

2015

A Genomic Portrait of Hepatitis C Virus and MicroRNA-122

Joseph M. Luna

Follow this and additional works at: http://digitalcommons.rockefeller.edu/student_theses_and_dissertations



Part of the [Life Sciences Commons](#)

Recommended Citation

Luna, Joseph M., "A Genomic Portrait of Hepatitis C Virus and MicroRNA-122" (2015). *Student Theses and Dissertations*. Paper 287.

This Thesis is brought to you for free and open access by Digital Commons @ RU. It has been accepted for inclusion in Student Theses and Dissertations by an authorized administrator of Digital Commons @ RU. For more information, please contact mcsweej@mail.rockefeller.edu.



A GENOMIC PORTRAIT OF HEPATITIS C VIRUS AND MICRORNA-122

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Joseph M. Luna

June 2015

A GENOMIC PORTRAIT OF HEPATITIS C VIRUS AND MICRORNA-122

Joseph M. Luna, Ph.D.

The Rockefeller University 2015

Hepatitis C virus (HCV) uniquely requires the liver specific microRNA-122 (miR-122) for replication, yet global effects on endogenous microRNA (miRNA) targets during infection are unexplored. In this body of work, we employed high-throughput sequencing and crosslinking immunoprecipitation (HITS-CLIP) experiments of human Argonaute (AGO) during HCV infection. We demonstrate robust AGO binding on the 5' untranslated region of HCV RNA at known and predicted miR-122 sites, thereby establishing conclusive biochemical evidence of endogenous miR-122 action on HCV RNA that firmly agrees with previous genetic evidence. We further characterize novel AGO binding on HCV RNA to determine its dependence on miR-122, miRNAs generally, replication competence and time. These results establish an unbiased interaction landscape between HCV RNA and cellular miRNAs, mostly miR-122.

On the human transcriptome, we observed reduced AGO binding and functional mRNA de-repression of miR-122 targets during virus infection. This miR-122 "sponge" effect was relieved and redirected to miR-15 targets by swapping the miRNA tropism of the virus. Single-cell expression data from reporters containing miR-122 sites showed significant de-repression during HCV infection depending on expression level and site number. Based on these results, we describe a quantitative mathematical model of HCV induced miR-122

sequestration and propose that such miR-122 inhibition by HCV RNA may result in global de-repression of host miR-122 targets. This in turn may provide an environment fertile for the long-term oncogenic potential of HCV.

This last point presented a fitting entree into miR-122 biology, given its known tumor suppressive activity in the liver. To conclude this work, we performed AGO-CLIP in miR-122 knockout mouse livers as well as in human liver samples, to determine the *in vivo* targetome for this miRNA across two species. Surprisingly, we discovered widespread and non-canonical miR-122 binding throughout the transcriptome. Furthermore, a substantial fraction of this binding was not conserved between mouse and human transcriptomes, despite the fact that miR-122 is highly conserved. These results, in concert with AGO-CLIP in HCV infected cells, point to a model where HCV may have evolved the use of miR-122 for its high abundance and its well buffered capacity to be inhibited with minimal detrimental effects to the host, and perhaps benefits for the virus.

In sum, this thesis reveals how miR-122 is redistributed in the cell following HCV infection. As a molecular mechanism, chronic inhibition of miR-122 by HCV RNA is proposed to impact, and may very well help induce, the complex constellation of liver diseases that characterize this infection in humans.

To my grandparents

Che non men che saper dubbiar m'aggrada

"For doubting pleases me as much as knowing"

-Dante, *Inferno* XI

ACKNOWLEDGEMENTS

This work was conducted under the mentorship of Robert B. Darnell and Charles M. Rice, and I am profoundly grateful to both of them for their guidance, patience, and openness. I thank Bob for his unwavering optimism and zeal for unraveling the mysteries of RNA in our genomes, and for his warmth. I thank Charlie for his deep and wide-ranging insight, his good humor, all the free wine, and for turning an interest in RNA viruses into a calling. These two men have tremendously altered the course of my life.

I heartily thank the members of my committee, Luciano Marraffini and Paul Bienaisz (Chair), for their support and critical eye over the years. I thank Peter Sarnow, my external examiner, for making the cross-country trek and for ongoing longdistance mentorship. I'm eternally grateful to the Dean's Office and the Rockefeller graduate program for the extraordinary opportunity to study here.

My enduring gratitude goes to: Troels Scheel for his enthusiasm, experimental brilliance, his generosity as a friend, and for editing this thesis; Tal Danino, for his mathematical artistry; Aldo Mele and Jak Fak, for RNA molecular biology gymnastics training; Chaolin Zhang as a bioinformatics guru; and Kalpana Ghoshal at OSU for turning my attention to *in vivo* work. Most of all, I am grateful to past and present members of the Darnell and Rice labs, you are all daily inspirations to me, and much before you make the scientific endeavor captivating, you make it fun. I single out Bill Schneider for lending his crisp editorial eye to this written document.

I thank Walther Mothes, Pradeep Uchil, and Thomas Biederer for inspiring me to go to graduate school. Carol Moberg and Jim Darnell for authoring excellent books, and for showing me how rich we become by remembering the past. Jean-Laurent Casanova for his admonishment at having never read the primary works of Pasteur, my eyes have been opened. Jen Darnell, Peggy MacDonald and Alex Ploss have provided invaluable advice and guidance throughout. I'm grateful to my late cat, Ignatius "Iggy" Montgomery, for teaching me to live in the moment. Sid Venkatesh, James Letts, Emily Lowry, Rudy Bellani, Alex Bolze, Simon Elsässer, Fabio Casadio, and Brad Rosenberg have been classmates, excellent mentors, frequent drinking buddies, and great friends. Outside of Rockefeller, Rodrigo Fuentes, Cesar Rodarte, and Peter Feigenbaum have been doing the same since college—I treasure you guys.

Lastly, I thank *la familia*. My mother Sylvia Baptista, my father Joe (the elder and wiser) Luna, my step-mother Kristelle Woods-Luna, and my siblings Shelby, Brandon, and Erin; you have all helped me grow in innumerable ways. Thank you for always being there. I'm grateful to my parents-in-law, Gail and Joe Windmuller, for unyielding support. Most of all, I thank my lovely wife Kristen, without whom, none of what I do is possible. She is an aesthete in the truest sense, an inspiring teacher, gifted writer, excellent editor, and the most loving of companions. She creates a world that I most happily inhabit.

To anyone who reads this thesis, thank you. I hope that you enjoy reading it as much as I've enjoyed writing it.

TABLE OF CONTENTS

Dedication	iii
Acknowledgements	iv
Table of Contents	vi
List of Figures	xii
List of Tables	xv
List of General Abbreviations	xvi
 Chapter 1: Introduction	
Non-A, non-B viral hepatitis	1
An overview of hepatitis C virus	3
The HCV lifecycle	7
HCV entry	8
HCV translation, protein processing and RNA replication	11
HCV assembly and release	12
HCV interactions with the innate immune response	13
HCV interactions with protein host factors	15
Host protein interactions with HCV RNA	17
MicroRNAs, small RNAs and silencing	18
MicroRNA target recognition, function and expression	22
Mechanisms of microRNA mediated control	27

miR-122: a historical overview	31
Possible roles in liver development	32
Roles in lipid and iron metabolism	33
Roles in circadian rhythm inputs to metabolic control	35
As a biomarker for liver damage	35
Tumor suppressive activities and loss in HCC	37
Roles in other cancers	39
Roles in HBV infection: a case for viral restriction	40
HCV and miR-122	41
miR-122 binding sites on the HCV genome	42
Proposed functions for miR-122 on HCV	44
Cell permissiveness and HCV exceptions to miR-122	46
Antagonizing miR-122 as an anti-viral therapy	47
Viruses and miRNAs	48
Uncovering host viral interactions at the RNA level	52
HiTS-CLIP, a method for global interrogation of protein-RNA interactions	54
Aims of this thesis	58
 Chapter 2: Materials and Methods	
Introduction	59
Selection, construction and culture of cell lines	59
Huh-7.5 TetON cell construction	60

CRISPR deletion of miR-122 in Huh-7.5 cells	61
CRISPR deletion of Drosha in Huh-7.5 cells	63
CRISPR deletion of Dicer in Huh-7.5 cells	65
HCV cloning and virus production	66
Reporter assay cloning	69
HCV assays	70
RNA assays	72
Protein assays	73
HTS methods	76
Experimental designs	76
Sample generation	78
Sample harvest	78
Ago-CLIP	79
PolyG-CLIP	89
RNAseq	94
Illumina sequencing	94
CLIP Bioinformatics	94
RNAseq Bioinformatics	99
Conservation and statistical testing	102
Modelling	102

Chapter 3: AGO-CLIP studies of HCV RNA

Introduction	107
Experimental setup	108
An AGO binding map on HCV RNA	114
AGO-CLIP timecourse studies of HCV	120
An HCV resistant to miR-122 antagonism engages miR-122	123
CRISPR disruption of small RNA biogenesis - part I	125
Probing non miR-122 dependent AGO binding on HCV RNA	129
The U3 virus replicates in the absence of miR-122 and miRNAs	132
Swapping miRNA tropism of HCV: an attempt with miR-21	134
Swapping miRNA tropism of HCV to use miR-15	137
CRISPR disruption of small RNA biogenesis - part II	144
Testing HCV dependence on Dicer: a role for pre-miR-122?	149
Chapter 3 summary and discussion	152

Chapter 4: AGO-CLIP studies of infected host cells

Introduction	157
AGO-CLIP miRNA abundance measurements - correcting ligation bias	157
A tour of AGO binding in Huh7.5 cells	165
Hepatic target examples	168
HCC target and miR-122 examples	171

HCV infection functionally reduces AGO binding on miR-122 targets: the case for a sponge	176
HCV sponges miR-122 form other genic regions	179
Comparisons with Targetscan: using miR-122 sponge to validate targets.....	183
Confirmation of miR-122 targets with LNA inhibition and knockouts	184
Validation of HCV induced miR-122 sequestration in bulk assays	188
Single cell validation - experimental setup and initial results	190
Single cell validation of cellular 3'UTRs and under physiologic miR-122 concentrations	194
A quantitative model of miR-122 sponging	200
Chapter 4 summary and discussion	203
On the likelihood of an HCV:miR-122 sponge in liver	203
On functional roles for the HCV:miR-122 sponge	210
 Chapter 5: AGO-CLIP studies of mouse and human liver	
Introduction	214
Experimental rationale	215
Experimental setup	218
Identifying miR-122 dependent peaks	223
Global analysis of miR-122:AGO peaks in WT and KO mouse liver	226
A brief survey of miR-122 dependent binding beyond 3'UTRs	231
miR-122 3'UTR targets are poorly conserved	237

miR-122 regulates a diverse set of metabolic proteins	240
Establishing a miR-122 target network in human liver	243
Intersecting miR-122 target networks between mouse and human	252
Chapter 5 summary and discussion	258
Redefining miRNA function: the Seitz hypothesis and target conservation	259
The miR-122 knockout molecular phenotype and HCV: proposed connections	263
 Chapter 6: Concluding remarks	266
Appendix of oligos	269
References	276

LIST OF FIGURES

Fig. 1.1 HCV infection natural history	4
Fig. 1.2 Global HCV genotype distribution	6
Fig. 1.3 HCV Genome Organization	8
Fig. 1.4 HCV Life Cycle	10
Fig. 1.5 Small RNA biogenesis	20
Fig. 1.6 miRNA seed and repression features	23
Fig. 1.7 Mechanisms of miRNA control	28
Fig. 1.8 A comprehensive scheme of AGO interactions during virus infections	53
Fig. 2.1 CRISPR deletion of miR-122	62
Fig. 2.2 CRISPR deletion of Drosha	64
Fig. 2.3 CRISPR deletion of Dicer	66
Fig. 2.4 AGO-CLIP overview	80
Fig. 2.5 AGO-CLIP cloning strategies	81
Fig. 2.6 RNase titration for AGO-CLIP	83
Fig. 2.7 Ribonucleotide incorporation between two polyA polymerases	90
Fig. 2.8 Analysis pipeline for AGO-CLIP	96
Fig. 2.9 Analysis pipeline for RNAseq	99
Fig. 2.10 A model for miR-122 dynamics with HCV RNA	103
Fig. 3.1 Infection frequency for HCV in Huh-7.5 cells	109
Fig. 3.2 Autoradiogram for HCV infected cells	111
Fig. 3.3 1st PCR results for AGO-CLIP library cloning	113
Fig. 3.4 Multiplex PCR results for AGO-CLIP library cloning	114
Fig. 3.5 An argonaute binding map of HCV RNA	115
Fig. 3.6 AGO binding timecourse on HCV RNA	122
Fig. 3.7 An HCV mutant resistant to miR-122	124
Fig. 3.8 A CRISPR based targeting strategy to delete miR-122 and Drosha	126
Fig. 3.9 AGO binding maps on HCV RNA in the absense of miRNAs	130
Fig. 3.10 Conserved miR-122 binding in NS5B coding sequence	131
Fig. 3.11 U3 virus replication and CLIP in absence of miRNAs	133
Fig. 3.12 Exchanging miRNA tropism to use miR-21	135
Fig. 3.13 Exchanging miRNA tropism to use miR-15	138
Fig. 3.14 m15 virus characterization	139
Fig. 3.15 m15 virus characterization in non-reporter genomes	140

