RNA from chromatin fraction was directly isolated by Trizol extraction. Taqman qPCR was performed to measure the relative abundance of chromatin bound and nucleoplasmic pri-miRNAs.

APPENDIX A
List of primers used for BAC transgenic cell line

Drosha	5’- agagcatcaagagagagagccagatgagactgaagacatcaagaaagattatgatattccaactactgca-3’ 5’- cagtcacagttactgagcaagtaaatactccacacttgcatgccctccttcagaagaactcgtcaagaag-3’
HP1BP3	5-’gggcaagggcaaatccaccatgaagaagtctttcagagtgaaaaaggattatgatattccaactactgca-3’ 5’-gaaaataagattttgaatttcatcatgataccctttttttcctataaaattcagaagaactcgtcaagaag-3’
H1.5	5’- aaaacctaaagctgcaaaggccaagaaggcggctgccaaaaagaaggattatgatattccaactactgca-3’ 5’- tctgaaaagagcctttggggctttgttgcggttttcacacgccagcttctcagaagaactcgtcaagaag-3’

APPENDIX B
siRNAs for Histone isoform knockdown

Primer Name	Sequence 5' - 3'
HP1BP3-F	TTTCCGAAGAAAGAGCCAGA
HP1BP3-R	GGACCCTCTCTGCTTCACAG
H1.5-F	GGCTGTGGCTGCTTCTAA
H1.5-R	CAGCTTAATGCGGCTGTTATTC
H1.2-F	AAAGAGCGTAGCGGAGTTTC
H1.2-R	CAGGCTCTTGAGACCAAGTTT
H1F0-F	TGTCCTCAAGCAGACCAAAG
H1F0-R	TGAAGGCCACTGACTTCTTG
H1X-F	ACCTACCTCAAGTACTCGATCA
H1X-R	TCTTGCGGTTGAGCTTGAA

APPENDIX C
Primers for ChIP targeting miRNA transcription loci

Human Hela and U2OS cells

miRNA	Forward Primer	Reverse Primer
let-7a	TGGGATGAGGTAGTAGGTTGTA	CTATCACGTTAGGAAAGACAGTAGAT
miR-7	CAGCATGGGACATGGGTATAG	GATAAAGTACAGGACAGCCAGTTA
miR-16	CCTTACTTCAGCAGCACAGTTA	TGCCTTAGCAGCACGTAAATA
miR-17	GTGCTTACAGTGCAGGTAG	GAGGCAGCTGTCACCATAAT
miR-21	CACCTTGTCGGGTAGCTTATC	GGATATGGATGGTCAGATGAAAGA
miR-33	CATTTGCTCCAGCGGTTTG	CAAACCGCTGGAGCAAATG
miR-1	CTCTTCTAGACCTTGCCCTAAC	CCAGACATAGCACTACCACAA
miR-9	CAAGGGTGACCATCTGCTAAT	CCTCTGAAAGAAGGTGATCCAG
miR-31	CCTTACTTCCTGTCCTGCTATG	CCTAGGTGTGTCCAAGGAATAG
let-7e	TCCCTGTCTGTCTGTCTGT	CTAGGAGGCCGTATAGTGATCT
miR-96	TGCCGTGGGTTGGAAAG	TCACAACCTGCAGCAGAAA
miR-99a	GATCCGATCTTGTGGTGAAGTG	TGTTGAACGGCACTGTGTATAG
miR-30b	CTGGGAGGTGGATGTTTACTT	AGCCTCTGTATACTATTCTTGCC
miR-181	AACGCTGTCGGTGAGTTT	GAATTCTGAGCACCATGGAGTA
let-7i	CTGGCTGAGGTAGTAGTTTGTG	CAGCACTAGCAAGGCAGTAG
miR-143	CAGTTGGGAGTCTGAGATGAAG	CACTCTGTCCTTCCTGTTATGG
miR-222	GCCAGTGTAGATCCTGTCTTTC	CTGCCCAATAATCTCTCTCAGG
miR-25	ATTGCACTTGTCTCGGTCTG	GACACCCTTGTTCTGGCTTTA
miR-193a	CGAACTCCGAGGATGGGA	CCGAGAACTGGGACTTTGTAG
miR-23a	ACATTGCCAGGGATTTCCA	CCTGCTCACAAGCAGCTAA
miR-22	ATGCCCTGCTCAGATCTTTC	CTTTAGCTGGGTCAGGACATAA

