THE UNIVERSITY of ADELAIDE

Micro-RNAs in cancer: novel origins and sequence variation

Feng Yu

M.Sc., B.Sc.

This thesis is submitted in fulfilment of the requirements for the Doctor of Philosophy

School of Medicine

The University of Adelaide

Adelaide SA, Australia

March 2016

Table of Contents

Table of Contents I

Overview III

Publications VI

Acknowledgements VII

Declaration IX

Chapter 1 1

1. THE INTERACTION BETWEEN P53 AND NON-CODING RNAS FROM NOVEL SOURCES 2

1.1. The p53 pathway – a key tumour suppressive pathway 2

1.2. Non-coding RNAs -- a new route by p53 to regulate the transcriptome 4

1.2.1. miRNAs are vital cancer regulators 6

1.2.2. MicroRNAs directly regulated by p53. 7

1.2.3. p53 regulates the transcription of miRNAs 9

1.2.4. p53 regulates miRNA maturation 11

1.2.5. The reciprocal interaction of p53 and its target miRNAs 12

1.2.6. Clinical implications of p53-regulated miRNAs 12

1.2.7. p53 regulation of small non-coding RNAs processed by RNA polymerase I and III 14

1.2.8. Potential sources of novel miRNAs under p53 regulation 15

1.3. Post- transcriptional roles of long non-coding RNAs (lncRNAs) in the p53 downstream network 17

1.3.1. Introduction to lncRNAs 17

1.3.2. Post-transcriptional regulatory roles of lncRNAs 18

1.3.3. Regulatory roles of lncRNAs directly targeted by p53 20

1.4. Conclusion 23

2. ISOMIRS – NATURALLY EXISTING ISOFORMS OF MIRNAS 24

2.1. The maturation of miRNA and its heterogeneity 24

2.1.1. The existence of 5′ and 3′ isomiRs 28

2.2. Sources of miRNA heterogeneity 29

2.2.1. 5′ and 3′ heterogeneity derived from the biogenesis of isomiRs 30

2.2.2. 3′ heterogeneity derived from trimming and nucleotidyl transferring 31

2.2.3. Heterogeneity derived from RNA editing 32

2.3. Functional impacts of miRNA heterogeneity 34

2.3.1. Alteration in target recognition 35

2.3.2. Alteration in AGO loading 36

2.3.3. Alteration in miRNA stability 37

2.4. Discussion 38

3. SUMMARY OF THE CHAPTER 39
REFERENCES 41
Chapter 2 69
PRELUDE 70
ABSTRACT 74
INTRODUCTION 75
RESULTS 77
Identification of p53 regulated snoRNAs 77
p53 repressed snoRNAs are processed into miRNAs 85
sno-miR-28 functions as a miRNA 92
sno-miR-28 alters p53 protein stability through TAF9B 97
sno-miR-28 is upregulated in breast tumours and promotes breast cell proliferation 101
DISCUSSION 104
snoRNAs are processed into miRNAs 104
sno-miR-28 is a novel member of the p53 regulatory network 106
sno-miR-28 is overexpressed in breast tumour samples and plays an oncogenic role in breast cells 107
MATERIALS AND METHODS 108
ACKNOWLEDGEMENTS 124
REFERENCES 125
Chapter 3 136
PRELUDE 137
ABSTRACT 142
INTRODUCTION 143
RESULTS 147
Endogenous expression of miR-222 isoforms 147
miR-222 isoforms significantly alter cell morphology 148
All four miR-222 species inhibit cell proliferation 152
Longer miR-222 isoforms are associated with increased apoptotic activity 157
The differential apoptotic activity of longer 222 isomiRs is not attributable to interferon responses 160
PIK3R3 (p55) is involved in miR-222CUCU-induced apoptosis 163
DISCUSSION 171
MATERIALS AND METHODS 175
REFERENCES 187
Conclusion 198
REFERENCES 203

Overview

Non-coding RNAs have become a hot topic of cancer research in recent decades, with miRNAs being probably the most active area of research. However, the functions of short non-coding RNA (ncRNA) fragments commonly existing in cells remains less understood. Herein we explore the novel ncRNAs which play a role in cancer progression.

The first part of the thesis focuses on miRNAs processed from novel sources that have a role in p53 regulation. p53 is a master tumour repressor that participates in vast regulatory networks, including feedback loops involving microRNAs (miRNAs) that regulate p53 and that themselves are direct p53 transcriptional targets. A group of polycistronic miRNA-like non-coding RNAs is derived from small nucleolar RNAs (sno-miRNAs) that are transcriptionally repressed by p53 through their host gene, *SNHG1*. Among them, sno-miR-28 is the most abundant sno-miRNA bound to the AGO (Argonaute) proteins and it directly targets TAF9B, thereby de-stabilizing p53. Therefore, a positive a regulatory loop was observed comprising p53, SNHG1, sno-miR-28 and TAF9B, which influences p53 stability and downstream p53-regulated pathways. In addition, SNHG1, SNORD28 and sno-miR-28 are all significantly upregulated in breast tumours and the overexpression of sno-miR-28 promotes breast

epithelial cell proliferation and colony formation. This research has broadened our knowledge of the crosstalk between small non-coding RNA pathways and p53 regulation.

The second part of the thesis investigates naturally existing isoforms of miR-222 that play pro-apoptotic roles. Alternative processing at the 3′ end of miRNAs has been broadly observed, producing variable lengths of miRNA mature forms. Deep-sequencing of various tissues and tumours, combined with sequencing of AGO-bound miRNAs from cell lines, indicates a variable proportion of endogenous miR-222 that is extended by one to five nucleotides at the 3′ end. We demonstrated that the 3′ heterogeneity of miR-222 possesses dramatic implications for the phenotype of miR-222 transfected cells, with longer isoforms driving apoptosis in addition to the proliferation inhibition bestowed by both the short (canonical) and longer forms. Further investigation revealed intrinsic apoptotic events exhibiting a positive correlation to the length of miR-222 isoforms, but not the specific 3′ sequence. However, the apoptosis failed to be correlated to interferon immunoresponse, and the longer miR-222 isoform exhibits identical targeting activity as the canonical miR-222. Widespread disruption of the expression of key PI3K-AKT components was observed upon miR-222CUCU transfection. A PI3K regulatory subunit, PIK3R3, was of particular interest, as siPIK3R3 phenocopied miR-222CUCU in terms of

apoptotic effects and inhibition of PI3K-AKT gene expression. Given the high prevalence of 3' variance in many other miRNAs, the functional impact of miR-222 isoforms reveals another layer of miRNA regulation that has implications for cancer therapy.

Taken together, the existence of miRNAs processed from novel sources and isomiRs has added to the complexity of our knowledge about miRNA regulation. These areas are much less explored than conventional miRNA processing and regulation, but their profound molecular biological and physiological implications suggest an unexplored layer of miRNA biology and may shed light to the contemporary research of cancer progression.

Publications

Original Research Papers

- Feng Yu, Cameron P. Bracken, Katherine A. Pillman, David M. Lawrence, Gregory J. Goodall, David F. Callen, Paul M. Neilsen. p53 represses the oncogenic sno-miR-28 derived from a snoRNA. PLoS One. 2015; 10(6): e0129190.
- Sun D, Yu F, Ma Y, Zhao R, Chen X, Zhu J, Zhang CY, Chen J, Zhang J. "Sun D and Yu F contributed equally to this work." MicroRNA-31 activates Ras pathway and functions as an oncogenic microRNA in human colorectal cancer by repressing RAS p21 GTPase activating protein 1 (RASA1). Journal of Biological Chemistry. 2013 Mar 29; 288(13):9508-18.

Presentations

- Poster: Naturally existing isoforms of miR-222 play distinct pro-apoptotic roles. The 7th Barossa Meeting on cell signalling in cancer biology and therapy, 18-21 Nov. 2015
- Poster: Naturally existing isoforms of miR-222 play distinct pro-apoptotic roles. ASMR SA Annual Scientific Meeting, 3 June 2015.
- Poster: p53 represses the oncogenic sno-miR-28 derived from a snoRNA. Florey International Postgraduate Research Conference, 25 September 2014.

Acknowledgements

A major research projects such as this would not have been completed without the collaboration of many different people, who have made this work possible. I would like to extend my gratitude and appreciation to the following:

Firstly, I sincerely wish to thank my supervisors, Dr Cameron Bracken and Dr Paul Neilsen. With their daily encouragement and continuous support, I was led into the world of scientific research. Their unsurpassed knowledge and unwavering enthusiasm for science constantly inspired me to work harder and think more. I doubt that I will be able to convey my appreciation fully, but I own Dr Cameron Bracken and Dr Paul Neilsen my eternal gratitude and could not have asked for a better mentor to guide me through this long journey.

I am extremely grateful for my supervisors Prof David Callen and Prof Gregory Goodall. They were the ones who pointed me the direction of my study and offered me a perfect environment to conduct my scientific research. Thank you for all your support and academic guidance over the years.

I would like to thank the bioinformatics people, Katherine Pillman, John Toubia, David Lawrence and Anna Tsykin, who have contributed valuable parts in our research. Without them, we would have never been able to expand our research to such a depth.

I would like to thank all my colleagues in the laboratories, in particular Alaknanda Alaknanda, who has collaborated with me intensively and offered indispensable help in flow cytometry and cell experiments. In particular, thank you to Renee Schulz, our former laboratory manager with huge experience, who supported me to establish my laboratory skills and helped me greatly for the experimental techniques in my first two years.

Special thanks to the University of Adelaide for providing me with a scholarship and comfortable research environment during my work towards the thesis.

Last, I dedicate my thesis to my family, in particular my wonderful partner Qian Yu, for his patience and unconditional support. He helped me both emotionally and physically in the daily life throughout the last two years.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Feng Yu

29 Mar 2016

Chapter 1

Introduction:

Novel non-coding RNAs in p53 regulation and miRNA isoforms

1. The interaction between p53 and non-coding RNAs from novel sources

1.1. The p53 pathway – a key tumour suppressive pathway

The p53 protein plays a vital role in cancer biology. Termed as "guardian of the genome", p53 is widely recognised as an important tumour suppressor, although recent studies have revealed an ever-expanding complexity in its regulatory roles that now include metabolism, autophagy, apoptosis, cell migration and invasion, cell cycle arrest, and DNA repair [1-5]. p53 is activated following DNA damage or aberrant oncogene expression, resulting in rapid stabilisation of p53 protein levels [5]. This is supported by the prevalent *TP53* mutation or p53 protein inactivation in various tumours [6, 7].

p53 is a transcription factor [8], and the normal function of p53 as a tumour suppressor is to react to stress through two distinctive aspects: it can eliminate cells with DNA damage or those undergoing aberrant proliferation, but p53 can also maintain cell survival by triggering cell cycle arrest, DNA repair, and other pro-survival pathways [9]. Numerous p53-regulated genes have been discovered to play a role in cancer, many of which are regulators of apoptosis or cell cycle arrest [10-12]. Many mechanisms exist within cells to regulate the p53 transcriptional program. In addition to locus-specific *cis*-regulatory elements, regulation of p53 activity is also achieved via a vast number of post-

translational modifications, covalent and non-covalent p53 binding partners and p53 response elements of variable binding affinity [13]. These p53-mediated transcriptional activation functions include transcriptional activation in DNA binding, transcription initiation, and transcription elongation. In RNA pol II transcription regulation, p53 recognises target sequences by binding to p53 response elements (REs). Menendez *et al*. concluded that a generally accepted consensus sequence for p53 binding is composed of two 10-base decamers and a spacer as follows: RRRCWWGYYY…n…RRRCWWGYYY (in which R is a purine, Y is a pyrimidine, W is an A or T and the spacer is <3 for functional REs) [9]. p53 binds DNA in a tetrameric complex by recognising the RRRCW or WGYYY sequence. There is also limited sequence-independent DNA binding, which can be increased by non B-form DNA structures [14]. Basically p53 REs can be located anywhere within the target gene locus, but they are most commonly found in the promoter at varying distances upstream (e.g., *CDKN1A*, *NOXA*) from the transcription start site (TSS).

Furthermore, recent reports also suggest that p53 can sometimes mediate transcriptional repression, which is much less understood than transcriptional activation [13]. It represses multiple cell cycle regulation genes by different mechanisms [15]. For example, Koumenis *et al*. reported that p53 promotes apoptosis by interacting with transcriptional co-repressor mSin3A upon induction of hypoxia stress [16]. Briefly, depending on the DNA binding

domain of p53, its transcriptional repression is achieved by not only directly binding to gene promoters, but also direct interference with the basal transcription machinery, and alteration of chromatin structure [15].

1.2. Non-coding RNAs -- a new route by p53 to regulate the transcriptome

Recently, ncRNAs (non-coding RNAs) have become a hot topic in molecular regulatory research. Knowledge about the regulatory roles of ncRNAs is rapidly emerging, and next-generation sequencing across various prokaryote and eukaryote species has revealed a positive correlation between the regulatory potential of non-coding RNAs and developmental complexity [17]. Ponting *et al*. pointed out that about 1% of the genome is protein coding, which suggests a vast proportion of the genome could be ncRNAs [18]. Researchers have so far discovered various mechanisms of the post-transcriptional regulation of ncRNAs including regulation of the translation machinery [19, 20], the stability or processing of mRNAs [21], and alternative splicing [22]. However, the systematic research of miRNAs (micro-RNAs) and other ncRNAs was only indicated in the 1990's and there are still numerous unanswered questions [23, 24]. Furthermore, a substantial proportion of the previous ncRNA function research was conducted on a case-by-case basis, and the selection of potential candidates could be biased within artificial selection

procedures based on the conventional detection methods, *e.g.*, microarray and RT-PCR. This indicates a wider range of potential functional ncRNAs which are potential for further identification using RNA deep sequencing techniques.

Perhaps to-date the best characterized family of ncRNAs are miRNAs, which are usually 20~23nt in length [25]. The maturation of miRNAs is a complex process, and the most commonly used model involves the transcription of the encoding genes into primary (pri)-miRNAs, the processing initially by Drosha into precursor (pre)-miRNAs, and next by Dicer into their mature form [26]. In human cells, miRNAs are recruited to Argonaute (AGO) proteins, which, together with other accessory proteins, form the RNA-induced silencing complex (RISC) [27]. miRNA plays a critical role in negative regulation of gene expression by recognizing complementary sites on target messenger RNAs (mRNAs) and thereby serve as the target recognition component of the RISC which degrades mRNAs, or inhibits their translation, or does both [28]. The influence of miRNAs on gene expression is widespread, as it has been estimated that more than 60% of human protein coding genes may be conserved targets of microRNAs [29]. Therefore, miRNAs are involved in the modulation of various critical biological processes including cell proliferation, differentiation, apoptosis, tumorigenesis, and the immune response [30].

1.2.1. miRNAs are vital cancer regulators

miRNAs have been studied most intensively in the field of oncological research, and emerging evidences reveal that dysregulation of certain miRNAs that regulate the translation of oncogenes and tumour suppressors is involved in the pathogenesis of cancers [31]. miRNA profiling in cancer has found that oncogenic miRNAs (oncomiRs) are usually up-regulated in cancers, playing their roles by inhibiting tumour suppressor genes or activating oncogenes. In contrast, tumour suppressive miRNAs are often down-regulated, inducing tumour-suppressor genes or repressing oncogenes. For example, miR-17-92 has been validated to target the oncogenic HIF-1α in lung tumours using an isobaric tagging reagent, iTRAQ [32]. Likewise, miR-21 has also been identified to target PTEN, TPM1, and PDCD4 in breast tumours both *in vitro* and *in vivo* [33]. On the other hand, miRNAs can also be tumour suppressors, e.g., miR-15a and miR-16 have been identified to target Bcl-2 in chronic lymphocytic leukaemia in cell experiments [34], and the let-7 family have been discovered to target Ras in lung tumours in both cell experiments and C. elegans transgenic reporter analysis [35]. Furthermore, miRNA expression profiling experiments have demonstrated a global decrease of miRNAs in tumour tissues compared to normal tissues, which probably reflects the less differentiated states of the tumour cells or their higher proliferation rates [36, 37]. Another possibility is

that reduced miRNA levels are selected during tumorigenesis because this phenomenon itself provides proliferative or survival advantage [38].

1.2.2. MicroRNAs directly regulated by p53.

Recent studies have revealed that miRNAs play an important role in the p53 network, providing additional regulatory patterns. They are regulated either through altered transcription of their host genes, or modifications in the maturation procedure [39].

A summary of known miRNAs directly regulated by p53 are listed in Table 1, and those that have been well-characterised will be discussed in further details below.

miRNAs directly targeted by p53	Examples of the Downstream Targets of miRNAs listed	General roles in cell physiology
miR-34 family [40-44]	BCL-2, WNT1, MET, CDK4, CDK6, cyclin E2, E2F3, MYC, SIRT1, E2F3, NOTCH1, PDGFRA, and MYCN [42, 45-50]	Apoptosis [40], senescence [42, 51], cell cycle arrest [42, 46], and repression of proliferation [52, 53], cancer stemness [54, 55], metastasis, invasion [56] and chemoresistence [57-59]

miR-145 (65)	MYC34, CDK4, CDK6 (66)	G1/S phase transition (66)
miR-200 family [60] [61]	ZEB1, ZEB2, JAG1, MAML2, MAML3	Epithelial–mesenchymal transition (EMT) [60, 62-64]
miR-192 family [60] [61]	ZEB1, ZEB2, JAG1, MAML2, MAML3, MDM2, TYMS [60, 62-66]	EMT [60, 62-64], Cell cycle arrest [61]
miR-15a, miR-16-1 [67, 68]	Cyclin E [69, 70]	G1/S phase transition [69, 70]
miR-107 [71]	HIF-1β, CDK6, p130 [71, 72]	Angiogenesis [72], G1 phase progression and G1/S transition (77)
miR-224 [73]	SMAD4 [73, 74]	Proliferation [73]
miR-17-92 cluster [75]	p21 [76]	cell cycle progression [76], Hypoxia-induced apoptosis [75], Decrease of estrogen receptor [77]
miR-22 [78]	Cyr61 [78]	Inflammation [78]
miR-149-3p [79]	GSK3 α [79]	Apoptosis [79]
miR-605 [80]	MDM2 [80]	Apoptosis [80]
miR-1204 [81]		Cell death [81]
miR-1915 [82]	Bcl-2 [82]	Apoptosis [82]

Table 1.2.2: The miRNAs that are directly targeted by p53 and their general roles in cell activity.

1.2.3. p53 regulates the transcription of miRNAs

Most miRNAs are located in introns of transcriptional units that are either protein-coding or non-coding [83]. Therefore, their expression could be regulated by p53 if their host genes harbour p53 REs. For example, several groups reported that p53 can directly regulate the expression of the miR-34 family by binding to p53 REs in the promoter of their host gene [40-44]. miR-34a and miR-34b/c share the same seed sequence, and have been identified to be master tumour suppressors targeting various critical cellular signaling pathways, e.g., PI3K, Notch, p53, Wnt, E2F, and MAPK [42, 45-50]. Through these targets, the miR-34 family influence major cancer-related functions, such as apoptosis [40] , senescence [42, 51], cell cycle arrest [42, 46], and repression of proliferation [52, 53], cancer stemness [54, 55], metastasis, invasion [56] and chemoresistence [57-59].

p53 was also reported to transactivate the genes that encode two miR-200 subfamilies, miR-200c/141 and miR-200a/200b/429, through direct binding to their promoters [60, 63]. The mature miRNA products of the miR-200 family are involved in the regulation of epithelial–mesenchymal transition (EMT) by targeting ZEB1 and ZEB2, and they also target Notch signalling by directly

repressing the expression of Jagged 1 (JAG1) and the mastermind co-activators MAML2 and MAML3 [60, 62-64].

Other miRNAs that are regulated by p53 binding to p53 REs in their promoters include miR-15a, miR-16, miR-107, miR-192, miR-215, and miR-155 [39, 84]. To be specific, miR-15a and miR-16-1 have been discovered to be processed at an increasing rate as p53 is activated [67]. They are encoded by an intron of the DLEU2 mRNA, whose gene was shown to be a transcriptional target of p53 [68]. These miRNAs target various cell cycle regulators such as cyclin E, and contribute to the p53-induced cell cycle arrest at the G1/S phase transition [69, 70]. Another p53 regulated miRNA, miR-107, is encoded in an intron of the p53-induced pantothenate kinase 1 (*PANK1*) gene [71, 72]. miR-107 targets expression of the hypoxia inducible factor-1β (HIF-1β), the cell cycle activator CDK6, and the anti-mitogenic p130 (also known as RBL2) [71, 72]. Yan et al. reported that p53 represses the expression of the miR-17-92 cluster by binding to the proximal region of its promoter, which contributes to the induction of hypoxia-induced apoptosis [75]. Notably, two members of this cluster, miR-20a and miR-17-5p, which belong to the miR-106b family, have been previously reported to directly target a vital cell cycle regulator, p21 [76]. Interestingly, melanoma tissues having a low level of miR-149-3p expression also show lower expression of p53, higher levels of GSK3α, and lower levels of Mcl-1, and a regulatory pathway between p53, miR-149-3p, GSK3α, and Mcl-1 was

demonstrated [79]. Another miRNA, miR-145, which downregulates MYC34 and other G1 phase and S phase regulators such as CDK4 and CDK6, is targeted by p53 in a similar manner [85, 86]. As p53 itself directly represses the expression of MYC [87], these miRNAs may offer a robust backup mechanism.

1.2.4. p53 regulates miRNA maturation

In addition to the transcriptional regulation of miRNA host genes, p53 can also regulate miRNA processing. p53 was found to directly target DDX5 (the DEAD-box RNA helicase p68) and enhance its interaction with the Drosha complex, thereby regulating the processing of specific pri-miRNAs to pre-miRNAs (e.g., miR-16-1, miR-143, and miR-145), and eventually affecting the mature miRNA biogenesis [67]. Recently, p53 has been found to induce the processing between pri-miR-1915 and pre-miR-1915, thereby inhibiting Bcl-2 in DNA damage response [82], and it also promotes the processing of miR-18a to repress estrogen receptor-α [77]. Moreover, Hermeking *et al*. pointed out recently that p53 can influence the recognition of miRNAs to their targets by regulation of RNA binding proteins [88]. For example, RBM (RNA-binding-motif protein), which competes with miRNAs for binding to 3′UTRs of mRNAs, is induced by p53 [89].

1.2.5. The reciprocal interaction of p53 and its target miRNAs

miRNAs are often associated with feed-back and feed-forward motifs and p53 target miRNAs are also involved in similar regulatory patterns [90, 91]. For example, the miR-34 family mediates tumour suppression through a positive feedback loop involving p53 and MDM4 [92]; 15 miRNAs, including the miR-106b/93/25 cluster, miR-17-92 cluster and the miR-106a-92 cluster, are repressed by p53 and involved with E2F in a feed-forward loop promoting proliferation [93]; whereas miR-192, 194 and 215 are involved in the p53-MDM2 auto-regulatory loop [66]. These discoveries, while revealing a deeper layer of the p53 signalling network, also highlight the complexity of miRNA regulation.

1.2.6. Clinical implications of p53-regulated miRNAs

Molecular biological roles of p53-regulated miRNAs have been extensively identified, suggesting that they also exert widespread influence *in vivo*. The best characterised p53-regulated miRNAs in human cancers are probably miR-34 and miR-200c.

The function of p53 seems to be intact in miR-34 deficient mice with normal development [94], and therefore miR-34 is perhaps largely redundant in normal tissues. However, the tumour suppressive role of miR-34 is prominent in tumour models, since miR-34a is downregulated by 100%, and miR-34b*/c by 72%, in human epithelial ovarian cancer tissues with p53 mutations [95]. In a transgenic mouse model, Cheng et al. discovered co-operation of miR-34 and p53 to supress prostate tumour development [96]. In addition, various mouse tumour xenograft models have revealed significant tumour suppressive roles of miR-34a *in vivo* in non-small cell lung cancer [52], prostate cancer [54], melanoma [97], pancreatic cancer [98], and lymphoma [99] using various delivery methods, which suggests potential therapeutic applications.

Similarly, miR-200c has been discovered to play a significant role in patient survival of stage I epithelial ovarian cancer [100], and a mouse xenograft model has shown the miR-200 family promotes MET (mesenchymal-epithelial transition) and invasion [101]. Furthermore, miR-192 and -215 are both downregulated in colon tumours [61], lung cancer [102], renal childhood neoplasms [103], and colon tumours [104, 105], while miR-192 is downregulated in hepatocellular carcinoma [106] and miR-215 is downregulated in renal cell carcinoma [107]. Consistent with this, a mouse

xenograft model showed slower tumour development following a combined treatment of miR-192, 194, and 215 [66].

1.2.7. p53 regulation of small non-coding RNAs processed by RNA polymerase I and III

p53 regulation is not restricted to RNA pol II transcription and various evidence suggests p53 is a vital player in RNA pol I and RNA pol III transcriptional regulation [108]. Indeed, significant increases of rRNA and tRNA biosynthesis were observed in fibroblasts from p53-knockdown mice [109]. This area is less well characterized than p53 regulation of RNA pol II, but recent clues have shown its great importance.

The role of p53 in the regulation of RNA pol I has been demonstrated by Zhai et al. They showed that p53 interferes with the assembly of the RNA pol I transcription machinery on the RNA pol I promoter, resulting in transcriptional repression [108]. The repression of RNA pol I transcription inhibits 18S, 5.8S, and 28S rRNA expression which compromises ribosome biogenesis and causes nucleolar stress [110]. In addition, the repression of 18S, 5.8S, and 28S rRNA biosynthesis also represses the cell cycle protein cyclin D1, stimulates the CDK

inhibitor p27, represses PIM1, and decreases the expression of E2F-1 [111-113]. All these changes in molecular pathways lead to inhibition of cell growth [114].

Similarly, p53 can also inhibit RNA pol III. Crighton et al. discovered that p53 binds to TBP and prevents promoter occupancy by TFIIIB, thereby interfering with the RNA Pol III transcription machinery [115]. The same authors also demonstrated that p53 is not bound directly to tRNA genes *in vivo.* This is supported by up to 10-fold selective upregulation of specific tRNAs in breast tissues compared with normal tissues, which increases the translation efficiency of oncogenes [116]. In addition, the repression of RNA pol III also affects the biosynthetic capacity of cells and thereby repressing cell growth [117]. Moreover, the widespread regulation of basic RNA and protein synthesis mechanisms highlights the role of p53 as a master cancer regulator.

1.2.8. Potential sources of novel miRNAs under p53 regulation

The widespread regulatory spectrum of p53, which involves various RNA pol I, II and III-transcribed RNAs, led to the question whether non-coding RNAs can be processed into functional miRNAs that play a role in the p53 regulatory network. Small RNA fragments have been found originating from not only

rRNAs and tRNAs, but also small nucleolar RNAs (snoRNAs), small nuclear RNAs (snRNAs) and other non-coding RNAs [118-121]. Importantly, many of these products are loaded to the Argonaute (AGO) proteins, which are a central part of the post-transcriptional regulation activity of the RISC [118, 122-124]. In addition, physiological functions have been attributed to these small RNA fragments, e.g., regulation in stress related to hypoxia and starvation, transcription, and miRNA-like gene inhibition [125-129]. While the small RNAs derived from tRNAs, 5S rRNA, yRNAs, vtRNAs (vault RNAs) and snRNAs generally do not serve miRNA-like targeting roles, small RNAs processed from C/D box snoRNAs are capable to function like canonical miRNAs [130].

snoRNAs are a highly evolutionarily conserved class of RNAs, which are present throughout the Eukaryotes. They were traditionally known to function as ribonucleoprotein (RNP) complexes to guide the modification and processing of other non-coding RNAs such as rRNAs or snRNAs in the nucleus through RNA: RNA antisense interaction [131]. Recent immunoprecipitation of AGO1 and AGO2 proteins revealed several smaller processed products of snoRNAs function similarly to miRNAs, being termed snoRNAs-derived microRNAs (sno-miRNAs, sno-miRs) [132]. Ender *et al*. reported that the snoRNA ACA45 is processed into small 20- to 25-nt-long RNAs that stably associate with AGO proteins [132]. The processing is independent of the

Drosha/DGCR8 complex but requires Dicer. They also identified a cellular target mRNA, CDC2L6, whose 3'-UTR (untranslated region) was regulated by this small RNA, indicating that snoRNA-derived small RNAs can function like miRNAs. Furthermore, Ono *et al*. reported that computational analyses identified 84 intronic miRNAs that are encoded within either box C/D snoRNAs, or in precursors showing similarity to box C/D snoRNAs [133]. Nevertheless, whether p53 regulates the expression of functional sno-miRNAs is still to be explored.

Due to the limits of the traditional ncRNA profiling methods by pre-selecting gene sets, as aforementioned, the distribution of snoRNAs and sno-miRNAs and their regulatory roles are poorly understood. The next generation deep sequencing techniques, which have a wider and unbiased output range, can hopefully help us to discover more about the nature and functions.

1.3. Post- transcriptional roles of long non-coding RNAs (lncRNAs) in the p53 downstream network

1.3.1. Introduction to lncRNAs

Although miRNA regulation is a hot topic of recent research, miRNA is not the only family of non-coding RNAs which have been discovered to play a role in the p53 regulatory network. Other vital non-coding RNAs involved in p53 regulation include long non-coding RNAs (lncRNAs), since more and more evidence has revealed lncRNA to have an important role on p53-induced cell cycle regulation. LncRNAs are longer than 200 nucleotides in length which can be transcribed from either protein-coding or intergenic regions of the genome, from introns or exons or a combination of both, and from either the sense or anti-sense strands [134]. The ENCODE program annotated 9277 lncRNA genes producing 14,880 transcripts [134]. Most lncRNAs are transcribed by RNA polymerase II, use similar splicing signals as mRNAs, and are normally 5'-capped and poly-A tailed [135, 136]. In addition, according to the authors, when compared to mRNAs, lncRNAs have less exons (most frequently 2 exons), and are relatively shorter in length. Furthermore, on average, they have lower expression levels, less evolutionary conservation, and lower stability than mRNAs [134]. However, Clark and Mattick discovered in mouse cells that a significant proportion of lncRNAs have high stability [137], suggesting biological roles.

1.3.2. Post-transcriptional regulatory roles of lncRNAs

Most lncRNAs are relatively more enriched in the nucleus [134], and ascribed basic functions include high-order chromosomal dynamics, telomere biology, subcellular structural organisation, and chromatin remodelling [138]. However, there is still a number of lncRNAs being enriched in the cytoplasm, and their regulatory roles are now being revealed [139-141]. For example, Zhang et al. discovered that LINC-ROR directly interacts with heterogeneous nuclear ribonucleoprotein I (hnRNP I) and represses p53 activation upon DNA damage, which results in inhibition of p53-mediated cell cycle arrest and apoptosis [24]. Similarly, lncRNA UCA1 interacts with hnRNP I and suppresses p27, thereby promoting breast tumour growth [142]. Recent evidence shows that several lncRNAs have a vital role in post-transcriptional regulation [138, 143, 144]. As Huarte *et al*. demonstrated, numerous lncRNAs are key constituents of the p53 pathway [145]. A considerable number of lncRNAs that have been systematically identified to act in a *trans* manner in post transcriptional regulation by complementary-recognition, but another class called eRNA (enhancer RNA) has been shown to function in *cis* in the p53 pathway to mediate p53-induced activation at distant locations [145-148]. In addition, using an approach of chromatin immunoprecipitation together with next-generation sequencing (ChIP-seq) investigation, 23 long intergenic non-coding RNAs (lincRNAs) have been found to be upregulated by p53 [149].

1.3.3. Regulatory roles of lncRNAs directly targeted by p53

One example of lncRNAs targeted by p53 is LincRNA-p21 [145]. The authors screened the expression of 400 lincRNAs using a microarray in p53 induction cell models. After validating a significant role of lincRNA-p21 in p53-induced apoptosis, the authors discovered that knockdown of either lincRNA-p21 or p53 resulted in a significant overlapping cohort of differentially expressed genes, most of which are repressed. To investigate the mechanism of lincRNA-p21 regulation, the authors discovered that its sense-strand associates with the heterogeneous nuclear ribonucleoprotein K (hnRNP-K), which is a component of a repressor complex that acts in the p53 pathway and exerts a global effect. The binding is dependent on a conserved 780 nt 5′ region of lincRNA-p21. Knockdown of lincRNA-p21 significantly reduced the binding of hnRNP-K to the promoter regions of genes that are normally repressed by lincRNA-p21 or p53, but the precise mechanisms of how lincRNA-p21 regulates specific loci remains to be defined. Notably, lincRNA-p21 is significantly downregulated in colorectal tumour tissues compared with normal tissue [150]. However, the same authors found no significant correlation of lincRNA-p21 with p53 status in these tumours.

Furthermore, another lncRNA directly regulated by p53 was also discovered in 2011 [146]. Using a custom microarray for the genome areas near several critical cell cycle regulators at a 5 nt resolution, Hung *et al*. identified PANDA (p21 Associated ncRNA DNA Damage Activated) as a functional lncRNA induced by p53 upon DNA damage. Like lincRNa-p21, the encoding region of PANDA is also located upstream of the *CDKN1A* (p21) gene. Although PANDA is induced by p53, it is not a linked transcript of *CDKN1A*, and its expression is independent of p21. PANDA mediates anti-apoptotic functions and knockdown of this lncRNA abolished the repression on a series of canonical apoptosis activators, including APAF1, BIK, FAS and LRDD. In addition, PANDA represses the occupancy of transcription factor NFYA at its target genes *CCNB1*, *FAS*, *BBC3* (*PUMA*) and *PMAIP1* (*NOXA*), resulting in repression of apoptosis.

Recently Liu *et al*. discovered a lncRNA, namely loc285194, which is directly induced by p53 [147]. The authors initialled a screen using an RT-PCR array to examine the expression of 83 lncRNAs after p53 activation and eventually validated that loc285194 is a direct target of p53. Since loc285194 inhibits tumour cell growth both *in vivo* and *in vitro*, the authors suggested that it is a p53 downstream tumour suppressor. To explain the regulatory mechanism of loc285194, the authors pointed out that it binds to the same RISC with miR-211 and both RNAs repress each other in a way similar to the microRNA-mediated

silencing of protein-coding genes. This is supported by the discovery that loc285194 is downregulated in colon cancer whereas miR-211 is upregulated. In addition, their results are also supported by a more recent report that low expression of loc285194 is associated with poor prognosis of colon cancer [151].

Another emerging family of longer non-coding RNAs in general are eRNAs (enhancer RNAs). In contrast to the abovementioned *trans*-regulation of lncRNAs, eRNAs tend to work in a *cis* manner [148]. However, our knowledge of eRNAs is rather limited. They are non-protein-coding, non-polyadenylated, and are generally shorter than mRNAs in length. Kim *et al*. reported that RNA pol II transcribes eRNAs bi-directionally within enhancer domains, and the expression of eRNAs at the enhancer sites positively correlates with the nearby protein-coding genes [152]. According to the recent work of Melo *et al*., eRNAs are widely involved in p53 regulation [148]. When exploring the transcription enhancement of p53 via distant locations by binding to p53-bound enhancer regions (p53BERs), the authors discovered that eRNAs were transcribed from p53BERs in a p53 dependent manner by RNA pol II. In this case eRNAs mediate the enhancement activity to p53 and regulate multiple distant genes at various genomic positions probably through pre-existing chromatin conformations. For instance, one of the targets identified by the authors, DUSP4, is at ~430kb from the p53BER that is responsible for its regulation.

1.4. Conclusion

There are a large number of ncRNAs that have already been identified to be vital regulators of biological significance, which are involved in major molecular pathways. In particular, recent discoveries of regulatory miRNAs and lncRNAs in the p53 pathway have been shown to regulate a crucial layer of the p53 regulatory network. Since the techniques to identify such ncRNAs have to this date been limited by the specialised approaches, it is suggested that there exists additional p53 regulatory networks that are yet to be identified. In Chapter 2, we will use next generation deep-sequencing to explore miRNAs derived from snoRNAs and their regulatory roles in the p53 network.

