

Minireview

# MicroRNA Target Recognition: Insights from Transcriptome-Wide Non-Canonical Interactions

Heeyoung Seok<sup>1</sup>, Juyoung Ham<sup>2</sup>, Eun-Sook Jang<sup>3</sup>, and Sung Wook Chi<sup>1\*</sup>

**MicroRNAs (miRNAs) are small non-coding RNAs (~22 nucleotides) regulating gene expression at the post-transcriptional level. By directing the RNA-induced silencing complex (RISC) to bind specific target mRNAs, miRNA can repress target genes and affect various biological phenotypes. Functional miRNA target recognition is known to majorly attribute specificity to consecutive pairing with seed region (position 2-8) of miRNA. Recent advances in a transcriptome-wide method of mapping miRNA binding sites (Ago HITS-CLIP) elucidated that a large portion of miRNA-target interactions *in vivo* are mediated not only through the canonical “seed sites” but also via non-canonical sites (~15-80%), setting the stage to expand and determine their properties. Here we focus on recent findings from transcriptome-wide non-canonical miRNA-target interactions, specifically regarding “nucleation bulges” and “seed-like motifs”. We also discuss insights from Ago HITS-CLIP data alongside structural and biochemical studies, which highlight putative mechanisms of miRNA target recognition, and the biological significance of these non-canonical sites mediating marginal repression.**

## INTRODUCTION

microRNAs (miRNAs) are single stranded non-coding RNA molecules of ~22 nucleotides (nt) that regulate gene expression via post-transcriptional and/or translational repression (Ambros, 2004). Primary miRNA (pri-miRNAs) are transcribed in the nucleus by RNA polymerase II or III, where ~70 nt stem-loop miRNA precursors (pre-miRNAs) are subsequently excised by the microprocessor complex containing the RNase III enzyme Drosha, and exported to the cytoplasm via Exportin-5 (Kim et al., 2009). Dicer, another RNase III enzyme, further processes

pre-miRNAs to produce mature miRNAs, the final product being a ~22 base-pair duplex with 2 nt-long 3' overhangs (He and Hannon, 2004). Then, one strand of the mature miRNA is loaded onto Argonaute (Ago, also known as Eif2c), a core protein of the RNA-induced silencing complex (RISC). miRNA forms base pairs with a target mRNA as a guide for Ago binding and to direct the specificity of the RISC effector, decreasing target mRNA levels and/or its translation (Fabian et al., 2010), where mRNA destabilization is the dominant mechanism (Eichhorn et al., 2014; Guo et al., 2010).

miRNAs are abundant in the mammalian genome (more than 2000 human miRNAs are currently reported in miRBase) (Kozomara and Griffiths-Jones, 2014) and their regulatory role is essential, affecting various biological phenomena (Kim, 2005; Sim et al., 2014). Supporting evidence derives from the fact that a lethal phenotype during early development was observed in Dicer1-null (Bernstein et al., 2003) or Ago2-null (Liu et al., 2004) mice, and various biological defects were also reported after losses of individual miRNAs (Park et al., 2010). In addition, alterations of miRNA regulation are related to many diseases such as neurological disorders (Hebert and De Strooper, 2009), various types of cancer (Croce, 2009), and cardiovascular diseases (Olson, 2014). Importantly, all of these defects were ultimately caused by a dysregulation in target gene expression. Therefore, identification of miRNA targets is the key for understanding miRNA function. However, the limitation here is our ability to delineate a general principle for identification of specific RNA targets upon which miRNAs act. The problem stems from the observation that most of miRNA target sites have partial complementarity (Ambros, 2004).

## CANONICAL TARGET SITES: SEED PAIRING RULES

In contrast to plants, a near-perfect base pairing of miRNA to its target is rare in animals, making it a challenge to predict the target sites (Bartel, 2009). However, initial prediction attempts provided evidence that local short stretches ( $\geq 6$  nt) of consecutive base-pairing significantly contribute to target recognition (John et al., 2004; Krek et al., 2005; Lewis et al., 2003; Stark et al., 2003). Conceptually termed as the “nucleus,” the short consecutive matches could initiate a miRNA-target duplex, followed by the propagation of partial annealing that may further stabilize miRNA-target hybridization (Filipowicz, 2005; Rajewsky, 2006). Intriguingly, nuclei were further found to be typically located in the 5' end region of miRNAs called the “seed”, enabling the prediction of miRNA target sites (Lewis et al., 2003).

Functional miRNA-target interactions are known to majorly

<sup>1</sup>Division of Life Sciences, College of Life Sciences and Biotechnology, Korea University, Seoul 02841, Korea, <sup>2</sup>Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences and Technology, Sungkyunkwan University, Seoul 06351, Korea, <sup>3</sup>EncodeGEN Co. Ltd., Seoul 06329, Korea

\*Correspondence: chi13@korea.ac.kr

Received 25 January, 2016; revised 4 April, 2016; accepted 4 April, 2016; published online 27 April, 2016

