The regulation of microRNAs in cancer through novel transcriptional and post-transcriptional mechanisms

Emily C Knouf

A dissertation

submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

James Olson, Chair

Muneesh Tewari

Patrick Paddison

Walter Ruzzo

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

The regulation of microRNAs in cancer through novel transcriptional and post-transcriptional mechanisms

Emily C Knouf

Chair of the Supervisory Committee:
Dr. James Olson
Professor, Department of Pediatrics

MicroRNAs are small non-coding RNA molecules that serve as important regulators of gene expression. While understanding of the functional roles of miRNAs in both normal physiology and disease has rapidly expanded, the regulation of these molecules remains a largely open question. We have investigated the regulation of mRNAs at multiple steps of their biogenesis. We began by studying the transcriptional regulation of miRNAs in ovarian cancer, and developed a computational pipeline by which to identify putative transcription factor: miRNA interactions. We found that the miR-200 family of miRNAs is regulated by p63 and p73, two members of the p53 transcription factor family. The miR-200 miRNAs serve as potent regulators of the epithelial-mesenchymal transitions that influence tumor invasion; therefore, identifying positive regulators may facilitate the future modulation of these miRNAs for therapeutic applications. We have also examined the transcriptional regulation of miR-210, which is potently induced under low oxygen conditions. We found that under normoxic conditions miR-210 is rapidly activated by HIF-1α in response to cell density, demonstrating the importance of a cell's microenvironment in shaping its miRNA expression patterns.

We next examined the downstream regulation of miRNAs at the post-transcriptional level. We discovered that miRNAs frequently show 3' non-templated nucleotide additions, forming a large number of miRNA isomiRs that expand the diversity of the miRNA transcriptome. We identified multiple nucleotidyl transferase enzymes that are responsible for these modifications in a miRNA-specific fashion, including three enzymes—TUT1, MTPAP, and ZCCHC6—not previously known to modify miRNAs. Finally, we have investigated the functional effects these 3' additions and their nucleotidyl transferase regulators exert on miRNA activity. We found 3' additions are neither universally stabilizing nor destabilizing; instead, certain nucleotide additions, such as increased 3' uridylation, are associated with reduced miRNA abundance. Additionally, we identified two nucleotidyl transferases, TUT1 and PAPD4, which broadly maintain the expression of miRNAs. Taken together, our work reveals several novel mechanisms of regulation for miRNAs that are influential in tumor initiation and spread. While some factors control the expression of many miRNAs, other proteins alter the abundance of a subset of miRNAs to enable a precise regulatory response.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

MicroRNA biogenesis and function

MicroRNAs (miRNAs) provide a unique and ubiquitous layer of post-transcriptional regulation of gene expression. MicroRNAs are short (~22 nucleotide), endogenous non-protein-encoding RNA molecules that regulate the stability and/or translation of messenger RNAs (mRNAs) via partially complementary base-pairing interactions. MicroRNAs are most frequently transcribed by RNA polymerase II as long hairpin structures known as primary miRNAs (pri-miRNAs), which are cleaved by the Drosha enzyme and exported into the cytoplasm as 60-70 nt precursor miRNAs (1). After removal of the hairpin loop by the Dicer enzyme complex, the miRNA duplex is unwound and loaded into the RNA- induced silencing complex (RISC), which facilitates the translational repression or destabilization of partially complementary mRNAs (2).

MicroRNAs can repress multiple targets with specificity and efficacy and can quickly respond to changing cellular conditions. These qualities allow miRNAs to serve as important regulators of normal development and differentiation (3). The first miRNA to be discovered was the *C. elegans* miRNA, lin-4, which is required for the completion of larval development (4). In humans, miRNAs also establish and maintain the fate and behaviors of cells. The miR-223 miRNA is an essential component in hematopoiesis, and is necessary for the differentiation of granulocytes (5). Similarly, the miR-200 family of miRNAs maintains the epithelial state of cells by repressing the ZEB transcription factors responsible for epithelial-mesenchymal transitions (6-8). Across multiple species,

miRNAs have been demonstrated to serve as conserved regulatory hubs capable of both driving widespread gene expression changes and also buffering the cell from unwanted changes in transcriptional activity (9).

While miRNAs modulate many aspects of normal physiology, they are frequently dysregulated in many diseases. MicroRNAs show aberrant expression in a variety of cancers (10), which in general feature a widespread downregulation of miRNAs (11). MicroRNAs can be used to classify tumors of unknown origin (11), in addition to serving as tissue and blood-borne biomarkers for the diagnosis and prognosis of multiple cancer types (12). The miRNA processing enzyme Dicer functions as a haploinsufficient tumor suppressor (13), and in certain cancers downregulation of Dicer is associated with poor prognosis (14). Some miRNAs act as oncogenes, such as the miR-17/92 cluster, which represses PTEN—a potent inhibitor of the PI3K/Akt pathway that promotes the survival and proliferation of cancer cells (15,16). MicroRNAs can also serve as tumor suppressors, as demonstrated by the let-7 family, which represses both the Myc and Ras oncogenes (17,18). Other miRNAs, such as the miR-10a/b and miR-200 families, govern cancer metastasis via regulation of the transcription factors that direct epithelial-to-mesenchymal transitions (EMTs) (8,19). Taken together, these studies suggest that miRNAs can act as potent effectors of pathways that promote or prevent cancer.

Regulation of microRNA expression

Because of the important roles miRNAs play in both normal physiology and disease, we sought to understand the mechanisms by which they are regulated. MicroRNA regulation is a multifaceted process that can involve modulation of miRNA

transcription, processing, stability and targeting. While some forms of regulation, such as protection in the RISC complex, affect the majority of miRNAs, other miRNAs are individually regulated by an ensemble of unique protein cofactors. For example, transcription factors provide sequence-specific activation or repression of certain miRNAs. The tumor suppressor p53 promotes transcription of the miR-34 family of miRNAs, which in turn induces cell cycle arrest and further amplifies the tumor-suppressive functions of p53 (20). Additionally, the proto-oncogene Myc activates transcription of the oncogenic miR-17/92 cluster, and represses expression of multiple tumor suppressor miRNAs such as miR-34a and the let-7 family (21,22). An online database of curated miRNA and transcription factor interactions from the literature currently features over 200 transcription factors regulating several hundred miRNAs across 16 species (23). However, the transcriptional regulation of most miRNAs has been poorly understood, and previous investigations have been limited by the lack of annotation of miRNA primary transcripts and their transcription factor binding sites.

Regulation of miRNA expression also occurs during the processing of the primary miRNA into the mature species. The association of auxiliary proteins with the Drosha and Dicer complexes can modulate miRNA biogenesis. Drosha processing is enhanced by interactions with the p68/p72 helicases, which upregulate the processing of a subset of miRNAs (24). Furthermore, association of the p53 transcription factor with p68 also enhanced Drosha cleavage of a smaller number of miRNAs with tumor-suppressive roles (25). While some protein cofactors exert widespread effects, other regulatory mechanisms are restricted to an individual miRNA. For example, binding of the protein hnRNP A1 to the terminal loop of miR-18a promotes Drosha processing, and may drive

preferential expression of this miRNA over the five other miRNAs encoded in a single polycistronic transcript (26). Other proteins affect miRNA biogenesis by modifying the sequence of the precursor and primary miRNAs enzymatically. For instance, the Lin28 protein binds to precursor members of the let-7 family and promotes uridylation by the nucleotidyl transferase ZCCHC11, which prohibits subsequent Dicer processing (27,28). Together these recent reports have identified a large ensemble of new miRNA regulators, but remaining questions include determining their specificity and understanding the cellular contexts in which these regulators are active.

While understanding of the transcriptional and post-transcriptional mechanisms of miRNA regulation has shown substantial gains, the factors governing the fate of mature miRNAs in humans remain largely unknown. MicroRNAs in general display very long half-lives of multiple days (29), although a subset of miRNAs in neurons in *Drosophila* have been demonstrated to show more rapid turnover (30). Recent work in *C. elegans* has identified a 5'-3' exonuclease, XRN-2, that degrades mature miRNAs (31), while in plants the 3'-5' nuclease SDN-1 was shown to affect miRNA turnover (32). In humans, neither the nuclease responsible for degradation or conserved signals for miRNA decay has been identified. One potential mechanism for miRNA decay involves the recognition of certain modifications of the 3' end of a miRNA. Studies in plants and algae have demonstrated that 3' nucleotide additions can modulate miRNA stability (33-35), although in humans the conservation of these regulatory mechanisms has not been addressed. Ultimately, the factors governing whether a miRNA can successfully repress its target—or instead be subjected to decay—remain an important and unanswered question.

New modes of miRNA regulation

In the following studies, we describe a series of novel regulatory mechanisms that modulate the expression of miRNAs that play notable roles in cancer. Chapter Two describes our development of a genomic pipeline to identify potential relationships between transcription factors and miRNAs. Using this approach, we have found a particularly intriguing regulatory network for the miR-200 family of miRNAs. The miR-200 family consists of five miRNAs (miR-200a, -200b, -200c, -429, and -141), which are co-expressed from two genetic loci (8,36). These five paralogous miRNAs share very similar sequences, particularly at the 5' seed region of the miRNA that is most important in determining their repertoire of mRNA targets. The miR-200 family is required for normal development and maintains the epithelial state of cells via repression of the mesenchymal transcription factors Zeb1 and Zeb2 (37). The miR-200 family is also dysregulated in many cancer types, although the direction varies across tissues. In prostate, bile duct, and ovarian cancers, miR-200 levels are high; while in advanced breast carcinomas, gastric, and renal carcinomas, miR-200 levels are low compared to normal tissue (8,38-40). The disparate expression patterns of the miR-200 family may result from differing cellular requirements for tumor establishment versus metastasis (41). Thus, low levels of miR-200 may be needed for cancer cell migration and invasion, while high miR-200 levels may enable the emergence of a solid tumor (42). Recent *in vivo* work has shown that forced expression of miR-200 in lung cancer cells implanted into mice prevents EMT and metastasis, which raises the possibility that therapeutically modulating miR-200 levels may decrease tumor aggressiveness (43). Together, these

studies have suggested that the miR-200 family may serve as a master regulator of both tumor formation and metastasis.

Despite the importance of the miR-200 family in cancer, the mechanisms of its regulation are not understood. In an earlier study, we found that both primary and mature miR-200 family members are increased in ovarian tumors compared to normal surface epithelium (44), which demonstrates that altered transcription may drive the aberrant expression patterns. Zeb1 and Zeb2 can repress transcription of the miR-200 family, forming a double-negative feedback loop (45). However, no positive regulators of the family have been identified. Through a combination of computational and experimental methods, we describe the activation of miR-200 miRNAs by the p63 and p73 transcription factors in ovarian carcinoma. Our work reveals a new role for these less understood members of the p53 family. These studies also demonstrate our successful application of a genomic pipeline by which miRNA and transcription factor associations can be identified in different systems featuring differential miRNA expression.

Chapter Three describes an additional new form of transcriptional regulation of miRNAs. We have investigated the effects of the cellular microenvironment on miRNA expression patterns. While many miRNAs are ubiquitously expressed, other miRNAs are specifically induced in response to various stresses. We have examined the hypoxia-regulated miRNA, miR-210, which is dramatically induced by the HIF-1α transcription factor in response to low oxygen conditions (46,47). MicroRNA-210 represses multiple regulators of proliferation and survival and promotes the successful adaptation of a cell to a hypoxic environment (48). Surprisingly, we have found this miRNA is also expressed at high levels in normoxic conditions across a panel of cancer cell lines. We demonstrate

that the induction of miR-210 in cell culture occurs in response to increased cell density. Density-dependent changes in mature miRNA expression have only been described in a single report, which attributed the effect to increased Drosha processing (49). Here, we describe an independent mechanism of regulation: We found that increased cell density specifically drives the transcriptional activation of miR-210 by HIF-1α, which serves as a highly sensitive sensor of the cellular microenvironment independently of global hypoxic conditions. This new mechanism of activation may have clinical significance, as miR-210 overexpression is associated with poor prognosis in several cancer types (50,51). While the functional effects of this induction are still under investigation, the potent transcriptional activation of miR-210 underscores the importance of localized changes in a cell's environment in affecting miRNA expression patterns.

In addition to examining the transcriptional activation of miRNAs, we sought to understand novel forms of regulation for mature miRNAs. Next-generation sequencing has facilitated an era of rapid miRNA discovery in recent years. While the majority of canonical miRNAs have likely been identified in humans, improved throughput and accuracy of second generation sequencing pipelines has enabled the identification of a number of new miRNA variants (52,53). In Chapter 4, we demonstrate that the miRNA transcriptome in humans displays much more variation than was previously hypothesized. In addition to the canonical miRNA sequences, we have identified a large number of miRNA "isomiRs" that display sequence heterogeneity at the 3' end of the miRNA. We have investigated both the origin of 3' non-templated nucleotide additions and their potential functional effects on miRNA stability. In Chapter 4, we demonstrate that 3' additions are specifically regulated by multiple members of the nucleotidyl

transferase family. In Chapter 5, we discuss the changes in abundance that may result from these modifications. We identified specific additions, such as 3' U, which are associated with reduced miRNA expression. We also found several nucleotidyl transferases, including TUT1, PAPD4, and MTPAP, which act as broad regulators of miRNA abundance. While understanding of the mechanisms of miRNA degradation is still at its nascent stage, this study provides a foundation for the identification of specific signals for miRNA decay versus stabilization.

Taken together, these studies reveal multiple new forms of regulation for miRNAs. MicroRNAs have been demonstrated to play influential roles in human physiology. These ubiquitous regulatory molecules provide an important layer of post-transcriptional control of gene expression. While the therapeutic potential of miRNAs is still in the early phases of investigation, discovering mechanisms of miRNA regulation will facilitate understanding of how miRNAs control the proliferation and survival of normal and malignant cells.

References

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

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

3. Rosa, A. and Brivanlou, A.H. (2009) microRNAs in early vertebrate development. *Cell Cycle*, **8**.

4. 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.

5. Fazi, F., Rosa, A., Fatica, A., Gelmetti, V., De Marchis, M.L., Nervi, C. and Bozzoni, I. (2005) A minicircuitry comprised of microRNA-223 and transcription factors NFI-A and C/EBPalpha regulates human granulopoiesis. *Cell*, **123**, 819-831.

6. Park, S.-M., Gaur, A.B., Lengyel, E. and Peter, M.E. (2008) The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev*, **22**, 894-907.

7. Korpal, M., Lee, E.S., Hu, G. and Kang, Y. (2008) The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *J Biol Chem*, **283**, 14910-14914.

8. Gregory, P.A., Bert, A.G., Paterson, E.L., Barry, S.C., Tsykin, A., Farshid, G., Vadas, M.A., Khew-Goodall, Y. and Goodall, G.J. (2008) The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol*, **10**, 593-601.

9. Ebert, M.S. and Sharp, P.A. (2012) Roles for microRNAs in conferring robustness to biological processes. *Cell*, **149**, 515-524.

10. Calin, G.A. and Croce, C.M. (2006) MicroRNA signatures in human cancers. *Nat Rev Cancer*, **6**, 857-866.

11. Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A. *et al.* (2005) MicroRNA expression profiles classify human cancers. *Nature*, **435**, 834-838.

12. Mitchell, P.S., Parkin, R.K., Kroh, E.M., Fritz, B.R., Wyman, S.K., Pogosova-Agadjanyan, E.L., Peterson, A., Noteboom, J., O'Briant, K.C., Allen, A. *et al.*

(2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci USA*, **105**, 10513-10518.

13. Kumar, M.S., Pester, R.E., Chen, C.Y., Lane, K., Chin, C., Lu, J., Kirsch, D.G., Golub, T.R. and Jacks, T. (2009) Dicer1 functions as a haploinsufficient tumor suppressor. *Genes Dev*, **23**, 2700-2704.

14. Merritt, W.M., Lin, Y.G., Han, L.Y., Kamat, A.A., Spannuth, W.A., Schmandt, R., Urbauer, D., Pennacchio, L.A., Cheng, J.-F., Nick, A.M. *et al.* (2008) Dicer, Drosha, and outcomes in patients with ovarian cancer. *N Engl J Med*, **359**, 2641-2650.

15. Olive, V., Bennett, M.J., Walker, J.C., Ma, C., Jiang, I., Cordon-Cardo, C., Li, Q.J., Lowe, S.W., Hannon, G.J. and He, L. (2009) miR-19 is a key oncogenic component of mir-17-92. *Genes Dev*, **23**, 2839-2849.

16. Mu, P., Han, Y.C., Betel, D., Yao, E., Squatrito, M., Ogrodowski, P., de Stanchina, E., D'Andrea, A., Sander, C. and Ventura, A. (2009) Genetic dissection of the miR-17~92 cluster of microRNAs in Myc-induced B-cell lymphomas. *Genes Dev*, **23**, 2806-2811.

17. Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D. and Slack, F.J. (2005) RAS is regulated by the let-7 microRNA family. *Cell*, **120**, 635-647.

18. Sampson, V.B., Rong, N.H., Han, J., Yang, Q., Aris, V., Soteropoulos, P., Petrelli, N.J., Dunn, S.P. and Krueger, L.J. (2007) MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. *Cancer Research*, **67**, 9762-9770.

19. Ma, L., Teruya-Feldstein, J. and Weinberg, R.A. (2007) Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature*, **449**, 682-688.

20. He, L., He, X., Lim, L.P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D. *et al.* (2007) A microRNA component of the p53 tumour suppressor network. *Nature*, **447**, 1130-1134.

21. Chang, T.C., Yu, D., Lee, Y.S., Wentzel, E.A., Arking, D.E., West, K.M., Dang, C.V., Thomas-Tikhonenko, A. and Mendell, J.T. (2008) Widespread microRNA repression by Myc contributes to tumorigenesis. *Nat Genet*, **40**, 43-50.

22. O'Donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V. and Mendell, J.T. (2005) c-Myc-regulated microRNAs modulate E2F1 expression. *Nature*, **435**, 839-843.

23. Wang, J., Lu, M., Qiu, C. and Cui, Q. (2010) TransmiR: a transcription factor-microRNA regulation database. *Nucleic acids research*, **38**, D119-122.

24. Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K., Mihara, M., Naitou, M., Endoh, H., Nakamura, T. *et al.* (2007) DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs. *Nat Cell Biol*, **9**, 604-611.

25. 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.

26. 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.

27. 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.

28. 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. *Nat Struct Mol Biol*, **16**, 1021-1025.

29. Bail, S., Swerdel, M., Liu, H., Jiao, X., Goff, L.A., Hart, R.P. and Kiledjian, M. (2010) Differential regulation of microRNA stability. *RNA*, **16**, 1032-1039.

30. Krol, J., Busskamp, V., Markiewicz, I., Stadler, M.B., Ribi, S., Richter, J., Duebel, J., Bicker, S., Fehling, H.J., Schübeler, D. *et al.* (2010) Characterizing light-regulated retinal microRNAs reveals rapid turnover as a common property of neuronal microRNAs. *Cell*, **141**, 618-631.

31. Chatterjee, S. and Großhans, H. (2009) Active turnover modulates mature microRNA activity in Caenorhabditis elegans. *Nature*, **461**, 546-549.

32. Ramachandran, V. and Chen, X. (2008) Degradation of microRNAs by a family of exoribonucleases in Arabidopsis. *Science*, **321**, 1490-1492.

33. Ibrahim, F., Rymarquis, L.A., Kim, E.-J., Becker, J., Balassa, E., Green, P.J. and Cerutti, H. (2010) Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. *Proc Natl Acad Sci USA*.

34. Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B. and Chen, X. (2012) The Arabidopsis Nucleotidyl Transferase HESO1 Uridylates Unmethylated Small RNAs to Trigger Their Degradation. *Current biology : CB*, **22**, 689-694.

35. Ren, G., Chen, X. and Yu, B. (2012) Uridylation of miRNAs by HEN1 SUPPRESSOR1 in Arabidopsis. *Current biology : CB*, **22**, 695-700.

36. Park, S.M., Gaur, A.B., Lengyel, E. and Peter, M.E. (2008) The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev*, **22**, 894-907.

37. Christoffersen, N., Silahtaroglu, A., Orom, U., Kauppinen, S. and Lund, A. (2007) miR-200b mediates post-transcriptional repression of ZFHX1B. *RNA*, **13**, 1172-1178.

38. Nakada, C., Matsuura, K., Tsukamoto, Y., Tanigawa, M., Yoshimoto, T., Narimatsu, T., Nguyen, L.T., Hijiya, N., Uchida, T., Sato, F. *et al.* (2008) Genome-wide microRNA expression profiling in renal cell carcinoma: significant down-regulation of miR-141 and miR-200c. *J Pathol*, **216**, 418-427.

39. Iorio, M.V., Visone, R., Di Leva, G., Donati, V., Petrocca, F., Casalini, P., Taccioli, C., Volinia, S., Liu, C.G., Alder, H. *et al.* (2007) MicroRNA signatures in human ovarian cancer. *Cancer Research*, **67**, 8699-8707.

40. Du, Y., Xu, Y., Ding, L., Yao, H., Yu, H., Zhou, T. and Si, J. (2009) Down-regulation of miR-141 in gastric cancer and its involvement in cell growth. *J Gastroenterol*, **44**, 556-561.

41. Olson, P., Lu, J., Zhang, H., Shai, A., Chun, M.G., Wang, Y., Libutti, S.K., Nakakura, E.K., Golub, T.R. and Hanahan, D. (2009) MicroRNA dynamics in the stages of tumorigenesis correlate with hallmark capabilities of cancer. *Genes Dev*, **23**, 2152-2165.

42. Gregory, P.A., Bracken, C.P., Bert, A.G. and Goodall, G.J. (2008) MicroRNAs as regulators of epithelial-mesenchymal transition. *Cell Cycle*, **7**, 3112-3118.

43. Gibbons, D.L., Lin, W., Creighton, C.J., Rizvi, Z.H., Gregory, P.A., Goodall, G.J., Thilaganathan, N., Du, L., Zhang, Y., Pertsemlidis, A. *et al.* (2009) Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression. *Genes Dev*, **23**, 2140-2151.

44. Bendoraite, A., Knouf, E.C., Garg, K.S., Parkin, R.K., Kroh, E.M., O'Briant, K.C., Ventura, A.P., Godwin, A.K., Karlan, B.Y., Drescher, C.W. *et al.* (2010) Regulation of miR-200 family microRNAs and ZEB transcription factors in ovarian cancer: evidence supporting a mesothelial-to-epithelial transition. *Gynecologic Oncology*, **116**, 117-125.

45. Bracken, C.P., Gregory, P.A., Kolesnikoff, N., Bert, A.G., Wang, J., Shannon, M.F. and Goodall, G.J. (2008) A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Research*, **68**, 7846-7854.

46. Huang, X., Ding, L., Bennewith, K.L., Tong, R.T., Welford, S.M., Ang, K.K., Story, M., Le, Q.-T. and Giaccia, A.J. (2009) Hypoxia-inducible mir-210

regulates normoxic gene expression involved in tumor initiation. *Mol Cell*, **35**, 856-867.

47. Fasanaro, P., D'Alessandra, Y., Di Stefano, V., Melchionna, R., Romani, S., Pompilio, G., Capogrossi, M.C. and Martelli, F. (2008) MicroRNA-210 modulates endothelial cell response to hypoxia and inhibits the receptor tyrosine kinase ligand Ephrin-A3. *J Biol Chem*, **283**, 15878-15883.

48. Chan, S.Y. and Loscalzo, J. (2010) MicroRNA-210: A unique and pleiotropic hypoxamir. *Cell cycle (Georgetown, Tex)*, **9**.

49. Hwang, H.W., Wentzel, E.A. and Mendell, J.T. (2009) Cell-cell contact globally activates microRNA biogenesis. *Proc Natl Acad Sci USA*, **106**, 7016-7021.

50. Jung, E.J., Santarpia, L., Kim, J., Esteva, F.J., Moretti, E., Buzdar, A.U., Di Leo, A., Le, X.F., Bast, R.C., Jr., Park, S.T. *et al.* (2012) Plasma microRNA 210 levels correlate with sensitivity to trastuzumab and tumor presence in breast cancer patients. *Cancer*, **118**, 2603-2614.

51. Volinia, S., Galasso, M., Sana, M.E., Wise, T.F., Palatini, J., Huebner, K. and Croce, C.M. (2012) Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. *Proc Natl Acad Sci USA*, **109**, 3024-3029.

52. Cummins, J.M., He, Y., Leary, R.J., Pagliarini, R., Diaz, L.A., Jr., Sjoblom, T., Barad, O., Bentwich, Z., Szafranska, A.E., Labourier, E. *et al.* (2006) The colorectal microRNAome. *Proc Natl Acad Sci U S A*, **103**, 3687-3692.

53. Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H. and Bartel, D.P. (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell*, **127**, 1193-1207.

Chapter 2

An integrative genomic approach identifies p73 and p63 as activators of miR-200 microRNA family transcription

The following text has been modified from a manuscript published in Nucleic Acids Research in 2012. Figure numbers have been updated to reflect the formatting of the dissertation. Supplemental figures and tables are available with open access at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3258134/bin/supp_40_2_499__index.html

We would also like to acknowledge the contributions of our co-authors:

Knouf EC*, Garg KS*, Arroyo JD*, Correa Y, Sarkar D, Parkin RK, Wurz K, O'Briant KC, Godwin AK, Urban ND, Ruzzo WL, Gentleman R, Dresher CW, Swisher EM, Tewari M. An integrative genomic approach identifies p73 and p63 as activators of miR-200 microRNA family transcription. *Nucleic Acids Research* **40**, 499-510 (2012).

Abstract

Although microRNAs (miRNAs) are important regulators of gene expression, the transcriptional regulation of miRNAs themselves is not well understood. We employed an integrative computational pipeline to dissect the transcription factors responsible for altered miRNA expression in ovarian carcinoma. Using experimental data and computational predictions to define miRNA promoters across the human genome, we identified transcription factors with binding sites significantly over-represented among miRNA genes overexpressed in ovarian carcinoma. This pipeline nominated transcription factors of the p53/p63/p73 family as candidate drivers of miRNA overexpression. Analysis of data from an independent set of 253 ovarian carcinomas in The Cancer Genome Atlas (TCGA) showed that p73 and p63 expression is significantly correlated with expression of miRNAs whose promoters contain p53/p63/p73 family binding sites. In experimental validation of specific miRNAs predicted by the analysis to be regulated by p73 and p63, we found that p53/p63/p73 family binding sites modulate promoter activity of miRNAs of the miR-200 family, which are known regulators of cancer stem cells and epithelial-mesenchymal transitions. Furthermore, in chromatin immunoprecipitation studies both p73 and p63 directly associated with the miR-200b/a/429 promoter. This study delineates an integrative approach that can be applied to discover transcriptional regulatory mechanisms in other biological settings where analogous genomic data are available.

Introduction

Regulation of gene expression at the post-transcriptional level is governed in part by microRNAs (miRNAs), which are ~22 nucleotide non-protein-encoding RNAs that modulate the stability and/or translation of messenger RNAs (mRNAs) via partially complementary base-pairing interactions (1). Most microRNAs are transcribed by RNA polymerase II (2), and miRNA expression can be regulated by transcription factor binding sites present in their promoters (3-7). However, for the majority of miRNAs, promoters have not been defined and the transcription factor binding sites upstream of these miRNA loci have not been experimentally tested.

Dysregulation of miRNA expression is common in human disease and contributes to pathology, since miRNAs regulate significant disease-relevant processes such as cell division, differentiation, and apoptosis (8,9). In addition, in certain cancer contexts, the pattern of miRNA expression captures important features of the developmental origin of cancers (10) and may predict the course of disease (11). However, the mechanisms underlying miRNA dysregulation are not clear, in part because the transcriptional regulation of most miRNAs is not well characterized.

In this study, we implemented an integrative computational approach to dissect the transcriptional regulation of miRNAs. We focused on the dysregulation of miRNAs in ovarian carcinoma of the serous histologic sub-type, which has a high mortality and accounts for approximately two-thirds of ovarian carcinomas. Although a subset of miRNAs dysregulated in ovarian carcinomas is associated with changes in genomic copy number and epigenetic modifications, for many miRNAs additional, unknown

mechanisms appear to contribute to the reprogramming of miRNA expression (12,13). We therefore sought to discover the transcription factors that may drive the dysregulation of miRNAs in ovarian carcinoma. We implemented a computational pipeline to annotate miRNA transcription start sites (TSS) and putative promoter regions, and then to identify the transcription factors with binding sites enriched in the promoters of overexpressed miRNAs in ovarian carcinoma. This approach yields putative regulatory interactions between transcription factors and miRNA promoters for subsequent experimental validation.

We report here that the best candidate driver of miRNA overexpression in ovarian carcinoma is the p53/p63/p73 family of transcription factors. Although p53 has been shown to transactivate several miRNAs, including the miR-34 family (14-17), the transcriptional regulation of miRNA genes by p73 and p63 has not been well-described. Further analysis using data from The Cancer Genome Atlas suggested that, in ovarian carcinoma, p73 and p63 are primarily responsible for the altered expression of miRNAs with p53 family binding sites. We experimentally validated our approach by confirming that p73 and p63 directly regulate transcription of the miR-200 family, a novel target predicted by our analysis that is an important regulator of epithelial-to-mesenchymal transitions and of the cancer stem cell phenotype (18-22).

This study illustrates how an integrative computational analysis can identify new regulatory interactions between transcription factors and miRNAs. We also provide a resource by defining putative miRNA promoters and associating transcription factor binding sites with these miRNA promoters on a genome-wide scale, and we discuss how our approach is broadly applicable to dissect transcription factor-miRNA regulatory

networks in other systems where miRNA expression data from distinct physiologic states are available.

Methods

Clinical materials

Normal primary human ovarian surface epithelial (HOSE) cell and serous ovarian carcinomas from which RNA was analyzed in this study have been previously described (23), and are described in Supplemental Information. HOSE specimens were obtained under a protocol that was approved by the Research Review Committee and Internal Review Board at the Fox Chase Cancer Center. All ovarian carcinoma tissue specimens were collected from the Pacific Ovarian Cancer Research Consortium Repository under a protocol approved by the Fred Hutchinson Cancer Research Center Institutional Review Board.

Analysis of miRNA microarray data generated in the current study

We analyzed miRNA microarray data generated on a locked nucleic acid (LNA) probe-based platform designed to profile 480 human miRNAs (Exiqon, Inc.) (23). Array quality control and data normalization was performed as previously described (23). We defined expressed miRNAs as the miRNAs expressed in at least two of the four normal or at least four of the 16 serous ovarian carcinomas. Nonspecific filtering was performed as previously described (23), and genes that showed low variability across samples were removed from the analysis, which yielded 181 genes for downstream analyses. Differentially expressed miRNAs were identified by considering miRNAs to be over- or

underexpressed in carcinoma compared to normal if $\log_2$ fold change >1.0 and adjusted p-value <0.01.

Identification of transcription factors commonly expressed in serous ovarian carcinoma

We analyzed data from Hendrix et al. (24), where Affymetrix GeneChip Human Genome U133 Array Set HG-U133A microarrays were used to profile gene expression across a panel of 4 normal ovaries and 41 serous ovarian carcinoma tissues. The carcinoma tissues comprised the following stages: stage 1A (n=1), stage 1C (n=3), stage 2 (n=2), stage 2C (n=2), stage 3 (n=2), stage 3B (n=1), 3C (n=25), 3D (n=1) and 4 (n=5) and the following grades: well-differentiated (n=2), moderately differentiated (n=5), poorly differentiated (n=18), grade information not available (n=16). All samples were selected to contain at least 70% malignant epithelial cells. We considered genes to be expressed if they showed expression (as determined by Affymetrix PMA (Present/Marginal/Absent) calls) in at least two out of four normal samples or eight out of 41 carcinoma samples. We identified 83 transcription factors expressed in normal ovary and/or serous ovarian carcinoma tissue specimens by intersecting expressed genes with all transcription factors in tfbsConsFactors files downloaded from the UCSC genome browser (25,26). Out of these 83 transcription factors, 79 corresponded to binding sites found upstream of miRNAs overexpressed in ovarian carcinoma.

Annotation of putative miRNA promoters and their transcription factor binding sites

miRNA coordinates were downloaded from miRBase (version 10) (27) and used to map miRNAs to their genomic locations as described in Supplemental Information. To assign TSS for each miRNA locus, we used RefSeq, AceView, ESTs, and Eponine predictions downloaded from the UCSC genome browser (hg 18 version of the genome assembly) (26). If both 5' and 3' ESTs were available from the same clone and formed a transcript containing the miRNA, the miRNA was considered expressed by this transcript and its TSS was the 5' end of the EST.