2. IsomiRs – naturally existing isoforms of miRNAs

2.1. The maturation of miRNA and its heterogeneity

miRNAs are well established to be critical regulatory molecules in human cells, and function as 21-24 nt guides to recognize mRNA sequences by sequence complementary. This is mediated by the AGO proteins which function as a central effector of the RISC (RNA-induced silencing complex) [153, 154]. The conventional processing of miRNAs involves multiple proteins that are organized and regulated [154]. The generation of miRNAs first involves transcription of the primary miRNA (pri-miRNA) from an independent host gene, an intronic area, or other sequences. The pri-miRNA is folded into hairpins and then cleaved by Drosha to form precursor miRNAs (pre-miRNAs) of 55-70 nt in length, followed by Dicer to cleave both strands at about 22nt from the 5′ and 3′ ends of the pre-miRNA hairpin to produce the mature form [154, 155]. The cleavage of the precursors by Drosha and Dicer is not always precise and the products are often heterogeneous in length [156]. These naturally occurring heterogenic miRNA species are collectively termed "isomiRs".

The specificity of miRNA cleavage is dependent on a variety of factors including sequence motifs around the cleavage sites on both the pri-miRNA and pre-miRNA, and also the secondary structure of the miRNA precursors [157]. The imprecise processing of Drosha and Dicer gives rise to heterogeneity of both the 5′ and 3′- ends of miRNAs. Surprisingly, the miRNA sequences annotated in miRBase are not always the most abundant forms and the relative proportion of isomiR expression levels are often variable in different tissues [158-160] (Table 2.1). In addition, the proportions of isoforms coded by the same miRNA gene are highly variable across different cell and tissue types [161].

miR-222 isoforms	**Reads per million**
CAGCUACAUCUGGCUACUGGGUCU****	0.412
CAGCUACAUCUGGCUACUGGGUCUC***	0.244
CAGCUACAUCUGGCUACUGGGUCUCU**	0.21
CAGCUACAUCUGGCUACUGGGU******	0.152
CAGCUACAUCUGGCUACUGGG*******	0.0507
CAGCUACAUCUGGCUACUGGGUC*****	3.79
*AGCUACAUCUGGCUACUGGGUC*****	913
*AGCUACAUCUGGCUACUGGGUCUCU**	621

*AGCUACAUCUGGCUACUGGG*******	181
*AGCUACAUCUGGCUACUGG********	20.8
*AGCUACAUCUGGCUAC**********	32.7
*AGCUACAUCUGGCUACUG*********	7.4
AGCUACAUCUGGCUACUGGGU***	8.94
*AGCUACAUCUGGCUA**********	40.7
*AGCUACAUCUGGCUACU**********	10.7
*AGCUACAUCUGGCUACUGGGUCUC***	2.17
*AGCUACAUCUGGCUACUGGGUCU****	3.48
AGCUACAUCUGGCUACUGGGUCUCUG	0.294
*AGCUACAUCUGGCUACUGGGUCUCUGA	0.0899
GCUACAUCUGGCUACUGGGUCU**	10.7
GCUACAUCUGGCUACUGGGUCUC*	4.95
GCUACAUCUGGCUACUGGGU****	6.47
GCUACAUCUGGCUACUGGGUC***	2.45
GCUACAUCUGGCUACUGGGUCUCU	1.91
GCUACAUCUGGCUACUGGG*****	0.113
GCUACAUCUGGCUAC********	0.128
GCUACAUCUGGCUACU********	0.128
CUACAUCUGGCUACUGGGUCU*	10.3
CUACAUCUGGCUACUGGGUCUC	2.02

CUACAUCUGGCUACUGGGU***	3.48
***CUACAUCUGGCUACUGGGUCUCU**	1
CUACAUCUGGCUACUGGGUC**	1.22
CUACAUCUGGCUACUGGG****	0.82
****UACAUCUGGCUACUGGGUCU****	5.3
****UACAUCUGGCUACUGGGUCUC***	0.406
****UACAUCUGGCUACUGGGUC*****	3.64
****UACAUCUGGCUACUGGGUCUCU**	0.263
****UACAUCUGGCUACUGGGU******	1.02
*****ACAUCUGGCUACUGGGUCU****	4.45
*****ACAUCUGGCUACUGGGUC*****	5.05
*****ACAUCUGGCUACUGGGU******	3.54
*****ACAUCUGGCUACUGGG*******	0.331
******CAUCUGGCUACUGGGUCU****	17.3
******CAUCUGGCUACUGGGU******	6.79
******CAUCUGGCUACUGGGUC*****	4.02
******CAUCUGGCUACUGGGUCUCU**	0.154
******CAUCUGGCUACUGGGUCUC***	0.468
*******AUCUGGCUACUGGGUCU****	17.9
*******AUCUGGCUACUGGGUC*****	24.6
*******AUCUGGCUACUGGGU******	19.3

*******AUCUGGCUACUGGGUCUC***	2.77
*******AUCUGGCUACUGGGUCUCU**	6.06
********UCUGGCUACUGGGUCU****	12.7
********UCUGGCUACUGGGUC*****	8.67
********UCUGGCUACUGGGUCUC***	1
*********CUGGCUACUGGGUCU****	26.6
*********CUGGCUACUGGGUCUC***	5.86
**********GGCUACUGGGUCUCU**	2.27

Table 2.1. 5′ and 3′ isoforms of miR-222.

The isoforms of miR-222 and their expression levels as observed from RNA deep-sequencing are listed according to the miRBase (Release 21: June 2014) [162], in which the currently annotated miR-222 sequence is AGCUACAUCUGGCUACUGGGU. Asterisks (*) indicate blanks.

2.1.1. The existence of 5′ and 3′ isomiRs

The 5′ variation of miRNAs (5′ isomiRs) probably has a more prominent impact on function though it represents a rarer class of isomiRs compared with 3′ variants. 5′-variants have been reported among miR-142-3p, isomiR-9-1 and

isomiR-101 [161, 163, 164].

The 3′ variation of miRNAs, however, is more abundant than 5′ variation, but has not been thoroughly researched [165-167]. A major technical obstacle preventing progresses in investigating 3′ isomiR variation is that the conventional stem-loop RT-PCR often cross-detects isomiRs with 3′ variation [168]. However, as stem-loop PCR techniques were constantly refined [168], and RNA deep sequencing technologies were developed, it has now become apparent that the addition of nucleotides at the 3′ end of miRNAs is not a random event, and the expression patterns of these isomiRs suggest functional and evolutionary relationships [165, 169]. It has been found in *Drosophila* that there are 3′ variations of miR-282 and miR-312 with abundant 3′ adenosine extensions [170].

2.2. Sources of miRNA heterogeneity

Several lines of evidence suggest that isomiRs are generated in an organized way. By examining the RNA deep-sequencing data in human cell lines, mouse embryonic stem cells, and *Drosophila* head tissue, Burroughs et al. discovered that the variation of isomiRs is globally present and conserved across these species [171]. In addition, analysis of published small RNA libraries revealed

that the relative proportion of isomiRs is altered accompanying the differentiation of human embryonic stem cells [166]. In addition, miR-8 is the most abundant isoform in the embryos while uridine-additions are frequently observed in female heads [170], and the fraction of uridylated Let-7 precursors increases with development and peaks in adult mouse tissues [172]. These three reports above are consistent with the regulation of isomiR expression during development. Another example can be found in the report of Etebari et al. that the isomiR abundance is subject to change upon dengue virus infection in mosquitos, which suggests possible regulation upon the biological stimuli related to virus replication [173]. Various bioinformatic analysis indicates that the heterogeneity of miRNAs is biologically relevant of different species and life activities [171, 174-176].

2.2.1. 5′ and 3′ heterogeneity derived from the biogenesis of isomiRs

The major processing pathway of miRNA biosynthesis, as described above in Section 2.2, is a significant source of both 5′ and 3′ miRNA variation [177]. *In vitro* RNA cleavage experiments using synthetic pri- and pre-miRNAs have revealed that both Dicer and Drosha cleavages are responsible for miRNA variability [178]. Several reports indicate that miRNA cleavage is less precise at

the Dicer stage of processing [178, 179], though Wu et al. reported in their model that Drosha was the dominating factor for miRNA heterogeneity [180]. In addition, Dicer partner proteins have been found to bind to Dicer and induce alternative cleavage. The *Drosophila* Dicer partner protein Loqs-PA tunes Dicer to generate a 21mer miR-307a whereas the other Dicer partner protein, Loqs-PB, introduces a 23mer miR-307a isoform that exhibits distinctive target specificity [181]. The diversity of Dicer cleavage is shown to be related to the precursor hairpin structure, e.g. internal bulges [173, 178].

2.2.2. 3′ heterogeneity derived from trimming and nucleotidyl transferring

Exoribonucleases are a significant source of AGO-bound miRNAs with varying 3′ ends. *Nibbler* (*Nbr*) is an example which interacts with AGO1 and trims the *Drosophila miR-34* within the RISC [182], with sequencing revealing more than a quarter of *Drosophila* miRNAs are trimmed by Nibbler at the 3′ end [182, 183]. Notably, when *Nibbler* is mutated, not all 3′ isomiRs are eliminated, which suggests *Nibbler* is not the only driving mechanism of miRNA trimming. Similarly, a 3′-5′ exonuclease, *QIP*, is involved in the maturation of *Neurospora crassa* miR-1 and trims its 3′ end in co-operation with the RNA exosome [184]. Less research has been conducted in the area of human ribonucleases involved

in miRNA diversity, but homologues of *Nibbler* and *QIP* may also exist in the human genome, playing similar roles as their *Drosophila* counterparts.

In addition to the removal of nucleotides by exoribonucleases, RNA-specific ribonucleotidyl transferases can add nucleotides to miRNAs. These transferases are a diverse family of template-independent polymerases which add 3′ ribonucleotide extensions to RNA [185]. Several are extensively involved in human miRNA processing, e.g., GLD-2 (PAPD4) [166, 171, 186], PAPD5 [166, 171], MTPAP, ZCCHC6, ZCCHC11, and TUT1 [166]. These ribonucleotidyl transferases mostly function as uridyltransferases or adenyltransferases or both, which is consistent with the bioinformatic analysis that indicates most un-templated additions at the 3′ end are uridines or adenosines [187-189]. A functional implication of ribonucleotidyl transfer has been revealed through PAPD5-mediated adenylation which results in destabilization of miR-21. Especially interesting is that PAPD5-mediated 3′ adenylation is disrupted in cancer and other proliferative diseases, implying that this may be a mechanism contributing to cancer progression [190].

2.2.3. Heterogeneity derived from RNA editing

RNA editing can also contribute to miRNA heterogeneity. The most frequent is adenosine deamination, which results in adenosine to inosine (A to I) editing, and is mediated by the adenosine deaminase acting on RNA (ADAR) enzymes [191]. Both pri-miRNAs and pre-miRNAs are possible substrates, and Kawahara Y et al. conducted an *in vitro* miRNA processing assay using recombinant Drosha-DGCR8 and Dicer-TRBP which suggested pri-miRNA editing may interfere with the miRNA maturation steps [192]. The same authors also estimated that about 16% of the pri-miRNAs in humans are subject to A to I editing [192]. However, their theory was challenged by more recent findings that RNA editing in mature miRNAs is rare [193], and previous findings were interfered by the background noise from sequencing or mismatches between homologues [189, 193].

Nevertheless, though not being as prevalent as initially proposed, there are still confirmed examples of functional impacts of RNA editing. miR-376 is subject to tissue specific A to I editing and a frequently edited site located in its seed region that is likely to be critical for its target recognition capability [194, 195]. The proportion of the edited and non-edited isoforms of miR-376 gradually changes during development as editing increases, which is consistent with editing-driven developmental regulation through isomiRs [196]. Another example is miR-455-5p which has a different set of targets after being edited by

ADAR, abolishing its inhibitory function of CPEB1. The edited form of miR-455-5p suppresses melanoma progression [197]. A comprehensive investigation carried out by Peng Z et al. found 44 editing sites in miRNAs and an additional ~2500 editing sites within 3'-UTRs that may affect their miRNA binding properties [198]. The editing pattern of miRNAs was distinctive from that of the broad transcriptome, as the A to I editing is much less frequent compared to the 93% overall rate of miRNA editing, and other types of editing were also detected, including G (guanidine) to A (adenosine), U (uridine) to A, T to G, C (cytosine) to U, etc. [198] However, while bioinformatic approaches are being commonly employed, more experimental evidences are still required to fully explore the role of RNA editing in isomiRs.

2.3. Functional impacts of miRNA heterogeneity

A major interest in isomiR research is the functional implications of isomiRs and how the heterogeneity of miRNAs affects common physiological activities, e.g. metabolism and cancer. A few examples of functional analysis of isomiR biogenesis have been discussed above, showing that individual isomiRs may vary in their sets of target genes or in other regulatory roles. However, more data is still required to build up a comprehensive view of the functional properties of isomiRs.

2.3.1. Alteration in target recognition

As aforementioned, 5′-variation of miRNAs may alter the seed-region that is determined by its position relative to the 5′ end of the miRNA, typically based on the nucleotides positioned at 2-8nt [199]. However, non-canonical recognition patterns of miRNAs have been discovered and a perfect match at the seed region is not the only way for a miRNA to regulate its targets [200-203]. On this basis, Manzano M et al. divided 5′ isomiRs into two distinct classes with divergent or convergent target repertoires, based on 5′ sequence features [163]. The first class with divergent target repertoires is represented by naturally co-expressed 5′ isomiRs of miR-142-3p which recognise not only shared targets involved in actin dynamics but also extended sets of binding targets. In contrast, two miR-142-3p 5′ isomiRs originated from herpes virus had overlapping binding sets. Additionally, the 5′ isomiR of miR-9 targets DNMT3B and NCAM2 instead of CDH1, which is a target of the canonical miR-9 [161, 175], however 5′-isomiR-101 recognizes the same target set as its main form [164]. Similar reports include differential mRNA targeting by two of miR-133a's prevalent 5′ isomiRs [204], and knockdown of both mouse *miR-223* and its 5′ isomiR depressed the same cohort of target mRNAs despite the seed shift [189].

In addition, RNA editing can also alter the miRNA seed region, with the abovementioned example of miR-455-5p undergoing A to I editing which reversed its oncogenic properties. Similarly, the A to I editing of miR-376 contributes to uric acid homeostasis [195].

In contrast to 5′ variations, the 3′ variations of miRNAs do not change their canonical seed recognition sites. Nevertheless, miRNA target research using newly developed approaches revealed that the 3′ region of miRNAs should not be neglected in terms of targeting function, and vital target recognition roles could be closely related to motifs located near the 3′ region [205]. Burroughs et al. reported that 3′ adenine addition reduces the effectiveness of miRNA targeting mRNA transcripts [171]. This suggests 3′ variations of miRNAs may also interfere with miRNA-target recognition and more experimental investigation is required to explore their functional implications.

2.3.2. Alteration in AGO loading

The Argonaute (AGO) proteins are the essential components of the RISC, and the human AGO family has eight members, with four of which (AGO1 through 4) being ubiquitously expressed and are the most extensively researched [206].

AGO1 to 4 proteins have partially overlapping but non-redundant roles in miRNA-mediated gene silencing, e.g., the RNA endonuclease activity is exclusively associated with AGO2 [206-208]. Since isomiRs are loaded preferentially to different AGO proteins, the alteration in AGO loading may present a factor of their functional impacts [209-212].

A significant amount of AGO-loading research has been in plants and invertebrates. In *Arabidopsis*, a 5′ U directs miRNAs to AGO1, a 5′ A to AGO2/4, and a 5′ C to AGO5. Changing the 5′ terminal nucleotide of a miRNA redirects it to a different AGO complex, thus altering its biological function [209-211]. In *Drosophila*, AGO1 prefers a U at the 5′ end, and AGO2 favours a C [213-215]. Furthermore, the 3′ variation of miRNAs can also affect AGO loading, inferred from the loss of AGO2-loading specificity by the removal of a 3′ G of miR-182 as this isomiR of miR-183 associates to all of AGO1, 2 and 3 [118]. However, whether this phenomenon of isomiR-183 AGO loading was a reason or a consequence of 3′ modification of miR-182 is not clear, since this differential loading could also be explained as the miRNA being modified subsequent to AGO binding [216, 217].

2.3.3. Alteration in miRNA stability

The stability of miRNAs is a vital factor for their observed physiological function and has complex regulatory mechanisms which are largely unknown [218, 219]. The discovery that miR-553 instability was mediated by both its seed region and 3' end sequence suggested that alteration of the 3' sequence may dramatically change miRNA stability [220]. In terms of isomiRs, GLD2, the adenosine deaminase that induces addition of A, has been found to stabilize miR-122 [186]. However, a word of caution should be raised as there are also conflicting examples, e.g., the addition of terminal uridines to the miR-26 3' end did not significantly alter its stability [221].

2.4. Discussion

Our knowledge about isomiRs is still far from complete, and the earlier research involving miRNAs did not explicitly distinguish between naturally existing isomiRs. Most previous research did not specify the exact sequence or target sequences when they report the use of miRNA mimics/pre-miRNAs, or miRNA inhibitors, which may lead to ambiguity when interpreting the potential difference in isomiR behaviours. The relative proportion of homologous isomiRs is variable, subject to regulation upon development, infection and are tissue specific [166, 170, 172, 173]. Moreover, in extreme cases, the annotated form of certain miRNAs is not the most abundant isoform, e.g.,

the annotated form of miR-222 is 21nt in length, which is expressed less than 1/50 compared to that of miR-222C and miR-222CUCU (Table 2.1) [158-160]. The functional research of miR-222 appears to be controversial, since miR-222 has been repeatedly reported to be tumour suppressive in castration-resistant prostate cancer [222], non-small cell lung cancer [223], gastrointestinal stromal tumours [224] and breast cancer [225], but has also been extensively shown to be an oncogenic miRNA [226-229]. Nevertheless, rarely these reports clearly specified the miR-222 species being used for transfection and the miRNA quantitation methods using stem-loop PCR can often cross-detect isomiR variants [168]. Furthermore, it is worth noting that the 3' variation, which has once been considered functionally neutral and neglected, may significantly alter the molecular biological and physiological functions of the miRNA. Therefore, these findings suggest a major limitation in published miRNA research. Chapter 3 of this thesis will explore the functional implication of isomiRs using miR-222 as a model.

3. Summary of the chapter

The existence of miRNAs processed from novel sources and isomiRs, which will be investigated in detail in the next chapters, has added to the complexity of our knowledge about miRNA regulation. These areas are much less explored

than conventional miRNA processing and regulation. The profound molecular biological and physiological implications of isomiRs suggests a novel layer of miRNA biology that will influence the interpretation of previous molecular biology research and provide new networks of molecular pathways in cancer.

References

1. Efeyan, A. and M. Serrano, *p53: guardian of the genome and policeman of the oncogenes.* Cell Cycle, 2007. **6**(9): p. 1006-10.
2. Lane, D.P., *Cancer. p53, guardian of the genome.* Nature, 1992. **358**(6381): p. 15-6.
3. Crighton, D., et al., *DRAM, a p53-induced modulator of autophagy, is critical for apoptosis.* Cell, 2006. **126**(1): p. 121-34.
4. Roger, L., G. Gadea, and P. Roux, *Control of cell migration: a tumour suppressor function for p53?* Biology of the Cell, 2006. **98**(3): p. 141-152.
5. Vousden, K.H. and D.P. Lane, *p53 in health and disease.* Nat Rev Mol Cell Biol, 2007. **8**(4): p. 275-83.
6. Vogelstein, B., D. Lane, and A.J. Levine, *Surfing the p53 network.* Nature, 2000. **408**(6810): p. 307-10.
7. Levine, A.J., *Introduction: The Changing Directions of p53 Research.* Genes Cancer, 2011. **2**(4): p. 382-4.
8. Helton, E.S. and X. Chen, *p53 modulation of the DNA damage response.* J Cell Biochem, 2007. **100**(4): p. 883-96.
9. Menendez, D., A. Inga, and M.A. Resnick, *The expanding universe of p53 targets.* Nat Rev Cancer, 2009. **9**(10): p. 724-37.

10. Espinosa, J.M. and B.M. Emerson, *Transcriptional regulation by p53 through intrinsic DNA/chromatin binding and site-directed cofactor recruitment.* Mol Cell, 2001. **8**(1): p. 57-69.

11. Nakano, K. and K.H. Vousden, *PUMA, a novel proapoptotic gene, is induced by p53.* Mol Cell, 2001. **7**(3): p. 683-94.

12. Ragimov, N., et al., *Wild-type but not mutant p53 can repress transcription initiation in vitro by interfering with the binding of basal transcription factors to the TATA motif.* Oncogene, 1993. **8**(5): p. 1183-93.

13. Beckerman, R. and C. Prives, *Transcriptional regulation by p53.* Cold Spring Harb Perspect Biol, 2010. **2**(8): p. a000935.

14. Kim, E. and W. Deppert, *The versatile interactions of p53 with DNA: when flexibility serves specificity.* Cell Death Differ, 2006. **13**(6): p. 885-9.

15. Ho, J. and S. Benchimol, *Transcriptional repression mediated by the p53 tumour suppressor.* Cell Death Differ, 2003. **10**(4): p. 404-8.

16. Koumenis, C., et al., *Regulation of p53 by hypoxia: dissociation of transcriptional repression and apoptosis from p53-dependent transactivation.* Mol Cell Biol, 2001. **21**(4): p. 1297-310.

17. Mattick, J.S., *RNA regulation: a new genetics?* Nat Rev Genet, 2004. **5**(4): p. 316-23.

18. Ponting, C.P., P.L. Oliver, and W. Reik, *Evolution and functions of long noncoding RNAs.* Cell, 2009. **136**(4): p. 629-41.

19. Wang, H., et al., *Dendritic BC1 RNA: functional role in regulation of translation initiation.* J Neurosci, 2002. **22**(23): p. 10232-41.

20. Wang, H., et al., *Dendritic BC1 RNA in translational control mechanisms.* J Cell Biol, 2005. **171**(5): p. 811-21.

21. Lim, L.P., et al., *Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs.* Nature, 2005. **433**(7027): p. 769-73.

22. Yan, M.D., et al., *Identification and characterization of a novel gene Saf transcribed from the opposite strand of Fas.* Hum Mol Genet, 2005. **14**(11): p. 1465-74.

23. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. **75**(5): p. 843-54.

24. Zhang, A., et al., *The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage.* Cell Res, 2013. **23**(3): p. 340-50.

25. Calin, G.A. and C.M. Croce, *MicroRNA signatures in human cancers.* Nat Rev Cancer, 2006. **6**(11): p. 857-66.

26. Bartel, D.P., *MicroRNAs: Target Recognition and Regulatory Functions.* Cell, 2009. **136**(2): p. 215-233.

27. Cenik, E.S. and P.D. Zamore, *Argonaute proteins.* Curr Biol, 2011. **21**(12): p. R446-9.

28. Griffiths-Jones, S., et al., *miRBase: microRNA sequences, targets and gene nomenclature.* Nucleic Acids Res, 2006. **34**(Database issue): p. D140-4.

29. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs.* Genome Res, 2009. **19**(1): p. 92-105.

30. Shenouda, S.K. and S.K. Alahari, *MicroRNA function in cancer: oncogene or a tumor suppressor?* Cancer Metastasis Rev, 2009. **28**(3-4): p. 369-78.

31. Garzon, R., G.A. Calin, and C.M. Croce, *MicroRNAs in Cancer.* Annu Rev Med, 2009. **60**: p. 167-79.

32. Taguchi, A., et al., *Identification of hypoxia-inducible factor-1 alpha as a novel target for miR-17-92 microRNA cluster.* Cancer research, 2008. **68**(14): p. 5540-5.

33. Zhu, S., et al., *MicroRNA-21 targets tumor suppressor genes in invasion and metastasis.* Cell research, 2008. **18**(3): p. 350-9.

34. Cimmino, A., et al., *miR-15 and miR-16 induce apoptosis by targeting BCL2.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(39): p. 13944-9.

35. Johnson, S.M., et al., *RAS is regulated by the let-7 microRNA family.* Cell, 2005. **120**(5): p. 635-47.

36. Kwak, P.B., S. Iwasaki, and Y. Tomari, *The microRNA pathway and cancer.* Cancer Sci, 2010. **101**(11): p. 2309-15.

37. Lu, J., et al., *MicroRNA expression profiles classify human cancers.* Nature, 2005. **435**(7043): p. 834-8.

38. Ventura, A. and T. Jacks, *MicroRNAs and cancer: short RNAs go a long way.* Cell, 2009. **136**(4): p. 586-91.

39. Feng, Z., et al., *Tumor suppressor p53 meets microRNAs.* J Mol Cell Biol, 2011. **3**(1): p. 44-50.

40. Chang, T.C., et al., *Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis.* Mol Cell, 2007. **26**(5): p. 745-52.

41. Corney, D.C., et al., *MicroRNA-34b and MicroRNA-34c are targets of p53 and cooperate in control of cell proliferation and adhesion-independent growth.* Cancer Res, 2007. **67**(18): p. 8433-8.

42. He, L., et al., *A microRNA component of the p53 tumour suppressor network.* Nature, 2007. **447**(7148): p. 1130-4.

43. Raver-Shapira, N., et al., *Transcriptional activation of miR-34a contributes to p53-mediated apoptosis.* Mol Cell, 2007. **26**(5): p. 731-43.

44. Tazawa, H., et al., *Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells.* Proc Natl Acad Sci U S A, 2007. **104**(39): p. 15472-7.

45. Cole, K.A., et al., *A functional screen identifies miR-34a as a candidate neuroblastoma tumor suppressor gene.* Mol Cancer Res, 2008. **6**(5): p. 735-42.

46. Hermeking, H., *The miR-34 family in cancer and apoptosis.* Cell Death Differ, 2010. **17**(2): p. 193-9.

47. Yamakuchi, M., M. Ferlito, and C.J. Lowenstein, *miR-34a repression of SIRT1 regulates apoptosis.* Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13421-6.

48. Du, R., et al., *Hypoxia-induced down-regulation of microRNA-34a promotes EMT by targeting the Notch signaling pathway in tubular epithelial cells.* PLoS One, 2012. **7**(2): p. e30771.

49. Welch, C., Y. Chen, and R.L. Stallings, *MicroRNA-34a functions as a potential tumor suppressor by inducing apoptosis in neuroblastoma cells.* Oncogene, 2007. **26**(34): p. 5017-22.

50. Silber, J., et al., *miR-34a repression in proneural malignant gliomas upregulates expression of its target PDGFRA and promotes tumorigenesis.* PLoS One, 2012. **7**(3): p. e33844.

51. Christoffersen, N.R., et al., *p53-independent upregulation of miR-34a during oncogene-induced senescence represses MYC.* Cell Death Differ, 2010. **17**(2): p. 236-45.

52. Wiggins, J.F., et al., *Development of a lung cancer therapeutic based on the tumor suppressor microRNA-34.* Cancer Res, 2010. **70**(14): p. 5923-30.

53. Lodygin, D., et al., *Inactivation of miR-34a by aberrant CpG methylation in multiple types of cancer.* Cell Cycle, 2008. **7**(16): p. 2591-600.

54. Liu, C., et al., *The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44.* Nat Med, 2011. **17**(2): p. 211-5.

55. de Antonellis, P., et al., *MiR-34a targeting of Notch ligand delta-like 1 impairs CD15+/CD133+ tumor-propagating cells and supports neural differentiation in medulloblastoma.* PLoS One, 2011. **6**(9): p. e24584.

56. Kim, N.H., et al., *p53 and microRNA-34 are suppressors of canonical Wnt signaling.* Sci Signal, 2011. **4**(197): p. ra71.

57. Bader, A.G., *miR-34 - a microRNA replacement therapy is headed to the clinic.* Front Genet, 2012. **3**: p. 120.

58. Kojima, K., et al., *MiR-34a attenuates paclitaxel-resistance of hormone-refractory prostate cancer PC3 cells through direct and indirect mechanisms.* Prostate, 2010. **70**(14): p. 1501-12.

59. Akao, Y., et al., *Dysregulation of microRNA-34a expression causes drug-resistance to 5-FU in human colon cancer DLD-1 cells.* Cancer Lett, 2011. **300**(2): p. 197-204.

60. Kim, T., et al., *p53 regulates epithelial-mesenchymal transition through microRNAs targeting ZEB1 and ZEB2.* J Exp Med, 2011. **208**(5): p. 875-83.

61. Braun, C.J., et al., *p53-Responsive micrornas 192 and 215 are capable of inducing cell cycle arrest.* Cancer Res, 2008. **68**(24): p. 10094-104.

62. Vallejo, D.M., E. Caparros, and M. Dominguez, *Targeting Notch signalling by the conserved miR-8/200 microRNA family in development and cancer cells.* Embo J, 2011. **30**(4): p. 756-69.

63. Chang, C.J., et al., *p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs.* Nat Cell Biol, 2011. **13**(3): p. 317-23.

64. Brabletz, S., et al., *The ZEB1/miR-200 feedback loop controls Notch signalling in cancer cells.* Embo J, 2011. **30**(4): p. 770-82.

65. Boni, V., et al., *miR-192/miR-215 influence 5-fluorouracil resistance through cell cycle-mediated mechanisms complementary to its post-transcriptional thymidilate synthase regulation.* Mol Cancer Ther, 2010. **9**(8): p. 2265-75.

66. Pichiorri, F., et al., *Downregulation of p53-inducible microRNAs 192, 194, and 215 impairs the p53/MDM2 autoregulatory loop in multiple myeloma development.* Cancer Cell, 2010. **18**(4): p. 367-81.

67. Suzuki, H.I., et al., *Modulation of microRNA processing by p53.* Nature, 2009. **460**(7254): p. 529-33.

68. Fabbri, M., et al., *Association of a microRNA/TP53 feedback circuitry with pathogenesis and outcome of B-cell chronic lymphocytic leukemia.* Jama, 2011. **305**(1): p. 59-67.

69. Ofir, M., D. Hacohen, and D. Ginsberg, *MiR-15 and miR-16 are direct transcriptional targets of E2F1 that limit E2F-induced proliferation by targeting cyclin E.* Mol Cancer Res, 2011. **9**(4): p. 440-7.

70. Cai, C.K., et al., *miR-15a and miR-16-1 downregulate CCND1 and induce apoptosis and cell cycle arrest in osteosarcoma.* Oncol Rep, 2012. **28**(5): p. 1764-70.

71. Bohlig, L., M. Friedrich, and K. Engeland, *p53 activates the PANK1/miRNA-107 gene leading to downregulation of CDK6 and p130 cell cycle proteins.* Nucleic Acids Res, 2011. **39**(2): p. 440-53.

72. Yamakuchi, M., et al., *P53-induced microRNA-107 inhibits HIF-1 and tumor angiogenesis.* Proc Natl Acad Sci U S A, 2010. **107**(14): p. 6334-9.

73. Liang, M., et al., *Transcriptional cooperation between p53 and NF-kappaB p65 regulates microRNA-224 transcription in mouse ovarian granulosa cells.* Mol Cell Endocrinol, 2013. **370**(1-2): p. 119-29.

74. Yao, G., et al., *MicroRNA-224 is involved in transforming growth factor-beta-mediated mouse granulosa cell proliferation and granulosa cell function by targeting Smad4.* Mol Endocrinol, 2010. **24**(3): p. 540-51.

75. Yan, H.L., et al., *Repression of the miR-17-92 cluster by p53 has an important function in hypoxia-induced apoptosis.* EMBO J, 2009. **28**(18): p. 2719-32.

76. Ivanovska, I., et al., *MicroRNAs in the miR-106b family regulate p21/CDKN1A and promote cell cycle progression.* Mol Cell Biol, 2008. **28**(7): p. 2167-74.

77. Li, C.L., et al., *Elevated p53 promotes the processing of miR-18a to decrease estrogen receptor-alpha in female hepatocellular carcinoma.* Int J Cancer, 2014.

78. Lin, J., et al., *A novel p53-miR-22-Cyr61 axis in synovial cells regulates inflammation in Rheumatoid Arthritis.* Arthritis Rheum, 2013.

79. Jin, L., et al., *MicroRNA-149*, a p53-responsive microRNA, functions as an oncogenic regulator in human melanoma.* Proc Natl Acad Sci U S A, 2011. **108**(38): p. 15840-5.

80. Xiao, J., et al., *miR-605 joins p53 network to form a p53:miR-605:Mdm2 positive feedback loop in response to stress.* EMBO J, 2011. **30**(24): p. 5021.

81. Barsotti, A.M., et al., *p53-Dependent induction of PVT1 and miR-1204.* J Biol Chem, 2012. **287**(4): p. 2509-19.

82. Nakazawa, K., N. Dashzeveg, and K. Yoshida, *Tumor suppressor p53 induces miR-1915 processing to inhibit Bcl-2 in the apoptotic response to DNA damage.* FEBS J, 2014.

83. Ozsolak, F., et al., *Chromatin structure analyses identify miRNA promoters.* Genes Dev, 2008. **22**(22): p. 3172-83.

84. Neilsen, P.M., et al., *Mutant p53 drives invasion in breast tumors through up-regulation of miR-155.* Oncogene, 2012.

85. Sachdeva, M., et al., *p53 represses c-Myc through induction of the tumor suppressor miR-145.* Proc Natl Acad Sci U S A, 2009. **106**(9): p. 3207-12.

86. Zhu, H., et al., *EGFR signals downregulate tumor suppressors miR-143 and miR-145 in Western diet-promoted murine colon cancer: role of G1 regulators.* Mol Cancer Res, 2011. **9**(7): p. 960-75.

87. Ho, J.S., et al., *p53-Dependent transcriptional repression of c-myc is required for G1 cell cycle arrest.* Mol Cell Biol, 2005. **25**(17): p. 7423-31.

88. Hermeking, H., *MicroRNAs in the p53 network: micromanagement of tumour suppression.* Nat Rev Cancer, 2012. **12**(9): p. 613-26.

89. Leveille, N., et al., *Selective inhibition of microRNA accessibility by RBM38 is required for p53 activity.* Nat Commun, 2011. **2**: p. 513.

90. Vera, J., et al., *MicroRNA-regulated networks: the perfect storm for classical molecular biology, the ideal scenario for systems biology.* Adv Exp Med Biol, 2013. **774**: p. 55-76.

91. Ren, Z.J., et al., *Mir-509-5p joins the Mdm2/p53 feedback loop and regulates cancer cell growth.* Cell Death Dis, 2014. **5**: p. e1387.

92. Okada, N., et al., *A positive feedback between p53 and miR-34 miRNAs mediates tumor suppression.* Genes Dev, 2014.

93. Brosh, R., et al., *p53-Repressed miRNAs are involved with E2F in a feed-forward loop promoting proliferation.* Mol Syst Biol, 2008. **4**: p. 229.

94. Concepcion, C.P., et al., *Intact p53-dependent responses in miR-34-deficient mice.* PLoS Genet, 2012. **8**(7): p. e1002797.

95. Corney, D.C., et al., *Frequent downregulation of miR-34 family in human ovarian cancers.* Clin Cancer Res, 2010. **16**(4): p. 1119-28.

96. Cheng, C.Y., et al., *miR-34 cooperates with p53 in suppression of prostate cancer by joint regulation of stem cell compartment.* Cell Rep, 2014. **6**(6): p. 1000-7.

97. Chen, Y., et al., *Nanoparticles modified with tumor-targeting scFv deliver siRNA and miRNA for cancer therapy.* Mol Ther, 2010. **18**(9): p. 1650-6.

98. Pramanik, D., et al., *Restitution of tumor suppressor microRNAs using a systemic nanovector inhibits pancreatic cancer growth in mice.* Mol Cancer Ther, 2011. **10**(8): p. 1470-80.

99. Craig, V.J., et al., *Systemic microRNA-34a delivery induces apoptosis and abrogates growth of diffuse large B-cell lymphoma in vivo.* Leukemia, 2012. **26**(11): p. 2421-4.

100. Marchini, S., et al., *Association between miR-200c and the survival of patients with stage I epithelial ovarian cancer: a retrospective study of two independent tumour tissue collections.* Lancet Oncol, 2011. **12**(3): p. 273-85.

101. Burk, U., et al., *A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells.* EMBO Rep, 2008. **9**(6): p. 582-9.

102. Feng, S., et al., *MicroRNA-192 targeting retinoblastoma 1 inhibits cell proliferation and induces cell apoptosis in lung cancer cells.* Nucleic Acids Res, 2011. **39**(15): p. 6669-78.

