ChemComm

A microRNA-gated thgRNA platform for multiplexed activation of gene expression in mammalian cells

Journal:	*ChemComm*
Manuscript ID	CC-COM-03-2022-001478.R1
Article Type:	Communication

COMMUNICATION

A microRNA-gated thgRNA platform for multiplexed activation of gene expression in mammalian cells

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Victoria M. Hunt and Wilfred Chen*

To effectively reprogram cellular regulatory networks towards desired phenotypes, it is critical to have the ability to provide precise gene regulation in a spatiotemporal manner. We have previously engineered toehold-gated guide RNA (thgRNA) to enable conditional activation of dCas9-mediated transcriptional upregulation in mammalian cells using synthetic RNA triggers. Here, we demonstrate that microRNA (miR)-gated thgRNAs can be transcribed by type II RNA polymerase to allow multiplexed transcriptional activation using both mRNA and miR. Activation is achieved only by proper miR-mediated processing of the flanking 5' cap and 3' poly A tail and hairpin unblocking by mRNA via strand displacement. This new AND-gate design is exploited to elicit conditional protein degradation based on induced expression of a specific ubiquibody. This new strategy may find many new applications in an RNA-responsive manner.

Recent advances in CRISPR-Cas9 technology have revolutionized our ability to perform precise genome editing and gene regulation.[1,2] Cas9 is an RNA-guided endonuclease that can be programmed for targeted DNA binding and cleavage.[3] Nuclease-null Cas9 (dCas9) proteins have been repurposed as highly specific DNA binders and have been exploited for transcriptional activation and repression by fusing the appropriate effector domains.[4-7] Since dCas9 protein targets DNA in an RNA-responsive manner, switchable gRNA motifs have been created to enable dynamic gene regulation. Recently, we reported a new class of dCas9 regulators containing a toehold-gated sgRNA (thgRNA) that is inactive in the absence of an RNA trigger.[8] Upon expression of the trigger, the thgRNA is unblocked by toehold-mediated strand displacement[9] and the spacer sequence is free to guide DNA binding.

While this new strategy has been successfully demonstrated for conditional gene regulation in mammalian cells,[8,10-12] most activations were achieved using synthetic RNA triggers and only one example of using *survivin* mRNA has been demonstrated.[12] However, the precise activation fold in protein expression was not reported. Our lab has recently demonstrated that the level of thgRNA activation is more than 3-fold lower for RNA triggers expressed using the CMV promoter versus the hU6 promoter.[13] This result indicates that the rapid translocation of RNA triggers expressed using the type II polymerase reduces duration of strand displacement within the nucleus for efficient thgRNA activation. This limitation renders full-length mRNA ineffective as a trigger for activation.

Since mRNA are rapidly exported into the cytosol for protein synthesis, a new strategy that enables efficient strand displacement to occur within the cytosol is needed. While gRNAs have been expressed by a CMV promoter or other type II promoters, early reports observed a lack of function as these gRNAs still contain the flanking 5' cap and poly A tail, which hinder their incorporation into dCas9.[4,14-16] To bypass this limitation, there were approaches where gRNAs were processed by self-cleaving ribozymes and/or tRNAs.[17-19] Additionally, gRNAs could be expressed by a type II promoter and then processed using the Cas6/Cys4 system to cleave long strands of RNA into mature sgRNAs.[20] Unfortunately, these processing approaches are constitutive in nature and do not provide conditional control. A new conditional approach has been reported by exploiting the use of flanking microRNA (miR) binding sites on the 5' and 3' end of the gRNA for processing.[21] When the specific miR is present, the pre-gRNA is processed into mature gRNA in the cytosol. miRs are intriguing targets for induction of CRISPR/Cas gene regulation because they are expressed in specific tissues and cell types, and their levels can undergo dynamic changes in various stages of development or disease.[22]

In this work, we designed a new class of miR-gated thgRNAs by interfacing miR-mediated processing with toehold-mediated strand displacement to overcome the hurdle of using full-length mRNAs in the cytosol for conditional activation. We discovered the design principles of this new approach and demonstrated its broad applicability to different trigger sequences including the use of full-length mCherry mRNA for activation. By using both miR-21 and an RNA trigger as inputs, several AND gate designs were used successfully to activate GFP expression as well conditional protein degradation. For all gRNAs used in this work, see Table S1.