Fig. 3.16 AGO-CLIP with m15 virus	142
Fig. 3.17 m15 replication in Δ miR-122 cells	144
Fig. 3.18 A CRISPR based targeting strategy to delete Dicer	146
Fig. 3.19 HCV replication in Δ Dicer cells	151
Fig. 4.1 Ligation dependent bias in miRNA-CLIP measurements	159
Fig. 4.2 CIMS analysis of miRNAs from AGO-CLIP	161
Fig. 4.3 miRNA bias correction for AGO-CLIP	163
Fig. 4.4 Host cluster annotation in Huh-7.5 cells	167
Fig. 4.5 AGO-CLIP binding to notable hepatic transcripts	169
Fig. 4.6 AGO-CLIP binding to HCC genes and ncRNA	172
Fig. 4.7 AGO-CLIP identifies known miR-122 targets	174
Fig. 4.8 AGO-CLIP identifies novel miR-122 targets	175
Fig. 4.9 A common assay to detect miRNA and HCV RNA levels	176
Fig. 4.10 HCV infection de-represses endogenous miR-122 targets	178
Fig. 4.11 HCV infection effects on othe genic regions	180
Fig. 4.12 HCV infection effects on other miRNAs	182
Fig. 4.13 Comparisons between AGO-CLIP and bioinformatic prediction	184
Fig. 4.14 AGO-CLIP in LNA treated and Δ miR-122 cells	185
Fig. 4.15 Exchanign miRNA tropism redirects miRNA sequestration	187
Fig. 4.16 qPCR validation of RNA de-repression	189
Fig. 4.17 Luciferase validation of RNA de-repression	190
Fig. 4.18 A single-cell assay for miRNA activity	192
Fig. 4.19 Single cell fluorescent reporter measurements of synthetic 3'UTRs	194
Fig. 4.20 Single cell fluorescent reporter measurements of cellular 3'UTRs	195
Fig. 4.21 Boosting miR-122 levels in Huh-7.5 cells to mimic human liver	197
Fig. 4.22 Validation of miR-122 sponging by HCV under <i>in vivo</i> -like levels of miR-122	198
Fig. 4.23 Validation of miR-122 sponging under altered miR-122 levels	199
Fig. 4.24 Quantitative modeling of miR-122 sequestration by HCV	202
Fig. 4.25 Potential mechanisms of miR-122 activity despite invariant mature levels	208
Fig. 5.1 Autoradiogram of AGO-CLIP in miR-122 KO mice	219
Fig. 5.2 Analysis of miRNA abundance in miR-122 KO mice	220
Fig. 5.3 Mouse cluster annotation by biologic complexity	222
Fig. 5.4 Mouse cluster annotation between WT and KO mice	222

Fig. 5.5 Identification of a novel G-bulged miR-122 targeting motif	224
Fig. 5.6 Detailed annotaiton of miR-122 dependent clusters	225
Fig. 5.7 Global AGO binding changes functionally derepress 3'UTR targets.....	227
Fig. 5.8 Widespread canonical and non-canonical miR-122 targeting throughout the transcriptome.....	230
Fig. 5.9 miR-122 dependent binding to coding exons.....	233
Fig. 5.10 miR-122 dependent binding to 5'UTRs.....	236
Fig. 5.11 miR-122 CLIP targets are poorly conserved in mice.....	239
Fig. 5.12 Gene ontology analysis of de-repressed miR-122 targets	242
Fig. 5.13 miR-122 qPCR in human liver tumor and normal tissue	244
Fig. 5.14 Autoradiogram of AGO-CLIP in human liver.....	244
Fig. 5.15 Analysis of miRNA abundance in human liver	246
Fig. 5.16 Human cluster annotation by biologic complexity.....	247
Fig. 5.17 Human cluster annotation between normal and tumor liver tissue	248
Fig. 5.18 Global reduction in AGO binding to miR-122 targets in liver tumor.....	249
Fig. 5.19 miR-122 CLIP targets are poorly conserved in humans.....	251
Fig. 5.20 miR-122 target overlaps between human and mouse.....	254
Fig. 5.21 Non-conserved common targeting for the ALPL transcript.....	255
Fig. 5.22 Conserved targeting of IL1RN in mouse and human.....	257
 Fig. 6.1 A concise model for HCV:miR-122 sponge induced liver disease.....	 267

LIST OF TABLES

Table 5.1.....	244
----------------	-----

LIST OF GENERAL ABBREVIATIONS

3'UTR	3 prime untranslated region
5'UTR	5 prime untranslated region
AGO	Argonuate
AGO-CLIP	Argonaute HiTS-CLIP
ASO	Anti-sense oligo
BC	Biologic complexity
CDS	Coding sequence, used synonymously with Exon
ceRNA	Competing endogenous RNA
CIMS	Crosslink induced mutations
circRNA	Circular RNA
CLIP	Cross-linking immunoprecipitation
CRISPR	Clustered regularly interspaced short palindromic repeats
DAA	Direct acting antiviral
dsRNA	Double stranded RNA
GO	Gene ontology
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HiTS-CLIP	High-throughput sequencing CLIP (also CLIP-seq)
HTA	Host targeting agent
Huh-7.5	Human hepatoma cell line 7.5
IFN	Interferon
IRES	Internal ribosome entry site
ISG	Interferon stimulated gene
ITAF	IRES-transactivating factor
LD	Lipid droplet
LNA	Locked nucleic acid
lncRNA	Long non-coding RNA
LVP	Lipovirion particle
miRNA	microRNA
miR-122	microRNA-122
mRNA	messenger RNA
NGS	Next generation sequencing
NANBH	Non-A non-B hepatitis
PAMP	Pathogen associated molecular pattern

PAR-CLIP	Photoactivatable-Ribonucleoside CLIP
RNA	Ribonucleic Acid (aka “Life’s indispensable molecule”)
RNABP	RNA binding protein
RNAi	RNA interference
RNAseq	RNA sequencing
SGR	Sub-genomic replicon
SHAPE	Selective 2'-hydroxyl acylation analyzed by primer extension
siRNA	Small interfering RNA
TA	Target abundance
TE	Transposable element

Chapter 1: Introduction

In 1989, Qui-Lim Choo, Michael Houghton and colleagues at the Chiron corporation identified a positive sense, ~10kb RNA of exogenous origin that was closely associated with non-A, non-B hepatitis (NANBH) and renamed Hepatitis-C Virus (HCV) (Choo et al., 1989). With a molecular handle for identifying HCV, work quickly progressed in establishing a causal role in NANBH, identifying anti-HCV antibodies in NANBH patients, and linking HCV with end-stage liver disease and hepatocellular carcinoma (HCC) in humans (Alter et al., 1989; Kiyosawa et al., 1990; Kuo et al., 1989). Despite comprising its own genus among the *flaviviridae* family of viruses, HCV was later found to exhibit significant sequence heterogeneity and was grouped into at least seven major genotypes with numerous subtypes (Bukh et al., 1993; Smith et al., 2014). The domestication of HCV to enable detailed laboratory investigation in tissue culture was a long and laborious process that ultimately succeeded with the development of sub-genomic replicons, the isolation of highly permissive cell lines, and the isolation of adaptive HCV variants that lead to the establishment of viral clones capable of producing infectious virus particles in cell culture (Blight et al., 2002; Lindenbach et al., 2005; Lohmann et al., 1999; Wakita et al., 2005). In the quarter century since its discovery, strategies to cure HCV have progressed dramatically from interferon therapy, with ~50% cure rates, to direct acting anti-viral agents (DAAs) with cure rates in excess of 95% across a variety of viral genotypes (Sadler and Lee, 2015). These stunning developments, only apparent in the past few years

and currently ongoing, were coincident and dependent upon breakthroughs in basic studies of HCV in virus labs the world over, to whom full credit as a field is due. For anyone who has borne witness to the discovery and cure of HCV within the span of one scientific lifetime, it is tempting to declare an end to this viral scourge in the age of highly effective curative therapy. And yet, much remains to be done. It is, in this student's eyes, only the beginning.

An overview of hepatitis C virus

As of 2014, HCV remains a major global public health burden with an estimated 185 million people ~2.8% of the world's population) currently infected (Mohd Hanafiah et al., 2013). In the United States, between 3.4 and 4.4 million people are estimated to be infected with HCV (Chak et al., 2011). While most infections are asymptomatic, symptoms during an acute HCV infection include hepatitis and jaundice (Hoofnagle, 1997; Villano et al., 1999). Between 10-40% of HCV infections spontaneously resolve while the majority of cases result in a chronic infection (Figure 1.1) (Thomas et al., 2000; Thomas, 2013). Persistent HCV infections can result in an array of outcomes, including cirrhosis and HCC, which manifest over decades in 2-30% of patients (Bruno et al., 1997). Complications from cirrhosis and related end-stage liver diseases (ESLDs) combined with HCV-associated HCC have made HCV the primary cause of liver transplants in the U.S. (Davis et al., 2010). Moreover, HCV-induced cirrhosis constitutes a top risk factor for HCC (Herbst and Reddy, 2013).

While HCV infections are globally prevalent, they are unevenly distributed across regions and age groups. HCV is the largest proximal cause of liver transplantations in industrialized nations where infection prevalence remains highest among marginalized populations such as the incarcerated, injection drug users, and the homeless (Chak et al., 2011). Global prevalence is highest in developing nations, particularly in the Middle-East and Asia (Mohd Hanafiah et al., 2013).

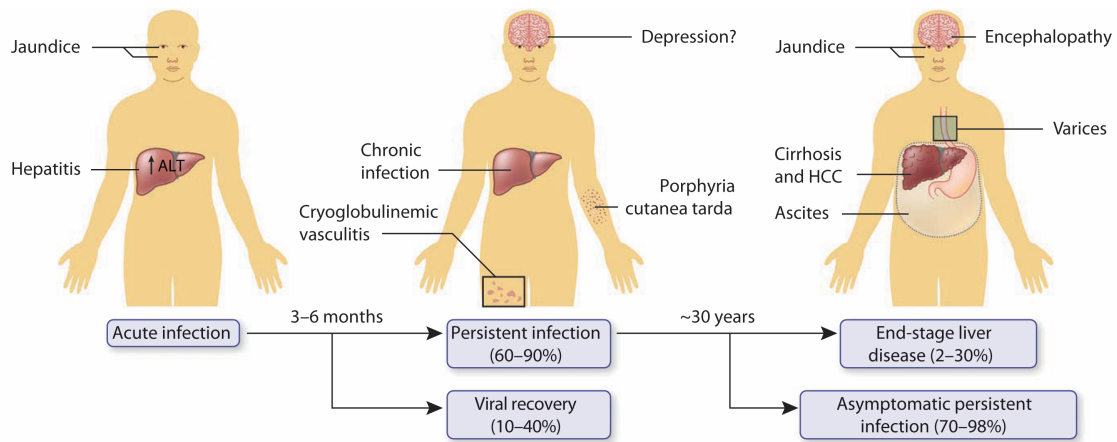


Figure 1.1. HCV infection natural history. An acute infection may result in hepatitis and jaundice, which in 10-40% of cases resolves within 3-6 months. Unresolved infections persist for decades in 60-90% of cases and can result in extra-hepatic complications. A wide range of cases (2-30%) progress to end stage liver disease including cirrhosis and HCC. ALT: Alanine transaminase. Figure from (Thomas, 2013).

Egypt is estimated to have the highest HCV prevalence among nations, where up to 50% of persons born before 1960 are affected. The history of HCV infection in Egypt exemplifies global HCV transmission in the era before mass blood surveillance: mass public health injection campaigns to eradicate schistosomiasis from the 1950s to 1970s resulted in millions of Egyptians unknowingly becoming infected with HCV. These mid-20th century surges in infection, in Egypt and elsewhere, were subsequently linked to unsafe injection and unsterile treatment practices that largely no longer occur (Thomas, 2013). Combined with routine blood screening in place since the early 1990s, incidences of HCV infection from contaminated blood products are exceedingly remote (Chak et al., 2011).

Despite tremendous advances in stemming new HCV infections from blood products, the epidemiology of HCV in the 20th century presents a continued challenge given that the infection is largely asymptomatic and presents with liver disease decades after the initial infection (Thomas et al., 2000). It is estimated that in the U.S., persons born between 1946 and 1965, the so-called "baby boom" generation, will likely bear the brunt of HCV related complications. Peak projections estimate over a quarter million HCV liver-related deaths between 2020 to 2029 (Davis et al., 2010). Concurrently, estimates on the need for liver transplantations are expected to sharply rise in the coming decade (Biggins et al., 2012). These sobering statistics combined with their projected economic and quality-of-life impacts, even in the age of curative DAAs, make continued scientific efforts at understanding the role of HCV in promoting liver disease all the more relevant (Thomas, 2013).

The uneven global distribution of HCV is reflected in the distribution of the various genotypes of the virus (Figure 1.2) (Messina et al., 2015). At both the nucleotide and amino acid level, genotypes differ around 30%, subtypes around 20%, and isolates within a subtype between 2-10% (Simmonds et al., 2005). Of the six major known genotypes, Genotype 1 predominates in the Americas and western Europe, accounting for 70% and between 50-70% of cases in each region, respectively. Genotype 2 is prevalent in the Americas, sub-Saharan Africa, and Japan, where genotype 1 is also present.

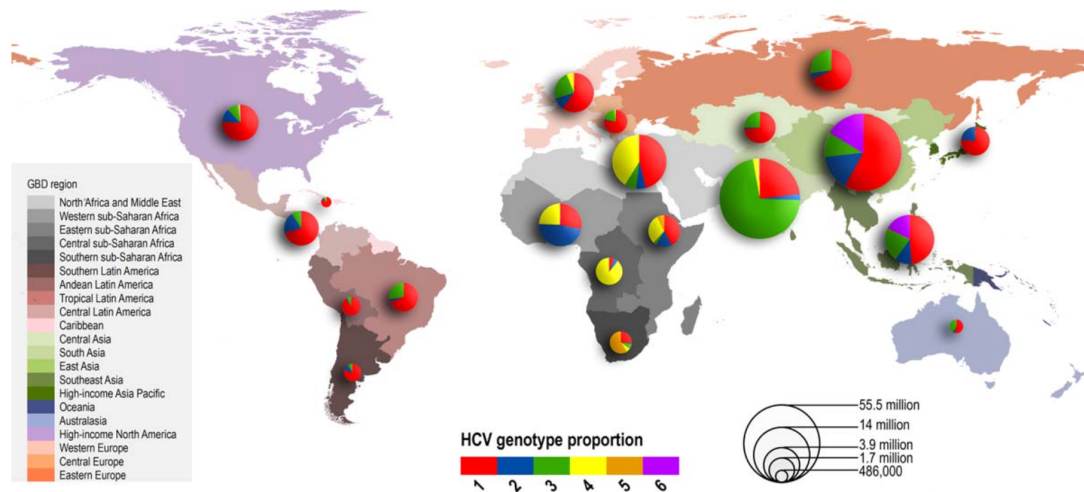


Figure 1.2. Relative prevalence of each HCV genotype by Global Burden of Disease (GBD) region. Size of pie charts is proportional to the number of seroprevalent cases as estimated by (Mohd Hanafiah et al., 2013). Figure from (Messina et al., 2015).