103. Senanayake, U., et al., *miR-192, miR-194, miR-215, miR-200c and miR-141 are downregulated and their common target ACVR2B is strongly expressed in renal childhood neoplasms.* Carcinogenesis, 2012. **33**(5): p. 1014-21.

104. Chiang, Y., et al., *microRNA-192, -194 and -215 are frequently downregulated in colorectal cancer.* Exp Ther Med, 2012. **3**(3): p. 560-566.

105. Karaayvaz, M., et al., *Prognostic significance of miR-215 in colon cancer.* Clin Colorectal Cancer, 2011. **10**(4): p. 340-7.

106. Yang, J., et al., *Analysis of sequence variations in 59 microRNAs in hepatocellular carcinomas.* Mutat Res, 2008. **638**(1-2): p. 205-9.

107. Khella, H.W., et al., *miR-192, miR-194 and miR-215: a convergent microRNA network suppressing tumor progression in renal cell carcinoma.* Carcinogenesis, 2013. **34**(10): p. 2231-9.

108. Zhai, W. and L. Comai, *Repression of RNA polymerase I transcription by the tumor suppressor p53.* Mol Cell Biol, 2000. **20**(16): p. 5930-8.

109. Budde, A. and I. Grummt, *p53 represses ribosomal gene transcription.* Oncogene, 1999. **18**(4): p. 1119-24.

110. Deisenroth, C. and Y. Zhang, *Ribosome biogenesis surveillance: probing the ribosomal protein-Mdm2-p53 pathway.* Oncogene, 2010. **29**(30): p. 4253-60.

111. Li, J., et al., *Down-regulation of pescadillo inhibits proliferation and tumorigenicity of breast cancer cells.* Cancer Sci, 2009. **100**(12): p. 2255-60.

112. Iadevaia, V., et al., *PIM1 kinase is destabilized by ribosomal stress causing inhibition of cell cycle progression.* Oncogene, 2010. **29**(40): p. 5490-9.

113. Donati, G., et al., *Selective inhibition of rRNA transcription downregulates E2F-1: a new p53-independent mechanism linking cell growth to cell proliferation.* J Cell Sci, 2011. **124**(Pt 17): p. 3017-28.

114. Donati, G., L. Montanaro, and M. Derenzini, *Ribosome biogenesis and control of cell proliferation: p53 is not alone.* Cancer Res, 2012. **72**(7): p. 1602-7.

115. Crighton, D., et al., *p53 represses RNA polymerase III transcription by targeting TBP and inhibiting promoter occupancy by TFIIIB.* EMBO J, 2003. **22**(11): p. 2810-20.

116. Pavon-Eternod, M., et al., *tRNA over-expression in breast cancer and functional consequences.* Nucleic Acids Res, 2009. **37**(21): p. 7268-80.

117. Stein, T., et al., *RNA polymerase III transcription can be derepressed by oncogenes or mutations that compromise p53 function in tumours and Li-Fraumeni syndrome.* Oncogene, 2002. **21**(19): p. 2961-70.

118. Burroughs, A.M., et al., *Deep-sequencing of human argonaute-associated small RNAs provides insight into miRNA sorting and reveals argonaute association with RNA fragments of diverse origin.* Rna Biology, 2011. **8**(1): p. 158-177.

119. Zywicki, M., K. Bakowska-Zywicka, and N. Polacek, *Revealing stable processing products from ribosome-associated small RNAs by deep-sequencing data analysis.* Nucleic Acids Research, 2012. **40**(9): p. 4013-4024.

120. Kawaji, H., et al., *Hidden layers of human small RNAs.* Bmc Genomics, 2008. **9**.

121. Li, Z.H., et al., *Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs.* Nucleic Acids Research, 2012. **40**(14): p. 6787-6799.

122. Flores, O., et al., *Differential RISC association of endogenous human microRNAs predicts their inhibitory potential.* Nucleic Acids Research, 2014. **42**(7): p. 4629-4639.

123. Yamakawa, N., et al., *Novel functional small RNAs are selectively loaded onto mammalian Ago1.* Nucleic Acids Research, 2014. **42**(8): p. 5289-5301.

124. Burroughs, A.M., et al., *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, 2011. **8**(1): p. 158-77.

125. Haussecker, D., et al., *Human tRNA-derived small RNAs in the global regulation of RNA silencing.* Rna-a Publication of the Rna Society, 2010. **16**(4): p. 673-695.

126. Ivanov, P., et al., *Angiogenin-Induced tRNA Fragments Inhibit Translation Initiation.* Molecular Cell, 2011. **43**(4): p. 613-623.

127. Jochl, C., et al., *Small ncRNA transcriptome analysis from Aspergillus fumigatus suggests a novel mechanism for regulation of protein synthesis.* Nucleic Acids Research, 2008. **36**(8): p. 2677-2689.

128. Sobala, A. and G. Hutvagner, *Small RNAs derived from the 5 end of tRNA can inhibit protein translation in human cells.* Rna Biology, 2013. **10**(4): p. 553-563.

129. Yamasaki, S., et al., *Angiogenin cleaves tRNA and promotes stress-induced translational repression.* Journal of Cell Biology, 2009. **185**(1): p. 35-42.

130. Thomson, D.W., et al., *Assessing the gene regulatory properties of Argonaute-bound small RNAs of diverse genomic origin.* Nucleic Acids Research, 2015. **43**(1): p. 470-481.

131. Matera, A.G., R.M. Terns, and M.P. Terns, *Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs.* Nat Rev Mol Cell Biol, 2007. **8**(3): p. 209-20.

132. Ender, C., et al., *A human snoRNA with microRNA-like functions.* Mol Cell, 2008. **32**(4): p. 519-28.

133. Ono, M., et al., *Identification of human miRNA precursors that resemble box C/D snoRNAs.* Nucleic Acids Res, 2011. **39**(9): p. 3879-91.

134. Derrien, T., et al., *The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression.* Genome Res, 2012. **22**(9): p. 1775-89.

135. Guttman, M., et al., *Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals.* Nature, 2009. **458**(7235): p. 223-7.

136. Khalil, A.M., et al., *Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression.* Proc Natl Acad Sci U S A, 2009. **106**(28): p. 11667-72.

137. Clark, M.B. and J.S. Mattick, *Long noncoding RNAs in cell biology.* Semin Cell Dev Biol, 2011. **22**(4): p. 366-76.

138. Mercer, T.R., M.E. Dinger, and J.S. Mattick, *Long non-coding RNAs: insights into functions.* Nat Rev Genet, 2009. **10**(3): p. 155-9.

139. Panning, B., J. Dausman, and R. Jaenisch, *X chromosome inactivation is mediated by Xist RNA stabilization.* Cell, 1997. **90**(5): p. 907-16.

140. Rinn, J.L., et al., *Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs.* Cell, 2007. **129**(7): p. 1311-23.

141. Yang, L., et al., *lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs.* Nature, 2013.

142. Huang, J., et al., *Long non-coding RNA UCA1 promotes breast tumor growth by suppression of p27 (Kip1).* Cell Death Dis, 2014. **5**: p. e1008.

143. Amaral, P.P. and J.S. Mattick, *Noncoding RNA in development.* Mamm Genome, 2008. **19**(7-8): p. 454-92.

144. Li, C.H. and Y. Chen, *Targeting long non-coding RNAs in cancers: Progress and prospects.* Int J Biochem Cell Biol, 2013.

145. Huarte, M., et al., *A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response.* Cell, 2010. **142**(3): p. 409-19.

146. Hung, T., et al., *Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters.* Nat Genet, 2011. **43**(7): p. 621-9.

147. Liu, Q., et al., *LncRNA loc285194 is a p53-regulated tumor suppressor.* Nucleic Acids Res, 2013. **41**(9): p. 4976-87.

148. Melo, C.A., et al., *eRNAs are required for p53-dependent enhancer activity and gene transcription.* Mol Cell, 2013. **49**(3): p. 524-35.

149. Idogawa, M., et al., *Identification and analysis of large intergenic non-coding RNAs regulated by p53 family members through a genome-wide analysis of p53-binding sites.* Hum Mol Genet, 2014. **23**(11): p. 2847-57.

150. Zhai, H., et al., *Clinical significance of long intergenic noncoding RNA-p21 in colorectal cancer.* Clin Colorectal Cancer, 2013. **12**(4): p. 261-6.

151. Qi, P., et al., *Low expression of LOC285194 is associated with poor prognosis in colorectal cancer.* J Transl Med, 2013. **11**(1): p. 122.

152. Kim, T.K., et al., *Widespread transcription at neuronal activity-regulated enhancers.* Nature, 2010. **465**(7295): p. 182-7.

153. Ha, M. and V.N. Kim, *Regulation of microRNA biogenesis.* Nat Rev Mol Cell Biol, 2014. **15**(8): p. 509-24.

154. Krol, J., I. Loedige, and W. Filipowicz, *The widespread regulation of microRNA biogenesis, function and decay.* Nat Rev Genet, 2010. **11**(9): p. 597-610.

155. Park, J.E., et al., *Dicer recognizes the 5 ' end of RNA for efficient and accurate processing.* Nature, 2011. **475**(7355): p. 201-U107.

156. Neilsen, C.T., G.J. Goodall, and C.P. Bracken, *IsomiRs--the overlooked repertoire in the dynamic microRNAome.* Trends Genet, 2012. **28**(11): p. 544-9.

157. Starega-Roslan, J., et al., *Sequence features of Drosha and Dicer cleavage sites affect the complexity of isomiRs.* Int J Mol Sci, 2015. **16**(4): p. 8110-27.

158. Vaz, C., et al., *Analysis of the microRNA transcriptome and expression of different isomiRs in human peripheral blood mononuclear cells.* BMC Res Notes, 2013. **6**: p. 390.

159. Zhang, H., et al., *Comparisons of isomiR patterns and classification performance using the rank-based MANOVA and 10-fold cross-validation.* Gene, 2015. **569**(1): p. 21-26.

160. Loher, P., E.R. Londin, and I. Rigoutsos, *IsomiR expression profiles in human lymphoblastoid cell lines exhibit population and gender dependencies.* Oncotarget, 2014. **5**(18): p. 8790-802.

161. Tan, G.C., et al., *5 ′ isomiR variation is of functional and evolutionary importance.* Nucleic Acids Research, 2014. **42**(14): p. 9424-9435.

162. Kozomara, A. and S. Griffiths-Jones, *miRBase: integrating microRNA annotation and deep-sequencing data.* Nucleic Acids Res, 2011. **39**(Database issue): p. D152-7.

163. Manzano, M., et al., *Divergent target recognition by coexpressed 5 ′-isomiRs of miR-142-3p and selective viral mimicry.* Rna, 2015. **21**(9): p. 1606-1620.

164. Llorens, F., et al., *A highly expressed miR-101 isomiR is a functional silencing small RNA.* Bmc Genomics, 2013. **14**.

165. Guo, L., et al., *A Comprehensive Survey of miRNA Repertoire and 3 ′ Addition Events in the Placentas of Patients with Pre-Eclampsia from High-Throughput Sequencing.* Plos One, 2011. **6**(6).

166. Wyman, S.K., et al., *Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity.* Genome Research, 2011. **21**(9): p. 1450-1461.

167. Lee, L.W., et al., *Complexity of the microRNA repertoire revealed by next-generation sequencing.* Rna-a Publication of the Rna Society, 2010. **16**(11): p. 2170-2180.

168. Schamberger, A. and T.I. Orban, *3 ′ IsomiR Species and DNA Contamination Influence Reliable Quantification of MicroRNAs by Stem-Loop Quantitative PCR.* Plos One, 2014. **9**(8).

169. Guo, L., et al., *Consistent isomiR expression patterns and 3′ addition events in miRNA gene clusters and families implicate functional and evolutionary relationships.* Molecular Biology Reports, 2012. **39**(6): p. 6699-6706.

170. Fernandez-Valverde, S.L., R.J. Taft, and J.S. Mattick, *Dynamic isomiR regulation in Drosophila development.* Rna-a Publication of the Rna Society, 2010. **16**(10): p. 1881-1888.

171. Burroughs, A.M., et al., *A comprehensive survey of 3 ′ animal miRNA modification events and a possible role for 3 ′ adenylation in modulating miRNA targeting effectiveness.* Genome Research, 2010. **20**(10): p. 1398-1410.

172. Newman, M.A., V. Mani, and S.M. Hammond, *Deep sequencing of microRNA precursors reveals extensive 3 ′ end modification.* Rna-a Publication of the Rna Society, 2011. **17**(10): p. 1795-1803.

173. Etebari, K., et al., *Dengue virus infection alters post-transcriptional modification of microRNAs in the mosquito vector Aedes aegypti.* Scientific Reports, 2015. **5**.

174. Wang, S.Q., et al., *Entropy-Based Model for MiRNA Isoform Analysis.* Plos One, 2015. **10**(3).

175. Tan, G.C. and N. Dibb, *IsomiRs have functional importance.* Malaysian Journal of Pathology, 2015. **37**(2): p. 73-81.

176. Sablok, G., et al., *isomiRex: Web-based identification of microRNAs, isomiR variations and differential expression using next-generation sequencing datasets.* Febs Letters, 2013. **587**(16): p. 2629-2634.

177. Kuchenbauer, F., et al., *In-depth characterization of the microRNA transcriptome in a leukemia progression model.* Genome Research, 2008. **18**(11): p. 1787-1797.

178. Starega-Roslan, J., et al., *Structural basis of microRNA length variety.* Nucleic Acids Research, 2011. **39**(1): p. 257-268.

179. Zhou, H.J., et al., *Deep annotation of mouse iso-miR and iso-moR variation.* Nucleic Acids Research, 2012. **40**(13): p. 5864-5875.

180. Wu, H.Q., et al., *Alternative Processing of Primary microRNA Transcripts by Drosha Generates 5 ′ End Variation of Mature microRNA.* Plos One, 2009. **4**(10).

181. Fukunaga, R., et al., *Dicer Partner Proteins Tune the Length of Mature miRNAs in Flies and Mammals (vol 151, pg 533, 2012).* Cell, 2012. **151**(4): p. 912-912.

182. Liu, N., et al., *The Exoribonuclease Nibbler Controls 3 ′ End Processing of MicroRNAs in Drosophila.* Current Biology, 2011. **21**(22): p. 1888-1893.

183. Han, B.W., et al., *The 3 ′-to-5 ′ Exoribonuclease Nibbler Shapes the 3 ′ Ends of MicroRNAs Bound to Drosophila Argonaute1.* Current Biology, 2011. **21**(22): p. 1878-1887.

184. Xue, Z.H., et al., *Reconstitution of an Argonaute-Dependent Small RNA Biogenesis Pathway Reveals a Handover Mechanism Involving the RNA Exosome and the Exonuclease QIP.* Molecular Cell, 2012. **46**(3): p. 299-310.

185. Martin, G. and W. Keller, *RNA-specific ribonucleotidyl transferases.* Rna-a Publication of the Rna Society, 2007. **13**(11): p. 1834-1849.

186. Katoh, T., et al., *Selective stabilization of mammalian microRNAs by 3 ′ adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2.* Genes & Development, 2009. **23**(4): p. 433-438.

187. Westholm, J.O., et al., *Common and distinct patterns of terminal modifications to mirtrons and canonical microRNAs.* Rna-a Publication of the Rna Society, 2012. **18**(2): p. 177-192.

188. Berezikov, E., et al., *Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence.* Genome Research, 2011. **21**(2): p. 203-215.

189. Chiang, H.R., et al., *Mammalian microRNAs: experimental evaluation of novel and previously annotated genes.* Genes & Development, 2010. **24**(10): p. 992-1009.

190. Boele, J., et al., *PAPD5-mediated 3 ′ adenylation and subsequent degradation of miR-21 is disrupted in proliferative disease.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(31): p. 11467-11472.

191. Nishikura, K., *Functions and Regulation of RNA Editing by ADAR Deaminases.* Annual Review of Biochemistry, Vol 79, 2010. **79**: p. 321-349.

192. Kawahara, Y., et al., *Frequency and fate of microRNA editing in human brain.* Nucleic Acids Research, 2008. **36**(16): p. 5270-5280.

193. de Hoon, M.J.L., et al., *Cross-mapping and the identification of editing sites in mature microRNAs in high-throughput sequencing libraries.* Genome Research, 2010. **20**(2): p. 257-264.

194. Alon, S., et al., *Systematic identification of edited microRNAs in the human brain.* Genome Research, 2012. **22**(8): p. 1533-1540.

195. Kawahara, Y., et al., *Redirection of silencing targets by adenosine-to-inosine editing of miRNAs.* Science, 2007. **315**(5815): p. 1137-1140.

196. Ekdahl, Y., et al., *A-to-I editing of microRNAs in the mammalian brain increases during development.* Genome Research, 2012. **22**(8): p. 1477-1487.

197. Shoshan, E., et al., *Reduced adenosine-to-inosine miR-455-5p editing promotes melanoma growth and metastasis.* Nature Cell Biology, 2015. **17**(3): p. 311-+.

198. Peng, Z.Y., et al., *Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome.* Nature Biotechnology, 2012. **30**(3): p. 253-+.

199. Brennecke, J., et al., *Principles of MicroRNA-target recognition.* Plos Biology, 2005. **3**(3): p. 404-418.

200. Chi, S.W., et al., *Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps.* Nature, 2009. **460**(7254): p. 479-486.

201. Hafner, M., et al., *Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP.* Cell, 2010. **141**(1): p. 129-141.

202. Betel, D., et al., *Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites.* Genome Biology, 2010. **11**(8).

203. Wang, X.W., *Composition of seed sequence is a major determinant of microRNA targeting patterns.* Bioinformatics, 2014. **30**(10): p. 1377-1383.

204. Humphreys, D.T., et al., *Complexity of Murine Cardiomyocyte miRNA Biogenesis, Sequence Variant Expression and Function.* Plos One, 2012. **7**(2).

205. Helwak, A., et al., *Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding.* Cell, 2013. **153**(3): p. 654-665.

206. Meister, G., et al., *Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs (Reprinted from Molecular Cell, vol 15, pg 185-197, 2004).* Cell, 2007. **131**(4): p. 50-62.

207. Liu, J.D., et al., *Argonaute2 is the catalytic engine of mammalian RNAi.* Science, 2004. **305**(5689): p. 1437-1441.

208. Su, H., et al., *Essential and overlapping functions for mammalian Argonautes in microRNA silencing.* Genes & Development, 2009. **23**(3): p. 304-317.

209. Mi, S.J., et al., *Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5 ′ terminal nucleotide.* Cell, 2008. **133**(1): p. 116-127.

210. Montgomery, T.A., et al., *Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation.* Cell, 2008. **133**(1): p. 128-141.

211. Takeda, A., et al., *The mechanism selecting the guide strand from small RNA duplexes is different among Argonaute proteins.* Plant and Cell Physiology, 2008. **49**(4): p. 493-500.

212. Ha, M. and V.N. Kim, *Regulation of microRNA biogenesis.* Nature Reviews Molecular Cell Biology, 2014. **15**(8): p. 509-524.

213. Czech, B., et al., *An endogenous small interfering RNA pathway in Drosophila.* Nature, 2008. **453**(7196): p. 798-U7.

214. Czech, B., et al., *Hierarchical Rules for Argonaute Loading in Drosophila.* Molecular Cell, 2009. **36**(3): p. 445-456.

215. Ghildiyal, M., et al., *Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway.* Rna-a Publication of the Rna Society, 2010. **16**(1): p. 43-56.

216. Juvvuna, P.K., et al., *Argonaute identity defines the length of mature mammalian microRNAs.* Nucleic Acids Research, 2012. **40**(14): p. 6808-6820.

217. Elkayam, E., et al., *The Structure of Human Argonaute-2 in Complex with miR-20a.* Cell, 2012. **150**(1): p. 100-110.

218. Bail, S., et al., *Differential regulation of microRNA stability.* Rna-a Publication of the Rna Society, 2010. **16**(5): p. 1032-1039.

219. Li, Y., et al., *Genome-Wide Analysis of Human MicroRNA Stability.* Biomed Research International, 2013.

220. Rissland, O.S., S.J. Hong, and D.P. Bartel, *MicroRNA Destabilization Enables Dynamic Regulation of the miR-16 Family in Response to Cell-Cycle Changes.* Molecular Cell, 2011. **43**(6): p. 993-1004.

221. Jones, M.R., et al., *Zcchc11-dependent uridylation of microRNA directs cytokine expression.* Nature Cell Biology, 2009. **11**(9): p. 1157-U258.

222. Goto, Y., et al., *MicroRNA expression signature of castration-resistant prostate cancer: the microRNA-221/222 cluster functions as a tumour*

suppressor and disease progression marker. British Journal of Cancer, 2015. **113**(7): p. 1055-1065.

223. Yamashita, R., et al., *Growth inhibitory effects of miR-221 and miR-222 in non-small cell lung cancer cells.* Cancer Medicine, 2015. **4**(4): p. 551-564.

224. Koelz, M., et al., *Down-regulation of miR-221 and miR-222 correlates with pronounced Kit expression in gastrointestinal stromal tumors.* International Journal of Oncology, 2011. **38**(2): p. 503-511.

225. Gan, R., et al., *Downregulation of miR-221/222 enhances sensitivity of breast cancer cells to tamoxifen through upregulation of TIMP3.* Cancer Gene Therapy, 2014. **21**(7): p. 290-296.

226. Felli, N., et al., *MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(50): p. 18081-18086.

227. Zhao, J.J., et al., *MicroRNA-221/222 Negatively Regulates Estrogen Receptor alpha and Is Associated with Tamoxifen Resistance in Breast Cancer.* Journal of Biological Chemistry, 2008. **283**(45): p. 31079-31086.

228. Visone, R., et al., *MicroRNAs (miR)-221 and miR-222, both overexpressed in human thyroid papillary carcinomas, regulate p27Kip1 protein levels and cell cycle.* Endocr Relat Cancer, 2007. **14**(3): p. 791-8.

229. Wurz, K., et al., *MiR-221 and MiR-222 Alterations in Sporadic Ovarian Carcinoma: Relationship to CDKN1B, CDKN1C and Overall Survival*. Genes Chromosomes & Cancer, 2010. **49**(7): p. 577-584.

Chapter 2

p53 represses the oncogenic sno-miR-28 derived from a snoRNA

Feng Yu[1, 2, 3], Cameron P. Bracken[2, 3*], Katherine A. Pillman[4, 5], David M. Lawrence[4, 5] Gregory J. Goodall[2, 3], David F. Callen[1, 2], Paul M. Neilsen[1, 2, 6]

[1]Centre for Personalized Cancer Medicine, University of Adelaide, SA, Australia

[2]Discipline of Medicine, University of Adelaide, Adelaide, SA, Australia

[3]Centre for Cancer Biology, SA Pathology, Adelaide, SA, Australia

[4]ACRF Cancer Genomics Facility, Centre for Cancer Biology, SA Pathology, Adelaide, Australia

[5]School of Molecular and Biomedical Science, University of Adelaide, Adelaide, Australia

[6]Swinburne University of Technology, Kuching, Sarawak, Malaysia.

*Corresponding author, Email: Cameron.Bracken@health.sa.gov.au

Prelude

Analysis of expression data from p53 induction in cancer cell lines revealed that six polycistronic C/D box snoRNAs were repressed by wild type p53. This observation led to further exploration which identified and characterized a novel miRNA that was processed from a snoRNA. The cellular activities of sno-miR-28 was then investigated.

This research has been published in Plos One (2015) and this chapter is the published version of the manuscript.

Contribution of the candidate: Figure 1A, B, C, Figure 2A, C, D, Figure 3, 4 and 5 all panels, Supplementary Figure 1 and 2, Supplementary Table 3 and 4, and writing the manuscript.

Statement of Authorship

Title of Paper	p53 represses the oncogenic sno-miR-28 derived from a snoRNA
Publication Status	Published
Publication Date	10 June 2015

Candidate

Name of Author	Feng Yu		
Contribution to the Paper	Performed analysis on cell and tissue experiments, interpreted data, wrote the manuscript.		
Signature		Date	23/2/2016

By signing the Statement of Authorship, each author certifies that the candidate's stated contribution to the publication is accurate, and permission is granted for the candidate in include the publication in the thesis.

Name of Author	Cameron Bracken		
Contribution to the Paper	Supervised development of work, helped in data interpretation and manuscript preparation.		
Signature		Date	1/3/16

Name of Author	Katherine Pillman		
Contribution to the Paper	Performed bioinformatical analysis of data, helped in data interpretation.		
Signature		Date	23/2/2016

Name of Author	David Lawrence		
Contribution to the Paper	Performed bioinformatical analysis of data, helped in data interpretation.		
Signature		Date	26/2/16

Name of Author	Gregory Goodall		
Contribution to the Paper	Supervised development of work, helped in data interpretation and manuscript evaluation.		
Signature		Date	1/3/16

Name of Author	David Callen		
Contribution to the Paper	Supervised development of work, helped in data interpretation and manuscript preparation.		
Signature		Date	24/2/16

Name of Author	Paul Neilsen		
Contribution to the Paper	Supervised development of work, helped in data interpretation and manuscript preparation.		
Signature		Date	

Name of Author	David Lawrence		
Contribution to the Paper	Performed bioinformatical analysis of data, helped in data interpretation.		
Signature		Date	

Name of Author	Gregory Goodall		
Contribution to the Paper	Supervised development of work, helped in data interpretation and manuscript evaluation.		
Signature		Date	

Name of Author	David Callen		
Contribution to the Paper	Supervised development of work, helped in data interpretation and manuscript preparation.		
Signature		Date	

Name of Author	Paul Neilsen		
Contribution to the Paper	Supervised development of work, helped in data interpretation and manuscript preparation.		
Signature		Date	25/02/2016

Abstract

p53 is a master tumour repressor that participates in vast regulatory networks, including feedback loops involving microRNAs (miRNAs) that regulate p53 and that themselves are direct p53 transcriptional targets. We show here that a group of polycistronic miRNA-like non-coding RNAs derived from small nucleolar RNAs (sno-miRNAs) are transcriptionally repressed by p53 through their host gene, *SNHG1*. The most abundant of these, sno-miR-28, directly targets the p53-stabilizing gene, TAF9B. Collectively, p53, SNHG1, sno-miR-28 and TAF9B form a regulatory loop which affects p53 stability and downstream p53-regulated pathways. In addition, SNHG1, SNORD28 and sno-miR-28 are all significantly upregulated in breast tumours and the overexpression of sno-miR-28 promotes breast epithelial cell proliferation. This research has broadened our knowledge of the crosstalk between small non-coding RNA pathways and roles of sno-miRNAs in p53 regulation.

Introduction

The p53 tumour suppressor plays a pivotal role in the prevention of oncogenic transformation as highlighted by the fact that over half of all tumours have mutations in *TP53*. Cellular insults such as DNA damage or aberrant oncogene expression engage the p53 pathway, resulting in rapid stabilization of p53 protein levels. Upon activation, p53 functions as a sequence-specific transcription factor to either activate or repress the expression of target genes. The global landscape of p53 transcriptional regulation is vast and complex, with gene expression profiling studies demonstrating that thousands of genes rapidly alter expression upon p53 activation [1-3]. In addition to direct regulation, p53 also imparts a substantial amount of transcriptional regulation through indirect mechanisms. For example, Nikulenkov *et al.* recently applied a combination of ChIP-Seq (chromatin immunoprecipitation combined DNA sequencing) and RNA-Seq (RNA Sequencing) to demonstrate that, although activation of p53 resulted in the altered transcription levels of over 4500 genes, less than 10% of these genes were directly bound by p53 [2]. Hence, a major component of the p53 transcriptional network is mediated through various indirect effectors or non-protein coding regulators.

miRNAs constitute one of the largest families of trans-acting gene regulators. The discovery that p53 can regulate miRNAs, coupled with the observation that many effects of p53 are indirect, suggests that they could be significant effectors in the p53 transcriptome [4]. Recently, p53 has been shown to transcriptionally activate or repress the expression of several miRNAs, including miR-17-92 cluster [5], miR-22 [6], the miR-34 family [7-11], miR-145 [12], miR-192 family [13,14], miR-149 [15], miR-200 family [16,17], miR-605 [18], miR-1204 [19], miR-509 [20], and miR-1915 [21]. Hence, by regulating a miRNA-based network, p53 could modulate an extensive downstream transcriptome. Other families of non-coding RNAs are also emerging as novel entities in the downstream p53 pathway, such as long non-coding RNAs (lncRNAs) [22] and various other pol I/III transcripts including tRNA and rRNAs [23,24]. The extent to which these and other non-coding RNAs participate in the p53 pathway are currently not well understood.

Herein, we have examined the ability of p53 to regulate miRNA-sized transcripts processed from non-coding small nucleolar RNAs (snoRNAs). Our findings demonstrate a role for p53 in the repression of a family of polycistronic C/D box snoRNAs (SNORDs), of which at least one is processed into an operative miRNA which feeds back to repress TAF9B-mediated stabilization of p53 and promote cell proliferation.

Results

Identification of p53 regulated snoRNAs

Given the existing links between p53 and various non-coding RNAs, we hypothesized that p53 may also regulate snoRNAs, as snoRNAs have been linked to carcinogenesis [25,26]. To examine this, we performed snoRNA expression profiling following activation of the p53 signalling network. This was conducted by Affymetrix gene expression profiling in two models of p53 activation to increase the chance of identifying *bona fide* snoRNA targets of p53 that were not restricted to specific cell type, cell origin or mode of p53 activation. These two cell-based models of wild-type p53 activation were (i) an inducible p53 system in the p53 null H1299 cell line, which was previously characterized at various doses (0-2.5 μg/ml) and time points (0-96 hours) for gene expression analysis [3], and (ii) activation of endogenous wild-type p53 through Nutlin-3a treatment of the WE-68 cell line (Fig. 1A, Supplementary Table S4), which was previously characterized at various doses and time points for cell cycle and gene expression analysis [69]. Wild type p53-inducible H1299 cells were treated with 2.5μg/ml of Ponasterone A (PonA) for 24 hours, and compared with non-induced cells. WE-68 cells were treated with 10 nM Nutlin-3a for 16 hours, compared with non-induced cells. Due to space limits, only a

few representative RNA transcripts are presented in Fig. 1B as examples from the microarray to indicate the anticipated responses of these cells to p53 activation, including mRNAs, miRNAs and snoRNAs. According to previous studies, FAS, PUMA (BBC3), CDKN1A, MDM2, RRM2B (p53R2), BAX, CCNG1, TLR3 and MIR34A (miR-34a host gene) are upregulated upon p53 activation [27]; ACTB (β-actin), GAPDH, PSMB4 and C1orf43 are usually stably expressed [28]; whereas E2F1 [27], CCNE2 [29], POLD1 [30], CDCA8 [31], FBXO5 [32], PLK4 [33], BRCA1 [34], CCNB1 [35] and MIRH1 (miR-17-92 host gene) [36] are directly or indirectly repressed upon p53 activation (Fig.1B). Clustering analysis was employed to indicate the relative closeness of the reactive patterns how these genes respond to p53 activation in both cell lines.

In a combined analysis of both the H1299 and WE-68 cell lines using Affymetrix microarray, we identified six snoRNAs that were most significantly (fold change > 1.085 or <0.915, $p<0.05$) regulated by p53 between induced (H1299 treated with PonA and WE-68 treated with Nutlin) and uninduced cells (Fig. 1B). Interestingly, all six snoRNAs were clustered to MIRH1, a polycistronic miRNA host gene of the miR-17-92 cluster that is transcriptionally repressed by p53 [36], and this indicates probable similarity in their reactive patterns upon p53 activation (Fig. 1B). There were no snoRNAs significantly activated by p53 in either system, suggesting that the role of p53 in the regulation of snoRNAs may be restricted to that of transcriptional repression. Five of these p53-repressed snoRNAs were encoded within the same polycistronic gene,

SNHG1 (SNORD host gene 1), a precursor of a family of C/D box snoRNAs (Table 1, Fig. 2A, Supplementary Table S1 and S2). All eight SNHG1-encoded snoRNAs were repressed by p53, with five reaching statistical significance (fold change > 1.085 or <0.915, $p<0.05$) (Table 1). Consistent with this, upon activation of p53 in both H1299 and WE-68 cell lines, expression levels of SNHG1, the SNORD host gene, was significantly reduced (Fig. 1C). We next investigated if SNHG1 was transcriptionally repressed by p53. Analysis of the promoter region of *SNHG1* (~5kB upstream of initiation site of *SNHG1*) using p53 Scan [37] identified a putative p53 responsive element (RE). ChIP analysis using the H1299 p53 inducible system indicated modest but reproducible recruitment of p53 to this response element (Fig. 1D). Collectively, these data implicate *SNHG1*, and its processed snoRNAs, as p53-repressed targets.

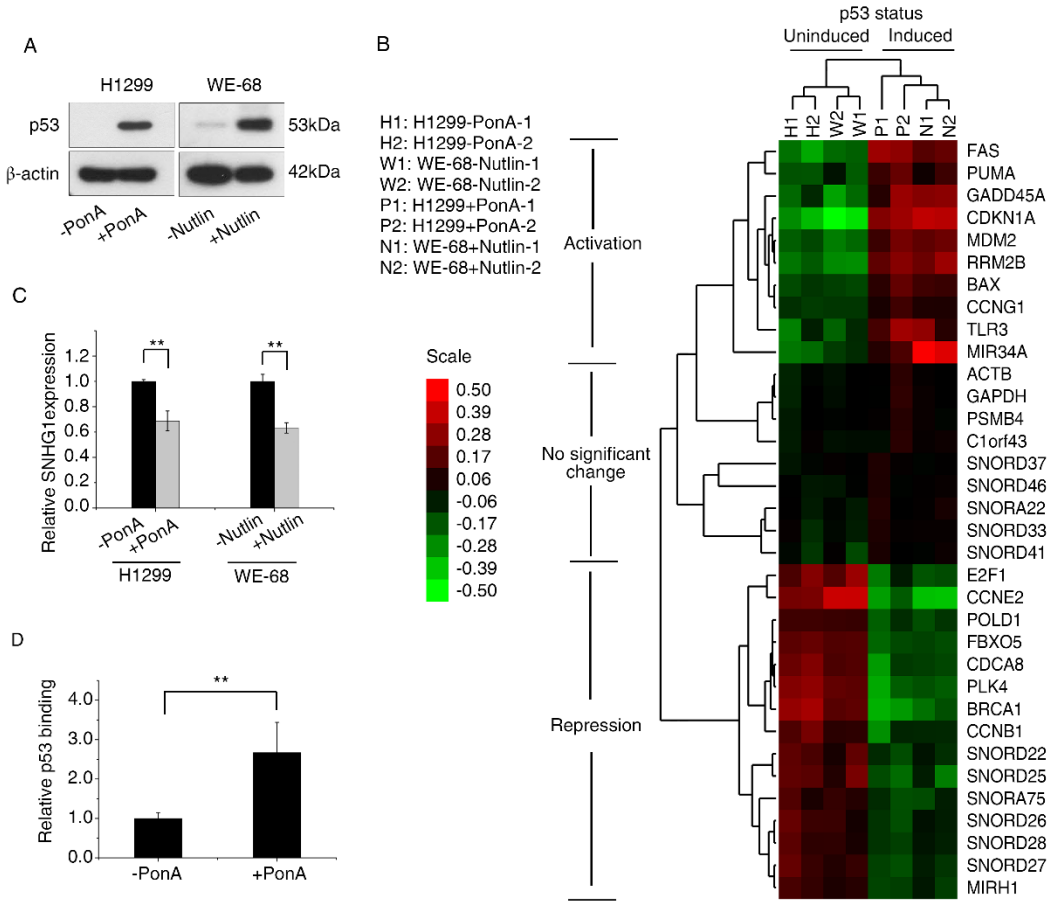


Fig. 1. Identification of p53 regulated snoRNAs.