We examined the putative promoters of each miRNA/miRNA locus for conserved transcription factor binding sites defined in the UCSC files tfbsConsSites and tfbsConsFactors (26), as described in Supplemental Information. The conservation of binding sites is based on human/mouse/rat multiple alignments and the score is computed using the Transfac Matrix Database (v7.0) (28).

Enrichment of transcription factor binding sites in promoters of overexpressed miRNA loci

We performed a Fisher's exact test for each of the 79 expressed transcription factors to compare the number of transcription factor binding sites in the promoters of overexpressed miRNA loci versus control miRNA loci. To assess the false discovery rate, predicted q-values were estimated from p-values (derived from Fisher's exact test) by using the *qvalue* package in R (29).

Correlations between expression of transcription factors and expression of miRNAs in TCGA dataset

We analyzed data from 253 serous ovarian carcinoma specimens from TCGA. Serous ovarian carcinoma tissue specimens in TCGA are generally from patients with advanced stage and high-grade disease. Clinical data on stage and grade was available on 217 of the 253 specimens we studied, and confirmed that the vast majority of these were high-grade (202 were poorly-differentiated and 13 were moderately-differentiated) and advanced stage (Stage IIB (n=2), Stage IIC (n=4), Stage IIIA (n=2), Stage IIIB (n=9), Stage IIIC (n=161) and Stage IV (n=39)). Data on percentage malignant epithelium was available for 202 of the 253 specimens we studied: > 75% malignant epithelial cells (n=150 specimens), ≥ 50% and < 75% (n=43 specimens), < 50% (n=9 specimens). To derive correlations between transcription factor and miRNA expression, we retrieved Agilent gene expression microarray and miRNA microarray data from the TCGA for 253 serous ovarian carcinomas, for which both gene expression (run on AgilentG4502A_07_3) and miRNA expression profiling data were available (http://cancergenome.nih.gov/index.asp). Agilent Level 1 microarray expression data were used for both miRNA and mRNA analyses. We used 447 miRNAs for calculating correlations; all of these miRNAs had at least one transcription factor binding site associated with their promoters as determined by our method. All TCGA Level-1 files used to calculate correlations are listed in Table S1. Correlation between the expression of the miRNAs and transcription factors was assessed by calculating Pearson correlation coefficients. Bioconductor packages ‘limma’ and ‘ltm’ were used to process the raw signal data and to calculate p-values for correlation between the expression of miRNAs and transcription factors respectively (30,31).

Sequencing of *TP53*

All coding and flanking regulatory regions for exons 4-10 of *TP53* were sequenced as previously described (32).

Cell culture

The 2008 ovarian carcinoma cell line was maintained in RPMI (Invitrogen) with 10% fetal bovine serum (Atlanta Biologicals) under 5% CO_2. All transfections were performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions.

Chromatin immunoprecipitation

Chromatin immunoprecipitation was performed on 2008 ovarian carcinoma cells as previously described (33). Sheared chromatin was incubated with 12 μg of antibody overnight at 4°C and then incubated with 40 μL prewashed Protein G agarose beads (Millipore). Antibodies used for immunoprecipitation included 4A4 (Santa Cruz) for p63, 259A (Imgenex) for p73, and normal mouse IgG (Santa Cruz) to control for nonspecific binding. qPCR was performed as described in Supplemental Information using the primers listed in Table S2. Triplicate immunoprecipitations were performed for each antibody, using independent batches of chromatin.

Promoter assays

Firefly luciferase promoter activity reporter plasmids were generated by PCR amplification of the region of interest from human genomic DNA (BioLine) followed by subcloning into the MluI and BglII sites of the pGL3-Enhancer vector (Promega), as described in Supplemental Information. Primers used are in Table S3. To assay promoter activity, 2008 cells were plated in 96-well plates at $8x10^3$ cells/well 24 h prior to

transfection. The cells were co-transfected with 100 ng of each firefly luciferase promoter reporter vector, or pGL3-Enhancer as a control, and 25 ng of pRL-TK (Promega). Twenty-four hours after transfection, firefly and *Renilla* luciferase activity was measured with the Dual-Luciferase Reporter Assay System (Promega). Promoter activity was calculated as the ratio of firefly to *Renilla* luciferase signal.

Results

Overview of computational pipeline for the identification of transcription factors regulating miRNA expression

We implemented a computational pipeline to study the transcriptional regulation of miRNA expression (**Figure 2.1**). Briefly, the pipeline starts with miRNA differential expression data as input and classifies miRNAs into two categories: miRNAs that are overexpressed in the samples of interest and control miRNAs that are expressed but not overexpressed. Putative promoters of both the overexpressed and control miRNAs are then defined via genomic mapping and annotation of transcription start sites using a series of criteria described below. After identifying the transcription factors expressed in the samples based on mRNA expression data, the number of predicted binding sites for each expressed transcription factor in the promoters of both the overexpressed miRNA and control miRNA loci is determined. The output of the pipeline is transcription factors with binding sites overrepresented in the promoters of the overexpressed miRNAs, and these regulatory predictions can then be experimentally validated.

Identification of a set of miRNAs commonly differentially expressed in serous ovarian carcinoma

As input to our pipeline, we defined miRNAs that are differentially expressed in serous ovarian carcinomas compared to normal specimens. To avoid bias, we used multiple studies to identify miRNAs that are commonly altered in ovarian carcinoma. We previously profiled miRNA expression in primary serous ovarian carcinomas from 16 individuals and in cultured normal human ovarian surface epithelial (HOSE) cells from 4 other individuals using miRNA microarrays (23). We analyzed that dataset here to identify miRNAs significantly dysregulated in serous ovarian carcinoma as compared to normal. We found 43 miRNAs to be overexpressed and 20 to be underexpressed in ovarian carcinoma ($\log_2$ fold change >1.0 and adjusted p-value <0.01) (**Table S4**). We then compared the results of this analysis and six other ovarian carcinoma miRNA profiling studies (13,34-38). We found 30 miRNAs to be significantly overexpressed in two or more of the entire set of seven studies (**Table S5**), which should provide a more reliable list of miRNAs commonly overexpressed in serous ovarian carcinomas. We also identified a set of 18 commonly underexpressed miRNAs (**Table S6**), although these did not yield statistically significant over-representation of any specific transcription factor binding sites in subsequent analysis and are therefore not discussed further.

In order to define a set of control miRNAs for comparison, we used data from our microarray analysis of ovarian carcinoma and normal specimens to identify 238 expressed miRNAs, which includes the 30 commonly overexpressed miRNAs defined above. We then converted this list of expressed miRNAs to a list of miRNA loci, where each locus corresponds to a single promoter, which may drive expression of multiple

miRNAs. Specifically, we mapped each miRNA to its genomic position, and miRNAs that are encoded at multiple locations were assigned to each locus separately. Because many miRNAs are found in clusters in single polycistronic transcripts (39,40), we annotated all miRNAs within a genomic cluster as a single locus. After accounting for the expansion of the number of miRNA locations as a result of miRNAs encoded at multiple loci, and contraction of this number because of miRNAs in clusters with shared promoters, the 238 miRNAs corresponded to 151 loci representing a miRNA or miRNA cluster that is expressed in serous ovarian carcinomas and/or normal samples (**Table S7**). We divided the 151 loci into two classes: those that encode overexpressed miRNAs and those that do not. If a single miRNA within a cluster was overexpressed, the entire miRNA locus was classified into the overexpressed class for the purpose of this analysis. From this approach, we annotated 30 loci (miRNAs or miRNA clusters) that were overexpressed in serous ovarian carcinomas, and we used the remaining 121 loci as the control set in all further analyses (**Table S8**).

Annotation of miRNA transcription start sites and identification of candidate transcription factors controlling miRNA expression

To define putative promoter regions for subsequent analysis, we first annotated the transcriptional start site (TSS) of each overexpressed or control miRNA locus (**Table S7**). We found 43% of miRNAs to be located within and in the same orientation as a RefSeq gene (41). The TSS for these miRNAs was assumed to be the same as for the host gene, as it has been shown that miRNAs within host genes are generally co-transcribed from a shared promoter (42,43). For miRNA genes that did not match to RefSeq, we used

AceView, which provides comprehensive transcriptional evidence from full length cDNAs and ESTs (44), to identify the TSS of 38% of the miRNAs. We next used predictions by Eponine (45) to define the TSS of 7% of the miRNAs and EST clones to locate the TSS of 3% of the miRNAs. For the remaining 9% of miRNAs whose TSS could not be found by the above methods, the position 500 bp upstream of the miRNA was taken as the TSS. In the case of miRNAs that lie in genomic clusters, the TSS of the most 5' miRNA was assigned to all miRNAs in the cluster, because such miRNAs are expressed as a single primary transcript from a shared promoter (40).

To ensure our analysis would involve only transcription factors commonly expressed in serous carcinoma and/or normal ovarian tissue, we used mRNA microarray data from serous ovarian carcinomas and from normal ovarian tissues available in the Gene Expression Omnibus (GSE6008) (24) to identify 79 transcription factors expressed in normal ovary and/or serous ovarian carcinomas. For each of the 151 miRNA (or miRNA cluster) loci identified above, we examined a region extending from 5000 bp upstream to 500 bp downstream of the TSS for predicted transcription factor binding sites (**Table S8**). If a miRNA was less than 500 bp downstream from the TSS, we instead used the distance between the miRNA and TSS for the downstream boundary of the region of interest. We only included binding sites conserved between human, mouse and rat, as annotated by the UCSC genome browser (26), for the 79 expressed transcription factors.

To determine which transcription factors are likely to drive miRNA overexpression in ovarian carcinoma, we performed a Fisher's exact test for each of the 79 transcription factors to compare the number of binding sites in the putative promoters of the 30 overexpressed miRNA loci with the number of sites for the 121 control miRNA

loci (**Table S9**). We found an over-representation of binding sites for the tumor protein 53 (p53) family, early growth response 2 (EGR2), and SP1 in the promoters of miRNAs overexpressed in serous ovarian carcinoma (q-value (29) <0.05, **Table 2.1**). The most significant q-value, 0.005, was found for the p53 family, which consists of three proteins (i.e., p53, p63, and p73) that recognize the same consensus binding site (46,47). We identified 15 p53 family binding sites among the 30 overexpressed miRNA promoter loci, versus 17 such sites in the 121 control miRNA promoter loci (**Tables 2.1 and S8**). Thus, this analysis suggests that the three members of the p53 family, as well as EGR2 and SP1, may be drivers of miRNA dysregulation in serous ovarian carcinoma.

Correlation between expression of transcription factors and expression of miRNAs in The Cancer Genome Atlas (TCGA) data

To determine if expression of the candidate miRNA transcription factors we identified correlates with expression of miRNAs with binding sites for those transcription factors, next we employed a larger, independent dataset from The Cancer Genome Atlas (TCGA), which provides comprehensive genomic profiling data from high-grade serous ovarian carcinomas. We selected samples from TCGA for which both miRNA and mRNA data were available for the same carcinoma tissue specimen. There were 253 such samples available. We calculated Pearson correlations between the expression of each of the five transcription factors identified above and the expression of each of 447 miRNAs in the 253 TCGA samples. We then compared correlation coefficients between the miRNAs whose promoters contain binding sites for a given transcription factor and the remaining miRNAs lacking binding sites. We found no significant difference in the

correlation of *SP1* and miRNAs with or without SP1 binding sites (data not shown). However, the remaining four transcription factors (i.e., *TP53*, *TP63*, *TP73*, and *EGR2*) showed significantly higher expression correlations with miRNAs bearing their respective binding sites versus those without (p-value <0.01, *t*-test) (**Figure 2.2**). This analysis suggests that our pipeline yields functional regulators of miRNA expression. Interestingly, we found significant negative correlations between the expression of both *EGR2* probes and the expression of miRNAs with EGR2 binding sites, which suggests that EGR2 may serve as a transcriptional repressor (**Figure 2.2D**).

All three members of the p53 family featured probes that showed positive, significantly higher correlations with miRNAs bearing p53 family binding sites than those without (**Figure 2.2A-C**). Three out of seven *TP63* probes and one out of three *TP53* probes showed lower correlations that were not statistically significant (data not shown), which may be related to sequence-specific variation in probe performance. Of the three members of the p53 family, *TP73* expression showed the strongest positive correlations with miRNAs bearing p53 family binding sites (**Figure 2.2C**). Taken together, these results suggest that members of the p53 family, and in particular *TP73*, may drive the overexpression of a panel of miRNAs in ovarian carcinoma.

Overexpression of miRNAs with p53 binding sites in their promoters is not associated with the mutational status of *TP53* in serous ovarian carcinomas

While expression of p53 mRNA appeared to correlate with miRNA expression, regulation of p53 occurs primarily at the protein level. Therefore, we sought to determine whether changes in the activity of the p53 protein may affect the expression of miRNAs

bearing p53 family binding sites. Mutation is the leading cause of p53 dysfunction in serous ovarian carcinoma, as 60-70% of serous ovarian cancers are known to have *TP53* mutations (48,49), which in general ablate the ability of p53 to transactivate its target genes. However, because mutations to *TP53* can result in a dominant negative protein that alters the activity of the other p53 family members, we examined whether *TP53* mutational status affects expression of miRNAs overexpressed in ovarian carcinoma. Somatic mutation data for ovarian carcinomas in TCGA were not available. We therefore determined the p53 mutation status of the 16 ovarian carcinoma specimens that we previously used for identifying differentially expressed miRNAs. We sequenced all coding and flanking regulatory regions for exons 4-10 of *TP53,* which include mutation hotspots and also coincide with the most highly conserved region of the gene (50). We found non-synonymous *TP53* mutations accompanied by loss of the wild-type allele in nine out of our 16 ovarian carcinoma specimens (**Table S10**). Comparing the expression of miRNAs in the *TP53* mutant versus wild-type specimens demonstrated no significant difference in the expression of the 17 overexpressed miRNAs that are encoded within the 15 overexpressed loci with p53 binding sites (**Figure S1**). While these results do not exclude p53 as a possible driver of differential miRNA expression, this analysis suggests that p53 is not the primary regulator of these miRNAs in ovarian carcinoma.

p73 and p63 as putative positive regulators of miRNA expression in ovarian cancer

Our miRNA expression correlation analyses from TCGA data suggested that p73, and potentially p63, may drive the overexpression of miRNAs in ovarian carcinoma (**Figures 2.2B and 2.2C**). To identify specific miRNAs that may be regulated by p73 or

p63 for further study, we used the p-values from the Pearson correlations to rank the miRNAs that were overexpressed in ovarian carcinoma and contained p53 family binding sites. We found multiple miRNAs with significant positive correlation with *TP73* or *TP63* expression (**Table 2.2**). A complete list of miRNAs showing positive correlations (Pearson correlation coefficient >0.25 and p-value <0.05) with all p73 and p63 probes is provided in **Tables S11** and **S12**. To assess the biological relevance of these correlation values, we also determined the correlations of 88 host gene-intronic miRNA pairs using the same dataset of 253 TCGA samples. Intronic miRNAs (i.e., lying in host genes on the same strand) are generally co-transcribed with their host gene and therefore show strong correlations in expression with their host gene. Among miRNA-host gene pairs analyzed by the same criteria (correlation coefficients >0.25 and p-values <0.05), host gene-intronic miRNA pairs showed an average correlation of 0.37 (±0.16). The correlations between the expression of p73 and p63 and miRNAs in **Table 2.2** range between 0.25-0.39, and thus reflect a similar magnitude of correlation coefficients as observed with the host gene-intronic miRNA correlations.

p73 and p63 directly regulate miR-200 transcription

To experimentally validate our computational predictions, we examined the transcriptional regulation of the miR-200 family of miRNAs. Expression of all five members of the miR-200 family (miR-200a, miR-200b, miR-200c, miR-141, and miR-429) was positively

correlated with p73 and p63 (all with p-values $<2.2\text{x}10^{-5}$), with correlation coefficients ranging from +0.318 to +0.389 for p73, and from +0.263 to +0.322 for p63 (**Table 2.2**).

Our previous work has implicated the miR-200 family in ovarian carcinoma pathogenesis (51), and the miR-200 family has a well-established role in governing epithelial-to-mesenchymal transitions (EMTs) and repressing tumor-initiating cells (i.e., cancer stem cells) (18-20,22) in a variety of other cancer contexts.

The miR-200 family miRNAs are transcribed from two chromosomal clusters: miR-200b/a/429 from chromosome 1, and miR-200c/141 from chromosome 12. The promoters of both miR-200 clusters contain p53 family binding sites. We first tested whether the predicted p53 family binding sites in the miRNA promoters modulate miR-200 transcription by employing a luciferase reporter system. We cloned a 2 kb fragment of the miR-200b/a/429 promoter, which contains one predicted p53 family binding site, upstream of the luciferase gene in the pGL3-Enhancer vector. Similarly, we also produced a miR-200c/141 promoter construct by cloning a 1.5 kb region of the miR-200c/141 promoter containing two predicted p53 binding sites into the vector. We then generated mutant versions of each promoter construct by mutating the consensus nucleotides in each predicted p53 family binding site, with the miR-200c/141 mutant promoter featuring mutations at both of the two predicted binding sites (**Figure 2.3A**). We transfected the wild-type and mutated constructs into 2008 ovarian carcinoma cells. Transfection with the wild-type promoter constructs led to increased luciferase activity as compared to the empty vector lacking a promoter, indicating that the minimal miR-200 promoters were functional (data not shown). To determine if the p53 family binding sites contribute to promoter activity, we compared the luciferase activity of the wild-type versus mutated promoters (**Figure 2.3B).** We found that mutation of the p53 binding sites of the miR-200b/a/429 and miR-200c/141 promoters each resulted in an approximately

40% decrease in luciferase activity (p-value <0.003, *t*-test), consistent with the notion that members of the p53 family directly regulate miR-200 transcription.

We next sought to determine which p53 family members are physically associated with the miR-200 promoters in ovarian carcinoma cells. We performed chromatin immunoprecipitations on the 2008 cell line using antibodies directed against p53, p63 or p73, as well as normal IgG as a control for nonspecific binding. We did not find enrichment of p53 at several known targets, including *p21* and *MDM2*, nor at the miR-200 promoter regions, which is consistent with Western blots revealing that 2008 cells do not express stable p53 protein (data not shown). In the p73 and p63 immunoprecipitations, however, we found a strong enrichment of the miR-200b/a/429 p53 family binding site relative to the IgG control (**Figures 2.3C** and **2.3D**). The levels of enrichment were similar to those seen at the p21 promoter, a known transcriptional target of the entire p53 family (46,52). As a further negative control for nonspecific enrichment, we assayed the immunoprecipitates for a region 1 kb downstream of the 3' end of miR-429, the last miRNA of the miR-200b/a/429 cluster. No enrichment was found at this control region, indicating that we were specifically enriching the promoter of the miR-200b/a/429 locus. When the locus on chromosome 12 corresponding to the miR-200c/141 promoter was analyzed in the p73 and p63 chromatin immunoprecipitates, no enrichment was observed compared to the IgG control (data not shown). Although such negative results are difficult to interpret, this observation may be a false negative due to the limitations of the sensitivity of chromatin immunoprecipitation, such as the limited set of promoter regions assayed by qPCR as well as potentially limited antibody accessibility related to chromatin structure or interference from transcriptional co-factors. As an

additional negative control, we assayed the promoter of miR-29a/b-1, which is overexpressed in ovarian carcinomas yet in our computational analysis was not found to have p53 family transcription factor binding sites. Our ChIP analysis found no enrichment of p63 or p73 at the miR-29a/b-1 promoter (**Figure S2**). Taken together, our results demonstrate that both p73 and p63 associate with the miR-200b/a/429 promoter and may also contribute to the expression of the miR-200c/141 locus in ovarian carcinoma.

Discussion

Our study demonstrates how an integrative computational approach can take advantage of increasingly comprehensive genomic profiling datasets to understand the transcriptional regulation of miRNAs. Although miRNA transcriptional regulation mechanisms have been investigated for specific miRNAs in the past, genome-scale studies have been limited, partially due to the lack of well-annotated miRNA primary transcripts and their corresponding transcriptional start sites. MicroRNAs are excised through a series of processing steps from primary transcripts, which are highly variable in length and can be several kilobases long; identifying the TSS for miRNAs at a genome scale is therefore not trivial. Although experimental methods using chromatin signatures to identify miRNA promoters have been useful in identifying approximate locations of miRNA TSS in some cases (53,54), such approaches have their limitations because miRNA expression is highly tissue specific and therefore experimental data applies to a relatively limited subset of miRNAs in specific cell line or tissue contexts. Our genome-wide approach relies heavily on transcript information derived from ESTs, both from

EST clusters (AceView) and from ESTs derived from a common clone, to provide transcript-based evidence for defining the TSS of intergenic miRNAs. AceView is one of the most comprehensive assemblies of human transcriptome data, drawing cDNA sequences from dbEST, Genbank and RefSeq and representing a broad panel of tissue and cell type contexts. There are currently more than 8 million ESTs in dbEST and notably approximately 2 million 5' capped ESTs in Genbank submitted by Kimura *et al* (55). An important aspect of our study, therefore, is the genome-wide definition of presumptive miRNA TSS and promoters using such data, which provides a resource for further investigations of miRNA transcriptional regulation.

The key biological finding of this work is the identification of the p63 and p73 transcription factors as significant activators of miRNA overexpression in ovarian carcinoma. We found conserved transcription factor binding sites for the p53/p63/p73 family upstream of 17 miRNA loci overexpressed in serous ovarian carcinoma, suggesting p63 and/or p73 may coordinately regulate an ensemble of miRNAs that could significantly modify the cellular gene expression landscape via repression of a cohort of mRNA targets. We also observed an over-representation of conserved *EGR2* binding sites upstream of miRNA loci overexpressed in ovarian carcinoma, with 10 overexpressed miRNA loci containing *EGR2* binding sites. In the TCGA ovarian carcinoma dataset, *EGR2* expression was significantly negatively correlated with the expression of miRNAs containing *EGR2* binding sites. *EGR2* has been suggested to be a tumor suppressor, based on the fact that it is frequently underexpressed in human cancers (including ovarian) and cancer cell lines (56,57). We hypothesize that the 10 miRNA loci

mentioned above are normally repressed by *EGR2*, but that decreased *EGR2* expression in ovarian carcinoma leads to their overexpression.

It is important to note that both p63 and p73 can be expressed as multiple isoforms that were not distinguishable on the expression platforms involved in this study. The isoforms of p63 and p73 can be divided into two classes: transactivating TAp63/TAp73 isoforms and the inhibitory ΔNp63/ΔNp73 isoforms. In ovarian carcinoma, multiple inhibitory isoforms of p73 and p63 are overexpressed and can be associated with poor prognosis (58,59). In contrast, transactivating TAp73 increases the response rate to chemotherapy in BRCA-1-associated ovarian carcinoma (60). We hypothesize that p73 and p63 may drive the expression of multiple miRNAs that contribute to tumor aggressiveness and/or treatment response. For instance, low levels of miR-200 family miRNA expression are associated with poor treatment response in ovarian carcinoma (61); thus, TAp73-mediated activation of miR-200 could serve as a mechanism for increased sensitivity to chemotherapy. Future studies will be required to determine the role of specific isoforms of p73 and/or p63 in determining miRNA transcription in ovarian carcinoma.

In experimental validation of the pipeline predictions, we focused on the miR-200 family because of its important roles in tumor development and progression. Members of the miR-200 family inhibit epithelial-to-mesenchymal transition and suppress tumor invasion by directly repressing the transcription factors Zeb1 and Zeb2 (18-20). Additionally, the miR-200 family has been shown to repress Bmi1 and Suz12, two essential components of polycomb repressor complexes that are responsible for the maintenance of tumor-initiating cells (21,22). Previous work on miR-200 regulation has largely focused on mechanisms of repression, including transcriptional inhibition and epigenetic

modifications (18,62,63); mechanisms of transcriptional activation of this miRNA family are not well-understood. Our results with p63 and p73 shed light onto positive transcriptional regulation of the miR-200 family miRNAs, which has potential therapeutic value in the development of approaches aimed at modifying miR-200 family miRNA expression for the treatment of diverse forms of human cancer.

Although our work focused on miR-200 transcriptional regulation by p73 and p63, recent studies have identified different mechanisms by which the p53 family can modulate miR-200 expression. p63 can serve as a transcriptional regulator of *Dicer* to modulate miRNA maturation, including processing of the miR-200 family (64). Given our observation of p63 binding to the miR-200b/a/429 promoter, these studies suggest that p63 can act both directly and indirectly to alter miRNA expression, although the mechanism may be cellular context-dependent. Additionally, a recent report found that in mammary epithelial cells, p53 serves as a transcriptional activator of miR-200c, but not of the miR-200b/a/429 cluster (17). In ovarian carcinomas, *TP53* is commonly mutated and thus unlikely to drive the overexpression of miR-200 that we detected in our samples; however, *TP53* mutations can frequently yield dominant negative proteins that interfere with the actions of p63 and p73 through promoter competition and through the formation of inactive heterotetrameric complexes (65). Although we did not find a correlation between the *TP53* mutational status of ovarian carcinomas and their expression of miRNAs with p53 family binding sites, we cannot exclude the role of p53 in the transcriptional regulation of the miR-200 family. In contrast, our results indicate that p63 and p73 are involved in overexpression of this miRNA family in ovarian carcinoma. Taken together with the report of p53-driven miR-200c expression in mammary epithelial

cells, our study suggests that cellular context determines the relative importance of different p53 family members on the transcriptional activation of the miR-200 family.

Through our integrative analysis, we identified transcriptional mechanisms that govern miRNA expression in ovarian carcinoma and also found novel activators of miRNAs implicated in tumorigenesis. Our results indicate that p73 and potentially also p63 contribute to broad dysregulation of miRNA expression in ovarian carcinoma. Although we have directed our analysis here toward ovarian cancer, our approach can be broadly applied to other cancer types (e.g., The Cancer Genome Atlas datasets are expected to expand to include nearly 25 cancer types), or any biological state(s) in which mRNA and differential miRNA expression data are available (e.g., expression profiles across stages of development or across different cell differentiation states). Because different cancer and tissue types show unique mRNA and miRNA expression patterns, utilizing our computational pipeline with other systems may reveal distinct transcription factors that regulate miRNA activity in varied biological contexts.

List of Abbreviations

miRNA: microRNA; mRNA: messenger RNA; HOSE: human ovarian surface epithelial; TSS: transcription start site; TF: transcription factor; EST: expressed sequence tag; EMT: epithelial-mesenchymal transition; qPCR: quantitative PCR; TCGA: The Cancer Genome Atlas; RQ: relative quantification.

Authors' contributions

ECK, KSG, and JDA wrote the manuscript. KSG designed and implemented the computational pipeline and TCGA data analyses, DS performed the Affymetrix mRNA microarray analysis and RG and WLR supervised parts of the computational analyses. ECK, JDA, YC, and RKP performed the experimental validation. KW and EMS sequenced *TP53,* and EMS edited the manuscript. CWD, AKG, KCO and NDU helped procure and/or process specimens, and CWD helped design the study. MT designed and supervised the study and helped to write the manuscript. All authors read and approved the final manuscript.

Supplementary Data

Supplementary data are available at NAR online: Supplemental Information, Supplemental Figures S1-S2, and Supplemental Tables S1-S12.

Funding

This work was supported by the National Institutes of Health [PHS NRSA T32 GM07270 from NIGMS to ECK, Interdisciplinary Training in Cancer Research Grant CA80416 to KG, R01 CA140323 to AKG]; the National Cancer Institute [Chromosome Metabolism and Cancer Training Grant NCI CA09657 to JDA, the Pacific Ovarian Cancer Research Consortium Ovarian SPORE Award P50CA083636 to NDU and MT, contract #HHSN261200800001E (caBIG In Silico Research Center of Excellence) to MT]; the

American Cancer Society [Fellowship to JDA]; the Department of Defense [Ovarian Cancer Career Development Award to MT] and by the Jaconette L. Tietze Young Scientist Award [to MT]. The content of this publication is solely the responsibility of the authors and does not necessarily reflect the views or policies of the NIH or Department of Health and Human Services or other funders, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Acknowledgements

We thank Y. Cao for advice pertaining to chromatin immunoprecipitation experiments.

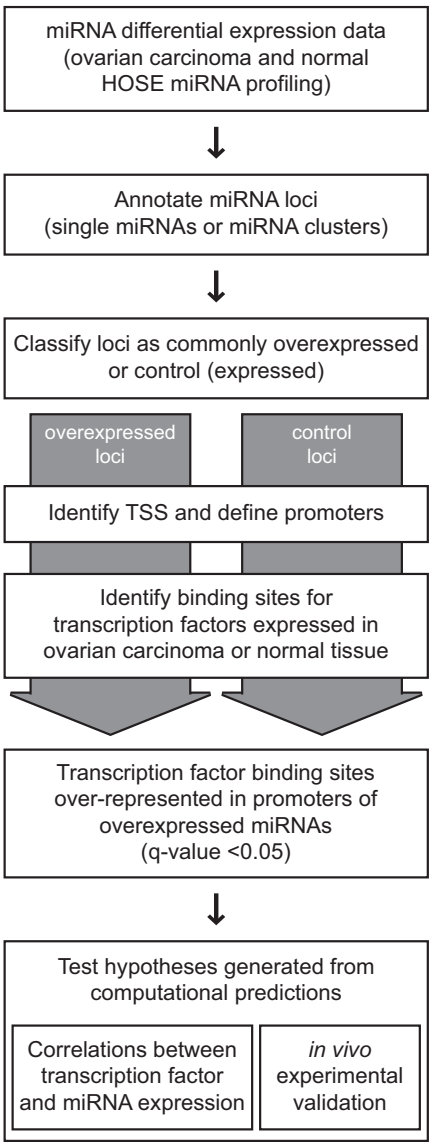

Figure 2.1. Analysis pipeline for identifying candidate transcription factors responsible for miRNA overexpression.
MicroRNA and mRNA expression profiling data were used to identify transcription factor binding sites over-represented in the promoters of miRNAs overexpressed in ovarian carcinoma. Transcription factor binding sites were based upon Transfac position weight matrices. (TSS: transcriptional start site, HOSE: human ovarian surface epithelial cells)

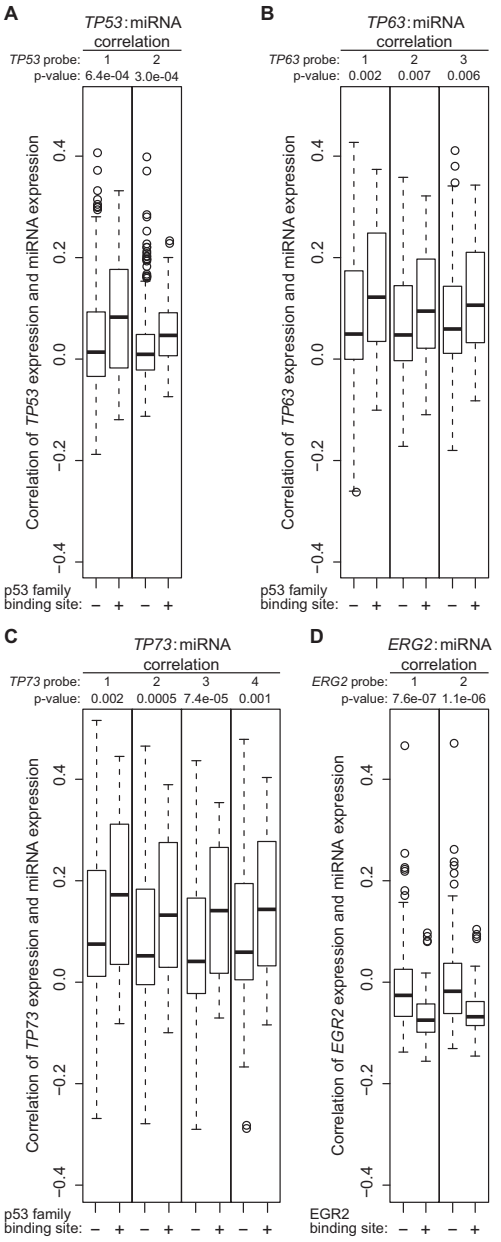

Figure 2.2. Expression of miRNAs with transcription factor binding sites is correlated with transcription factor expression in TCGA ovarian carcinoma specimens.