Genotype 3 is mostly found in Europe, southern and western Asia, while genotypes 4 and 5 are prevalent in central and South Africa. Genotype 6 mostly affects East and Southeast Asia. A recent and novel genotype 7, with only a few cases, appears to be present in central Africa (Murphy et al., 2014). While all genotypes appear to have similar pathogenic features, some differ in their association with certain complications and their abilities to be studied in the lab. For example, genotype 3 has been associated with an increased risk of liver steatosis and HCC compared to other genotypes (Adinolfi et al., 2001; Nkontchou et al., 2011; Rubbia-Brandt et al., 2000). Historically, Genotype 1 has been both the most prevalent and difficult to treat (Hoofnagle and Seeff, 2006). Genotype 2 viruses have been the most proximate source of cell culture adapted HCVs: a

Japanese genotype 2 isolate (JFH-1) was the progenitor of fully infectious HCV cell culture systems (Lindenbach et al., 2005; Wakita et al., 2005).

With the advent of DAAs providing highly curative, interferon-free, treatment regimens that exhibit pan-genotypic effectiveness with minimal side effects, it is conceivable that viral genotyping will cease to be relevant for HCV therapy (Messina et al., 2015). And yet, with a plurality of outcomes for HCV infections, from asymptomatic to severe liver disease, environmental factors such as genotypic differences, combined with host-genetics in the dynamically regulated and aging liver only highlight how much there is to learn about HCV disease causation. For our purposes, this introduction will start small and at the source by turning our attention to the virus.

The HCV life cycle

HCV is an enveloped, spherical virus particle between 50-80nm in diameter (Catanese et al., 2013b). The virion is studded with E1 and E2 glycoprotein heterodimers embedded in the lipid bilayer, which encloses a nucleocapsid consisting of the core protein and a positive sense RNA genome approximately 9.6kb in length. The RNA genome consists of highly structured 5' and 3' untranslated regions (UTRs) that flank a single viral open reading frame (ORF) which encodes a single polyprotein of around 3000 amino acids that is processed into ten distinct components comprised of structural proteins (Core, E1, and E2), p7 channel, and the non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) (Figure 1.3) (Lindenbach and Rice, 2005).

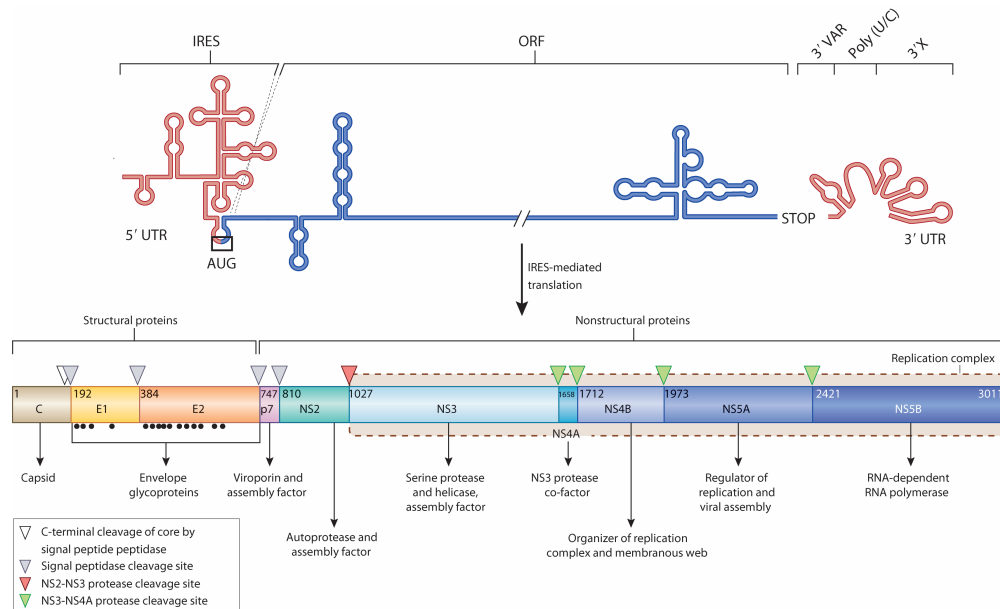


Figure 1.3. HCV genome organization and polyprotein processing. The HCV RNA genome (top) contains highly structures 5' and 3' UTRs (red) which flank a single long ORF (blue). IRES-mediated translation of the ORF leads to a polyprotein (bottom) that is co- and post-translationally processed into ten viral proteins. Core protein maturation involves a cellular signal peptide peptidase cleavage of a C-terminal signal peptide (white triangle) and cleavage from E1 by the cellular signal peptidase, which also cleaves E1, E2 and p7 from the polyprotein (gray triangles). In an autocleavage mechanism requiring a dimer to make up the composite active site, the NS2-NS3 protease cleaves itself (red triangle). The NS3 protease located in the first one-third of NS3, assisted by its membrane bound cofactor, NS4A, cleaves the remaining proteins NS3, NS4A, NS4B, NS5A and NS5B (green triangles). Glycosylation of the envelope proteins (black dots) and the functions of the individual HCV proteins are indicated. Figure adapted from (Scheel and Rice, 2013).

HCV entry

Among the most well studied aspects of HCV biology, viral entry presents a multitude of unique features and outstanding questions (Lindenbach and Rice, 2013). HCV particles typically exist as lipoviroparticles (LVPs) in infected hosts,

so described due to the association of the virion with low-density and very-low-density lipoproteins (LDL and VLDL) (Andre et al., 2002; Merz et al., 2011)(Figure 1.4). In addition to association with ApoE and ApoB, LVPs engineer a "Trojan horse" for HCV by coercing lipoprotein secretion and adsorption mechanisms for productive entry into cells, and may help shield the virus from neutralization (Bartenschlager et al., 2011). To present a summarized model of HCV entry (comprehensively reviewed by (Lindenbach and Rice, 2013)), LVPs are thought to first make contact with host cells via low affinity engagement of the LDL receptor (LDLR) and glycosaminoglycans (GAGs), before E1-E2 interactions engage co-receptors Scavenger Receptor class B member 1 (SR-B1) and Cluster of Differentiation 81 (CD81) with higher affinity promoted in part by epidermal growth factor receptor (EGFR) and ephrin receptor type A2 (EphA2). The CD81 bound HCV particles are then thought to traffic laterally to tight-junctions where Claudin-1 (CLDN1) and Occluding (OCLN) play critical roles in late stages of entry thorough interactions with the NPC1L1 cholesterol absorption receptor.

At this point, LVP uptake occurs via clathrin mediated endocytosis, likely into endosomes wherein low pH alongside CD81 priming is sufficient to trigger fusion of the viral envelope with the host membrane (Sharma et al., 2011). Although E2 is predicted to be a class II fusion protein, recent structures on highly related pestivirus E2 has led to speculation that E1 may be fusogenic (Li et al., 2013e; Omari et al., 2013).

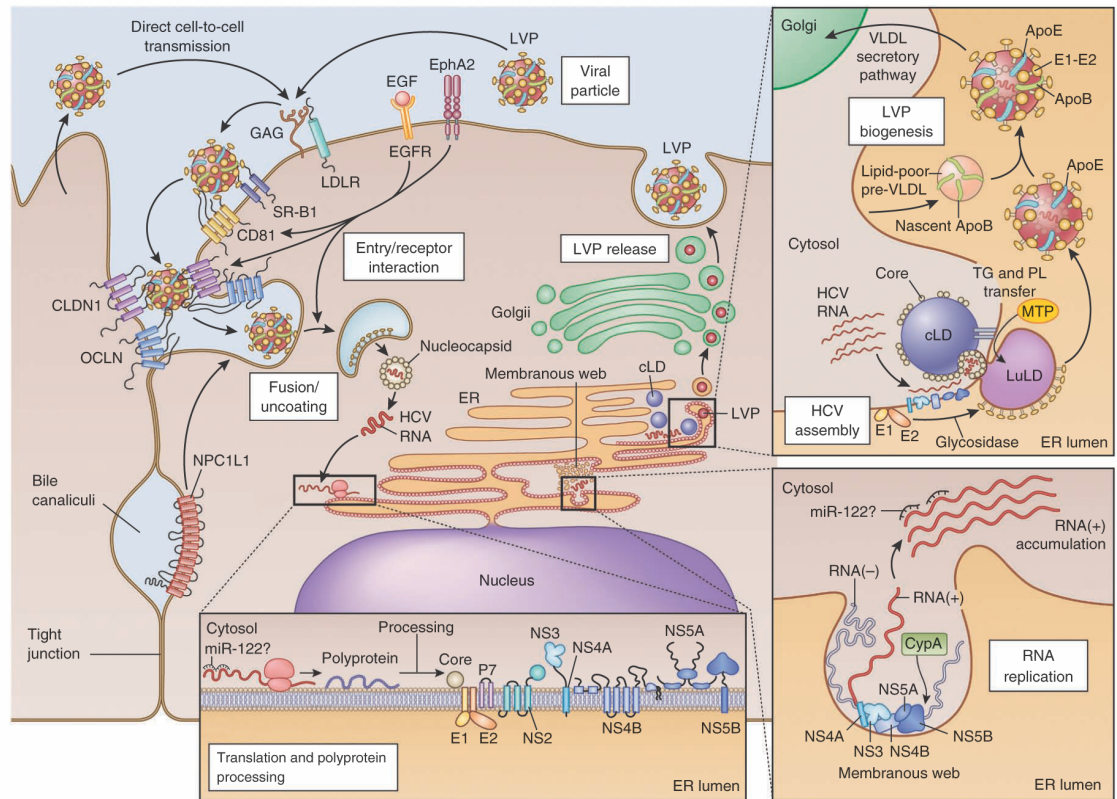


Figure 1.4. The HCV life cycle. Interaction of extracellular HCV LVPs with cellular surface receptors initiates the entry process, which can also occur from direct cell-to-cell transmission. After pH-dependent fusion and uncoating, the incoming HCV genome is translated and the resulting polyprotein is processed (bottom inset and Figure 1.3). HCV NS proteins set up ER-derived membrane spherules (membranous web, bottom right inset) to initiate the process of RNA replication. The spatiotemporal contribution of miR-122 binding to the HCV genome is not yet fully understood, and miR-122 presence is indicated with '?'. In the assembly and release process (top right inset), core protein is transferred from cytoplasmic Lipid Droplets (cLD) to form nucleocapsids are loaded with viral RNA, with the help of NS5A, NS2, NS3-NS4A and perhaps p7. It is not clear whether the RNA is transiently located on the cLD. Virion morphogenesis is coupled to the VLDL secretion pathway through luminal lipid droplets (LuLD), and particles are produced as LVPs. EphA2, ephrin receptor type A2; GAG, glycosaminoglycans; PL, phospholipids; TG, triglycerides; MTP, microsomal transfer protein. Figure adapted from (Scheel and Rice, 2013).

Further work on the E2 ectodomain has recently resolved a novel IgG like fold that does not appear to undergo structural rearrangements in low pH (Khan et al., 2014). The late stages of entry, especially fusion and viral uncoating remain poorly understood.

HCV translation, protein processing and RNA replication

As mentioned above, the HCV genome is highly structured with critical RNA elements in the 5' and 3' UTRs (Lohmann, 2013). Following uncoating and release of viral genome into the cytoplasm, the pioneer round of translation of HCV RNA is initiated by an Internal Ribosome Entry Site (IRES) located in the 5'UTR (Figure 1.3 and 1.4) (Fraser and Doudna, 2007). Translation begins on the IRES through direct and RNA structure-guided recognition of the 40S ribosomal and eukaryotic initiation factor 3 (eIF3), thus bypassing the need for the cap-structure and related binding factors (Tsukiyama-Kohara et al., 1992; Wang et al., 1993). The AUG start codon is directly placed into the peptidyl (P) site of the 40S subunit (Reynolds et al., 1995; Rijnbrand et al., 1996). The resulting HCV polyprotein is co- and post-translationally cleaved by the viral NS2-NS3 and NS3-NS4A proteases and by cellular proteases (signalase and signal peptide peptidase), respectively into ten HCV proteins (Figure 1.3) (Lindenbach and Rice, 2005). The NS4B and NS5A proteins are believed to induce ER rearrangements that result in membrane spherules which in aggregate are termed the "membranous web" (Figure 1.4) (Egger et al., 2002; Romero-Brey et al., 2012).

RNA replication occurs in this membranous web, which is thought to offer protection from dsRNA sensing and damaging reactive oxidants. As the RNA-dependent RNA polymerase, NS5B forms the core of the HCV replicase complex, which in concert with the NS3 helicase, NS4A, and multi-functional NS5A, synthesizes anti-sense RNA genomes from sense templates (Lohmann, 2013). Anti-sense genomes serve as the template for continual and successive production of sense genomes, which can then undergo translation, assembly and export, or RNA replication. The assembly, stability and decay of the replicase is thought to be a highly dynamic process, and so determining the precise spatiotemporal roles for replicase components has been difficult, and much remains to be explored (Lohmann, 2013).

HCV assembly and release

Virus assembly and release is tightly coupled with host cell lipid biosynthesis (Lindenbach and Rice, 2013). Upon protease cleavage and liberation from the HCV polyprotein, the mature core protein critically relocates to cytoplasmic lipid droplets (cLDs) (Barba et al., 1997; Miyanari et al., 2007). Nucleocapsid formation involves the interaction of core with NS5A, though the precise mechanism for delivery of the RNA genome to sites of assembly is poorly understood (Lindenbach and Rice, 2013). Close juxtaposition of sites of RNA replication and virion assembly are thought to facilitate nucleocapsid formation alongside NS2 coordination of assembly through interactions with E1 and E2, p7,

NS3, and NS5A (Jones et al., 2007; Phan et al., 2009). E1 and E2, once processed, proceed through the ER where they assume a type 1 membrane topology. Further post-translational processing of E1 and E2 takes place to include the addition of N-linked sugar moieties and their trimming with glycosidases (Lavie et al., 2007). Nucleocapsids forming at cLDs link up with mature E1 and E2 by budding into luminal lipid droplets (luLDs) where the virion completes maturation and is secreted via the VLDL pathway (Gastaminza et al., 2008). As virus particles transit through this secretory pathway, they are thought to be protected from low pH exposure by the p7 ion channel (Wozniak et al., 2010). Once secreted from the cell as LVPs, the process then begins anew (Bartenschlager et al., 2011).