**Keywords:** argonaute, CLIP, microRNA, non-canonical targets

	Types	Base pairing	mRNA : miRNA	references
<b>A</b> Widespread canonical types	7-mer seed (2-8)	5' - AACUCACAACCA <b>CUCAGGGA</b> -3' 3' AGUGUGAACUCCA <b>GAGUCCCU</b> 5'	<i>lin-14</i> 3' UTR <i>lin-4</i> ( <i>C.elegans</i> )	Lee et al, 1993 Wightman et al, 1993
	6-mer seed (2-7)	5' - UAAGUUUCGUGUUGCAA <b>GAAACAA</b> -3' 3' AGUGCGCUCGGCUUG <b>CUUGUUU</b> 5'	<i>Mtpn</i> 3' UTR miR-375 (Mouse)	Poy et al, 2004
	offset 6-mer seed (3-8)	5' --UCA--AAA-- <b>CUCAGGA</b> --3' 3' AGUGUGAAACUCCA <b>GAGUCCCU</b> 5'	<i>lin-14</i> 3' UTR <i>lin-4</i> ( <i>C.elegans</i> )	Lee et al, 1993 Wightman et al, 1993
		near perfect site (3' compensatory)	5' - CCACAACAUGAA <b>ACUGCCU</b> A -3' 3' GGGUUGUUGUACUU <b>UGAUGGAU</b> 5'	<i>Hoxb8</i> 3' UTR miR-196 (Human)
<b>B</b> A few observed non-canonical types	3' compensatory site (bulge in seed)	5' - UUUUAUACAACCGUUC <b>UAC<sup>A</sup>CUCA</b> -3' 3' UUGAUAUGUUGG - AUG <b>AUG GAG</b> U -5'	<i>lin-41</i> 3' UTR <i>let-7</i> ( <i>C.elegans</i> )	Vella et al, 2004
	wobble in seed	5' - GUCUGAUUCAG - AA <b>GGGCUCA</b> -3' 3' UGUCCUAACUCCCC <b>CCC</b> GGGA 5'	<i>Nanog</i> CDS miR-296 (Mouse)	Tay et al, 2008
	seedless elements (mismatch in seed)	5' - GUGGGUGUCU - CUGGG <b>CUGAACCA</b> -3' 3' GACA - AGGACGACUU <b>GACUCGGU</b> 5'	<i>E2f2</i> 3' UTR miR-24 (Human)	Lal et al, 2009
		centered site	5' - AGUUUU <b>UCAGUCUGAUAA</b> CUAU -3' 3' AGUUGU <b>AGUCAGACUAUU</b> CGAU 5'	<i>Gstm</i> 3' UTR miR-21 (Human)
	<b>C</b> Widespread non-canonical types	nucleation bulge site	5' - CUCCUCAUUGUA <b>GUG<sup>G</sup>CCUUA</b> -3' 3' CCGUAAGUGGCG <b>CAC GGAAU</b> 5'	<i>Mink1</i> 3' UTR miR-124 (Human)
seed-like motif		5' - UGGGGAUAGUGUUAAL <b>CGUAAU</b> U -3' 3' ACCCCUCAAU - -GCGU <b>GCUUUA</b> A 5'	<i>Gimap3</i> 3' UTR miR-155 (Mouse)	Loeb et al, 2012

**Fig. 1.** Canonical and non-canonical miRNA target sites. Representative examples of each type are indicated as widespread canonical (A), a few observed non-canonical (B), and widespread non-canonical types (C). Key binding regions are highlighted in red and subtly contributing regions are in purple. Solid lines indicate Watson-Crick base pairing and dots indicate G:U wobble pairs.

require as few as 6-nt matches within the seed region (position 2-8, Fig. 1A) (Bartel, 2009). There are possible 6-mers (positions 1-6, 2-7, and 3-8), 7-mers (positions 2-8 and 1-7), and 8-mer (position 1-8) matches in the seed. Otherwise, a 6-mer match to position 3-8 is called an “offset 6-mer seed” because of its position and a marginal effect on repression (Friedman et al., 2009). Such canonical seed sites were initially known by early biological studies (Lee et al., 1993; Poy et al., 2004; Wightman et al., 1993), which were further validated by microarray experiments that detected enrichment of seed matches in miRNA-dependent transcripts showing repression (Grimson, 2007; Lim, 2005), and also by bioinformatics analyses, which found widespread conservation of seed sites in 3' untranslated regions (3' UTRs) in multi-genome sequences (Lewis et al., 2005; Xie et al., 2005). Seed-pairing rules have been informative in prediction and analysis of canonical seed sites, often in combination with evolutionary conservation (Friedman et al., 2009; Lewis et al., 2005), secondary structure (Long et al., 2007), or neighboring context information (Grimson, 2007). However, since a 6 nt match presents on average every ~4,000 nt, likely to be occurred often by chance, such strategies still suffer from both false-positive (~40-66%) and false-negative predictions (~50-70%) (Mourelatos, 2008) even in the usage of microarray or proteomic approaches (Baek et al., 2008; Selbach et al., 2008). Furthermore, seed-pairing rules cannot identify non-canonical target sites, which have been reported as functional (Brodersen and Voinnet, 2009).