(A) Wild type p53-inducible H1299 cells were treated with 2.5μg/ml of Ponasterone A (PonA) for 24 hours, and compared with non-induced cells. WE-68 cells were treated with 10nM Nutlin-3a for 16 hours, compared with non-induced cells. Western blots for p53 (and β-actin as a loading control) are shown. **(B)** Based on microarray profiling, a dendrogram generated by cluster analysis shows the separation of p53 uninduced cells from induced cells, and separation of representative genes activated, repressed or not significantly changed by p53. **(C)** SNHG1 expression levels were determined by RT-PCR

after p53 induction in H1299 and WE-68 cells. **(D)** ChIP assay, using the anti-p53 antibody DO-1, was performed to determine relative p53 occupancy in p53-induced H1299 cells (+PonA). RT-PCR results show relative p53 occupation at upstream of the SNHG1 promoter, and p53 null H1299 (-PonA) was used as a negative control. ** $p<0.01$ versus controls for all experiments.

Gene Name	Accession Number	p value induced/ uninduced	Chromosome Location	Precursor Transcript
SNORD22	NR_000008	0.021988	11q13	SNHG1
SNORD25	NR_002565	0.019725		
SNORD26	NR_002564	0.036236		
SNORD27	NR_002563	0.018451		
SNORD28	NR_002562	0.02692		
SNORA75	NR_002921	0.010118	2q37.1	NCL

Table 1. snoRNAs repressed by p53 in H1299 and WE-68 cells

Affymetrix gene expression profiling identified six snoRNAs that were repressed in common in H1299 and WE-68 cell lines as wild-type p53 was induced. Statistical significance of these snoRNAs is characterized by p-values that compare p53 induced (H1299 treated with PonA and WE-68 treated with Nutlin) versus uninduced cells, with their host genes listed.

Gene Name	Accession Number	Fold Repression	Location	Precursor Transcript
SNORD22	NR_000008	2.19	11q13	**SNHG1**
SNORD25	NR_002565	2.69		
SNORD26	NR_002564	3.12		
SNORD27	NR_002563	2.64		
SNORD28	NR_002562	2.74		
SNORD30	NR_002561	2.11		
SNORD31	NR_002560	1.91		
SNORD74 **SNORD75** **SNORD78**	NR_002579 NR_003941 NR_003944	1.84 1.81 1.53	1q25.1	**SNHG2**
SNORD50A **SNORD50B**	NR_002743 NR_003044	1.50 2.09	6q14.3	**SNHG5**
SNORD42B **SNORD4** **SNORD4B**	NR_000013 NR_000010 NR_000009	1.87 1.63 2.16	17q11	**RPL23A**
SNORA75	NR_002921	2.07	2q37.1	**NCL**
SNORA61	NR_002987	1.66	1p35.5	**SNHG12**
SNORD34	NR_000019	1.77	19q13.3	**RPL13A**
SNORA45	NR_002977	2.07	11p15.4	**RPL27A**
SNORA64	NR_002326	1.55	16p13.3	**RPS2**
SNORD73A	NR_000007	1.74	4q31.2	**RPS3A**
SNORA9	NR_002952	1.68	7p13	**C7orf40**
SNORA56	NR_002984	1.99	Xq28	**DKC1**
SNORD96A	NR_002592	1.61	5q35.3	**GNB2L1**
SNORA71B	NR_002910	1.57	20q11.23	**LOC388796**
SNORA55	NR_002983	1.67	1p34.3	**PABPC4**
SNORD45C	NR_003042	1.55	1p31.1	**RABGGTB**
SNORA40	NR_002973	1.62	11q21	**-**
SNORD3B-2	NR_003924	2.68	17p11.2	**-**
SNORD5	**NR_003033**	**1.70**	**11q21**	**-**

S1 Table. snoRNAs repressed by p53 in H1299 cells

As wild-type p53 was induced in H1299 cells, Affymetrix gene expression profiling identified a list of snoRNAs that were repressed. These snoRNAs are listed with their host genes.

Gene Name	Accession Number	Fold Repression	Location	Precursor Transcript
SNORD22	NR_000008	1.53	11q13	**SNHG1**
SNORD25	NR_002565	2.09		
SNORD26	NR_002564	1.50		
SNORD27	NR_002563	1.59		
SNORD28	NR_002562	1.51		
SNORA75	**NR_002921**	**1.56**	**2q37.1**	**NCL**

S2 Table. snoRNAs repressed by p53 in WE-68 cells

When wild-type p53 was induced in WE-68 cells, Affymetrix gene expression profiling identified a list of snoRNAs that were repressed. These snoRNAs are shown with their host genes.

p53 repressed snoRNAs are processed into miRNAs

In addition to their well characterized function guiding the enzymatic modification of ribosomal RNA, snoRNAs may also be processed into smaller miRNA-sized molecules which are capable of binding Argonaute (AGO) and exerting miRNA-like effects. These are variously termed sno-miRNAs or snoRNA-derived small RNAs (sdRNAs) [38-40]. Such processing is both widespread and evolutionarily conserved [38], leading us to ask whether the products of SNHG1 may also serve miRNA-like roles.

In order to investigate this, we analysed data from HITS-CLIP (high throughput sequencing of crosslinked and immunoprecipitated RNA) performed in MDA-MB-231 breast cancer cells [41]. Briefly, this technique works by the immunoprecipitation of AGO, a key component of the miRNA processing machinery, with which bound small RNAs are also co-immunoprecipitated. Associated small RNAs are then identified by deep sequencing. We found evidence of miRNA-sized (~17-26nt) molecules derived from all characterized snoRNAs within SNHG1 (Fig. 2A), with a specific small RNA mapping to the 3′ end of SNORD28 being the most abundant followed by a small RNA originating from the 5′ end of SNORD25. The SNORD28- and SNORD25- derived small RNAs were sequenced within the library 630 and 365 times respectively, placing them at the level of other moderately-expressed well characterized miRNAs such as miR-155 and miR-34a (Fig. 2A). Interestingly, although a small RNA derived from SNORD28 was the most abundant small RNA recruited to AGO, SNORD29 was by far the most abundantly expressed SNHG1-derived snoRNA, demonstrating high processing and/or Ago-binding selectivity (Fig 2A). In addition, the precise 5′ termini of the SNORD-28 derived small RNA further indicates the strong specificity of RNA processing and AGO binding (Fig. 2B).

To determine if sno-miR-25 and 28 are really expressed in abundance *in vivo*, we measured the endogenous expression levels of sno-miR-25 and sno-miR-28 in normal and malignant breast tissues, and compared to miR-155 which has a moderate expression level in breast tissues according to miRBase [42]. TaqMan assay showed sno-miR-28 has a higher *in vivo* expression level than miR-155, whereas the expression of sno-miR-25 is extremely low (Supplementary Fig. S2 A). For this reason, we focused our research on SNORD28 and sno-miR-28. We confirmed the PCR efficiencies and specificities of the TaqMan assays for these RNAs, demonstrating the sno-miR-28 TaqMan assay is 16 times more specific to sno-miR-28 than to SNORD28 (Supplementary Fig. S1).

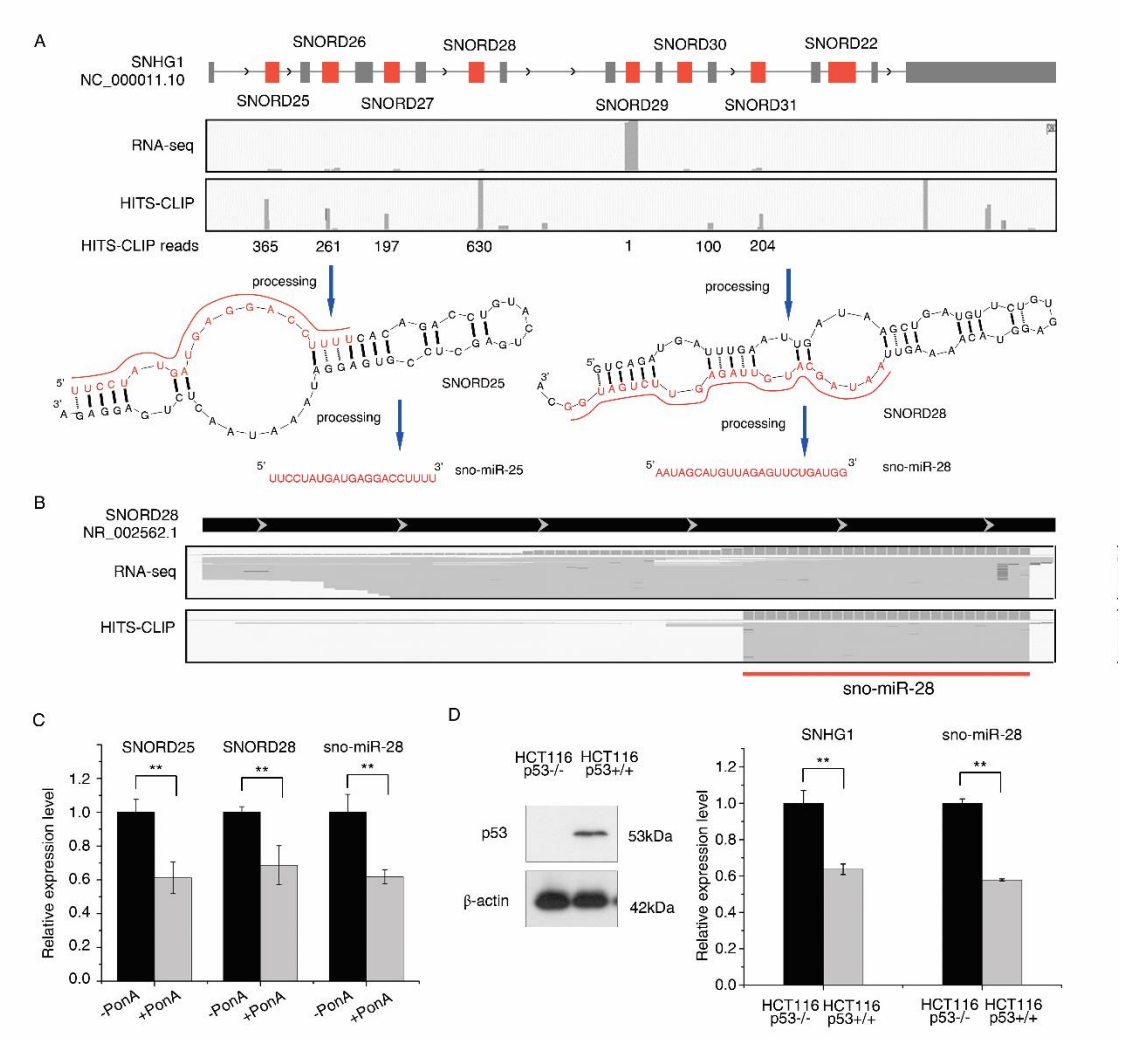
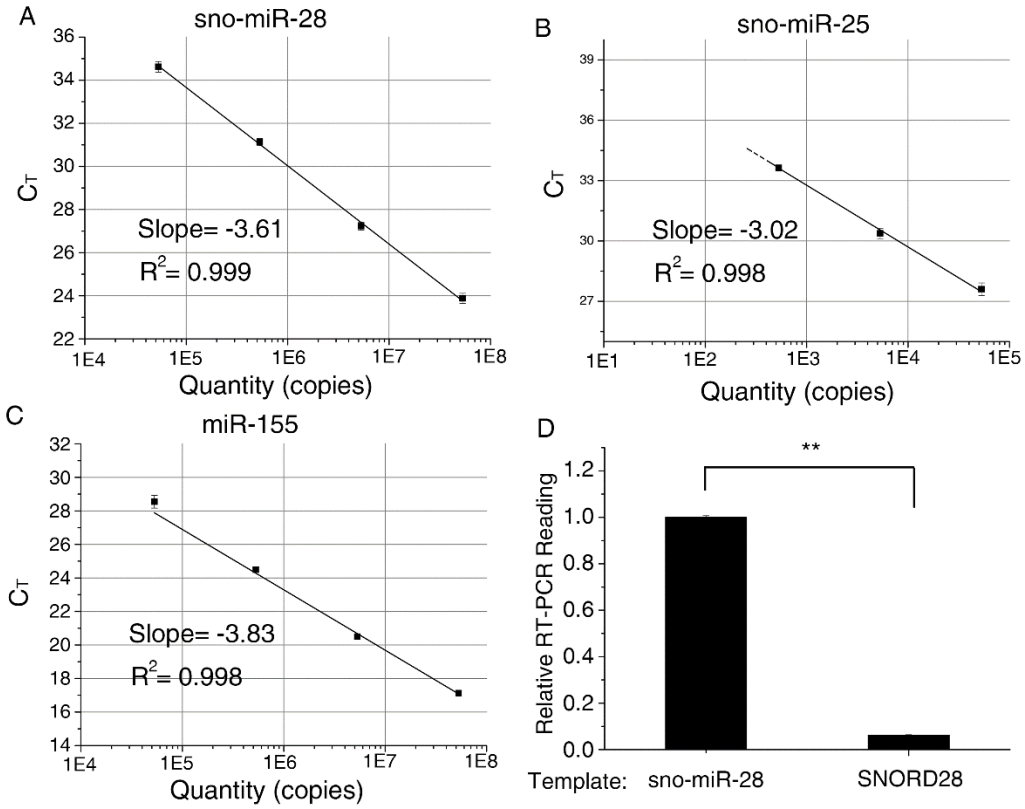


Fig. 2. p53 repressed snoRNAs are processed into miRNAs.

(A) SNHG1 is processed into snoRNAs including SNORD25 and SNORD28. The top panel shows RNA deep-sequencing and HITS-CLIP (high throughput sequencing of crosslinked and immunoprecipitated RNA) results at the SNHG1 genomic loci. Predicted stem-loop folding of SNORD25 and SNORD28 are shown. The regions marked in bold are processed into sno-miRNAs which can bind to Argonaute proteins, which was confirmed by RNA deep-

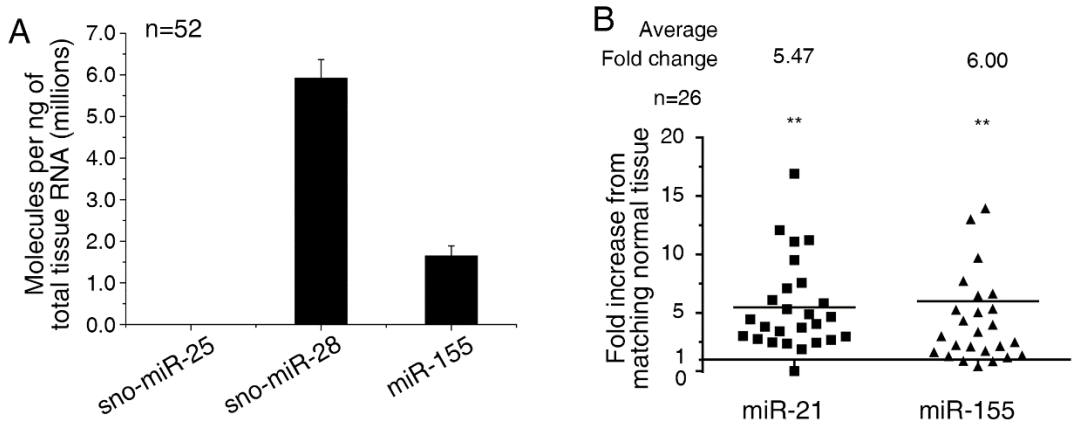
sequencing and HITS-CLIP results. The solid lines between chains represent hydrogen bonds between adenine (A)-uracil (U) pairs and guanine (G)-cytosine (C) pairs, whereas dashed lines represent G-U pairing. **(B)** RNA-seq and HITS-CLIP mapping reads across the SNORD28 region is shown indicating precise binding of sno-miR-28 to AGO **(C)** p53 was induced by PonA treatment in inducible H1299 cells, and the expression levels of SNORD25, SNORD28 and sno-miR-28 were determined using TaqMan assay and RT-PCR. Expression levels of SNORD25, SNORD28 and sno-miR-28 were shown in induced or uninduced cells. **(D)** Isogenic HCT116 -/-p53 and HCT116 +/+p53 cell lines were used to investigate the relation of SNHG1 and sno-miR-28 expression levels with p53. Left: p53 protein expression in the HCT116 isogenic cell lines was shown by Western blot and β-actin was used as a loading control. Right: SNHG1 and sno-miR-28 expression was determined by RT-PCR as shown. ** $p<0.01$ versus controls for all experiments.



S1 Fig. Characterization of snoRNA and sno-miRNA RT-PCR

(A, B, C) Standard curves were shown for absolute quantitation of endogenous sno-miR-28, sno-miR-25 using custom TaqMan assays with miR-155 included as a positive control. Synthesized sno-miR-28 and miR-155 were prepared at a series of known concentrations at 4nM, 0.4nM, 0.04nM, and 0.004 nM, while synthesized sno-miR-25 was prepared at 4pM, 0.4pM, 0.04pM, and 0.004 pM to match its low endogenous expression detected in tissues. RT-PCR was performed subsequently and CT values were plotted to log10 of the miRNA concentrations. The point representing 0.004pM of sno-miR-25 was not plotted

because no CT value was detectable at this concentration. **(D)** sno-miR-28 TaqMan RT-PCR was performed using synthesized SNORD28 or sno-miR-28 as templates. RT-PCR readings are shown to evaluate the specificity of the sno-miR-28 TaqMan assay.



S2 Fig. Evaluation of endogenous expression of miRNAs.

(A) The standard curves in Supplementary Fig. S1 A, B, C were used to define a standard curve for the absolute expression of miR-155, sno-miR-28 and sno-miR-25 in 26 breast tumour tissues and paired normal adjacent breast tissues using TaqMan assay and RT-PCR. The average expression levels of sno-miR-28, miR-155, and sno-miR-25 are shown in millions of molecules per nanogram (ng) of total RNA. **(B)** miR-21 and miR-155 expression was determined as

positive controls using Taqman assay and RT-PCR in breast tumours compared with paired normal tissue.

Since SNORD25, SNORD28 and sno-miR-28 are all processed from SNHG1, we hypothesized their expression may be affected through SNHG1 upon p53 activation. Indeed, activation of p53 in H1299 cells resulted in significant downregulation of the expression levels of SNORD25, SNORD28 and sno-miR-28 (Fig. 2C). We also demonstrated that this regulatory axis is not restricted to any specific p53 activation models using the HCT116 isogenic cell line system. Indeed, HCT116 (*TP53*$^{+/+}$) cells express significantly lower levels of SNHG1 and sno-miR-28 than the p53 null HCT116 (*TP53*$^{-/-}$). Taken together, these results confirm that the SNHG1-sno-miR-28 axis is negatively regulated by p53.

sno-miR-28 functions as a miRNA

Since previous studies have demonstrated miRNA-like functions for sno-miRNAs [38-40,43-45], we employed a bioinformatics approach to explore potential sno-miR-28 targets. As predicted by TargetScan Custom 5.1, TAF9B (transcription initiation factor TFIID subunit 9B), BHLHE41 (class E basic helix-loop-helix protein 41) and TGFBR2 (transforming growth factor beta receptor

II) were identified among the putative targets of sno-miR-28 (Fig. 3A, Supplementary Fig. S3). Several (~10) candidate mRNAs were investigated upon overexpression or inhibition of sno-miR-28 (data not shown), and TAF9B was associated with the greatest level of repression in response to exogenous sno-miR-28. In addition, RNA folding analysis predicted that TAF9B has a moderate-to-high level of hybridization energy binding to sno-miR-28 (ΔG= -21.0 kcal/mol) (Fig. 3A). [46]

TAF9B was deduced to be a target of sno-miR-28 not only by bioinformatics analysis, but also by the inverse correlation between sno-miR-28 and TAF9B expression. Following overexpression of sno-miR-28, endogenous TAF9B mRNA and protein expression levels were significantly reduced in H1299 cells (Fig. 3B, C). To confirm that sno-miR-28 directly interacts with TAF9B's 3'-UTR, we overexpressed sno-miR-28 along with a psiCHECK2 luciferase reporter containing the TAF9B 3'UTR fused to the 3' end of the Renilla luciferase gene. We observed that the Renilla luciferase expression was inhibited by sno-miR-28 overexpression, and the repression was abolished by mutation of the proposed sno-miR-28 recognition site (Fig. 3D). Taken together, these results demonstrate that sno-miR-28 directly mediates repression of TAF9B through a canonical miRNA binding site.

In order to help verify our observations are not restricted to a specific cell type, the relation between sno-miR-28 and TAF9B was then investigated in MCF10A cells, an immortalized, non-transformed breast epithelial cell line. Transfection of sno-miR-28 mimics downregulates TAF9B mRNA and protein, consistent with sno-miR-28 also functioning like a canonical miRNA in breast epithelial cells (Fig. 3E, F). Furthermore, a Locked Nucleic Acid (anti-sno-miR-28 LNA) was used to inhibit endogenous sno-miR-28 expression in MCF10A cells and, consistently, TAF9B mRNA and protein expression was increased (Fig. 3E, F). Taken together, this indicates TAF9B is subject to regulation by the endogenous sno-miR-28.

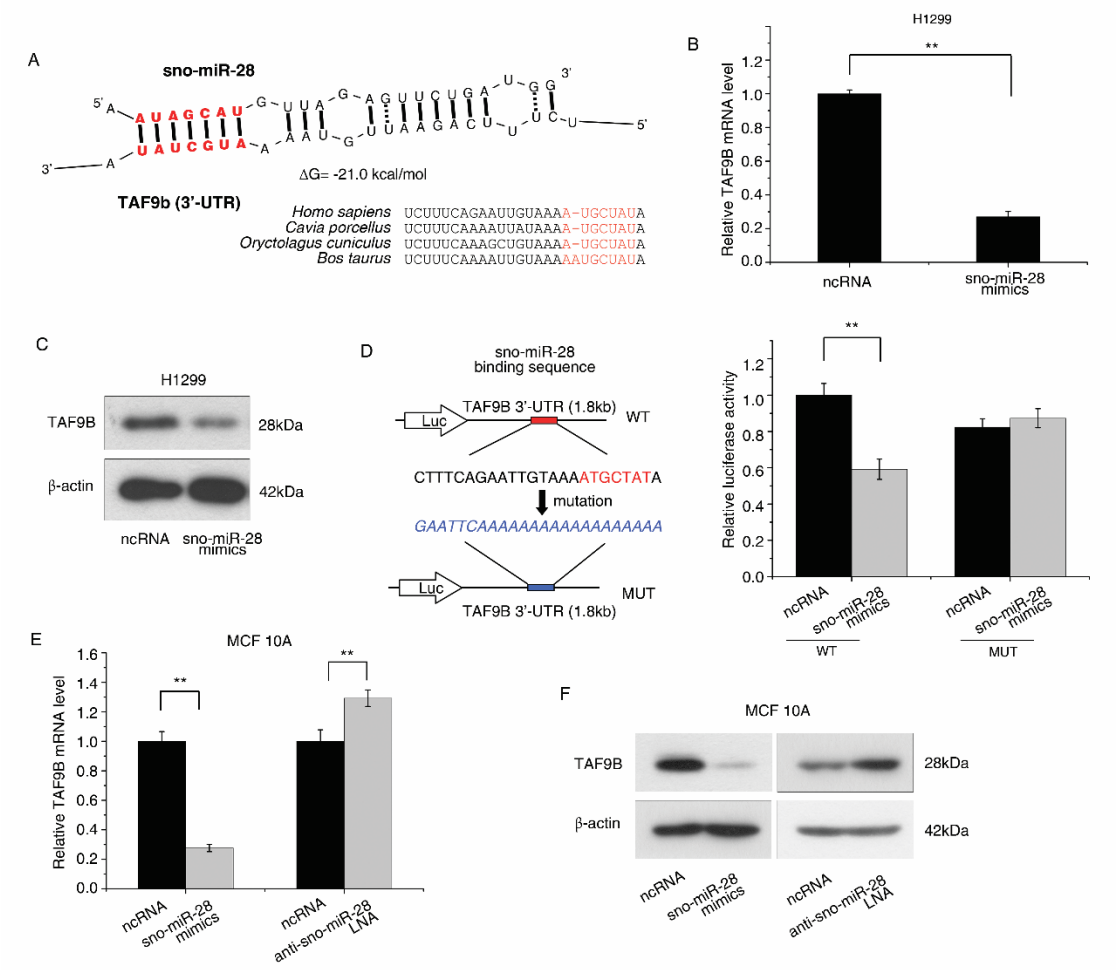
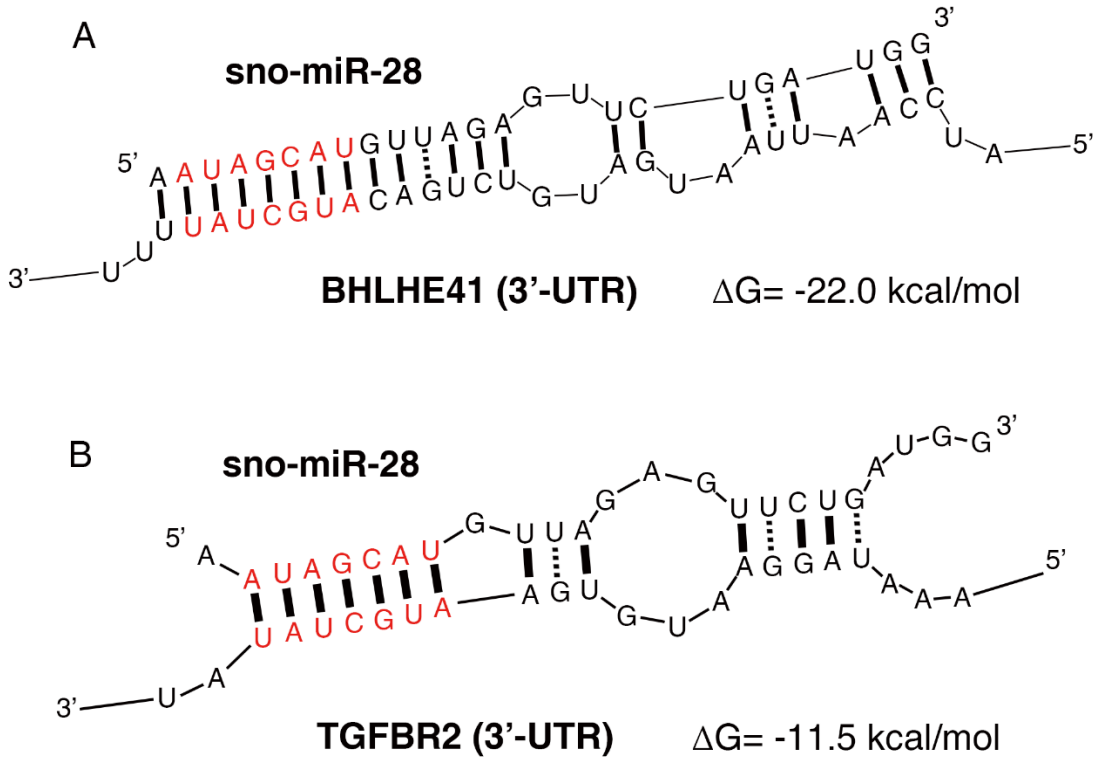


Fig. 3. sno-miR-28 functions as a miRNA.

(A) Proposed sno-miR-28 binding site within the TAF9B 3'UTR. The seed-recognition site is marked in bold; hypothesized duplexes formed by the interaction of TAF9B and sno-miR-28 are illustrated, and the predicted free energy of the hybrid is indicated. Conservation of the seed region across 4 species is also indicated. **(B, C)** sno-miR-28 (or negative control RNA, ncRNA) was overexpressed in H1299 cells. TAF9B mRNA and protein levels were determined by RT-PCR and Western blot, respectively. **(D)** Using a dual-

luciferase reporter system, H1299 cells were co-transfected with sno-miR-28 mimics (or negative control RNA), and psiCHECK2 luciferase reporter plasmids with either wild type (WT) or mutated TAF9B 3'-UTR (MUT) cloned at downstream of the Renilla luciferase gene (Luc). Relative luciferase activities are shown. **(E, F)** sno-miR-28 was either overexpressed (mimics) or inhibited (LNA) in MCF10A cells. TAF9B mRNA levels were determined by RT-PCR (E), and protein expression was determined by Western blot (F). ** $p<0.01$ versus controls for all experiments, and β-actin was included as a loading control for all Western blots.



S3 Fig. Hybridization of sno-miR-28 binding to BHLHE41 (A) and TGFBR2 (B).

Bioinformatics analysis showed the binding of sno-miR-28 to BHLHE41 and TGFBR2 3′-UTR in a similar manner as other miRNAs binding to their targets. The schematic graphs are made to show the proposed binding sites for sno-miR-28. The seed-recognizing sites are marked in red; hypothesized duplexes formed by the interaction of the binding sites of the 3′-UTR of TAF9B and sno-miR-28 are illustrated, and the predicted free energy of the hybrids were indicated. The solid lines between two chains represent hydrogen bonds between adenine (A)-uracil (U) pairs and guanine (G)-cytosine (C) pairs, whereas dashed lines represent G-U pairings.

sno-miR-28 alters p53 protein stability through TAF9B

TAF9B functions as a subunit of TFIID (transcription initiation factor II D) and TFTC (TATA-binding Protein-free TAF-containing) complexes. It also acts as a p53 co-activator, stabilizing p53 possibly by competing for Mdm2 binding [47,48]. We therefore reasoned that via its direct regulation of TAF9B, sno-miR-28 may indirectly regulate p53. To investigate this, we examined p53 protein

levels after sno-miR-28 overexpression in H1299 cells and found that sno-miR-28 downregulated p53 protein but not RNA (fig. 4A, B), suggesting the sno-miR-28 and TAF9B regulation of p53 may function at the protein level. In addition, sno-miR-28 overexpression also significantly repressed multiple p53 regulated genes including CDKN1A (p21), RRM2B, CCNG1, FAS and HDM2 in induced H1299 cells (Fig. 4A, C).

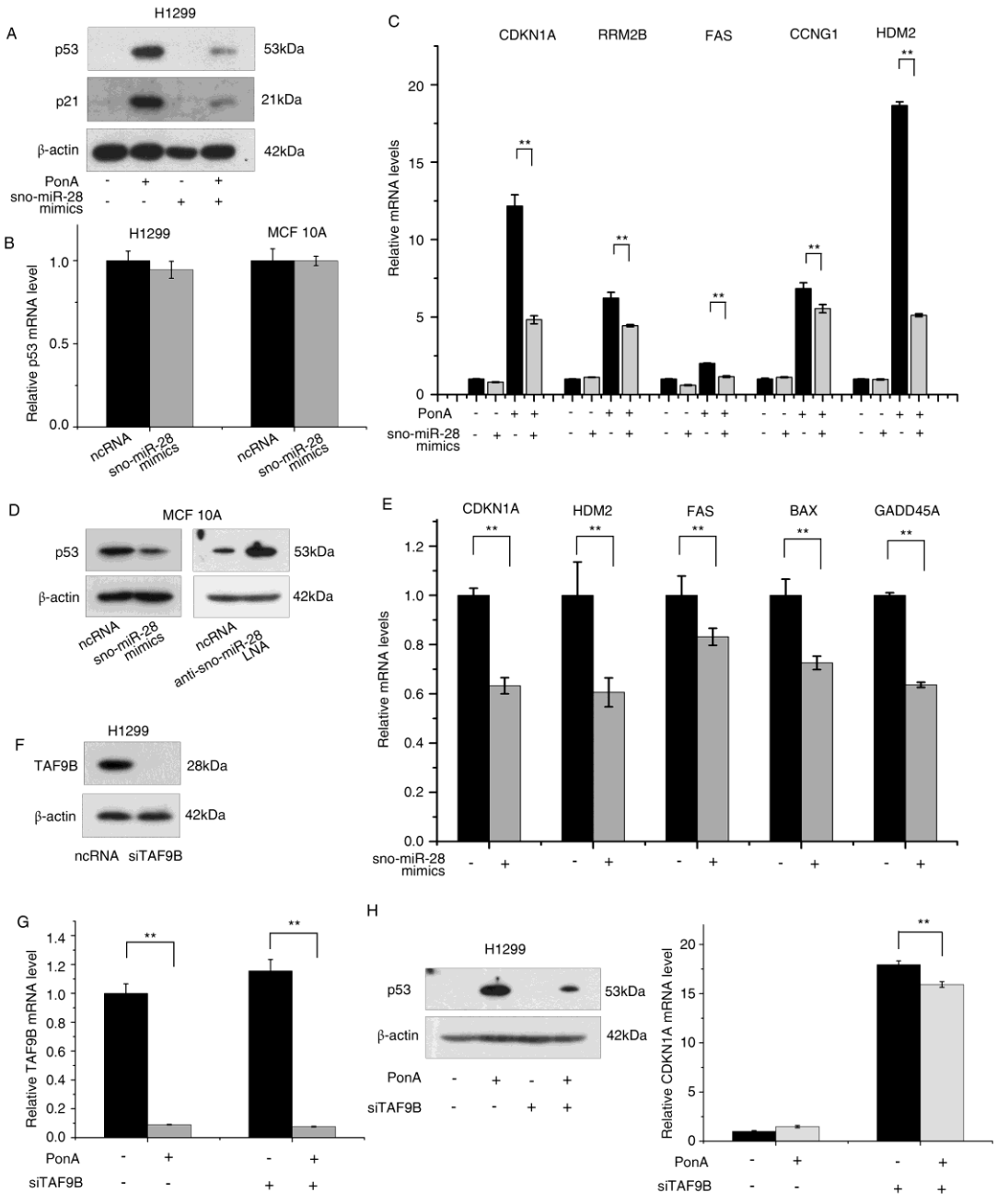


Fig. 4. sno-miR-28 alters p53 protein stability through TAF9B and forms a feed-forward loop between p53, sno-miR-28, and TAF9B.

(A) Western blots are shown to determine the protein levels of p53 and p21 after overexpression of sno-miR-28 and/or induction of p53 by Ponasterone A (PonA) in inducible H1299 cell line, compared with cells transfected with

negative control RNA (ncRNA) and/or uninduced H1299 cells. (B) p53 mRNA levels in MCF10A and PonA-treated H1299 cells after overexpression of sno-miR-28 are shown by RT-PCR, compared with scrambled negative control (ncRNA). (C) The mRNA expression levels of p53 downstream targets are shown after overexpression of sno-miR-28 (or negative control) in H1299 cells upon p53 activation. (D) sno-miR-28 (or negative control) was either expressed (mimic) or inhibited (LNA) in MCF10A cells and the protein expression of p53 was determined by Western blot. (E) The mRNA expression levels of p53 downstream targets are shown after overexpression of sno-miR-28 in MCF10A cell line, compared with ncRNA. (F, G) TAF9B was knocked using an siRNA in H1299 cells compared with a negative control. Successful knockdown is shown at both the protein (F) and mRNA (G) levels. (H) CDKN1A expression was determined by RT-PCR after TAF9B was knockdown. ** $p<0.01$ versus controls for all experiments. β-actin is included as a loading control for all Western blots. Minus marks (-) in sno-miR-28 mimic and anti-sno-miR-28 LNA experiments represent negative control transfections using a scrambled ncRNA.

To confirm that the sno-miR-28-TAF9B-p53 regulatory axis is not restricted to a specific cell type or model of p53 activation, we also investigated this pathway in MCF10A cells. As seen in H1299 cells, sno-miR-28 overexpression reduced

p53 protein but not mRNA in MCF10A cells, while inhibition of sno-miR-28 restored p53 protein levels (Fig. 4B, D). In addition, overexpression of sno-miR-28 repressed the mRNA levels of various p53 downstream regulators including CDKN1A, HDM2, FAS, BAX and GADD45A (Fig. 4E) highlighting the prevalent influence of sno-miR-28 in p53 signalling.

Consistent with sno-miR-28 stabilizing p53 protein via its regulation of TAF9B, we found that siRNA-mediated knockdown of TAF9B phenocopied the effect of sno-miR-28, reducing the levels of p53 protein and CDK1A mRNA (Fig. 4F-H).

sno-miR-28 is upregulated in breast tumours and promotes breast cell proliferation

Our findings indicate that sno-miR-28 participates in a feed forward loop with p53, whereby p53 represses sno-miR-28 via SNHG1, whilst sno-miR-28 directly targets the TAF9B 3'UTR to negatively regulate p53 stability (Fig. 5A). The reciprocal negative association with p53 implies sno-miR-28 might have an oncogenic role. Consistent with this, SNHG1 is upregulated in gastric cancer [49]. To further investigate, in a cohort of 26 pairs of matched malignant and

non-malignant breast tissues samples, we found that SNHG1, SNORD28 and sno-miR-28 were all significantly upregulated in breast tumours (Fig. 5B). Two well-characterized oncogenic miRNAs, miR-21 and miR-155, that were used as positive controls were similarly upregulated (Supplementary Fig. S2 B) [50,51]. In further agreement with an oncogenic role for sno-miR-28, we also found that sno-miR-28 expression promoted the proliferation (Fig. 5C) and colony forming capacity (Fig. 5D) of MCF10A cells. Taken as a whole, these results demonstrate that sno-miR-28, which targets TAF9B, antagonizes p53 protein levels and is capable of playing an oncogenic role to accelerate breast cell proliferation and colony formation.