HCV interactions with the innate immune response

HCV presents a thought-provoking case in that the standard of care prior to DAAs was interferon treatment, a central molecule in the innate immune response. And yet the mechanisms by which IFN therapy failed to work were mostly unknown and continue to be a major focus of current research (Horner and Gale, 2013). One theme centers on unique features of the HCV RNA genome. Unlike normal cellular mRNAs, the HCV genome lacks a 5' cap-structure. And like all RNA viruses, HCV replication goes through an intermediate phase where double stranded RNA (dsRNA) is produced. As such uncapped RNAs or dsRNA products are rarely found in normal cellular cytoplasm, these

molecules constitute pathogen associated molecular patterns (PAMP) for which the cell has evolved sensing and effector mechanisms to mount an anti-viral response (Beutler et al., 2007). Early reports linking innate immunity to HCV infection focused on protein kinase R and RNaseL, two classical interferon stimulated genes (ISGs), as effector molecules counteracted by HCV proteins or RNA structures (Han and Barton, 2002; Taylor et al., 1999). Still, how HCV evaded innate immune sensing was a mystery. Clarity emerged in 2003, when Foy, Gale and colleagues established a role for NS3/4A protease activity as critical for preventing IRF3 activation and the downstream IFN response (Foy et al., 2003). A flurry of subsequent reports established that NS3/4A antagonizes a molecule dubbed the mitochondrial antiviral signaling protein (MAVS, aka IPS-1, VISA, and Cardif), a key player downstream of retinoid acid-inducible gene-I (RIG-I) (Meylan et al., 2005). In this manner, NS3/4A cleavage of MAVS blunts RIG-I activation of the interferon response. MAVS cleavage has been observed in patients where importantly, the presence of MAVS cleavage fragments correlated with lower levels of IFN activation (Bellecave et al., 2010).

The story of MAVS is presented here as one of a growing list, which includes TRIF (also cleaved by NS3/4A), a renewed interest in the sensing and effector activities of PKR, and the enigmatic clinical and predictive associations of HCV with IFNL4 (Horner and Gale, 2013; Prokunina-Olsson et al., 2013). The fact that HCV proteins have evolved functions that antagonize the host innate immune response is not surprising, particularly given what is known for other

viruses, notably HIV-1, for which each accessory protein seems all but tailor-made to counter a host restriction factor (Malim and Bieniasz, 2012). Work utilizing ISG screens has provided a compelling platform to systematically address these issues in HCV and in other viruses (Diamond and Schoggins, 2013; Dittmann et al., 2015; Schoggins et al., 2011; Wilson et al., 2012). What bears keeping in mind for HCV is that the complete picture of its regulation and subversion of innate immunity is unclear and that future surprises lie in store.

HCV interactions with protein host factors

As detailed above, HCV depends extensively on the host cell to infect, replicate and spread, is subject to host innate immune control, and has evolved mechanism to evade innate immune detection. While too numerous to list here, a vast number of host factors influencing each aspect of the viral life cycle have been identified using RNAi-screening and mass spectroscopy methods (Shulla and Randall, 2012). HCV interactions with the host can be broadly classified as either pro- or anti-viral, on the basis of functional studies aimed monitoring the effect of factor perturbation on measureable virus outputs: RNA, protein and infectious virus. Many pro-viral factors have been the subject of detailed mechanistic investigation, as potential therapeutic targets complementary to DAA centered approaches. A few examples are highlighted here:

Cyclophilin A (CypA), a peptidyl-prolyl *cis-trans* isomerase required for HCV replication through its interaction with NS5A, is a prime example, as it was

discovered after observing that the immunosuppressive drug cyclosporin A (CsA), which targets CypA, inhibited HCV replication in cell culture (Watashi, 2003). While the precise role for CypA in HCV replication is still not clear, CsA derivatives are highly effective at antagonizing the critical NS5A-CypA interaction across all genotypes, and importantly, pose a high genetic barrier to drug resistance *in vivo* (Flisiak et al., 2012)

Diacylglycerol acyltransferase-1 (DGAT1) is an ER resident, triglyceride synthesizing enzyme, that was found to be critical for HCV particle formation (Herker et al., 2010). DGAT1 appears to regulate core trafficking via NS5A to lipid droplets, such that DGAT1 knockdown or pharmacologic inhibition prevents virus assembly (Camus et al., 2013; Herker et al., 2010).

Found in several siRNA screens, phosphatidylinositol-4 kinase III- α (PI4KIII α) is an ER associated lipid kinase that hijacked by NS5A and seems to be essential for the integrity of the membranous web (Berger et al., 2009; Reiss et al., 2011). Later work identified that HCV inhibitors thought to target NS5A were actually directed to PI4KIII α . While promising, transient ablation of PI4KIII α lead to severe liver pathology in mice, so clearly more work is needed to determine if strategies aimed at inhibiting this enzyme are viable (Vaillancourt et al., 2012).

The mechanisms of how these host proteins facilitate HCV replication have been and are being worked squarely with therapeutic potential in mind. The overall motivation is that by developing host targeting agents (HTAs), problems

with viral resistance associated with prototypical DAAs can largely be circumvented, and in select cases perhaps eliminated (Scheel and Rice, 2013).

Host protein interactions with HCV RNA

As the master blueprint for HCV, the viral RNA genome must interact with many host RNA binding proteins (RNABPs) to facilitate its tripartite role as template for replication, translation, and RNA packaging into virions. It also presents a central node whose targeted disruption theoretically presents an insurmountable barrier for virus resistance. A number of proteins bind the 3'UTR of HCV RNA to include heterogeneous nuclear ribonucleoprotein C (HNRNPC), ELA V-like RNA-binding protein 1 (ELAVL1, or HuR), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as well as polypyrimidine tract binding protein (PTBP1/2) (Ito and Lai, 1999; Luo, 1999; Petrik et al., 1999; Spångberg et al., 2000). As the latter would suggest, the poly-U/C tract of HCV RNA as well as the highly conserved 3' X-tail serve as binding sites for the above proteins though by and large, the functional importance of these protein-RNA interactions is far from clear. Most were found to interact with *in vitro* transcribed HCV RNA in cell lysates, so it's possible that many of these founding results are artifacts. Proteins binding the 5'UTR include poly(rC)-binding protein 2 (PCBP2), the La antigen, polypyrimidine tract binding protein (PTBP1/2), heterogeneous nuclear ribonucleoprotein L (HNRNPL), insulin-like growth factor 2 binding protein 1 (IGF2BP1), and U6 snRNA-associated Sm-like protein (LSM1-7) (Ray and Das,

2011; Rosenfeld and Racaniello, 2005; Scheller et al., 2009; Weinlich et al., 2009). These proteins may act as IRES-transactivating factors (ITAFs) as their specific depletion impairs HCV translation in cell culture and in general they are thought to be cofactors for a variety of IRESes (King et al., 2010). Some ITAFs bind the 3'UTR in addition, as shown for PCBP2, which may promote translation via circularization of the viral genome (Wang et al., 2011). Circularization of viral genomes is a common theme for regulating translation versus RNA replication in flaviviruses, however its role on HCV biology is unclear outside of *in vitro* systems. With the advent of CLIP methods to be discussed below, the functional genomic characterization of these and other protein-RNA interactions on a viral genome is increasingly within reach and represents a frontier ripe for advancement. For now, I focus your attention on one class of very well studied RNABPs, the Argonautes (AGOs), and their unique association with HCV RNA. But before that, we must explore a unique feature of this protein as being programmed by small snippets of RNA, called microRNAs, that allow AGOs to play critical roles in a multitude of biologic processes.

MicroRNAs, small RNAs and silencing

MicroRNAs (miRNAs) are a class of small (~22nt) noncoding RNAs (ncRNAs) that regulate gene expression by sequence-specific targeting of mRNAs causing mRNA degradation or translational repression (Bartel, 2009). Initially identified as non-coding RNAs, *lin-4* and *let-7*, which control the timing of

larval development in the worm *C. elegans*, miRNAs were soon found to represent an abundant class of small endogenous RNAs with a variety of regulatory roles throughout diverse metazoans (Lagos-Quintana et al., 2001; Lee and Ambros, 2001; Lee et al., 1993; Reinhart et al., 2000). In addition, miRNAs have also been found in plants, green algae, and viruses (Griffiths-Jones et al., 2008).

In general, miRNAs are but one class of a number of small endogenous RNAs that include small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs) (Ambros et al., 2003; Aravin et al., 2007). These small RNA types broadly contribute to the phenomena of RNA silencing, whereby small RNAs act as guides for PIWI domain containing RNA nucleases, predominantly the Argonautes (AGOs), to repress gene expression. However, miRNAs differ considerably from other small RNAs in their mode of biogenesis (Figure 1.5) (Bartel, 2004). miRNAs are transcribed by RNA polymerases II and III into primary miRNA (pri-miRNA), which undergoes processing into pre-miRNA hairpins through the action of the Microprocessor complex composed of the RNase III enzyme Drosha and Di-George syndrome critical region 8 (DGCR8) proteins (Gangaraju and Lin, 2009). For a minority of miRNAs, Drosha independent pre-miRNA production can occur from miRNA genes present in introns, called "mirtrons", for which the spliceosome functions analogous to Drosha (Okamura et al., 2007; Ruby et al., 2007).

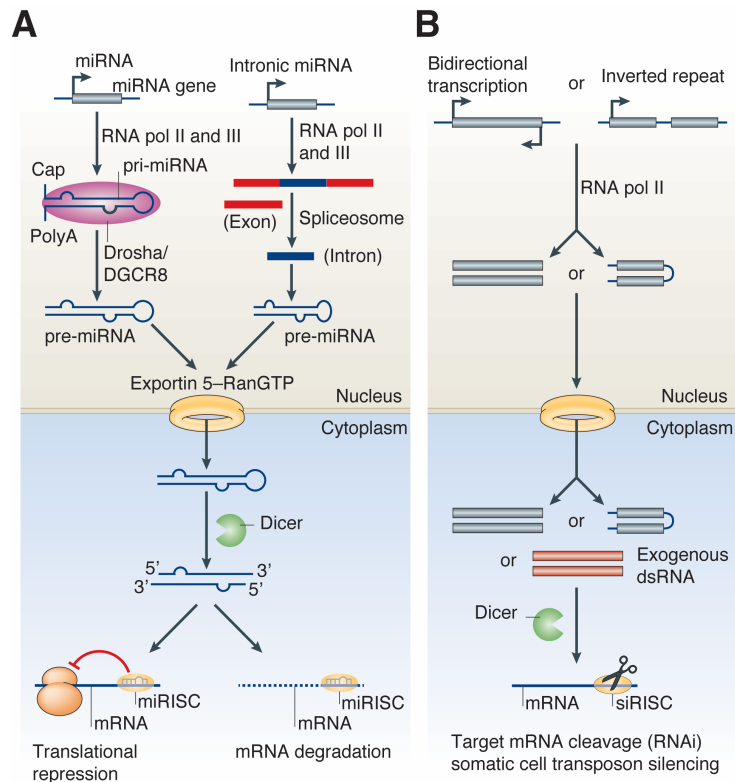


Figure 1.5. Canonical miRNA and siRNA biogenesis in metazoans. **(A)** MicroRNA (miRNA) genes are transcribed by RNA polymerases II and III into pri-miRNAs. These are processed into pre-miRNAs in the nucleus by the Microprocessor complex, which contains the RNase III enzyme Drosha and the double-stranded RNA-binding protein DGCR8. Some miRNAs, called mirtrons, have also been shown to be generated from introns that bypass Drosha requirement. Pre-miRNAs are then transported into the cytoplasm by the karyopherin exportin-5 and RanGTP, where they are further processed by the RNase III enzyme Dicer. This results in double-stranded 20–25nt intermediates with 2nt overhangs on the 3' end. One of the RNA strands is then loaded by Dicer into an RNA-induced silencing complex (RISC), which contains an Argonaute family member, that then targets the 3' untranslated region of the target mRNAs by an imperfect match between the miRNA and the mRNA, to induce mRNA degradation or repress translation. **(B)** The small interfering RNA (siRNA) pathway initiates with either bidirectional transcription or transcription of an inverted repeat that results in a double-stranded precursor. This precursor is also processed in the cytoplasm by Dicer and is loaded into a small interfering RISC (siRISC) complex intended for target mRNA degradation. Alternatively, dsRNA from an exogenous source (viral, transfected, careless graduate students, etc) can be processed directly from the cytoplasm. Figure modified from (Gangaraju and Lin, 2009).

Pre-miRNAs are exported from the nucleus through the action of exportin-5 and RanGTP (Bohnsack et al., 2004). Once in the cytoplasm, pre-miRNAs undergo processing by the RNase III enzyme Dicer, generating a double stranded 20-25nt intermediate that gets loaded onto AGO to form a RNA-induced silencing complex (RISC) (Hutvagner and Zamore, 2002). The miRNA bound RISC represses through base pairing of the miRNA with the 3'UTR of the mRNA target, and subsequently promotes mRNA decay or translational repression, via partially understood mechanisms (Filipowicz et al., 2008).

The small interfering RNA (siRNA) pathway initiates with either bidirectional transcription or transcription of an inverted repeat that results in a double-stranded precursor. This precursor is also processed in the cytoplasm by Dicer and is loaded into a small interfering RISC (siRISC) complex intended for target mRNA degradation (Carthew and Sontheimer, 2009). Alternatively, dsRNA from an exogenous source can be processed directly from the cytoplasm and loaded onto RISC (MacKay et al., 2014). While this pathway is present in mammalian cells, it is not thought to compose an appreciable anti-viral response (see below).

In theory, the capacity for siRISC or miRISC to silence target mRNAs without nuclear control or input offers a highly flexible and rapid platform for an immune response to foreign nucleic acids (Carthew and Sontheimer, 2009). Indeed, RNA interference (RNAi) as described above forms that backbone of the innate immune response in both plants and invertebrates (Baulcombe, 2004).

Along similar lines, CRISPR systems in bacteria and archaea provide somewhat analogous protection mechanisms from invading bacteriophage DNA or RNA, with the added benefits of being both adaptive and heritable (Bondy-Denomy and Davidson, 2014). Given the above, vertebrates appear to be outliers with elaborate protein based immune responses, centered largely on interferon signaling, that induce the production of hundreds of effector proteins called interferon stimulated genes (ISGs) (Schoggins and Rice, 2011). The relationship between a robust protein based and an ancient, though ancillary RNA-based immune response in vertebrates will be discussed in a following section. It suffices here to mention that recent findings have only reinvigorated the discussion on the anti-viral functions of RNAi in mammalian cells (Burgess, 2013; Li et al., 2013c; Maillard et al., 2013; Sagan and Sarnow, 2013; Tanguy and Miska, 2013).