We have previously designed thgRNA, g2v2thg-F,[13] containing a 5nt toehold region, 15nt loop, and a fully hybridized stem region that is activated by a specific region of the mCherry mRNA. A short synthetic trigger was shown to activate GFP expression from a minimum CMV promoter up to 15-fold based on CRISPRa. Unfortunately, we did not detect any activation in the presence of full-length mCherry mRNA (Fig. S1). It is clear that mRNA translocation into cytosol is too fast for efficient strand displacement and thgRNA activation.

To address this limitation, we further tested to see if we could exploit the miR-inducible approach to provide conditional processing

[a] *Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE, USA. E-mail: wilfred@udel.edu; Fax:+1 30 2831 1048;Tel:+1 30 2831 1048*

† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

and activation of thgRNA in the presence of cytoplasmic full-length mRNA. In this design, thgRNAs are expressed from a CMV promoter, which results in their localization to the cytosol where full-length RNAs reside. Proper processing can only occur if the appropriate miRs are present to enable cleavage of the flanking 5' cap and poly A tail for dCas9 complexation (Fig. 1A).

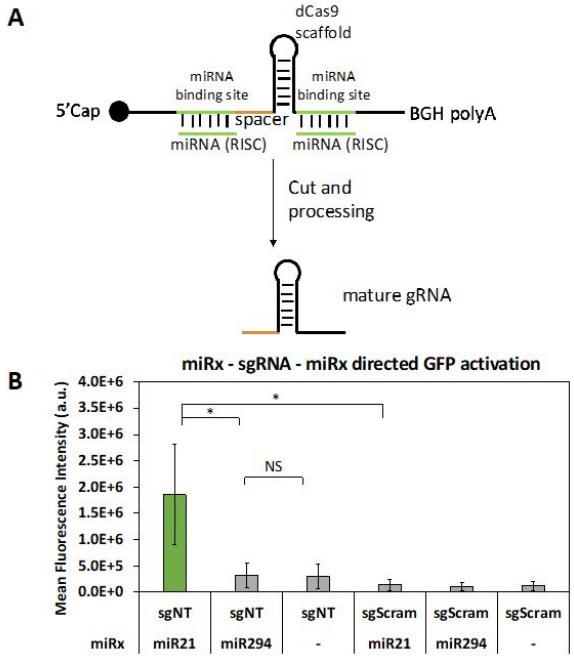

Figure 1. (A) Depiction of miR-inducible CRISPR/Cas9. A pre-gRNA is flanked by two miR binding sites. After specific miR binding, mature gRNA is formed by RISC-mediated cleavage. (B) Confirmation of miRx-sgRNA-miRx behavior in HeLa cells. Only highly expressed miR21 was able to induce processing and gRNA activation. No difference in GFP expression was observed when miR294 was used or when sgNT was expressed directly from a CMV promoter or when scrambled gRNA (sgScram) was used as a negative control. Data are the average of three or more biological replicates (separately transfected wells), and the error is the standard deviation (SD) of the mean. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, NS = no statistically significant difference.

We first confirmed that the miR-gated sgRNA approach could function effectively in HeLa cells. Using a previously developed sgRNA design that can activate GFP expression in HeLa cells, we inserted flanking sequences of either miR21 or miR294 to both the 5' and 3'end to generate miR21-sgNT-miR21 and miR294-sgNT-miR294 (Fig. 1A). These two miRs were chosen because miR21 is known to be overexpressed in HeLa,[23] while miR294 is expressed at negligible levels.[24] Using cells expressing a scrambled sgRNA as the background control, clear activation was observed only with miR21-sgNT-miR21, resulting in a 13-fold increase in GFP expression (Fig. 1B). A small level of activation was observed with miR294-sgNT-miR294 likely due to a small degree of unprocessed sgRNA complexing with dCas9. This low GFP level is similar to that observed with cells expressing sgRNA from a CMV promoter, highlighting the importance of removing both end modifications for proper sgRNA activation.