## EVIDENCE OF NON-CANONICAL TARGET SITES

Since seed-pairing rules are widely adopted, there has been an unintentional bias to study only the canonical seed matches, overlooking the non-canonical targets. Nevertheless, several biological studies have functionally validated that perfectly matched miRNA seeds are neither necessary nor sufficient for all functional miRNA-target interactions (Brodersen and Voinnet, 2009). For example, supplementary components in near-perfect sites compensate for imperfect seed matches and are functional for target cleavage (miR-196 for *Hoxb8*, Fig. 1B) (Yekta et al., 2004). In *C. elegans*, *let-7* functionally recognizes the bulge (Fig. 1B) as well as the wobble (G:U pairing) in seed sites of *lin-41* with 3' compensatory pairings (Vella et al., 2004). *lsey-6* can also tolerate wobble pairing in the seed to downregulate its target, *cog-1* (Didiano and Hobert, 2006). In mammals, *Nanog*, *Oct4* and *Sox2* genes, well-known regulators generating induced pluripotent stem (iPS) cells, have been shown to contain functional wobble (miR-296 for *Nanog*, Fig. 1B) and bulge pairing sites for cognate miRNA seeds (miR-134, miR-296, and miR-470), harboring a few cases in their coding sequences (CDS) (Tay et al., 2008). In contrast to a fruit fly genetic study where 3' compensatory sites were shown to be as functional as the 5' dominant canonical seed sites (Brennecke et al., 2005), such non-canonical miRNA target sites were revealed to be rare in mammals (less than ~5%) (Friedman et al., 2009) and their effects were modest (Grimson, 2007; Wee et al.,

2012).

When seed-pairing rules were applied, putative miRNA targets from microarrays that showed miRNA-dependent repression were often demonstrated to have high false-negatives, implicating prevalent usage of non-canonical target sites. In lieu of this, microarray analysis of miR-24-transfected K562 cells found that several miR-24 targets are repressed through non-canonical sites, named “seedless” recognition elements (Fig. 1B) (Lal et al., 2009). In addition, “centered sites”, comprising 11-12 consecutive base-pairing to the center of miRNA, were also identified by the analysis of microarray data where neither perfect seed nor 3' compensatory pairing was observed (Fig. 1B) (Shin et al., 2010). However, the limitation of such studies is that, lacking information on precise binding sites, they are unable to distinguish between direct and indirect miRNA targets.

### TRANSCRIPTOME-WIDE MIRNA BINDING SITES IN VIVO

Uncertainty in direct miRNA target sites necessitates the development of experimental methods capable of recovering miRNAs physically associated with their targets (Easow et al., 2007). Initially, biochemical isolation of miRNA-mRNA complexes via Ago protein-specific immunoprecipitation was attempted in order to purify mRNAs bound by Ago-miRNA (Easow et al., 2007; Hammell, 2008; Hendrickson et al., 2008; Karginov, 2007). However, the integrity of the approach was questioned because of the possible high background caused by nonspecific RNA-protein interactions, especially mediated by *in vitro* rearrangements in the Ago-RNA complex (Mili and Steitz, 2004; Riley et al., 2012b). However, the cross-linking and immunoprecipitation (CLIP) method that uses ultraviolet (UV) irradiation to crosslink RNA-protein complexes in direct contact within approximately single Angstrom distances in living cells, can secure specific RNA-protein interactions by allowing stringent purification (Ule, 2003). Combined with high-throughput sequencing (HITS-CLIP) (Licatalosi, 2008), it was successfully applied to Ago (Ago HITS-CLIP) to produce transcriptome-wide information of miRNA binding sites (Chi et al., 2009). The novelty lies not only in providing direct miRNA target sites, but also in accomplishing high target specificity (~93%), low false-positives (~13-27%), and low false-negatives (~15-25%) (Chi et al., 2009). Furthermore, by analyzing crosslinking-induced mutation sites (CIMS), footprints of Ago-miRNA binding regions (~45-62 nt) became available at single nucleotide resolution (Zhang and Darnell, 2011). A modification of this method, PAR-CLIP (photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation), used a nucleoside analog that improved the efficiency of UV crosslinking and subsequent mutations (Hafner et al., 2010), and iCLIP (individual-nucleotide resolution UV crosslinking and immunoprecipitation) analyzed truncated cDNAs at the position of the protein-RNA crosslink sites (Konig et al., 2010).

Ago HITS-CLIP was the first to offer a general means of mapping precise miRNA target sites and has been widely applied to cultured cells (Haecker et al., 2012; Hafner et al., 2010; Kim et al., 2015; Kishore et al., 2011; Leung et al.; Loeb et al., 2012; Riley et al., 2012a; Xue et al., 2013), tissues (Boudreau et al., 2014; Chi et al., 2009; Kameswaran et al., 2014), and even to a whole organism (*C. elegans*) (Grosswendt et al., 2014; Zisoulis et al., 2010). It also provides an opportunity to analyze general properties of miRNA-target interactions in an unbiased and transcriptome-wide manner. Intriguingly, Ago HITS-CLIP analysis revealed that not all identified direct Ago-

target interactions follow canonical seed-pairing rules (Chi et al., 2012; Chi et al., 2009). Indeed, seed matches for ~90% of Ago-bound miRNAs only explain ~73% of Ago-mRNA interactions in the mouse brain, and the rest of ~27% are orphans with no predicted seed matches, strongly suggesting that non-canonical miRNA target sites might be prevalent *in vivo*.