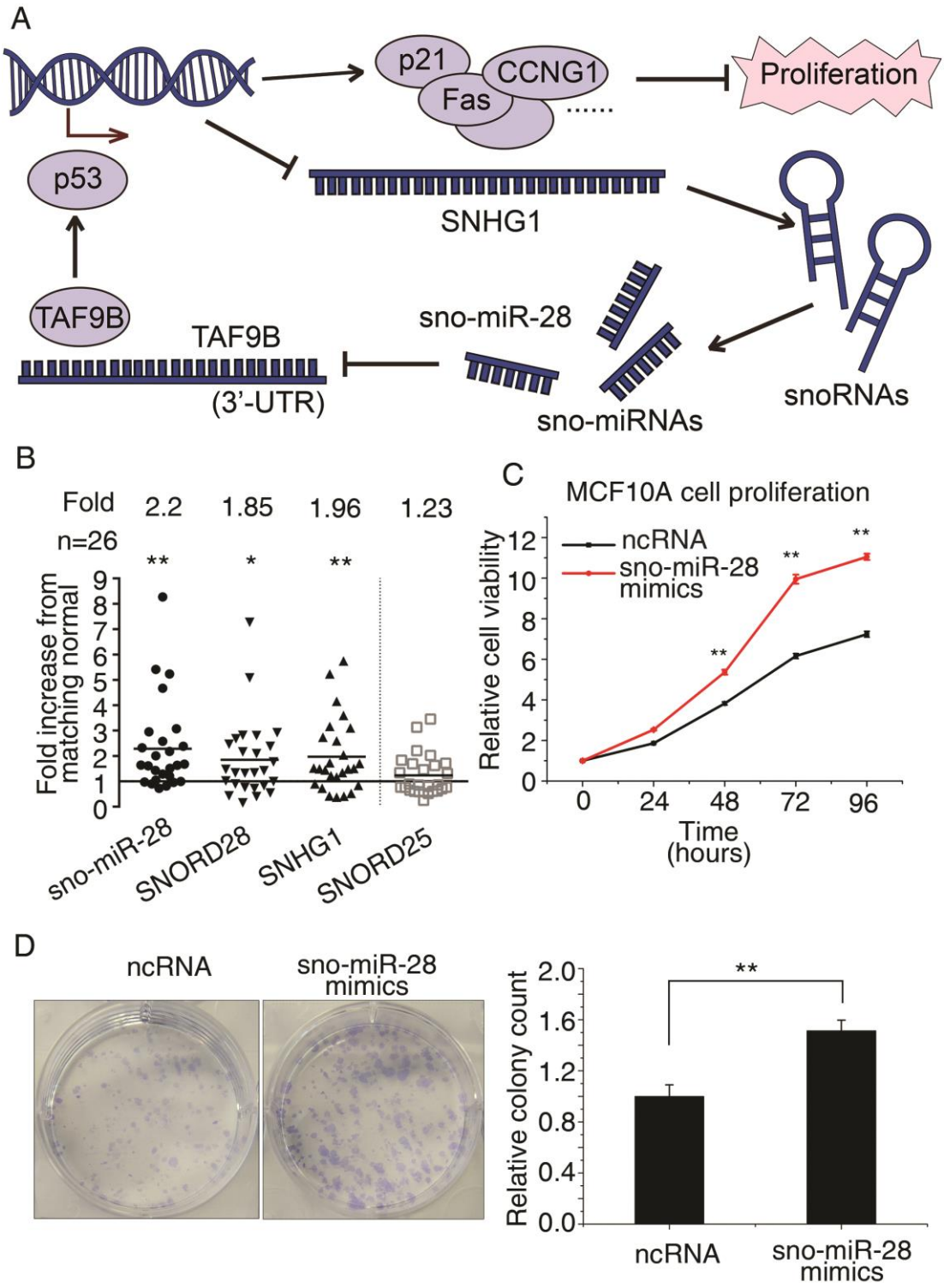


Fig. 5. sno-miR-28 is over-expressed in breast tumours.

(A) The feed-forward loop between p53, sno-miR-28, and TAF9B is shown as proposed. (B) sno-miR-28, SNORD28, and SNORD25 expression levels were

determined using TaqMan assay in breast tumours compared with paired normal adjacent tissues, while SNHG1 expression levels were determined by RT-PCR. In this part of figure, RNA expression levels are shown as the ratio relative to normal tissue expression; e.g., 1 represents equal expression to normal tissues. (C) MCF10A cell proliferation is shown after overexpression of sno-miR-28 compared with negative control RNA (ncRNA). (D) Colony formation assay in MCF10A cells after expression of sno-miR-28 or a negative control RNA (ncRNA). Representative images are included on the left, and relative quantitation of the graph is on the right. ** $p<0.01$ and * $p\leq0.1$ versus controls for all experiments.

Discussion

snoRNAs are processed into miRNAs

Through the use of both microarrays and the analysis of endogenous small RNAs bound to AGO, we discovered that a number of polycistronic snoRNAs are processed into miRNA-sized molecules from a common host gene (SNHG1). Interestingly, we noticed from whole-cell RNA-seq that SNORD29 is far more

abundant than SNORD28, but in contrast, the sno-miRNA derived from the SNORD28 region is the most abundantly recruited to AGO (Fig. 2A). The 5′ end of sno-miR-28 is also very precisely processed, indicating that the processing and recruitment of snoRNA-derived products to AGO can be highly selective *[52].* In addition, up-regulation of SNORD28 but not SNORD25 in breast cancer tissues (Fig. 5B) also indicates existence of extensive post-transcriptional regulation. These evidences strongly imply the presence of a *bonda fide* RNA product, rather than random RNA degradation debris. This provides another example of concomitant post-transcriptional regulation of miRNA levels by regulating their precursors. For instance, LIN28 regulates miR-9 by inducing degradation of its precursor [53], and blocks processing of let-7 by co-transcriptional binding [54]. In accordance with our data, other researchers have also found smaller processed products of snoRNAs to have miRNA-like functions. For example, snoRNA ACA45 is processed into small RNAs 20-25 nt in length, one of which regulates the 3′-UTR of CDC2L6 mRNA by stably associating with AGO proteins [44], and 11 box C/D sno-miRNAs were found to have efficient gene silencing function [39]. In addition, computational analysis has identified 84 intronic miRNAs that are encoded within either box C/D snoRNAs, or in precursors showing similarity to box C/D snoRNAs [40]. Whilst functions for most of these small RNAs have not been determined, the expanded use of deep sequencing technologies have led to the extensive profiling of small RNA fragments derived from a number of sources, including

snoRNAs, tRNAs and rRNAs, and the characterization of their dependence on Drosha or Dicer for their production [55].

sno-miR-28 is a novel member of the p53 regulatory network

While miRNAs have been known to play vital roles in the p53 pathway, our research has revealed a novel miRNA regulatory pathway based on a sno-miRNA in p53 regulation. We discovered that p53 transcriptionally regulates the host gene of sno-miR-28 which targets TAF9B. TAF9B is a subunit of TFIID and TAFIIC, functioning as a stabilizer and co-activator of the p53 protein, and has been reported to be essential for cell viability [47,48]. In addition, TAF9B has been previously discovered to play a role in transcriptional repression and silencing [56]. The regulation of TAF9B by sno-miR-28, and the reciprocal repression of the sno-miR host gene by p53, suggests a role for sno-miR-28: p53 feedback in cancer, which supports recent discoveries that a large number of miRNAs interact with the p53 network as an alternative mechanism of the tumour-suppressor activity of p53 [5-17,19-21,57]. The complex regulatory loop involving p53, SNHG1 and TAF9B is also reminiscent of the feed-back and feed-forward motifs with which miRNAs are frequently associated [20,58].

Furthermore, we noticed that sno-miR-28 actually represses CDKN1A mRNA more effectively than merely knocking down TAF9B using a siRNA (Fig. 4 A,

C, E, H), and this might be explained by the fact that miRNAs usually have a large number of targets, many of which have synergistic functions in the same pathway. For instance, miR-34a targets E2F3 [59], CCNE2 and CDK4 [9], CCND1 and CDK6 [60] in cell cycle regulation, and no single target can fully cover its entire function. We anticipate there will be a large number of additional sno-miR-28 targets awaiting discovery.

sno-miR-28 is overexpressed in breast tumour samples and plays an oncogenic role in breast cells

The regulatory role of sno-miR-28 is further confirmed by our expression profiling studies which relate SNHG1, SNORD28 and sno-miR-28 to breast tumours. Interestingly, SNHG1 was also reported as one of the 5 most significantly upregulated long non-coding RNAs (lncRNAs) in gastric cancer among the 9294 lncRNAs detected [61]. The upregulation of sno-miRNA-28 and SNHG1 in tumours and its relationship to p53 parallels our understanding of other miRNA regulatory loops that closely interact with the p53 pathway and play significant roles in p53-mediated apoptosis and cell cycle arrest. For example, the miR-34 family mediates tumour suppression through a positive feedback loop involving p53 and MDM4 [62]; 15 miRNAs, including the miR-106b/93/25 cluster, miR-17-92 cluster and the miR-106a-92 cluster, are repressed

by p53 and involved with E2F in a feed-forward loop promoting proliferation [63]; whereas miR-192, 194 and 215 are involved in the p53-MDM2 auto-regulatory loop [64]. Whilst our research has revealed yet another feed-forward loop involving p53 and a miRNA, it is surprising that this regulatory loop is based on a snoRNA-derived miRNA, thereby building upon the novel regulatory roles of snoRNAs in cancer.

Collectively, this research builds on a small but growing knowledge base of small RNAs from "unfamiliar" sources playing important biological roles. We believe more novel targets of sno-miR-28 will be identified, and functions for more miRNAs derived from novel sources will be explored. From this point of view, developments of knowledge in cancer regulatory networks and their interaction with small non-coding RNAs may eventually offer insights into new approaches to cancer treatment.

Materials and Methods

HITS-CLIP (High throughput sequencing of crosslinked and immunoprecipitated RNA):

MDA-MB-231 cells were suspended in cold PBS by scraping cross-linked at 254 nM using a Stratalinker. Cell pellets were lysed (0.1% SDS, 0.5% deoxycholate, 0.5% NP-40 with protease inhibitors, Roche, Indianapolis, IN USA) for 10 mins on ice followed by RQ1 DNAse (Promega, Madison, WI USA) at 37°C for 15 mins with shaking. RNAse A/T1 (Ambion, Grand Island, NY USA) was then added for 8 minutes, prior to the addition of EDTA (30 mM). Pellets were then spun (30,000 rpm) and the lysate subjected to immunoprecipitation for 2 h with a pan-anti-AGO antibody (2A8) conjugated to protein-A dynabeads (Invitrogen, Grand Island, NY USA) using bridging rabbit anti-mouse IgG (Jackson ImmunoResearch Laboratories, West Grove, PA USA). Pellets were then successively washed (0.1% SDS, 0.5% deoxycholate, 0.5% NP40 in 1 × PBS; 0.1% SDS, 0.5% deoxycholate, 0.5% NP40 in 5 × PBS; 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 0.5% NP40) and on-bead phosphatase treatment performed for 30 mins with Antarctic phosphatase (New England Biolabs, NEB, Ipswich, MA USA) in the presence of superasin RNAse inhibitor (Ambion). The 3′ RNA linker (CAGACGACGAGCGGG) was labelled with P^{32} using T4-PNK (NEB) and ligated on-bead for 1 h at 16°C with T4 RNA ligase (Fermentas, Thermo Fisher Scientific, Pittsburgh, PA USA). Beads were then washed as previous and treated with PNK to ligate the 5′ RNA linker (AGGGAGGACGAUGCGGxxxG, with "X" representing different nucleotides for barcoding). Beads were resuspended in 4 × LDS Novex loading buffer with 4% B-mercaptoethanol, incubated at 70°C for 10 mins and the supernatant loaded on Novex NuPAGE

4-12% Bis-Tris acrylamide gels (Bio-Rad, Hercules, CA USA). After running, the Ago-RNA complexes were then transferred to nitrocellulose and exposed to film at -80°C for 3 days. Complexes running at ~110 kDa were then excised with a scalpel and resuspended (100 mM Tris pH 7.5, 50 mM NaCl, 10 mM EDTA, 4 mg/ml proteinase K) for 20 min at 37°C. The sample was incubated for an additional 20 minutes in the presence of 3.5 M urea and RNA isolated by a phenol-chloroform extraction. Samples were then run on a 10% denaturing (1:19) polyacrylamide gel and exposed to film with an intensifying screen at -80°C for 5 days. A thin band corresponding to the Ago-miRNA (~110 kDa) was excised, crushed and eluted at 37°C for 1 h (1M NaOAc, pH 5.2, 1 mM EDTA). RNA was precipitated overnight with ethanol, centrifuged and dried, and then resuspended in 8 µl H_2O. Next, reverse transcription primer was added and reverse transcription performed using SuperScript III (Invitrogen). After that, PCR was performed with the above primer and a reverse primer for 25 cycles, and PCR product was run on a 10% native (1:29) polyacrylamide gel, stained with Sybr Gold (Qiagen, Valencia, CA USA) and bands excised over a UV light box. Following this, the DNA was precipitated using isopropanol and a final 10-cycle PCR performed with the HITS-CLIP primers listed in Supplementary Table S3. Reactions were subsequently run on 2% metaphor agarose/TBE gels and bands (~115 bp) excised corresponding to the linker sequence + miRNA CLIP tag. Samples were finally purified using quick-spin columns (Qiagen) and subjected to Illumina sequencing (Illuminia, San Diego, CA USA). Using an in

house Perl script, reads were filtered for average quality and for homopolymeric tracts exceeding 12 nt, trimmed of linker sequence fragments and separated by barcode. The bowtie program [65] was used to align resulting 17 to 30 nt reads to the human genome. The dataset has been previously published [41] and has been deposited at the NCBI (National Centre for Biotechnology Information) SRA (Sequence Read Archive) SRP045204, BioProject PRJNA257235.

Cell lines and reagents

HCT116 (human colon cancer) and its *TP53*$^{-/-}$ derivative were supplied by B. Vogelstein [66]; MDA-MB-231 (human breast cancer) and MCF10A (human breast epithelial) cells were purchased from ATCC (American Type Culture Collection, Manassas, VA USA); wild-type p53-inducible H1299 (mentioned as H1299 below) human non-small cell lung cancer cell line was generated as previously described [67] and was characterized in Fig. 1A using Western blot; WE-68 human Ewing's sarcoma cell line was a kind gift from Prof F. van Valen (Department of Orthopaedic Surgery, Westfälische-Wilhelms-University, Germany) as previously described [68]. HCT116, MDA-MB-231 and H1299 cell lines were cultured in DMEM (Sigma-Aldrich, St. Louis, USA) and WE-68 cells were cultured on type-1 collagen-coated plates (Iwaki, Thermo Fisher, Newport, UK) in RPMI-1640 (Sigma-Aldrich). The media above were

supplemented with 10% (v/v) fetal calf serum and 10mmol/L HEPES. MCF10A cells were cultured in DMEM/F12 medium (Sigma-Aldrich) supplemented with 5% horse serum, 20ng/mL EGF, 0.5μL/ml Hydrocortisone, 100ng/ml Cholera toxin, 10μg/mL Insulin. All media were supplemented with 2mmol/L L-Glutamine, 100 IU/mL penicillin and 100 μg/ml streptomycin, and all cell lines were cultured at 37 °C in a humidified atmosphere at 5% CO_2. All cell lines have been under regular tests every 1-2 months using MycoAlert mycoplasma detection kit (Lonza, Melbourne, VIC Australia), which ensured that all cell lines being used for experiments were safe from mycoplasma contamination.

Nutlin-3a (4-[[(4S,5R)-4,5-bis(4-chlorophenyl)-4,5-dihydro-2-[4-methoxy-2-(1-methylethoxy)phenyl]-1H-imidazol-1-yl]carbonyl]-2-piperazinone) was purchased from Cayman Biochemicals (Ann Arbor, MI USA), and Ponasterone A (PonA) was purchased from Invitrogen.

Breast tumour and paired normal tissues

Breast tumour and paired normal adjacent tissue samples were obtained from 26 breast cancer patients who have received surgery at the Royal Adelaide Hospital (Adelaide, Australia) between 2003 and 2011. Samples were stored in RNAlater solution (Ambion) for RNA stabilization before RNA extraction using miRNeasy kit (Qiagen) following the manufacturer's protocol. RNA concentration was determined using a ND-1000 NanoDrop spectrometer

(Thermo Scientific, Wilmington, DE USA). This research has been approved by the Institutional Review Boards of the University of Adelaide and SA Pathology. All samples were gathered according to the institutional review board-approved protocol and the written informed consent from each patient. Relevant clinical data was retrieved from patients' records including human epidermal growth factor 2 (HER2), estrogen receptor (ER), progesterone receptor (PR) and proliferation index (MIB-1) status.

miRNA overexpression and inhibition

For miRNA overexpression studies, approximately 3×10^5 cells/well were seeded in 6-well plates (for RT-PCR or Western blot) and 1×10^5 cells/well were seeded in 24-well plates (for luciferase assay, proliferation assay and colony formation). Transfections were done 24 hours post-seeding with 50 nM Genepharma miRNA mimics (Genepharma, Shanghai, China) or negative control RNA (ncRNA) (Genepharma) using Lipofectamine RNAiMAX reagent (Invitrogen) following the manufacturer's instructions. The cells were harvested 72 hours post-transfection for RT-PCR, Western blot, or luciferase assay.

For miRNA inhibition studies, approximately 3×10^5 cells/well were seeded in six-well plates. The first transfection was done 24 hours post-seeding with 50 nM miRVana miRNA inhibitor (Ambion) or 50 nM miRVana miRNA inhibitor negative control RNA (ncRNA) (Ambion) using Lipofectamine RNAiMAX reagent (Invitrogen) following the manufacturer's instructions. Cells were harvested 72 hours post-transfection for further experiments.

The efficiency of miRNA transfection was not examined, because according to Thompson et al., perhaps the only accurate way to estimate the outcomes of miRNA mimic transfection is Argonaute HITS-CLIP [70], a method which is extremely time- and labour- consuming. The same authors pointed out that a simple RT-PCR after transient transfection of miRNA mimics does not accurately determine their actual effects. Therefore, measuring transfection efficiency is beyond the scale of the current research.

RNA interference

RNA interference of TAF9B was performed using siTAF9B (AGUAUGAACCAAGGGUUAUAA) (GenePharma). The transfection procedure using Lipofectamine RNAiMAX reagent (Invitrogen) is the same as described above for miRNA overexpression and inhibition.

Luciferase Assay

For the luciferase assay, we cloned the TAF9B 3′-UTR (WT) downstream of the Renilla luciferase gene in the psiCHECK2 dual-luciferase vector (Promega) (for primers, see Supplementary Table S3). The Firefly luciferase gene (which is expressed from the same vector from an HSV-TK promoter) was used as an internal reference. For the mutated TAF9B 3′-UTR (MUT) construct, a 24 bp mutation was introduced at the proposed sno-miR-28 binding site (CTTTCAGAATTGTAAAATGCTATA to GAATTCAAAAAAAAAAAAAAAAAA) (Fig. 3D, left panel) (for primers, see Supplementary Table S3). Approximately 1×10^5 H1299 cells/well in 24-well plates. After 24 hours, in each well we co-transfected 0.4 ng/μL of either WT or MUT luciferase constructs with 8.33 nM of either sno-miR-28 mimics (Genepharma) or negative control RNA (ncRNA) (Genepharma), using Lipofectamine 2000 (Invitrogen) following the manufacturer's instructions. Cells were harvested 72 hours post-transfection, and Renilla and Firefly luminescence were measured by a GloMax 20/20 Luminometer (Promega) following the manufacturer's instructions. A ratio of Renilla/Firefly luminescence intensity was used to indicate the relative luciferase expression activity.

Western blot and ChIP analysis

Cells were rinsed with PBS and lysed in lysis buffer (50 mmol/L Tris-HCl, pH 7.5, 250 mmol/L NaCl, 1% Triton X-100 and 1x protease inhibitors (Roche, South San Francisco, CA)) on ice for 30 min. Insoluble components of cell lysates were removed by centrifugation for 10 min at 4°C, 12,000 *g*, and protein concentration was measured using a bicinchoninic acid protein assay kit (Pierce, Thermo Scientific, Radnor, PA USA). Protein extracts were resolved using SDS PAGE electrophoresis on 10% polyacrylamide gels and electro-transferred to Hybond-C Extra nitrocellulose membranes (GE Healthcare Life Sciences, Pittsburgh, PA USA).

For quantification, p53 protein was probed by mouse monoclonal p53 antibody (DO-1) (Santa Cruz, 1:1000) (Dallas, TX USA); TAF9B was probed by rabbit monoclonal TAF9B antibody (Abcam, at 1:1000) (Cambridge, MA USA); p21 was probed by mouse monoclonal $p21^{WAF1}$ Ab-3 antibody (Thermo Scientific, 1:250); Equal loading was confirmed by blotting of β-actin antibody (Sigma-Aldrich, 1:2000). Chemiluminescent detection of protein was done using appropriate secondary antibodies conjugated with horseradish peroxidise (GE Healthcare) and the enhanced chemiluminescence kit (GE Healthcare) according to the manufacturer's instructions.

ChIP analysis was performed as previously described [57] using mouse p53 antibody (DO-1) (Santa Cruz). The uninduced H1299 cells were used as a p53-null control. Levels of specific promoter DNAs were determined by real-time PCR using specific primers. A negative genomic region which does not contain any p53 responsive element was used as a negative control, and a pair of primers targeting the p53 recognition site of *CDKN1A* was used as a positive control (Supplementary Table S3). Data presented is the mean of three independent biological replicates.

RNA extraction and Real-time PCR

RNA (>250nt) extraction from cells was performed using the RNAeasy mini kit (Qiagen) using on-column RNase-free DNase digestion according to the manufacturer's instructions. After extraction, RNA concentration was measured using a ND-1000 NanoDrop spectrometer (Thermo Scientific). Thereafter, 1μg of total RNA was reverse-transcribed into cDNA using Moloney Murine Leukaemia Virus (M-MLV) Reverse Transcriptase (Promega) with random 6'mer primers or oligo-dT primer (Promega) under the following the manufacturer's instructions. Real-time PCR reaction was performed on a CFX Connect Real-Time PCR Detection System (Bio-Rad) using iTaq Universal SYBR Green Supermix (Bio-Rad). The RT-PCR program was: 95°C for 3min, then start cycles consisting 95°C for 10s and 57/60/61°C for 60s (depending on

the primers used) for 40 cycles. After the reactions were complete, the C_T values were determined using automated threshold settings. In this study, RNA (>250bp) expression was normalized to PSMB4 (Supplementary Table S3).

Small RNAs (including miRNAs and sno-miRNAs) were extracted using the miRNeasy mini kit (Qiagen) following the manufacturer's instructions. After extraction, total RNA concentration was determined using a ND-1000 NanoDrop spectrometer (Thermo Scientific). For quantification of small RNAs, TaqMan gene expression assays were obtained from Applied Biosystems (Grand Island, NY USA), including TaqMan assays for miR-21, miR-155, miR-16, miR-24 and U6 snRNA, and Custom TaqMan assays for SNORD25, SNORD28, sno-miR-25, sno-miR-28. Reverse transcription was performed using a TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems) and specific Reverse Transcription stem-loop primers provided in the TaqMan RNA assays following the manufacturer's instructions. Real-time PCR reaction was performed on a Rotor Gene 6000 Real-Time PCR Machine (Qiagen) using specific TaqMan RNA assays following the manufacturer's instructions. Small RNA expression in cell samples was normalized to U6 snRNA, whereas small RNA expression in tissue samples was normalized to averaged relative expression level of a group of pooled normalizers: miR-24, U44, U48 and U6.

To determine the endogenous expression levels of sno-miRNAs in breast tissues, we performed absolute quantitation of sno-miR-28 and sno-miR-28 in

breast tissues, and miR-155 was used as a positive control. Synthesized sno-miR-28 and miR-155 miRNA oligos were prepared at a gradient of known concentrations at 4 nM, 0.4 nM, 0.04 nM and 0.004 nM, while synthesized sno-miR-25 was prepared at 4 pM, 0.4 pM, 0.04 pM and 0.004 pM to match its lower expression in tissues. RT-PCR was performed subsequently and C_T values were plotted to $\log_{10}$ of the miRNA concentrations respectively to create a standard curve (Supplementary Fig. S1 A, B, C), which was used to determine the endogenous expression levels of sno-miR-28, sno-miR-25 and miR-155 in 26 breast tumour tissues and paired-matching non-malignant breast tissues (Supplementary Fig. S2 A).

Since sno-miR-28 is processed from a region of SNORD28 near its 3′ end (Fig. 2A), there is possibility of mispriming in stem-loop PCR. To make sure that our sno-miRNA expression data is not a result of mispriming, we conducted sno-miR-28 TaqMan RT-PCR using 2nM of synthesized SNORD28 mimics or sno-miR-28 mimics as templates.

Target gene	Forward primer (5'-3')	Reverse Primer (5'-3')	Description
p53	CCCCTCCTGGCCCCTGTCATCTTC	GCAGCGCCTCACAACCTCCGTCAT	For RT-PCR detection of p53 mRNA
TAF9B	CCGCGGATGACGAGTGGCTG	TCAGCACGACACTGGATTGCCA	For RT-PCR detection of TAF9B mRNA
SNHG1	CGCGCACGTTGGAACCGAAG	GCTGGCCCTTTGAGCCAAGCA	For RT-PCR detection of SNHG1 mRNA
SNORD25	TGATGAGGACCTTTTCACAGACCT	TCAGAGTTATTTATCCTCACGGAGC	For RT-PCR detection of SNORD25
SNORD28	AGCTGATGTTCTGTGAGGTACA	GCCATCAGAACTCTAACATGCTATT	For RT-PCR detection of SNORD28
CDKN1A	TGGACCTGGAGACTCTCAGGGTCG	TTAGGGCTTCCTCTTGGAGAAGATC	For RT-PCR detection of p21 mRNA
RRM2B	GCCAGGACTCACTTTTTCCA	TCAGGCAAGCAAAGTCACAG	For RT-PCR detection of p53R2 mRNA
CCNG1	GTCCCATTGGCAACTGACTT	TGACATGCCTTCAGTTGAGC	For RT-PCR detection of CCNG1 mRNA
FAS	ATGCTGGGCATCTGGACCCT	GCCATGTCCTTCATCACACAA	For RT-PCR detection of FAS mRNA
HDM2	TCTACAGGGACGCCATCGA	CTGATCCAACCAATCACCTGAA	For RT-PCR detection of HDM2 mRNA
ChIP negative site	GCTGTGGTGTGTAGGAAGCAGCAG	AAAGGAGGCCAAAGTCGGGCAATG	Negative control for p53 ChIP
SNHG1 p53-RE	AGACAATGGAGAGGCACTGAACTGA	GGTCAGGTCCCACTTCAGCTCACAA	For ChIP PCR of p53-RE in SNHG1 promoter
CDKN1A ChIP	GTGGCTCTGATTGGCTTTCTG	CTGAAAACAGGCAGCCCAAG	CDKN1A ChIP primers For cloning of TAF9B 3'-UTR into the psiCHECK2 vector
TAF9B 3'-UTR	GCAGCTAGCAAGGAAAGTTGGTTTTGA	GCTGCGGCCGCTATTAACCTTGGAACAAA	
TAF9B-mut-1	TGATAACTCTTCCATGTTTCACATC	GCAATAATTCCTTACAGTCATTTTTTTTTTTTTTTTTGAATTCAATTTTCACATATACCAATT	For introducing a mutation at the sno-miR-28 recognition sequence in the TAF9B 3'-UTR cloned in the psiCHECK2 vector
TAF9B-mut-2	AATTGGTATATGTGAAAATTGAATTCAAAAAAAAAAAAAAAAAATGACTGTAAGGAATTATTGC	CAAACCCTAACCACCGCTTA	

S3 Table. Primer list

Primers used in this study are listed by targets and usage.

Cell proliferation assay and colony formation

MCF10A breast cells were plated in 24-well plates at about 1 x 10^5 cells/well and transfected according to the procedure described above as miRNA overexpression. At 72 hours post-transfection, the cells were harvested for proliferation or colony formation assays.

For proliferation assay, the cells were re-plated in 96-well plates in 5 replicate wells for each experimental group (i.e., ncRNA, sno-miR-28 mimics) at 2000 cells/well and were incubated in normal culture conditions for 6 hours to attach. After that, these cells were processed at 0, 24, 48, 72, 96 hour time points using the CellTiter-Glo Luminescent Cell Viability Assay Kit (Promega) according to the manufacturer's instructions. Fluorescence was measured using a LUMIstar Galaxy luminometer (BMG Labtech). The fluorescence readings in each experimental group were normalized against the 0-hour point.

For colony formation, cells were re-plated in 6-well plates at 500/1000/2000 cells/well and were incubated in normal culturing conditions for 7 days. Then the cells were fixed in methanol for 15 minutes, and stained for 1 hour in 1:20

Giemsa Stain, Modified Solution (Sigma-Aldrich). Afterwards, colonies were photographed by a digital camera and counted.

Microarray profiling

An Affymetrix Human Gene 1.0 ST array containing 234 annotated snoRNAs (Affymetrix, Santa Clara, CA USA) was used to identify snoRNAs that were differentially expressed following the manufacturer's instructions as previously described [57,69]. "Fold change" refers to the ratio of expression in induced (H1299 treated with PonA and WE-68 treated with Nutlin) versus uninduced cells. Statistical criteria for microarray: fold change > 1.085 or <0.915, $p<0.05$.

S4 Table. Significant genes in the microarray (attached as a standalone file)

All genes significantly regulated in both H1299 and WE-68 cell lines upon p53 activation are listed with fold change and p values. "Fold change" refers to the ratio of expression in induced (H1299 treated with PonA and WE-68 treated with Nutlin) versus uninduced cells. Statistical criteria for microarray: fold change > 1.085 or <0.915, $p<0.05$.

The S4 Table is also available for open access at the Plos One website, as per it was published:

http://journals.plos.org/plosone/article/asset?unique&id=info:doi/10.1371/journal.pone.0129190.s007

Statistical analysis

Results are given as mean of at least three independent experiments ± standard deviation. Student's t-test was performed using replicate values to indicate significance. Values of $p<0.05$ were considered statistically significant (as labelled with ** in figures), while values of $p<0.1$ were indicated by *.

Abbreviations: RT-PCR, real-time quantitative PCR following reverse transcription; ChIP, chromatin immunoprecipitation; PBS, phosphate-buffered saline; cDNA, complementary DNA; ChIP-Seq, chromatin immunoprecipitation combined DNA sequencing; ncRNA, negative control RNA; RNA-Seq, RNA Sequencing; lncRNA, long non-coding RNA; snoRNA, small nucleolar RNA; miRNA, micro-RNA; SNORD, polycistronic C/D box snoRNA; sdRNA, snoRNA-derived small RNA; sno-miRNA, snoRNA-derived micro-RNA; siRNA, small interference RNA; LNA, Locked Nucleic Acid; HITS-CLIP, high throughput sequencing of crosslinked and

immunoprecipitated RNA; 3′-UTR, 3′-untranslated region; p53 RE, p53 responsive element; AGO, Argonaute; RISC: RNA-induced silencing complex; PonA: Ponasterone A.

Acknowledgements

We thank Renee B. Schulz, Rebecca C. Haycox, Alaknanda Adwal, Kathleen I. Pishas, Daniel W. Thomson and Jan M. Szubert for technical assistance.

Funding

This research received funding from National Health and Medical Research Council of Australia (NHMRC) grant 44112162, NHMRC grant GNT1034633, and from the National Breast Cancer Foundation Australia. C. Bracken is supported by a Fellowship from the National Breast Cancer Foundation Australia, and F. Yu is supported by Adelaide Scholarships International.

References

1. Beckerman R, Prives C (2010) Transcriptional regulation by p53. Cold Spring Harb Perspect Biol 2: a000935.

2. Nikulenkov F, Spinnler C, Li H, Tonelli C, Shi Y, et al. (2012) Insights into p53 transcriptional function via genome-wide chromatin occupancy and gene expression analysis. Cell death and differentiation 19: 1992-2002.

3. Neilsen PM, Noll JE, Suetani RJ, Schulz RB, Al-Ejeh F, et al. (2011) Mutant p53 uses p63 as a molecular chaperone to alter gene expression and induce a pro-invasive secretome. Oncotarget 2: 1203-1217.

4. Shi M, Liu D, Shen B, Guo N (2010) Helpers of the cellular gatekeeper-miRNAs dance in P53 network. Biochim Biophys Acta 1805: 218-225.

5. Yan HL, Xue G, Mei Q, Wang YZ, Ding FX, et al. (2009) Repression of the miR-17-92 cluster by p53 has an important function in hypoxia-induced apoptosis. The EMBO journal 28: 2719-2732.

6. Lin J, Huo R, Xiao L, Zhu X, Xie J, et al. (2013) A novel p53-miR-22-Cyr61 axis in synovial cells regulates inflammation in Rheumatoid Arthritis. Arthritis and rheumatism.

7. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, et al. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129: 1311-1323.

8. Corney DC, Flesken-Nikitin A, Godwin AK, Wang W, Nikitin AY (2007) MicroRNA-34b and MicroRNA-34c are targets of p53 and cooperate in control of cell proliferation and adhesion-independent growth. Cancer Res 67: 8433-8438.

9. He L, He X, Lim LP, de Stanchina E, Xuan Z, et al. (2007) A microRNA component of the p53 tumour suppressor network. Nature 447: 1130-1134.

10. Raver-Shapira N, Marciano E, Meiri E, Spector Y, Rosenfeld N, et al. (2007) Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. Mol Cell 26: 731-743.

11. Tazawa H, Tsuchiya N, Izumiya M, Nakagama H (2007) Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells. Proc Natl Acad Sci U S A 104: 15472-15477.

12. Sachdeva M, Zhu S, Wu F, Wu H, Walia V, et al. (2009) p53 represses c-Myc through induction of the tumor suppressor miR-145. Proc Natl Acad Sci U S A 106: 3207-3212.

13. Georges SA, Biery MC, Kim SY, Schelter JM, Guo J, et al. (2008) Coordinated regulation of cell cycle transcripts by p53-Inducible microRNAs, miR-192 and miR-215. Cancer Res 68: 10105-10112.

14. Braun CJ, Zhang X, Savelyeva I, Wolff S, Moll UM, et al. (2008) p53-Responsive micrornas 192 and 215 are capable of inducing cell cycle arrest. Cancer Res 68: 10094-10104.

15. Jin L, Hu WL, Jiang CC, Wang JX, Han CC, et al. (2011) MicroRNA-149*, a p53-responsive microRNA, functions as an oncogenic regulator in human melanoma. Proc Natl Acad Sci U S A 108: 15840-15845.

16. Kim T, Veronese A, Pichiorri F, Lee TJ, Jeon YJ, et al. (2011) p53 regulates epithelial-mesenchymal transition through microRNAs targeting ZEB1 and ZEB2. J Exp Med 208: 875-883.

17. Chang CJ, Chao CH, Xia W, Yang JY, Xiong Y, et al. (2011) p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. Nat Cell Biol 13: 317-323.

18. Xiao J, Lin H, Luo X, Wang Z (2011) miR-605 joins p53 network to form a p53:miR-605:Mdm2 positive feedback loop in response to stress. The EMBO journal 30: 5021.

19. Barsotti AM, Beckerman R, Laptenko O, Huppi K, Caplen NJ, et al. (2012) p53-Dependent induction of PVT1 and miR-1204. The Journal of biological chemistry 287: 2509-2519.