Since the original design requires flanking miR-targeting sequences in both the 5' and 3' end, we next investigated if it was possible to create additional logic gates by individually altering the respective ends. We designed constructs miR21-sgNT-miR294, where the 3' polyA tail is intact and miR294-sgNT-miR21, which maintains the flanking 5' cap, to evaluate the levels of GFP activation. Both sgRNA designs were capable of activating GFP expression partially as compared to the dually processed miR21-sgNT-miR21 (Fig. S2). While both miR binding sites are important for fully functional mature sgRNA, an OR gate design could be constructed if the level of partial activation is sufficient to yield desired results.

After confirming the behavior of the miR-inducible gRNA approach in HeLa cells, we further tested whether this strategy could be combined with thgRNA to provide a new miR-gated thgRNA approach for conditional activation using RNA triggers that are transported to the cytosol (Fig. 2A). This combination provides the feasibility to create an AND gate architecture where both miRs and mRNAs serve as inputs to activate gene regulation. We added two flanking miR binding sites to thgNT-F (Fig. S4A), a thgRNA design with the best signal to background ratio in HeLa cells.[13] A synthetic trigger was placed behind the miR-gated sgRNA and is released by a sandwiching hammerhead (HH) and hepatitis delta virus (HDV) ribozyme (Fig. S3). Surprisingly, miR21-thgNT-F-miR21 showed a high level of background leak, while thgNT-F by itself has negligible background (Fig. S5A).[13] Since the toehold sequence has been shown to interact with the loop region of the thgRNA, thgNT-F was designed specifically with a short 5-nt toehold region to minimize background leak. Because cleavage by the RISC complex occurs between the RNA targets that pair with bases 10 and 11 of the miR strand,[25] there is a 11nt fragment remaining on both the 5' and 3' ends (Fig. S4B). We suspect that addition of the extra 11nt flanking sequence to the 5nt toehold led to increased non-specific interaction with the flexible loop. This is further supported by observing background leak when the two flanking miR21 binding sites were added to thgNT-E (Fig. S4C), which contains no toehold (Fig. S5B).

Since we cannot eliminate the remaining 11nt fragment after miR processing, one way to address the background leak is to eliminate the flexible loop. We have shown that thgNT-H (Fig. S4D), which contains no flexible loop, has no background leak but also minimal activation.[13] When we adapted thgNT-H with the miR-gated design, only a small leak of 3% was observed for miR21-thgNT-H miR21 (Fig. 2B). This result confirms that the additional 11nt flanking sequence is unable to destabilize the thgRNA structure without the flexible loop. We observed a 7-fold increase in GFP level when both inputs are present (Fig. 2B). Since the presence of either the synthetic trigger or miR-21 alone is insufficient for activation, the current design indeed operates as an AND-gate circuit. To our knowledge, this is the first time multiple RNAs have been used to create an AND gate architecture for dCas9-guided transcriptional activation in mammalian systems.

Finally, we investigated the combined use of miR and mRNA for conditional activation. We designed miR-gated g2v2thg-H targeting region G2 of the mCherry mRNA. Using a synthetic trigger expressed by the CMV promoter targeting the G2 region, we observed a fold change of 2.1 (Fig. 3A). The small leak may reflect other non-specific interactions between the extra 11nt and the stem region in the OFF state. It is likely that the leak could be minimized by either extending the stem region or shortening the toehold region of the thgRNA to minimize regions available for non-specific interaction. We next tested to see whether full-length mcherry mRNA can elicit the same level of activation (Fig S7). Activation in GFP expression was observed only for the miR21-gated g2v2thg-H with mCherry co-expression (Fig. 3B). The need for both mCherry mRNA and miR for activation was demonstrated as no increase in GFP expression was detected for cells expressing mi294-gated g2v2thg-H even with mCherry co-expression. The level of activated GFP expression is similar to that achieved using the synthetic trigger, suggesting our ability to provide full activation.