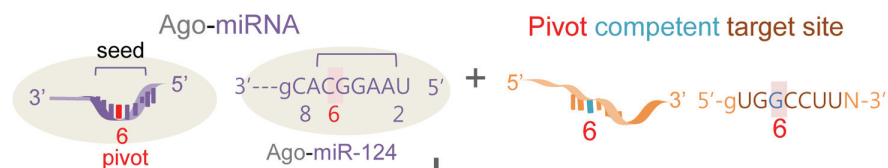
### WIDESPREAD NON-CANONICAL INTERACTIONS: NUCLEATION BULGES AND SEED-LIKE MOTIFS

Non-canonical miRNA-target sites called “nucleation bulges” were identified by analyzing Ago HITS-CLIP “orphan clusters” (Chi et al., 2012). Initially, G-bulge sites for miR-124 were found to be abundant in the mouse brain, where the target sites matched to the seed (positions 2-7) contained a bulged-out G nucleotide corresponding to position between 5 and 6 of the miRNA (miR-124 for *Mink1*, Fig. 1C). Upon further analysis, a general rule governing such non-canonical interactions, called the “pivot pairing rule”, was proposed (Fig. 2) (Chi et al., 2012). This rule dictates that the nucleotide composition in the bulge position should be determined by the base-pairing competency to a nucleotide in position 6 of the miRNA, named “pivot” to confer the thermodynamic stability on the consecutive 5 base pairs of nucleation (termed “nucleation bulge”; Figs. 2A-2C). Otherwise, the base-pair-mediated interaction of only the 4 nt resulting from the non-competent nucleotide in the pivot (non-nucleation bulge) is unstable and not functional. The “pivot pairing rule” well accommodates a canonical interaction mode through nucleation pairing in position 2-6, which is followed by the propagation to position 6-8 of the miRNA (Fig. 2D).

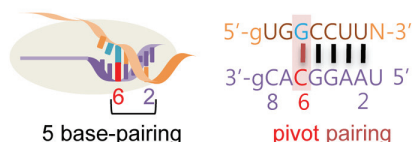
Application of the pivot pairing rule successfully decoded the non-canonical nucleation bulge sites, comprising ≥15% of all Ago-miRNA-mRNA interactions in the mouse brain (Chi et al., 2012). Nucleation bulge sites were also observed in Ago HITS-CLIP analyses performed in the human brain (Boudreau et al., 2014) and several cell lines (Hafner et al., 2010). In addition, their sequences are evolutionally conserved (Chi et al., 2012). The pivot pairing rule improved both quality and quantity of miRNA target sites in their identification (Stefani and Slack, 2012) since it can serve as a general rule that can be incorporated in any computational analysis (Kim et al., 2013).

In addition to nucleation bulges, “seed-like motifs” that contain mismatches in seed pairing were found by examining differential Ago HITS-CLIP binding sites in miR-155 deficient T cells (miR-155 for *Gimap3*, Fig. 1C), indicating that ~20% of Ago-miR-155 binding sites contain seed-like motifs (Loeb et al., 2012). Recent variants of Ago HITS-CLIP, which can directly sequence the binding sites together with miRNAs by inducing ligation between miRNAs and their target RNA sites, also confirmed the widespread occurrence of non-canonical “seed-like motifs” including some “nucleation bulges” (Grosswendt et al., 2014; Helwak et al., 2013; Moore et al., 2015). Initially, CLASH (cross-linking, ligation, and sequencing of hybrids) identified that ~60% of seed interactions are “seed-like motifs” in HEK293T cells (Helwak et al., 2013). Soon after, modified PAR-CLIP and analyses of miRNA-target chimeras in Ago HITS-CLIP showed that ~30% of miRNA-target chimeras were also “seed-like motifs” (Grosswendt et al., 2014). More recently, CLEAR-CLIP (covalent ligation of endogenous Argonaute-bound RNAs) also reported the same observation (Moore et al., 2015). Although most of the non-canonical interactions were identified by Ago HITS-CLIP, it would be difficult to explain their occurrence through globally applicable rules. However, non-canonical “nucleation bulges” could be governed by a general

**A** Ago-miRNA and target interaction



**B** Stable transitional nucleation

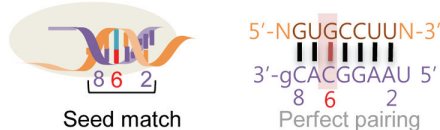


**C** Nucleation bulge interaction



Non-canonical target repression

**D** Seed match interaction



Canonical target repression

**Fig. 2.** The pivot pairing rule and the transitional nucleation model. (A) Ago-miRNA recognizes the pivot (position 6) competent target sites. (B) Pivot competent target sites induce stable transitional nucleation (5 consecutive pairings in position 2-6). (C-D) Transitional nucleation consequently initiates the formation of nucleation bulge (C) or seed match (D) interactions, depending on their sequences.

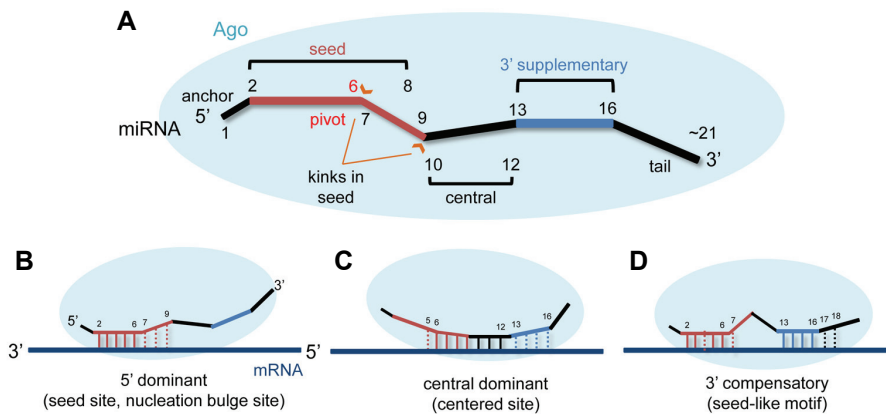
rule of “pivot pairing”, leading to further insights into miRNA-target recognition within other similar transcriptome-wide non-canonical interactions.