20. Ren ZJ, Nong XY, Lv YR, Sun HH, An PP, et al. (2014) Mir-509-5p joins the Mdm2/p53 feedback loop and regulates cancer cell growth. Cell Death Dis 5: e1387.

21. Nakazawa K, Dashzeveg N, Yoshida K (2014) Tumor suppressor p53 induces miR-1915 processing to inhibit Bcl-2 in the apoptotic response to DNA damage. FEBS J.

22. Liu Q, Huang J, Zhou N, Zhang Z, Zhang A, et al. (2013) LncRNA loc285194 is a p53-regulated tumor suppressor. Nucleic acids research 41: 4976-4987.

23. Budde A, Grummt I (1999) p53 represses ribosomal gene transcription. Oncogene 18: 1119-1124.

24. Zhai W, Comai L (2000) Repression of RNA polymerase I transcription by the tumor suppressor p53. Mol Cell Biol 20: 5930-5938.

25. Thorenoor N, Slaby O (2014) Small nucleolar RNAs functioning and potential roles in cancer. Tumour Biol.

26. Su H, Xu T, Ganapathy S, Shadfan M, Long M, et al. (2014) Elevated snoRNA biogenesis is essential in breast cancer. Oncogene 33: 1348-1358.

27. Vousden KH, Prives C (2009) Blinded by the Light: The Growing Complexity of p53. Cell 137: 413-431.

28. Eisenberg E, Levanon EY (2013) Human housekeeping genes, revisited. Trends Genet 29: 569-574.

29. Spurgers KB, Gold DL, Coombes KR, Bohnenstiehl NL, Mullins B, et al. (2006) Identification of cell cycle regulatory genes as principal targets of p53-mediated transcriptional repression. J Biol Chem 281: 25134-25142.

30. Li B, Lee MY (2001) Transcriptional regulation of the human DNA polymerase delta catalytic subunit gene POLD1 by p53 tumor suppressor and Sp1. J Biol Chem 276: 29729-29739.

31. Date DA, Jacob CJ, Bekier ME, Stiff AC, Jackson MW, et al. (2007) Borealin is repressed in response to p53/Rb signaling. Cell Biol Int 31: 1470-1481.

32. Wiebusch L, Hagemeier C (2010) p53- and p21-dependent premature APC/C-Cdh1 activation in G2 is part of the long-term response to genotoxic stress. Oncogene 29: 3477-3489.

33. Nakamura T, Saito H, Takekawa M (2013) SAPK pathways and p53 cooperatively regulate PLK4 activity and centrosome integrity under stress. Nat Commun 4: 1775.

34. Xu X, Qiao W, Linke SP, Cao L, Li WM, et al. (2001) Genetic interactions between tumor suppressors Brca1 and p53 in apoptosis, cell cycle and tumorigenesis. Nat Genet 28: 266-271.

35. Kho PS, Wang Z, Zhuang L, Li Y, Chew JL, et al. (2004) p53-regulated transcriptional program associated with genotoxic stress-induced apoptosis. J Biol Chem 279: 21183-21192.

36. Yan HL, Xue G, Mei Q, Wang YZ, Ding FX, et al. (2009) Repression of the miR-17-92 cluster by p53 has an important function in hypoxia-induced apoptosis. EMBO J 28: 2719-2732.

37. Smeenk L, van Heeringen SJ, Koeppel M, van Driel MA, Bartels SJJ, et al. (2008) Characterization of genome-wide p53-binding sites upon stress response. Nucleic Acids Research 36: 3639-3654.

38. Taft RJ, Glazov EA, Lassmann T, Hayashizaki Y, Carninci P, et al. (2009) Small RNAs derived from snoRNAs. RNA 15: 1233-1240.

39. Brameier M, Herwig A, Reinhardt R, Walter L, Gruber J (2011) Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. Nucleic Acids Res 39: 675-686.

40. Ono M, Scott MS, Yamada K, Avolio F, Barton GJ, et al. (2011) Identification of human miRNA precursors that resemble box C/D snoRNAs. Nucleic Acids Res 39: 3879-3891.

41. Bracken CP, Li X, Wright JA, Lawrence DM, Pillman KA, et al. (2014) Genome-wide identification of miR-200 targets reveals a regulatory network controlling cell invasion. EMBO J 33: 2040-2056.

42. Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42: D68-73.

43. Kawaji H, Nakamura M, Takahashi Y, Sandelin A, Katayama S, et al. (2008) Hidden layers of human small RNAs. BMC Genomics 9: 157.

44. Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, et al. (2008) A human snoRNA with microRNA-like functions. Mol Cell 32: 519-528.

45. Saraiya AA, Wang CC (2008) snoRNA, a novel precursor of microRNA in Giardia lamblia. PLoS Pathog 4: e1000224.

46. Reuter JS, Mathews DH (2010) RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics 11: 129.

47. Frontini M, Soutoglou E, Argentini M, Bole-Feysot C, Jost B, et al. (2005) TAF9b (formerly TAF9L) is a bona fide TAF that has unique and overlapping roles with TAF9. Mol Cell Biol 25: 4638-4649.

48. Lu H, Levine AJ (1995) Human TAFII31 protein is a transcriptional coactivator of the p53 protein. Proc Natl Acad Sci U S A 92: 5154-5158.

49. Cao WJ, Wu HL, He BS, Zhang YS, Zhang ZY (2013) Analysis of long non-coding RNA expression profiles in gastric cancer. World journal of gastroenterology : WJG 19: 3658-3664.

50. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, et al. (2005) MicroRNA gene expression deregulation in human breast cancer. Cancer Res 65: 7065-7070.

51. Mattiske S, Suetani RJ, Neilsen PM, Callen DF (2012) The oncogenic role of miR-155 in breast cancer. Cancer Epidemiol Biomarkers Prev 21: 1236-1243.

52. Thomson DW, Bracken CP, Goodall GJ (2011) Experimental strategies for microRNA target identification. Nucleic Acids Res 39: 6845-6853.

53. Nowak JS, Choudhury NR, de Lima Alves F, Rappsilber J, Michlewski G (2014) Lin28a regulates neuronal differentiation and controls miR-9 production. Nat Commun 5: 3687.

54. Van Wynsberghe PM, Kai ZS, Massirer KB, Burton VH, Yeo GW, et al. (2011) LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans. Nat Struct Mol Biol 18: 302-308.

55. Burroughs AM, Ando Y, de Hoon MJ, Tomaru Y, Suzuki H, et al. (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: 158-177.

56. Chen Z, Manley JL (2003) In vivo functional analysis of the histone 3-like TAF9 and a TAF9-related factor, TAF9L. J Biol Chem 278: 35172-35183.

57. Neilsen PM, Noll JE, Mattiske S, Bracken CP, Gregory PA, et al. (2013) Mutant p53 drives invasion in breast tumors through up-regulation of miR-155. Oncogene 32: 2992-3000.

58. Vera J, Lai X, Schmitz U, Wolkenhauer O (2013) MicroRNA-regulated networks: the perfect storm for classical molecular biology, the ideal scenario for systems biology. Adv Exp Med Biol 774: 55-76.

59. Welch C, Chen Y, Stallings RL (2007) MicroRNA-34a functions as a potential tumor suppressor by inducing apoptosis in neuroblastoma cells. Oncogene 26: 5017-5022.

60. Sun F, Fu H, Liu Q, Tie Y, Zhu J, et al. (2008) Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest. FEBS Lett 582: 1564-1568.

61. Cao WJ, Wu HL, He BS, Zhang YS, Zhang ZY (2013) Analysis of long non-coding RNA expression profiles in gastric cancer. World J Gastroenterol 19: 3658-3664.

62. Okada N, Lin CP, Ribeiro MC, Biton A, Lai G, et al. (2014) A positive feedback between p53 and miR-34 miRNAs mediates tumor suppression. Genes Dev.

63. Brosh R, Shalgi R, Liran A, Landan G, Korotayev K, et al. (2008) p53-Repressed miRNAs are involved with E2F in a feed-forward loop promoting proliferation. Mol Syst Biol 4: 229.

64. Pichiorri F, Suh SS, Rocci A, De Luca L, Taccioli C, et al. (2010) Downregulation of p53-inducible microRNAs 192, 194, and 215 impairs the p53/MDM2 autoregulatory loop in multiple myeloma development. Cancer Cell 18: 367-381.

65. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25.

66. Bunz F, Dutriaux A, Lengauer C, Waldman T, Zhou S, et al. (1998) Requirement for p53 and p21 to sustain G2 arrest after DNA damage. Science 282: 1497-1501.

67. Noll JE, Jeffery J, Al-Ejeh F, Kumar R, Khanna KK, et al. (2012) Mutant p53 drives multinucleation and invasion through a process that is suppressed by ANKRD11. Oncogene 31: 2836-2848.

68. van Valen F (1999) Ewing's Sarcoma Family of Tumors. Human Cell Culture 1: 55-85.

69. Pishas KI, Neuhaus SJ, Clayer MT, Schreiber AW, Lawrence DM, et al. (2014) Nutlin-3a efficacy in sarcoma predicted by transcriptomic and epigenetic profiling. Cancer Res 74: 921-931.

70. Thomson DW, Bracken CP, Szubert JM, Goodall GJ. On Measuring miRNAs after Transient Transfection of Mimics or Antisense Inhibitors. Plos One. 2013;8(1).

Chapter 3

Naturally existing isoforms of miR-222 play pro-apoptotic roles

Feng Yu[1, 3], Katherine A. Pillman[1, 2], Corine T. Neilsen[1], John Toubia[1, 2], David M. Lawrence[1, 2], Anna Tsykin[1, 2], David F. Callen[3], Gregory J. Goodall[*, 1, 4, 5 and] Cameron P. Bracken[*, 1, 4, 5].

[1]Centre for Cancer Biology, SA Pathology, Adelaide, SA, Australia

[2]ACRF Cancer Genomics Facility, Centre for Cancer Biology, SA Pathology, Adelaide, Australia

[3]School of Medicine, Discipline of Medicine, University of Adelaide, SA, Australia

[4]School of Molecular and Biomedical Science, University of Adelaide, SA, Australia

[5]School of Pharmacy and Medical Sciences, University of South Australia

[*]Corresponding Authors

Prelude

This chapter utilises next generation sequencing to explore the expression and function of miRNA isoforms using miR-222 as an example. Naturally existing isoforms prevalently expressed in normal and cancer tissues are described with significant functional impacts in cell physiology and apoptosis. These findings have implications regarding previous published studies on the roles of miRNAs in cancer.

This chapter is presented as a manuscript in preparation for submission to a journal for publication.

Contribution of the candidate: Figure 1A, Figure 2A, B, C, Figure 3A, B, C, Figure 4 all panels, Figure 5C, Figure 6B, C, Figure 7 all panels, Supplementary Figure 1, Supplementary Figure 2B, both supplementary tables, and writing the manuscript.

Statement of Authorship

Title of Paper	Naturally existing isoforms of miR-222 play distinct pro-apoptotic roles
Publication Status	Manuscript prepared for publication

Candidate

Name of Author	Feng Yu		
Contribution to the Paper	Performed analysis on cell and tissue experiments, interpreted data, wrote the manuscript.		
Signature		Date	23 Feb 2016

By signing the Statement of Authorship, each author certifies that the candidate's stated contribution to the publication is accurate, and permission is granted for the candidate in include the publication in the thesis.

Name of Author	Katherine Pillman		
Contribution to the Paper	Performed bioinformatical analysis, interpreted data		
Signature		Date	23/2/2016

Name of Author	Corine Neilsen		
Contribution to the Paper	Performed analysis on cell experiments, interpreted data.		
Signature		Date	

Name of Author	John Toubia		
Contribution to the Paper	Performed bioinformatical analysis, helped in data interpretation.		
Signature		Date	26/02/16

Name of Author	David Lawrence		
Contribution to the Paper	Performed bioinformatical analysis, helped in data interpretation.		
Signature		Date	26/02/16

Name of Author	Anna Tsykin		
Contribution to the Paper	Performed bioinformatical analysis, helped in data interpretation.		
Signature		Date	29/02/16

Name of Author	David Callen		
Contribution to the Paper	Supervised development of work		
Signature		Date	24/2/16

Name of Author	Gregory Goodall		
Contribution to the Paper	Supervised development of work, helped in data interpretation and manuscript evaluation.		
Signature		Date	1/3/16

Name of Author	Cameron Bracken		
Contribution to the Paper	Supervised development of work, helped in data interpretation and manuscript preparation.		
Signature		Date	1/3/16

Statement of Authorship

Title of Paper	Naturally existing isoforms of miR-222 play distinct pro-apoptotic roles
Publication Status	Manuscript prepared for publication

Candidate

Name of Author	Feng Yu		
Contribution to the Paper	Performed analysis on cell and tissue experiments, interpreted data, wrote the manuscript.		
Signature		Date	

By signing the Statement of Authorship, each author certifies that the candidate's stated contribution to the publication is accurate, and permission is granted for the candidate in include the publication in the thesis.

Name of Author	Katherine Pillman		
Contribution to the Paper	Performed bioinformatical analysis, interpreted data		
Signature		Date	

Name of Author	Corine Neilsen		
Contribution to the Paper	Performed analysis on cell experiments, interpreted data.		
Signature		Date	26/02/2016

Abstract

Deep-sequencing reveals extensive variation in the sequence of endogenously expressed microRNAs (termed "isomiRs") in human cell lines and tissues, especially in relation to the 3′ end. From the immunoprecipitation and sequencing of the microRNA-binding protein Argonaute, we also note the extensive presence of 3′-isomiRs, including for miR-222 where the majority of endogenously expressed miR-222 is extended by 1-5 nucleotides compared to the canonical sequence. We demonstrate this 3′ heterogeneity has dramatic implications for the phenotype of miR-222 transfected cells, with longer isoforms promoting apoptosis in a size (but not 3′ sequence) -dependent manner. The apoptosis induced by longer miR-222 isomiRs failed to be correlated with an interferon response, but did correspond with the widespread inhibition of RNA and protein expression of key PI3K-AKT components. A PI3K regulatory subunit, PIK3R3, was of particular interest, as its knockdown phenocopied the expression of longer 222 isoforms in terms of apoptosis and the inhibition of other PI3K-AKT genes. As this work demonstrates the capacity for 3′ isomiRs to mediate differential functions, we contend more attention needs to be given to 3′ variance given the widespread prevalence of this class of isomiR.

Introduction

miRNAs are widely known to be vital regulatory molecules in human cells, playing widespread roles in virtually all cellular processes including proliferation, migration, development, immunity, and metabolism [1]. MicroRNAs function as 20-24 nt guides that recognize complementary mRNA sequences and act to bring effector proteins [Argonaute (AGO) and other components of the RNA-induced silencing complex (RISC)] to targets to silence gene expression through a combination of mRNA destabilisation and translational inhibition [2, 3]. The processing of miRNAs typically requires the sequential activities of the RNAses Drosha and Dicer, processing double-stranded primary miRNA (pri-miRNA) transcripts into single stranded ~22nt mature microRNAs. The activities of both enzymes are regulated through additional proteins and constrained by the physical characteristics of the pri-miRNAs being processed [3]. With that said however, the cleavage of Drosha and Dicer is not always precise and additional mechanisms exist to edit the double stranded pri-miRNAs, or to lose or add nucleotides post- cleavage, resulting in a range of heterogeneous products [4]. These naturally occurring heterogenic miRNA species are collectively termed “isomiRs”.

The specificity of miRNA cleavage is dependent on a host of factors including sequence motifs around the cleavage sites on both the pri-miRNA and pre-miRNA, and also the secondary structure of the miRNA precursors [5]. Surprisingly, the annotated miRNA sequences, according to the miRBase [6] and GenBank [7], are not always the most abundant forms e.g., canonical miR-222 is annotated as AGCUACAUCUGGCUACUGGGU (miRBase/GenBank), but this is less abundant than isomiRs bearing 1 and 4 nt 3′ extensions (Figure 1A). The expression patterns of isomiRs are also variable between cells and are subject to regulation during development and under such conditions as infection [8-14]. Differential isomiR expression is also reported in cancer. For example, the relative abundance of isomiRs are related to changes in differentiation status in chronic myeloid leukaemia [8] and the expression profile of isomiRs in lymphoblastoid cell lines were shown to be both gender-specific and population specific [10]. miR-21, miR-221, miR-155, miR-30e and miR-25 have different isomiR expression patterns between breast cancer and normal tissues [15], and seven isomiRs, miR-10b, -30d, -93, -143, -181a, -182, and -744, were sub-grouped to develop an biomarker panel for coronary artery disease [16]. IsomiR-specific biological functions have also been recently demonstrated: e.g., miR451a.1 has been found to repress cell migration and invasion (thus functioning as a tumour suppressor), while miR-451a was associated with melanotic phenotypes [17]. Similarly, the 5′-variants of miR-142-3p and miR-9 have diverse target cohorts [18-20]. These examples both

concern the phenomenon of "seed-shifting" whereby 5′ variance causes repositioning of primary target interaction site (the "seed") which is defined through binding to AGO relative to the 5′ end of the miRNA. This leads to differential targeting profiles. 5′ variation is relatively uncommon however, with the majority of isomiRs displaying 3′ variance. The extent to which this regulates miRNA function is not well explored.

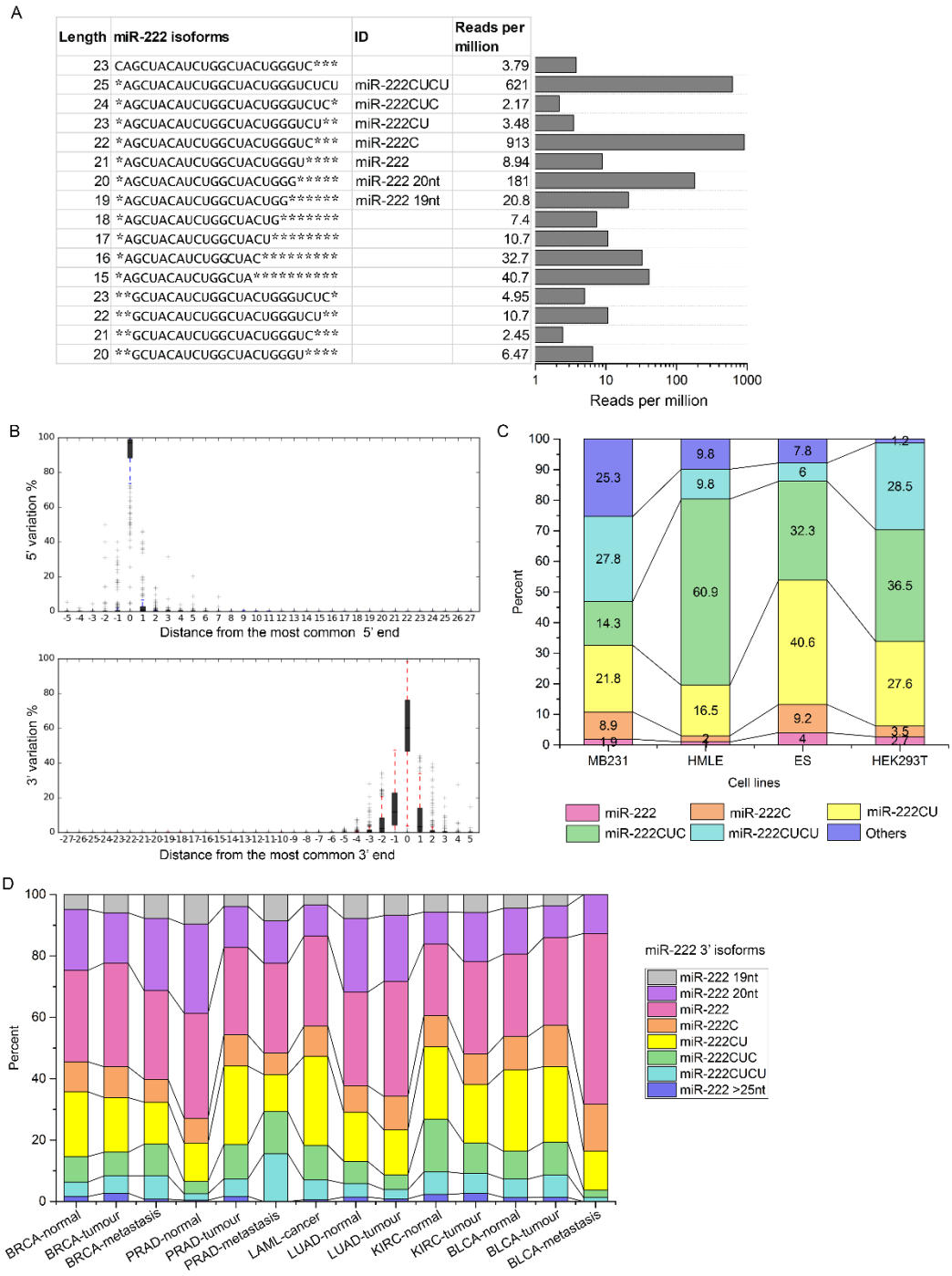


Figure 1: miRNA isoforms are expressed in abundance endogenously.

(A) The major isoforms of miR-222 and their expression levels (reads per million) were collected from the RNA deep-sequencing data published by the

miRBase (Release 21: June 2014). (B) The distribution of 5′ and 3′ variations of 227 miRNAs is shown classified by the distance from the most common end; data was obtained from the Cancer Genome Atlas (TCGA). (C) The expression of miR-222 isoforms in cell lines was determined by HITS-CLIP (High throughput sequencing of crosslinked and immunoprecipitated RNA). ES: embryonic stem cells. (D) Deep-sequencing data of miR-222 isoforms was obtained from the TCGA to characterize their proportion of expression in different tissue and cancer types. BRCA: breast cancer; PRAD: prostate adenocarcinoma; LAML: acute myeloid leukaemia; LUAD: lung adenocarcinoma; KIRC: kidney renal clear cell carcinoma; BLCA: bladder urothelial carcinoma.

Results

Endogenous expression of miR-222 isoforms

To investigate the overall expression pattern of isomiRs, deep-sequencing data of tissue samples was retrieved from the Cancer Genome Atlas (TCGA) to characterize the distribution of 5′ and 3′ variations of 227 of the most abundant miRNAs. Bioinformatic analysis revealed that 3′ heterogeneity is the dominating form of isomiRs compared to 5′ variation, both in terms of the

proportion and the differential size range of 3′:5′ isomiRs (Figure 1B). As miRNAs function within the context of AGO binding, we examined the isomiR profile of human MDA-MB-231 cells from which we immunoprecipitated AGO and sequenced the associated small RNA pool. From this, we selected miR-222 for further study as it is of significance to cancer [21-27] and it represents one of the most 3′ variable miRNAs in our library, with 98% of endogenous AGO-bound miR-222 having 3′ extensions of between 1 and 4 nucleotides (Figure 1C). Similar trends were observed within the small RNA sequencing of additional cell lines (Figure 1C) and from the TCGA of normal and matched tumour samples (Figure 1D), demonstrating a high proclivity for miR-222 3′ variance across multiple contexts.

miR-222 isoforms significantly alter cell morphology

To investigate their biological roles, we transfected miRNA mimics for miR-222 isoforms between 21nt and 25nt in length along with size-matched negative control RNAs (nc21 and nc25). We noted morphological alteration including blebbing and cell shrinkage that became progressively more severe with the transfection of larger miR-222 isoforms, a phenotype suggestive of apoptosis. No morphological changes were reported with the transfection of either 21nt or 25nt scrambled controls (Figure 2A). The observation was further supported

by immunofluorescent staining of β-actin and β-tubulin which indicated alteration of cytoskeletal structure as a result of miR-222CUCU transfection (Supplementary Fig. 2). In addition, miR-222GAGA, which is a synthetic 25nt miR-222 isomiR with the 3′ extension being mutated from CUCU into GAGA, exerts similar morphological changes in MCF10A cells to the endogenous sequence (Figure 2B). We therefore conclude that the changes in morphology are not due to the exact sequence of the 3′ extension, but occur in response to length. Further investigation revealed the effects of longer miR-222 isomiRs are not restricted to MCF10A cells, but are also present in multiple cell lines from various origins, e.g., KGN, MB468, MCF7 and T47D (Figure 2C).

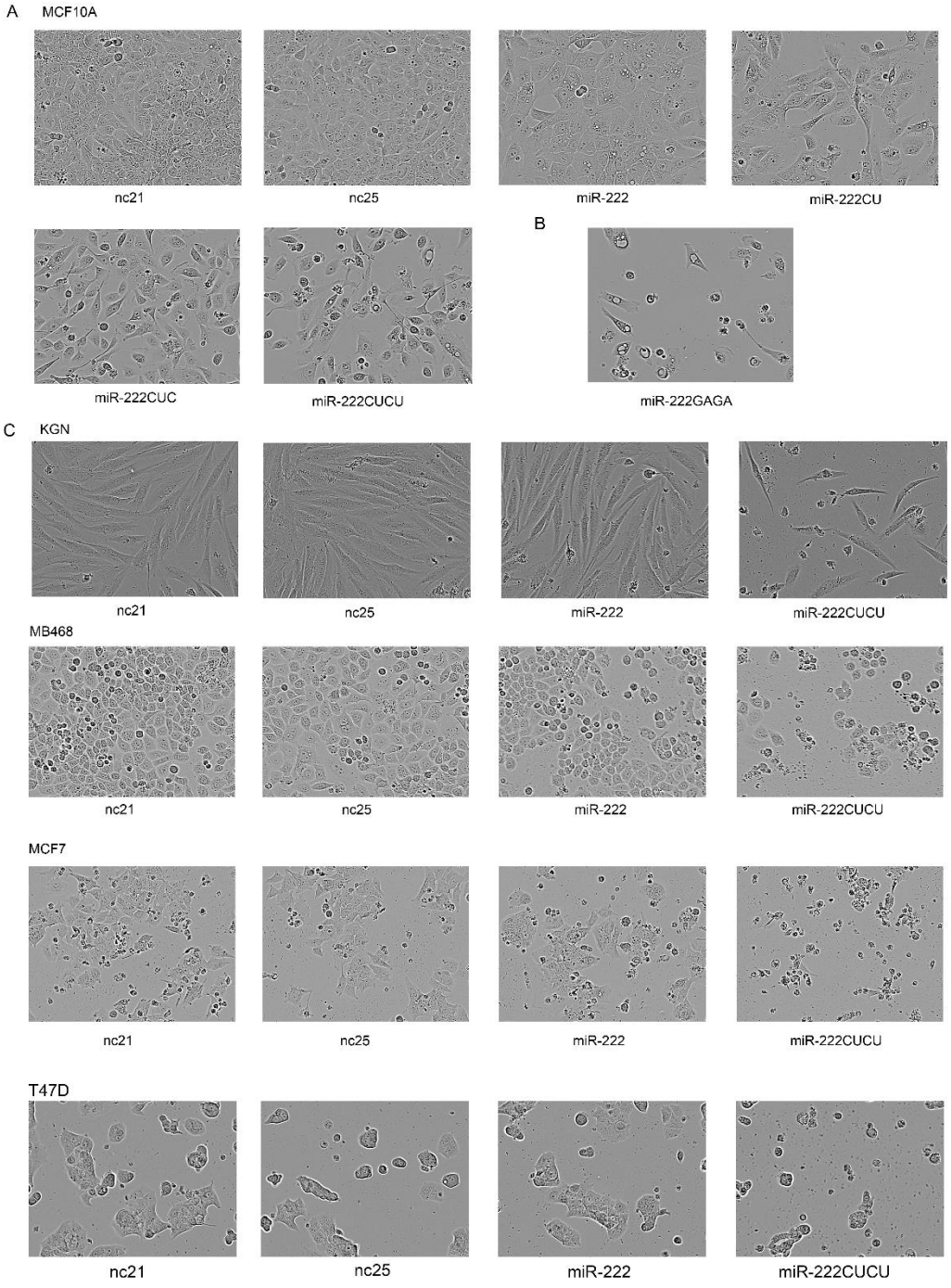
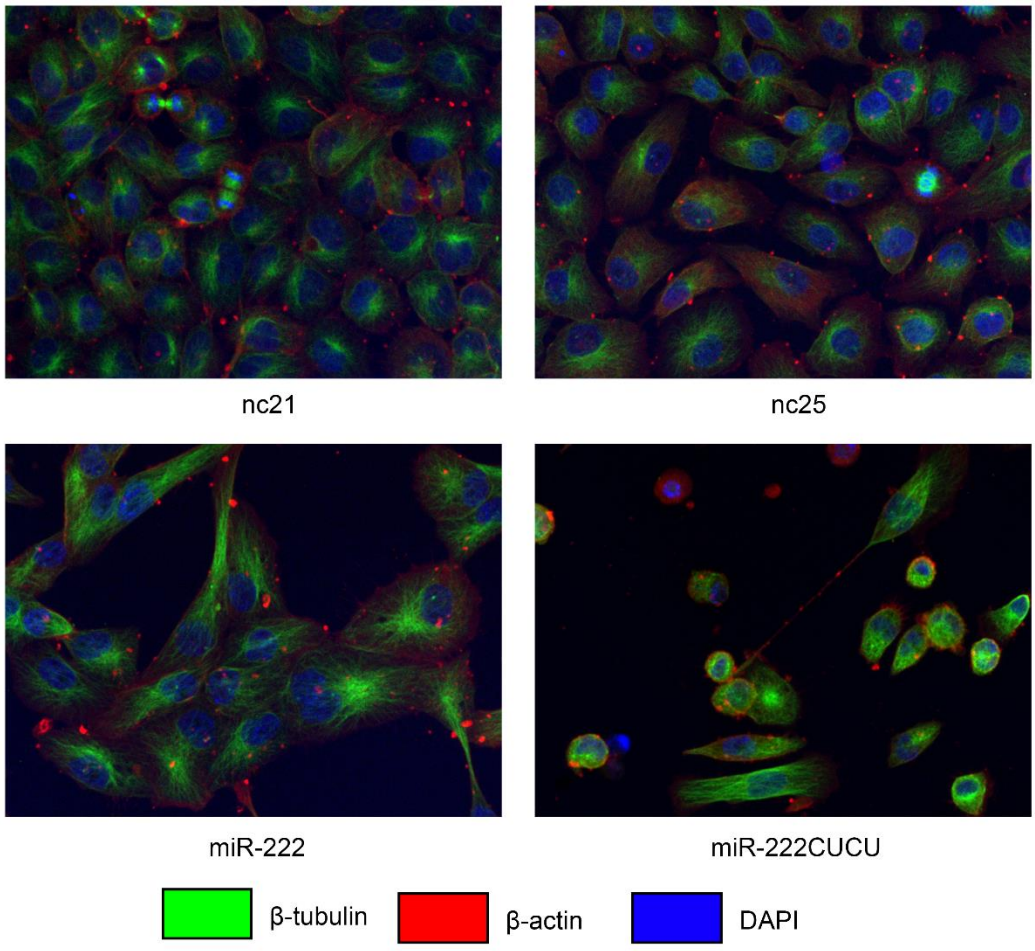


Figure 2: miR-222 and its isoforms alter cell morphology.

(A) MCF10A cells were photographed after transfection of miR-222, miR-222CU, miR-222CUC and miR-222CUCU respectively, compared with 21 and 25nt scramble controls (nc21 and nc25). (B) MCF10A cells are photographed after miR-222GAGA transfection. (C) Photos of KGN, MB468, MCF7 and MB231 cells are shown post-transfection of miR-222, miR-222CUCU and their respective controls.



Supplementary Figure 1:

Cell are visualized by immunofluorescence of β-tubulin (green), β-actin (red) post-transfection of miR-222 short and long forms, along with the 21 and 25nt controls (nc21 and nc25). DAPI (blue) was used to stain the DNA.

All four miR-222 species inhibit cell proliferation

The role of miR-222 remains controversial as some researchers report tumour suppressive roles [21-27], whereas others reported that it is oncogenic [28-31]. A possible reason is that different isoforms of miR-222 may be used between different studies by researchers previously unaware of functional differences. This may be likely as studies rarely report the sequences of transfected microRNAs and only recently have 3' variations been reported to confer functional consequence [18, 19, 32-34]. Notably, miR-222 has been repeatedly reported to inhibit cell proliferation [22, 24, 25]. To further investigate the biological function of miR-222 isoforms, all four variants (21, 23, 24 and 25nt) were transfected into MCF10A cells and proliferation assessed by confluence (Incucyte), DNA-based cell count (CyQUANT assay) and NAD(P)H-based proliferation assay (MTS assay) (Figure 3A, B, Supplementary Fig. 2A, Supplementary Video 1). All miR-222 isomiRs effectively suppressed proliferation as indicated by lower confluence and cell counts, with miR-222CUCU having the strongest effects, most likely a result of the

commencement of apoptosis (Figure 2A). Similarly, both miR-222 and miR-222CUCU repressed the speed of wound closure, with the longer isomiR again exerting stronger effects (Supplementary Fig. 2B). We also examined MCF10A cell cycle status using flow cytometry, however no significant difference was observed (data not shown). Consistent with a generally anti-proliferative role for miR-222, both miR-222 and 222CUCU (but not nc21 and nc25), induced the mRNA and protein expression of CDKN1A (p21), a universal inhibitor of cyclin kinases and mediator of cell cycle delay [35, 36] (Figure 3E). Coupled with previous data, this indicates all isoforms of miR-222 play anti-proliferative roles in MCF10A cells, with an apoptotic activity additional to progressively longer forms.

Supplementary Video 1 (attached as a standalone video file):

Incucyte videos were taken of MCF10A cells to visualize biological impacts of 222 isomiRs on cell proliferation and morphology, compared to 21 and 25nt scramble controls (nc21 and nc25). The filming initiated at 16 hours post-transfection, and this time point was shown as "day 0" on the video's scale bar.

miRNAs function through the post-transcriptional suppression of target mRNAs, typically interacting with the 3'UTR of targets via their 5' seed sequence. Motifs located near the 3' region however have been found to confer target selection in some circumstances [37] and 3' adenine addition reportedly reduces the effectiveness of miRNA targeting [38]. To formally test whether 3'UTR extension alters targeting potential, 3'UTRs of the top 25 predicted miR-222 targets were cloned downstream of luciferase reporters and the potential for differential targeting investigated with co-transfection of either miR-222 (21nt) or the CUC addition (24nt). Consistent with the capacity of the CUCU-GAGA mutation to still drive apoptosis (Figure 2B), we see no differential targeting capacity between miR-222 isoforms indicating differential effects are not likely mediated through a discrepancy between the general targeting capacities of the isomiRs (Figure 3 D). Combined with the fact that longer isomiRs are still capable of binding AGO (Figure 1C), these data suggest differential mechanisms are unlikely to result from basic properties such as stability, AGO-loading or generalised target selection.