MicroRNA target recognition, function and expression

At the simplest level, miRNA function is governed within the AGO protein. On AGO, a miRNA encounters its mRNA target via Watson-Crick base pairing to a "seed" region on the mRNA (positions 2-7 from the miRNA 5' end) (Figure 1.6A) (Bartel, 2009). Each miRNA is predicted to regulate hundred of targets, primarily by engaging 3'UTRs (Lewis et al., 2005). With hundreds of known and predicted miRNA genes in the human genome, and comparable numbers in mice and other mammals, the array of possible regulatory events is astounding. However, three key principles are worth emphasizing.

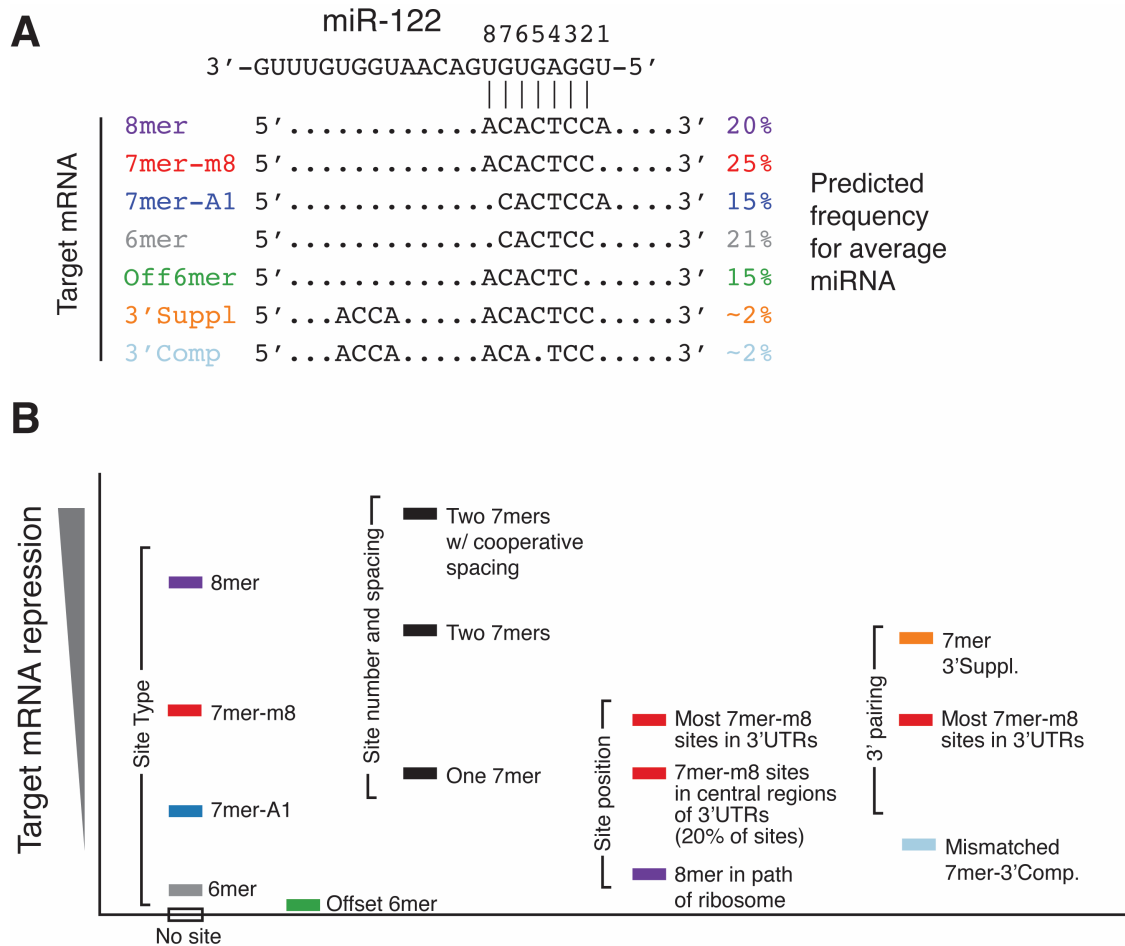


Figure 1.6. miRNA seed characteristics and repression features. **(A)** Schematic of miRNA (miR-122) engagement of mRNA targets displaying all major classes of miRNA seed pairing: 8mer, 7mer-m8, 7mer-A1, 6mer, offset 6mer, and 3' supplementary or compensatory pairing, which can occur with all of the above seeds, but is illustrated here with 7mer-m8. The predicted frequency of each seed class for conserved mRNA targets for a prototypical and conserved miRNA is shown at right from calculations by (Friedman et al., 2009). **(B)** Target mRNA repression is schematized along the y-axis where repressive potential of seed site types, site number and positioning, and 3' supplementary pairing is displayed graphically. Panel B modified from (Bartel, 2009)

First, binding energetics and the location of mRNA binding sites are important in terms of delineating functional repression. The strength of target mRNA repression is directly proportional to seed energetic stability, as summarized in Figure 1.6B. Practically, this means that 8mers are on average more repressive than 7mers, which are more repressive than 6mer seeds, and so on. 7mer-m8 seeds are typically more repressive than 7mer-A1 seeds, as the first miRNA position does not engage in base pairing with mRNA; instead the mRNA binding pocket in AGO structurally prefers an adenosine to allow seed binding to occur (Elkayam et al., 2012; Schirle et al., 2014). Supplementary 3' pairing of a miRNA to its target beyond nucleotides 10-12 can boost repression largely by increasing base pairing stability, though these are considered rare (Bartel, 2009). By and large, miRNA binding to 3'UTRs is more repressive than binding to coding exons, presumably due to competition of the translocating ribosome, though this has not been directly tested. As might be expected, multiple miRNA seeds per mRNA are more repressive than single seeds, where notably, cooperative repression is observed if miRNA sites are within 40nt of one another (Grimson et al., 2007; Saetrom et al., 2007).

Second, functional conservation tends to matter for both the miRNA and its mRNA binding seed targets. Many miRNA families—i.e. a collection of miRNAs that share the same seed—are broadly conserved throughout metazoans while others are found in only a few lineages (Altuvia et al., 2005). miRNAs related by family typically regulate the same sets of genes in various organisms, such as

miR-142-3p in myeloid cells or miR-1 in muscle, though this is not always the case (Lee et al., 2007). In broad terms, the conservation of a miRNA binding site on its target mRNA can be viewed as evidence for its functional selection. This is particularly true for miRNAs with critical roles in developmental transitions or stem cell maintenance (Gangaraju and Lin, 2009).

Lastly, one point that tends to be overlooked is that functional miRNA:mRNA interactions can only take place if both species are co-expressed. Like mRNAs, miRNA expression ranges dramatically from ubiquitous to highly tissue specific. In their current iterations, most miRNA prediction methods overwhelmingly fail to account for co-expression of a miRNA and its target, providing instead a static list of every possible interaction. In principle, miRNA:mRNA co-expression is highly related to the conservation status of a miRNA binding site, since a selected-for site will be most "functional" upon encountering its expressed cognate miRNA. Conversely, one can also think of the depletion of sites from messages co-expressed with miRNAs. For a prototypical transcriptome, 3'UTR targets contained on average about half as many conserved and non-conserved binding sites for expressed miRNAs as expected by chance. In other words, large swaths of the transcriptome appear to be *selectively avoiding* regulation by a co-expressed miRNA (Farh et al., 2005). Furthermore, the evolutionary pressure to avoid miRNA sites is specifically detected in "housekeeping genes" which might explain their shorter 3'UTRs in animals compared to the longer 3'UTRs of orthologous transcripts in plants and

fungi, which lack the same miRNAs (Stark et al., 2005). Considered a major driver of 3'UTR evolution, the phenomena of selective avoidance has given rise to the term miRNA "antitargets" to describe mRNAs lacking co-expressed miRNA sites (Bartel and Chen, 2004; Farh et al., 2005). This negatively impacts miRNA prediction in that many bona-fide, but tissue-specific or species-specific binding sites are missed as false negatives.

The situation described above makes it unsurprising that the bioinformatic emphasis on conservation of miRNA seeds has limits. While conservation of a miRNA or its target binding site has historically been a key predictive and computationally searchable hallmark for phenotype screening, more recently it has become clear that both non-conserved and non-canonical targeting are frequent occurrences with functional importance (Frankel et al., 2014; Loeb et al., 2012). The emphasis on miRNA and target conservation is historically justifiable given the blind frequency of miRNA seeds on target transcripts (on average one 7mer every 10kb, multiplied by ~1000 miRNAs). Reliance on conservation and strict seed definitions dramatically reduced the searchable space on a transcriptome and with it, the number of false positives. In this vein, the advent of CLIP based NGS methods, to be discussed later in this chapter, is a further refinement to target prediction, and arguably supercedes the reliance on conservation, seed definition and co-expression by providing direct biochemical evidence of a miRNA engaging its target in a native setting (Chi et al., 2009; Grosswendt et al., 2014; Hafner et al., 2010)(Moore, et al. 2015. Submitted).

To sum, miRNAs likely influence the expression or evolution of nearly all mammalian mRNAs, via direct interactions or inversely through selective avoidance, as well as through a myriad targeting rules (Bartel, 2009). In so doing, we can readily appreciate the ability by which new regulatory events in 3'UTRs can rapidly emerge within organisms. As 3'UTRs drift to include or exclude miRNA seeds, or as miRNAs become expressed in certain conditions to regulate previously unperturbed gene networks, the stage is set for functional modulation in species and tissue specific realms towards the evolution of new regulatory features.

Mechanisms of miRNA mediated control

But what of more proximal functions for miRNAs within individual cells? As strictly post-transcriptional modifiers of gene expression, miRNAs are unique in that they exert regulatory effects at the level of transcriptional noise (<2 fold repression in most cases) as defined by inter-individual variation in gene expression (Cheung et al., 2003). This modest regulation provides an added dimension to finely tune mRNAs to generate a rich palette of gene expression (Figure 1.7A-B). This concept is similar to a dimmer switch or rheostat, which forms an apt analogy to explore the various classes of miRNA targets under changing environments (Figure 1.7C) (Bartel and Chen, 2004). In the rheostat model, regulation can proceed in one of three ways: switch-like, tuning, and neutral. Switch-like regulation is the most familiar in that it is binary; the expression of a miRNA "turns off" the expression of its target protein by inhibiting

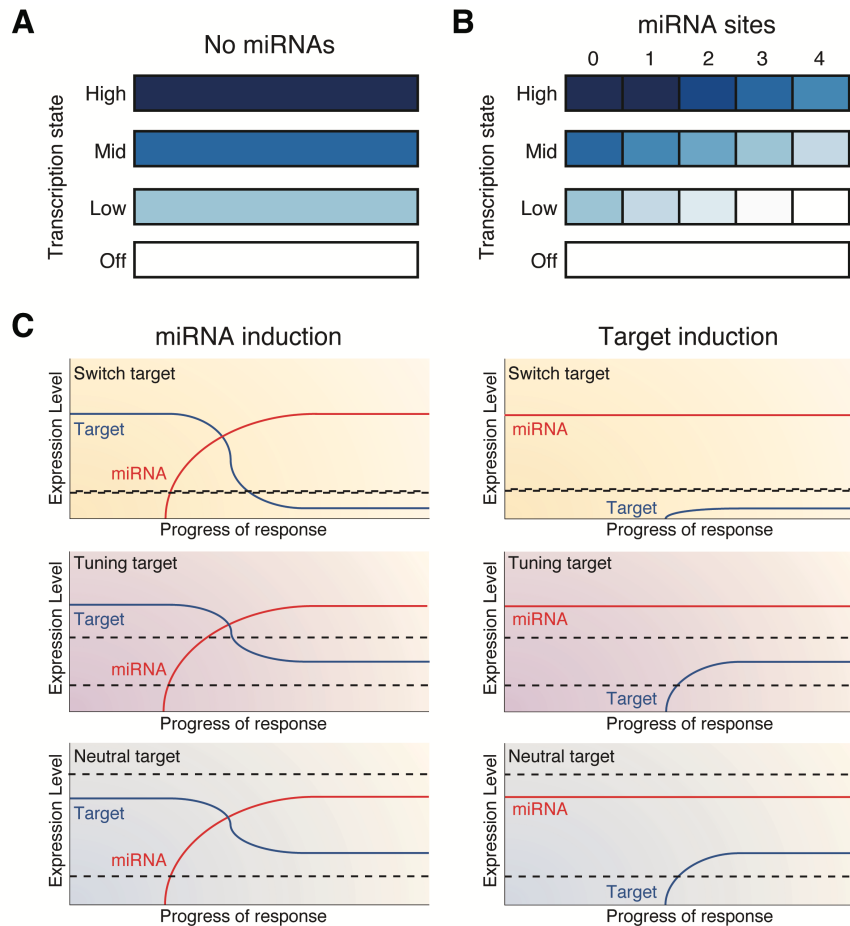


Figure 1.7. Mechanisms of miRNA control. **(A)** Transcriptional regulation can provide a variety of distinct expression levels, illustrated graphically as various shades of blue. **(B)** Additive miRNA action adds an additional and post-transcriptional layer of control to produce a richer and more complex expression palette. **(C)** Three categories of miRNA targets are presented. Upper dashed line indicates the upper limit of desired protein expression, lower line demonstrates the threshold under which the protein is no longer active; in the top panels, these lines overlap indicating that these thresholds are one and the same and act in a binary on-off fashion. On the left, miRNA expression induction can engage like a switch and turn off the expression of a protein (top); can tune protein expression to within an ideal range (middle) or may have a neutral effect (bottom). On the right, target induction can be highly dampened by a constitutively expressed miRNA (top), can be titrated to an optimal expression (middle), or have a neutral effect (bottom). Panels A and B adapted from David Bartel's *iBiology* "Intro to miRNAs" Youtube series. Panel C is modified from (Bartel and Chen, 2004).

its mRNA. Similarly, a constitutively expressed miRNA can silence aberrant mRNA expression as a means to dampen transcriptional noise. In the tuning mode, miRNA expression can lower protein amounts to some desired level, or can dampen an overshooting transcriptional event. Lastly, miRNA action can exert neutral effects where a protein is repressed but still within an optimal range for activity. Importantly, examples of each of these versions of miRNA targeting have been described (Bartel and Chen, 2004).

The optimal expression ranges for proteins whose mRNAs undergo miRNA regulation--the dashed lines in the figure--are crucial to the model of miRNA activity presented in Figure 1.7C. Numerous factors can influence the location of these lines, and by extension, the capacity for protein expression to be buffered from miRNAs. Here we focus on two related factors: the expression difference between a miRNA and its target, and the cumulative effect of all expressed targets of a miRNA (target abundance).