(A-D) Data from 253 TCGA serous ovarian carcinomas were used to calculate Pearson correlations between expression of the indicated transcription factors and expression of 447 miRNAs. Boxplots show the correlations using mRNA expression data from two *TP53* probes **(A)**, three *TP63* probes **(B)**, four *TP73* probes **(C)**, and two *EGR2* probes **(D)**. miRNAs were classified by the presence (+) or absence (-) of the indicated transcription factor binding site in the miRNA promoter. For each probe, *t*-tests were used to compare mean correlation values of the two miRNA classes, and p-values computed from the *t*-tests are shown.

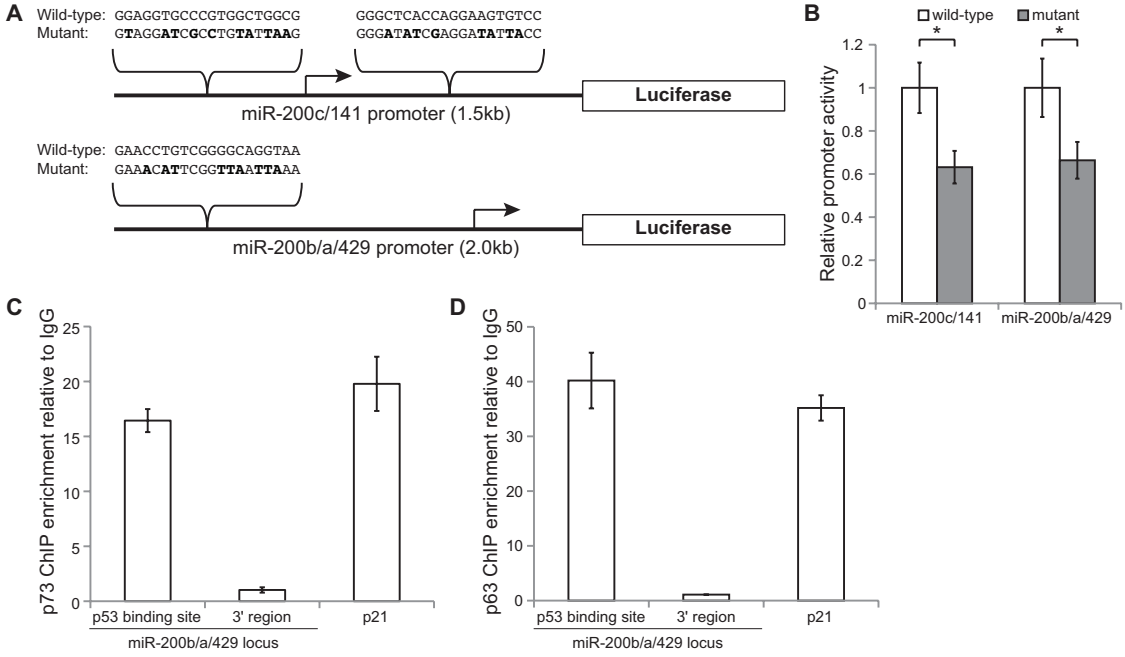

Figure 2.3. p73 and p63 directly regulate transcription of the miR-200 miRNA families.

(A) Schematic of the miR-200c/141 and miR-200b/a/429 promoter fragments cloned upstream of the luciferase gene in the pGL3-Enhancer vector. To generate a mutant version of each promoter, the p53 family consensus binding sites were mutated at the nucleotides shown in bold. **(B)** Activity of the indicated mutant and wild-type miR-200 promoter constructs transfected into 2008 ovarian carcinoma cells. Promoter activity is expressed relative to the wild-type construct for each promoter, and bars represent the mean relative activity ± SD of four replicates. * indicates p-value <0.003, *t*-test. **(C-D)** Chromatin immunoprecipitation of p73 **(C)** and p63 **(D)** from 2008 cells. qPCR was used to assess the enrichment of the indicated regions relative to an IgG control chromatin immunoprecipitation. Enrichment of the p53 family binding sites and negative control regions 1 kb 3' to the miRNAs at the miR-200b/a/429 loci are shown. Enrichment of the p21 promoter served as a positive control. Bars represent mean enrichment ± SD of technical replicates from a representative experiment that was repeated three times.

Table 2.1. Transcription factor binding sites enriched in promoters of miRNAs overexpressed in serous ovarian carcinoma.
MicroRNAs were classified into 30 loci overexpressed in ovarian carcinoma and a control group of 121 loci expressed in carcinoma and/or normal HOSE samples. The number of promoters with binding sites for a given transcription factor was compared between the two classes. A Fisher's exact test was performed for each of the 79 transcription factors that are expressed in carcinoma and/or normal samples. p-values computed by Fisher's exact test were converted to q-values. Transcription factors with p-value <0.05 are shown, while those with q-values <0.05 are in bold.

Transcription factor binding site	Promoters of over-expressed miRNA		Promoters of expressed miRNAs		p-value	q-value
	with site	without site	with site	without site		
p53/p63/p73	**15**	**15**	**17**	**104**	**6.83E-05**	**0.0054**
EGR2	**10**	**20**	**9**	**112**	**0.0006**	**0.0246**
SP1	**10**	**20**	**10**	**111**	**0.0011**	**0.0283**
NFKB1	10	20	16	105	0.0145	0.2856
RELA	9	21	14	107	0.0206	0.2954
NR2F1	6	24	7	114	0.0233	0.2954
EGR3	8	22	11	110	0.0262	0.2954
PAX4	14	16	31	90	0.0429	0.4036
XBP1	7	23	10	111	0.0460	0.4036

Table 2.2. Correlations between *TP73* or *TP63* and miRNA expression in TCGA ovarian carcinomas.
Correlation coefficients for expression of miRNAs with *TP73* expression (*left*) and *TP63* expression (*right*) across 253 TCGA ovarian carcinomas were calculated. MicroRNAs listed showed positive correlation coefficients >0.25, p-values <0.05, and contain p53/p63/p73 binding sites in their promoters. MicroRNAs are ranked in order of increasing p-value. The miR-200 family miRNAs, indicated in bold, had significant positive correlations with *TP73* and *TP63* expression. Correlation coefficients shown were calculated using representative probes for *TP73* (A_23_P74081) and *TP63* (A_32_P114473).

TP73			*TP63*		
miRNA	corr	p-value	miRNA	corr	p-value
miR-107	0.389	1.39E-10	**miR-200a**	0.322	1.64E-07
miR-200a	0.389	1.47E-10	miR-29c	0.320	1.91E-07
miR-29c	0.383	2.77E-10	miR-30e	0.316	2.94E-07
miR-23a	0.373	8.72E-10	miR-30c	0.311	4.50E-07
miR-24	0.371	1.17E-09	miR-107	0.310	5.06E-07
miR-200c	0.364	2.36E-09	**miR-429**	0.294	1.99E-06
miR-141	0.355	6.12E-09	**miR-141**	0.293	2.09E-06
miR-29b	0.346	1.57E-08	miR-24	0.287	3.50E-06
miR-429	0.337	4.03E-08	miR-103	0.283	4.66E-06
miR-103	0.332	6.62E-08	miR-29b	0.278	7.10E-06
miR-200b	0.318	2.31E-07	**miR-200c**	0.275	9.19E-06
miR-28-5p	0.303	9.31E-07	**miR-200b**	0.263	2.20E-05
miR-30e	0.303	9.41E-07	miR-93	0.257	3.57E-05
miR-212	0.302	9.77E-07			
miR-93	0.301	1.06E-06			
miR-7	0.295	1.85E-06			
miR-106b	0.291	2.46E-06			
let-7c	0.289	2.88E-06			
miR-20a	0.284	4.62E-06			
miR-130a	0.278	7.05E-06			
miR-34a	0.276	8.11E-06			
miR-27a	0.274	9.62E-06			
miR-135b	0.271	1.25E-05			
miR-219-5p	0.266	1.77E-05			
miR-19a	0.259	3.02E-05			

References

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

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

3. Chang, T.C., Yu, D., Lee, Y.S., Wentzel, E.A., Arking, D.E., West, K.M., Dang, C.V., Thomas-Tikhonenko, A. and Mendell, J.T. (2008) Widespread microRNA repression by Myc contributes to tumorigenesis. *Nat Genet*, **40**, 43-50.

4. Xi, Y., Shalgi, R., Fodstad, O., Pilpel, Y. and Ju, J. (2006) Differentially regulated micro-RNAs and actively translated messenger RNA transcripts by tumor suppressor p53 in colon cancer. *Clin Cancer Res*, **12**, 2014-2024.

5. Chang, T.C., Wentzel, E.A., Kent, O.A., Ramachandran, K., Mullendore, M., Lee, K.H., Feldmann, G., Yamakuchi, M., Ferlito, M., Lowenstein, C.J. *et al.* (2007) Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Mol Cell*, **26**, 745-752.

6. He, L., He, X., Lim, L.P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D. *et al.* (2007) A microRNA component of the p53 tumour suppressor network. *Nature*, **447**, 1130-1134.

7. Schanen, B.C. and Li, X. (2011) Transcriptional regulation of mammalian miRNA genes. *Genomics*, **97**, 1-6.

8. Calin, G.A. and Croce, C.M. (2006) MicroRNA signatures in human cancers. *Nat Rev Cancer*, **6**, 857-866.

9. Latronico, M.V., Catalucci, D. and Condorelli, G. (2008) MicroRNA and cardiac pathologies. *Physiol Genomics*, **34**, 239-242.

10. Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A. *et al.* (2005) MicroRNA expression profiles classify human cancers. *Nature*, **435**, 834-838.

11. Calin, G.A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S.E., Iorio, M.V., Visone, R., Sever, N.I., Fabbri, M. *et al.* (2005) A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med*, **353**, 1793-1801.

12. Zhang, L., Huang, J., Yang, N., Greshock, J., Megraw, M.S., Giannakakis, A., Liang, S., Naylor, T.L., Barchetti, A., Ward, M.R. *et al.* (2006) microRNAs exhibit high frequency genomic alterations in human cancer. *Proc Natl Acad Sci U S A*, **103**, 9136-9141.

13. Zhang, L., Volinia, S., Bonome, T., Calin, G.A., Greshock, J., Yang, N., Liu, C.G., Giannakakis, A., Alexiou, P., Hasegawa, K. *et al.* (2008) Genomic and epigenetic alterations deregulate microRNA expression in human epithelial ovarian cancer. *Proc Natl Acad Sci USA*, **105**, 7004-7009.

14. Georges, S.A., Biery, M.C., Kim, S.Y., Schelter, J.M., Guo, J., Chang, A.N., Jackson, A.L., Carleton, M.O., Linsley, P.S., Cleary, M.A. *et al.* (2008) Coordinated regulation of cell cycle transcripts by p53-Inducible microRNAs, miR-192 and miR-215. *Cancer Research*, **68**, 10105-10112.

15. Song, B., Wang, Y., Kudo, K., Gavin, E.J., Xi, Y. and Ju, J. (2008) miR-192 Regulates dihydrofolate reductase and cellular proliferation through the p53-microRNA circuit. *Clin Cancer Res*, **14**, 8080-8086.

16. Hermeking, H. (2007) p53 enters the microRNA world. *Cancer Cell*, **12**, 414-418.

17. Chang, C.J., Chao, C.H., Xia, W., Yang, J.Y., Xiong, Y., Li, C.W., Yu, W.H., Rehman, S.K., Hsu, J.L., Lee, H.H. *et al.* (2011) p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nat Cell Biol*, **13**, 317-323.

18. Burk, U., Schubert, J., Wellner, U., Schmalhofer, O., Vincan, E., Spaderna, S. and Brabletz, T. (2008) A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep*, **9**, 582-589.

19. Gregory, P.A., Bert, A.G., Paterson, E.L., Barry, S.C., Tsykin, A., Farshid, G., Vadas, M.A., Khew-Goodall, Y. and Goodall, G.J. (2008) The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol*, **10**, 593-601.

20. Park, S.M., Gaur, A.B., Lengyel, E. and Peter, M.E. (2008) The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev*, **22**, 894-907.

21. Iliopoulos, D., Lindahl-Allen, M., Polytarchou, C., Hirsch, H.A., Tsichlis, P.N. and Struhl, K. (2010) Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells. *Mol Cell*, **39**, 761-772.

22. Shimono, Y., Zabala, M., Cho, R.W., Lobo, N., Dalerba, P., Qian, D., Diehn, M., Liu, H., Panula, S.P., Chiao, E. *et al.* (2009) Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell*, **138**, 592-603.

23. Sarkar, D., Parkin, R., Wyman, S., Bendoraite, A., Sather, C., Delrow, J., Godwin, A.K., Drescher, C., Huber, W., Gentleman, R. *et al.* (2009) Quality assessment and data analysis for microRNA expression arrays. *Nucleic Acids Res*, **37**, e17.

24. Hendrix, N.D., Wu, R., Kuick, R., Schwartz, D.R., Fearon, E.R. and Cho, K.R. (2006) Fibroblast growth factor 9 has oncogenic activity and is a downstream target of Wnt signaling in ovarian endometrioid adenocarcinomas. *Cancer Res*, **66**, 1354-1362.

25. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome Res*, **12**, 996-1006.

26. Rhead, B., Karolchik, D., Kuhn, R.M., Hinrichs, A.S., Zweig, A.S., Fujita, P.A., Diekhans, M., Smith, K.E., Rosenbloom, K.R., Raney, B.J. *et al.* (2010) The UCSC Genome Browser database: update 2010. *Nucleic Acids Res*, **38**, D613-619.

27. Griffiths-Jones, S., Saini, H.K., van Dongen, S. and Enright, A.J. (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res*, **36**, D154-158.

28. Matys, V., Fricke, E., Geffers, R., Gössling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E., Kel-Margoulis, O.V. *et al.* (2003) TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res*, **31**, 374-378.

29. Storey, J.D. and Tibshirani, R. (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A*, **100**, 9440-9445.

30. Smyth, G.K. (2005), *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. Springer, New York, pp. 397-420.

31. Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J. *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*, **5**, R80.

32. Galic, V., Willner, J., Wollan, M., Garg, R., Garcia, R., Goff, B.A., Gray, H.J. and Swisher, E.M. (2007) Common polymorphisms in TP53 and MDM2 and the relationship to TP53 mutations and clinical outcomes in women with ovarian and peritoneal carcinomas. *Genes Chromosomes Cancer*, **46**, 239-247.

33. Nelson, J.D., Denisenko, O. and Bomsztyk, K. (2006) Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nat Protoc*, **1**, 179-185.

34. Dahiya, N., Sherman-Baust, C.A., Wang, T.L., Davidson, B., Shih Ie, M., Zhang, Y., Wood, W., 3rd, Becker, K.G. and Morin, P.J. (2008) MicroRNA expression and identification of putative miRNA targets in ovarian cancer. *PLoS One*, **3**, e2436.

35. Iorio, M.V., Visone, R., Di Leva, G., Donati, V., Petrocca, F., Casalini, P., Taccioli, C., Volinia, S., Liu, C.G., Alder, H. *et al.* (2007) MicroRNA signatures in human ovarian cancer. *Cancer Res*, **67**, 8699-8707.

36. Nam, E.J., Yoon, H., Kim, S.W., Kim, H., Kim, Y.T., Kim, J.H., Kim, J.W. and Kim, S. (2008) MicroRNA expression profiles in serous ovarian carcinoma. *Clin Cancer Res*, **14**, 2690-2695.

37. Wyman, S.K., Parkin, R.K., Mitchell, P.S., Fritz, B.R., O'Briant, K., Godwin, A.K., Urban, N., Drescher, C.W., Knudsen, B.S. and Tewari, M. (2009) Repertoire of microRNAs in epithelial ovarian cancer as determined by next generation sequencing of small RNA cDNA libraries. *PLoS ONE*, **4**, e5311.

38. Yang, H., Kong, W., He, L., Zhao, J.J., O'Donnell, J.D., Wang, J., Wenham, R.M., Coppola, D., Kruk, P.A., Nicosia, S.V. *et al.* (2008) MicroRNA expression profiling in human ovarian cancer: miR-214 induces cell survival and cisplatin resistance by targeting PTEN. *Cancer Res*, **68**, 425-433.

39. Tanzer, A. and Stadler, P.F. (2004) Molecular evolution of a microRNA cluster. *J Mol Biol*, **339**, 327-335.

40. Baskerville, S. and Bartel, D.P. (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*, **11**, 241-247.

41. Pruitt, K.D., Tatusova, T. and Maglott, D.R. (2005) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res*, **33**, D501-504.

42. Saini, H.K., Griffiths-Jones, S. and Enright, A.J. (2007) Genomic analysis of human microRNA transcripts. *Proc Natl Acad Sci U S A*, **104**, 17719-17724.

43. Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L. and Bradley, A. (2004) Identification of mammalian microRNA host genes and transcription units. *Genome Res*, **14**, 1902-1910.

44. Thierry-Mieg, D. and Thierry-Mieg, J. (2006) AceView: a comprehensive cDNA-supported gene and transcripts annotation. *Genome Biol*, **7 Suppl 1**, S12 11-14.

45. Down, T.A. and Hubbard, T.J. (2002) Computational detection and location of transcription start sites in mammalian genomic DNA. *Genome Res*, **12**, 458-461.

46. Kaghad, M., Bonnet, H., Yang, A., Creancier, L., Biscan, J.C., Valent, A., Minty, A., Chalon, P., Lelias, J.M., Dumont, X. *et al.* (1997) Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. *Cell*, **90**, 809-819.

47. Schmale, H. and Bamberger, C. (1997) A novel protein with strong homology to the tumor suppressor p53. *Oncogene*, **15**, 1363-1367.

48. Havrilesky, L., Darcy, M., Hamdan, H., Priore, R.L., Leon, J., Bell, J. and Berchuck, A. (2003) Prognostic significance of p53 mutation and p53

overexpression in advanced epithelial ovarian cancer: a Gynecologic Oncology Group Study. *J Clin Oncol*, **21**, 3814-3825.

49. Leitao, M.M., Soslow, R.A., Baergen, R.N., Olvera, N., Arroyo, C. and Boyd, J. (2004) Mutation and expression of the TP53 gene in early stage epithelial ovarian carcinoma. *Gynecol Oncol*, **93**, 301-306.

50. Nigro, J.M., Baker, S.J., Preisinger, A.C., Jessup, J.M., Hostetter, R., Cleary, K., Bigner, S.H., Davidson, N., Baylin, S., Devilee, P. *et al.* (1989) Mutations in the p53 gene occur in diverse human tumour types. *Nature*, **342**, 705-708.

51. Bendoraite, A., Knouf, E.C., Garg, K.S., Parkin, R.K., Kroh, E.M., O'Briant, K.C., Ventura, A.P., Godwin, A.K., Karlan, B.Y., Drescher, C.W. *et al.* (2010) Regulation of miR-200 family microRNAs and ZEB transcription factors in ovarian cancer: evidence supporting a mesothelial-to-epithelial transition. *Gynecologic Oncology*, **116**, 117-125.

52. Westfall, M.D., Mays, D.J., Sniezek, J.C. and Pietenpol, J.A. (2003) The Delta Np63 alpha phosphoprotein binds the p21 and 14-3-3 sigma promoters in vivo and has transcriptional repressor activity that is reduced by Hay-Wells syndrome-derived mutations. *Mol Cell Biol*, **23**, 2264-2276.

53. Ozsolak, F., Poling, L.L., Wang, Z., Liu, H., Liu, X.S., Roeder, R.G., Zhang, X., Song, J.S. and Fisher, D.E. (2008) Chromatin structure analyses identify miRNA promoters. *Genes Dev*, **22**, 3172-3183.

54. Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J. *et al.* (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell*, **134**, 521-533.

55. Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H. *et al.* (2006) Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res*, **16**, 55-65.

56. Unoki, M. and Nakamura, Y. (2001) Growth-suppressive effects of BPOZ and EGR2, two genes involved in the PTEN signaling pathway. *Oncogene*, **20**, 4457-4465.

57. Unoki, M. and Nakamura, Y. (2003) EGR2 induces apoptosis in various cancer cell lines by direct transactivation of BNIP3L and BAK. *Oncogene*, **22**, 2172-2185.

58. Concin, N., Becker, K., Slade, N., Erster, S., Mueller-Holzner, E., Ulmer, H., Daxenbichler, G., Zeimet, A., Zeillinger, R., Marth, C. *et al.* (2004) Transdominant DeltaTAp73 isoforms are frequently up-regulated in ovarian

cancer. Evidence for their role as epigenetic p53 inhibitors in vivo. *Cancer Res*, **64**, 2449-2460.

59. Marchini, S., Marabese, M., Marrazzo, E., Mariani, P., Cattaneo, D., Fossati, R., Compagnoni, A., Fruscio, R., Lissoni, A.A. and Broggini, M. (2008) DeltaNp63 expression is associated with poor survival in ovarian cancer. *Ann Oncol*, **19**, 501-507.

60. Ibrahim, N., He, L., Leong, C.O., Xing, D., Karlan, B.Y., Swisher, E.M., Rueda, B.R., Orsulic, S. and Ellisen, L.W. (2010) BRCA1-associated epigenetic regulation of p73 mediates an effector pathway for chemosensitivity in ovarian carcinoma. *Cancer Res*, **70**, 7155-7165.

61. Leskela, S., Leandro-Garcia, L.J., Mendiola, M., Barriuso, J., Inglada-Perez, L., Munoz, I., Martinez-Delgado, B., Redondo, A., de Santiago, J., Robledo, M. *et al.* (2011) The miR-200 family controls beta-tubulin III expression and is associated with paclitaxel-based treatment response and progression-free survival in ovarian cancer patients. *Endocr Relat Cancer*, **18**, 85-95.

62. Vrba, L., Jensen, T.J., Garbe, J.C., Heimark, R.L., Cress, A.E., Dickinson, S., Stampfer, M.R. and Futscher, B.W. (2010) Role for DNA methylation in the regulation of miR-200c and miR-141 expression in normal and cancer cells. *PLoS ONE*, **5**, e8697.

63. Bracken, C.P., Gregory, P.A., Kolesnikoff, N., Bert, A.G., Wang, J., Shannon, M.F. and Goodall, G.J. (2008) A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Research*, **68**, 7846-7854.

64. Su, X., Chakravarti, D., Cho, M.S., Liu, L., Gi, Y.J., Lin, Y.L., Leung, M.L., El-Naggar, A., Creighton, C.J., Suraokar, M.B. *et al.* (2010) TAp63 suppresses metastasis through coordinate regulation of Dicer and miRNAs. *Nature*, **467**, 986-990.

65. Brosh, R. and Rotter, V. (2009) When mutants gain new powers: news from the mutant p53 field. *Nat Rev Cancer*, **9**, 701-713.

Chapter 3

Cell density dependent regulation of miR-210 by HIF-1α

Abstract

Under hypoxic conditions, microRNA-210 is dramatically upregulated and facilitates a cell's adaption to the low oxygen environment via modulation of cellular respiration and apoptosis. While miR-210 has been well characterized as a hypoxic-specific miRNA, other changes in a cell's microenvironment that may drive miR-210 induction are not understood. We found that miR-210 is broadly expressed in a panel of cancer cell lines cultured under normoxic conditions. To investigate potential new mechanisms of miR-210 activation, we conducted *in vitro* time course experiments to assess miR-210 levels at increasing cell densities. We found that primary miR-210 transcription is upregulated by HIF-1α in response to increased cell confluency. The induction is conserved across multiple cell types, and does not require cell adhesion. Together, these studies suggest that miR-210 regulation is highly sensitive to localized changes in the cellular microenvironment. Additionally, we have identified a novel mechanism by which cell density yields differential expression of a single miRNA. Our work emphasizes that caution must be used in controlling for confluency in cell culture experiments. While the functional effects of miR-210 induction in high density culture are still under investigation, our results suggest that miR-210 may facilitate a cell's ability to tolerate variable environmental conditions.

Introduction

MicroRNAs can serve as important tools for a cell's adaptation to changing conditions in the microenvironment. While some miRNAs are ubiquitously expressed, others are rapidly induced in response to various stresses. In response to reduced oxygen levels, miR-210 has been demonstrated to show striking upregulation across a wide panel of human tissues (1-3). The induction of miR-210 primarily results from transcriptional activation by the hypoxia-inducible factor (HIF) proteins (2-4). HIFs are heterodimeric transcription factors that consist of a constitutively active subunit, HIF-β, in addition to an oxygen-sensitive subunit (HIF-1α, 2α, or 3α, depending on the cell-type) (5). The three HIF-α subunits feature partially overlapping targets and expression patterns (6). In epithelial cells, the best-characterized component is the HIF-1α isoform, which is a highly unstable protein under normoxic conditions and is rapidly degraded in the proteasome following hydroxylation of conserved proline resides via the VHL E3 ubiquitin ligase; however, as oxygen levels decrease, HIF-1α hydroxylation is reduced, which prevents HIF degradation and allows protein accumulation (7). HIF-1α stabilization facilitates the formation of the HIF heterodimers that induce the expression of many genes necessary for a cell's successful adaption to the hypoxic environment, including the miRNA miR-210 (2).

MicroRNA-210 targets multiple regulators of cellular proliferation, apoptosis, and metabolism. Under conditions of low oxygen, miR-210 induction represses the iron-sulfur cluster assembly proteins (ISCU1/2) necessary for oxidative respiration and facilitates a shift to increased glycolysis (8-10). MicroRNA-210 also prevents the induction of apoptosis under hypoxic conditions, potentially via suppression of several

factors that promote cell death, including the apoptosis-inducting factor [mitochondrion-associated 3] (AIFM3) (11) and caspase-8 associated protein 2 (12). The receptor tyrosine kinase ligand Ephrin-A3 (EFNA3) is also a well-documented miR-210 target, and may enable miR-210 to modulate cell motility and angiogenesis (2,13). Interestingly, miR-210 has conflicting roles in tumorigenesis: Overexpression of miR-210 in a mouse xenograft model delayed tumor initiation (3); however, in many human cancers, miR-210 levels are elevated, which may reflect the divergent roles miR-210 may play in different stages of cancer initiation and metastasis. In breast cancer, increased levels of miR-210 are associated with increased invasiveness and poor prognosis (14,15). Levels of miR-210 circulating in the plasma of breast cancer patients are associated with reduced sensitivity to the drug trastuzumab, which raises the possibility that miR-210 could also modulate chemosensitivity (16).

Although miR-210 is best known as a miRNA transcriptionally induced under low oxygen conditions, other mechanisms of regulation that could be important in disease pathology have not been described. During the course of experiments designed to explore the induction and degradation of miR-210 in hypoxia, we were surprised to find that miR-210 is an abundantly expressed miRNA in multiple cancer cell lines under standard tissue culture conditions. We found that miR-210 levels are highly sensitive to cell density, with HIF-1α serving as a potent inducer of miR-210 under normoxic conditions in response to increased cell number. Hwang et al. have examined density-dependent expression changes in mature miRNA, and found that higher cell density correlated with approximately 2-fold higher expression of most miRNAs across multiple cancer cell lines (17). The authors attributed this increase to enhanced Drosha processing

of the primary miRNA transcripts in confluent cultures, as both precursor and mature miRNA levels were increased. Here, we report an independent mechanism of density-dependent miRNA regulation that is specific to miR-210. We found that primary miRNA transcription is robustly induced by HIF-1α in response to small increases in cell confluency. These results underscore the importance of controlling for equivalent cell densities when assessing differentially regulated miRNAs. Additionally, these studies reveal the sensitivity and precision of HIF-1α's regulation of miR-210 levels, which may enable the successful adaption of a cell to its microenvironment.

Methods

Cell Culture

HCT-116 cells were maintained in McCoy's media (Invitrogen) with 10% FBS (Atlanta Biologicals). For time course experiments, HCT-116 were seeded at 50,000 cells per well and lysed at 2, 48, and 96 hours post-seeding. For cell density experiments, HCT-116 cells were seeded at various concentrations in 6-well plates in 3 mL media. 24 h post seeding, the media was changed. At 48 h, the cells were washed in 1X PBS (Gibco), and lysed in 700 uL Qiazol (Qiagen). IMR-90 and 293T cells were prepared as described for HCT-116, using Eagle's Minimum and DMEM media (Invitrogen) respectively. For the suspension cell line, THP-1, cells were seeded at increasing density in untreated 6 well-plates with 1×10^5 - 1×10^6 cells per well. 48 h post seeding, cells were pelleted by centrifugation, washed in PBS, and lysed in 700 µL Qiazol.

To attain conditioned media from cells of varying density, HCT-116 cells were seeded in 6-well plates at low density (1×10^5 cells per well) or high density (1×10^6 cells per well)

in 3 mL media. Recipient cells were seeded at three different densities: 1 x 10^5 cells per well, 3 x 10^5 cells per well, or 6 x 10^5 cells per well. 24 h post plating, the conditioned media from multiple wells of low and high density donor wells was pooled, filtered through 0.45 uM filters to remove cells, and then applied to the recipient cells. 24 h post addition of the conditioned media, recipient cells were harvested for RNA.

RNA isolation and qRT-PCR

RNA was isolated with the miRNeasy kit (Qiagen) according to the manufacturer's instructions, and diluted to equal concentration. qRT-PCR was performed using Taqman assays (Applied Bioystems) for miR-210, miR-15a, pri-miR-210, and pri-miR-15a/16 and their control assays, RNU-48 and GUSB. Relative expression was determined by using RNU-48 or GUSB to compute delta CT values, and calculating delta delta CT to the lowest density sample. For primary miRNA profiling, Taqman assays were obtained from Applied Biosystems for 70 different miRNAs. qPCR primers and probes were diluted (1.25 uL water per 0.25 uL primer-probe mixture) and aliquoted into 384-well plates in an arrayed manner with duplicate wells per assay. For each sample of interest, a single 100 uL reverse transcription reaction was performing using the cDNA High Capacity Reverse Transcription (RT) Kit (Applied Biosystems) using 574 ng RNA. Each RT reaction was then diluted with 100 uL water and 500 uL 2X Master mix from Applied Biosystems. 3.5 uL of the diluted sample was added to each well of the 384-well plates, and amplification was performed on a Viia7 qPCR machine (Applied Bioystems). Technical replicates on each plate were averaged, and any miRNA with an average cycle threshold (CT) > 34 across all samples, or with an undetermined CT in ≥ 10 samples was removed. Delta CT values were calculated to the *ACTB* housekeeping gene present on

every card, and the relative expression of each miRNA compared to the expression at the earliest time point was computed.

Suppression of *HIF-1α* and *EPAS-1* in HCT-116

HCT-116 cells were seeded at 2 x 10^6 cells per 6 cm dish. 24 h post seeding, the cells were transfected with siRNAs for *HIF-1α* and *EPAS-1* (*HIF-2α*), in addition to non-targeting control siRNAs (Dharmacon On-Target+ Smartpools). Biological replicates of each siRNA were performed. For the knockdown of both *HIF-1α* and *EPAS-1*, 50 nM of each siRNA was used. For all other knockdowns, 50 nM of siRNA was supplemented with 50 nM non-targeting siRNA to ensure all wells contained a final siRNA concentration of 100 nM. All transfections were performed using Lipofectamine RNAiMAX (Invitrogen) and Optimem low-serum media (Invitrogen) according to the manufacturer's instructions. 6 h post transfection, the wells were washed 1X in PBS, trypsinized, and re-seeded into high and low density wells. 48 h post transfection, all wells were lysed in Qiazol as described above.

Results

MicroRNA-210 can be potently induced in response to hypoxic conditions via activation by the HIF transcription factors (2-4). While investigating the potential induction of miR-210 under low oxygen conditions, we were surprised to observe that miR-210 is abundantly expressed in a variety of cell lines under seemingly normoxic conditions. In a panel of 22 different cancer cell lines, all featured miR-210 levels of greater than 1 x 10^6 copies per ng of total RNA **(Figure 3.1)**. The ubiquity of miR-210 expression suggests that the miRNA may be activated by alternative mechanisms.

Because several of the cell lines used for this analysis were prepared from nearly confluent plates of cells, we asked whether cell density could affect miR-210 expression levels. We performed time course experiments with HCT-116 colon cancer cells, which displayed intermediate miR-210 expression levels in our panel of cell lines and grow rapidly under standard tissue culture conditions. Using qRT-PCR, we compared miR-210 expression in wells of HCT-116 cells grown for increasing time points post plating. We found an induction in miR-210 of approximately four-fold by 96 hours (**Figure 3.2A**). This induction was in stark contrast to the miRNA miR-15a, which did not show a change in expression over time (**Figure 3.2A**). To investigate the potential mechanism of this increase, we also measured primary miR-210 levels, as rapid transcriptional activation of miR-210 has been well documented. We found that primary miR-210 displayed an even stronger increase in expression of over 10 fold by 96 hours (**Figure 3.2B**). To determine whether this induction was unique to miR-210, we profiled the expression of 70 other primary miRNA transcripts in a time course of four days (**Figure 3.2C**). We found that the average primary transcript expression was slightly increased over time; however, primary miR-210 showed a striking induction of over 30-fold in the four day time course. Therefore, at increasing time points, and potentially as the density of the wells is increased, miR-210 levels are dramatically raised.