Drosophila S2 cells

	Sequence (5'-3')
pri-miR-1 F	GTTAGCCGCGTTGTGGAAAATC
Pri-miR-1 R	CATTTCATTACGGTTCTACTTCTG
pri-miR-8 F	AGAACTTTGAGCTTCCTCTGGC
pri-miR-8 R	TTTGGTGCTGCTGCTGCTGTTG
pri-miR-276a F	AAAAGGGAAACGCGCTGCCAAG
pri-miR-276a R	CGTTTGTCCAGCGTTTTCTCATC
pri-miR-305 F	GAAATGCTCGCAGGCGAGTCC

pri-miR-305 R	GTTGAACACTTGTATCGGTCGC
pri-miR-317 F	ACGGTTTGTGTCTCTGCTGAGC
pri-miR-317 R	CTGTGGGGCATTCTCGTTATCC
pri-bantam F	CGCTCAGATGCAGATGTTGTTG
pri-bantam R	TCGACCATCGGAATGTGGAATG
RP49 F	CCAAGCACTTCATCCGCCACCA
RP49 R	GCGGGTGCGCTTGTTCGATCC
GAPDH F	GGCAACTTCTGCGAAACG A
GAPDH R	CGCAGCACCTTGCCATACTT

APPENDIX D
RNA probes used in Northern Blot

Human miRNA

miRNA	Probe Sequence
Anti-let-7a	5′-AACUAUACAACCUACUACCUCA-3′
Anti-miR-16	5′-CGCCAAUAUUUACGUGCUGCUA-3′
Anti-miR-21	5′-UCAACAUCAGUCUGAUAAGCUA-3′
Anti-miR-23a	5’ AAAUCCCAUCCCCAGGAACCCC-3’
Anti-5S	5′-CCGACCCTGCTTAGCTTCCGAGATCA-3′

Drosophila miRNA

5.0S (DNA)	CAACACGCGCTGTTCCCAAGCGGT
Bantam (RNA)	AGUCAAACCAAAUCGAAAACCGG
miR-1 (RNA)	UAUUGAAUGCAAGGAAGCAUGG
miR-8	UCUAAUGCUGCCCGGUAAGAUG
miR-305	CAGAGCACCUGAUGAAGUACAAU

APPENDIX E
PCR primers used for pri-miRNA detection in S2 cells

	Sequence (5'-3')
pri-miR-1 F	GTTAGCCGCGTTGTGGAAAATC
Pri-miR-1 R	CATTTCATTACGGTTCTACTTCTG
pri-miR-8 F	AGAACTTTGAGCTTCCTCTGGC
pri-miR-8 R	TTTGGTGCTGCTGCTGCTGTTG
pri-miR-276a F	AAAAGGGAAACGCGCTGCCAAG
pri-miR-276a R	CGTTTGTCCAGCGTTTTCTCATC
pri-miR-305 F	GAAATGCTCGCAGGCGAGTCC
pri-miR-305 R	GTTGAACACTTGTATCGGTCGC
pri-miR-317 F	ACGGTTTGTGTCTCTGCTGAGC
pri-miR-317 R	CTGTGGGGCATTCTCGTTATCC
pri-bantam F	CGCTCAGATGCAGATGTTGTTG
pri-bantam R	TCGACCATCGGAATGTGGAATG
RP49 F	CCAAGCACTTCATCCGCCACCA
RP49 R	GCGGGTGCGCTTGTTCGATCC
GAPDH F	GGCAACTTCTGCGAAACG A
GAPDH R	CGCAGCACCTTGCCATACTT

BIBLOGRAPHY

Adachi, J., and Hasegawa, M. (1992). Amino acid substitution of proteins coded for in mitochondrial DNA during mammalian evolution. Idengaku zasshi *67*, 187-197.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research *25*, 3389-3402.

Ambros, V. (2004). The functions of animal microRNAs. Nature *431*, 350-355.

Ballarino, M., Pagano, F., Girardi, E., Morlando, M., Cacchiarelli, D., Marchioni, M., Proudfoot, N.J., and Bozzoni, I. (2009). Coupled RNA processing and transcription of intergenic primary microRNAs. Mol Cell Biol *29*, 5632-5638.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nat Rev Genet *15*, 163-175.

Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. Nature genetics *35*, 215-217.

Brugiolo, M., Herzel, L., and Neugebauer, K.M. (2013). Counting on co-transcriptional splicing. F1000prime reports *5*, 9.

Bustin, M., Catez, F., and Lim, J.H. (2005). The dynamics of histone H1 function in chromatin. Molecular cell *17*, 617-620.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740-744.

Cheng, T.L., Wang, Z., Liao, Q., Zhu, Y., Zhou, W.H., Xu, W., and Qiu, Z. (2014). MeCP2 suppresses nuclear microRNA processing and dendritic growth by regulating the DGCR8/Drosha complex. Developmental cell *28*, 547-560.