Single cell work from the van Oudenaarden and Sharp labs revealed that miRNA repression is not uniform across cells and is driven by the expression level and number of miRNA binding sites in target mRNAs (Mukherji et al., 2011). Consequently, miRNAs can generate gene expression thresholds. At one extreme, a target can be dramatically silenced if its mRNA is low relative to the expression of its regulating miRNA, i.e. it resembles switch-like behavior. At the other extreme, high target mRNA expression vastly overpowers the inhibitory effect of a low expressed miRNA (neutral). The middle ground, where miRNA

abundance and target expression converge, serves a tuning function by balancing mRNA abundance with repressive miRNA pressure. Adding multiple miRNA binding sites per transcript shifts the threshold in mathematically predictable ways (Mukherji et al., 2011). In these manners, target expression and the number of miRNA sites appear to govern expression thresholds for individual miRNA regulated transcripts.

While the above considers the expression difference between one miRNA and one target, it is important to highlight the aggregate effects of all the targets of a particular miRNA, and their role in influencing the gene expression thresholds for one another. Also termed "target abundance" (TA), this parameter has been incorporated into miRNA target prediction methods to aid in discovering functional sites (Garcia et al., 2011). Nevertheless, TA implementation typically relies exclusively on counting the number of conserved miRNA targeting events in 3'UTRs; thus, real target abundance (the total number of miRNA sites expressed in a transcriptome) is likely under-estimated. Recent demonstrations that long non-coding RNAs (lncRNAs) and circular RNAs (circRNAs) bind miRNAs and impart no repressive function on a coding transcript would agree with TA under-estimation. Moreover, these data point to an entirely new layer of miRNA regulation by competing endogenous RNAs (ceRNAs) (Hansen et al., 2013; Liu et al., 2014b; Memczak et al., 2013; Salmena et al., 2011; Wang et al., 2013b).

ceRNAs (including miRNA sponges) will be discussed in Chapters 5. For now, it suffices to say that miRNAs have continued to reveal their mysteries in the age of high-throughput sequencing, proving that they are far from settled and predictable regulators. And as we'll see, their role in virus biology is only beginning to be revealed.

miR-122

The liver specific miR-122 is the central miRNA player for our purposes. A confluence of events that lead to its discovery are worth mentioning. As a distinct miRNA entity, miR-122 was initially cloned in 2002 from a survey of small RNAs from mice and notably displayed expression only in the liver, where it was also the most abundant (Lagos-Quintana et al., 2002). At that time, the only known database sequence match for the miR-122 came from woodchuck liver, as a result of work done over a decade earlier. In the mid-1980's, Buendia and colleagues at the Pasteur Institute were studying genetic rearrangements that resulted in HCC from chronic woodchuck hepatitis virus (WHV) infections. Closely related to carcinogenic hepatitis B virus (HBV), WHV was and remains an excellent tractable model for studying hepadnavirus induced liver tumors (Tennant and Gerin, 2001). The Pasteur group found that the majority of tumors were the result of a rearrangement and enhanced expression of *N-* or *c-myc*, due to viral insertional mutagenesis (Moroy et al., 1986). In one unique case, they characterized a chromosomal translocation in which the 5' end of a previously

unknown gene recombined with the second exon of *c-myc*, resulting in a 50-fold increase in *c-myc* expression (Etiemble et al., 1989; Moroy et al., 1989). Termed the *hcr* locus, the unknown gene was found to encode a highly abundant polyadenylated RNA around 4.7kb in length (Moroy et al., 1989). Curiously, the RNA was essentially non-coding as the longest ORF was only 37aa in length. The link between the mature miR-122 sequence and its parent *hcr* gene was finally made in 2003-2004, while Chang, Taylor and colleagues were studying the effects of Dicer on hepatitis delta virus (HDV) infections and confirmed that miR-122 was highly abundant in rodent livers and in human hepatoma (Huh-7) cells (Chang et al., 2003). Shortly thereafter, they confirmed *hcr* gene as the source for miR-122, and provided the first functional evidence of its ability to down-regulate an mRNA target, cationic acid transporter protein (CAT-1, aka SLC7A1), with a clear seed dependence (Chang et al., 2004). Much has been learned since then about miR-122 in terms of its role in many, if not most, aspects of liver biology. While its association with HCV is most germane to the present work (see next section), a comprehensive account of all other miR-122 functions is presented here.

Possible roles in liver development

miR-122 expression begins as early as e12.5 in mice where its expression is promoted by hepatocyte nuclear factor 6 (HNF6), a hepatocyte-specific transcription factor (Laudadio et al., 2012; Xu et al., 2010). miR-122, in turn, forms a positive feedback loop by enhancing the expression of HNF6 and other

liver-enriched transcription factors (LETFs) (Laudadio et al., 2012). HNF6 along with other LETFs participate in a network of transcriptional regulation essential for terminal hepatocyte differentiation (Kymizi et al., 2006). More directly, miR-122 also downregulates CUTL1, a transcription factor that represses genes required for hepatocyte differentiation (Xu et al., 2010). It should be noted however, that these tissue culture based results are largely based on miR-122 overexpression, and may be artifactual (Wang et al., 2010). Moreover, the miR-122 knockout mouse, both total and liver specific, is viable and displays no major defects in anatomical or functional liver development (Hsu et al., 2012; Tsai et al., 2012).

Lipid metabolism

Studies in mice following miR-122 antagonism with anti-sense oligonucleotides (ASOs) or locked nucleic acids (LNAs) revealed altered lipid profiles characterized by reduced plasma cholesterol levels, increased hepatic fatty-acid oxidation, and a decrease in hepatic cholesterol and fatty-acid synthesis rates (Elmén et al., 2008; Esau et al., 2006; Krützfeldt et al., 2005). Interestingly, while a number of direct miR-122 targets were identified in these studies using microarrays, a substantial fraction of up and down-regulated genes did not contain miR-122 sites, indicating substantial and indirect roles for miR-122 activity. Indeed several genes with crucial to lipid metabolism (ACC1, ACLY, SCD1, SREBP2 and HMGCR) were down-regulated (Esau et al., 2006; Krützfeldt et al., 2005). Consistent with these findings, miR-122 antagonism decreased

serum cholesterol levels in chimpanzees (Lanford et al., 2010). Recent work with the miR-122 knock-out mouse has linked AGPAT1 and CIDEA as direct targets of miR-122 (Hsu et al., 2012). CIDEA, also called FSP27, regulates triglyceride storage and fatty acid oxidation within lipid droplets (Keller et al., 2008; Vila-Brau et al., 2013). In the liver, CIDEA is induced during the initial stages of fasting, but decreases during late fasting (Vila-Brau et al., 2013); this oscillation is notable in light of circadian miR-122 activity (discussed below). AGPAT1 is mostly involved in regulating the energy state in adipose and muscle tissues by stimulating the conversion of fatty acids into triglycerides (Ruan and Pownall, 2001; Takeuchi and Reue, 2009). It is thought that miR-122 repression of CIDEA and AGPAT1 inhibits the storage of triglycerides in liver tissue, and may account for the increased triglyceride secretion observed in miR-122 KO mice (Hsu et al., 2012). These two relatively well-studied examples are likely the tip of the iceberg for the role of miR-122 in hepatic lipid metabolism, as we shall soon see in our own work using CLIP in miR-122 knockout mice.

Iron metabolism

miR-122 depletion was also observed to cause iron deficiency in mice where it was found that miR-122 directly targets the hemochromatosis (Hfe) and hemojuvelin (Hjv) mRNAs for suppression (Castoldi et al., 2011). These proteins activate hepcidin, a hormone that controls systemic iron levels, which in turn binds to and induces the degradation of ferroportin, an iron efflux channel expressed in hepatocytes (Castoldi and Muckenthaler, 2012; Nemeth et al.,

2004). Consequently, by suppressing the activators of hepcidin, miR-122 prevents the degradation of ferroportin, which permits proper iron uptake in the liver (Castoldi and Muckenthaler, 2012).

Circadian rhythm inputs to metabolic control

In addition to being controlled by HNF transcription factors, expression of the miR-122 locus oscillates in a circadian manner and is under the control of two retinoid-related orphan receptor elements (ROREs) upon which the oscillating REV-ERB α transcriptional repressor acts to start and stop expression (Gatfield et al., 2009; Ueda et al., 2002). In the mouse liver, mature miR-122 levels remain relatively constant, but pre- and pri-miR-122 levels fluctuate such that levels of these intermediates are lowest at Zeitgeber Time (ZT) 8-12 and highest at ZT24. These periods correspond to the peak and minimal expression of REV-ERB α respectively; in REV-ERB α knock-out mice, near constant levels of pre- and pri-miR-122 were observed (Gatfield et al., 2009). Many miR-122 targets also demonstrate circadian oscillations in expression that are consistent with miR-122 expression, where a few targets, such as PPAR β /d and Nocturnin play key roles in the liver circadian clock (Gatfield et al., 2009; Kojima et al., 2010). These results are particularly puzzling, given that mature levels of miR-122 do not fluctuate. Much remains to explore on this front.

As a biomarker for liver damage

The world of serum miRNA profiling as non-invasive biomarkers for a variety of disease is vast, sometimes oversold, and often contradictory. While

comprehensive and rigorous studies to determine the utility of miR-122 as a biomarker for a variety of complex liver diseases has not yet been performed, one trend is seemingly apparent: in the many small scale studies, miR-122 levels are elevated in serum upon liver damage or disease. Correlating with disease scoring and severity, miR-122 levels were elevated in human and murine models of liver damage from alcohol and HBV (Zhang et al., 2010). Elevated serum miR-122 has been identified as a potential biomarker for liver damage from many sources, including acute hepatotoxicity, chronic HBC and HCV infections, HCC-associated liver damage, non-alcoholic fatty liver disease, acetaminophen poisoning, and liver transplant rejection (Cermelli et al., 2011; Ding et al., 2012; Farid et al., 2012; Hu et al., 2013; Starkey Lewis et al., 2011). It should be noted however, that elevated alanine transaminase (ALT) in serum, a classic marker for liver injury, was observed in these studies alongside increased serum miR-122. Thus, it's likely that miR-122 in the blood reflects hepatic cytotoxicity more than anything else. Still, there is some evidence for miR-122 as a marker for specific liver injuries, which opens the door for delineating functional differences in miR-122 secretion. A study using a mouse model of liver damage demonstrated that circulating miR-122 could be used to differentiate between causes of liver damage: drug-induced injury led to miR-122 partitioning to the protein fraction of mouse serum, while alcohol and inflammation-induced liver injury led to miR-122 circulation in exosomes (Bala et al., 2012).

Tumor suppressor activities and loss in HCC

miR-122 has been implicated as a tumour suppressor in several different conditions and model systems, and is notably and typically lost in HCC (Bai et al., 2009; Coulouarn et al., 2009; Hsu et al., 2012; Kutay et al., 2006; Tsai et al., 2009). In the HCC cell lines HepG2 and Hep3B, which express little miR-122, restoration of miR-122 promotes apoptosis and is thought to inhibit proliferation by suppressing Wnt/ β -catenin signaling (Xu et al., 2012). Furthermore, loss of miR-122 expression in HCC cell lines correlates with increased cell migration and invasion, while restoration of miR-122 reversed these effects (Coulouarn et al., 2009). Although a comprehensive evaluation of the role miR-122 targets in tumor suppression has not yet been performed (but is the focus of Chapter 5), numerous studies have begun to link miR-122 tumor suppressor activity to specific targets. miR-122 appears to stabilize the p53 tumor suppressor by regulating Cyclin G1, and consequently sensitizes cells to doxorubicin chemotherapy (Fornari et al., 2009). The muscle specific isoenzyme of pyruvate kinase, PKM2, is a miR-122 target that is upregulated in HCC, is inversely correlated with miR-122 expression, and is thought to contribute to altered energy metabolism in HCC cells (Jung et al., 2011). miR-122 knockout mice are prone to tumor development with age: after progressive liver pathology to include fibrosis and steatosis, many mice spontaneously develop HCC after a year in age, and phenocopy the higher male-to-female disease incidence observed in humans (Hsu et al., 2012; Tsai et al., 2012). These studies demonstrated altered

lipid metabolism consistent with findings with LNA or ASO treated mice and furthermore characterized increased liver inflammation and infiltration of IL-6 and TNF producing cells as well as linked a dysregulation of a miR-122 target KLF6, to fibrosis (Hsu et al., 2012; Tsai et al., 2012). Importantly, both studies showed that miR-122 loss was reversible: a *myc* model of liver tumorigenesis could be rescued by over expressing miR-122, and miR-122 loss could be corrected by overexpressing microsomal triglyceride transport protein (MTTP), which largely restored lipid metabolism (Hsu et al., 2012; Tsai et al., 2012). It should be noted that these rescue experiments, especially in the *myc* system, were devoid of the chronic inflammation that usually accompanies and often precedes HCC development, thus it can be argued that miR-122 tumor suppressor activity is partially inflammation independent. Still, other studies suggest otherwise. A subset HCC patients bear a polymorphism (rs3783553) in a putative miR-122 binding site in the 3' UTR of IL-1 α . As IL-1 α is implicated in carcinogenesis and tumour growth, as well as in controlling tumor immunity, miR-122 regulation of IL-1 α may provide a link to liver inflammation upon miR-122 loss (Gao et al., 2009). A similar argument has been made for the myeloid cell chemoattractant CCL2, also a miR-122 target, which when derepressed alters the immune cell profile of the liver (Hsu et al., 2012). The multi-valent role of miR-122 in liver cancer reflects its role in maintaining liver homeostasis, yet clearly there is much more to learn.

Roles in other cancers

Although miR-122 is not expressed in high levels in other tissues, it has been implicated in cancers of other tissue types. Compared to normal tissues, miR-122 was found to be downregulated in gastrointestinal primary tumours and cell lines (Wang et al., 2009). This study further linked miR-122 to APC-mediated growth inhibition, as miR-122 sequestration had promoted growth; a similar result to those mentioned above in HCC cell lines (Wang et al., 2009). In a rat model of colorectal cancer, lowered miR-122 in healthy tissues was predictive of carcinogenesis elsewhere in the colon, and could be monitored ex vivo from fecal colonocytes present in the stool (Kunte et al., 2012). In breast cancer, miR-122 was identified in serum deep sequencing of stage II-III patients, with increased levels of miR-122, and presumably liver damage, correlating with and later predicting non-responsiveness to neoadjuvant chemotherapy and metastatic outcomes (Wu et al., 2012). More recent work has begun to establish a causal role for circulating miR-122 to reprogram glucose metabolism of the pre-metastatic niche in breast cancer (Fong et al., 2015). Lastly, in anaplastic thyroid cancer, miR-122 levels increase in thyroid cancer cell lines that respond to treatment with a PAX8/PPAR γ fusion protein, and subsequently are play a tumor suppressive role, both in vitro and in xenografts (Reddi et al., 2013). This would suggest that tumor suppressor activity for miR-122 might extend to other tissues, however, some caution is warranted in interpreting these findings given that miR-122 levels in non-hepatic cells is very low.