To demonstrate that this AND-gated strategy could be used to elicit conditional cellular responses, we repurposed the design to

activate expression of a bifunctional ubiquibody GS2-IpaH9.8 to target GFP for conditional degradation (Fig. 4A and Fig. S8).[26] As shown in Fig. 4B, miR21-sgNT-miR21 resulted in 60% knockdown of GFP as compared to cells expressing a miR21-gated scrambled sgRNA. A small level of knockdown was observed with miR294-sgNT-miR294, an observation consistent with the low level of background leak with GFP expression. The feasibility of using an AND gate design to provide conditional protein degradation was further demonstrated. Compared to the control miR294-thgNT-H-miR294, the two-input system (miR21-thgNT-H-miR21 HH trigH* HDV) showed a 65% knockdown in GFP level (Fig 4C). Only a small reduction in GFP was detected either by removing the hairpin blocking alone or by using miR21 processing alone, again confirming the ability to elicit GFP knockdown using an AND gate architecture.

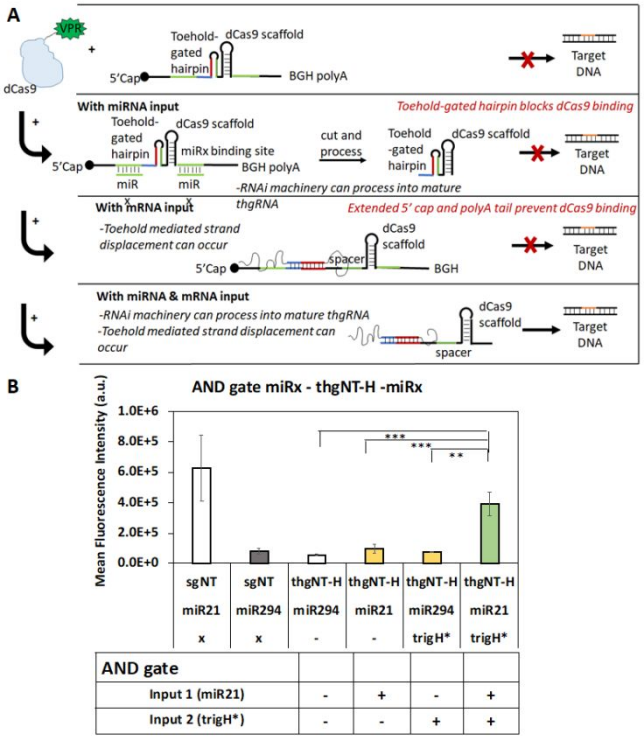

Figure 2. (A) Schematic of the AND gate architecture using miR and mRNA as dual inputs. Both inputs are required (miR21 and mRNA) for proper processing of the 5' and 3' ends and to unblock the spacer region. The activated miR-gated thgRNA can complex with dCas9 for targeted upregulation of GFP expression. (B) Induced GFP expression using the AND gate designs (miRx-thgNT-H-miRx). We observed negligible leak when neither input or only a single input was present (miR21 or trigH*) but observed a 7-fold increase when both inputs were present (miR21 and trigH*). Data are the average of three or more biological replicates (separately transfected wells), and the error is the standard deviation (SD) of the mean. * p<0.05, ** p<0.01, *** p<0.001.

