

Insight into the function of microRNAs and other small RNAs of diverse origin

Daniel Wyville Thomson

BBiotech (Hons) Flinders University

Thesis submitted for the award of the degree of

Doctor of Philosophy

School of Medicine

Faculty of Health Sciences

University of Adelaide

Gene Regulation Laboratory

Centre for Cancer Biology

IMVS, SA Pathology

2014

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma 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 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 catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Daniel Thomson

Date

Table of Contents

Declaration	v
Table of Contents	vii
Abstract	xi
Thesis layout.....	xiii
Preamble	xv
Acknowledgements	xvii
Abbreviations	xxi
CHAPTER 1.....	1
1 Introduction	2
1.1 MicroRNAs: A Historical Perspective.....	2
1.2 Canonical microRNA biosynthesis.....	3
1.3 MicroRNA targeting	5
1.4 MicroRNA mechanism of action.....	5
1.5 MiRNA target networks.....	7
1.6 Argonaute Proteins	8
1.7 Non-canonical miRNA biosynthesis.....	10
1.8 The RNA Degradome.....	10
References.....	16
CHAPTER 2.....	21
2 Experimental Strategies for microRNA Target Identification.....	23
2.1 Abstract.....	24
2.2 Introduction	24
2.3 Gene-specific experimental validation of miRNA targets	27
2.4 Experimental miRNA target screening techniques	28
2.5 Caveats of miRNA over-expression	29
2.6 Gene Expression Analysis.....	30
2.7 Immunoprecipitation of RISC components	31

2.8	HITS-CLIP.....	33
2.9	PAR-CLIP	34
2.10	Biotin tagged miRNA.....	36
2.11	Detection of direct cleavage targets.....	36
2.12	Reverse transcription of targets	38
2.13	Proteomic Approaches	38
2.13.1	SILAC	38
2.13.2	2D-DIGE	39
2.14	Translation profiling.....	40
2.15	Finding multiple miRNAs that target a single mRNA	41
2.16	Concluding Remarks	42
2.17	References	44
CHAPTER 3	55
3	On measuring miRNAs after transient transfection of mimics or antisense inhibitors	57
3.1	Preface: Abolishing a Common Pitfall- published examples of the inappropriate use of qPCR for microRNA mimic and inhibitor quantitation	58
3.2	Abstract	61
3.3	Introduction.....	62
3.4	Results and Discussion.....	62
3.5	Material and Methods.....	74
3.5.1	Quantitative PCR following miRNA mimic transfection or addition post lysis.....	74
3.5.2	Argonaute : miRNA immunoprecipitation	74
3.5.3	Bioinformatics.....	76
3.5.4	Fluorescent oligonucleotide transfection, LysoTracker transfection, and immunofluorescence microscopy	76
3.5.5	miRNA inhibitor doping experiment.....	77
3.6	Acknowledgements	78
3.7	Additional unpublished work.....	79
3.7.1	Measurement of microRNA mimic abundance by qPCR is not representative of microRNA functionality – microRNA mimic titration	79

3.7.2	Measurement of microRNA mimic abundance by qPCR is not representative of microRNA functionality – microRNA mimic time course.....	81
3.7.3	Methods: Additional unpublished work - Luciferase reporter assay with miRNA mimic transfection titration	83
3.8	References	84
CHAPTER 4.....		85
	Statement of Authorship.....	86
4	Assessing the gene regulatory properties of Argonaute-bound small RNAs of diverse genomic origin	89
4.1	Abstract.....	90
4.2	Introduction.....	91
4.3	Results.....	94
4.3.1	Small RNAs of diverse origin bind Argonaute.....	94
4.3.2	Assessing small RNAs from diverse sources for miRNA-like activity	97
4.3.3	Testing the sensitivity of luciferase reporter assays	101
4.3.4	AGO-binding is a better indicator of functionality than total RNA abundance	103
4.4	Discussion.....	107
4.5	Supplementary Figures	112
4.5.1	Supplementary Figure 4.1	112
4.5.2	Supplementary Figure 4.2	114
4.5.3	Supplementary Figure 4.3	115
4.5.4	Supplementary Figure 4.4	118
4.5.5	Supplementary Figure 4.5	119
4.5.6	Supplementary Figure 4.6	120
4.5.7	Supplementary Figure 4.7	121
4.6	Materials and Methods.....	123
4.6.1	Argonaute : miRNA immunoprecipitation.....	123
4.6.2	Reporter construct design and cloning	123
4.6.3	Luciferase reporter assay	123
4.6.4	Bioinformatic Analyses.....	124
4.7	Primers annealed for reporter cloning.....	124
4.8	REFERENCES	128
CHAPTER 5.....		135