Davis, B.N., Hilyard, A.C., Lagna, G., and Hata, A. (2008). SMAD proteins control DROSHA-mediated microRNA maturation. Nature *454*, 56-61.

Davis, B.N., Hilyard, A.C., Nguyen, P.H., Lagna, G., and Hata, A. (2010). Smad proteins bind a conserved RNA sequence to promote microRNA maturation by Drosha. Molecular cell *39*, 373-384.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.

Di Carlo, V., Grossi, E., Laneve, P., Morlando, M., Dini Modigliani, S., Ballarino, M., Bozzoni, I., and Caffarelli, E. (2013). TDP-43 regulates the microprocessor complex activity during in vitro neuronal differentiation. Molecular neurobiology *48*, 952-963.

Djuranovic, S., Nahvi, A., and Green, R. (2012). miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. Science *336*, 237-240.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Forstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., and Zamore, P.D. (2005). Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS Biol *3*, e236.

Fromm, B., Billipp, T., Peck, L.E., Johansen, M., Tarver, J.E., King, B.L., Newcomb, J.M., Sempere, L.F., Flatmark, K., Hovig, E., *et al.* (2015). A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome. Annual review of genetics *49*, 213-242.

Gagnon, K.T., Li, L., Chu, Y., Janowski, B.A., and Corey, D.R. (2014). RNAi factors are present and active in human cell nuclei. Cell reports *6*, 211-221.

Ganesan, G., and Rao, S.M. (2008). A novel noncoding RNA processed by Drosha is restricted to nucleus in mouse. RNA *14*, 1399-1410.

Garfinkel, B.P., Arad, S., Le, P.T., Bustin, M., Rosen, C.J., Gabet, Y., and Orly, J. (2015a). Proportionate Dwarfism in Mice Lacking Heterochromatin Protein 1 Binding Protein 3 (HP1BP3) Is Associated With Alterations in the Endocrine IGF-1 Pathway. Endocrinology *156*, 4558-4570.

Garfinkel, B.P., Melamed-Book, N., Anuka, E., Bustin, M., and Orly, J. (2015b). HP1BP3 is a novel histone H1 related protein with essential roles in viability and growth. Nucleic acids research *43*, 2074-2090.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.

Gromak, N., Dienstbier, M., Macias, S., Plass, M., Eyras, E., Caceres, J.F., and Proudfoot, N.J. (2013). Drosha regulates gene expression independently of RNA cleavage function. Cell reports *5*, 1499-1510.

Guil, S., and Caceres, J.F. (2007). The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. Nature structural & molecular biology *14*, 591-596.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*, 835-840.

Haase, A.D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A., and Filipowicz, W. (2005). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO reports *6*, 961-967.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nature structural & molecular biology *16*, 1021-1025.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science *293*, 1146-1150.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes & development *18*, 3016-3027.

Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell *125*, 887-901.

Han, J., Pedersen, J.S., Kwon, S.C., Belair, C.D., Kim, Y.K., Yeom, K.H., Yang, W.Y., Haussler, D., Blelloch, R., and Kim, V.N. (2009). Posttranscriptional crossregulation between Drosha and DGCR8. Cell *136*, 75-84.

Hayashihara, K., Uchiyama, S., Shimamoto, S., Kobayashi, S., Tomschik, M., Wakamatsu, H., No, D., Sugahara, H., Hori, N., Noda, M., *et al.* (2010). The middle region of an HP1-binding protein, HP1-BP74, associates with linker DNA at the entry/exit site of nucleosomal DNA. The Journal of biological chemistry *285*, 6498-6507.

He, L., and Hannon, G.J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet *5*, 522-531.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138*, 696-708.

Hergeth, S.P., and Schneider, R. (2015). The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. EMBO reports *16*, 1439-1453.

Hildebrand, A., Remmert, M., Biegert, A., and Soding, J. (2009). Fast and accurate automatic structure prediction with HHpred. Proteins *77 Suppl 9*, 128-132.

Hiragami, K., and Festenstein, R. (2005). Heterochromatin protein 1: a pervasive controlling influence. Cellular and molecular life sciences : CMLS *62*, 2711-2726.

Hutvagner, G., Simard, M.J., Mello, C.C., and Zamore, P.D. (2004). Sequence-specific inhibition of small RNA function. PLoS Biol *2*, E98.

Hutvagner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science *297*, 2056-2060.

Iwakawa, H.O., and Tomari, Y. (2015). The Functions of MicroRNAs: mRNA Decay and Translational Repression. Trends in cell biology *25*, 651-665.