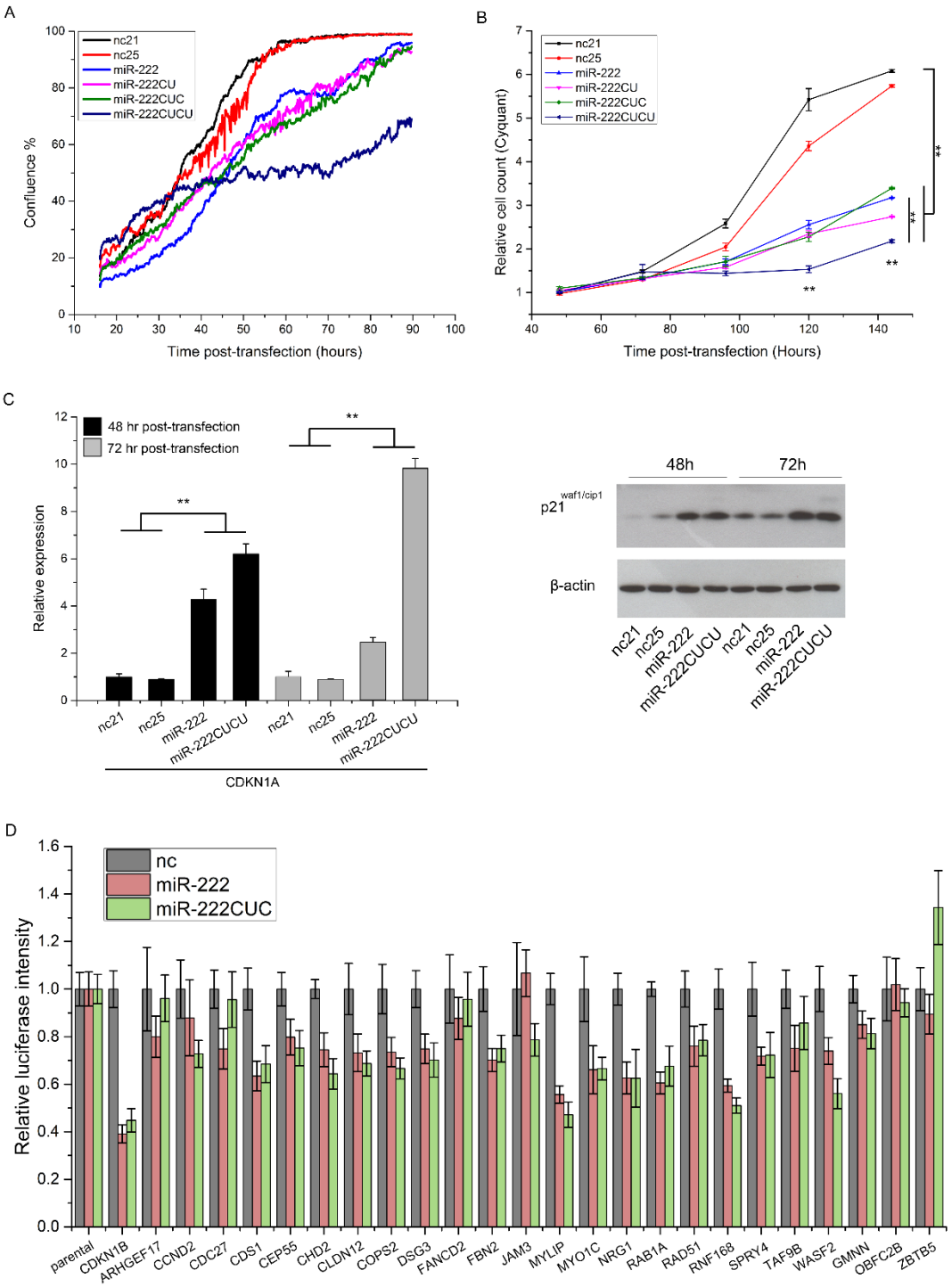
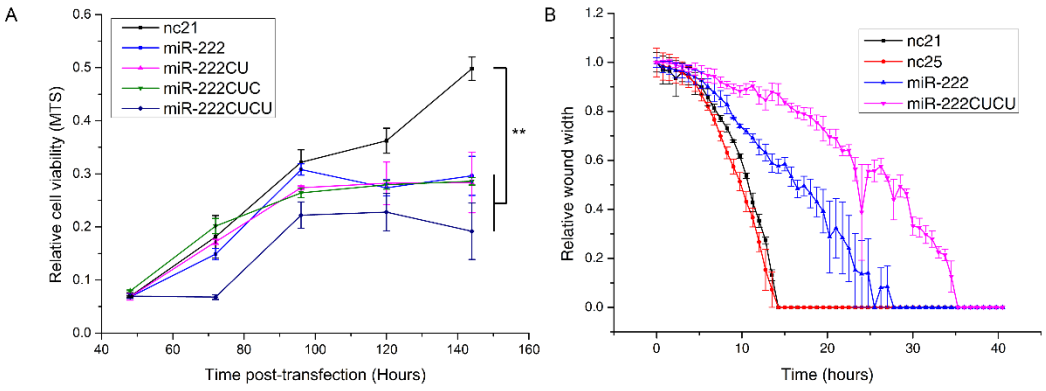


Figure 3: miR-222 and its isomiRs inhibit proliferation and share the same target recognition potential.

(A) The confluence of MCF10A cells was monitored to evaluate the influence of 222 isomiRs on cell proliferation, compared to 21 and 25nt scramble controls (nc21 and nc25). (B) The numbers of MCF10A cells were measured by the CyQUANT proliferation kit to evaluate the influence of 222 isomiRs on cell proliferation, compared to 21 and 25nt controls (nc21 and nc25). (C) The expression of CDKN1A (p21) was determined at both the mRNA and protein levels, 48 and 72 hr post-transfection of miR-222, miR-222CUCU and the negative controls. (D) The relative luciferase activity was measured of 25 most potential targets of the miR-222 seed sequence with the highest scores according to TargetScan, following transfection of scramble negative control RNA (nc) and miR-222 long and short isoforms. The parental vector was used as a negative control. ** $p<0.05$.



Supplementary Figure 2:

(A) MTS cell proliferation assay was performed to measure the MCF10A cell viability post-transfection of miR-222 short and long forms, along with the 21 and 25nt controls (nc21 and nc25). (B) Wound closure assay was performed under the same transfection conditions.

Longer miR-222 isoforms are associated with increased apoptotic activity

To further investigate the cell death phenotype which was observed with longer miR-222 isoforms, we performed flow cytometry to evaluate the proportion of unviable cells after transfection. Consistent with morphological observations, progressively longer miR-222 isomiRs displayed a progressively stronger tendency to promote cell death, from 17% by miR-222, to nearly 42% by miR-222CUCU (Figure 4A). Interestingly, the synthetic 25nt miR-222 isoform, miR-222GAGA (Figure 2B), exhibited an identical apoptotic activity to that of miR-222CUCU, which again suggested that the apoptosis was due to the length but not sequence of the 3′ extension. In further confirmation of apoptotic cell death, 72 hours post-transfection, miR-222CUCU-transfected cells displayed cleavage of both Caspase 3 and PARP, activation of the apoptosis-promoter Puma (BBC3) and downregulation of the apoptosis inhibitor Bcl-xL (BCL2L1) (Figure 4B). Further, both apoptosis promoters Puma

and Noxa were also upregulated at the RNA level (Figure 4C). Collectively, these observations indicate the triggering of the intrinsic apoptosis pathway [39] upon expression of longer miR-222 isomiRs.

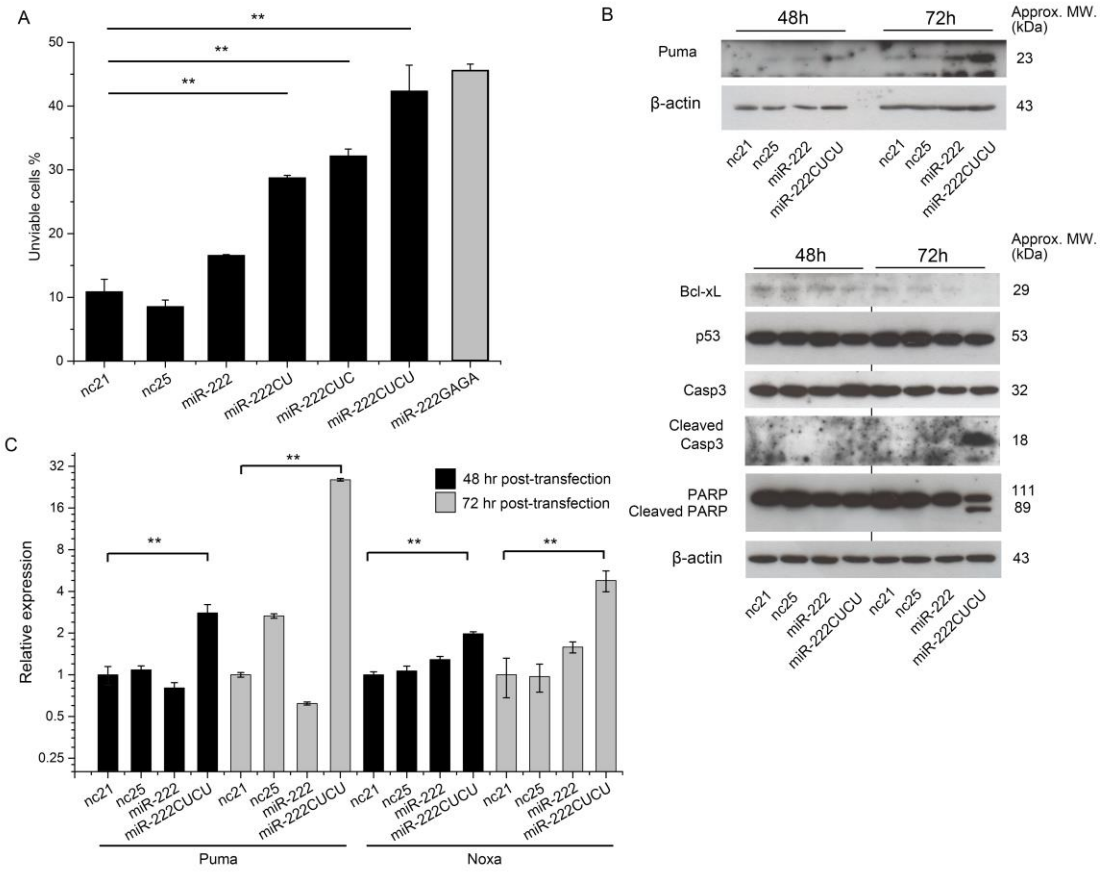


Figure 4: The longer miR-222 isomiRs induce apoptosis.

(A) The proportion of unviable cells were determined after transfection of the miR-222 isoforms varying from 21 to 25nt, along with the 21and 25nt scramble controls (nc21 and nc25). (B) Immunoblots are shown to demonstrate the expression of Puma (BBC3), Bcl-xL (BCL2L1), p53 (TP53), and caspase 3 and PARP cleavage following transfection of miR-222 short and long forms, along with the 21 and 25nt controls. (C) RT-PCR reveals the mRNA expression of PUMA and NOXA following transfection of miR-222 short and long forms, along with the 21 and 25nt controls. All experiments in this figure used MCF10A cells. ** $p<0.05$.

The differential apoptotic activity of longer 222 isomiRs is not attributable to interferon responses

Longer double-stranded (ds-) RNAs are known to trigger an interferon immune response, evolved as a defensive mechanism against viral infection [41]. Recently, longer double-stranded miRNA mimics were reported to inhibit cell proliferation and induce apoptosis independent of the seed sequence through the activation of an interferon response [42]. Throughout our work, the inclusion of a size-matched 25nt ds-negative control (nc25) fails to elicit anti-proliferative or pro-apoptotic responses which argues against a non-specific interferon response as a mechanism to explain the activities of miR-222CUCU. To further address this, we performed RT-PCR of 21 key IFN-responsive genes and showed no activation after the transfection of any miR-222 isoform. A double-stranded positive control (siBlunt27) which promotes an interferon response through direct interaction with RIG-1 (retinoic acid-inducible gene 1) [43, 44] successfully promoted the expression of the majority of these genes (Figure 5A). In addition, analysis of the expression of 131 interferon-responsive genes [45] from our deep-sequencing data indicated very little difference between the transfection of miR-222 or miR-222CUCU, again arguing against global interferon activation as an explanatory mechanism for miR-222CUCU effects (Figure 5B).

The cell physiological effects we report for miR-222 were re-produced using double-stranded microRNA mimics from 2 separate sources (Ambion and GenePharma). We also synthesised single-stranded (ss-) miRNA mimics to offer an alternative overexpression model which again re-capitulated previous observations, indicating reduced proliferation by all miR-222 isoforms and the promotion of apoptosis by miR-222CUCU but not canonical miR-222 (Figures 5C). It is worth noting that transfection of higher levels of single stranded mimics are required due to their known poor efficiency of AGO-loading and hence, targeting activity [40].

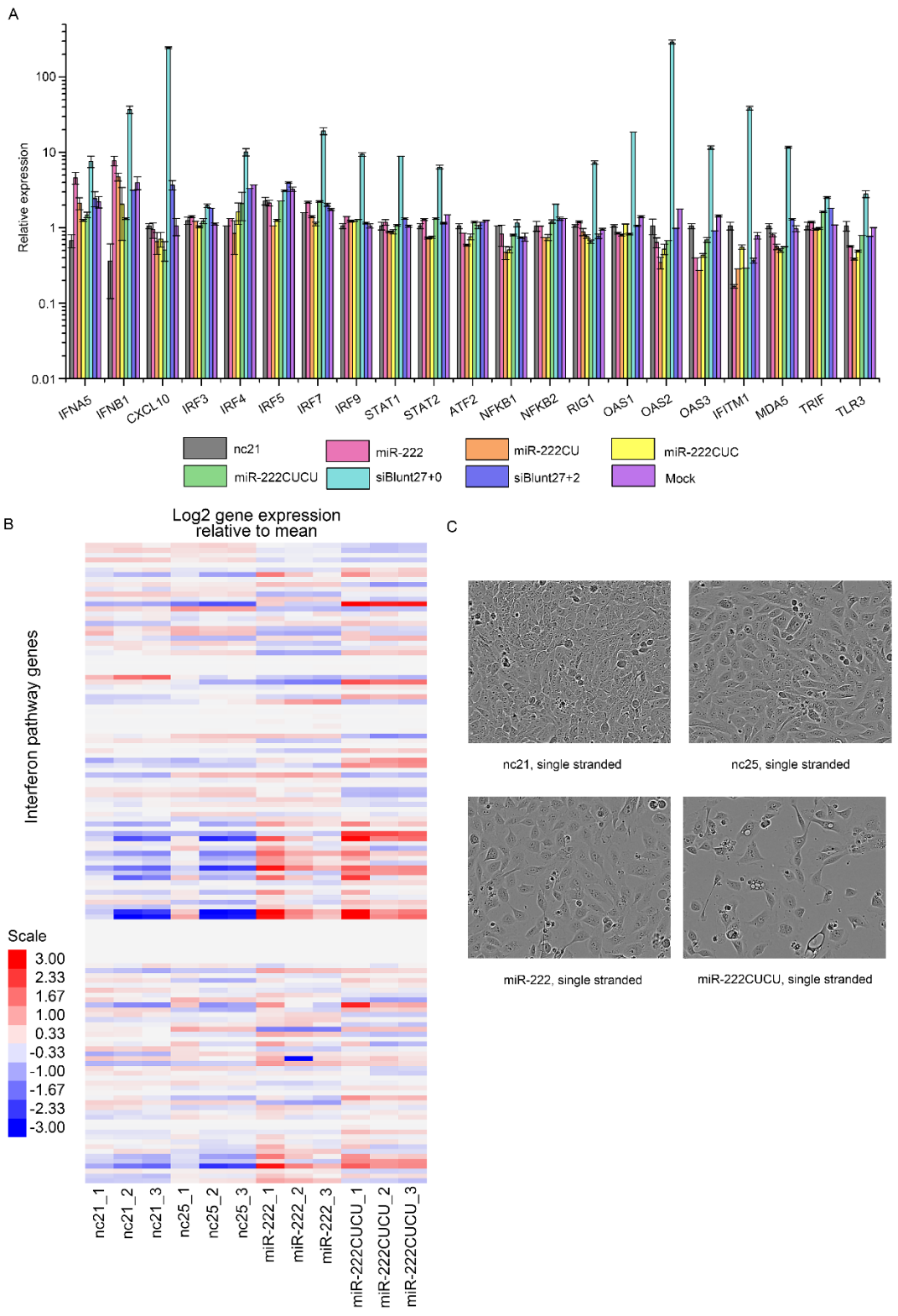

A
Relative expression
100
10
1
0.1
0.01
IFNA5
IFNB1
CXCL10
IRF3
IRF4
IRF5
IRF7
IRF9
STAT1
STAT2
ATF2
NFKB1
NFKB2
RIG1
OAS1
OAS2
OAS3
IFITM1
MDA5
TRIF
TLR3
nc21
miR-222
miR-222CU
miR-222CUC
miR-222CUCU
siBlunt27+0
siBlunt27+2
Mock
B
Log2 gene expression relative to mean
Interferon pathway genes
Scale
3.00
2.33
1.67
1.00
0.33
-0.33
-1.00
-1.67
-2.33
-3.00
nc21_1
nc21_2
nc21_3
nc25_1
nc25_2
nc25_3
miR-222_1
miR-222_2
miR-222_3
miR-222CUCU_1
miR-222CUCU_2
miR-222CUCU_3
C
nc21, single stranded
nc25, single stranded
miR-222, single stranded
miR-222CUCU, single stranded

Figure 5: Length dependent double-stranded miRNA IFN response was not the primary reason for apoptosis induced by longer 222 isoforms.

(A) The miR-222 forms varying from 21 to 25nt were transfected and a series of IFN (interferon) pathway markers were investigated to compare with the 21nt scramble control (nc21), IFN positive control (siBlunt27+0), IFN negative control (siBlunt27+2) and mock transfection. (B) RNA deep-sequencing was performed on MCF10A cells transfected by miR-222, miR-222CUCU, and negative control RNAs of 21 and 25nt (nc21 and nc25). The expression 131 genes that are IFN-responsive were visualized in a heat map. (C) Single-stranded miR-222 and miR-222CUCU mimics were transfected along with 21nt and 25nt scramble controls (nc21 and nc25), and photos are shown to characterize the cell morphology. ** $p<0.05$.

PIK3R3 (p55) is involved in miR-222CUCU-induced apoptosis

Analysis of deep sequencing data from miR-222 or control transfected cells indicated changes in the expression of large numbers of genes within our pairwise comparisons (Figure 6A). In order to identify genes and pathways through which miR-222CUCU might be exerting pro-apoptotic functions, we analysed these gene lists using IMPaLA (Integrated Molecular Pathway Level Analysis) [46], looking for genetic networks that are both consistently

modulated by miR-222CUCU in comparison to miR-222, nt21 or nc25, and which are also not modulated between the pairwise comparisons of miR-222, nc21 and nc25. We found the PI3K-Akt, a key survival pathway [47], to be one of the most strongly represented pathways altered by miR-222CUCU ($p=6.65x10^{-6}$) (Supplementary Table 1). RT-PCR and western blotting confirmed many key components (including PIK3CA, PIK3CB, PIK3R3, AKT1, AKT2 and AKT3) were repressed at mRNA and protein levels upon miR-222CUCU transfection (Figure 6B, C), which verifies global disruption of this vital survival pathway. However, this is different from an IFN response, since the IFNs are commonly found to regulate Akt phosphorylation, rather than total Akt protein expression [48, 49].

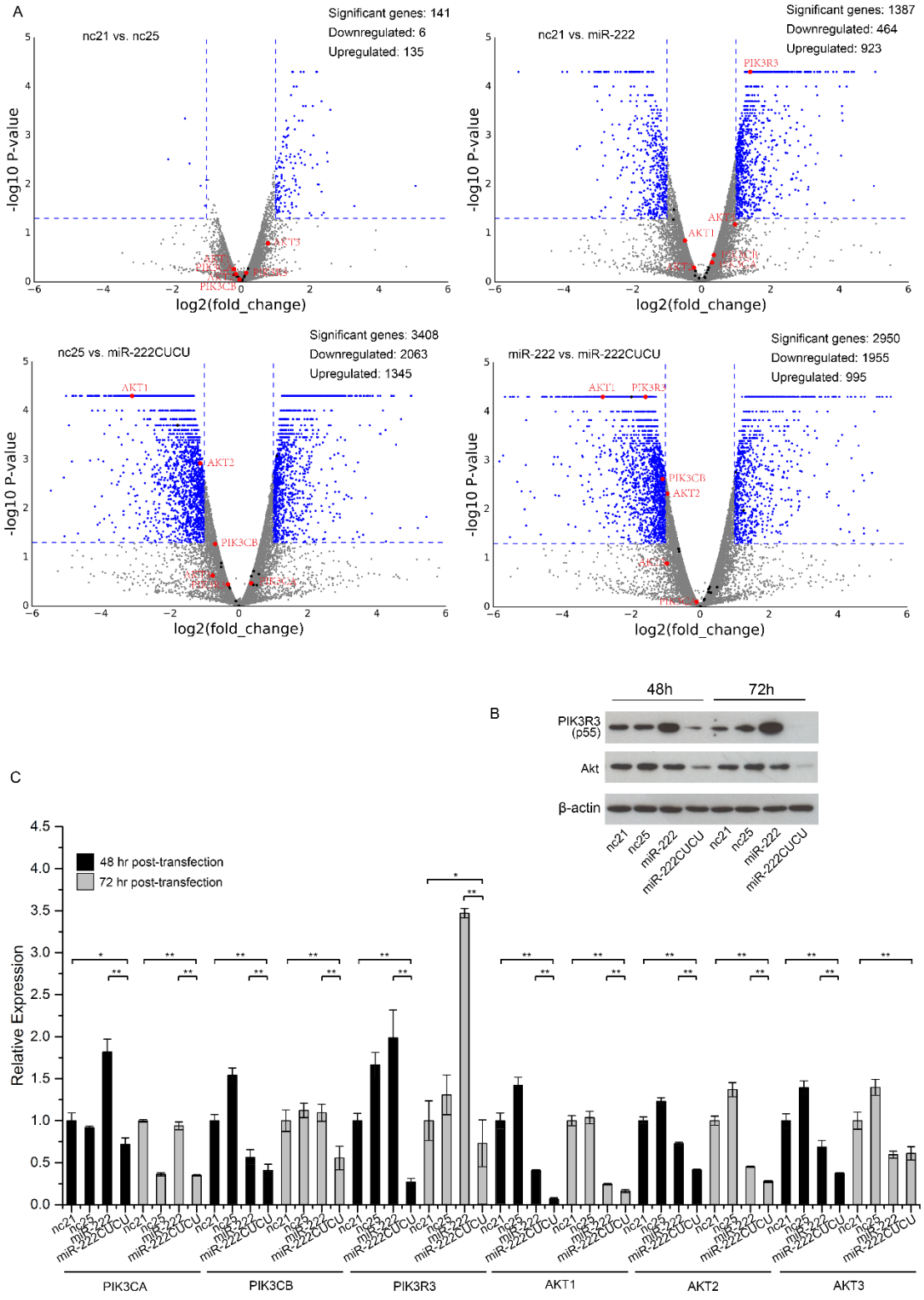


Figure 6: The expression of key PIK3-AKT genes is disrupted.

(A) Volcano plots were generated based on our deep-sequencing result. The expression data was plotted in comparison in 4 ways: 21nt control (nc21) vs. 25nt control (nc25), nc21 vs. miR-222, nc25 vs. miR-222CUCU, and miR-222 vs. miR-222CUCU. A cohort of housekeeping genes was marked as black dots, while red dots represent genes of interest in this research (for gene lists, see the Methods section). The criteria of genes being plotted was $p < 1$, whereas the statistical cut-off was $p > 0.05$ and fold change > 2 or <0.5. (B) The PIK3R3 (p55) and Akt protein expression was visualized by immunoblots, following transfection of miR-222, miR-222CUCU and negative control RNAs of 21 and 25nt (nc21 and nc25). (C) The expression of key PI3K-AKT members was measured by RT-PCR following the same transfection settings. * $p<0.1$, ** $p<0.05$.

Among these genes, PIK3R3 (p55) was of particular interest as it is functionally upstream of Akt and is known to be upregulated in breast, liver, prostate, ovarian, gastric cancers and glioblastoma [50-52]. Furthermore, high PIK3R3 expression promotes migration and invasion [50-52] whilst PIK3R3 knockdown represses proliferation, migration, invasion, and importantly triggers apoptosis [50-53]. Therefore, we designed a siRNA targeting PIK3R3 to evaluate whether it could reproduce the phenotypic effects of miR-222CUCU. Transfection of siPIK3R3, which successfully inhibited PIK3R3 mRNA and protein expression (Figure 7A, B), phenocopied miR-222CUCU by four aspects:

(i) siPIK3R3 inhibited the expression of key PI3K-Akt genes, i.e., PIK3CA, PIK3CB, AKT1, 2 and 3 (Figure 7A, B); (ii) The MCF10A cells displayed a blebbing morphology indicative of apoptosis (Figure 7C); (iii) Cell confluence was significantly reduced after siPIK3R3 treatment (Figure 7D); and (iv) siPIK3R3 induced cell death to a similar extent as miR-222CUCU overexpression (Figure 7E). Together, these data suggest PIK3R3 plays a vital role in miR-222CUCU-mediated apoptosis.

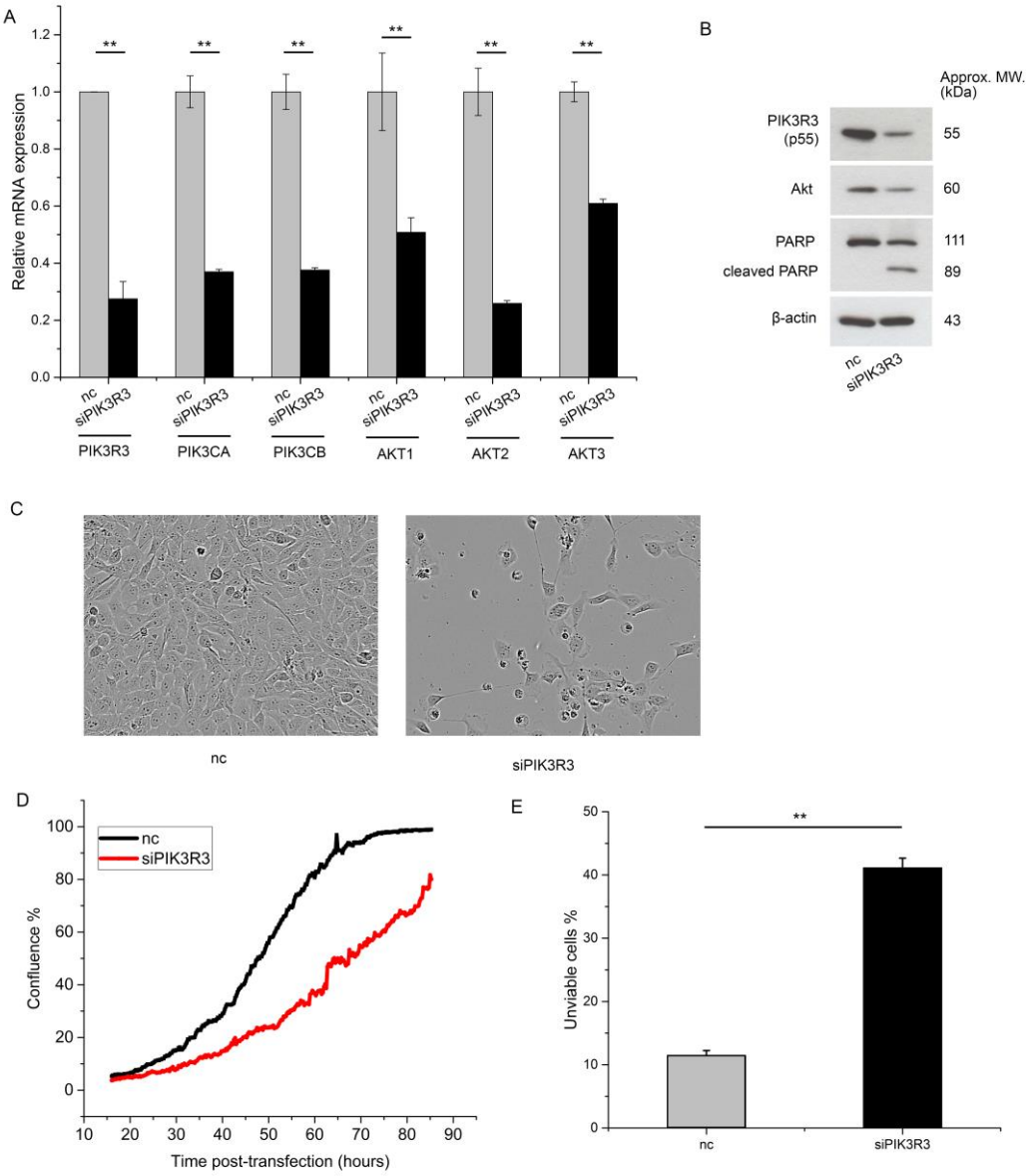


Figure 7: PIK3R3 is a vital mediator of miR-222CUCU-induced apoptosis.

(A, B) The expression of PIK3R3 and key PI3K-AKT members were measured by RT-PCR and Western blot after transfection of the PIK3R3 siRNA, compared with a negative control RNA (nc). (C) MCF10A cells were photographed following the same transfection settings to demonstrate the cell morphology.

(D) The cell confluence was monitored following the same transfection conditions. (E) Cell viability was determined using flow cytometry to illustrate the percentage of living cells following the same transfection settings.

Pathway name	Overlapping genes	All genes	p value	q value
Metabolism	383	1427	1.32E-12	5.15E-09
Nuclear Receptors Meta-Pathway	99	307	9.37E-08	0.000183
Non-integrin membrane-ECM interactions	22	42	0.00000233	0.00304
EGFR1	126	447	0.00000665	0.0065
Metabolism of lipids and lipoproteins	136	497	0.0000151	0.0118
Peroxisome - Homo sapiens	32	81	0.0000301	0.0144
Diseases of glycosylation	70	226	0.0000318	0.0144
Metabolism of carbohydrates	77	256	0.0000406	0.0144
Myoclonic epilepsy of Lafora	77	256	0.0000406	0.0144
Glycogen storage diseases	77	256	0.0000406	0.0144
TGF Beta Signaling Pathway	24	55	4.43E-05	0.0144
TGF Beta Signaling Pathway	24	55	4.43E-05	0.0144
Laminin interactions	13	22	5.42E-05	0.0152
Pre-NOTCH Processing in Golgi	9	12	5.46E-05	0.0152
Glucocorticoid Receptor Pathway	27	67	8.24E-05	0.0215

Supplementary Table 1:

The 15 most significantly altered pathways are listed according to the IMPaLA analysis of our deep-sequencing of cells treated by miR-222CUCU compared to miR-222.

Discussion

With the advent of deep sequencing, high isomiR abundance is now recognised, though few studies address their functional significance, especially when the variation is not at the 5′ end.

By analysing data human cancer data (TCGA), we confirmed both that the prevalence of 3′ isomiRs far outweigh 5′ isomiRs, and that the extent of variation (number of variable nucleotides) per miRNA is greater at the 3′ end. This is in agreement with previously published research [11, 32, 54, 55], e.g., 296 out of 445 miRNAs show 3′ variation, whereas only 39 show 5′ changes in a series of human lymphoblastoid cell lines [56]. The same authors also pointed out that (as we also find) the 5′ ends show a narrower range (+/- 1nt) of modifications compared to the 3′ ends (+/- 3nt). 3′ isomiRs are prevalent across all species [20, 38, 57, 58] and individual miRNAs have been reported in various contexts to display particularly high 3′ variability (i.e.: miR-222, miR-282 and

miR-312 in *Drosophila)* [12]. Similarly, miR-222 3′ variation is also abundant in HeLa cells [59]. Alternative processing of Dicer and Drosha are considered as a major source of miRNA heterogeneity [60-63], which is related to the secondary structures of the pri- and pre-miRNAs [14, 61]. RNA editing [64, 65], trimming [66, 67] and nucleotidyl transfer [11, 38, 68, 69] are also vital contributing factors. In the specific case of miR-222, it is unclear whether 3′ variability is a result of variable Drosha cleavage or the variable "nibbling" activities of exonucleases. Regardless of the mechanism by which isomiRs are generated however, the natural abundance of 3′ variants suggest a need for careful consideration of any functional impacts. This is especially the case as (i) the most expressed version is not always the canonical "standard" form, and (ii) many researchers do not consider the potential for differential effects and often do not note the sequence of the miRNA mimics they use. Furthermore, the varieties and expression patterns of isomiRs suggest functional and evolutionary relationships [32, 33].

Our interest was drawn to miR-222 due to the high degree of natural 3′ variability in our AGO-immunoprecipitated sequencing samples and was further stimulated by the dramatic (and different) morphological impacts after miRNA mimic transfection in multiple cell lines. Of particular interest is the controversial nature of the role of miR-222. It has been repeatedly reported to be a tumour suppressive miRNA in various oncogenic models: prostate cancer

[21, 25], non-small cell lung cancer [22], breast cancer [23], primary effusion lymphoma and Kaposi sarcoma [26]. On the other hand, however, multiple reports also indicate oncogenic roles [28-31].

One possible explanation might lie in the differential activities of miR-222 isoforms, to which most authors are probably unaware. As a result, most publications listed "miR-222" in general without specifying the isomiR sequences being used. We investigated cell confluence, proliferation and wound healing post-transfection of miR-222 and its isoforms, and demonstrated that many of the functions overlap. The ability of progressively longer isomiRs to progressively drive apoptosis, however, indicates vital differences. The luciferase assay confirmed that the longer isoform did not exert a global difference in 3′-UTR target recognition in comparison with miR-222 as all targets we tested were seed-based and this target recognition motif (9-15 nt) remains intact despite the 3′ extension. Whilst rare examples exist for target specification involving the 3′ end, the fact that the apoptosis-promoting effect of miR-222 remains intact despite mutation of the extended nucleotides (Figure 2B) strongly argue for the fact that standard canonical targeting is an unlikely explanation for differential roles. Consistent luciferase effects further imply generalised properties such as stability and AGO-loading are equally unlikely candidates for mechanistic differences.

Importantly, we indicate that the apoptosis promotion of miR-222CUCU is independent of an interferon response or p53 engagement. According to the previous reports, miR-222 targets PUMA [70], which could be seen from our results as miR-222 repressed PUMA mRNA (Figure 4C). In contrast, miR-222CUCU dramatically induces PUMA at both the mRNA and protein levels (Figure 4B, C), suggesting an overwhelming signalling mechanism in the opposite direction. Similar cases are present; e.g., adenosine-to-inosine editing of the oncogenic miR-455-5p abolishes its inhibitory regulation of CPEB1, repressing melanoma growth and metastasis *in vivo* [71].

Whilst we are yet to fully characterise the mechanism responsible, it is interesting to note that the key survival signalling pathway, PI3K-Akt, was the pathway we noted to be most consistently different between cells transfected with miR-222CUCU. Among the multiple PI3K-Akt factors affected, we note among them PIK3R3 is an important upstream factor known to modulate the activity of PIK3CA and Akt. It is also active at downstream of miR-193a-3p, -5p and miR-511, mediating cancer metastasis and cell growth [72, 73]. It has been less explored compared to the other PIK3/Akt members, but still plays vital roles to promote proliferation and survival, since knockdown of PIK3R3 represses proliferation and migration and induces apoptosis in cell lines [50-

53]. Silencing of PIK3R3 also attenuates Rb1 (Retinoblastoma 1) phosphorylation and decreases the expression of both cyclin D1 and PCNA [51]. We observed siPIK3R3-related suppression of PIK3CA, PIK3CB and AKT expression, which could explain the apoptosis and inhibition of cell confluence [74-76]. Therefore, PIK3R3 may be one of the vital mediators of the disruption of PIK3-Akt pathway by miR-222CUCU.

In conclusion, this research demonstrates the prevalent existence of miRNA variants and experimentally demonstrates functional differences between isoforms, thereby representing one of very few investigations into the significance of 3′ variance. This work therefore adds to our understanding of miR-222, provides a possible explanation for the mutually exclusive roles reported for miR-222 in literature and reveals a further layer of miRNA regulation which must now also be considered.

Materials and Methods

Cell lines and reagents

MCF10A (breast epithelial), MDA-MB-468, MCF7, T47D (breast cancer) and KGN (granulosa) cells were purchased from the ATCC (American Type Culture Collection, Manassas, VA USA), and all cell lines have been under regular mycoplasma monitoring every 1-2 months using a MycoAlert mycoplasma detection kit (Lonza, Melbourne, VIC Australia), ensuring that all cell lines being used for experiments were free from mycoplasma contamination.

MCF10A cells were cultured in DMEM/F12 medium (Sigma-Aldrich, St. Louis, USA) supplemented with 5% horse serum (Invitrogen, Grand Island, NY USA), 20ng/mL EGF, 0.5μL/ml Hydrocortisone, 100ng/ml Cholera toxin, 10μg/mL Insulin (Sigma-Aldrich). MB231, MB468 and KGN cell lines were cultured in DMEM (Sigma-Aldrich), whereas MCF7 and T47D were cultured in RPMI-1640 (Sigma-Aldrich), both supplemented with 10% (v/v) fetal calf serum (Invitrogen).