5	Concluding Remarks	137
5.1	Summary of work	137
5.2	Final comments and future direction	139
	REFERENCES	141

Abstract

MicroRNAs are small non-coding RNA molecules which have become appreciated over the last decade as important regulators of gene expression. Functional microRNAs (miRNAs) are bound to Argonaute, the core protein of the RNA induced silencing complex (RISC). They control gene expression by guiding RISC to target mRNAs causing RNA degradation or translational repression. Recent advances in next-generation sequencing technology have greatly increased the pace of molecular research. Nowhere is this more evident than in the field of RNA and microRNA with over 34,000 being identified (miRBase, 2014). However, a challenge of the field is to follow this rate of discovery with functional validation, which is often laborious and expensive. The underlying aim of this thesis is to better understand the functionality of miRNAs and other small RNAs of endogenous and exogenous origin. Crucial to this aim was the application of Argonaute HITsCLIP (High throughput sequencing of cross-linking immunoprecipitation) as a high throughput method of assessing the function of microRNAs, which was supported with other genomic, molecular and cellular approaches.

Beginning with an introduction to the field of microRNAs (Chapter 1) and review of the experimental and genomic techniques available for this purpose (Chapter 2), Chapter 3 presents the use of Argonaute HITsCLIP to assess the functionality of microRNA mimic and inhibitors. The findings reveal that the majority of transfected RNA is not functionally bound to Argonaute but rather resides in non-functional pools within lysosomes and associated vesicles. This has important implications for the use of these common tools for miRNA over-expression or inhibition.

Chapter 4 aims to interrogate theories of non-canonical miRNA biogenesis. Using RNA-seq in MDA-MB-231 breast cancer cells we observed defined small RNAs on a transcriptome-wide scale including small RNAs from tRNAs, snoRNAs, vault RNAs, yRNAs, and rRNAs. Many of which are specifically Argonaute bound. To test the ability of these identified small RNAs to silence target RNAs we used reporter assays with perfectly antisense target sites cloned down stream of a luciferase reporter and antisense oligonucleotides for small RNA knockdown. From these analyses, we have concluded that the repertoire of small RNAs produced from hairpin secondary structures is not restricted to canonical miRNA precursors and many of these are specifically enriched in Argonaute immunoprecipitates. However, we emphasise that a miRNA's functionality is relative to its abundance within RISC, and the vast majority of non-canonical miRNAs such as those produced from snoRNAs are only comparable in abundance to lowly expressed miRNAs. Although we do identify novel non-canonical miRNAs which are of moderate abundance in these cells and have the ability to repress the targets at endogenous expression. We also present that there are highly abundant small RNAs which have been postulated to act as miRNAs, such as small RNAs from y3-RNA, Chr6.tRNA166, 5s-rRNAs and U3 snoRNAs but do not repress their target despite being as abundant as miRNAs which do.

In summary, this work provides insight to the function of small RNAs from diverse genomic origins as well as transfected siRNAs.

Thesis layout

The thesis presented herein contains a combination of 3 published articles, additional supporting work not included in the manuscripts, an introduction and discussion. Collectively, these works provide an insight into the biosynthesis and function of microRNAs and other small RNAs of diverse origin.