To more directly test the effects of cell density on miR-210, we plated a 10-fold dilution series of HCT-116 cells and assayed miR-210 expression after 48 h of culture. We found that pri-miR-210 expression showed a strong increase in higher density cultures, while the expression of the control miRNA pri-miR-15a/16-1 did not change (**Figure 3.2D**). Together, these experiments demonstrate a strong correlation in cell

density and miR-210 expression in HCT-116 cells. We expanded these experiments to 293T and non-transformed IMR-90 cells, and determined the relationship between cells number and miR-210 expression. Samples with higher cell density showed a strong increase in pri-miR-210 expression (**Figure 3.3A-B**), indicating that the induction of miR-210 is not cell line dependent. However, HCT-116, 293T, and IMR-90 are all adherent cell lines. We thus also performed our analysis on a suspension cell line in order to address whether cell adhesion is required for miR-210 induction. In the acute monocytic leukemia cell line, THP-1, we found a strong increase in primary miR-210 in higher density cultures (**Figure 3.3C**), which indicated that the response is independent of cell adhesion.

We next investigated the mechanism of the rapid induction of miR-210. The regulation of miR-210 has been investigated in depth under hypoxic settings, in which HIF-1α and in some cell lines, EPAS-1 (known as HIF-2α), induces miR-210 (2-4). Yet the mechanism of miR-210 activation under normoxic settings, in which HIF-1/2 α is thought to be rapidly degraded, has not been investigated. One possibility to explain the induction is that a soluble factor could be mediating the upregulation of miR-210. However, when conditioned media from either high or low density cells was applied to recipient cells for 24 h (**Figure 3.4A**), miR-210 expression was not altered by application of conditioned media from high density cells. The miR-210 levels of the recipient cells were entirely dependent on the density in which they had been seeded (**Figure 3.4**). We conducted additional experiments in which the length of incubation with conditioned media was extended to 72 h, and again observed no changes in miR-210 (data not shown).

Together, these experiments suggest that miR-210 induction is not mediated by a soluble factor.

We thus re-focused our attention on the HIF transcription factors to determine whether stabilization of these proteins in high-density culture may explain the induction of miR-210. Previous studies have demonstrated that confluent cultures yield areas of localized hypoxia that can activate HIF-1α transcriptional activity (18), but the potential effects on miR-210 have not been examined. We used siRNAs against HIF-1α and EPAS-1 (HIF-2α) individually and in combination with one another to suppress HIF expression in HCT-116 cells. The transfected cells were split into low and high density plates 6 h post transfection, and then assayed two days later to determine if suppression of these factors abrogated the induction seen in high density cells. To confirm that our suppression of HIF-1α and HIF-2α was functionally effective, we used qRT-PCR to assay the expression of the *CA9* and *BNIP3* transcripts, which are well-characterized transcriptional targets of the HIF family. Both *CA9* and *BNIP3* showed strong induction in the high-density cells treated with non-targeting siRNAs (**Figure 3.5C-D**), indicating that HIF proteins are stabilized in the high-density samples. When HIF-1α was suppressed with siRNAs, however, the induction of *CA9* and *BNIP3* was reduced by approximately two-fold. In this cell line, suppression of EPAS-1 (HIF-2α) did not significantly alter the expression of these targets, which is consistent with previous reports showing HIF-1α to be the primary form expressed in epithelial cells. We next examined the effects of HIF suppression on miR-210: We found that suppression of HIF-1α decreased the induction of both primary and mature miR-210 in the high-density plates by approximately 40-50% (**Figures 3.5A-B**). Suppression of both HIF-1α and

HIF-2α did not further block induction of pri-miR-210. Therefore, in HCT-116 cells, HIF-1α appears to serve as the dominant driver of miR-210 induction in high density cultures.

Discussion

We have found that miR-210 displays robust induction in response to increased cell density. Our results demonstrate that increased cell confluency, even in cases of fairly sparse cultures, creates localized microenvironments sufficient for HIF-1α activation of miR-210. Thus, miR-210 appears to act as a highly sensitive indicator of altered oxygen levels, although we cannot exclude that HIF-1α is stabilized via an additional, oxygen-independent mechanism. While this is the first demonstration of the induction of miR-210 in high density cultures, the HIF proteins have previously been shown to be stabilized by increased cell confluency. In MDA-MD-231 breast cancer cells, increasing cell density led to accumulation of the HIF-1α protein and increased transcriptional activity; Interestingly, shaking of the cultured cells abrogated the induction, likely as a result of the disruption of the oxygen gradient of the pericellular environment (18). The mechanism of HIF-1α activation remains a controversial topic. Previous studies have come to contrasting conclusions regarding the importance of soluble factors in mediating HIF activation. While one study has reported conditioned media is sufficient for the stabilization of HIF-1α (18), we did not observe miR-210 induction in response to the application of conditioned media from dense cultures; rather, we found the induction to be largely dependent upon HIF-1α levels, although we cannot rule out that additional factors may also be involved.

The impact of HIF induction of miR-210 in response to high cell density remains an open question. We hypothesize miR-210 upregulation may have substantial effects on the chemosensitivity of cancer cells. Poor drug penetrance of densely packed cancer cells is considered a major impediment of successful chemotherapy (19,20). *In vitro*, confluence-mediated resistance to chemotherapy has been described in murine and human mammary cells and colon cancer cells (18,21-23). Additionally, suppression of HIF-1α has been shown to increase the chemosensitivity of high-density cultures (18), which suggests that downstream effectors of HIF may play important roles in cancer cell survival. MicroRNA-210 has been shown to regulate cell death and proliferation in response to hypoxia; therefore, future work is needed to address the contributions of miR-210 in mediating confluence-dependent chemoresistance.

Finally, these studies reveal a new mechanism of density-dependent miRNA regulation. While a previous report described global increases in the processing of miRNAs in response to increased cell confluency (17), we have discovered an independent mechanism of activation that is specific to a single miRNA. The unique induction of primary-miR-210 transcription emphasizes that cell confluency must be monitored to avoid potential misinterpretations of the mechanisms of miR-210 differential expression. Additionally, these studies suggest that in addition to serving as a sensor for hypoxic areas within tissues, miR-210 displays rapid activation within localized areas of high concentrations of cells. The tumor microenvironment features gradients in oxygen availability and cell density. Thus, our studies predict that miR-210 may exert a larger impact within the heterogeneous tumor environment than has been previously suggested.

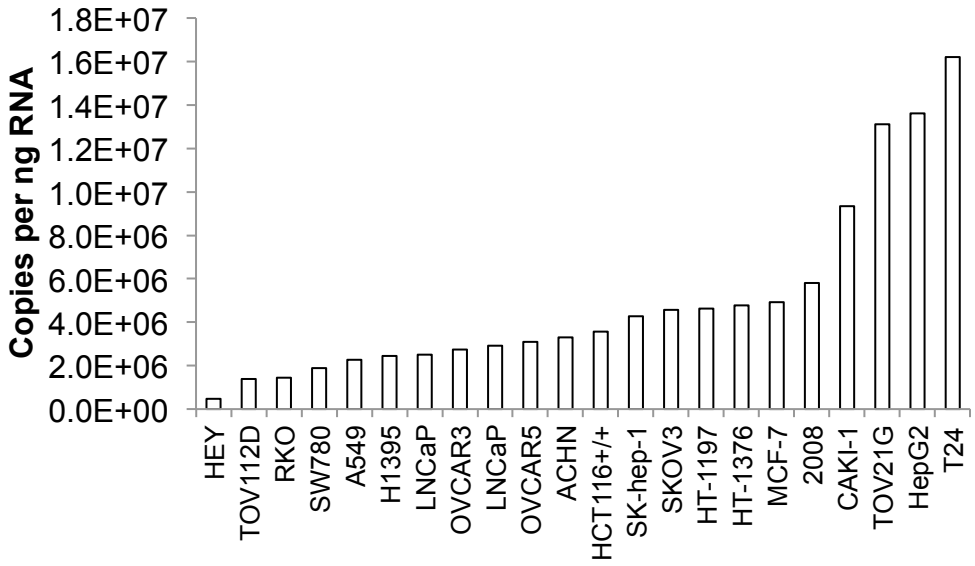

Figure 3.1. MicroRNA-210 is abundantly expressed in cancer cell lines.
Bars depict the copies of miR-210 per ng RNA used as input for qRT-PCR with Taqman assays. Cell lines cultured under normoxic conditions all featured readily detectable levels of miR-210, although the expression levels varied approximately 10-fold across our panel of cell lines.

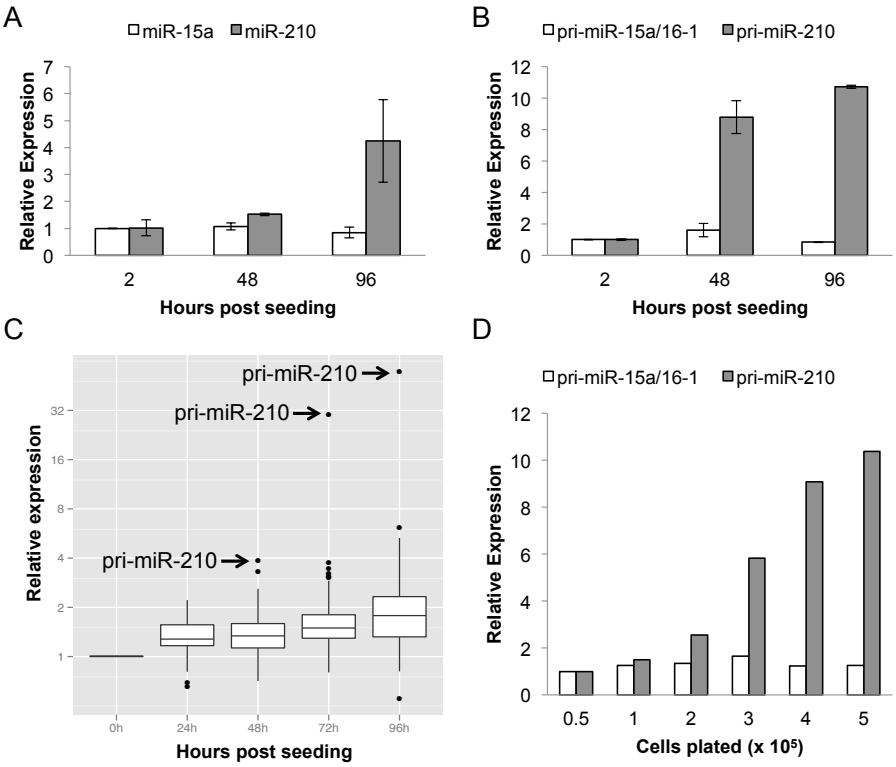

Figure 3.2. miR-210 is induced with increasing cell density.
(A) qRT-PCR was used to quantify mature miRNA expression in HCT-116 cells at increasing time points post seeding. Bars represent the average expression of miR-210 and miR-15a in duplicate samples +/- SD. **(B)** Average expression of primary miRNA transcripts in the samples from (A) as determined by qRT-PCR. **(C)** The expression of primary miRNA transcripts were profiled by qRT-PCR at increasing time points post cell plating. The boxplots show the expression of 48 primary transcripts relative to their expression at the earliest time point. MicroRNA-210 shows dramatic induction over time, and appears as a distant outlier compared to the other miRNAs. **(D)** Primary miRNA levels in HCT-116 cells were plated at increasingly cell densities and lysed 48 hours post seeding. qRT-PCR PCR was used to assess the expression of pri-miR-210 in addition to a control miRNA, pri-miR-15a/-16-1.

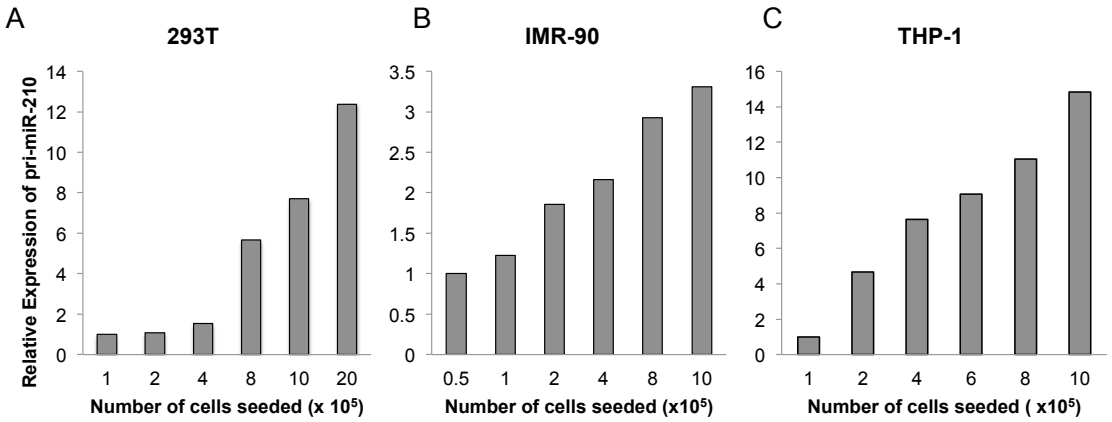

Figure 3.3. miR-210 levels are elevated with increasing cell density in a panel of adherent and non-adherent cell lines.
qRT-PCR was used to assay primary miR-210 levels 48 h post-seeding of increasing number of **(A)** 293T cells, **(B)** IMR-90 non-transformed lung fibroblasts, or **(C)** the non-adherent THP-1 monocytic leukemia cell line.

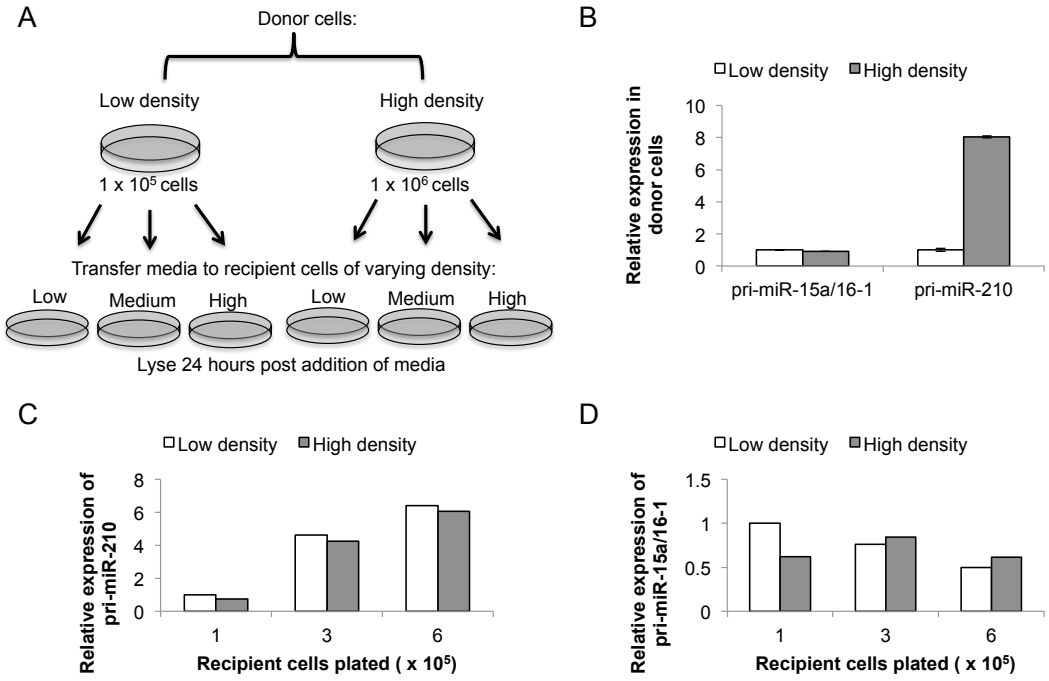

Figure 3.4. miR-210 induction at high cell densities is not mediated by a soluble factor.

(A) Scheme of media swap experiment. The conditioned media from donor cells plated at high or low density 24 h before media removal was added to recipient plates of varying densities. **(B)** Confirmation of high levels of miR-210 in the high density donor cells at the time of media harvesting. **(C)** Relative expression of pri-miR-210 in the recipient cells treated with the conditioned media from low or high density donor cells. **(D)** Relative expression of pri-miR-15a/16 in the recipient cells. MicroRNA-15a/16-1 was used as a control miRNA not induced by cell density.

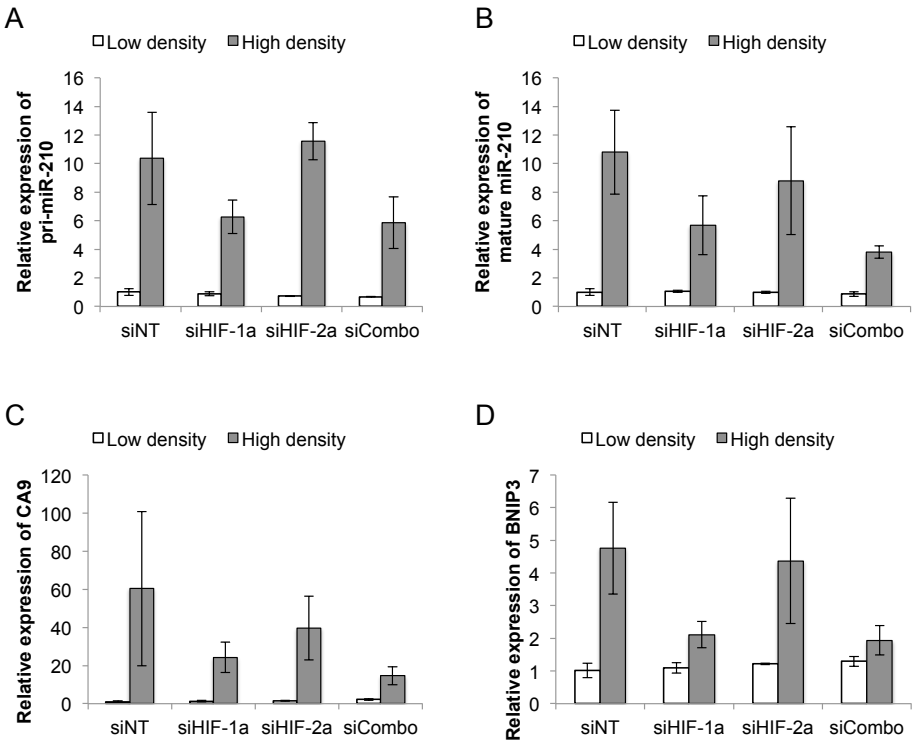

Figure 3.5. Suppression of HIF-1α abrogates induction of miR-210 and additional HIF targets in high-density cultures.
HCT-116 cells were treated with non-targeting siRNAs (siNT), or siRNAs against HIF-1α, HIF-2α (EPAS-1), or a combination of both HIF siRNAs. After transfection, the cells were split into high and low density wells and assayed by qRT-PCR for changes in gene and miRNA expression. Bars represent average values of biological replicates +/- SD. **(A)** Relative expression of primary miR-210 following HIF knockdown. **(B)** Levels of mature miR-210. **(C)** Induction of carbonic anhydrase IX, a known target of HIF-1α, in response to high cell density is reduced by the suppression of HIF-1α. **(D)** Induction of the HIF target *BNIP3* is also decreased following HIF-1α suppression.

References

1. Giannakakis, A., Sandaltzopoulos, R., Greshock, J., Liang, S., Huang, J., Hasegawa, K., Li, C., O'Brien-Jenkins, A., Katsaros, D., Weber, B.L. *et al.* (2008) miR-210 links hypoxia with cell cycle regulation and is deleted in human epithelial ovarian cancer. *Cancer Biol Ther*, **7**, 255-264.

2. Huang, X., Ding, L., Bennewith, K.L., Tong, R.T., Welford, S.M., Ang, K.K., Story, M., Le, Q.-T. and Giaccia, A.J. (2009) Hypoxia-inducible mir-210 regulates normoxic gene expression involved in tumor initiation. *Mol Cell*, **35**, 856-867.

3. Zhang, Z., Sun, H., Dai, H., Walsh, R.M., Imakura, M., Schelter, J., Burchard, J., Dai, X., Chang, A.N., Diaz, R.L. *et al.* (2009) MicroRNA miR-210 modulates cellular response to hypoxia through the MYC antagonist MNT. *Cell cycle (Georgetown, Tex)*, **8**, 2756-2768.

4. Crosby, M.E., Kulshreshtha, R., Ivan, M. and Glazer, P.M. (2009) MicroRNA regulation of DNA repair gene expression in hypoxic stress. *Cancer Research*, **69**, 1221-1229.

5. Wenger, R.H. (2002) Cellular adaptation to hypoxia: O2-sensing protein hydroxylases, hypoxia-inducible transcription factors, and O2-regulated gene expression. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, **16**, 1151-1162.

6. Wiesener, M.S., Jurgensen, J.S., Rosenberger, C., Scholze, C.K., Horstrup, J.H., Warnecke, C., Mandriota, S., Bechmann, I., Frei, U.A., Pugh, C.W. *et al.* (2003) Widespread hypoxia-inducible expression of HIF-2alpha in distinct cell populations of different organs. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, **17**, 271-273.

7. Weidemann, A. and Johnson, R.S. (2008) Biology of HIF-1alpha. *Cell death and differentiation*, **15**, 621-627.

8. Favaro, E., Ramachandran, A., McCormick, R., Gee, H., Blancher, C., Crosby, M., Devlin, C., Blick, C., Buffa, F., Li, J.-L. *et al.* (2010) MicroRNA-210 regulates mitochondrial free radical response to hypoxia and krebs cycle in cancer cells by targeting iron sulfur cluster protein ISCU. *PLoS ONE*, **5**, e10345.

9. Chen, Z., Li, Y., Zhang, H., Huang, P. and Luthra, R. (2010) Hypoxia-regulated microRNA-210 modulates mitochondrial function and decreases ISCU and COX10 expression. *Oncogene*.

10. Chan, S.Y., Zhang, Y.Y., Hemann, C., Mahoney, C.E., Zweier, J.L. and Loscalzo, J. (2009) MicroRNA-210 controls mitochondrial metabolism during hypoxia by

repressing the iron-sulfur cluster assembly proteins ISCU1/2. *Cell Metab*, **10**, 273-284.

11. Yang, W., Sun, T., Cao, J., Liu, F., Tian, Y. and Zhu, W. (2012) Downregulation of miR-210 expression inhibits proliferation, induces apoptosis and enhances radiosensitivity in hypoxic human hepatoma cells in vitro. *Experimental cell research*, **318**, 944-954.

12. Kim, H.W., Haider, H.K., Jiang, S. and Ashraf, M. (2009) Ischemic preconditioning augments survival of stem cells via miR-210 expression by targeting caspase-8-associated protein 2. *J Biol Chem*, **284**, 33161-33168.

13. Fasanaro, P., D'Alessandra, Y., Di Stefano, V., Melchionna, R., Romani, S., Pompilio, G., Capogrossi, M.C. and Martelli, F. (2008) MicroRNA-210 modulates endothelial cell response to hypoxia and inhibits the receptor tyrosine kinase ligand Ephrin-A3. *J Biol Chem*, **283**, 15878-15883.

14. Rothé, F., Ignatiadis, M., Chaboteaux, C., Haibe-Kains, B., Kheddoumi, N., Majjaj, S., Badran, B., Fayyad-Kazan, H., Desmedt, C., Harris, A.L. *et al.* (2011) Global microRNA expression profiling identifies MiR-210 associated with tumor proliferation, invasion and poor clinical outcome in breast cancer. *PLoS ONE*, **6**, e20980.

15. Volinia, S., Galasso, M., Sana, M.E., Wise, T.F., Palatini, J., Huebner, K. and Croce, C.M. (2012) Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. *Proc Natl Acad Sci USA*, **109**, 3024-3029.

16. Jung, E.J., Santarpia, L., Kim, J., Esteva, F.J., Moretti, E., Buzdar, A.U., Di Leo, A., Le, X.F., Bast, R.C., Jr., Park, S.T. *et al.* (2012) Plasma microRNA 210 levels correlate with sensitivity to trastuzumab and tumor presence in breast cancer patients. *Cancer*, **118**, 2603-2614.

17. Hwang, H.W., Wentzel, E.A. and Mendell, J.T. (2009) Cell-cell contact globally activates microRNA biogenesis. *Proc Natl Acad Sci USA*, **106**, 7016-7021.

18. Fang, Y., Sullivan, R. and Graham, C.H. (2007) Confluence-dependent resistance to doxorubicin in human MDA-MB-231 breast carcinoma cells requires hypoxia-inducible factor-1 activity. *Exp Cell Res*, **313**, 867-877.

19. Tredan, O., Galmarini, C.M., Patel, K. and Tannock, I.F. (2007) Drug resistance and the solid tumor microenvironment. *J Natl Cancer Inst*, **99**, 1441-1454.

20. Grantab, R., Sivananthan, S. and Tannock, I.F. (2006) The penetration of anticancer drugs through tumor tissue as a function of cellular adhesion and packing density of tumor cells. *Cancer research*, **66**, 1033-1039.

21. Garrido, C., Ottavi, P., Fromentin, A., Hammann, A., Arrigo, A.P., Chauffert, B. and Mehlen, P. (1997) HSP27 as a mediator of confluence-dependent resistance to cell death induced by anticancer drugs. *Cancer research*, **57**, 2661-2667.

22. Dimanche-Boitrel, M.T., Genne, P., Duchamp, O. and Chauffert, B. (1994) Confluence dependent resistance (CDR) to doxorubicin and E-cadherin expression in murine mammary cells. *Cancer Lett*, **85**, 171-176.

23. Dimanche-Boitrel, M.T., Pelletier, H., Genne, P., Petit, J.M., Le Grimellec, C., Canal, P., Ardiet, C., Bastian, G. and Chauffert, B. (1992) Confluence-dependent resistance in human colon cancer cells: role of reduced drug accumulation and low intrinsic chemosensitivity of resting cells. *International journal of cancer. Journal international du cancer*, **50**, 677-682.

Chapter 4

Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity

The following text has been modified from a manuscript published in Genome Research in 2011. Figure numbers have been updated to reflect the formatting of the dissertation. Supplemental figures and tables are available with open access at: http://genome.cshlp.org/content/21/9/1450/suppl/DC1

The data from this study have been submitted to the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under accession no. GSE26970.

I would also like to acknowledge the contributions of our co-authors:

Wyman SK*, **Knouf EC***, Parkin RK, Fritz BR, Lin DW, Dennis LM, Krouse MA, Webster PJ, Tewari M. Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Res.* **21**, 1450-1461 (2011).

Abstract

Modification of microRNA sequences by the 3' addition of nucleotides to generate so-called “isomiRs” adds to the complexity of miRNA function, with recent reports showing that 3' modifications can influence miRNA stability and efficiency of target repression. Here we show that the 3' modification of miRNAs is a physiological and common post-transcriptional event that shows selectivity for specific miRNAs and is observed across species ranging from *C. elegans* to human. The modifications result predominantly from adenylation and uridylation, and are seen across tissue types, disease states, and developmental stages. To quantitatively profile 3' nucleotide additions, we developed and validated a novel assay based on NanoString Technologies' nCounter platform. For certain miRNAs, the frequency of modification was altered by processes such as cell differentiation, indicating that 3' modification is a biologically regulated process. To investigate the mechanism of 3' nucleotide additions, we used RNA interference to screen a panel of eight candidate miRNA nucleotidyl transferases for 3' miRNA modification activity in human cells. Multiple enzymes, including MTPAP, PAPD4, PAPD5, ZCCHC6, ZCCHC11, and TUT1, were found to govern 3' nucleotide addition to miRNAs in a miRNA-specific manner. Three of these enzymes–MTPAP, ZCCHC6 and TUT1–have not previously been known to modify miRNAs. Collectively, our results indicate that 3' modification observed in next generation small RNA sequencing data is a biologically relevant process, and identify enzymatic mechanisms that may lead to new approaches for modulating miRNA activity *in vivo*.

Introduction

MicroRNAs (miRNAs) are small non-coding RNAs that play a regulatory role in gene function by binding to the 3' untranslated region (UTR) of target mRNAs. Although the biogenesis and turnover of miRNAs has been the subject of much recent study (1,2), mechanisms of post-transcriptional regulation of mature miRNAs are still not well understood. A few recent reports have suggested that post-transcriptional addition of nucleotides to the 3' end of miRNAs is a mechanism for regulation of miRNA activity. Studies in plants and *C. elegans* have identified examples in which such modifications influence miRNA stability (2-5). In humans, effects on miRNA stability and on mRNA target repression have both been observed for specific miRNAs. For example, miR-122 was shown to be adenylated by the RNA nucleotidyl transferase PAPD4 (also known as GLD-2) in humans and mice, resulting in an increase in the stability of the miRNA (6). In contrast, miR-26a was shown to be uridylated by the RNA nucleotidyl transferase ZCCHC11, which had no effect on miRNA stability but rather reduced the ability of miR-26a to inhibit its mRNA target (7). Taken together, these studies indicate that understanding the prevalence and mechanisms of miRNA 3' nucleotide additions will be important for understanding the post-transcriptional regulation of miRNA function, and could also inform strategies for therapeutic modulation of miRNA activity.

We and others have observed additions of nucleotides to the 3' end of miRNAs on a global scale in animal miRNA high-throughput sequencing studies (8-14). For example, in our own high-throughput sequencing of small RNAs from human embryonic stem cells (hESC) and ovarian cancers (9,10), we noticed substantial sequence variation at the 3' end of the known miRNAs, frequently manifesting as nucleotide additions not matching

the human genome sequence (i.e., 3' non-templated additions (3' NTA)). However, it was not clear to what extent these were artifacts of cDNA library preparation as opposed to physiologic nucleotide additions. More recently, such miRNA variants resulting from 3' NTA (so-called “isomiRs”) have been described in more detail in small RNA next generation sequencing studies in humans, mice, and flies (15-28). In one recent study (19), the enzyme PAPD4 was implicated as a miRNA nucleotidyl transferase, based on the finding that PAPD4 knockdown led to a decrease in the abundance of multiple miRNA variants bearing a non-templated 3' adenosine. However, not all miRNA variants were affected and the extent of reduction was highly variable across miRNAs, suggesting that other enzymes may also be responsible for miRNA additions. Thus, the mechanisms of 3' nucleotide additions on a global scale are still not well understood, even though the biological impact of such post-transcriptional modifications may be substantial.

In the present study, we systematically tested the extent to which miRNA 3' NTA observed in sequencing datasets are physiologic additions versus artifacts of cDNA library preparation, and analyzed the additions on a global scale using multiple second generation sequencing (i.e., 454 and Illumina) datasets in human (9,10,29,30), mouse (31), and *C. elegans* (32). To further investigate the 3' end additions using an orthogonal platform ideally suited for quantitative analysis and high-throughput processing of samples, we adapted the nCounter Gene Expression Assay from NanoString Technologies (33) to discriminate between and provide quantitative expression profiles of miRNA 3' variants. The nCounter assay involves the hybridization of fluorescently-labeled, barcoded probes to the miRNAs of interest, which are then scanned and counted to quantify miRNA expression. We used the nCounter assay to profile 3' end nucleotide

additions during cell differentiation, which revealed that specific miRNAs undergo a change in 3' NTA with differentiation, suggesting that 3' NTA is a physiologically regulated process. We also conducted a systematic siRNA-based knockdown screen of RNA nucleotidyl transferases, using the nCounter assay as a readout, which identified multiple enzymes responsible for miRNA 3' additions in human cells. Our results demonstrate the prevalence and mechanism of miRNA additions in humans, and introduce a new platform capable of quantitatively profiling miRNA 3' NTA variants across a variety of biological samples.

Methods

Preparation of samples and library construction.