Jiang, F., Ye, X., Liu, X., Fincher, L., McKearin, D., and Liu, Q. (2005). Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila. Genes Dev *19*, 1674-1679.

Kadener, S., Rodriguez, J., Abruzzi, K.C., Khodor, Y.L., Sugino, K., Marr, M.T., 2nd, Nelson, S., and Rosbash, M. (2009). Genome-wide identification of targets of the drosha-pasha/DGCR8 complex. Rna *15*, 537-545.

Katoh, K., and Standley, D.M. (2014). MAFFT: iterative refinement and additional methods. Methods in molecular biology *1079*, 131-146.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol *10*, 126-139.

Kim, Y.K., and Kim, V.N. (2007). Processing of intronic microRNAs. EMBO J *26*, 775-783.

Kittler, R., Pelletier, L., Ma, C., Poser, I., Fischer, S., Hyman, A.A., and Buchholz, F. (2005). RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells. Proceedings of the National Academy of Sciences of the United States of America *102*, 2396-2401.

Kittler, R., Zhou, J., Hua, S., Ma, L., Liu, Y., Pendleton, E., Cheng, C., Gerstein, M., and White, K.P. (2013). A comprehensive nuclear receptor network for breast cancer cells. Cell reports *3*, 538-551.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science *293*, 2269-2271.

Knuckles, P., Vogt, M.A., Lugert, S., Milo, M., Chong, M.M., Hautbergue, G.M., Wilson, S.A., Littman, D.R., and Taylor, V. (2012). Drosha regulates neurogenesis by controlling neurogenin 2 expression independent of microRNAs. Nature neuroscience *15*, 962-969.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. Curr Biol *14*, 2162-2167.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology *10*, R25.

Le Douarin, B., Nielsen, A.L., Garnier, J.M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. (1996). A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. The EMBO journal *15*, 6701-6715.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., *et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.

Lee, Y., Hur, I., Park, S.Y., Kim, Y.K., Suh, M.R., and Kim, V.N. (2006). The role of PACT in the RNA silencing pathway. The EMBO journal *25*, 522-532.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004a). MicroRNA genes are transcribed by RNA polymerase II. The EMBO journal *23*, 4051-4060.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004b). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69-81.

Liang, C., Wang, Y., Murota, Y., Liu, X., Smith, D., Siomi, M.C., and Liu, Q. (2015). TAF11 Assembles the RISC Loading Complex to Enhance RNAi Efficiency. Molecular cell *59*, 807-818.

Liang, C., Xiong, K., Szulwach, K.E., Zhang, Y., Wang, Z., Peng, J., Fu, M., Jin, P., Suzuki, H.I., and Liu, Q. (2013). Sjogren syndrome antigen B (SSB)/La promotes global

microRNA expression by binding microRNA precursors through stem-loop recognition. The Journal of biological chemistry *288*, 723-736.

Liao, Q., Shen, J., Liu, J., Sun, X., Zhao, G., Chang, Y., Xu, L., Li, X., Zhao, Y., Zheng, H., *et al.* (2014). Genome-wide identification and functional annotation of Plasmodium falciparum long noncoding RNAs from RNA-seq data. Parasitology research *113*, 1269-1281.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.

Liu, Q., and Paroo, Z. (2010). Biochemical principles of small RNA pathways. Annu Rev Biochem *79*, 295-319.

Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.E., Smith, D.P., and Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science *301*, 1921-1925.

Liu, X., Jiang, F., Kalidas, S., Smith, D., and Liu, Q. (2006). Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. Rna *12*, 1514-1520.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95-98.

Ma, L., Reinhardt, F., Pan, E., Soutschek, J., Bhat, B., Marcusson, E.G., Teruya-Feldstein, J., Bell, G.W., and Weinberg, R.A. (2010). Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model. Nature biotechnology *28*, 341-347.

Magrane, M., and Consortium, U. (2011). UniProt Knowledgebase: a hub of integrated protein data. Database : the journal of biological databases and curation *2011*, bar009.

Maniataki, E., and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. Genes & development *19*, 2979-2990.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., *et al.* (2015). CDD: NCBI's conserved domain database. Nucleic acids research *43*, D222-226.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 *17*.

Millan-Arino, L., Islam, A.B., Izquierdo-Bouldstridge, A., Mayor, R., Terme, J.M., Luque, N., Sancho, M., Lopez-Bigas, N., and Jordan, A. (2014). Mapping of six somatic linker histone H1 variants in human breast cancer cells uncovers specific features of H1.2. Nucleic acids research *42*, 4474-4493.

Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I., and Proudfoot, N.J. (2008). Primary microRNA transcripts are processed co-transcriptionally. Nature structural & molecular biology *15*, 902-909.