Since virtually any protein could be targeted for degradation by tuning the binding moiety on the ubiquibody,[27] this AND-gated activation strategy provides a highly flexible framework to target many intracellular proteins for conditional degradation. To show this versatility, we replaced the GFP-targeting GS2 monobody with a different monobody, Nsa5,[26] in order to target the SH2 domain of SHP2 for conditional degradation. Simple tracking of degradation by Nsa5-IpaH9.8 was done using a SH2-GFP fusion. More than 75% knockdown was observed in the presence of both inputs, while only a modest decrease was detected with either one of the two inputs (Fig. S9). The higher knockdown level is consistent with the improved degradation efficiency of the Nsa5-IpaH9.8 ubiquibody.[13]

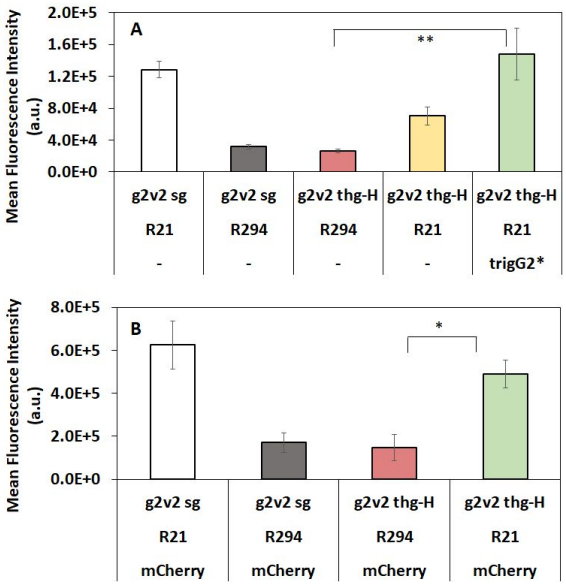

Figure 3. Extending the miR-gated thgRNA design to target mCherry mRNA. (A) A synthetic trigger was first used to evaluate the design. We observed a leak of 36% in the presence of miR21 and a fold change of 2.1. in the presence of both inputs (miR21 and trigG2*). (B) The use of full-length mCherry mRNA as the trigger. Activation was detected only when both miR21 and mCherry mRNA were presented. Data is the average of three or more biological replicates (separately transfected wells), and the error is the standard deviation (SD) of the mean. * p<0.05, ** p<0.01, *** p<0.001.

In this work, we have successfully created thgRNAs that are responsive to endogenous mRNAs in mammalian cells. We showed that thgRNAs expressed from the hU6 promoter cannot interact efficiently with rapidly exported full-length mRNAs in the nucleus. To resolve this issue, miR-gated thgRNAs were designed that are responsive to miR-mediated RISC complex processing in order to allow proper interaction between processed thgRNA and mRNA in the cytosol. The additional miR-mediated processing step not only enables thgRNA activation by full-length mCherry mRNA but also the creation of an AND gate design requiring the presence of both miR and mRNA for activation. This new transcriptional reprogramming framework has the potential to be used as a highly flexible approach to modulate cellular protein content based on any combination of miR and mRNA.

We would like to acknowledge the funding support from NSF (MCB2013991, MCB 1817675, and CBE1803008).

Conflicts of interests

There are no conflicts to declare.

Notes and references

1. B. Wiedenheft, S. H. Sternberg and J. A. Doudna, *Nature*, 2012, **482**, 331-338.
2. P. Horvath and R. Barrangou, *Science*, 2010, **327**, 167-170.
3. R. Sapranauskas, G. Gasiunas, C. Fremaux, R. Barrangou, P. Horvath and V. Siksnys, *Nucleic acids research*, 2011, **39**, 9275-9282.
4. Lei S. Qi, Matthew H. Larson, Luke A. Gilbert, Jennifer A. Doudna, Jonathan S. Weissman, Adam P. Arkin and Wendell A. Lim, *Cell*, 2013, **152**, 1173-1183.