The thesis is organised into chapters where each chapter can be read as its own entity but fits into the narrative as a whole. Each chapter contains a relevant introduction, methods, results, discussion and references section where relevant. The chapter organisation of the thesis is as follows; Chapter 1 provides an introductory background into the field of microRNAs and other small RNAs. Chapter 2 presents a survey and summary of the experimental strategies for microRNA target identification (Thomson et al., 2011), of which many of the methods discussed have been employed in later chapters. Chapter 3 presents work assessing the functionality of transfected oligonucleotides as tools for microRNA over-expression and inhibition (Thomson et al., 2013), additional unpublished work is presented in support of this. Chapter 4 gives evidence of the biosynthesis of small RNAs from a range of constitutively expressed RNA precursors and presents an assessment of the functionality of these identified endogenous small RNAs (Thomson et al. 2014). Chapter 5 is a discussion of the work as a whole in the context of current genomics and RNA research.

Where work was produced in collaboration with other lab members, their contribution is acknowledged at the beginning of each section.

Preamble

At the time of commencing my PhD, advances in next-generation sequencing technologies were revealing new potential for genomics research. One seminal publication applied high-throughput sequencing of RNA co-immunoprecipitated with Argonaute, allowing the identification of miRNA targets on a genome wide scale (Chi et al., 2009). Prior to coming to the lab, the Goodall laboratory had developed interest in the miR-200 family of miRNAs after identifying their role in Epithelial to Mesenchymal Transition (EMT) (Gregory et al., 2008b). In work beginning during my honours year in the same laboratory, we aimed to identify new targets of the miR-200 family; I used miR-200 over-expression models with microarray analysis and began to develop an Argonaute immunoprecipitation technique. After seeing the success of the HITS-CLIP technique (Chi et al., 2009), the lab continued with this new technique in a project spearheaded by Dr Cameron Bracken with the intention of identifying transcriptome wide miR-200 regulated gene networks (Bracken et al., 2014). An alternate application of HITS-CLIP drew the attention of my work; using Argonaute HITS-CLIP selecting for and quantifying small RNAs. This approach allowed me to assess the function of small RNAs within the RNAi pathway.

Argonaute HITS-CLIP was a key technique making it possible to assess the functionality of transfected microRNA mimics. This was the first of a series of experiments used to draw important conclusions including that transfected RNA resides predominantly within cellular vesicles such as lysosomes and are commonly misinterpreted as functional at physiological levels by many in the field (Thomson et al., 2013).

I also applied Argonaute HITS-CLIP to assess the functionality of miRNAs and other endogenous small RNAs which are increasingly being observed in sequencing data by several groups, as well as our own. This work supported recent findings that small RNAs derived from small nucleolar RNAs (snoRNAs) are capable of acting as microRNAs and adds to the expanding list of non-canonical miRNAs. However importantly, I describe that microRNA function within a cell is dependent on sufficient expression levels and I compare the relative function of small RNAs derived from a variety of RNA precursors beyond canonical microRNAs (Thomson et al., 2014).

BRACKEN, C. P., LI, X., WRIGHT, J. A., LAWRENCE, D. M., PILLMAN, K. A., SALMANIDIS, M., ANDERSON, M. A., DREDGE, B. K., GREGORY, P. A., TSYKIN, A., NEILSEN, C., **THOMSON, D. W.**, BERT, A. G., LEERBERG, J. M., YAP, A. S., JENSEN, K. B., KHEW-GOODALL, Y. & GOODALL, G. J. 2014. Genome-wide identification of miR-200 targets reveals a regulatory network controlling cell invasion. *EMBO J*, 33, 2040-56.

BRACKEN, C. P., SZUBERT, J. M., MERCER, T. R., DINGER, M. E., **THOMSON, D. W.**, MATTICK, J. S., MICHAEL, M. Z. & GOODALL, G. J. 2011. Global analysis of the mammalian RNA degradome reveals widespread miRNA-dependent and miRNA-independent endonucleolytic cleavage. *Nucleic Acids Res*, 39, 5658-68.

CHI, S. W., ZANG, J. B., MELE, A. & DARNELL, R. B. 2009. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*, 460, 479-86.