Roles in HBV infection: a case for viral restriction?

In HBV infected patients, Wang et al. found a striking negative correlation between hepatic miR-122 levels and HBV viral load. Loss of miR-122 led to increased viral replication in HepG2 and Huh-7 cells, suggesting an anti-viral role for miR-122 in HBV (Wang et al., 2012b). Further work demonstrated that miR-122 suppresses Cyclin G1 (also mentioned above), which forms a complex with p53, itself a negative regulator of HBV transcription. Thus, virus driven suppression of miR-122 leads to increased Cyclin G1, sequestering anti-viral p53, and permitting HBV transcription (Wang et al., 2012b). In addition, miR-122 has been shown to directly bind to conserved HBV sequences, and is thought to directly suppress viral replication. In apparent response to this, miR-122 expression is inhibited by the HBx viral protein by binding PPAR γ , a transcriptional activator of miR-122 (Chen et al., 2011; Song et al., 2013). It has also been proposed that HBV transcripts harboring conserved miR-122 binding sites behave as miR-122 sponges, though some caution is warranted as these results are based solely on overexpression of HBV RNAs (Li et al., 2013a). As in HCV, miR-122 has been proposed as a serum marker for HBV associated liver damage and hepatocellular carcinoma, as a means of detecting active or occult HBV infections, and of predicting HBV disease progression in patients (Chen et al., 2012; Ding et al., 2012; Qi et al., 2011; Waidmann et al., 2012; Zhang et al., 2010). While the lack of robust HBV infection systems currently presents a barrier to identifying the precise role of miR-122 in natural HBV infections, the general

and compelling evidence that miR-122 expression does appear to inhibit HBV replication reveals that there is much more to learn.

HCV and miR-122: coming together as a fascinating outlier

And then there's HCV. In 2005, Jopling, Lemon, Sarnow and colleagues, mostly at Stanford, wondered why HCV replicons could grow in Huh-7 cells but not in HepG2 cells, despite the fact that both were HCC derived hepatoma cell lines. Starting with the observation by Chang et al. that miR-122 was found to be expressed in Huh-7 but not in HepG2 cells, the Stanford group posited that miR-122 could be regulating HCV expression (Chang et al., 2004; Jopling et al., 2006). After confirming miR-122 expression in liver and Huh-7, but not in HepG2 or HeLa cells, they found that suppression of miR-122 using ASOs inhibited HCV sub-genomic replicons (SGRs) in Huh-7 cells (Jopling et al., 2005). They then searched the the HCV genome for miR-122 seeds and began a mutagenesis study, and found remarkably that a miR-122 site in the 5' UTR, when mutated at various seed positions (p3, p3,4, and p6), resulted in dead SGRs (Jopling et al., 2005). The mutation could be rescued by transfected miR-122 variants bearing complementary mutations (Jopling et al., 2005). This genetic evidence formed the basis of subsequent studies of miR-122 and HCV, which from the outset represented a unique outlier on two fronts. Whereas miRNAs typically interact with the 3'UTRs of mRNAs to suppress gene expression, HCV engagement of a

miRNA at its 5'UTR was critical for viral RNA accumulation. Subsequent work to describe the unique mechanism at play is described here.

miR-122 binding sites on the HCV genome

In addition to the original miR-122 site (termed S1), a second miR-122 site (S2) located 10nt downstream of S1 in the 5'UTR is also critical for viral RNA accumulation, and both sites engage with auxiliary pairing at nucleotides 2-3nt for S1 and 30-31nt for S2 (Jopling et al., 2008; Machlin et al., 2011). Importantly, S1 and S2 sites are highly conserved across all known HCV genotypes (Jopling et al., 2006). Additionally, viruses closely related to HCV, such as GVB-V, and the recently identified non-primate hepacivirus (NPHV), retain two and one conserved miR-122 sites in their respective 5'UTRs (Burbelo et al., 2012; Kapoor et al., 2011; Sagan et al., 2013). The very recent isolation of additional rodent and perhaps bat hepaciviruses will likely add to this list, and may offer long sought after murine models of HCV infections (Drexler et al., 2013; Firth et al., 2014; Kapoor et al., 2013).

For HCV, S1 and S2 sites were found to be simultaneously required for viral RNA accumulation, where interestingly, swapping the S1 seed for a miR-21 binding site was non-viable (Jopling et al., 2008). More recent work has delved deeper into the base contacts proposed at S1/S2. Two studies have applied selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) analysis to measure RNA nucleotide flexibility in the HCV 5'UTR with or without miR-122 *in vitro*. In both studies, SHAPE and gel shift assays revealed additional base

contacts at the S2 site, resulting in higher energetic stability (Mortimer and Doudna, 2013; Pang et al., 2012). While the functional consequences of differential S1 versus S2 binding have yet to be worked out, one recent study looking for HCV variants that would be resistant to miR-122 sequestration found that a G to A mutation at position 28, between S1 and S2, was partially resistant to miR-122 inhibition (Israelow et al., 2014).

In addition to increased contacts at the S2 site, Pang et al. observed miR-122 dependent changes in SHAPE reactivity in the IRES domain 4, suggesting a 3rd albeit very non-canonical miR-122 site in the IRES (Pang et al., 2012). The relevance of this site for a fully infectious virus is unknown. Other miR-122 sites on the virus genome can be found in the NS5B coding sequence and interestingly, one very highly conserved site in the 3'UTR; however, none of these sites appear to be functionally important, at least in cell culture system settings, as mutagenesis of these sites showed no major effects on viral replication (Jopling et al., 2008; Nasheri et al., 2011).

Interactions of other miRNAs with the HCV genome

With the discovery of a positive role for miR-122 in HCV replication, much attention turned to the role, if any, of additional miRNAs on impacting HCV by directly engaging the viral RNA genome. In 2007, Pedersen, Chisari and colleagues reported that upon IFN β treatment in cell culture, miR-122 expression was repressed and that surprisingly, a number of IFN stimulated miRNAs could antagonize HCV RNA directly and repress viral gene expression

(Pedersen et al., 2007). Furthermore, over-expression of the five IFN induced "anti-viral" miRNAs reduced HCV RNA replication in a manner that was seed dependent, as chimeric viruses of isolates containing natural mutations in putative seeds were no longer repressed (Pedersen et al., 2007). Not surprisingly, there was much excitement surrounding these findings, as they seemed to merge the protein based ISG response with one that included miRNA function in an ancient and anti-viral RNAi-like mode (Beard and Helbig, 2008). Enthusiasm was soon tempered as no correlation between miR-122 levels and HCV viral load was found in patients undergoing IFN therapy. Furthermore, the absolute expression of the "anti-viral" set of miRNAs was dramatically low in both mouse liver and in Huh-7 cells, even with IFN treatment (SarasinFilipowicz et al., 2009). Despite these findings, numerous groups have continued to find additional miRNAs that antagonize the HCV genome (reviewed in (Singaravelu et al., 2014)). It must be stressed however, that as these results are overwhelmingly based on the overexpression of small RNAs and reporter constructs, artifactual findings can be considered likely.

miR-122 function on HCV: stabilizing RNA and promoting translation

As the HCV genome lacks a 5' cap structure, one early hypothesis was that the Ago:miR-122 complex may serve as a type of cap to protect the 5' end from exonuclease recognition. This appears to be the case. A study using replication defective genomes found that miR-122 transfection slowed the decay of HCV RNA, while miR-122 inhibition enhanced decay (Shimakami et al., 2012).

Furthermore, this activity was dependent on the presence of AGO2 (Shimakami et al., 2012; Wilson et al., 2011). AGO2:miR-122 complexes appear to protect HCV RNA from 5' exoribonucleases XRN1 and XRN2, as knockdown of these factors largely rescues viral replication after miR-122 antagonism (Li et al., 2013d; Sedano and Sarnow, 2014).

In addition to protecting the 5' end of viral RNA, two studies have shown that miR-122 promotes translation, albeit somewhat transiently. Using HCV RNA in a non-premissive rabbit reticulocyte lysate system, Henke et al. observed a 50% increase in HCV translation in the presence of miR-122 but not miR-124 (Henke et al., 2008). Similar results were observed using replication defective reporter genomes in Huh-7 cells and measuring shortly after RNA electroporation (Jangra et al., 2010).

miR-122 function on HCV: a replication and translation switch?

Very recent work from Masaki, Lemon and colleagues has posited a role for miR-122 as a regulator for positive strand replication versus translation. Using 5-ethynyl uridine incorporation into nascent RNA, miR-122 transfection was found to transiently increase HCV RNA synthesis and was dependent on protein synthesis (Masaki et al., 2015). Consistent with this, a small reduction in polysome associated HCV RNA was observed upon miR-122 transfection (Masaki et al., 2015). As the S2 miR-122 site overlaps with a PCBP2 binding site, thought to aid in the circularization of the HCV genome, a competition model has been proposed where miR-122 occupancy prevents PCBP2 induced

circularization, and subsequent translation in favor of RNA replication. How this effect on increased RNA replication squares with increased HCV translation observed with miR-122 in other HCV contexts remains to be determined.

Making cells permissive: miR-122 expression engineering

The initial impulse to explain the HCV permissiveness of Huh-7 but not HepG2 cells highlights the importance of miR-122 as important for cell tropism (Jopling et al., 2005). With this in mind, miR-122 expression has been engineered to make cells permissive to HCV in HepG2 cells and in mouse cells with limited success (Israelow et al., 2014; Lin et al., 2010; Narbus et al., 2011). It is generally thought that miR-122, alongside entry factors, is a major determinant for the hepatotropism of HCV. In support of this, ectopic expression of miR-122 can render numerous non-hepatic cell lines permissible for HCV infection, including 293T cells, Hec1B (uterus), MC-IXC (nerve), and RERF-LC-AI (lung) cells (Da Costa et al., 2012; Fukuhara et al., 2012). Whether the low levels of miR-122 in non-hepatic tissues might support an extra-hepatic reservoir of HCV is unknown at present.

HCV exceptions to miR-122

Despite the fact that miR-122 seeds sites are conserved in all HCV genotypes, mutants have been isolated that are largely resistant to miR-122 antagonism (Li et al., 2011). The most resistant variant in this study, called the U3 virus, replicated in the presence of miR-122 antagomir. This recombinant virus was found to have replaced stem loop I of the IRES with cellular U3

snoRNA, ablating the S1 site while maintaining S2 (Li et al., 2011). As mentioned above, the isolation of horse, rodent and bat hepaciviruses with between one and two miR-122 sites in various configurations raises the question as to their putative miR-122 dependence. Additional work has pinpointed miR-122 independence for the hepatotropic GB virus B (GBV-B), a close relative to HCV (Sagan et al., 2013).

Antagonizing miR-122 as an anti-viral therapy: a high resistance barrier

That HCV uses a small RNA as a critical host factor raised the immediate possibility that miR-122 antagonism could be a viable anti-HCV therapy (Jopling et al., 2005). Work has progressed systematically on this front, with the demonstration that miR-122 locked nucleic acid (LNA) inhibitors were well tolerated in mouse livers and in HCV infected chimpanzees (Elmén et al., 2008; Lanford et al., 2010). The resulting human compound, called miravirsen, represents a first-in-class therapy aimed at targeting a cellular miRNA (Scheel and Rice, 2013). Phase II studies in humans treated with miravirsen resulted in rapid drops in viral load for the duration of therapy and long-term follow-up has not resulted in any major adverse effects (Janssen et al., 2013; van der Ree et al., 2014). Importantly, no virus breakthrough or escape mutants were observed for the duration of treatment, highlighting the utility of targeting a small RNA host factor and furthermore suggestive of a high barrier to resistance (Janssen et al., 2013). While miravirsen resistant HCV variants have been uncovered *in vitro*, none have yet to be reported in patients.

Viruses and miRNAs: a novel axis of host pathogen interaction

The story of HCV and miR-122 is illustrative but certainly not unique. Early speculation that other viruses, particularly large and complex dsDNA viruses, might manipulate the miRNA pathway was confirmed early on with the discovery that Epstein Barr Virus (EBV) encodes miRNAs (Pfeffer et al., 2004). Subsequent work revealed that the majority of *herpesviridae* family members encode multiple miRNA genes (Cullen, 2013). Generally, natural viruses that encode miRNAs have a DNA component to their replication cycle, have full access to host pri-miRNA biogenesis machinery by replicating in the nucleus, and typically undergo long-term persistent infections (Kincaid and Sullivan, 2012). Viruses with DNA genomes (the Herpesvirus, Polyomavirus, Ascovirus, Baculovirus, Iridovirus, and Adenovirus families) and at least one member of the retrovirus family, bovine leukemia virus (BLV) have been shown to encode functional miRNAs (Kincaid and Sullivan, 2012). Most work has focused on the herpesviruses due to the sheer number (usually dozens) of miRNAs that each virus encodes as well as the relative ease by which miRNA expression can be measured.

In general, the functions of DNA virus encoded miRNAs are thought to relate to DNA virus lifestyles which, in the case of herpesviruses, have both latent and lytic phases. From the perspective of a latent herpesvirus infection, a miRNA presents a wonderfully subversive means to auto-regulate viral gene expression or to manipulate host gene expression, all in a non-immunogenic manner (Umbach and Cullen, 2009). This hypothesis has animated the herpesvirus field

to uncover such examples and in the process has helped shed light on long mysterious phenotypes. For instance, during latency, Herpes simplex virus 1 (HSV-1) expresses one non-coding transcript called the latency associated transcript (LAT) and little else (Stevens et al., 1987). The precise role of the LAT was mysterious until it was found to encode numerous miRNAs that targeted the HSV reactivating transcription factors ICP0 and ICP4 (Umbach et al. 2008). Thus LAT expression represses leaky expression of ICP0 and ICP4 which would otherwise compromise latency if minimally expressed. Supporting this, suppression of LAT results in spontaneous reactivation in ganglia (Du et al., 2011).