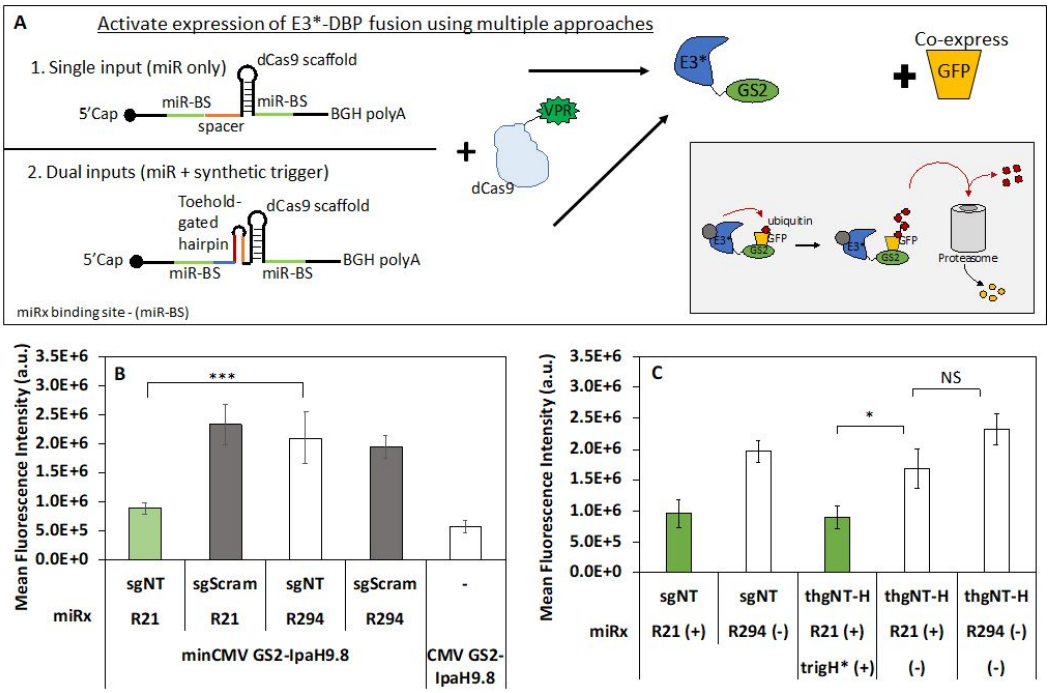

Figure 4. (A) Conditional protein degradation based on induced expression of a specific ubiquibody (E3*-DBP). Expression of E3*-DBP is conditionally upregulated when either (1) a native miR is used to correctly process sgRNA in the cytosol or (2) both a native miR and a cytosolic RNA trigger are used to activate thgRNA. (B) miR-directed conditional protein degradation. We observed 62% knockdown of GFP using this strategy as compared to 76% knockdown using sgRNA expressed from a hU6 promoter. (C) Conditional GFP degradation using an AND gate (miRx and synthetic trigger (trigH*)) design to activate GS2-IpaH9.8 expression. When both inputs are present (miR21 and trigH*) to activate thgRNA properly, we observed 62% knockdown of GFP, while minimal degradation was detected in the presence of only one of the two processing steps. Data are the average of three or more biological replicates (separately transfected wells), and the error is the standard deviation (SD) of the mean. * p<0.05, ** p<0.01, *** p<0.001.