GREGORY, P. A., BERT, A. G., PATERSON, E. L., BARRY, S. C., TSYKIN, A., FARSHID, G., VADAS, M. A., KHEW-GOODALL, Y. & GOODALL, G. J. 2008b. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol*, 10, 593-601.

THOMSON, D. W., BRACKEN, C. P. & GOODALL, G. J. 2011. Experimental strategies for microRNA target identification. *Nucleic Acids Res*, 39, 6845-53.

THOMSON, D. W., BRACKEN, C. P., SZUBERT, J. M. & GOODALL, G. J. 2013. On Measuring miRNAs after Transient Transfection of Mimics or Antisense Inhibitors. *PLoS One*, 8, e55214.

THOMSON, D. W., PILLMAN, K. A., ANDERSON, M., LAWRENCE, D. L., TOUBIA, J., GOODALL, G. J., & BRACKEN, C. P. 2014. Assessing the gene regulatory properties of Argonaute-bound small RNAs of diverse genomic origin. *Nucleic Acids Res*, 43(1):470-81.

Acknowledgements

I owe gratitude to my supervisors Dr Cameron Bracken and Prof Greg Goodall. As my honours year supervisor as well as my PhD supervisor, I can thank Cameron for guiding me through the basic to complex of laboratory skills. Cameron has put countless hours into demonstrating techniques, reading and rereading manuscripts and sharing his genuine enthusiasm and intrigue for research. Greg has had an underlying influence over my work, sharing his profound understanding of the field and firm leadership of the lab. Furthermore, I have gained support from the members of the Goodall Laboratory and the broader research community at the Centre for Cancer Biology (CCB). Several members of the Goodall laboratory deserve special thanks for generously sharing their time and wisdom. I was lucky to share laboratory space with Andrew Bert, a skillful molecular biologist and Dr. Phillip Gregory, both who never hesitated to offer their assistance. Two members of the ACRF genomic facility, David Lawrence and Dr. Katherine Pillman, must be thanked for guiding me through bioinformatics and analysis, but also for sharing the fun side of bioinformatics.

List of Figures

Figure 1-1: Canonical MicroRNA Biosynthesis.....	4
Figure 1-2: Crystal structure of yeast Argonaute with guide and target nucleotides.	9
Figure 2-1: Argonaute high throughput sequencing of cross-linking immunoprecipitation (HITS-CLIP) protocol, also known as CLIP-seq.	33
Figure 2-2: PARE (parallel analysis of RNA ends) protocol, also known as GMUCT (genome-wide mapping of uncapped transcripts) or degradome-seq.....	37
Figure 3-1: Measurement of miRNA by qRT-PCR after transient transfection with miRNA mimic.	63
Figure 3-2 Quantitation of functional transfected miRNA mimic by deep sequencing of RNA from Argonaute immunoprecipitation.....	65
Figure 3-3: Transfected siRNAs show little co-localisation with Argonaute.	67
Figure 3-4: Transfected siRNAs localise with and adjacent to lysosomes.	68
Figure 3-5: The passenger strand of an unmodified miRNA mimic is incorporated into RISC.	70
Figure 3-6: Measurement of miRNA by qPCR after antisense inhibitor transfection does not reliably measure the level of the cellular miRNA.	72
Figure 3-7: Measurement of microRNA mimic abundance by qPCR is not representative of microRNA functionality – microRNA mimic titration.....	81
Figure 3-8: Measurement of microRNA mimic abundance by qPCR is not representative of microRNA functionality - qPCR following microRNA mimic transfection time course	82
Figure 4-1 Argonaute binds small RNA from diverse sources	96
Figure 4-2 Comparing the abundance of AGO-bound small RNAs with their capacity to suppress reporters of fully complementary sequence:.....	98
Figure 4-3_ Higher sensitivity reporter assays do not demonstrate miRNA-like roles for other AGO-bound small RNAs.....	102
Figure 4-4_ Argonaute binding is a better indicator of miRNA function than total expression	104
Figure 4-5_ Argonaute selectively binds small RNAs.....	106
Supplementary Figure 4. 1: Representation of all ncRNA sequences irrespective of their genomic loci	113
Supplementary Figure 4. 2: Default parameters for different sequence aligners yield similar results	114