Preparation and sequencing of the cDNA libraries corresponding to the synthetic miRNA pool and normal prostate tissue small RNAs in this paper has already been described (29,30). Briefly, 473 human and 40 non-human miRNAs (4 *A. thaliana*, 26 *C. elegans*, and 10 *O. sativa*) were synthesized and pooled at equal concentrations of 2.14 nM to generate a synthetic miRNA library (29). The 40 non-human miRNAs at varying concentrations (50 nM to 190.7 fM) were 5’ phosphorylated by T4 polynucleotide kinase (Invitrogen) and spiked with radiolabeled 18-24 nt markers and gel purified. The resulting synthetic pools were spiked into normal prostate RNA derived from two individuals and the 18-24 nt fraction was gel purified. Normal prostate RNA was from histologically-confirmed normal prostate tissue obtained with informed consent under IRB supervision at the time of radical prostatectomy for clinically localized prostate cancer. Linker ligation and amplification of both the synthetic miRNA pool and prostate

RNA samples were performed as described in (34). Small RNA cDNA libraries were sequenced on the Genome Analyzer (Illumina) according to the manufacturer's instructions. Small RNA cDNA libraries for 454 sequencing from cultured prostate epithelial and stromal cells were generated as described previously (9). RNA from prostate epithelial cell and prostate stromal cell cultures was a generous gift from B. Knudsen. Small RNA cDNA libraries were sequenced on the 454 platform according to the manufacturer's instructions.

Processing of the sequencing data

Processing of the sequencing data was performed using a set of custom bioinformatics scripts (detailed in **Supplemental Methods**). Briefly, sequence reads were pruned of any relevant linker sequences and then compared to canonical miRNA sequences from miRBase (35-37). Any reads that exactly matched the canonical miRNA sequence (with or without additional nucleotides) contributed to the total matching reads (TMR) count and any with additional 3' nucleotides that did not match the precursor sequence contributed to the non-templated addition (NTA) count.

NanoString nCounter miRNA assay

Total RNA or synthetic miRNA pools (IDT; 30 pmol per oligonucleotide) were used as input for nCounter miRNA sample preparation reactions. All sample preparation was performed according to manufacturer's instructions with the following exception: for assays examining end-variant expression, nCounter Human miRNA Tag Reagent (which contains a mixture of tags and bridges, see **Figure 4.5**) was replaced with a custom

reagent containing the same nCounter Human miRNA tags mixed with a novel pool of bridge oligos, each designed to template the ligation of a particular 3' end variant to a specific tag (3' end variants are given in **Supplemental Table S6**). Two such pools were generated; only one variant for each miRNA was present in each pool, and the variants of a given miRNA species were always ligated to the same tag. Following ligation, sample preparation reactions were purified and diluted according to manufacturer's instructions. Hybridization reactions were performed according to manufacturer's instructions with 5 µl of the 5-fold diluted sample preparation reaction. All hybridization reactions were incubated at 65°C for a minimum of 18 h. Hybridized probes were purified and counted on the nCounter Prep Station and Digital Analyzer (NanoString) following the manufacturer's instructions. For each assay, a high-density scan (600 fields of view) was performed.

NanoString nCounter miRNA data analysis

For platform validation using synthetic oligonucleotides, NanoString nCounter miRNA raw data was normalized for lane-to lane variation with a dilution series of six spike-in positive controls. The sum of the six positive controls for a given lane was divided by the average sum across lanes to yield a normalization factor, which was then multiplied by the raw counts in each lane to give normalized values. The normalized counts for the synthetic oligonucleotide mixtures (canonical or variant) in each of the three bridge pools were used to determine the specificity of the assays. To assess accuracy, the relative abundance of the canonical, V1 and V2 counts for each miRNA in a mixture of five canonical miRNAs and their 3' variants (containing 60% canonical, 30% variant 1, 10%

variant 2) was determined. To assess the linear range of the assays, pools of the canonical, variant 1, or variant 2 RNA oligonucleotides were generated and diluted to yield three six-point standard curves. We assayed the ability of the three bridge pools to detect miRNA inputs ranging from 10^5 to 10^8 copies per reaction.

For the hESC and the nucleotidyl transferase knockdown datasets, NanoString nCounter miRNA raw data was normalized for lane-to-lane variation with a dilution series of six spike-in positive controls using the `vsn` R package (38). Using `vsn`, each lane of nCounter data is calibrated by an affine transformation, and then the data is transformed by a variance-stabilizing transformation. Once normalized, the data were filtered (described in detail for each dataset below), and then relative abundances of the canonical, V1 and V2 counts for each miRNA were calculated, representing the fraction of the total abundance that each of the three sequences contributed. Results from assays corresponding to four miRNAs were filtered out universally because of intrinsically high background (miR-16, miR-192 and miR-196a), or in one case because the assay design precluded distinguishing between two miRNA family members (miR-20a assay).

For the H1 hESC data, we had two biological replicates each for canonical, V1, and V2 for the undifferentiated hESC and the differentiated hESC. Using the averaged two lanes, the data were filtered to require a miRNA to have at least 50 counts in two of the three bridge pools (canonical, V1, or V2) for both the undifferentiated and the differentiated hESC sample. Ten miRNAs met these criteria. Differential addition of nucleotides was calculated based on differences in fractional abundance of the variants. *P*-values were calculated for each miRNA between the undifferentiated and differentiated samples using the biological replicates.

For the nucleotidyl transferase knockdown data, we had four biological replicates of our negative control (siCyclophilin) and two replicates of the nucleotidyl transferases knockdowns for each of the canonical, V1 and V2 lanes assayed. The data were filtered based on counts in the four siCyclophilin samples. If a miRNA had at least 50 counts in two of three assays (canonical, variant 1 or variant 2) for all four samples, it was retained. The counts were then converted to relative percentages of the canonical, variant 1 and variant 2 sequences. Relative percentages were compared between each of the knockdown samples (averaged across two lanes) and the four siCyclophilin lanes (averaged across four lanes). *P*-values were calculated for each miRNA comparing canonical, V1 and V2 results from the negative control siCyclophilin samples to the nucleotidyl transferase knockdown samples. The false discovery rate (FDR) was determined by the Benjamini-Hochberg method using the `multtest` R package (39,40).

RNA for stem cell differentiation experiment

RNA used was from undifferentiated H1 human embryonic stem cells, or from plates of H1 cells allowed to spontaneously differentiate (i.e., in a non-directed manner) for 7 or 9 days after the removal of fibroblast growth factor. The RNAs corresponded to aliquots from our earlier study in which the culture of the cells and RNA isolation was performed (9). 100 ng of RNA was used as input into the sample preparation reaction for the NanoString nCounter assay. Two biological replicates of undifferentiated and differentiated cells were run.

Cell culture, siRNA transfection, RNA isolation, and qRT-PCR

HCT-116 cells were maintained in McCoy's media (Gibco) with 10% FBS (Atlanta Biologicals Inc). Transfections were performed with Lipofectamine RNAiMax (Invitrogen) and 30 nM of ON-TARGET*plus* SMARTpool siRNAs (Dharmacon) directed against each enzyme of interest or the negative control gene, *Cyclophilin B*. Duplicate transfections for each enzyme knockdown, and quadruplicate transfections of siCyclophilin were performed. RNA was isolated with the miRNeasy RNA isolation kit (Qiagen) 72 hours post-transfection. cDNA was synthesized using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Expression analysis via qRT-PCR was performed with Taqman gene expression assays (Applied Biosystems), using expression of *GUSB* to normalize for variations in RNA input. 400 ng of RNA was used as input into the sample prep reaction for the NanoString nCounter assay.

Results

MicroRNA 3' non-templated nucleotide additions observed in next generation sequencing data are a physiologic phenomenon. We sought to determine the extent to which non-templated 3' addition of nucleotides to miRNAs observed in next generation small RNA sequencing datasets is a physiologic, *in vivo* phenomenon, as opposed to an artifact introduced during cDNA library preparation. We analyzed data from small RNA cDNA libraries prepared side-by-side and in triplicate from normal human prostate tissue RNA (30), and from a pool of 473 chemically synthesized human miRNAs (29), in order to compare 3' NTA frequency between the physiologically-derived versus chemically synthesized miRNAs (**Figure 4.1A**). Each of the libraries also had 40 non-human chemically synthesized miRNAs spiked in (listed in **Supplemental Methods**), with the

expectation that if miRNA nucleotide addition were an artifact of cDNA library preparation, the spiked-in synthetic miRNAs would undergo additions in both sets of libraries. Sequencing on the Illumina Genome Analyzer platform produced 4,946,602 high-quality reads from the triplicate synthetic miRNA library and 9,417,361 high-quality reads from the triplicate prostate library, which were then analyzed for 3' NTA using custom bioinformatics tools (details provided in **Supplemental Methods**).

We compared the synthetic miRNA sequencing data to the prostate tissue sequencing data, analyzing a set of 167 miRNAs that were shared between the two datasets and that passed filtering based on read abundance and inter-replicate variance criteria for these miRNAs (see **Supplemental Methods** for additional detail). For this set of 167 miRNAs, we calculated total matching reads (TMR) and percent non-templated nucleotide additions (%NTA) for each miRNA in both datasets. TMR is the sum of: (i) all reads that exactly match the canonical miRNA sequence from miRBase (35-37), plus (ii) any matching reads that have nucleotide additions, regardless of whether those additions are templated (i.e., match the human genome sequence and miRNA precursor sequence) or non-templated (i.e., do not match the human genome sequence and miRNA precursor sequence). %NTA is the fraction of TMR that has 3' nucleotide additions that do not match the precursor sequence of the miRNA (hence counting only non-templated 3' additions). We consider only non-templated additions here because our interest is in unambiguous identification of post-transcriptional modifications to mature miRNAs.

Figure 4.1B depicts the %NTA for the 167 shared, abundant miRNAs sorted by increasing %NTA in the prostate sequencing dataset. We observed a dramatic difference in the %NTA between the prostate tissue versus the chemically synthesized miRNA

datasets, with the synthetic miRNAs only rarely demonstrating nucleotide additions, while over half of the prostate miRNAs examined had some extent of 3' NTA. Notably, none of the 40 spiked-in synthetic non-human miRNAs showed any evidence of 3' additions in either the prostate or chemically synthesized miRNA datasets (**Figure 4.1B**). These data demonstrate that miRNA 3' non-templated nucleotide additions are predominantly a physiologic, *in vivo* process. In addition, we found no evidence of a relationship between miRNA abundance and the frequency of non-templated nucleotide addition (**Supplemental Fig. S1**), also consistent with additions being a physiologic process rather than an artifact of library preparation.

Survey of miRNA 3' NTA in multiple next generation sequencing datasets representing diverse tissue types and species. To more broadly characterize miRNA 3' NTA, we obtained and analyzed small RNA sequencing datasets representing a diverse set of tissues, disease states, and developmental stages (**Supplemental Table S1**). These included cultures of normal human ovarian surface epithelial (HOSE) cells and tissue from three histological subtypes of ovarian cancer (10); two prostate primary cell culture samples, one epithelial and one stromal; and undifferentiated and differentiated human embryonic stem cells (9). To assess the nature of miRNA 3' NTA across diverse species, we also examined miRNA 3' NTA in publicly available Illumina sequencing datasets from *M. musculus* (31) and *C. elegans* (32). For mouse, this consisted of 11 mouse cerebellum and medulloblastoma sequencing datasets representing diverse developmental stages, cell populations and genotypes. For *C. elegans*, we examined 12 small RNA sequencing datasets representing different developmental stages and genotypes. A

complete list of all sequencing datasets and the dataset identifiers used in this manuscript is given in **Supplemental Table S1**. Results of these analyses are presented below.

Nucleotide addition is miRNA specific. Analysis of miRNA sequencing datasets demonstrated that although 3' NTA occurs commonly, it is not universal and shows a strong predilection for specific miRNAs over others. In **Figure 4.2A**, histograms for three individual human datasets show that certain miRNAs have a very high frequency of additions (%NTA), while the majority of miRNAs are modified at a low frequency, if at all (histograms for six additional human datasets are shown in **Supplemental Fig. S2A**). A similar distribution of frequencies was observed when we analyzed 11 mouse cerebellum and medulloblastoma sequencing datasets representing several developmental stages and genotypes. The histogram for one of the 11 datasets analyzed is shown in **Figure 4.2B** and histograms for the other 10 are shown in **Supplemental Fig. S2B**. We also examined 12 *C. elegans* small RNA sequencing datasets representing different developmental stages and genotypes. In *C. elegans,* the overall frequency of 3' NTA observed was less than in human and mouse, but showed a similar distribution across sequencing datasets from seven developmental stages and multiple genotypes in which certain miRNAs show a high frequency of modification whereas the majority of miRNAs are infrequently or never modified. A representative histogram of miRNA %NTA for the L1 stage of *C. elegans* larval development is shown in **Figure 4.2C**, with histograms for all 12 datasets shown in **Supplemental Fig. S2C**.

To determine whether there are some miRNAs that consistently show a high degree of 3' NTA or consistently show an absence of it, we compared the %NTA of

miRNAs across nine human sequencing datasets from different cell lines, tissues, and disease states (**Supplemental Table S1**, datasets h1-h9). A set of 73 miRNAs was detected with 40 or more reads in at least five of the nine datasets. Among the set of 73, we observed that certain miRNAs are consistently frequently modified across diverse tissue and cell types, while others are consistently rarely modified. A boxplot of %NTA for the 10 most and 10 least modified of these 73 miRNAs is shown in **Figure 4.3A**. The 10 miRNAs which are most often highly modified in our human sequencing datasets are miR-100, -146a, -151-3p, -143, -331-3p, -23b, -24, -222, -199a-3p, -1308. **Supplemental Table S2** gives the total matching reads (TMR) and %NTA for all miRNAs in each of these nine datasets.

In mouse, there was a set of 81 miRNAs detected at 40 or more reads in at least five of the 11 sequencing datasets. The %NTA boxplot for the 10 most and 10 least modified of these 81 miRNAs is shown in **Figure 4.3B**. As observed in the human datasets, specific murine miRNAs were consistently highly modified across different biological samples. These miRNAs were only infrequently the mouse homologs of the most frequently modified human miRNAs, although this could be related to the expression of different sets of miRNAs between the mouse and human datasets analyzed. The TMR and %NTA for all miRNAs in each of the 11 mouse datasets is given in **Supplemental Table S3**. In *C. elegans*, we again saw that particular miRNAs were consistently among the most frequently modified miRNAs across developmental stages, as shown by the boxplot in **Figure 4.3C. Supplemental Table S4 gives** the TMR and %NTA for all miRNAs in each of the 12 *C. elegans* datasets**.**

3' NTA is predominantly the result of adenylation and uridylation in human, mouse and *C. elegans*. We next assessed specifically which nucleotides were added to generate 3' NTA miRNA variants. The fractional contribution of each type of addition (A, U, G, C and >1 nt) was calculated for each miRNA, and then averaged over all the miRNAs detected at an abundance of at least 10 reads in each dataset. **Figure 4.4** shows the distribution of nucleotide additions across the nine human sequencing datasets. We found that mono-adenylation (blue) or mono-uridylation (red) were by far the most common type of 3' NTA. To illustrate the distribution of additions on a per-miRNA basis, the complete list of nucleotide additions for abundant miRNAs (defined as 20 or more reads) in the normal prostate tissue sequencing dataset is given in **Supplemental Table S5.** Throughout the nine human datasets examined, the distribution of added nucleotides was remarkably uniform, with mono-adenylation typically accounting for approximately 50% of 3' NTA and mono-uridylation for approximately 25% of 3' NTA.

In mouse, we observed a similar distribution of predominant mono-adenylation and mono-uridylation (**Supplemental Fig. S3A**). In four of the mouse cerebellum sequencing datasets, however, an increased prevalence of the addition of a single guanine was observed (**Supplemental Fig. S3A**, datasets m3-m6 in **Supplemental Table S1**). These datasets were from specific mouse embryonic developmental stages (which were not represented in the human datasets) and were largely accounted for by three miRNAs (mmu-miR-16, mmu-miR-103 and mmu-miR-181a) that, although they were observed have an added guanine, had a low overall %NTA. Nonetheless, the results raise the possibility that although adenylation and uridylation are the most common modes of 3'

NTA, in specific contexts and for specific miRNAs additional types of 3' NTA may be activated, perhaps with different biological consequences. Across a panel of mouse tissues, Chiang *et al*. reported finding uridine to be the mostly commonly added nucleotide (21), which was calculated by summing all the reads with added nucleotides for each of the four individual nucleotides for *all* of the miRNAs, and then calculating the percent from the sum. When the fractions of added nucleotides for the Chiang *et al.* manuscript are re-calculated using our method described above, we found that adenosine is the most frequently added nucleotide for the mouse brain sequencing datasets of Chiang *et al*.

In small RNA sequencing datasets corresponding to diverse developmental stages of *C. elegans*, we again observed mono-adenylation and mono-uridylation to be the most common forms of 3' NTA **(Supplemental Fig. S3B)**. Interestingly, however, the predominant form of modification in *C. elegans* was mono-uridylation (~40%) followed by mono-adenylation (~25%) compared to primarily mono-adenylation (~40-50%) and secondly mono-uridylation (~25%) in human and mouse (**Supplemental Fig. S3A,B, Figure 4.4**). This suggests that although we found miRNA 3' nucleotide addition to be an evolutionarily conserved phenomenon, the degree to which specific mechanisms such as adenylation and uridylation are employed for generating 3' miRNA variants may differ depending upon the species.

Development of a quantitative platform to profile relative abundance of miRNA variants resulting from 3' nucleotide additions. In order to profile 3' non-templated nucleotide additions to miRNAs in a more quantitative and high-throughput manner than

is currently practical by next generation sequencing, we developed a novel version of NanoString's nCounter miRNA Expression Assay. NanoString's hybridization-based technology uses color-coded, fluorescently labeled probe pairs to barcode selected target molecules, which are then scanned and counted (33). The complexity of the barcodes, each containing one of four colors in each of six positions, allows a large diversity of target molecules present in the same sample to be individually distinguished during data collection. The NanoString miRNA assay uses an additional sample processing step to allow the detection of small RNAs (**Figure 4.5A**). Preparation of the small RNA samples involves the ligation of a specific DNA tag onto the 3' end of each mature miRNA; these tags are designed to normalize the melting temperatures of the targeted miRNA as well as to provide a unique identification for each miRNA species in the sample. Tagging is accomplished in a multiplexed ligation reaction using reverse-complement bridge oligonucleotides to direct the ligation of each miRNA to its designated tag. Following the ligation reaction, excess tags and bridges are removed, and the resulting material is hybridized with a panel of miRNA:tag-specific nCounter capture and barcoded reporter probes (**Figure 4.5B**). Following purification, each captured barcode is counted and tabulated in the nCounter assay (33).

The standard nCounter miRNA assay detects only those miRNAs with canonical 3' ends, as the tag ligation is a 3' end-specific event. We modified the standard assay to be able to measure two 3' variants each (arbitrarily designated variant 1 and variant 2). To measure the miRNA variants, we created pools of bridges that direct the tagging of the variant miRNA to the same tag as the canonical miRNA. Each sample is split in three and assayed separately with the canonical bridge pool, the variant 1 pool, and the variant 2

pool. By labeling each 3' variant of a given miRNA species with the same tag in three parallel assays, we eliminate any potential tag-to-tag variation and ensure a direct comparison of the relative levels of each miRNA variant. A titration of six synthetic miRNA spike-ins is used in each assay as a control for assay-to-assay variation.

In initial pilot studies, we assayed a set of five synthetic miRNAs (miR-15a, miR-15b, miR-125a-5p, miR-143 and miR-221) and two variants for each miRNA. We validated the specificity of the platform by assaying three pools of synthetic RNA oligonucleotides representing canonical, variant 1, and variant 2 versions of our five pilot miRNAs (**Figure 4.5C**). Each bridge pool (i.e., specific to canonical, variant 1 or variant 2) was used to assay the three mixtures of synthetic miRNA oligonucleotides (**Figure 4.5C**). We found that each bridge pool distinguished the miRNA species of interest with high specificity, with minimal background detection of the other variants.

In order to assess accuracy, we also assayed mixtures of synthetic miRNAs containing 60% canonical miRNA, 30% variant 1, and 10% variant 2 for the set of five miRNAs. In four of the five cases, the ratios of canonical to variant miRNAs observed in the assay matched closely the known composition of the synthetic mixtures (**Figure 4.5D**). Finally, to determine the linear range of the platform, we generated six-point standard curves for miR-15a and its two 3' variants. Each standard curve was assayed in its respective bridge pool, with between $1x10^5$ and $1x10^8$ input copies of miRNA per reaction **(Figure 4.5E)**. We found that all three assays (canonical, variant 1, and variant 2) yielded a strong, linear response to increases in concentration across our entire range of input values. These pilot studies demonstrate that the nCounter miRNA 3' variant assay provides a quantitative platform for studying miRNA 3' NTA.

Following our pilot experiments, we expanded the bridge pools to assay a set of 132 human miRNAs. The list of 132 miRNAs and their corresponding variants chosen for assay development is given in **Supplemental Table S6**. The miRNAs were chosen on the basis of showing an appreciable frequency of 3' NTA in our next generation small RNA sequencing datasets from human samples (**h1-h9, Supplemental Table S1**), as well as several miRNAs of special biological interest because of disease association or tissue specificity. The specific 3' NTA variants for analysis were chosen largely on the basis of being the most common variants for specific miRNAs observed in the human small RNA sequencing datasets.

The frequency of 3' NTA to specific miRNAs changes with differentiation of human embryonic stem cells, suggesting that post-transcriptional nucleotide addition is a physiologically regulated process in humans. To investigate whether miRNA 3' additions are a physiologically regulated process, we determined whether the frequency of miRNA 3' NTA varies with a change in biological state, using differentiation of embryonic stem cells as a model system. We employed the NanoString nCounter assay to profile 3' NTA in undifferentiated H1 human embryonic stem cells versus spontaneously differentiated cells derived from the same cell line. To ensure that we restricted our analysis to miRNAs and miRNA variants for which robust measurements were available, we filtered the data to consider only the miRNAs for which at least 50 counts were reliably obtained in at least two out of three bridge pools (i.e., canonical, variant 1, or variant 2) in both of the datasets (i.e., the undifferentiated and differentiated hESC datasets). This comprised 10 miRNAs and their corresponding variants (given in

Supplemental Table S7). We found that the fractional abundance of five miRNA variants changed significantly after differentiation ($p < 0.05$ in at least one variant assay for a given miRNA) (**Supplemental Table S7**). Depending on the specific miRNA, cell differentiation could be associated with either an increase or a decrease in variant abundance. The two most marked examples were miR-1246 (differentiation was associated with a significant increase in the fractional abundance of one of the variants and decrease in that of the canonical sequence; **Figure 4.6A)** and miR-455-3p (differentiation was associated with a decrease in fractional abundance of both the variants and an increase in that of the canonical sequence; **Figure 4.6B**). These results suggest that non-templated nucleotide additions to different miRNAs could be mediated by unique mechanisms. More importantly, our results demonstrate that miRNA 3' additions in human cells are dynamic and can vary in response to biological stimuli, indicating that 3' NTA is a physiologically regulated process.

Identification of nucleotidyl transferases responsible for miRNA 3' NTA. Previous studies have suggested that members of the RNA nucleotidyl transferase family, including PAPD4, PAPD5, and ZCCHC11, can mediate miRNA 3' additions in humans (6,7,19). However, RNA nucleotidyl transferases comprise a much larger family of enzymes, and the existing data on these three enzymes do not fully account for the widespread 3' NTA of miRNAs observed on a genomic scale. We hypothesized that additional nucleotidyl transferases are involved in mediating miRNA 3' non-templated nucleotide additions, and that even for PAPD4, PAPD5, and ZCCHC11 additional miRNA substrates may exist beyond ones that have been identified to date. To

understand the mechanism of 3' NTA to a range of different miRNAs, we used RNA interference to individually suppress expression of eight different nucleotidyl transferases in HCT-116 colon cancer cells, followed by quantitative analysis of canonical and variant miRNAs using the nCounter miRNA variant analysis platform developed above. These eight candidate miRNA 3' modifying enzymes represent all but four of the 12 known RNA nucleotidyl transferases in humans (41). These specific enzymes were chosen based on a high likelihood of having specific effects (e.g., poly(A) polymerase was excluded because of expected broad effects on polyadenylation), evidence for robust expression in HCT-116 cells (**Supplemental Table S8**), and availability of effective reagents for RNA interference. As a negative control, we performed RNA interference directed against Cyclophilin B (a commonly used negative control in siRNA experiments), which is not expected to affect miRNA nucleotide addition.

We confirmed the success of knockdown with quantitative reverse transcription-PCR (qRT-PCR) **(Figure 4.7A)**, which demonstrated at least 75% reduction in expression of each enzyme compared to cells transfected with negative control siRNA targeting Cyclophilin B**.** We also used qRT-PCR to assess the specificity of the nucleotidyl transferase knockdowns by measuring all eight nucleotidyl transferases in each knockdown experiment (**Supplemental Fig. S4**). Overall, we found the knockdown of each enzyme to be highly specific (i.e., knockdown of a given nucleotidyl transferase was not associated with a concomitant decrease in other nucleotidyl transferases; **Supplemental Fig. S4**). Interestingly, in a few cases knockdown of one enzyme was associated with an increase in expression of one or more other nucleotidyl transferase(s) (e.g., knocking down PAPD4 led to an increase in PAPD5, while TUT1 and MTPAP

knockdown led to an increase in PAPOLG), raising the possibility that compensation of function by one enzyme for another may occur in some cases. This is worth mention because it suggests that our results may underestimate the full effect of knockdown of individual nucleotidyl transferases on 3' NTA, because of such cases of potential compensatory increase in expression of other enzymes.

We used the NanoString platform to monitor changes in miRNA 3' NTA fractional abundance as a result of the suppression of each RNA nucleotidyl transferase. MicroRNAs were filtered by retaining only those that had a count of at least 50 for the canonical in all four of the negative control (i.e., siCyclophilin) knockdown lanes. We calculated the percentage of each variant miRNA as a fraction of the total abundance of the miRNA (i.e., counts obtained for canonical plus variant 1 counts plus variant 2 counts), which ensured that any decrease in a given variant observed with knockdown of specific nucleotidyl transferases would reflect a change in propensity for nucleotide addition, rather than being due to an overall decrease in all forms of the miRNA as a result of decreased transcription, for example. The relative percentages of each canonical and variant species found in the knockdown cells were then compared to negative control cells treated with siRNAs targeting Cyclophilin B. A *t*-test conducted across groups representing biological replicates was performed, followed by an adjustment for multiple hypothesis testing based on the false discovery rate (FDR) method of Benjamini and Hochberg (39,40) to identify nucleotidyl transferase knockdowns associated with a significant reduction (FDR < 10%) of specific miRNA 3' NTA variants. The most compelling cases were considered to be those knockdowns in which a statistically significant decrease in at least one 3' NTA variant was observed along with a

concomitant statistically significant increase in the fractional abundance of the canonical miRNA sequence.

We found that suppression of seven out of eight nucleotidyl transferases led to a significant reduction in at least one miRNA 3' variant (FDR < 10%). In addition to PAPD4, PAPD5, and ZCCHC11, which have previously been described as affecting miRNA additions (6,7,19), we found significant reductions in miRNA variants after suppression of MTPAP, PAPOLG, TUT1, and ZCCHC11 (**Supplemental Table S9)**. To highlight nucleotidyl transferases showing the most substantial effect on 3' NTA of specific miRNAs, we selected cases in which effects of enzyme knockdown on specific miRNA variants were both statistically significant (FDR < 10%) and had at least a 5% reduction in the fractional abundance of the variant **(Figure 4.7B)**. Loss of PAPD4 resulted in the most dramatic changes in miRNAs with 3' A additions: six miRNAs with a 3' A addition showed significant decreases, with some miRNA variants such as miR-92a and let-7b showing greater than 70% reduction **(Figure 4.7B).** Suppression of PAPD5 led to significant reductions in four miRNA 3' variants which all featured a 3' A, and one miRNA variant, miR-1246, with a dinucleotide 3' GA **(Figure 4.7B)**. The variant miR-1246 with a 3' GA also showed a significant reduction with the MTPAP knockdown. While both PAPD4 and PAPD5 suppression predominantly affected variants with a 3' A, knockdown of two enzymes (TUT1 and ZCCHC6) was associated with a decrease in 3' uridylation. Suppression of TUT1 led to a significant reduction in the variant of miR-200a with a 3' U, and loss of ZCCHC6 decreased expression of let-7e with a 3' U **(Figure 4.7B)**, which suggests that different enzymes may show specificity for certain nucleotide

additions. Taken together, our results suggest multiple enzymes affect the post-transcriptional modification of miRNAs by 3' NTA.

To validate the accuracy of the assays in **Figure 4.7B**, we tested the ability of our adapted nCounter miRNA variant profiling assay to distinguish between the canonical and variant forms of the miRNAs from **Figure 4.7B.** We profiled pools of chemically synthesized versions of the canonical and variant miRNAs of interest and again, the assay demonstrated excellent discrimination between miRNA 3' end nucleotide variants (**Supplemental Fig. S5**). We also tested the linear range of these miRNA assays by creating pools of synthetic miRNAs (containing the canonical or variant version of each miRNA) and generating 6-point standard curves assaying from $1x10^5$ to $1x10^8$ input copies of each miRNA or miRNA 3' variant per reaction. We assayed each dilution point of the standard curves in the appropriate bridge pool, and found that all assays showed a strong, linear response to changes in concentration (**Supplemental Fig. S6**).

Discussion

Our study provides confirmation on a broad scale that miRNA 3' non-templated nucleotide additions that have been observed in small RNA next generation sequencing data are a physiologic and biologically-regulated phenomenon in human cells, complementing and extending work presented in other reports (16,18-21,25-28). We find that across multiple species, these modifications are miRNA-specific and that adenylation and uridylation are the primary modes of 3' NTA, which is consistent with the previously described specificities of known RNA nucleotidyl transferase enzymes in their modification of other RNA substrates (41). We attempted to identify specific sequence

features in miRNAs that might govern modification. However, we did not find any motifs or sequence characteristics that showed significant association with modifications (data not shown). It is of interest, however, that the degree to which adenylation versus uridylation is employed for miRNA modification varies across species, with uridylation predominating among *C. elegans* miRNAs whereas adenylation is more prominent in mouse and human. Although the basis for this difference is not yet known, we speculate that cross-species differences in expression patterns of nucleotidyl transferases and their miRNA substrates may underlie this observation.

The development of a novel, bridged ligation-based assay permitted us to quantitatively profile 3' NTA variants (along with their corresponding canonical miRNA species) following knockdown of a panel of candidate miRNA nucleotidyl transferases in order to investigate the mechanism of 3' NTA. Prior studies had primarily implicated two nucleotidyl transferases, PAPD4 and ZCCHC11, and to a lesser extent, PAPD5, in miRNA 3' NTA (6,19). Strikingly, we found that knockdown of seven of the eight RNA nucleotidyl transferases we examined (including PAPD4, ZCCHC11 and PAPD5) was associated with a significant reduction in 3' NTA of one or more miRNA variants. Individual enzymes tended to show specificity for particular miRNAs, and the involvement of multiple additional enzymes points toward much more complexity in mechanisms of miRNA 3' NTA than has previously been appreciated. One caveat of our analysis is that we have examined only the steady-state levels of 3’ NTA. Thus, the mechanisms by which each enzyme modulates the observed miRNA 3’ NTA frequency—whether it be via altered kinetics of 3’ NTA or effects on miRNA degradation—remain to be determined. Remarkably, MTPAP is known to localize to the

mitochondria and to polyadenylate the 3' ends of mitochondrial transcripts (42) and thus it was surprising to find that suppression of this heretofore mitochondria-specific enzyme affected miRNA additions. TUT1, on the other hand, was found to regulate both 3' A and 3' U additions, which supports previous reports that describe TUT1 both as a uridyltransferase and as a nuclear poly(A) polymerase (43,44). Additionally, we identified the first known role for ZCCHC6, an enzyme with unknown substrates that is homologous to the miRNA-modifying enzyme ZCCHC11 (45).

Our study identified four new miRNA nucleotidyl transferases (MTPAP, TUT1, ZCCHC6 and PAPOLG) that have significant effects on 3' NTA to specific miRNAs, supporting the notion that 3' NTA represents a multifaceted layer of post-transcriptional regulation for miRNAs by an ensemble of miRNA-modifying enzymes. Our study likely represents a minimal estimate of the enzymes and miRNA substrates involved, as there was evidence in several cases of a compensatory increase in the expression of other nucleotidyl transferases after knockdown of a given enzyme, which could have obscured some of the effects on 3' NTA. Additionally, some of the enzymes that did not show an effect may be false negatives as a result of incomplete suppression of protein levels in siRNA experiments. Importantly, the finding that 3' NTA is a regulated and miRNA-specific process mediated by distinct enzymes raises the possibility that inhibiting particular RNA nucleotidyl transferases could provide a new strategy for modulating miRNA activity for therapeutic as well as research purposes. In this vein, the enzymes we have identified may be worth investigating as targets for small molecules to potentially modify miRNA function via inhibiting (or potentially enhancing) 3' NTA. Nucleotidyl transferases may also show differential expression in different disease states. In a brief

survey of Oncomine (46), for example, we found ZCCHC6 to be in the top 1% of genes overexpressed in prostate carcinoma versus normal tissue (47). Thus, we speculate that differential expression of nucleotidyl transferases could impact miRNA 3' NTA in different biological states. In the future, quantification of 3' NTA variants in samples from varied biological sources using the NanoString assay developed here and/or other technologies will facilitate understanding of the potential biological effects of these enzymes.