5. D. Bikard, W. Jiang, P. Samai, A. Hochschild, F. Zhang and L. A. Marraffini, *Nucleic acids research*, 2013, **41**, 7429-7437.
6. Luke A. Gilbert, Matthew H. Larson, L. Morsut, Z. Liu, Gloria A. Brar, Sandra E. Torres, N. Stern-Ginossar, O. Brandman, Evan H. Whitehead, Jennifer A. Doudna, Wendell A. Lim, Jonathan S. Weissman and Lei S. Qi, *Cell*, 2013, **154**, 442-451.
7. P. Perez-Pinera, D. D. Kocak, C. M. Vockley, A. F. Adler, A. M. Kabadi, L. R. Polstein, P. I. Thakore, K. A. Glass, D. G. Ousterout, K. W. Leong, F. Guilak, G. E. Crawford, T. E. Reddy and C. A. Gersbach, *Nat Meth*, 2013, **10**, 973-976.
8. K. H. Siu and W. Chen, *Nat Chem Biol*, 2019, **15**, 217-220.
9. R. P. Chen, D. Blackstock, Q. Sun and W. Chen, *Nature Chemistry*, 2018, **10**, 474-481.
10. M. H. Hanewich-Hollatz, Z. Chen, L. M. Hochrein, J. Huang and N. A. Pierce, *ACS Central Science*, 2019, **5**, 1241-1249.
11. Y. Hao, J. Li, Q. Li, L. Zhang, J. Shi, X. Zhang, A. Aldalbahi, L. Wang, C. Fan and F. Wang, *Angewandte Chemie International Edition*, 2020, **59**, 20612-20618.
12. Z.-M. Ying, F. Wang, X. Chu, R.-Q. Yu and J.-H. Jiang, *Angewandte Chemie International Edition*, 2020, **59**, 18599-18604.
13. V. M. Hunt and W. Chen, *ACS Synthetic Biology*, 2022, **11**, 397-405.
14. L. A. Gilbert, M. H. Larson, L. Morsut, Z. Liu, G. A. Brar and S. E. Torres, *Cell*, 2013, **154**.
15. M. L. Maeder, S. J. Linder, V. M. Cascio, Y. Fu, Q. H. Ho and J. K. Joung, *Nature methods*, 2013, **10**.
16. C. Xie, Y.-L. Chen, D.-F. Wang, Y.-L. Wang, T.-P. Zhang, H. Li, F. Liang, Y. Zhao and G.-Y. Zhang, *Scientific reports*, 2017, **7**, 5795.
17. Y. Gao and Y. Zhao, *Journal of Integrative Plant Biology*, 2014, **56**, 343-349.
18. K. Xie, B. Minkenberg and Y. Yang, *Proceedings of the National Academy of Sciences*, 2015, **112**, 3570-3575.
19. L. Xu, L. Zhao, Y. Gao, J. Xu and R. Han, *Nucleic acids research*, 2016, **45**, e28.
20. L. Nissim, Samuel D. Perli, A. Fridkin, P. Perez-Pinera and Timothy K. Lu, *Molecular cell*, 2014, **54**, 698-710.
21. X.-W. Wang, L.-F. Hu, J. Hao, L.-Q. Liao, Y.-T. Chiu, M. Shi and Y. Wang, *Nature Cell Biology*, 2019, **21**, 522-530.
22. P. Landgraf, M. Rusu, R. Sheridan, A. Sewer, N. Iovino, A. Aravin, S. Pfeffer, A. Rice, A. O. Kamphorst, M. Landthaler, C. Lin, N. D. Socci, L. Hermida, V. Fulci, S. Chiaretti, R. Foà, J. Schliwka, U. Fuchs, A. Novosel, R. U. Müller, B. Schermer, U. Bissels, J. Inman, Q. Phan, M. Chien, D. B. Weir, R. Choksi, G. De Vita, D. Frezzetti, H. I. Trompeter, V. Hornung, G. Teng, G. Hartmann, M. Palkovits, R. Di Lauro, P. Wernet, G. Macino, C. E. Rogler, J. W. Nagle, J. Ju, F. N. Papavasiliou, T. Benzing, P. Lichter, W. Tam, M. J. Brownstein, A. Bosio, A. Borkhardt, J. J. Russo, C. Sander, M. Zavolan and T. Tuschl, *Cell*, 2007, **129**, 1401-1414.
23. G. Du, D. Cao and L. Meng, *Molecular medicine reports*, 2017, **15**, 2713-2719.
24. Y. Yan, X. Yang, T. T. Li, K. L. Gu, J. Hao, Q. Zhang and Y. Wang, *Journal of genetics and genomics = Yi chuan xue bao*, 2017, **44**, 179-189.
25. A. J. Pratt and I. J. MacRae, *Journal of Biological Chemistry*, 2009, **284**, 17897-17901.
26. M. B. Ludwicki, J. Li, E. A. Stephens, R. W. Roberts, S. Koide, P. T. Hammond and M. P. DeLisa, *ACS Central Science*, 2019, **5**, 852-866.
27. R. P. Chen, A. S. Gaynor and W. Chen, *Biotechnology Advances*, 2019, **37**, 107446.