Supplementary Figure 4.3 117

supplementary Figure 4. 4: Over-expression of a sdRNA mimic indicates miRNA-like functionality . 118

supplementary Figure 4. 5: The abundance of miRNAs within cells partially indicates functionality 119

Supplementary Figure 4. 6. Location of snoRNA and tRNA-derived small RNAs 120

Supplementary Figure 4. 7. Profiles of various AGO-associated small RNAs 121

Abbreviations

2D-DIGE	two-dimensional differentiation in-gel electrophoresis
3'UTR	3' untranslated region
5'RLM-RACE	RNA ligase mediated –5' rapid identification of cDNA ends
A	Adenosine
AGO	Argonaute
AGO1	Argonaute 1
AGO2	Argonaute 2
AGO3	Argonaute 3
AGO4	Argonaute 4
BLAT	Basic Local Alignment Tool
Bp	base pairs
BSA	bovine serum albumin
C	Cytosine
cDNA	complementary DNA
CLIP	cross linking immunoprecipitation
CLIP-seq	cross linking immunoprecipitation sequencing
cm	Centimetre
CO ₂	carbon dioxide
Co-IP	co-Immunoprecipitation
<i>Ct</i>	<i>Cycle threshold</i>
DAPI	4'-6-Diamidino-2-phenylindole
DGCR8	DiGeorge critical Region 8
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EMT	Epithelial to mesenchymal transition
endo-shRNA	Endogenous short hairpin RNA
eRNA	Enhancer RNA
Exp5	Exportin-5
FCS	Foetal calf serum
FL	Firefly luciferase
G	Guanine
GFP	Green fluorescent protein
GMUCT	Genome-wide mapping of uncapped transcripts
H	Hours
HEK293	Human embryonic kidney 293
Hg18	Human genome version 18
Hg19	Human genome version 19
HITS-CLIP	High throughput sequencing of cross-linking immunoprecipitation
kb	Kilobase
lincRNA	long intergenic non coding RNA
LNA	locked nucleic acid
M	Molar

MDCK	Madin Darby canine kidney
mg	milligram
min	minute
miR	microRNA
miRNA	microRNA
miRNA*	microRNA star (passenger strand)
ml	millilitre
mM	millimolar
mm9	mus musculus genome version 9
mRNA	messenger RNA
NaCl	sodium chloride
ncRNA	non-coding RNA
neg	negative
ng	nanogram
nM	nano Molar
nt	nucleotide
OD	optical density
P	phosphate
PAR-CLIP	photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation
PARE	parallel analysis of RNA ends
PAZ	<i>Piwi</i> Argonaute and Zwillie domain
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde
pM	pico Molar
PNK	Polynucleotide <i>kinase</i>
Pre-miR	precursor microRNA
pri-miR	primary microRNA
PTEN	Phosphate and tensin homolog
qPCR	quantitative polymerase chain reaction
qRT-PCR	quantitative reverse transcription polymerase chain reaction
RISC	RNA induced silencing complex
RL	renilla luciferase
RNA	ribonucleic acid
RNAi	RNA interference
rpm	revolutions per minute
rRNA	ribosomal RNA
RT	reverse transcriptase
RT	reverse transcriptase
SDS	sodium dodecyl sulphate
shRNA	small hairpin RNA
SILAC	stable isotope labelling with amino acids in cell culture
siRNA	small interfering RNA
snoRNA	small nucleolar RNA
snRNA	small nuclear RNA

sRNA	small RNA
stRNA	small temporal RNA
T	Tyrosine
TBE	tris buffered saline
TE	Tris EDTA
tiRNA	transcription initiation RNAs
tRF	tRNA derived RNA fragments
tRNA	transfer RNA
TuD	tough decoy (constructs)
U	Uracil
UCSC	University of California Santa Cruz
ug	Microgram
ul	Microliter
uM	micromolar
UTR	untranslated region
UV	ultraviolet
vRNA	vault RNA
ZEB	Zinc finger E-Box binding homeobox