Quantification of 3' NTA variants distinct from canonical miRNAs in biological samples poses a problem for current technologies. Although next generation sequencing can certainly identify 3' variants, it is not currently an assay for which parameters such as quantitative precision, quantitative accuracy, linear range of quantification or absolute quantification are well established. Additionally, the throughput and turnaround time for analysis of large numbers of biological or clinical samples is still limited in practice. Although quantitative reverse transcription-PCR (qRT-PCR) approaches are routinely used for miRNA quantification, currently available qRT-PCR platforms are not equipped to measure and discriminate between variants with high accuracy (24,27). The NanoString assay we have introduced here is an important advance for enabling further studies in this area, as it permits quantitative analysis of miRNAs and variants on a platform that has sufficient throughput and precision to be routinely used for analysis of large numbers of biological samples.

Although 3' NTA can clearly have functional effects on animal miRNAs by increasing miRNA stability in the case of hsa-miR-122 (6), or affecting the efficacy of mRNA target repression in the case of hsa-miR-26a (7), investigation of the biological

functions of miRNA 3' NTA is still at a nascent stage. Future studies will be required to determine if 3’ NTA affects the expression of miRNA targets or the stability of miRNAs on a global scale. A recent study has noted an increase in uridylated miR-223 during miRNA decay (48). Although our analysis has evaluated only the steady-state levels of miRNA additions, we speculate that 3’ NTA may serve as either a signal for or a consequence of targeting of miRNAs for degradation. The functional effects of 3' NTA on miRNAs may be varied and miRNA-specific, making such modifications a potentially versatile mechanism for generating functional diversity and complexity in an otherwise relatively limited miRNA transcriptome. Our study has identified a multitude of new miRNA variants along with the nucleotidyl transferases that may mediate their genesis. These findings, along with our introduction of a quantitative platform for miRNA variant analysis, should enable further studies of the biological roles of miRNA 3' NTA.

Acknowledgements

This work was supported by Chromosome Metabolism Training Grant 5 T32 CA09657-16 (S.K.W.), a Rosetta Inpharmatics Fellowship in Molecular Profiling (S.K.W.), an American Cancer Society/Canary Foundation Postdoctoral Fellowship (S.K.W.), a Public Health Service National Research Service Award (T32 GM07270) from the National Institute of General Medical Sciences (E.C.K.), a Jaconnette L. Tietze Young Scientist Award (M.T.), and a Damon Runyon-Rachleff Innovation Award (M.T.). We thank the Fred Hutchinson Genomic Shared Resource for sequencing services, Merav Bar and Beatrice Knudsen for providing RNA samples, and Jason Arroyo, Ingrid Ruf and Lynn Amon for helpful advice and discussions.

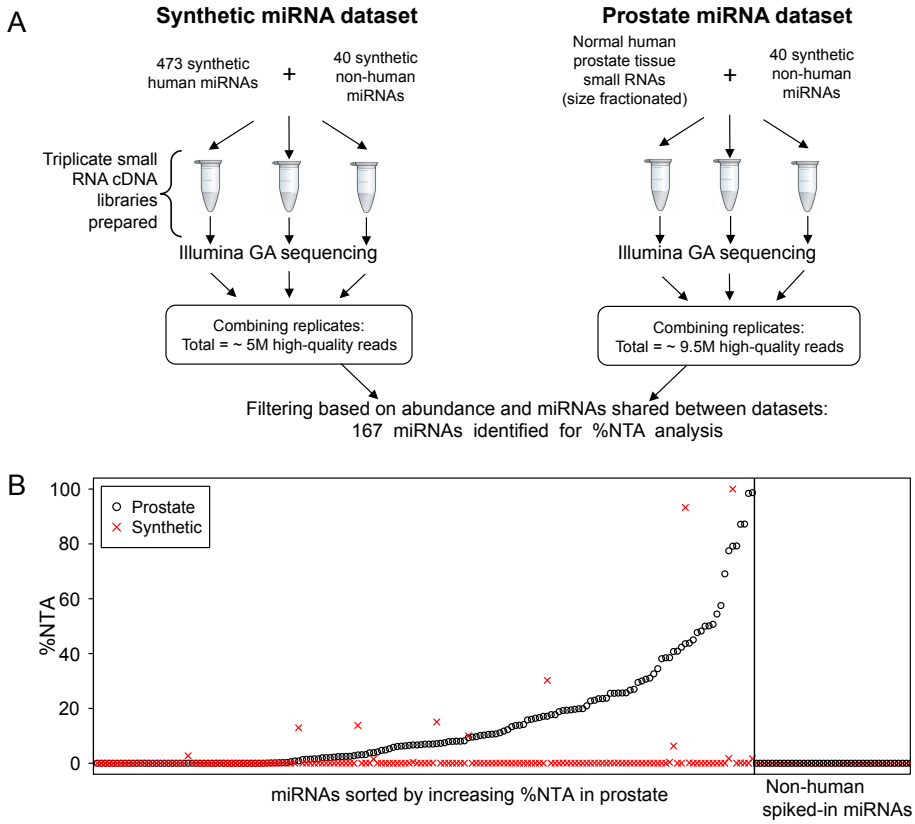

Figure 4.1. %NTA observed in Illumina Genome Analyzer sequencing of a pool of chemically synthesized miRNAs or of prostate tissue-derived miRNAs.
(A) The schematic describes the workflow for the generation and sequencing of triplicate cDNA libraries corresponding to a pool of chemically synthesized miRNAs and normal human prostate tissue. **(B)** (*Left*) The %NTA for 167 shared miRNAs present at 10 or more reads in both the chemically synthesized miRNA and prostate sequencing datasets are plotted along the y-axis. Normal prostate tissue data is plotted in black and the synthetic miRNA data in red. MicroRNAs are sorted along the x-axis with respect to increasing %NTA observed in prostate tissue sequencing data. Only seven of the 167 miRNAs analyzed demonstrated greater %NTA in the synthetic miRNA pool compared to its prostate tissue %NTA. In most cases, these represented an additional nucleotide identical to the terminal nucleotide expected for the canonical miRNA sequence (data not shown), which is consistent with a typical type of error expected with solid-phase RNA oligonucleotide synthesis. (*Right*) The %NTA for 40 non-human (a mix of plant and *C. elegans*) spiked-in miRNAs is plotted for both normal prostate tissue (black) and the synthetic miRNA (red) sequencing datasets.

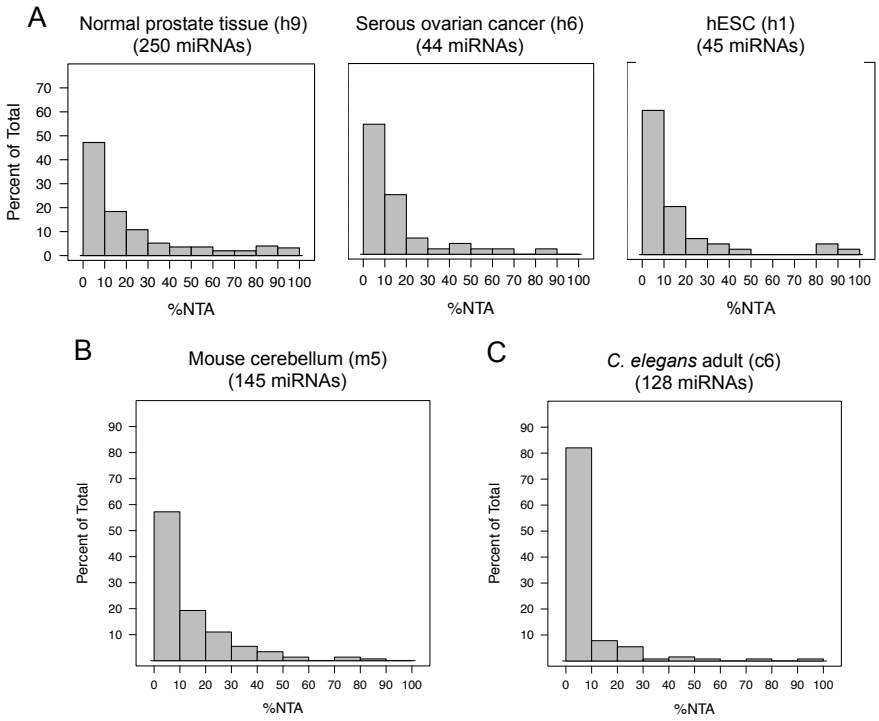

Figure 4.2. 3' non-templated nucleotide additions are miRNA specific.
(A) Histograms of %NTA observed in three representative human miRNA sequencing datasets (normal prostate tissue (h9), serous ovarian cancer (h6), and human embryonic stem cells (h1)). MicroRNAs were required to have 10 or more reads in a given dataset to be included, and the number of miRNAs that qualified is given in parentheses for each histogram. The y-axis scale is the same for all three human histograms. (**B-C)** Histograms of %NTA for representative datasets corresponding to mouse cerebellum or medulloblastoma (panel B, m5 in **Supplemental Table S1**) and the L1 developmental stage of *C. elegans* (panel C**,** c6 in **Supplemental Table S1**). MicroRNAs were required to have 10 or more reads in a given dataset to be included. Dataset identifiers are provided in parentheses for all histograms and refer to descriptions of these datasets provided in **Supplemental Table S1.**

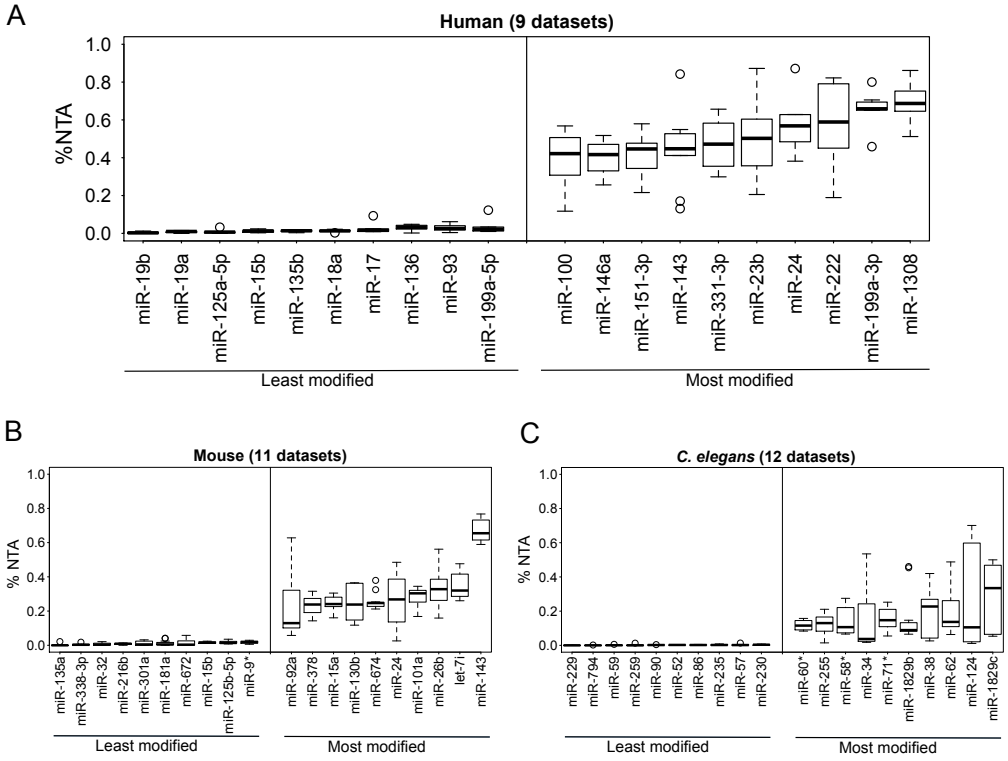

Figure 4.3. MicroRNA-specific 3' non-templated nucleotide addition is seen across small RNA sequencing datasets.
(A) A boxplot of the miRNAs demonstrating the lowest %NTA (i.e., least modified) and highest %NTA (i.e., most modified) across nine human sequencing datasets (datasets h1-h9 in **Supplemental Table S1**). MicroRNAs were required to have at least 40 reads in five or more of the 9 datasets. (**B-C**) Boxplots of the miRNAs demonstrating the lowest %NTA (i.e., least modified) and highest %NTA (i.e., most modified) across 11 mouse cerebellum or medulloblastoma datasets (panel **B**, datasets m1-m11 in **Supplemental Table S1**) and 12 *C. elegans* sequencing datasets (panel **C**, datasets c1-c12 in **Supplemental Table S1**). MicroRNAs were required to have at least 40 reads in five or more of the datasets for mouse and for *C. elegans* to be included in the analysis.

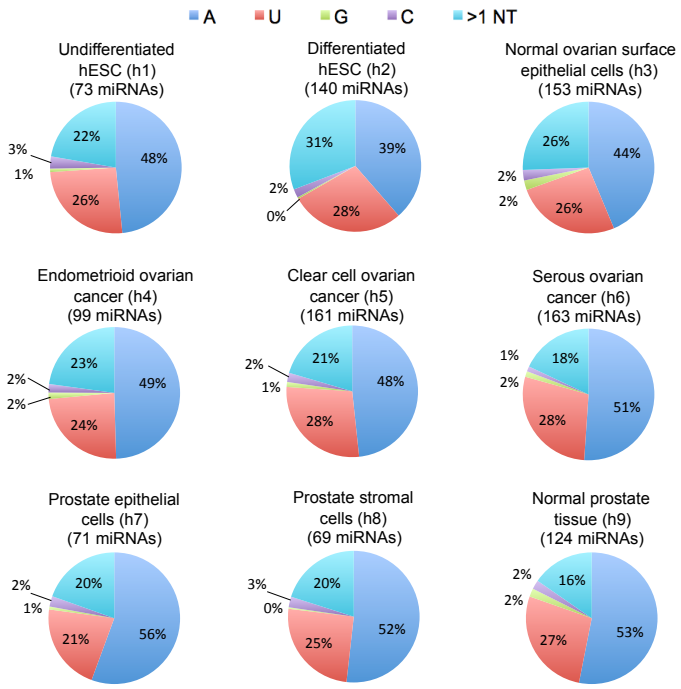

Figure 4.4. Distribution of added nucleotides observed in nine human small RNA sequencing datasets.
Distribution was calculated based on miRNAs with an abundance of at least 10 reads, and the number of miRNAs which qualified is given in parentheses for each pie chart. Fractional contribution of each nucleotide addition was calculated for each miRNA, and then averaged for all the miRNAs in each dataset. Pie charts indicate what fraction of total additions each mono-nucleotide represented (on average) for each dataset, and all multi-nucleotide additions were counted in the >1 category. Parenthetical dataset identifiers refer to descriptions provided in **Supplemental Table S1**.

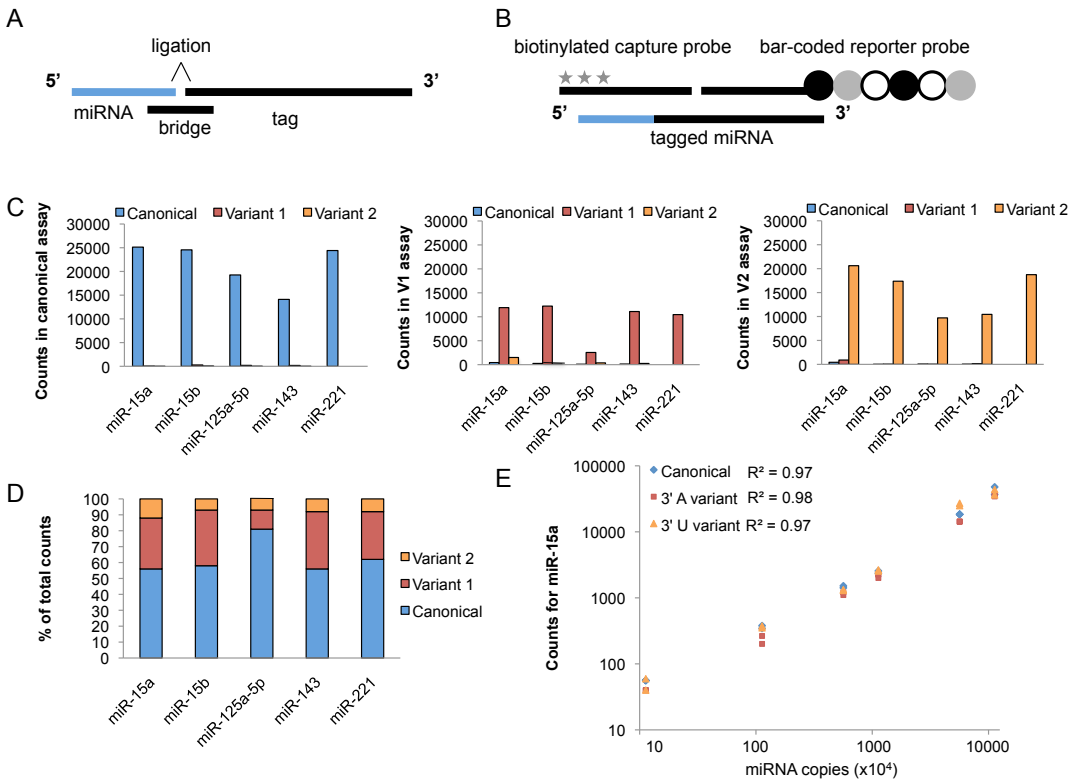

Figure 4.5. Development of the nCounter miRNA assay to detect miRNA 3' variants. **(A)** A schematic of the molecular components of the nCounter miRNA assay. The miRNA is shown in gray and DNA oligonucleotides in black. Each bridge oligonucleotide serves to template the 3' end ligation of a particular miRNA species to a sequence-specific tag. Custom bridges were designed to discriminate between the 3' end miRNA variants. **(B)** Following the removal of the excess tags and bridges, the tagged miRNAs are hybridized to specific capture and reporter probes attached to unique barcodes. The captured barcodes are individually resolved and counted in the NanoString nCounter assay system. The six-spot fluorescent barcode is represented by circles and the biotin capture moieties by stars. **(C)** Validation of the specificity of the nCounter assay. Three pools of synthetic RNA oligonucleotides, each containing canonical, variant 1, or variant 2 versions of five miRNAs were assayed. The graphs display the counts resulting when each of the three mixtures was individually assayed using the canonical (*left*), variant 1 (*center*), and variant 2 (*right*) bridge pools. **(D)** Validation of the accuracy of the nCounter assay. A mixture containing 60% canonical, 30% variant 1, and 10% variant 2 chemically synthesized miRNAs were assayed in each bridge pool. The relative abundance of each variant measured in the assay was then determined. **E.** Validation of the linear range of the nCounter assays for miR-15a. Standard curves of synthetic oligonucleotides corresponding to the canonical, 3' A (variant 1), or 3' U (variant 2) form of miR-15a were assayed in their appropriate bridge pool. The graph displays the counts resulting from technical duplicates of standard curves ranging from $1x10^5$ to $1x10^8$ input copies of miRNA per reaction.

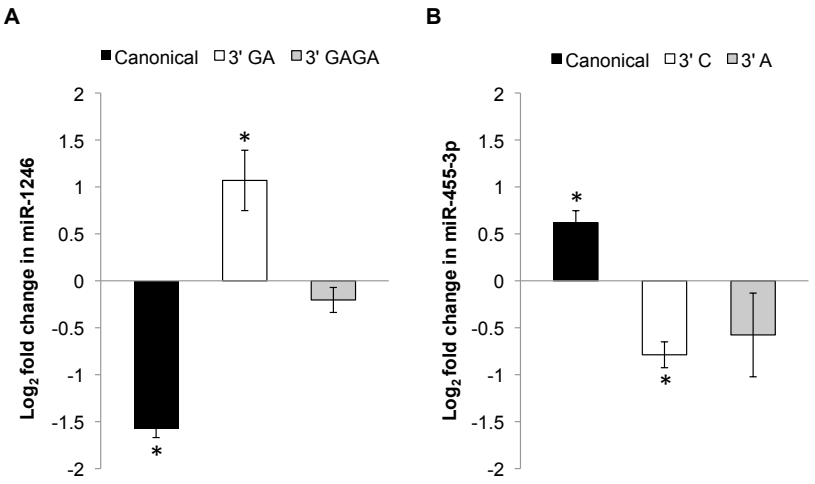

Figure 4.6. MicroRNA additions are altered in response to differentiation.
The NanoString nCounter assay was used to profile differential expression of miRNA 3' variants in undifferentiated versus differentiated H1 human embryonic stem cells. MicroRNAs were filtered by requiring at least 50 counts in at least two of the three assays (canonical, V1 or V2) in both the undifferentiated and differentiated hESC datasets. The plotted miRNAs are the two showing the greatest change in the fractional abundance of 3' variants following differentiation. Error bars represent ± SD of the fold change across two biological replicates and asterisks indicate p-values < 0.05. **Supplemental Table S7** provides data corresponding to all ten miRNAs that met filtering criteria, including three additional miRNAs showing significant changes in relative variant abundance. (**A)** The graph shows the fold change in the canonical and 3' variant forms of miR-1246. For miR-1246, differentiation is associated with a significant increase in the fractional abundance of the 3' GA variant, and a corresponding decrease in the canonical sequence. (**B)** For miR-455-3p, differentiation is associated with a significant decrease in the abundance of the 3' C variant and a corresponding increase in the canonical miRNA fractional abundance.

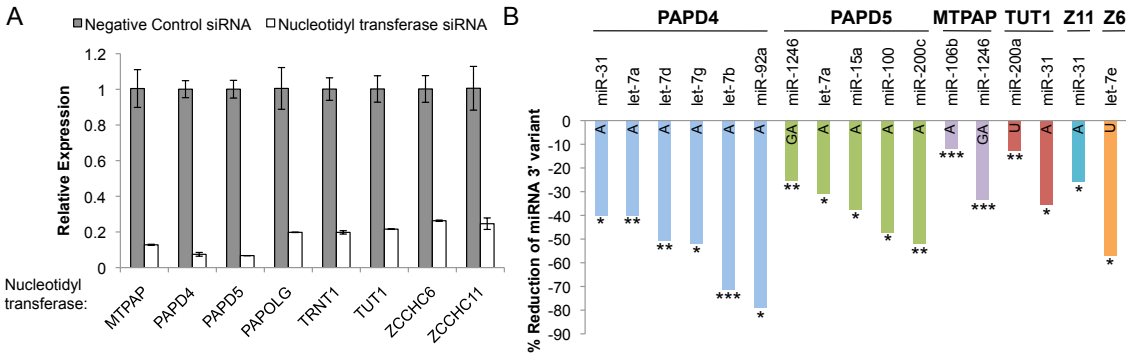

Figure 4.7. MicroRNA 3' non-templated nucleotide additions are regulated by multiple nucleotidyl transferase enzymes.
(A) qRT-PCR demonstrates the successful suppression of eight different nucleotidyl transferases in HCT-116 cells transfected with siRNAs individually targeting each enzyme. Bars represent mean expression ± SD of biological replicates relative to cells transfected with siRNAs against the Cyclophilin B negative control. Normalization was performed using qRT-PCR quantification of the endogenous control gene *GUSB* to control for variations in RNA input. (**B)** Bar graphs for each enzyme show percent reduction in miRNA 3' variants following suppression of each of the six nucleotidyl transferases indicated above the x-axis. MicroRNAs shown represent variants that 1) were significantly reduced (FDR < 10%) in the cells treated with siRNAs targeting the enzyme of interest versus those treated with the negative control, siCyclophilin and 2) showed at least a 5% absolute decrease in fractional abundance of the variant miRNA. The letter(s) within each bar for a given miRNA identifies which 3' variant displays significant reduction. Z11 refers to the enzyme ZCCHC11 and Z6 to ZCCHC6. *indicates miRNAs meeting criteria of FDR < 10%, **represents FDR < 5%, and ***represents FDR < 2%. The false discovery rate (FDR) is determined using the method of Benjamini and Hochberg. The change in fractional abundance and associated FDR for all canonical and 3' miRNA variants retained after filtering for abundance is given in **Supplemental Table S9**.

References

1. Krol, J., Loedige, I. and Filipowicz, W. (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet*, **11**, 597-610.

2. Chatterjee, S. and Großhans, H. (2009) Active turnover modulates mature microRNA activity in Caenorhabditis elegans. *Nature*, **461**, 546-549.

3. Lu, S., Sun, Y.H. and Chiang, V.L. (2009) Adenylation of plant miRNAs. *Nucleic Acids Res*, **37**, 1878-1885.

4. Ibrahim, F., Rymarquis, L.A., Kim, E.J., Becker, J., Balassa, E., Green, P.J. and Cerutti, H. (2010) Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. *Proc Natl Acad Sci U S A*, **107**, 3906-3911.

5. Ramachandran, V. and Chen, X. (2008) Degradation of microRNAs by a family of exoribonucleases in Arabidopsis. *Science*, **321**, 1490-1492.

6. Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., Baba, T. and Suzuki, T. (2009) Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev*, **23**, 433-438.

7. Jones, M.R., Quinton, L.J., Blahna, M.T., Neilson, J.R., Fu, S., Ivanov, A.R., Wolf, D.A. and Mizgerd, J.P. (2009) Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat Cell Biol*, **11**, 597-610.

8. Berezikov, E., Thuemmler, F., van Laake, L.W., Kondova, I., Bontrop, R., Cuppen, E. and Plasterk, R.H. (2006) Diversity of microRNAs in human and chimpanzee brain. *Nat Genet*, **38**, 1375-1377.

9. Bar, M., Wyman, S.K., Fritz, B.R., Qi, J., Garg, K.S., Parkin, R.K., Kroh, E.M., Bendoraite, A., Mitchell, P.S., Nelson, A.M. *et al.* (2008) MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. *Stem Cells*, **26**, 2496-2505.

10. Wyman, S.K., Parkin, R.K., Mitchell, P.S., Fritz, B.R., O'Briant, K., Godwin, A.K., Urban, N., Drescher, C.W., Knudsen, B.S. and Tewari, M. (2009) Repertoire of microRNAs in epithelial ovarian cancer as determined by next generation sequencing of small RNA cDNA libraries. *PLoS One*, **4**, e5311.

11. Berezikov, E., van Tetering, G., Verheul, M., van de Belt, J., van Laake, L., Vos, J., Verloop, R., van de Wetering, M., Guryev, V., Takada, S. *et al.* (2006) Many novel mammalian microRNA candidates identified by extensive cloning and RAKE analysis. *Genome Res*, **16**, 1289-1298.

12. Cummins, J.M., He, Y., Leary, R.J., Pagliarini, R., Diaz, L.A., Jr., Sjoblom, T., Barad, O., Bentwich, Z., Szafranska, A.E., Labourier, E. *et al.* (2006) The colorectal microRNAome. *Proc Natl Acad Sci U S A*, **103**, 3687-3692.

13. Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M. *et al.* (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*, **129**, 1401-1414.

14. Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H. and Bartel, D.P. (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell*, **127**, 1193-1207.

15. Ahn, H.W., Morin, R.D., Zhao, H., Harris, R.A., Coarfa, C., Chen, Z.J., Milosavljevic, A., Marra, M.A. and Rajkovic, A. (2010) MicroRNA transcriptome in the newborn mouse ovaries determined by massive parallel sequencing. *Mol Hum Reprod*, **16**, 463-471.

16. Fernandez-Valverde, S.L., Taft, R.J. and Mattick, J.S. (2010) Dynamic isomiR regulation in Drosophila development. *RNA*, **16**, 1881-1888.

17. Guo, L. and Lu, Z. (2010) Global expression analysis of miRNA gene cluster and family based on isomiRs from deep sequencing data. *Comput Biol Chem*, **34**, 165-171.

18. Martí, E., Pantano, L., Bañez-Coronel, M., Llorens, F., Miñones-Moyano, E., Porta, S., Sumoy, L., Ferrer, I. and Estivill, X. (2010) A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing. *Nucleic Acids Res*, **38**, 7219-7235.

19. Burroughs, A.M., Ando, Y., de Hoon, M.J., Tomaru, Y., Nishibu, T., Ukekawa, R., Funakoshi, T., Kurokawa, T., Suzuki, H., Hayashizaki, Y. *et al.* (2010) A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome Res*, **20**, 1398-1410.

20. Morin, R.D., O'Connor, M.D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A.L., Zhao, Y., McDonald, H., Zeng, T., Hirst, M. *et al.* (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res*, **18**, 610-621.

21. Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E. *et al.* (2010) Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev*, **24**, 992-1009.

22. Ruby, J.G., Stark, A., Johnston, W.K., Kellis, M., Bartel, D.P. and Lai, E.C. (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. *Genome Res*, **17**, 1850-1864.

23. Pantano, L., Estivill, X. and Marti, E. (2010) SeqBuster, a bioinformatic tool for the processing and analysis of small RNAs datasets, reveals ubiquitous miRNA modifications in human embryonic cells. *Nucleic Acids Res*, **38**, e34.

24. Wu, H., Neilson, J.R., Kumar, P., Manocha, M., Shankar, P., Sharp, P.A. and Manjunath, N. (2007) miRNA profiling of naive, effector and memory CD8 T cells. *PLoS ONE*, **2**, e1020.

25. Fehniger, T.A., Wylie, T., Germino, E., Leong, J.W., Magrini, V.J., Koul, S., Keppel, C.R., Schneider, S.E., Koboldt, D.C., Sullivan, R.P. *et al.* (2010) Next-generation sequencing identifies the natural killer cell microRNA transcriptome. *Genome Res*, **20**, 1590-1604.

26. Kuchenbauer, F., Morin, R.D., Argiropoulos, B., Petriv, O.I., Griffith, M., Heuser, M., Yung, E., Piper, J., Delaney, A., Prabhu, A.L. *et al.* (2008) In-depth characterization of the microRNA transcriptome in a leukemia progression model. *Genome Res*, **18**, 1787-1797.

27. Lee, L.W., Zhang, S., Etheridge, A., Ma, L., Martin, D., Galas, D. and Wang, K. (2010) Complexity of the microRNA repertoire revealed by next-generation sequencing. *RNA*, **16**, 2170-2180.

28. Berezikov, E., Robine, N., Samsonova, A., Westholm, J.O., Naqvi, A., Hung, J.H., Okamura, K., Dai, Q., Bortolamiol-Becet, D., Martin, R. *et al.* (2011) Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. *Genome Res*, **21**, 203-215.

29. Linsen, S.E., de Wit, E., Janssens, G., Heater, S., Chapman, L., Parkin, R.K., Fritz, B., Wyman, S.K., de Bruijn, E., Voest, E.E. *et al.* (2009) Limitations and possibilities of small RNA digital gene expression profiling. *Nat Methods*, **6**, 474-476.

30. Pomerantz, M.M., Beckwith, C.A., Regan, M.M., Wyman, S.K., Petrovics, G., Chen, Y., Hawksworth, D.J., Schumacher, F.R., Mucci, L., Penney, K.L. *et al.* (2009) Evaluation of the 8q24 prostate cancer risk locus and MYC expression. *Cancer Res*, **69**, 5568-5574.

31. Baek, D., Villén, 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.

32. Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S. *et al.* (2008) PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. *Mol Cell*, **31**, 67-78.

33. Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T. *et al.* (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol*, **26**, 317-325.

34. Mitchell, P.S., Parkin, R.K., Kroh, E.M., Fritz, B.R., Wyman, S.K., Pogosova-Agadjanyan, E.L., Peterson, A., Noteboom, J., O'Briant, K.C., Allen, A. *et al.* (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A*, **105**, 10513-10518.

35. Griffiths-Jones, S. (2004) The microRNA Registry. *Nucleic Acids Res*, **32**, D109-111.

36. Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A. and Enright, A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*, **34**, D140-144.

37. Griffiths-Jones, S., Saini, H.K., van Dongen, S. and Enright, A.J. (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res*, **36**, D154-158.

38. Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A. and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, **18 Suppl 1**, S96-104.

39. van der Laan, M.J., Dudoit, S. and Pollard, K.S. (2004) Augmentation procedures for control of the generalized family-wise error rate and tail probabilities for the proportion of false positives. *Stat Appl Genet Mol Biol*, **3**, Article15.

40. Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Royal Stat. Soc. B*, **57**, 289-300.

41. Martin, G. and Keller, W. (2007) RNA-specific ribonucleotidyl transferases. *RNA*, **13**, 1834-1849.

42. Tomecki, R., Dmochowska, A., Gewartowski, K., Dziembowski, A. and Stepien, P.P. (2004) Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase. *Nucleic Acids Res*, **32**, 6001-6014.

43. Mellman, D.L., Gonzales, M.L., Song, C., Barlow, C.A., Wang, P., Kendziorski, C. and Anderson, R.A. (2008) A PtdIns4,5P2-regulated nuclear poly(A) polymerase controls expression of select mRNAs. *Nature*, **451**, 1013-1017.

44. Trippe, R., Sandrock, B. and Benecke, B.J. (1998) A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. *Nucleic Acids Res*, **26**, 3119-3126.

45. Rissland, O.S., Mikulasova, A. and Norbury, C.J. (2007) Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol Cell Biol*, **27**, 3612-3624.

46. Rhodes, D.R., Kalyana-Sundaram, S., Mahavisno, V., Varambally, R., Yu, J., Briggs, B.B., Barrette, T.R., Anstet, M.J., Kincead-Beal, C., Kulkarni, P. *et al.* (2007) Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia*, **9**, 166-180.

47. Tomlins, S.A., Mehra, R., Rhodes, D.R., Cao, X., Wang, L., Dhanasekaran, S.M., Kalyana-Sundaram, S., Wei, J.T., Rubin, M.A., Pienta, K.J. *et al.* (2007) Integrative molecular concept modeling of prostate cancer progression. *Nat Genet*, **39**, 41-51.

48. Baccarini, A., Chauhan, H., Gardner, T.J., Jayaprakash, A.D., Sachidanandam, R. and Brown, B.D. (2011) Kinetic Analysis Reveals the Fate of a MicroRNA following Target Regulation in Mammalian Cells. *Curr Biol*, **21**, 369-376.

Chapter 5

The regulation of miRNA abundance by nucleotidyl transferases and 3' non-templated nucleotide additions

Abstract

Post-transcriptional modifications of several miRNAs with 3' non-templated nucleotide additions have been demonstrated to modulate their targeting and stability. However, the global effects of these 3' additions have not been examined in humans, and conserved signals that promote miRNA decay remain unknown. We previously reported that miRNA 3' additions are regulated by multiple nucleotidyl transferase enzymes. Here we examine the relationship between changes in addition and changes in abundance following the suppression of a panel of nucleotidyl transferases. We observed no global correlation between the magnitude of change in nucleotide addition and the change in miRNA abundance, indicating that 3' additions do not have uniform effects on miRNAs. However, the identity of the 3' nucleotide addition was associated with characteristic changes in abundance, as miRNAs with increased 3' uridylation showed significantly decreased expression. To determine whether nucleotidyl transferases yield enzyme-specific effects on miRNAs, we performed global miRNA profiling following enzyme suppression. Knockdown of two enzymes—PAPD4 and TUT1—led to decreased expression of many miRNAs, including numerous sequences that did not display altered 3' additions. The most substantial effects were observed following suppression of TUT1, which led to an extensive loss of miRNAs that was detected across multiple platforms

and cell lines. Our work demonstrates that a select group of nucleotidyl transferase enzymes broadly regulate the expression of miRNAs through mechanisms that may be dependent or independent of 3' nucleotide additions.

Introduction

MicroRNA expression levels are precisely regulated by multiple mechanisms, which enable miRNAs to serve as important regulators of many biological processes. Although recent progress has been made in understanding miRNA transcription and processing, the regulation of miRNA degradation is poorly understood in humans. Most studies report miRNAs to display remarkable stability, with reported half-lives for human miRNAs ranging from hours to several days (1,2). In a *D. melanogaster* model system, neuron miRNAs featured rapid turnover following transcriptional shutoff, indicating that tissue-specific factors may influence the decay process. (3). In plants and worms, exonucleases that degrade miRNAs have been identified (4,5), but in humans no global mechanisms for the decay or stabilization of miRNAs have been described.

One potential mechanism for miRNA stabilization is 3' nucleotide addition to mature miRNAs. In humans, miRNAs commonly feature non-templated additions of an adenosine (A) or uridine (U) at the 3' end of the molecule (6,7). Once thought to be an artifact of sequencing, 3' additions have now been demonstrated to be a conserved, biological phenomenon in humans, mice, plants, and flies (7-10). These miRNA isomiRs drastically expand the known miRNA transcriptome; however, the biological significance of 3' additions remains unclear. In plants, methylation and adenylation rendered miRNAs resistant to degradation (11,12). In contrast, uridylation by nucleotidyl transferase

enzymes led to increased degradation of miRNAs in both plants and algae (13-15). Taken together, these studies suggest that the modification of a miRNA with a single nucleotide addition can be sufficient to alter miRNA decay.

In humans, the role of 3' additions is poorly understood, as only a handful of reports have examined the effects of A and U additions on the stability of a small number of miRNAs. The addition of a 3' A to the miRNA miR-122 increased the stability of the miRNA (16). Conversely, the addition of a 3' U to miR-26a did not change the abundance of the miRNA, but instead prevented the miRNA from repressing its mRNA target (17). MicroRNA isomiRs are found in association with the Ago proteins, although recent evidence suggests that nucleotide additions may reduce miRNA association with two out of the four Ago proteins in humans (18). Thus, 3' additions may have miRNA-specific effects on controlling either the targeting or stability of a miRNA.

Non-templated 3' additions are regulated by multiple members of the nucleotidyl transferase family in humans. We and others have found that the PAPD4 enzyme governs the 3' A additions of a broad panel of miRNAs (6,7). We also identified multiple other enzymes—including TUT1, PAPD5, MTPAP, ZCCHC11, and ZCCHC6—that affected 3' additions in a miRNA-specific manner (6), which raises the possibility that 3' modifications by different enzymes may have distinct effects on miRNAs. In this study, we investigated the biological significance of miRNA 3' additions. We examined the relationships between the change in additions and change in total miRNA abundance, and did not find 3' additions to be either universally stabilizing or destabilizing. Instead, we found that certain changes in additions, such as increased 3' U addition, were associated with decreased miRNA abundance. We also found that two nucleotidyl transferases,

TUT1 and PAPD4, have broad effects on miRNA expression levels despite their seemingly restrictive regulation of 3' additions to specific miRNAs.

Methods

Cell culture and transfections

HCT-116 cells were obtained from Bert Vogelstein and maintained in McCoy's media (Invitrogen) with 10% FBS (Atlanta Biologicals). A549 and HeLa cells were obtained from ATTC and maintained in DMEM (Invitrogen) with 10% FBS. Transfections with biological replicates were performed in 6-well dishes with 62,500 cells transfected with 66 nM siRNA using Lipofectamine RNAiMax (Invitrogen) according to the manufacturer's instructions. Dharmacon On-Targetplus SMARTpool siRNAs were used for suppression of a panel of eight nucleotidyl transferases as previously described (6). SMARTpool siRNAs against the housekeeping gene *Cyclophilin B* was used as the negative control (Dharmacon). TUT1 siRNAs described by Mellman et al. (19) were synthesized by IDT with an unmodified backbone. Negative controls for these siRNAs included siCyclophilin-1 and non-targeting siRNAs (siAll-Stars) from Qiagen.

RNA isolation and qRT-PCR

72 hours post transfection with siRNAs, cells were washed with PBS and lysed in 700 uL Qiazol (Invitrogen). RNA was isolated using the miRNeasy RNA isolation kit (Qiagen), and normalized to equal concentration. Taqman assays from Applied Bioystems were used to confirm nucleotidyl transferase suppression. The expression of the endogenous control genes *GUSB* and *GAPDH* were used for normalization. For primary miRNA qRT-PCR profiling, Taqman assays were obtained from Applied Biosystems for 70

different primary miRNAs. qPCR primers and probes were diluted (1.25 μL water per 0.25 μL primer-probe mixture) and aliquoted into 384-well plates in an arrayed manner with duplicate wells per assay. For each sample of interest, a single 100 μL reverse transcription reaction was performed using the cDNA High Capacity Reverse Transcription Kit (Applied Biosystems) with 2.5 μg RNA. Each RT reaction was then diluted with 100 μL water and 500 μL 2X Master mix from Applied Biosystems. 3.5 μL of the diluted sample was added to each well of the 384-well plates, and amplification was performed on a Viia7 qPCR machine (Applied Bioystems). Technical replicates on each plate were averaged, and we removed any miRNA with an undetermined cycle threshold (CT) or with CT > 35 in any sample. This yielded 48 pri-miRNAs for subsequent analysis. Delta CT values were calculated using the average of three endogenous control genes run on each plate: GAPDH-FAM, ACTB-FAM, and GAPDH-VIC. Delta delta CT values to the siCyclophilin negative control cells were calculated, and relative expression values were determined.

NanoString analysis

For NanoString quantification of 3' additions, we performed additional analyses on data collected from previously described samples (6). MicroRNAs were filtered to include only those expressed with at least 50 counts (the total of both canonical and isomiR sequences) for our NanoString abundance analyses.

miRNA qRT-PCR arrays

Exiqon V2.0 Panel 1 Human miRNA arrays were used to assess global changes in miRNA levels. Reverse transcription was performed using the Universal cDNA Synthesis

kit (Exiqon) using equal amounts of RNA within each sample set. qPCR arrays were amplified with a Viia7 instrument (Applied Biosystems), and the data were imported into a single gene expression study and thresholded to 0.2 across all cards. CT values were filtered to maintain only reliably detected miRNAs (requiring CT < 35 across all samples). Delta CT values were then calculated relative to cells treated with negative control siRNAs, and relative expression values were computed.

Results

Nucleotidyl transferases modulate miRNA 3' additions

Multiple nucleotidyl transferase enzymes affect the 3' additions of miRNAs in an enzyme and miRNA specific manner. To assay miRNA 3' additions in a quantitative manner, we employed the nCounter miRNA platform from NanoString Technologies, which we previously adapted to assess the expression of approximately 130 canonical miRNAs and their two most common 3' variants (6). We have shown that suppression of seven different nucleotidyl transferase enzymes in HCT-116 colon cancer cells led to the significant reduction of specific miRNA 3' additions, with most enzymes mediating 3' A additions (6). We also found that multiple 3' additions are increased in response to enzyme suppression (**Figure 5.1**). Surprisingly, knockdown of the enzyme TUT1 led to significant increases in additions in 11 out of the 24 miRNAs that yielded robust isomiR expression with the NanoString platform. The diversity of additions added to the miRNAs was high. In contrast to the preponderance of 3' A additions that were lost following enzyme suppression, we saw significant increases in 3' U, UA, UU, and other additions (**Figure 5.1**). These results indicate that nucleotidyl transferases have variable

effects on miRNAs, with potentially some members of the family promoting and other members preventing 3' nucleotide additions.

Increased 3' U additions are associated with significant changes in miRNA abundance

The relationship between 3' additions and miRNA abundance has not been described on a global scale. Because 3' additions can modulate miRNA stability in other organisms, we hypothesized that these additions would affect the overall abundance levels of miRNAs in humans. We examined the expression of miRNAs that displayed increases and decreases in 3' additions following our suppression of nucleotidyl transferases. We predicted that miRNAs showing changes in additions would also display changes in miRNA abundance. However, when we compared the magnitude of change in miRNA 3' isomiRs with the change in total miRNA abundance, we found no correlation (**Figure 5.2A**). To determine whether gaining versus losing an addition would lead to distinct changes in miRNA expression, we also examined the overall abundance changes in each category of miRNA. MicroRNAs with increased additions following enzyme suppression showed a significant decrease in abundance compared to miRNAs with unchanged or decreased 3' addition (**Figure 5.2B**). However, as these categories of miRNAs included multiple nucleotide additions, we asked whether certain nucleotides would yield characteristic changes in abundance. To address these potential nucleotide-specific effects, we compared the abundance changes following the significant loss or gain of each nucleotide individually (**Figure 5.2C**). We found substantial decreases in the expression of miRNAs featuring increased 3' U addition, with most miRNAs showing a reduction of nearly 50% compared to miRNAs with unchanged additions ($p < 0.05$, t-test).

Increases in 3' A and 3' UA additions also led to significant reductions in miRNA abundance. Together, these results demonstrate that certain 3' additions are associated with consistent changes in miRNA abundance.

Enzyme-specific effects on miRNA abundance

We have used the suppression of a panel of nucleotidyl transferase enzymes to create perturbation in the pattern of miRNA 3' additions. Thus, we were also curious whether the nucleotidyl transferases that regulate these additions would have distinctive effects on miRNA expression patterns. We thus examined the changes in miRNA abundance resulting from suppression of each enzyme individually (**Figure 5.3**). Suppression of both TUT1 and PAPD4 led to decreased abundance of most miRNAs with altered 3' additions. However, TUT1 suppression led to nearly uniform increases in additions, while PAPD4 suppression yielded decreased additions of miRNAs **(Figure 5.3**). Suppression of the five other nucleotidyl transferases resulted in smaller changes in abundance, with the direction of the change varying across miRNAs. Together, these results suggest that 3' additions do not have uniform effects on miRNA stability; instead, changes in abundance are influenced by which enzyme is responsible for the addition.

Nucleotidyl transferases have broad effects on miRNA abundance levels

Having observed that several nucleotidyl transferases affect both the additions and expression of certain miRNAs, we sought to determine whether these enzymes would affect miRNA abundance more broadly. We profiled miRNA expression patterns following the individual suppression of a panel of eight nucleotidyl transferases in HCT-116 cells. To detect the total abundance of both canonical and isomiR miRNA species,

we used Exiqon qRT-PCR miRNA arrays to profile biological replicates for each enzyme. We predicted that a subset of miRNAs would feature significant changes, while miRNAs rarely modified by 3' additions would show unchanged expression levels. Instead, we found that suppression of the majority of the enzymes had very minimal effects on miRNA abundance (**Figure 5.4A, Table 5.1**). In contrast, when we suppressed PAPD4 and TUT1, we found substantial decreases in miRNA expression compared to cells treated with negative control siRNAs **(Figure 5.4A)**. The magnitude of decrease was modest, as the median expression of miRNAs in the siTUT1 cells was 0.60 and in the siPAPD4 cells was 0.71 compared to control cells (**Table 5.1**); yet, the decrease was seen in the vast majority of miRNAs that we examined (**Figure 5.4B)**, and yielded highly significant p-values in paired t-tests comparing the expression of each miRNA in enzyme suppressed versus control cells (paired t-test, $p < 10^{-20}$). Together, these results indicate that two nucleotidyl transferases—TUT1 and PAPD4—have broad effects on miRNA expression levels that extend beyond their effects on 3' additions.

To ensure that these results were not an artifact of the qPCR platform, we examined the global changes in miRNA abundance observed with the NanoString platform in an independent sample set of HCT-116 cells transfected with siRNAs against the nucleotidyl transferase enzymes. The NanoString platform, which does involve amplification, detected 65 miRNAs (expressed at greater than 50 counts) across all samples. In this subset of miRNAs, we found that TUT1 suppression produced substantial decreases in overall miRNA abundance, and that suppression of PAPD4 and MTPAP also led to modest miRNA loss (**Figure 5.5**). These results suggest that multiple

nucleotidyl transferases broadly regulate miRNA expression levels, with the TUT1 enzyme yielding the most significant effects.

TUT1 as a novel regulator of miRNA expression

To verify that the reductions in miRNA abundance seen upon nucleotidyl transferase suppression were not due to off-target effects of our siRNAs, we repeated these experiments with two additional TUT1 siRNAs that have previously been shown to potently suppress the TUT1 protein (19). Again we found that suppression of TUT1 led to decreased miRNA expression compared to cells treated with negative control siRNAs (**Figure 5.6**). We next tested whether TUT1 suppression would yield the same phenotype in additional cell lines. In A549 lung carcinoma cells, we suppressed TUT1 and used miRNA qRT-PCR arrays to quantify changes in miRNA abundance. We again found that suppression of TUT1 resulted in widespread decreases in miRNA expression levels (**Figure 5.7A**). We did not see a change in miRNA abundance following suppression of ZCCHC11, another nucleotidyl transferase, which confirms that this result is specific to the suppression of TUT1. We have also expanded this work to HeLa cells, which displayed decreased miRNA abundance following TUT1 loss, although the magnitude of the decrease was not as great as seen in HCT-116 and A549 cells (**Figure 5.7B**). In this cell line, suppression of ZCCHC11 decreased miRNA abundance to a similar level as seen after the suppression of TUT1. Together, these experiments demonstrate that TUT1 regulates miRNA levels across multiple cell lines, although the magnitude of the effect may potentially vary based on tissue-specific contributions of the various nucleotidyl transferase enzymes.

Potential mechanisms of TUT1

TUT1 could potentially affect miRNA levels via indirect or direct mechanisms. TUT1 has previously been described to act in the nucleus as a poly (A) and (U) polymerase for a variety of substrates. While TUT1 was originally described as exclusively modifying the U6 small nuclear RNA with a 3' U, more recently TUT1 (also called Star-PAP) was shown to regulate the cleavage and polyadenylation of a subset of mRNA transcripts (19,20). These two reports describe roles for TUT1 in the nucleus, which led us to hypothesize that TUT1 could potentially affect miRNA primary transcript abundance. Using qRT-PCR arrays to measure the expression of a panel of 70 primary transcripts, we found that TUT1 suppression did not lead to decreased expression of primary transcripts (**Figure 5.8A**), suggesting that TUT1 affects later stages of miRNA biogenesis. Because of the breadth of the TUT1-mediated effects, we also considered whether TUT1 could indirectly regulate miRNAs by affecting the expression of their processing enzymes. Yet, we found no change in the expression of *Dicer*, *Drosha*, and the Drosha co-factor *DGCR8* following TUT1 suppression (**Figure 5.8B**). Therefore, while the mechanism of TUT1 remains unknown, our preliminary evidence suggests that TUT1 may act on the mature miRNA to maintain miRNA homeostasis in the cell.

Discussion

Our work investigates the functional effects of miRNA 3' non-templated additions by nucleotidyl transferases on a broad scale. We found that 3' additions are associated with non-uniform effects on miRNA abundance, suggesting that miRNA post-transcriptional modifications have more complex roles in humans than have been

described in model organisms. While future *in vivo* assays for miRNA turnover will be needed to directly address the differential stability of miRNA isomiRs, these experiments suggest that 3' nucleotide modifications may affect miRNA abundance levels. miRNAs show substantial length heterogeneity, which raises the question of how single nucleotide additions can trigger variable rates of degradation. Future work to address the enzyme responsible for miRNA turnover in humans will enable a better understanding of the substrate specificity that may allow 3' additions to promote or prevent miRNA degradation.

We have also examined the role of nucleotidyl transferases in regulating miRNA abundance levels. This study expands previous investigations that examined a single nucleotidyl transferase, PAPD4 (7), to include analyses of a panel of enzymes we found to exert unique effects on miRNA additions. We have assayed miRNA expression levels with multiple platforms, including the NanoString platform that does not require sample amplification, to ensure quantitative detection of modest changes in miRNA levels. Our work suggests that multiple nucleotidyl transferases influence miRNAs through both 3' addition dependent and independent mechanisms. In particular, we identified two enzymes, TUT1 and PAPD4, that showed inverse effects on miRNA additions despite their very similar effects on miRNA abundance: We found that suppression of TUT1 led to the significant increases in additions in half of the miRNAs we examined, while suppression of the enzyme PAPD4 uniformly led to decreases in miRNA additions. However, individual knockdown of TUT1 and PAPD4 uniformly led to decreased expression of many miRNAs, demonstrating that global miRNA expression changes do not correlate with changes in 3' additions. We also identified miRNAs with unchanged

additions that nevertheless displayed dramatic changes in abundance. The poor association between additions and abundance raises several possibilities: nucleotidyl transferases may act indirectly to regulate miRNAs, or some miRNA isomiRs may exist only as transient intermediates not detectable by our platform.

Suppression of TUT1 yielded the most significant changes in miRNA expression levels. Across multiple cell lines, we found TUT1 suppression resulted in moderate but widespread decreases in abundance. To further assess the specificity of TUT1, we attempted to use the exogenous overexpression of TUT1 to rescue the miRNA repression phenotype seen upon introduction of TUT1 siRNAs. We generated a TUT1 overexpression construct in which synonymous mutations were introduced in the binding sites of two different TUT1 siRNAs. However, attempts at transient overexpression of TUT1 in multiple plasmid backbones were not successful (data not shown), which suggests that TUT1 expression may have deleterious effects. To circumvent these difficulties, we have relied upon our use of multiple independent siRNAs against TUT1 to assess the importance of TUT1 in controlling miRNA expression. It is also of note that out of the entire panel of nucleotidyl transferases examined, suppression of TUT1 resulted in the most substantial effects on miRNA levels. However, future studies are needed to address whether TUT1 acts through a direct or indirect manner. Our preliminary attempts to investigate the mechanism of TUT1 suggest that TUT1 acts downstream of miRNA transcription to modulate the level of the precursor or mature species. We have also ruled out the possibility that TUT1 regulates the expression of key miRNA processing enzymes. Because TUT1 suppression leads to increased miRNA addition, TUT1 may primarily act to remove or prevent 3’ additions by other nucleotidyl

transferases that could lead to miRNA degradation. Future investigations to determine if TUT1 associates with the Dicer or RISC complex, and whether mature miRNAs bind TUT1 will advance understanding of the mechanism of TUT1.

Studies of miRNA isomiRs have rapidly expanded and have led to the annotation of many types of modifications. However, the kinetics of miRNA additions remain poorly understood. We have previously noted that a subset of miRNAs in humans is nearly always modified across tissues, while other miRNAs never show additions (6). Yet we cannot exclude the possibility that 3' additions may be a transient property of some miRNAs, which prohibits detection on both the NanoString and next-generation sequencing platforms. One recent report has noted an increase in 3' uridylation during the decay of an exogenously expressed miRNA (21). Our work also demonstrated increased 3' uridylation in miRNAs showing decreased abundance following nucleotidyl transferase suppression. Together, these results raise the possibility that 3' additions may be an intermediate in the miRNA decay process. Although the mechanism of miRNA turnover is not understood, future investigations of the role of both nucleotidyl transferases and their 3' additions will facilitate understanding of the factors that determine the outcome of mature miRNAs.

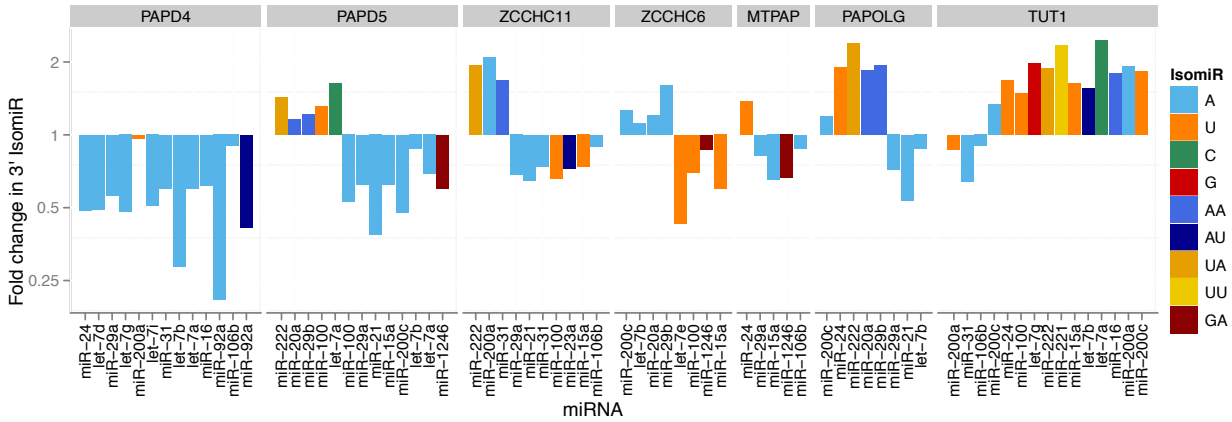

Figure 5.1. Nucleotidyl transferases mediate specific increases and decreases in miRNA 3' additions.

Following suppression of a panel of nucleotidyl transferases, we assessed the changes in miRNA additions with the NanoString nCounter miRNA profiling platform. Bars depict the fold change in a given miRNA isomiR, with the color of the bar indicating the 3' addition affected. The enzyme names above the bars indicate the suppressed nucleotidyl transferase. While some enzymes appear to primarily promote 3' additions, suppression of other enzymes such as TUT1 leads to increased 3' modification of specific miRNAs. All depicted miRNAs showed significant changes in additions compared to negative control cells with a false discovery rate < 10%.

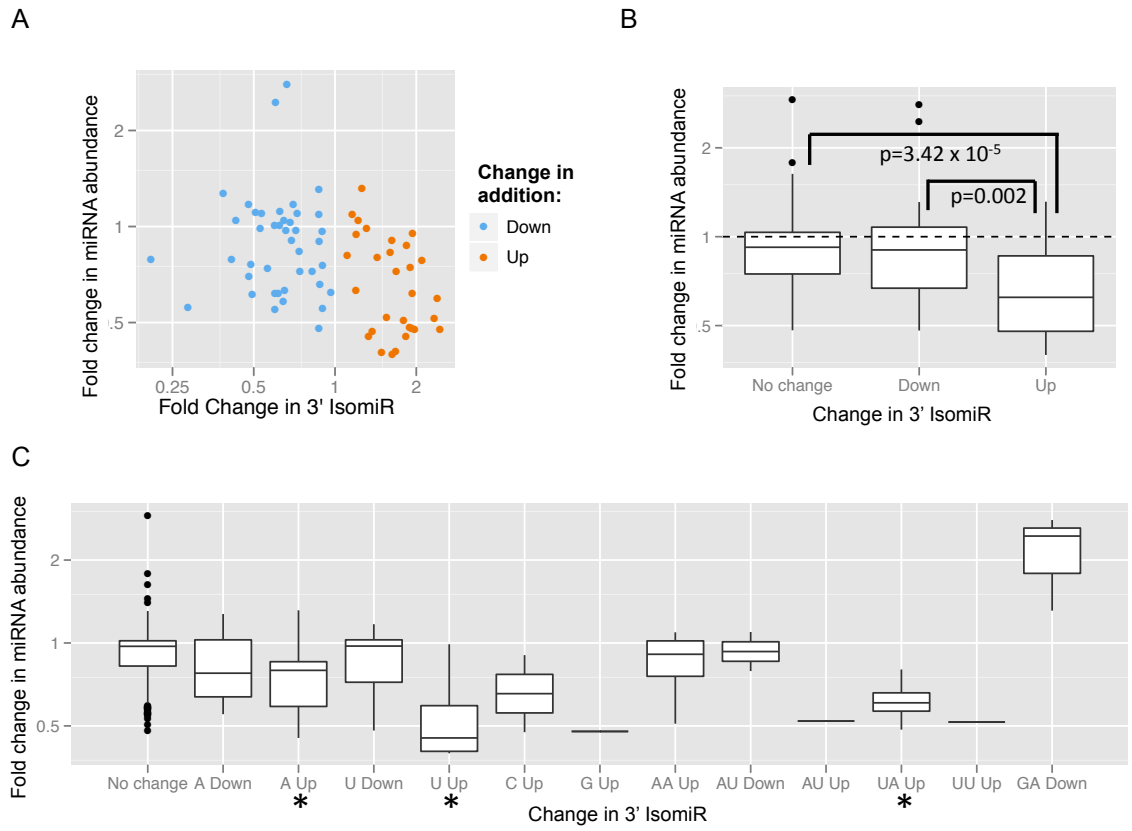

Figure 5.2. The relationship between the abundance and 3' additions of miRNAs following nucleotidyl transferase suppression.
MicroRNA 3' additions and total abundance levels were assessed with NanoString profiling following the suppression of nucleotidyl transferases in HCT-116 cells. **(A)** The scatterplot displays the relationship between the magnitude of change in 3' addition and the change in overall miRNA abundance. Blue points indicate miRNAs with decreased additions, while orange points designate increased 3' addition. We found no global correlation between changes in 3' additions and changes in abundance, indicating 3' additions do not have uniform effects on abundance. **(B)** Boxplots comparing the abundance changes of miRNAs showing increases, decreases, or no change in 3' additions following nucleotidyl transferase suppression. MicroRNAs showing significant increases in 3' additions feature a statistically significant decrease in abundance compared to miRNAs with unchanged or decreased additions (t-test, $p < 0.05$). **(C)** MicroRNAs with altered nucleotide additions show changes in abundance that are associated with the identity of the nucleotide and the direction of change. Asterisks signify a significant change in abundance compared to miRNAs with unaltered additions (t-test, $p < 0.05$).

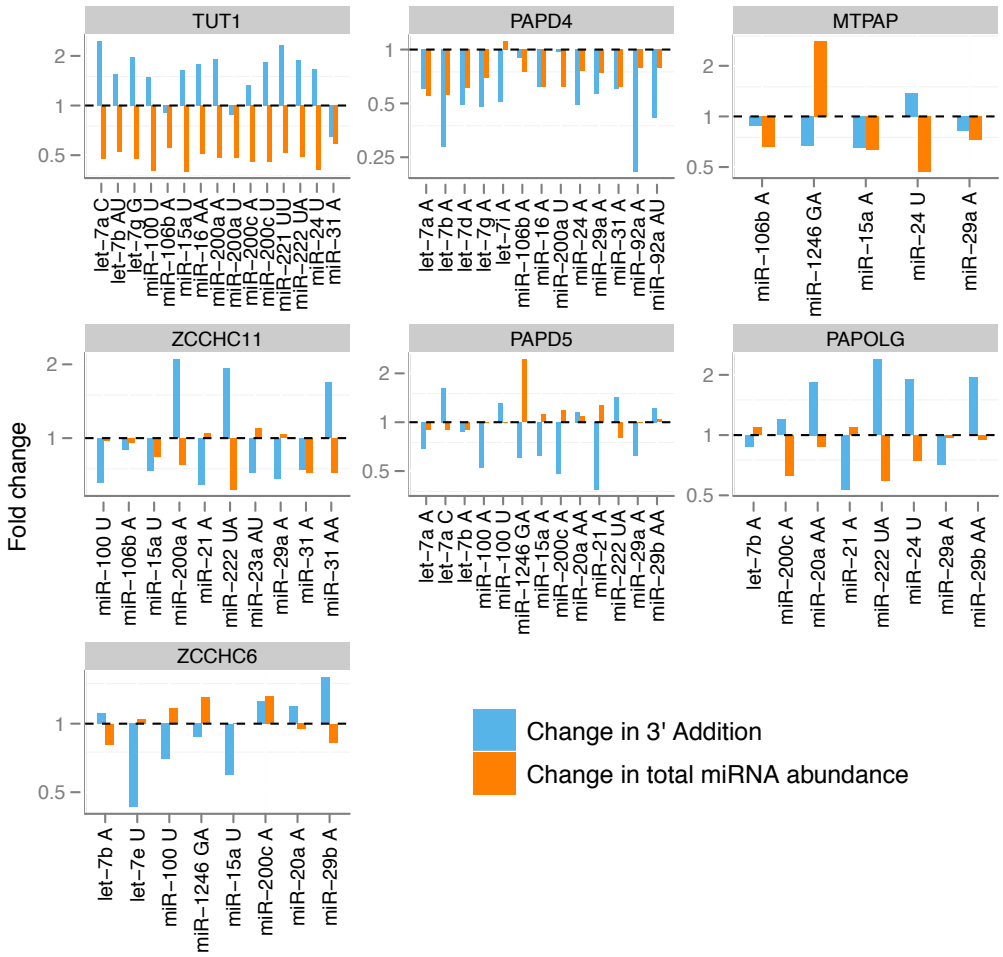

Figure 5.3. Nucleotidyl transferase suppression yields enzyme-specific changes in miRNA abundance and additions.
Bar graphs depict the changes in 3' additions and abundance of all miRNAs that displayed significant changes in additions following suppression of a given enzyme. Blue bars depict the changes in miRNA addition, and orange bars show the corresponding changes in total miRNA abundance. Suppression of both TUT1 and PAPD4 yield decreases in the abundance of most miRNAs, despite their contrasting effects on miRNA additions.