Morlando, M., Dini Modigliani, S., Torrelli, G., Rosa, A., Di Carlo, V., Caffarelli, E., and Bozzoni, I. (2012). FUS stimulates microRNA biogenesis by facilitating co-transcriptional Drosha recruitment. The EMBO journal *31*, 4502-4510.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. The Plant cell *2*, 279-289.

Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. Rna *14*, 1539-1549.

Nojima, T., Gomes, T., Grosso, A.R., Kimura, H., Dye, M.J., Dhir, S., Carmo-Fonseca, M., and Proudfoot, N.J. (2015). Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. Cell *161*, 526-540.

Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr Biol *12*, 1484-1495.

Paroo, Z., Ye, X., Chen, S., and Liu, Q. (2009). Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling. Cell *139*, 112-122.

Pawlicki, J.M., and Steitz, J.A. (2008). Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. The Journal of cell biology *182*, 61-76.

Pawlicki, J.M., and Steitz, J.A. (2009). Subnuclear compartmentalization of transiently expressed polyadenylated pri-microRNAs: processing at transcription sites or accumulation in SC35 foci. Cell Cycle *8*, 345-356.

Pham, J.W., Pellino, J.L., Lee, Y.S., Carthew, R.W., and Sontheimer, E.J. (2004). A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. Cell *117*, 83-94.

Poser, I., Sarov, M., Hutchins, J.R., Heriche, J.K., Toyoda, Y., Pozniakovsky, A., Weigl, D., Nitzsche, A., Hegemann, B., Bird, A.W., *et al.* (2008). BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. Nature methods *5*, 409-415.

Rasmussen, E.B., and Lis, J.T. (1993). In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. Proceedings of the National Academy of Sciences of the United States of America *90*, 7923-7927.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research *43*, e47.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Roque, A., Iloro, I., Ponte, I., Arrondo, J.L., and Suau, P. (2005). DNA-induced secondary structure of the carboxyl-terminal domain of histone H1. The Journal of biological chemistry *280*, 32141-32147.

Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005). Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. PLoS Biol *3*, e235.

Sakurai, T., Bai, H., Konno, T., Ideta, A., Aoyagi, Y., Godkin, J.D., and Imakawa, K. (2010). Function of a transcription factor CDX2 beyond its trophectoderm lineage specification. Endocrinology *151*, 5873-5881.

Schlissel, M.S., and Brown, D.D. (1984). The transcriptional regulation of Xenopus 5s RNA genes in chromatin: the roles of active stable transcription complexes and histone H1. Cell *37*, 903-913.

Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science *305*, 1434-1437.

Suzuki, H.I., Yamagata, K., Sugimoto, K., Iwamoto, T., Kato, S., and Miyazono, K. (2009). Modulation of microRNA processing by p53. Nature *460*, 529-533.

Thoma, F., Koller, T., and Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. The Journal of cell biology *83*, 403-427.

Thum, T., Catalucci, D., and Bauersachs, J. (2008). MicroRNAs: novel regulators in cardiac development and disease. Cardiovascular research *79*, 562-570.

Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A.D., Filipowicz, W., Ramos, A., Gherzi, R., and Rosenfeld, M.G. (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. Nature *459*, 1010-1014.

Tuschl, T., Zamore, P.D., Lehmann, R., Bartel, D.P., and Sharp, P.A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. Genes & development *13*, 3191-3197.

Van Wynsberghe, P.M., Kai, Z.S., Massirer, K.B., Burton, V.H., Yeo, G.W., and Pasquinelli, A.E. (2011). LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans. Nature structural & molecular biology *18*, 302-308.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320*, 97-100.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522-525.

Whitelaw, E., and Proudfoot, N. (1986). Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene. The EMBO journal *5*, 2915-2922.

Wienholds, E., Koudijs, M.J., van Eeden, F.J., Cuppen, E., and Plasterk, R.H. (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nature genetics *35*, 217-218.

Wu, H., Xu, H., Miraglia, L.J., and Crooke, S.T. (2000). Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. J Biol Chem *275*, 36957-36965.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev *17*, 3011-3016.

Yin, S., Yu, Y., and Reed, R. (2015). Primary microRNA processing is functionally coupled to RNAP II transcription in vitro. Scientific reports *5*, 11992.

Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q., and Tomari, Y. (2010). ATP-dependent human RISC assembly pathways. Nature structural & molecular biology *17*, 17-23.

Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.

Zaret, K.S., and Sherman, F. (1982). DNA sequence required for efficient transcription termination in yeast. Cell *28*, 563-573.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology *9*, R137.