**TRANSITIONAL NUCLEATION MODEL**

To explain the pivot pairing rule, a hypothetical phase named “transitional nucleation state” was proposed (Fig. 2) (Chi et al., 2012). Combining the concepts of a “nucleus” (Filipowicz, 2005; Rajewsky, 2006; Tomari and Zamore, 2005) and findings from structural studies for recognition mechanisms of Ago silencing complexes - nucleation, propagation and cleavage of target RNAs (Schirle et al., 2014; Wang et al., 2009), transitional nucleation is defined as a transient miRNA-target duplex with a 5-base-paired nucleation (position 2-6) (Fig. 2B). If the transitional nucleation becomes sufficiently stable to form, this state may be further transformed into a bulge formation where the originally matched pivot nucleotide in position 6 becomes bulged-out and subsequently extended to hybridization towards the 3' end of the miRNA (further than position 6, Fig. 2) (Chi et al., 2012). This model is also well supported by several structural studies of Ago (Elkayam et al., 2012; Nakanishi et al., 2012; Schirle and MacRae, 2012) where nucleotides poised for transitional nucleation (position 2-6) are particularly prearranged becoming A-form helical structures, which are susceptible for base pairing. Intriguingly, such A-form-like helical geometry is disrupted after a pivot (between position 6 and 7) formed by a kink resulting from the insertion of the amino acid isoleucine (I365) from the human Ago2 protein (Elkayam et al., 2012; Schirle and MacRae, 2012). Thus, in theory, any target site pairing to the seed region (either a seed match or a nucleation bulge) re-

quires a shift of this nuclear helix to overcome the kink. In support of this model, single-molecule analysis showed such step-wise processes whereby Ago2 initially scans for target sites using a small region (position 2-4) (Chandradoss et al., 2015) and subsequently mediates a rapid and stable binding to the seed region of a miRNA (Jo et al., 2015; Salomon et al., 2015), serving as a proofreading procedure for target recognition (Yao et al., 2015).

miRNA is reshaped by loading onto Ago, being divided into several functional domains-the anchor, seed, central, 3' supplementary, and tail regions (Fig. 3) (Salomon et al., 2015; Schirle et al., 2014; Wee et al., 2012). Importantly, the seed region has two prearranged continuous base stacking configurations (positions 2-6 and 7-9) caused by kinks at nucleotides 6-7 and 9-10 (Fig. 3A) (Elkayam et al., 2012; Schirle and MacRae, 2012; Schirle et al., 2014). Therefore, transitional nucleation starts pairing through helix 2-6 and subsequently propagates to helices 7-9, overcoming the kink at 6-7 for cases of 5' dominant interactions (Fig. 3B), such as the seed (Bartel, 2009) or nucleation bulge sites (Chi et al., 2012). The opposite may also happen for central dominant interactions (Shin et al., 2010) - the interaction could be initiated by pairing through helix 7-9 along with the central region (positions 10-12) and further expand up to the 3' supplementary region (positions 13-16) (Fig. 3C). In this case, crossing the barrier of the kink at position 9-10 may be required. For “seed-like motifs”, where seed sites contain mismatches, deletions, or wobble pairings (Grosswendt et al., 2014; Helwak et al., 2013; Loeb et al., 2012; Moore et al., 2015), transitional nucleation may require 3' compensatory interactions (Fig. 3D), which could be a general determinant of additional specificity for Ago binding as shown by CLEAR-CLIP



**Fig. 3.** Modes of miRNA-target recognition. (A) Schematic diagrams of functional miRNA domains structured by Ago. (B-D) Possible initial target recognition models for 5' dominant binding (seed site or nucleation bulge site, B), central dominant pairing (centered site, C), and 3'-compensatory interactions (seed-like motifs, D). Transitional nucleation pairings are indicated by bold lines. Non Watson-Crick base pairing, mismatches, and deletions are all indicated by dotted lines.

(Moore et al., 2015).

### INSIGHT FOR MARGINAL REPRESSION AND SIGNIFICANCE OF NON-CANONICAL INTERACTIONS

Biological systems initially generate marginally effective non-canonical regulations when existing biological mechanisms require alternative strategies to compensate for what major canonical pathways have been unable to accomplish. Followed by this notion, majority of non-canonical miRNA-target sites were shown to mediate gene repression at a modest level (Chi et al., 2012; Helwak et al., 2013; Lal et al., 2009; Loeb et al., 2012; Moore et al., 2015) only except for centered sites (Shin et al., 2010), which can trigger slicing activity of Ago but only exist as few in whole transcriptome. However, such modest repression, shown by which the most of non-canonical target sites including nucleation bulges for miRNAs, was often observed as insignificant in large-scale gene expression analyses (Agarwal et al., 2015). These were possibly because the marginal repression was confounded by the issue of cellular heterogeneity, variability derived from secondary effects of target repression, and differences in sensitivity and threshold used in the analyses, or non-canonical sites identified by Ago CLIP based methods could be the true interaction but may not be always functional as the consequence of transient bindings for searching targets (Chandradoss et al., 2015) or as requiring combinatorial occurrences of target sites (Krek et al., 2005). In fact, evaluation of target repression at individual single cell level (Moore et al., 2015) and gene expression analyses in combination with CLIP data (Chi et al., 2012; Grosswendt et al., 2014; Helwak et al., 2013; Loeb et al., 2012; Moore et al., 2015) did observe significant repression mediated by non-canonical sites albeit the effect is still marginal. Since CLIP data only indicate the direct bindings, they should be analyzed together with gene expression data to access the functionality. Future studies should be performed carefully to clear out such issues whenever they analyze marginal effects from non-canonical interactions.