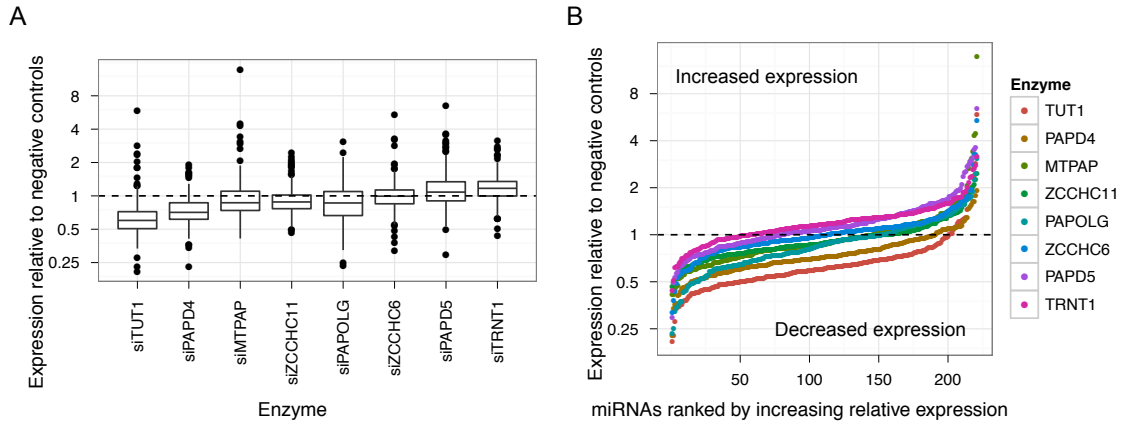

Figure 5.4. Global changes in miRNA expression following nucleotidyl transferase suppression.
Exiqon qRT-PCR miRNA arrays were used to assess the global changes in miRNAs resulting from the individual suppression of eight enzymes. **(A)** Boxplots display miRNA expression in the nucleotidyl transferase suppressed cells relative to the cells treated with negative control siRNAs. Suppression of both TUT1 and PAPD4 led to substantial decreases in miRNA abundance. **(B)** Suppression of TUT1 and PAPD4 decrease the abundance of the vast majority of miRNAs analyzed. Scatterplots show the relative expression of each miRNA in the the enzyme suppressed cells, with miRNAs ranked on the x-axis by increasing relative expression value. Any point that falls below the horizontal line at 1 indicates decreased miRNA expression compared to the controls.

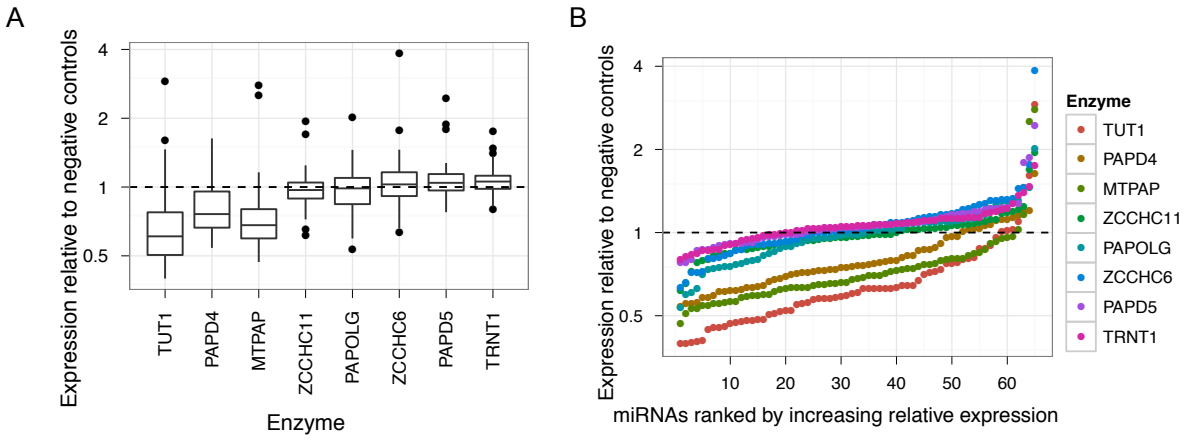

Figure 5.5. NanoString platform confirms widespread decreases in miRNA expression following TUT1 suppression.
We compared the the abundance of miRNAs detectable with at least 50 counts on the NanoString platform in the enzyme suppressed and control cells. **(A)** Boxplots show the relative miRNA expression in the enzyme suppressed versus control cells. **(B)** Scatterplots show the expression of each of the 65 miRNAs reliably detected across all samples, which are ranked by increasing relative expression on the x-axis. Any point that falls below the horizontal line at 1 indicates decreased miRNA expression compared to the controls.

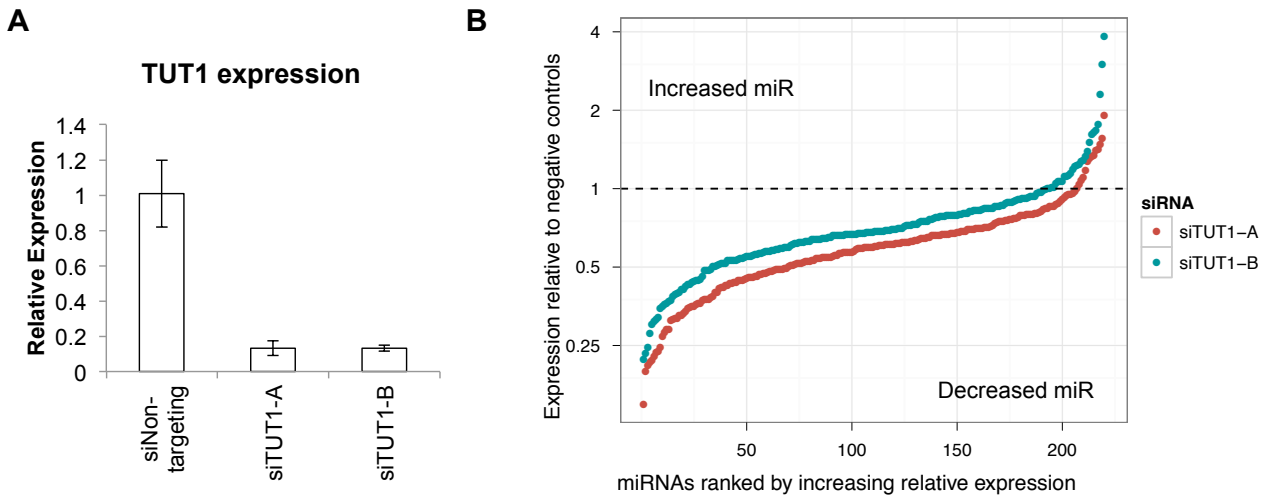

Figure 5.6. Validation with independent TUT1 siRNAs confirms the importance of TUT1 in maintaining miRNA expression levels.
Two additional TUT1 siRNAs were used to suppress TUT1 in HCT-116 cells. MicroRNA qRT-PCR arrays were then used to compare the global miRNA expression in enzyme suppressed and control cells. **(A)** qRT-PCR confirmation of TUT1 supppression in enzyme suppressed versus negative control cells. **(B)** Scatterplots depict miRNA expression following TUT1 suppression with each of the two new TUT1 siRNAs. Points below the dashed line indicate decreased miRNA abundance with TUT1 suppression compared to the negative control cells.

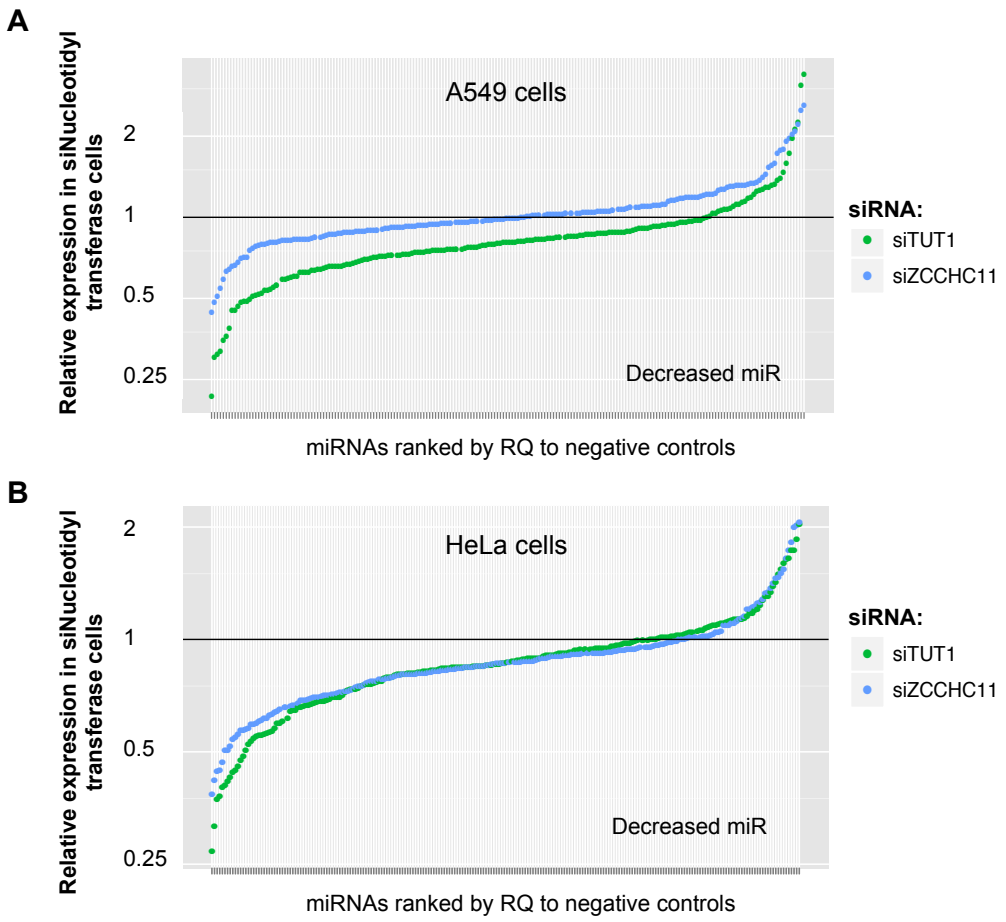

Figure 5.7. TUT1 suppression decreases miRNA expression in multiple cell lines. siRNAs against TUT1 or ZCCHC11, in addition to negative control siRNAs, were transfected into **(A)** A549 and **(B)** HeLa cells. The resulting changes in miRNA expression were then assessed with miRNA qRT-PCR arrays. Graphs depict the relative expression of miRNAs in the enzyme-suppressed versus negative control cells. Any point that falls below the horizontal line indicates decreased miRNA expression compared to controls. miRNAs are sorted by increasing relative quantification (RQ) values for each enzyme-suppressed sample.

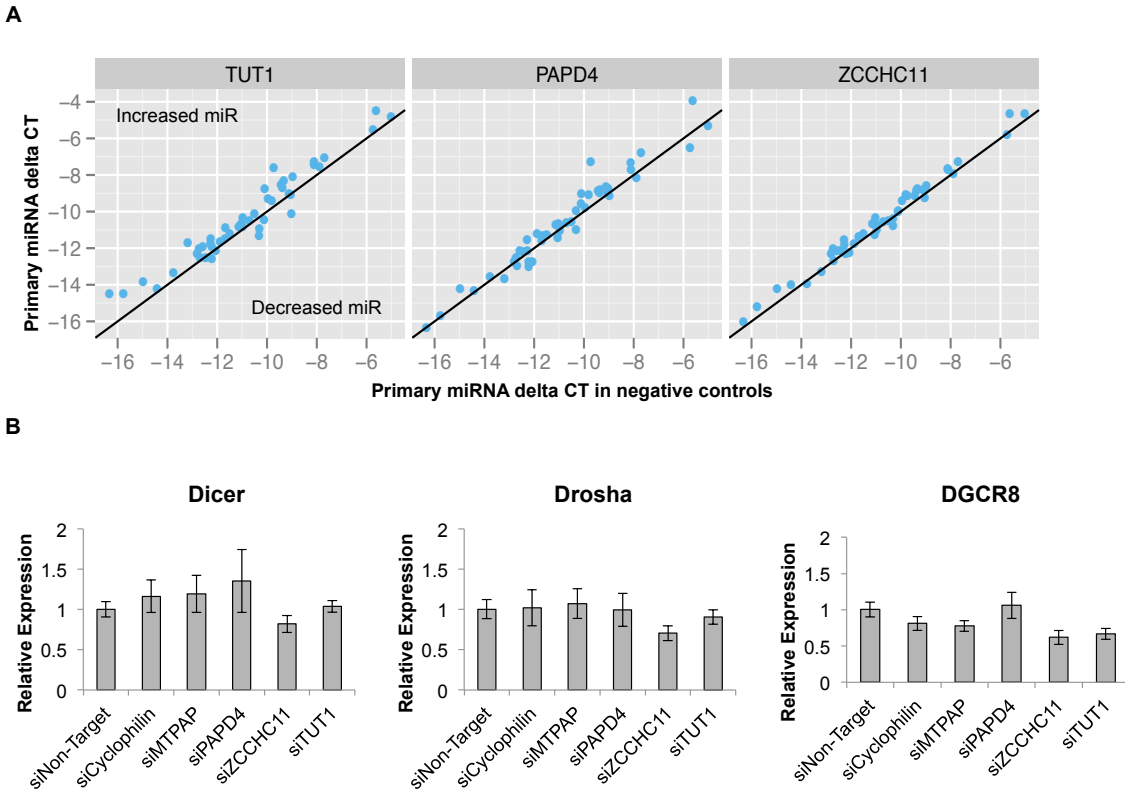

Figure 5.8. TUT1 does not regulate the expression of primary miRNA transcripts or miRNA processing components.
Following suppression of TUT1 in HCT-116 cells, two potential mechanisms for TUT1-mediated changes in miRNA levels were investigated. **(A)** The expression of 70 primary miRNA transcripts was measured with qRT-PCR. Scatterplots depict the delta CT values in the enzyme suppressed versus control cells. Suppression of TUT1, PAPD4, and ZCCHC11 all yielded slight increases in pri-miRNA abundance, indicating that changes in mature miRNA levels following TUT1 loss are not a result of upstream changes in the primary miRNA transcript abundance. **(B)** qRT-PCR was used to assess the expression of three miRNA processing components following enzyme suppression. Although the expression levels of these genes are slightly variable following nucleotidyl transferase knockdown, suppression of TUT1 does not yield unique changes that likely account for the global decreases in miRNA abundance.

Table 5.1. Suppression of nucleotidyl transferases affects miRNA abundance.
We profiled miRNA expression with Exiqon miRNA qRT-PCR arrays following suppression of a panel of nucleotidyl transferases. Suppression of TUT1 and PAPD4 led to the most substantial changes in miRNA expression compared to cells treated with negative control siRNAs.The table depicts the mean and median relative expression values (RQ) of 217 miRNAs robustly detected across all samples.

Enzyme	**Mean RQ**	**Median RQ**
TUT	0.70	0.60
PAPD4	0.77	0.71
PAPOLG	0.91	0.86
ZCCHC11	0.95	0.88
MTPAP	1.05	0.87
ZCCHC6	1.06	0.99
TRNT1	1.22	1.17
PAPD5	1.22	1.08

References

1. Bail, S., Swerdel, M., Liu, H., Jiao, X., Goff, L.A., Hart, R.P. and Kiledjian, M. (2010) Differential regulation of microRNA stability. *RNA*, **16**, 1032-1039.

2. Gantier, M.P., McCoy, C.E., Rusinova, I., Saulep, D., Wang, D., Xu, D., Irving, A.T., Behlke, M.A., Hertzog, P.J., Mackay, F. *et al.* (2011) Analysis of microRNA turnover in mammalian cells following Dicer1 ablation. *Nucleic acids research*, **39**, 5692-5703.

3. Krol, J., Busskamp, V., Markiewicz, I., Stadler, M.B., Ribi, S., Richter, J., Duebel, J., Bicker, S., Fehling, H.J., Schübeler, D. *et al.* (2010) Characterizing light-regulated retinal microRNAs reveals rapid turnover as a common property of neuronal microRNAs. *Cell*, **141**, 618-631.

4. Ramachandran, V. and Chen, X. (2008) Degradation of microRNAs by a family of exoribonucleases in Arabidopsis. *Science*, **321**, 1490-1492.

5. Chatterjee, S. and Großhans, H. (2009) Active turnover modulates mature microRNA activity in Caenorhabditis elegans. *Nature*, **461**, 546-549.

6. Wyman, S.K., Knouf, E.C., Parkin, R.K., Fritz, B.R., Lin, D.W., Dennis, L.M., Krouse, M.A., Webster, P.J. and Tewari, M. (2011) Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome research*, **21**, 1450-1461.

7. Burroughs, A.M., Ando, Y., de Hoon, M.J.L., Tomaru, Y., Nishibu, T., Ukekawa, R., Funakoshi, T., Kurokawa, T., Suzuki, H., Hayashizaki, Y. *et al.* (2010) A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome research*.

8. Morin, R.D., O'Connor, M.D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A.-L., Zhao, Y., McDonald, H., Zeng, T., Hirst, M. *et al.* (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome research*, **18**, 610-621.

9. Guo, L. and Lu, Z. (2010) Global expression analysis of miRNA gene cluster and family based on isomiRs from deep sequencing data. *Comput Biol Chem*, **34**, 165-171.

10. Fernandez-Valverde, S.L., Taft, R.J. and Mattick, J.S. (2010) Dynamic isomiR regulation in Drosophila development. *RNA*, **16**, 1881-1888.

11. Li, J., Yang, Z., Yu, B., Liu, J. and Chen, X. (2005) Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. *Curr Biol*, **15**, 1501-1507.

12. Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R. and Chen, X. (2005) Methylation as a crucial step in plant microRNA biogenesis. *Science*, **307**, 932-935.

13. Ren, G., Chen, X. and Yu, B. (2012) Uridylation of miRNAs by HEN1 SUPPRESSOR1 in Arabidopsis. *Current biology : CB*, **22**, 695-700.

14. Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B. and Chen, X. (2012) The Arabidopsis Nucleotidyl Transferase HESO1 Uridylates Unmethylated Small RNAs to Trigger Their Degradation. *Current biology : CB*, **22**, 689-694.

15. Ibrahim, F., Rymarquis, L.A., Kim, E.-J., Becker, J., Balassa, E., Green, P.J. and Cerutti, H. (2010) Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. *Proc Natl Acad Sci USA*.

16. Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S.-I., Baba, T. and Suzuki, T. (2009) Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev*, **23**, 433-438.

17. Jones, M.R., Quinton, L.J., Blahna, M.T., Neilson, J.R., Fu, S., Ivanov, A.R., Wolf, D.A. and Mizgerd, J.P. (2009) Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat Cell Biol*, **11**, 597-610.

18. Burroughs, A.M., Ando, Y., Hoon, M.L.d., Tomaru, Y., Suzuki, H., Hayashizaki, Y. and Daub, C.O. (2011) Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biol*, **8**.

19. Mellman, D.L., Gonzales, M.L., Song, C., Barlow, C.A., Wang, P., Kendziorski, C. and Anderson, R.A. (2008) A PtdIns4,5P2-regulated nuclear poly(A) polymerase controls expression of select mRNAs. *Nature*, **451**, 1013-1017.

20. Trippe, R., Sandrock, B. and Benecke, B.J. (1998) A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. *Nucleic Acids Res*, **26**, 3119-3126.

21. Baccarini, A., Chauhan, H., Gardner, T.J., Jayaprakash, A.D., Sachidanandam, R. and Brown, B.D. (2011) Kinetic Analysis Reveals the Fate of a MicroRNA following Target Regulation in Mammalian Cells. *Curr Biol*, **21**, 369-376.

Chapter 6

Conclusions

Advances in miRNA research

Since the discovery of miRNAs nearly twenty years ago (1), the field has rapidly expanded to include miRNA involvement in a multitude of normal biological processes and diseases. MicroRNAs represent a unique layer of regulation in the cell. These non-coding molecules act in the cytoplasm to modulate the levels of target molecules in a sequence-specific manner. While the effects of a single miRNA on one mRNA transcript are often modest, the combinatorial effects of a network of miRNAs regulating an ensemble of targets may allow for a tunable, large-scale regulatory response (2). The ubiquity of miRNAs in numerous biological processes and diseases has been well established; however, remaining questions include determining the drivers of miRNA expression and the impact these miRNAs networks may exert in the cell.

The advent of affordable, rapid methods of assaying miRNA levels has yielded a tremendous amount of information regarding miRNA expression patterns in humans. Next-generation sequencing studies enabled the discovery of over 1000 mature miRNAs in humans in addition to the identification of a multitude of miRNA isomiRs that further increase the diversity of the miRNA transcriptome. A plethora of miRNA profiling platforms, including qRT-PCR, microarrays, and the NanoString technology described in Chapter 4, have also facilitated the description of miRNA expression patterns in a variety of normal and disease states. Profiling miRNA expression levels has yielded promising results for the classification of tumors of unknown origin (3). Additionally, differential

expression of certain miRNAs in the blood or tissues of patients has been shown to supply clinically valuable diagnostic and prognostic information (4,5). The list of miRNAs associated with cancers is quickly growing; however, for the majority of these miRNAs, the cause and effect of their dysregulation is not known. An important and largely unanswered question in the field is: what genetic or epigenetic changes are responsible for altered miRNA expression patterns?

Novel forms of miRNA regulation

The regulation of miRNAs clearly involves multiple mechanisms. We have investigated several independent forms of regulation that span both transcriptional and post-transcriptional stages of miRNA biogenesis. Through our annotation of the putative promoters of miRNAs and the transcription factors that may regulate them, we have developed an integrated pipeline by which to identify tissue-specific drivers of miRNA dysregulation. Our discovery of the p63 and p73 transcription factors as activators of the miR-200 family underscores the importance in examining the potential intersections of miRNAs with cancer-relevant pathways. p63 and p73 are members of the p53 family of proteins that are frequently dysregulated in ovarian cancer and can be associated with poor prognosis (6,7). Our work suggests that one mechanism of p63 and p73's effects may involve widespread regulation of miRNAs. We identified a total of 17 miRNAs overexpressed in ovarian cancer with binding sites for the p53 family, which will facilitate investigations of the broader p53 family:miRNA regulatory networks that may impact tumor pathology. Future experiments to address global transcription factor binding events via chromatin immunoprecipitation followed by second-generation sequencing (ChIP-Seq) will enable verification of these new forms of transcriptional

regulation. Furthermore, although our study examined the associations of transcription factor and miRNAs in a single cancer type, the pipeline we have developed can be applied in any system featuring differential miRNA expression in order to identify potential drivers of miRNA dysregulation.

While our studies on the miR-200 family revealed new interactions between transcription factor and miRNAs, we also sought to understand the importance of the cellular microenvironment in the transcriptional activation of miRNAs. The environment of a cell can be greatly influenced *in vitro* and *in vivo* by variations in cell density, nutrient availability, pH, and many more factors that can lead to substantial changes in miRNA expression. In cell culture, experimental manipulations often lead to environmental changes that may yield off-target effects if conditions are not carefully controlled for. For example, one recent report demonstrated a global increase in mature miRNAs in response to increased cell density, which was attributed to enhanced Dicer processing (8). However, we have identified an independent mechanism of density-dependent miRNA regulation that instead occurs at the transcriptional level. We found that the expression of most primary miRNA transcripts modestly increased in response to higher cell density, but were surprised to find a dramatic outlier: The miRNA miR-210 showed rapid induction in response to increased cell density. We found this upregulation was mediated by HIF-1α—a protein that has previously been demonstrated to be activated in response to increased confluency (9). This finding may have *in vivo* significance, as the sensitivity of miR-210 induction may allow the miRNA, which represses several important regulators of proliferation and apoptosis, to enable cellular adaption within a relatively normoxic environment. While the role of miR-210 in

modulating chemosensitivity is under investigation, we hypothesize that miR-210 induction by HIF-1α in solid tumors may play an important role in treatment response. Taken together, our studies expand the role of miR-210 as a sensitive regulator of the cellular microenvironment and further elucidate the known transcriptional regulation of a miRNA with notable roles in cancer.

In addition to investigating the transcriptional regulation of individual miRNAs, we also discovered novel regulatory mechanisms that may broadly affect miRNA activity. Recent work from our lab and others has revealed a diversity of novel miRNA sequences that arise from post-transcriptional modifications. A wide range of auxiliary proteins can influence the sequence of miRNAs and potentially alter their ability to repress their canonical targets. Internal sequence modifications can be performed by the adenosine deaminase proteins, which initiate adenosine to inosine conversions that can prevent miRNA processing (10), in addition to changing the miRNA seed sequence—an important determinant of target specificity (11). Non-templated nucleotide additions to the 3' end of mature miRNAs also alter miRNA sequences. We have investigated these 3' miRNA isomiRs in order to determine their prevalence, regulation, and functional effects on miRNA activity. We have shown these 3' variants are not an artifact of sequencing, and instead are a biological phenomenon regulated by specific members of the nucleotidyl transferase family of enzymes. Our development of a new platform by which to assay miRNA isomiRs will facilitate future studies to examine these variants in different diseases or cellular states. Additionally, these studies reveal the dynamic regulation of mature miRNAs and dramatically expand the known repertoire of human miRNAs to include thousands of non-canonical sequences.

Our preliminary experiments suggest 3' additions may play a complex role in mediating miRNA expression patterns. While work from plants and algae has suggested certain 3' nucleotide additions can uniformly stabilize miRNAs (12-14), we instead have found 3' additions occur only on specific miRNAs and yield differential effects on stability. Surprisingly, several nucleotidyl transferases that regulate these 3' additions also appear to broadly affect miRNA abundance through mechanisms that may be independent of 3' additions. Biochemical analyses are now needed to determine the associations of nucleotidyl transferases such as TUT1 with the RISC complex and mature miRNAs. Our work advances understanding of the unknown factors that may control the fate of mature miRNAs in humans. Additionally, we have demonstrated that the complexity of the miRNA transcriptome in humans is much higher than was once hypothesized, with 3' additions potentially serving as a powerful new layer of regulation for miRNA expression.

Remaining challenges in miRNA biology

Ultimately, understanding of the role of miRNAs in cancers will benefit from an analysis of how a network of players, including proteins and small RNAs, act together to shape the behavior of a cell. The vast majority of research on miRNAs has examined the effect of one miRNA on one mRNA target irrespective of the multitude of upstream and downstream processes that determine when a miRNA is expressed and how expression of that miRNA may influence the cell. Most miRNAs appear to fine-tune the production of multiple proteins (15); thus, studies that focus only on the dynamics of one miRNA:mRNA interaction fail to reflect the biological role of these molecules. Future studies are needed to understand the placement of miRNAs within larger networks

containing transcription factors and other regulatory hubs that can amplify their effects and initiate large signaling cascades.

An additional layer of complexity arises because many miRNAs show tissue-specific expression patterns and functions. Because a single miRNA species is likely to target multiple transcripts, some miRNAs, such as miR-221/222, may act as tumor suppressors in certain cancers but in other cases have oncogenic roles (16,17). Thus, prior to the therapeutic application of miRNAs, it will be important to understand how they are involved in the known tissue-specific pathways that may modulate tumorigenesis. Recent work on the role of miRNAs in the multi-step development of pancreatic neuroendocrine tumors in mice has begun to define which miRNAs may contribute to the different stages of cancer progression (18). However, again the missing element of these studies is a demonstration of how the miRNAs work in concert with other tumor suppressors and oncogenes to yield a landscape conducive or refractory to cancer.

We have begun to address these questions by expanding the known regulatory network of the miR-200 family, which plays a complex role in promoting or preventing tumorigenesis and metastasis in different tissues. We have also explored the unique density-dependent induction of miR-210 in order to understand how this miRNA may affect the survival of cancer cells in the heterogeneous tumor microenvironment. Finally, with our investigations of the functional effects of miRNA 3' additions, we have examined a potential new avenue for the selective stabilization of miRNAs. Together, our work has identified novel regulatory relationships between miRNAs and proteins that may inform future studies to modulate the activity of individual miRNAs. By understanding the mechanisms of miRNA activation and repression, we may increase the

utility of miRNAs as potential diagnostic or therapeutic tools. Future work is now needed to address the interplay between these regulatory components in disease, and to understand the unique contributions of miRNAs to the gene expression patterns of the cell.

References

1. 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.

2. Ebert, M.S. and Sharp, P.A. (2012) Roles for microRNAs in conferring robustness to biological processes. *Cell*, **149**, 515-524.

3. Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A. *et al.* (2005) MicroRNA expression profiles classify human cancers. *Nature*, **435**, 834-838.

4. Volinia, S., Galasso, M., Sana, M.E., Wise, T.F., Palatini, J., Huebner, K. and Croce, C.M. (2012) Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. *Proc Natl Acad Sci USA*, **109**, 3024-3029.

5. Mitchell, P.S., Parkin, R.K., Kroh, E.M., Fritz, B.R., Wyman, S.K., Pogosova-Agadjanyan, E.L., Peterson, A., Noteboom, J., O'Briant, K.C., Allen, A. *et al.* (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci USA*, **105**, 10513-10518.

6. Marchini, S., Marabese, M., Marrazzo, E., Mariani, P., Cattaneo, D., Fossati, R., Compagnoni, A., Fruscio, R., Lissoni, A.A. and Broggini, M. (2008) DeltaNp63 expression is associated with poor survival in ovarian cancer. *Ann Oncol*, **19**, 501-507.

7. Concin, N., Becker, K., Slade, N., Erster, S., Mueller-Holzner, E., Ulmer, H., Daxenbichler, G., Zeimet, A., Zeillinger, R., Marth, C. *et al.* (2004) Transdominant DeltaTAp73 isoforms are frequently up-regulated in ovarian cancer. Evidence for their role as epigenetic p53 inhibitors in vivo. *Cancer Res*, **64**, 2449-2460.

8. Hwang, H.W., Wentzel, E.A. and Mendell, J.T. (2009) Cell-cell contact globally activates microRNA biogenesis. *Proc Natl Acad Sci USA*, **106**, 7016-7021.

9. Fang, Y., Sullivan, R. and Graham, C.H. (2007) Confluence-dependent resistance to doxorubicin in human MDA-MB-231 breast carcinoma cells requires hypoxia-inducible factor-1 activity. *Exp Cell Res*, **313**, 867-877.

10. Yang, W., Chendrimada, T.P., Wang, Q., Higuchi, M., Seeburg, P.H., Shiekhattar, R. and Nishikura, K. (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nature structural & molecular biology*, **13**, 13-21.

11. Kawahara, Y., Megraw, M., Kreider, E., Iizasa, H., Valente, L., Hatzigeorgiou, A.G. and Nishikura, K. (2008) Frequency and fate of microRNA editing in human brain. *Nucleic acids research*, **36**, 5270-5280.

12. Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B. and Chen, X. (2012) The Arabidopsis Nucleotidyl Transferase HESO1 Uridylates Unmethylated Small RNAs to Trigger Their Degradation. *Current biology : CB*, **22**, 689-694.

13. Ren, G., Chen, X. and Yu, B. (2012) Uridylation of miRNAs by HEN1 SUPPRESSOR1 in Arabidopsis. *Current biology : CB*, **22**, 695-700.

14. Ibrahim, F., Rymarquis, L.A., Kim, E.-J., Becker, J., Balassa, E., Green, P.J. and Cerutti, H. (2010) Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. *Proc Natl Acad Sci USA*.

15. Baek, D., Villén, 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.

16. Felli, N., Fontana, L., Pelosi, E., Botta, R., Bonci, D., Facchiano, F., Liuzzi, F., Lulli, V., Morsilli, O., Santoro, S. *et al.* (2005) MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation. *Proc Natl Acad Sci USA*, **102**, 18081-18086.

17. Galardi, S., Mercatelli, N., Giorda, E., Massalini, S., Frajese, G.V., Ciafrè, S.A. and Farace, M.G. (2007) miR-221 and miR-222 expression affects the proliferation potential of human prostate carcinoma cell lines by targeting p27Kip1. *J Biol Chem*, **282**, 23716-23724.

18. Olson, P., Lu, J., Zhang, H., Shai, A., Chun, M.G., Wang, Y., Libutti, S.K., Nakakura, E.K., Golub, T.R. and Hanahan, D. (2009) MicroRNA dynamics in the stages of tumorigenesis correlate with hallmark capabilities of cancer. *Genes Dev*, **23**, 2152-2165.

VITA

Emily C Knouf was born in Madison, Indiana. She graduated *magna cum laude* from Pomona College with a Bachelor of Arts degree in Molecular Biology in 2007. In 2012 she graduated with a Doctor of Philosophy in Molecular and Cellular Biology from the University of Washington. She performed research at the Fred Hutchinson Cancer Research Center under the supervision of Dr. Muneesh Tewari.