All abovementioned media were supplemented with 2 mmol/L L-Glutamine, 100 IU/mL penicillin and 100 μg/ml streptomycin, and all cell lines were cultured at 37 °C in a humidified atmosphere at 5% CO2.

miRNA and siRNA overexpression and inhibition

For miRNA overexpression studies, approximately 2.5 × 10^5 cells/well were seeded in 6-well plates or 0.5 x 10^5 cells/well in 24-well plates depending on the scale of cell numbers needed. Reverse-transfections was performed using miRVana (Ambion, Grand Island, NY USA) and GenePharma (GenePharma, Shanghai, China) miRNA mimics, paired with negative control RNAs from the same manufacturers. The transfection was performed using Lipofectamine RNAiMAX reagent (Invitrogen) following the manufacturer's instructions. The miRNA transfected was 50nM and cells were harvested 72 hr post-transfection, unless otherwise specified. The single-stranded miRNA mimics and paired negative controls were obtained from GenePharma and transfected at 200 nM following the same procedure.

The efficiency of miRNA transfection was not examined, because according to Thompson et al., perhaps the only accurate way to estimate the outcomes of miRNA mimic transfection is Argonaute HITS-CLIP [82], a method which is extremely time- and labour- consuming. The same authors pointed out that a simple RT-PCR after transient transfection of miRNA mimics does not accurately determine their actual effects. Therefore, measuring transfection efficiency is beyond the scale of the current research.

2’-Oxymethelated siRNAs and paired negative controls were purchased from GenePharma, following the same transfection procedure at 50 nM. siPIK3R3: CAACCUCGUUUCCUUACAAAU.

Luciferase assay

For the luciferase assay, we cloned the proposed regions of genes into downstream of the Renilla luciferase gene in the psiCHECK2 dual-luciferase vector (Promega). The Firefly luciferase gene expressed from the same vector from an HSV-TK promoter was used as an internal reference. Approximately 0.5×10^5 MCF10A cells were plated in each well of 24-well plates. After 24 hr, in each well we co-transfected 0.4 ng/μL of luciferase constructs with 10 nM of either miRNA mimics or negative control RNA (ncRNA) (Ambion), using Lipofectamine 2000 (Invitrogen) following the manufacturer’s instructions. Cells were harvested 72 hr post-transfection, and Renilla and Firefly luminescence were measured by a GloMax 20/20 Luminometer (Promega) following the manufacturer’s instructions. A ratio of Renilla/Firefly luminescence intensity was used to evaluate the relative luciferase expression activity.

Immunoblotting and antibodies

Cells were rinsed with PBS and lysed in lysis buffer (50 mmol/L Tris-HCl, pH 7.5, 250 mmol/L NaCl, 1% Triton X-100 and 1x protease inhibitors (Roche, South San Francisco, CA) on ice for 30 min. Insoluble components of cell lysates were removed by centrifugation for 10 min at 4°C, 12,000 g, and protein concentration was measured using a bicinchoninic acid (BCA) protein assay kit (Pierce, Thermo Scientific, Radnor, PA USA). Protein extracts were resolved using SDS PAGE electrophoresis on 8-15% polyacrylamide gels, depending on the molecular weights of the target proteins, and the proteins were electro-transferred to a Hybond- C Extra nitrocellulose membrane (GE Healthcare Life Sciences, Pittsburgh, PA USA). The membrane was then blocked for 30 min in blocking buffer which was made by dissolving 5% skim milk in TBST (50 mM Tris-HCl, 150 mM NaCl, 0.05% Tween 20, pH=7.6).

For quantification, the p53 protein was probed by Santa Cruz sc-126 at 1:1000 in the blocking buffer (Dallas, TX USA), $p21^{waf1/cip1}$ by Cell Signalling Technology (CST) #2947 at 1:1000 (Danvers MA), PI3K p55 by CST #11889 at

1:1000, Puma by ProSci #41232 at 1:500 (OneWorldLab), phospho-Akt by CST #4051 at 1:1000, Akt by CST #9272 at 1:1000, PARP by CST #9542 at 1:1000, Casp3 by CST #9662 at 1:1000, and BCL-xL by CST #2762 at 1:500. Equal loading was confirmed by blotting of β-actin (Sigma-Aldrich AC-15) or β-tubulin (Abcam ab6046), both at 1:2000. Chemiluminescent detection of protein was conducted using appropriate secondary antibodies conjugated with horseradish peroxidise (GE Healthcare) and the enhanced chemiluminescence kit (GE Healthcare) according to the manufacturer's instructions.

RNA extraction and real-time PCR

RNA (>250nt) extraction from cells was performed using a GenElute™ mRNA Miniprep Kit (Sigma) with on-column RNase-free DNase digestion (Sigma) according to the manufacturer's instructions. After extraction, RNA concentration was measured using a ND-1000 NanoDrop spectrometer (Thermo Scientific). Thereafter, 1μg of total RNA was reverse-transcribed into cDNA using Moloney Murine Leukaemia Virus (MMLV) Reverse Transcriptase (Promega) with random 6'mer primers (Promega) under the following the manufacturer's instructions. Real-time PCR reaction was performed on a CFX Connect Real-Time PCR Detection System (Bio-Rad) using a QuantiTect SYBR Green PCR Kit (QIAGEN). The RT-PCR program was: 95°C

for 3 min, then start cycles consisting 95°C for 10s and 57°C 60s (depending on the primers used) for 40 cycles. After the reaction was complete, the C_T values were determined using automated threshold settings.

The total RNA including small RNAs was extracted using the Trizol Reagent (Invitrogen) following the manufacturer's instructions. After extraction, total RNA concentration was determined using a ND-1000 NanoDrop spectrometer (Thermo Scientific). For quantification of small RNAs, TaqMan gene expression assays were obtained from Applied Biosystems (Grand Island, NY USA) and reverse transcription was performed using a TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems) using specific Reverse Transcription stem-loop primers provided with the TaqMan RNA assays following the manufacturer's instructions. Real-time PCR reaction was performed on the CFX Connect Real-Time PCR Detection System (Bio-Rad) using specific TaqMan RNA assays following the manufacturer's instructions. In this study, 18s rRNA was chosen as a housekeeping gene for the normalization of RNA expression.

The primers used for PCR are listed in Supplementary Table 2.

Primer ID	Sequence
PIK3R3-F	CGTACCATTCTGATTGAGGCT
PIK3R3-R	AAAAGCTGCTCTCCATCTCG
18s rRNA-F	GTAACCCGTTGAACCCCATT
18s rRNA-R	CCATCCAATCGGTAGTAGCG
PIK3CA	AGGCCACTGTGGTTGAATTGGGA
PIK3CA	CAGTGCACCTTTCAAGCCGCC
AKT1-F	TTATGGTAAAGGCAGGCGGC
AKT1-R	AACGTATACACCAGGCCAGC
AKT2-F	AGCTAGGTGACAGCGTACCA
AKT2-R	TGTACCCAATGAAGGAGCCG
AKT3-F	ATTTCTCCAAGTTGGGGGCTC
AKT3-R	CTCCCCTCTTCTGAACCCAA
CDKN1A-F	TGGACCTGGAGACTCTCAGGGTCG
CDKN1A-R	TTAGGGCTTCCTCTTGGAGAAGATC
PUMA-F	TGGGGGAGCTAAGAGTGTGT
PUMA-R	CCTTCCAGTGCCTAGTGTGG
NOXA-F	AGAGCTGGAAGTCGAGTGT
NOXA-R	GCACCTTCACATTCCTCTC

Supplementary Table 2:

The primers being used in this research are listed.

Incucyte scanning and cell proliferation assay

For Incucyte (Essen BioScience, by Millennium Science, Mulgrave VIC Australia) scans, the cells were reverse-transfected as abovementioned in 24 or 96-well plates which were loaded on the Incucyte for scanning over 72 hr. For proliferation assay, cells of the same number were reverse-transfected and cell numbers at different time points were measured using a CyQUANT proliferation kit (Invitrogen) or the CellTiter 96 Aqueous MTS cell proliferation assay kit (Promega) according to the manufacturer's instructions.

Cell cycle assay, viability assay and flow cytometry

The cell cycle assay was performed using PI (propidium iodide) staining and flow cytometry as previously described [77]. Briefly, cells were harvested, fixed in ice-cold ethanol (70%), and incubated overnight at 4°C. Cells were stained with propidium iodide solution (50 μg/mL; Sigma–Aldrich) and 100 μg/mL RNase A (Sigma–Aldrich) for 45 minutes at 37°C. DNA content was determined with the use of a FACS Calibur flow cytometer with cell cycle profiles analysed using FLOWJO software.

The cell viability assay was performed using 7AAD (7-aminoactinomycin D) staining as previously described [78]. Cells were harvested 72 hr post-treatment,

centrifuged at 1,300 × g, washed in phosphate-buffered saline (PBS), and stained with 7AAD solution (2 μg/mL) (7-amino-actinomycin-D, Invitrogen, Carlsbad, CA) for 10 min at RT (room temperature). Viable cells were determined using the same flow cytometry instrument set as abovementioned.

RNA Deep sequencing and HITS-CLIP (High throughput sequencing of crosslinked and immunoprecipitated RNA)

The RNA Deep sequencing and HITS-CLIP were performed as previously described [79]. Briefly, cell pellets were lysed and treated with RNAse A/T1 (Ambion) for 8 minutes and subjected to immunoprecipitation for 2 hr with a pan-anti-AGO antibody (2A8)) conjugated to protein-A dynabeads (Invitrogen) using bridging rabbit anti-mouse IgG (Jackson ImmunoResearch Laboratories, West Grove, PA USA). The beads were ligated with 3′ RNA linker (CAGACGACGAGCGGG) labelled with P32 and the 5′ RNA linker (AGGGAGGACGAUGCGGxxxG, with "X" representing different nucleotides for barcoding). After separation on acrylamide gels, PCR amplification and purification, the samples were subjected to Illumina sequencing (Illuminia, San Diego, CA USA). Using an in house Perl script, reads were filtered for average quality and for homopolymeric tracts exceeding 12 nt, trimmed of linker

sequence fragments and separated by barcode. The bowtie program [80] was used to align resulting 17 to 30 nt reads to the human genome.

A cohort of housekeeping genes was used to evaluate the normalization of the deep-sequencing results (marked as black dots), including ACTB (β-actin), GAPDH, TUBA (α-tubulin), TUBB (β-tubulin), and also a previously reported collection of C1orf43, CHMP2A, EMC7, GPI, PSMB2, PSMB4, RAB7A, REEP5, SNRPD3, VCP and VPS29 [81]. Red dots represent genes of interest in this research (PIK3R3, PIK3CA, PIK3CB, AKT1, AKT2 and AKT3). The criteria of genes being plotted on Figure 6A was $p < 1$ and expression > 0, whereas the statistical cut-off was $p > 0.05$ and fold change > 2 or <0.5.

Immunofluorescence

MCF10A cells attached to Thermanox coverslips (Nalge Nunc International, Naperville, IL) were fixed in 4% PFA (paraformaldehyde) for 15 min, and were blocked in 2% BSA (bovine serum albumin) for 30 min. The primary antibodies for β-actin and β-tubulin were the same as those listed in the Western blot section, both at a concentration of 1:500 and were incubated with the cells at 4°C overnight. Then the cells were labelled by fluorescent-tagged secondary

antibodies, goat anti-Mouse Alexa Fluor 488 IgG, and goat anti-rabbit Alexa Fluor 594 IgG (Invitrogen), both at 1:500. Eventually the coverslips were mounted in Vectasheld Antifade Mounting Medium with DAPI (4',6-Diamidino-2-Phenylindole, Dihydrochloride) (Vector Laboratories, Burlingame, CA) and visualized under a laser confocal microscope.

Statistical analysis

Results are given as mean of at least three independent experiments ± standard deviation. Student's t-test was performed using replicate values to indicate significance. Values of $p < 0.05$ were considered statistically significant (as labelled with ** in figures), while values of $p < 0.1$ were indicated by *.

References

1. Pasquinelli, A.E., MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. Nat Rev Genet, 2012. 13(4): p. 271-82.
2. Ha, M. and V.N. Kim, Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol, 2014. 15(8): p. 509-24.
3. Krol, J., I. Loedige, and W. Filipowicz, The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet, 2010. 11(9): p. 597-610.
4. Neilsen, C.T., G.J. Goodall, and C.P. Bracken, IsomiRs--the overlooked repertoire in the dynamic microRNAome. Trends Genet, 2012. 28(11): p. 544-9.
5. Starega-Roslan, J., et al., Sequence features of Drosha and Dicer cleavage sites affect the complexity of isomiRs. Int J Mol Sci, 2015. 16(4): p. 8110-27.
6. Griffiths-Jones, S., miRBase: microRNA sequences and annotation. Curr Protoc Bioinformatics, 2010. Chapter 12: p. Unit 12 9 1-10.
7. Benson, D.A., et al., GenBank. Nucleic Acids Research, 2015. 43(D1): p. D30-D35.

8. Vaz, C., et al., Analysis of the microRNA transcriptome and expression of different isomiRs in human peripheral blood mononuclear cells. BMC Res Notes, 2013. 6: p. 390.

9. Zhang, H., et al., Comparisons of isomiR patterns and classification performance using the rank-based MANOVA and 10-fold cross-validation. Gene, 2015. 569(1): p. 21-26.

10. Loher, P., E.R. Londin, and I. Rigoutsos, IsomiR expression profiles in human lymphoblastoid cell lines exhibit population and gender dependencies. Oncotarget, 2014. 5(18): p. 8790-802.

11. Wyman, S.K., et al., Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. Genome Research, 2011. 21(9): p. 1450-1461.

12. Fernandez-Valverde, S.L., R.J. Taft, and J.S. Mattick, Dynamic isomiR regulation in Drosophila development. Rna-a Publication of the Rna Society, 2010. 16(10): p. 1881-1888.

13. Newman, M.A., V. Mani, and S.M. Hammond, Deep sequencing of microRNA precursors reveals extensive 3 ' end modification. Rna-a Publication of the Rna Society, 2011. 17(10): p. 1795-1803.

14. Etebari, K., et al., Dengue virus infection alters post-transcriptional modification of microRNAs in the mosquito vector Aedes aegypti. Scientific Reports, 2015. 5.

15. Wu, X., et al., Comprehensive expression analysis of miRNA in breast cancer at the miRNA and isomiR levels. Gene, 2015. 557(2): p. 195-200.

16. James, A.M., et al., Seven Encapsulated and Nonencapsulated (iso)microRNAs in Human Samples Predict Coronary Atherosclerosis and Cardiac Events. Faseb Journal, 2015. 29.

17. Babapoor, S., et al., A Novel miR-451a isomiR, Associated with Amelanotypic Phenotype, Acts as a Tumor Suppressor in Melanoma by Retarding Cell Migration and Invasion. Plos One, 2014. 9(9).

18. Manzano, M., et al., Divergent target recognition by coexpressed 5 '-isomiRs of miR-142-3p and selective viral mimicry. Rna, 2015. 21(9): p. 1606-1620.

19. Tan, G.C., et al., 5 ' isomiR variation is of functional and evolutionary importance. Nucleic Acids Research, 2014. 42(14): p. 9424-9435.

20. Tan, G.C. and N. Dibb, IsomiRs have functional importance. Malaysian Journal of Pathology, 2015. 37(2): p. 73-81.

21. Goto, Y., et al., MicroRNA expression signature of castration-resistant prostate cancer: the microRNA-221/222 cluster functions as a tumour suppressor and disease progression marker. British Journal of Cancer, 2015. 113(7): p. 1055-1065.

22. Yamashita, R., et al., Growth inhibitory effects of miR-221 and miR-222 in non-small cell lung cancer cells. Cancer Medicine, 2015. 4(4): p. 551-564.

23. Gan, R., et al., Downregulation of miR-221/222 enhances sensitivity of breast cancer cells to tamoxifen through upregulation of TIMP3. Cancer Gene Therapy, 2014. 21(7): p. 290-296.

24. Felli, N., et al., MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation. Proceedings of the National Academy of Sciences of the United States of America, 2005. 102(50): p. 18081-18086.

25. Fuse, M., et al., Tumor suppressive microRNAs (miR-222 and miR-31) regulate molecular pathways based on microRNA expression signature in prostate cancer. J Hum Genet, 2012. 57(11): p. 691-9.

26. O'Hara, A.J., et al., Tumor suppressor microRNAs are underrepresented in primary effusion lymphoma and Kaposi sarcoma. Blood, 2009. 113(23): p. 5938-41.

27. Medina, R., et al., MicroRNAs 221 and 222 bypass quiescence and compromise cell survival. Cancer Res, 2008. 68(8): p. 2773-80.

28. Zhao, J.J., et al., MicroRNA-221/222 Negatively Regulates Estrogen Receptor alpha and Is Associated with Tamoxifen Resistance in Breast Cancer. Journal of Biological Chemistry, 2008. 283(45): p. 31079-31086.

29. Visone, R., et al., MicroRNAs (miR)-221 and miR-222, both overexpressed in human thyroid papillary carcinomas, regulate p27Kip1 protein levels and cell cycle. Endocr Relat Cancer, 2007. 14(3): p. 791-8.

30. Wurz, K., et al., MiR-221 and MiR-222 Alterations in Sporadic Ovarian Carcinoma: Relationship to CDKN1B, CDKN1C and Overall Survival. Genes Chromosomes & Cancer, 2010. 49(7): p. 577-584.

31. Wang, M., et al., Role and mechanism of miR-222 in arsenic-transformed cells for inducing tumor growth. Oncotarget, 2016.

32. Guo, L., et al., A Comprehensive Survey of miRNA Repertoire and 3 ' Addition Events in the Placentas of Patients with Pre-Eclampsia from High-Throughput Sequencing. Plos One, 2011. 6(6).

33. Guo, L., et al., Consistent isomiR expression patterns and 3' addition events in miRNA gene clusters and families implicate functional and evolutionary relationships. Molecular Biology Reports, 2012. 39(6): p. 6699-6706.

34. Llorens, F., et al., A highly expressed miR-101 isomiR is a functional silencing small RNA. Bmc Genomics, 2013. 14.

35. Nayak, K.B., et al., EVI1 targets Delta Np63 and upregulates the cyclin dependent kinase inhibitor p21 independent of p53 to delay cell cycle progression and cell proliferation in colon cancer cells. International Journal of Biochemistry & Cell Biology, 2013. 45(8): p. 1568-1576.

36. Fecteau, J.F., et al., Lenalidomide inhibits the proliferation of CLL cells via a cereblon/p21(WAF1/Cip1)-dependent mechanism independent of functional p53. Blood, 2014. 124(10): p. 1637-1644.

37. Helwak, A., et al., Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding. Cell, 2013. 153(3): p. 654-665.

38. Burroughs, A.M., et al., A comprehensive survey of 3 ' animal miRNA modification events and a possible role for 3 ' adenylation in modulating miRNA targeting effectiveness. Genome Research, 2010. 20(10): p. 1398-1410.

39. Inoue, S., et al., Ordering of caspases in cells undergoing apoptosis by the intrinsic pathway. Cell Death and Differentiation, 2009. 16(7): p. 1053-1061.

40. Chorn, G., et al., Single-stranded microRNA mimics. Rna-a Publication of the Rna Society, 2012. 18(10): p. 1796-1804.

41. Karpala, A.J., T.J. Doran, and A.G. Bean, Immune responses to dsRNA: implications for gene silencing technologies. Immunol Cell Biol, 2005. 83(3): p. 211-6.

42. Goldgraben, M.A., et al., Double-stranded microRNA mimics can induce length- and passenger strand-dependent effects in a cell type-specific manner. RNA, 2016. 22(2): p. 193-203.

43. Ishibashi, O., et al., Short RNA Duplexes Elicit RIG-I-Mediated Apoptosis in a Cell Type-and Length-Dependent Manner. Science Signaling, 2011. 4(198).

44. Marques, J.T., et al., A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. Nature Biotechnology, 2006. 24(5): p. 559-565.

45. Liu, S.Y., et al., Systematic identification of type I and type II interferon-induced antiviral factors. Proceedings of the National Academy of Sciences of the United States of America, 2012. 109(11): p. 4239-4244.

46. Kamburov, A., et al., Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. Bioinformatics, 2011. 27(20): p. 2917-2918.

47. Franke, T.F., et al., PI3K/Akt and apoptosis: size matters. Oncogene, 2003. 22(56): p. 8983-8998.

48. Su, X., et al., Interferon-gamma regulates cellular metabolism and mRNA translation to potentiate macrophage activation. Nat Immunol, 2015. 16(8): p. 838-49.

49. Zhu, L., et al., TSC1 controls macrophage polarization to prevent inflammatory disease. Nat Commun, 2014. 5: p. 4696.

50. Wang, G.H., et al., PIK3R3 Induces Epithelial-to-Mesenchymal Transition and Promotes Metastasis in Colorectal Cancer. Molecular Cancer Therapeutics, 2014. 13(7): p. 1837-1847.

51. Zhou, J., et al., Genetic and bioinformatic analyses of the expression and function of PI3K regulatory subunit PIK3R3 in an Asian patient gastric cancer library. Bmc Medical Genomics, 2012. 5.

52. Zhang, L., et al., Integrative genomic analysis of phosphatidylinositol 3 '-kinase family identifies PIK3R3 as a potential therapeutic target in epithelial ovarian cancer. Clinical Cancer Research, 2007. 13(18): p. 5314-5321.

53. Wang, G.H., et al., TGF- regulates the proliferation of lung adenocarcinoma cells by inhibiting PIK3R3 expression. Molecular Carcinogenesis, 2015. 54: p. E162-E171.

54. Lee, L.W., et al., Complexity of the microRNA repertoire revealed by next-generation sequencing. Rna-a Publication of the Rna Society, 2010. 16(11): p. 2170-2180.

55. Morin, R.D., et al., Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res, 2008. 18(4): p. 610-21.

56. Loher, P., E.R. Londin, and I. Rigoutsos, IsomiR expression profiles in human lymphoblastoid cell lines exhibit population and gender dependencies. Oncotarget, 2014. 5(18): p. 8790-8802.

57. Wang, S.Q., et al., Entropy-Based Model for MiRNA Isoform Analysis. Plos One, 2015. 10(3).

58. Sablok, G., et al., isomiRex: Web-based identification of microRNAs, isomiR variations and differential expression using next-generation sequencing datasets. Febs Letters, 2013. 587(16): p. 2629-2634.

59. Llorens, F., et al., Microarray and deep sequencing cross-platform analysis of the mirRNome and isomiR variation in response to epidermal growth factor. Bmc Genomics, 2013. 14.

60. Kuchenbauer, F., et al., In-depth characterization of the microRNA transcriptome in a leukemia progression model. Genome Research, 2008. 18(11): p. 1787-1797.

61. Starega-Roslan, J., et al., Structural basis of microRNA length variety. Nucleic Acids Research, 2011. 39(1): p. 257-268.

62. Zhou, H.J., et al., Deep annotation of mouse iso-miR and iso-moR variation. Nucleic Acids Research, 2012. 40(13): p. 5864-5875.

63. Wu, H.Q., et al., Alternative Processing of Primary microRNA Transcripts by Drosha Generates 5 ' End Variation of Mature microRNA. Plos One, 2009. 4(10).

64. Nishikura, K., Functions and Regulation of RNA Editing by ADAR Deaminases. Annual Review of Biochemistry, Vol 79, 2010. 79: p. 321-349.

65. Kawahara, Y., et al., Frequency and fate of microRNA editing in human brain. Nucleic Acids Research, 2008. 36(16): p. 5270-5280.

66. Liu, N., et al., The Exoribonuclease Nibbler Controls 3 ' End Processing of MicroRNAs in Drosophila. Current Biology, 2011. 21(22): p. 1888-1893.

67. Han, B.W., et al., The 3 '-to-5 ' Exoribonuclease Nibbler Shapes the 3 ' Ends of MicroRNAs Bound to Drosophila Argonaute1. Current Biology, 2011. 21(22): p. 1878-1887.

68. Martin, G. and W. Keller, RNA-specific ribonucleotidyl transferases. Rna-a Publication of the Rna Society, 2007. 13(11): p. 1834-1849.

69. Katoh, T., et al., Selective stabilization of mammalian microRNAs by 3 ' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes & Development, 2009. 23(4): p. 433-438.

70. Zhang, C.Z., et al., MiR-221 and miR-222 target PUMA to induce cell survival in glioblastoma. Molecular Cancer, 2010. 9.

71. Shoshan, E., et al., Reduced adenosine-to-inosine miR-455-5p editing promotes melanoma growth and metastasis. Nature Cell Biology, 2015. 17(3): p. 311-+.

72. Yu, T., et al., MicroRNA-193a-3p and -5p suppress the metastasis of human non-small-cell lung cancer by downregulating the ERBB4/PIK3R3/mTOR/S6K2 signaling pathway. Oncogene, 2015. 34(4): p. 413-423.

73. Cao, G., et al., MiR-511 inhibits growth and metastasis of human hepatocellular carcinoma cells by targeting PIK3R3. Tumor Biology, 2015. 36(6): p. 4453-4459.

74. Toker, A. and M. Yoeli-Lerner, Akt signaling and cancer: Surviving but not moving on. Cancer Research, 2006. 66(8): p. 3963-3966.

75. Yoeli-Lerner, M. and A. Toker, Akt/PKB signaling in cancer - A function in cell motility and invasion. Cell Cycle, 2006. 5(6): p. 603-605.

76. Guerreiro, A.S., et al., Targeting the PI3K p110 alpha Isoform Inhibits Medulloblastoma Proliferation, Chemoresistance, and Migration. Clinical Cancer Research, 2008. 14(21): p. 6761-6769.

77. Kumar, R., et al., FBXO31 is the chromosome 16q24.3 senescence gene, a candidate breast tumor suppressor, and a component of an SCF complex. Cancer Research, 2005. 65(24): p. 11304-11313.

78. Neilsen, P.M., et al., New 26S Proteasome Inhibitors with High Selectivity for Chymotrypsin-Like Activity and p53-Dependent Cytotoxicity. Acs Chemical Biology, 2013. 8(2): p. 353-359.

79. Yu, F., et al., p53 Represses the Oncogenic Sno-MiR-28 Derived from a SnoRNA. Plos One, 2015. 10(6).

80. Langmead, B., et al., Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol, 2009. 10(3): p. R25.

81. Eisenberg, E. and E.Y. Levanon, Human housekeeping genes, revisited (vol 29, pg 569, 2013). Trends in Genetics, 2014. 30(3): p. 119-120.

82. Thomson DW, Bracken CP, Szubert JM, Goodall GJ. On Measuring miRNAs after Transient Transfection of Mimics or Antisense Inhibitors. Plos One. 2013;8(1).

Conclusion

The human genome project revealed that non-protein-coding RNAs (ncRNAs), rather than protein-coding genes, comprise a dominant proportion of the human genome [1]. Research continues to reveal and assign biological functions to these ncRNAs, including miRNAs with which this project is concerned. Here, I extend our understanding of miRNA origin and function, both characterising new miRNAs from non-canonical sources, and identifying previously unknown functions for naturally occurring miRNA variants.

snoRNAs were traditionally known to function in ribonucleoprotein (RNP) complexes, guiding the modification and processing of non-coding RNAs. Recently however, a small number of papers have begun to assign another function as a potential source of miRNAs that result from the further processing of snoRNAs [2-5]. Although relatively abundant within deep sequencing, we know little regarding how many of these snoRNA-derived smaller RNAs have functional roles, either as miRNAs or in any other capacity. This research supports snoRNAs as an alternative source of miRNAs that can play vital regulatory roles. Our microarray screening revealed six C/D box snoRNAs (small nucleolar RNAs) that are repressed by p53, the master tumour suppressor, through transcriptional repression of their host gene. p53 is mostly

well-known for transcriptional activation, but recent reports have also revealed that p53 is capable of mediating transcriptional repression, which is much less understood than transcriptional activation [6]. Briefly, depending on the DNA binding domain of p53, its transcriptional repression is achieved by not only directly binding to gene promoters, but also direct interference with the basal transcription machinery, and alteration of chromatin structure [7]. The exact mechanism how p53 regulates the promoter of SNHG1 is not yet well understood, and promoter transcription factor analysis for SNHG1 can be a further study.

What was surprising was that one of the snoRNAs, SNORD28, was processed into a functional miRNA-like transcript that has the properties of a standard miRNA, adding to a small but growing number of examples whereby functional miRNAs are derived from non-canonical biogenesis pathways. snoRNAs are most localized in the nucleus. How the snoRNAs are incorporated to the miRNA maturation process and where the product miRNAs are localized can be a further study.

This work therefore defines a novel origin for a biologically significant miRNA, which has broader implications as small RNA fragments have been found originating not only from snoRNAs, but also rRNAs, tRNAs, small nuclear

RNAs (snRNAs) and other non-coding RNAs [8-11], revealing a much more complex network of ncRNA species than originally envisaged.

The second part of the thesis again focuses characterising on functional roles for something we and others again find abundant in deep-sequencing data: that naturally existing miRNAs exist as a range of sequence variants (called isomiRs). 5′ variations are reported to alter miRNA regulatory function as this site of variation exists near to the target recognition motif (known as the "seed"). However, 3′ heterogeneity is the dominating population of miRNA isoforms, though it is largely thought to be functionally neutral because it does not affect the seed region. However, new discoveries have been made, highlighting that the miRNA sequences outside the seed region are also essential components in target interaction. Helwak et.al. employed a CLASH (crosslinking, ligation, and sequencing of hybrids) approach to investigate miRNAs interacting with mRNAs bound to the Argonaute (AGO) proteins and discovered target-pairing motifs contributing to miRNA target interaction [12], may be affected by 3′ variation. In Chapter 3 of this thesis, miR-222 isoforms with 3′ extensions exhibit distinctive apoptotic roles in comparison to the canonical miR-222, but share proliferation-inhibiting functions. Notably, the 3′ miR-222 isoforms can be immunoprecipitated from the Argonaute (AGO) proteins and share identical 3′-UTR target recognition properties with the conventional miR-222, exhibiting identical target interaction properties as conventional miRNAs. Our

discovery of their role in the regulation of expression of PIK3R3 and other PIK3-Akt genes reveal a substantial part of the molecular mechanism of these long miR-222 isoforms. These findings showed that the long miR-222 isoform has its own regulatory roles in addition to those of miR-222.

However, the exact molecular mechanism how miR-222CUCU regulates PIK3R3 is still not so clear. According to several miRNA target programs (MicroT4, miRanda, miRbridge, miRDB, miRMap, miRNAMap, Pictar2 and TargetScan), PIK3R3 does not have any canonical 3'-UTR sites for miR-222 that is strong enough to cause the level of PIK3R3 expression changes reported in our experiments. In addition, EISA (Exon-Intron Split Analysis) of deep-sequencing data from our miR-222CUCU model revealed that PIK3R3 is regulated at the transcriptional level, yet no valid miR-222 binding elements could be located on the promoter of PIK3R3. Therefore, the miR-222CUCU regulation of PIK3R3 may be indirect through other genes. Some recent discoveries showed that the Argonaute proteins and miRNAs exist in the mammalian nucleus [13, 14] and may be capable of regulating gene expression by interaction with promoter regions [15, 16]. This will be an interesting area to explore in the future.

This research indicates that miRNA isoforms, especially the abundant 3' variants, can have major biological roles that were poorly characterized

previously. Since the expression of isomiRs is prevalent yet dynamic in various tissues and cell lines, implications are raised for previously published works where the roles and function of the 3′ isomiRs have not been fully appreciated.

Nevertheless, snoRNAs and miRNAs are only part of the ncRNA landscape that includes long-established structural RNAs (rRNAs, tRNAs etc.) and newer "hot areas" such as long non-coding RNAs (lncRNAs) [17-20], enhancer RNAs (eRNAs) [21-24] and circular RNAs (circRNA) [25-27]. There are numerous non-protein-coding RNA transcripts remaining to be characterized [28], and no-doubt much is remaining to be learned from those transcripts already characterised.

Overall, ncRNAs are emerging as a vital layer of functional transcripts, with complex regulatory patterns that play essential roles in cell physiology. As the research on ncRNAs advance, it is increasingly evident that their characteristics and functions are highly complex. From this point of view, extensive opportunities and challenges are both present in future ncRNA research, involving areas including development, cancer progression, metabolism, virus infection, and numerous other life activities. Ultimately, discoveries about miRNAs may contribute to the designing of novel drugs and therapies to combat cancer and other diseases.

References

1. Ponting, C.P., P.L. Oliver, and W. Reik, Evolution and functions of long noncoding RNAs. Cell, 2009. 136(4): p. 629-41.
2. Ono, M., et al., Analysis of Human Small Nucleolar RNAs (snoRNA) and the Development of snoRNA Modulator of Gene Expression Vectors. Molecular Biology of the Cell, 2010. 21(9): p. 1569-1584.
3. Taft, R.J., et al., Small RNAs derived from snoRNAs. RNA, 2009. 15(7): p. 1233-40.
4. Brameier, M., et al., Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. Nucleic Acids Res, 2011. 39(2): p. 675-86.
5. Ono, M., et al., Identification of human miRNA precursors that resemble box C/D snoRNAs. Nucleic Acids Res, 2011. 39(9): p. 3879-91.
6. Beckerman, R. and C. Prives, Transcriptional regulation by p53. Cold Spring Harb Perspect Biol, 2010. 2(8): p. a000935.
7. Ho, J. and S. Benchimol, Transcriptional repression mediated by the p53 tumour suppressor. Cell Death Differ, 2003. 10(4): p. 404-8.
8. Burroughs, A.M., et al., Deep-sequencing of human argonaute-associated small RNAs provides insight into miRNA sorting and reveals argonaute association with RNA fragments of diverse origin. Rna Biology, 2011. 8(1): p. 158-177.

9. Zywicki, M., K. Bakowska-Zywicka, and N. Polacek, Revealing stable processing products from ribosome-associated small RNAs by deep-sequencing data analysis. Nucleic Acids Research, 2012. 40(9): p. 4013-4024.

10. Kawaji, H., et al., Hidden layers of human small RNAs. Bmc Genomics, 2008. 9.

11. Li, Z.H., et al., Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. Nucleic Acids Research, 2012. 40(14): p. 6787-6799.

12. Helwak, A., et al., Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding. Cell, 2013. 153(3): p. 654-665.

13. Hwang, H.W., E.A. Wentzel, and J.T. Mendell, A hexanucleotide element directs microRNA nuclear import. Science, 2007. 315(5808): p. 97-100.

14. Marcon, E., et al., miRNA and piRNA localization in the male mammalian meiotic nucleus. Chromosome Research, 2008. 16(2): p. 243-260.

15. Huang, V. and L.C. Li, Demystifying the nuclear function of Argonaute proteins. Rna Biology, 2014. 11(1): p. 18-24.

16. Smalheiser, N.R. and O.L.A. Gomes, Mammalian Argonaute-DNA binding? Biology Direct, 2014. 9.

17. Derrien, T., et al., The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res, 2012. 22(9): p. 1775-89.

18. Guttman, M., et al., Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature, 2009. 458(7235): p. 223-7.

19. Khalil, A.M., et al., Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A, 2009. 106(28): p. 11667-72.

20. Idogawa, M., et al., Identification and analysis of large intergenic non-coding RNAs regulated by p53 family members through a genome-wide analysis of p53-binding sites. Hum Mol Genet, 2014. 23(11): p. 2847-57.

21. Huarte, M., et al., A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell, 2010. 142(3): p. 409-19.

22. Hung, T., et al., Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet, 2011. 43(7): p. 621-9.

23. Liu, Q., et al., LncRNA loc285194 is a p53-regulated tumor suppressor. Nucleic Acids Res, 2013. 41(9): p. 4976-87.

24. Melo, C.A., et al., eRNAs are required for p53-dependent enhancer activity and gene transcription. Mol Cell, 2013. 49(3): p. 524-35.

25. Lasda, E. and R. Parker, Circular RNAs: diversity of form and function. Rna-a Publication of the Rna Society, 2014. 20(12): p. 1829-1842.

26. Li, Z.Y., et al., Exon-intron circular RNAs regulate transcription in the nucleus. Nature Structural & Molecular Biology, 2015. 22(3): p. 256-264.

27. Jeck, W.R. and N.E. Sharpless, Detecting and characterizing circular RNAs. Nature Biotechnology, 2014. 32(5): p. 453-461.

28. Costa, F.F., Non-coding RNAs: new players in eukaryotic biology. Gene, 2005. 357(2): p. 83-94.