The modest effects from the widespread non-canonical sites are likely to be caused by reduced numbers of targets bound by Ago-miRNA (Chi et al., 2012; Grosswendt et al., 2014; Helwak et al., 2013; Loeb et al., 2012; Moore et al., 2015) probably due to low binding affinity. In support of this, any mismatches or wobbles in the seed region decrease target binding but enhance the turnover of the RISC complex (Wee et al., 2012), suggesting that non-canonical binding can induce intermediate

affinity without affecting the concentration of the Ago complex. Additionally, this may be a mechanism that provides an unbound Ago-miRNA complex to adjacent target sites, as shown by the single-molecule analysis where the lateral diffusion from weakly bound Ago promoted cooperation between neighboring target sites (Chandradoss et al., 2015). As a result, combination of canonical and non-canonical sites may provide a variety of spectra in the regulation of gene expression, enabling a fine-tuning of repression activity. Moreover, relative to canonical seed sites, non-canonical sites have modest sequence conservation across species (Chi et al., 2012; Grosswendt et al., 2014; Loeb et al., 2012; Moore et al., 2015), suggesting that they may be evolutionary intermediates under selective pressure for a shift towards high affinity seed sites. In addition, gene ontology analysis of Ago HITS-CLIP showed that the majority of non-canonical targets have similar functions to the canonical ones, although they are slightly different in detail (Chi et al., 2012; Loeb et al., 2012; Moore et al., 2015). This indicates that non-canonical targets may have a different biological function that needs to be acquired to improve or compensate for the canonical targets.

### CONCLUSION

Although several non-canonical miRNA-target sites were reported as functional, they did not receive much attention since they could not be definitely defined (Brodersen and Voinnet, 2009). However, Ago HITS-CLIP method, which can generate a precise transcriptome-wide map of miRNA target sites (Chi et al., 2009), unexpectedly revealed that large portion of miRNA-target interactions are non-canonical (Chi et al., 2012). Advances in Ago HITS-CLIP analyses further identified non-canonical "nucleation bulges" (Chi et al., 2012) and "seed-like motifs" (Grosswendt et al., 2014; Helwak et al., 2013; Loeb et al., 2012; Moore et al., 2015), expanding our knowledge in the understanding of miRNA targets and their functions. Moreover, the transitional nucleation model, yielded by the analytic process of explaining the pattern of nucleation bulges, offers a general molecular model that can be used to understand the mechanism of miRNA target recognition through seed regions (Chi et al., 2012). Extending this knowledge to applications of RNA silencing, modified siRNAs that contain abasic substitution in the pivot (position 6) were recently developed to completely eliminate miRNA-like off-target repression (Lee et al., 2015; Seok et al., 2016). Biological significance of the non-canonical

interactions could be interpreted as evolutionary intermediates with slightly distinct functions. However, the biological functions postulated here remain to be corroborated.

## ACKNOWLEDGMENTS

This work was supported by grants from the Korean Health Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (H115C3137) to S.W.C and National Research Foundation of Korea funded by the Ministry of Education (NRF-2014R1A1A2A16055016) to H.S.

## REFERENCES

- Agarwal, V., Bell, G.W., Nam, J.W., and Bartel, D.P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4.
- Ambros, V. (2004). The functions of animal microRNAs. *Nature* 431, 350-355.
- Baek, D., Villen, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. (2008). The impact of microRNAs on protein output. *Nature* 455, 64-71.
- Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215-233.
- 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. *Nat. Genet.* 35, 215-217.
- Boudreau, R.L., Jiang, P., Gilmore, B.L., Spengler, R.M., Tirabassi, R., Nelson, J.A., Ross, C.A., Xing, Y., and Davidson, B.L. (2014). Transcriptome-wide discovery of microRNA binding sites in human brain. *Neuron* 81, 294-305.
- Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. (2005). Principles of microRNA-target recognition. *PLoS Biol.* 3, e85.
- Brodersen, P., and Voinnet, O. (2009). Revisiting the principles of microRNA target recognition and mode of action. *Nat. Rev. Mol. Cell Biol.* 10, 141-148.
- Chandradoss, S.D., Schirle, N.T., Szczepaniak, M., MacRae, I.J., and Joo, C. (2015). A dynamic search process underlies microRNA targeting. *Cell* 162, 96-107.
- Chi, S.W., Zang, J.B., Mele, A., and Darnell, R.B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 460, 479-486.
- Chi, S.W., Hannon, G.J., and Darnell, R.B. (2012). An alternative mode of microRNA target recognition. *Nat. Struct. Mol. Biol.* 19, 321-327.
- Croce, C.M. (2009). Causes and consequences of microRNA dysregulation in cancer. *Nat. Rev. Genet.* 10, 704-714.
- Didiano, D., and Hobert, O. (2006). Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat. Struct. Mol. Biol.* 13, 849-851.
- Easow, G., Teleman, A.A., and Cohen, S.M. (2007). Isolation of microRNA targets by miRNP immunoprecipitation. *RNA* 13, 1198-1204.
- Eichhorn, S.W., Guo, H., McGeary, S.E., Rodriguez-Mias, R.A., Shin, C., Baek, D., Hsu, S.H., Ghoshal, K., Villen, J., and Bartel, D.P. (2014). mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. *Mol. Cell* 56, 104-115.
- Elkayam, E., Kuhn, C.D., Tocilij, A., Haase, A.D., Greene, E.M., Hannon, G.J., and Joshua-Tor, L. (2012). The structure of human argonaute-2 in complex with miR-20a. *Cell* 150, 100-110.
- Fabian, M.R., Sonenberg, N., and Filipowicz, W. (2010). Regulation of mRNA translation and stability by microRNAs. *Ann. Rev. Biochem.* 79, 351-379.
- Filipowicz, W. (2005). RNAi: the nuts and bolts of the RISC machine. *Cell* 122, 17-20.
- Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19, 92-105.
- Grimson, A. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* 27, 91-105.
- Grosswendt, S., Filipchuk, A., Manzano, M., Klironomos, F., Schilling, M., Herzog, M., Gottwein, E., and Rajewsky, N. (2014). Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Mol. Cell* 54, 1042-1054.
- 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.
- Haecker, I., Gay, L.A., Yang, Y., Hu, J., Morse, A.M., McIntyre, L.M., and Renne, R. (2012). Ago HITS-CLIP expands understanding of Kaposi's sarcoma-associated herpesvirus miRNA function in primary effusion lymphomas. *PLoS Pathog.* 8, e1002884.
- Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., et al. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141, 129-141.
- Hammell, M. (2008). mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. *Nat. Methods* 5, 813-819.
- He, L., and Hannon, G.J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* 5, 522-531.
- Hebert, S.S., and De Strooper, B. (2009). Alterations of the microRNA network cause neurodegenerative disease. *Trend Neurosci.* 32, 199-206.
- Helwak, A., Kudla, G., Dudnakova, T., and Tollervey, D. (2013). Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* 153, 654-665.
- Hendrickson, D.G., Hogan, D.J., Herschlag, D., Ferrell, J.E., and Brown, P.O. (2008). Systematic identification of mRNAs recruited to argonaute 2 by specific microRNAs and corresponding changes in transcript abundance. *PLoS One* 3, e2126.
- Jo, M.H., Shin, S., Jung, S.R., Kim, E., Song, J.J., and Hohng, S. (2015). Human argonaute 2 has diverse reaction pathways on target RNAs. *Mol. Cell* 59, 117-124.
- John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human MicroRNA targets. *PLoS Biol.* 2, e363.
- Kameswaran, V., Bramswig, N.C., McKenna, L.B., Penn, M., Schug, J., Hand, N.J., Chen, Y., Choi, I., Vourekas, A., Won, K.J., et al. (2014). Epigenetic regulation of the DLK1-MEG3 microRNA cluster in human type 2 diabetic islets. *Cell Metabol.* 19, 135-145.
- Karginov, F.V. (2007). A biochemical approach to identifying microRNA targets. *Proc. Natl. Acad. Sci. USA* 104, 19291-19296.
- Kim, V.N. (2005). Small RNAs: classification, biogenesis, and function. *Mol. Cells* 19, 1-15.
- 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, K.K., Ham, J., and Chi, S.W. (2013). miRTCat: a comprehensive map of human and mouse microRNA target sites including non-canonical nucleation bulges. *Bioinformatics* 29, 1898-1899.
- Kim, S., Seo, D., Kim, D., Hong, Y., Chang, H., Baek, D., Kim, V.N., Lee, S., and Ahn, K. (2015). Temporal landscape of microRNA-mediated host-virus crosstalk during productive human cytomegalovirus infection. *Cell Host Microbe* 17, 838-851.
- Kishore, S., Jaskiewicz, L., Burger, L., Hausser, J., Khorshid, M., and Zavolan, M. (2011). A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat. Methods* 8, 559-564.
- Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* 17, 909-915.
- Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 42, D68-73.
- Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., et al. (2005). Combinatorial microRNA target predictions. *Nat. Genet.* 37, 495-500.
- Lal, A., Navarro, F., Maher, C.A., Maliszewski, L.E., Yan, N., O'Day, E., Chowdhury, D., Dykxhoorn, D.M., Tsai, P., Hofmann, O., et al. (2009). miR-24 Inhibits cell proliferation by targeting E2F2, MYC, and other cell-cycle genes via binding to "seedless" 3'UTR microRNA recognition elements. *Mol. Cell* 35, 610-625.
- 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, H.S., Seok, H., Lee, D.H., Ham, J., Lee, W., Youm, E.M., Yoo, J.S., Lee, Y.S., Jang, E.S., and Chi, S.W. (2015). Abasic pivot

- substitution harnesses target specificity of RNA interference. *Nat. Commun.* **6**, 10154.
- Leung, A.K., Young, A.G., Bhutkar, A., Zheng, G.X., Bosson, A.D., Nielsen, C.B., and Sharp, P.A. (2011). Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. *Nat. Struct. Mol. Biol.* **18**, 237-244.
- Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. *Cell* **115**, 787-798.
- Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20.
- Licatalosi, D.D. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464-469.
- Lim, L.P. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769-773.
- 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.
- Loeb, G.B., Khan, A.A., Canner, D., Hiatt, J.B., Shendure, J., Damell, R.B., Leslie, C.S., and Rudensky, A.Y. (2012). Transcriptome-wide miR-155 binding map reveals widespread noncanonical microRNA targeting. *Mol. Cell* **48**, 760-770.
- Long, D., Lee, R., Williams, P., Chan, C.Y., Ambros, V., and Ding, Y. (2007). Potent effect of target structure on microRNA function. *Nat. Struct. Mol. Biol.* **14**, 287-294.
- Mili, S., and Steitz, J.A. (2004). Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* **10**, 1692-1694.
- Moore, M.J., Scheel, T.K., Luna, J.M., Park, C.Y., Fak, J.J., Nishiuchi, E., Rice, C.M., and Darnell, R.B. (2015). miRNA-target chimeras reveal miRNA 3'-end pairing as a major determinant of Argonaute target specificity. *Nat. Commun.* **6**, 8864.
- Mourelatos, Z. (2008). Small RNAs: the seeds of silence. *Nature* **455**, 44-45.
- Nakanishi, K., Weinberg, D.E., Bartel, D.P., and Patel, D.J. (2012). Structure of yeast Argonaute with guide RNA. *Nature* **486**, 368-374.
- Olson, E.N. (2014). MicroRNAs as therapeutic targets and biomarkers of cardiovascular disease. *Sci. Translational Med.* **6**, 239ps233.
- Park, C.Y., Choi, Y.S., and McManus, M.T. (2010). Analysis of microRNA knockouts in mice. *Hum. Mol. Genet.* **19**, R169-175.
- Poy, M.N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., Macdonald, P.E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P., et al. (2004). A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* **432**, 226-230.
- Rajewsky, N. (2006). MicroRNA target predictions in animals. *Nat. Genet.* **38**, S8-S13.
- Riley, K.J., Rabinowitz, G.S., Yario, T.A., Luna, J.M., Darnell, R.B., and Steitz, J.A. (2012a). EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. *EMBO J.* **31**, 2207-2221.
- Riley, K.J., Yario, T.A., and Steitz, J.A. (2012b). Association of Argonaute proteins and microRNAs can occur after cell lysis. *RNA* **18**, 1581-1585.
- Salomon, W.E., Jolly, S.M., Moore, M.J., Zamore, P.D., and Serebrov, V. (2015). Single-molecule imaging reveals that argonaute reshapes the binding properties of its nucleic acid guides. *Cell* **162**, 84-95.
- Schirle, N.T., and MacRae, I.J. (2012). The crystal structure of human Argonaute2. *Science* **336**, 1037-1040.
- Schirle, N.T., Sheu-Gruttadauria, J., and MacRae, I.J. (2014). Structural basis for microRNA targeting. *Science* **346**, 608-613.
- Selbach, M., Schwanhauss, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58-63.
- Seok, H., Jang, E.S., and Chi, S.W. (2016). Rationally designed siRNAs without miRNA-like off-target repression. *BMB Rep.* **49**, 135-136.
- Shin, C., Nam, J.W., Farh, K.K., Chiang, H.R., Shkumatava, A., and Bartel, D.P. (2010). Expanding the microRNA targeting code: functional sites with centered pairing. *Mol. Cell* **38**, 789-802.
- Sim, S.E., Bakes, J., and Kaang, B.K. (2014). Neuronal activity-dependent regulation of microRNAs. *Mol. Cells* **37**, 511-517.
- Stark, A., Brennecke, J., Russell, R.B., and Cohen, S.M. (2003). Identification of *Drosophila* microRNA targets. *PLoS Biol.* **1**, E60.
- Stefani, G., and Slack, F.J. (2012). A 'pivotal' new rule for microRNA-mRNA interactions. *Struct. Mol. Biol.* **19**, 265-266.
- Tay, Y., Zhang, J., Thomson, A.M., Lim, B., and Rigoutsos, I. (2008). MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* **455**, 1124-1128.
- Tomari, Y., and Zamore, P.D. (2005). Perspective: machines for RNAi. *Genes Dev.* **19**, 517-529.
- Ule, J. (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science* **302**, 1212-1215.
- Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K., and Slack, F.J. (2004). The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'primeUTR. *Genes Dev.* **18**, 132-137.
- Wang, Y., Juraneck, S., Li, H., Sheng, G., Wardle, G.S., Tuschl, T., and Patel, D.J. (2009). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **461**, 754-761.
- Wee, L.M., Flores-Jasso, C.F., Salomon, W.E., and Zamore, P.D. (2012). Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell* **151**, 1055-1067.
- Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell* **75**, 855-862.
- Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., and Kellis, M. (2005). Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**, 338-345.
- Xue, Y., Ouyang, K., Huang, J., Zhou, Y., Ouyang, H., Li, H., Wang, G., Wu, Q., Wei, C., Bi, Y., et al. (2013). Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. *Cell* **152**, 82-96.
- Yao, C., Sasaki, H.M., Ueda, T., Tomari, Y., and Tadakuma, H. (2015). Single-molecule analysis of the target cleavage reaction by the *Drosophila* RNAi enzyme complex. *Mol. Cell* **59**, 125-132.
- Yekta, S., Shih, I.H., and Bartel, D.P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594-596.
- Zhang, C., and Darnell, R.B. (2011). Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat. Biotechnol.* **29**, 607-614.
- Zisoulis, D.G., Lovci, M.T., Wilbert, M.L., Hutt, K.R., Liang, T.Y., Pasquinelli, A.E., and Yeo, G.W. (2010). Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* **17**, 173-179.