In addition to auto-regulation, herpesviruses have co-opted existing miRNA network for selfish ends. A notable example is Kaposi's sarcoma associated herpesvirus (KSHV), which encodes a miRNA called miR-K12-11 that acts as a miRNA analog to cellular miR-155 (Gottwein et al., 2007). In this manner, KSHV can subvert the miR-155 targetome in its native B-cell setting by expressing its own version, which in this case, appears to suppress innate immune, pro-apoptotic and transcription factors (Gottwein et al., 2007). Moreover, as miR-155 over-expression is thought to drive B-cell lymphomas, these data also provide compelling molecular evidence for the mechanism of KSHV induced B-cell transformation (Eis et al., 2005). Similar anti-apoptotic themes were also found among Epstein Barr virus (EBV) encoded miRNAs, in analyses for which I was able to take part (Riley et al., 2012a). As many EBV miRNAs have no host analog, these data were demonstrative of the unique

capability of a virus evolving a miRNA to target new sets of genes (Riley et al., 2012a). Moreover, recent evidence suggests that EBV induced lymphoma critical requires viral miRNAs to suppress apoptosis (Vereide et al., 2013).

The story is quite different for RNA viruses and is currently evolving. It was not widely accepted that naturally occurring ssRNA or dsRNA genome viruses encode functional miRNAs (Cullen, 2010). To a degree, the arguments supporting this claim make sense: a miRNA hairpin or dsRNA precursor produced by an RNA genome should by definition undergo Drosha or Dicer cleavage of either the genome, anti-genome or subgenomic RNAs and is likely to result in reduced fitness. Additionally, dsRNA or extant hairpins are obvious targets for PAMP detection. One early study failed to show any miRNAs from six well characterized RNA viruses (Parameswaran et al., 2010). Emerging exceptions can be found among retroviruses, which package an RNA genome into virions but uniquely go through a DNA stage in their lifecycle via reverse transcription and integration into the host genome. While early studies purportedly found small RNAs encoded in the HIV-1 genome in addition to a suppressor of RNA silencing, they could not be confirmed (Bennasser et al., 2005; Lin and Cullen, 2007; Triboulet et al., 2007). A more recent renewed focus on retroviral miRNAs has centered on bovine leukemia virus (BLV), which clearly encodes a miRNA that interestingly acts as a cellular miR-29 mimic (Kincaid et al., 2012). Both simian and bovine foamy viruses (SFV and BFV) were shown to encode functional miRNAs (Kincaid et al., 2014; Whisnant et al., 2014). Notably,

SFV possesses a unique mechanism for controlling Drosha processing by regulating RNA pol II versus pol III transcription of the viral miRNA; in this manner, the virus is thought to avoid the fitness penalty associated with Drosha cleavage (Kincaid et al., 2014). Moreover, one SFV miRNA appears to act as a mimic to miR-155, which is very reminiscent of the KSHV story presented above.

The emerging picture painted for retroviruses has not gone unnoticed by RNA virologists in re-evaluating the potential for other RNA viruses to encode miRNAs. While still too early to draw firm conclusions, miRNAs have been putatively discovered in flaviviruses such as West Nile virus (WNV) and Dengue virus (DENV) (Hussain and Asgari, 2014; Moon et al., 2012; Pijlman et al., 2008; Schnettler et al., 2012).

Perhaps the most interesting recent take on miRNAs as an "anti-viral" effector concerns the North American eastern equine encephalitis virus (EEEV), a highly virulent mosquito-borne alphavirus that results in a 30-70% mortality rate in humans. Unlike its South American counterpart, Venezuelan equine encephalitis virus (VEEV), EEEV is restricted from myeloid cells which in turn do not produce interferon. It was found that this EEEV restriction in immune cells was due to the presence of miR-142-3p sites, a myeloid specific miRNA, in the EEEV RNA genome (Trobaugh et al., 2013). By this mechanism, EEEV has evolved to be suppressed in myeloid cells specifically so as to evade the innate immune response. Moreover, the selective maintenance of miR-142-3p sites on the viral genome also seemed to confer a benefit for mosquito borne

transmission, so there appears to be a selective advantage for the miR-142-3p sequence in both host and vector backgrounds (Trobaugh et al., 2013). In sum, these findings turn the anti-viral miRNA hypothesis on its head, by placing the virus in control of its own tropism where miRNAs can be considered essential host factors required to negatively select virus replication from specific contexts.

Uncovering host-viral interactions at the RNA level

As the above discussion highlights, the role of small RNAs in viral infections is far from complete and is under current, active investigation. Nevertheless, we can consider a comprehensive framework that attempts to address the possible roles by which a miRNA (or siRNA) can aid or inhibit a virus, taking into account the host or viral source of the RNA (Figure 1.8) (inspired by the discussion in (Cullen, 2009)). In this framework, we consider AGO as the protein stage for an unfolding drama between host and viral RNAs. In normal cells, miRNA:mRNA interactions take place as usual, but upon virus infection, switching and tuning of a host response (anti-viral or otherwise) can occur. Viral coercion can take place in three direct forms: by encoding viral miRNAs, by permitting miRNA targeting of viral transcripts, and by the complete viral re-programming of AGO. As highlighted above, examples of each of these forms of regulation have been documented, especially in DNA viruses. Importantly, it must be noted that each of these forms of regulation do not have to occur in isolation and can influence one another. For instance, a viral miRNA

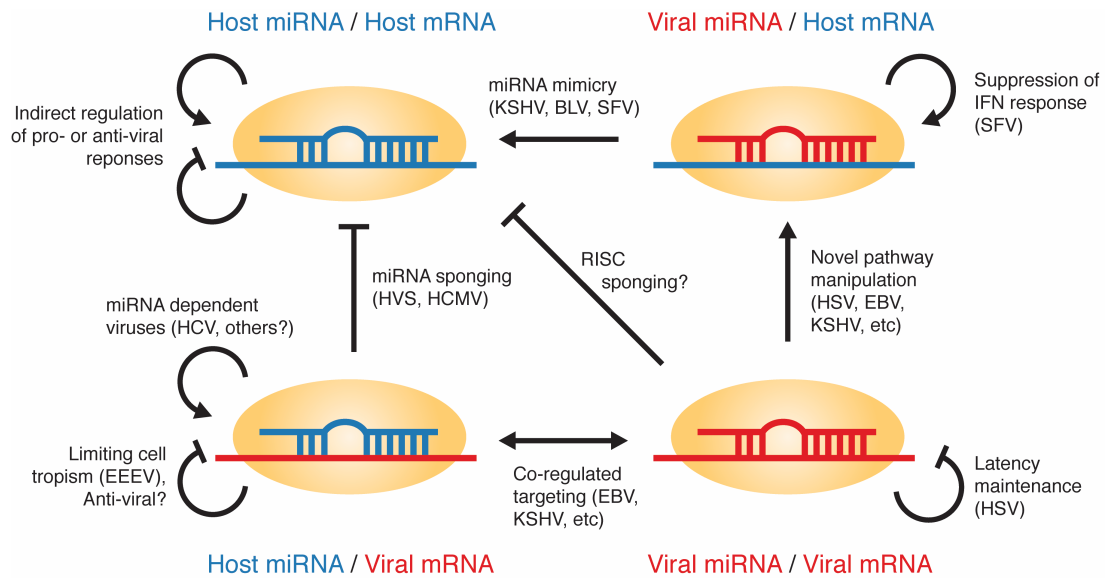


Figure 1.8. A comprehensive scheme of AGO associated miRNA:mRNA interactions between host and virus. In this model, four main interactions are considered: host responses to virus infection (top left), viral miRNA targeting of host mRNAs (top right), host miRNA targeting of viral mRNAs (bottom left), and viral auto-regulation (bottom right). Note that effects of one type of targeting may have effects on others. For instance, in miRNA sponging, an abundant viral transcript sequesters a cellular miRNA and would thus affect the normal host targeting of that miRNA. The obverse of this scenario are viral mimics of cellular miRNAs. For each scenario, relevant viruses that offer experimental support are highlighted (see text for details).

sponge would resemble a host miRNA:viral mRNA regulation on AGO, but its effects would be functionally exerted on inhibiting normal cellular miRNA function and de-repressing specific targets. This is consistent with a few examples from herpesviruses, which encode non-coding RNAs that serve as decoys for miRNA activity: *Herpesvirus saimiri* produces RNAs that were found to sequester miR-27, while human cytomegalovirus (HCMV) produces a stretch of intergenic RNA that functionally inhibits the miR-17 family (Cazalla et al., 2010; Lee et al., 2013).

Virus miRNAs that mimic a cellular miRNA are the obverse of this scenario (Kincaid et al., 2014). However, it must be stressed that these examples are all mostly anecdotal, relying on specific genetic characterization and reporter validation that misses the global picture of the RNA traffic on AGO in the infected cell. To circumvent this, we must turn our attention to genomics methods aimed at achieving a high resolution and unbiased map of global miRNA:mRNA interactions across the cell.

HiTS-CLIP, a method for global interrogation of protein-RNA interactions

High-throughput sequencing crosslinking and immunoprecipitation (HiTS-CLIP aka CLIP-seq) is a method that combines the stringent purification of RNAs bound to an immunoprecipitated RNABP and massively parallel sequencing of these RNAs by cDNA cloning and library construction (see Chapter 2). It is useful to consider HiTS-CLIP as a subset of total RNA-seq, where instead of purifying mRNA on the basis of a polyA tail (or negative selection of ribosomal RNAs), RNA is "filtered" through a protein. In this manner, the specific targets of an RNABP may be identified from the level of transcriptomes to nucleotides (Darnell, 2010).

The motivation behind CLIP based methods owes much to work aimed at understanding DNA-protein interactions in living cells. Early attempts to isolate DNA-protein complexes, often weakly bound in vivo, harbored the major concern of ensuring that physiologic complexes were preserved and not artifactually

created upon cell disruption (Dedon et al., 1991). Improved biochemical purification methods in the mid 1980's that incorporated a chemical crosslinking step greatly enhanced signal to noise (Gilmour and Lis, 1984; Kuo and Allis, 1999; Solomon and Varshavsky, 1985). In this manner, the DNA biochemists were able to ascertain physiologic DNA-protein complexes in a variety of systems, and thus underwrote the growth of the chromatin field.

The relative fragility of RNA relative to DNA is probably the most proximal cause for the lag between pioneering work with chromatin IP (or ChIP) and its protein-RNA equivalent in CLIP. Re-association artifacts in early RNA IP ("RIP") work were later confirmed to have confounded early work, presumably due to the "stickiness" of ssRNA relative to dsRNA (Brown et al., 2001; Darnell et al., 2001; 2005). Compellingly, RNA-protein artifacts were shown to be a major driver for RIP experiments with ELAV and AGO (Mili and Steitz, 2004; Riley et al., 2012b).

Again, crosslinking came to the rescue, but unlike the chromatin field, which primarily uses formaldehyde as a *general* and *reversible* crosslinking agent, UV crosslinking became the method of choice for CLIP precisely for its *specificity* and *irreversibility* (Darnell, 2010). Although the mechanism is incompletely understood, UV-mediated crosslinking is thought to involve UV light absorption by nucleic acid bases (Brimacombe et al., 1988). This in turn raises their grounded energy state to enable the formation of a covalent bond with molecules very closely opposed, within the order of ångströms (Fecko et al., 2007). Importantly, UV light (at 254nm) only induces protein-RNA and RNA-RNA

crosslinks, no protein-protein crosslinks are observed as they are with chemical reagents. Moreover, while the true figure is unknown, it does appear that the protein-RNA crosslinking reaction occurs on a minority of contact sites (estimated between 1-5%) (Darnell, 2010). The crosslink itself also does not appear to substantively interfere with the reverse transcription of RNAs with crosslinked amino acid moieties; instead RT tends to skip such bases, resulting in deletions that can be used to map RNABP binding sites to single nucleotide resolution (Zhang and Darnell, 2011). Most importantly, the covalent bond afforded by UV crosslinking enables the rigorous purification of protein-RNA complexes under stringent wash conditions.

The benefits provided by crosslinking RNA to protein enabled pioneering work by Ule, Jenson, Darnell and colleagues to map the RNA network regulated by NOVA in the mouse brain (Ule et al., 2003). What is prescient about this work is how clearly the idea was ahead of the sequencing technology: NOVA bound RNAs were cloned as a library (presumably with millions of individual binding events) but underwent TOPO cloning and Sanger sequencing. The result was a then heroic, but now measly ~340 CLIP "tags" from that mapped to NOVA regulated exons (Ule et al., 2003). Within three years, the "HiTS" was added with the adaptation of CLIP libraries for sequencing on early Illumina machines, in this way the NOVA regulated network expanded three orders of magnitude by tag number (>200K tags) (Licatalosi et al., 2008). Since then, numerous RNABPs have been studied at the systems level with CLIP to reveal transcriptome wide

roles in RNA regulation: HuC, Rbfox, RBM47, FMRP, MBNL2, FUS, PTBP2 to name but a few (Darnell et al., 2011; Ince-Dunn et al., 2012; Licatalosi et al., 2012; Nakaya et al., 2013; Vanharanta et al., 2014; Wang et al., 2012a; Weyn-Vanhentenryck et al., 2014).

Common among all of these studies is the study of "binary" protein-RNA interactions, that is, a protein bound to a single RNA target. Applying this method to a "ternary" interaction, namely AGO bound to miRNAs and mRNA targets, was achieved shortly before I joined the Darnell lab. Unequivocally, everything to be presented in this thesis is based on work done by Sung Wook Chi, Julie Zang, Aldo Mele & Robert Darnell, who in 2009 published results obtained from applying HiTS-CLIP to AGO (AGO-CLIP) and revealing a miRNA:mRNA interaction map in the mouse brain (Chi et al., 2009). It was immediately apparent to me, given the anecdotal and largely un-resolved role of small RNAs in virus infections that this could be a technique to steer towards questions in virology.

Aims of this thesis

The preceding sections build a case for studying virus infections with AGO HiTS-CLIP and using HCV as a model system, given its well-known association with miR-122 as a positive control. The overall goal of this thesis has been to generate a physical and testable miRNA:mRNA interaction map for an HCV infection genome wide. Moreover, the study of HCV raises an intriguing possibility, given that HCV causes liver cancer and critically requires miR-122 for its replication. Combined with the observation that miR-122 knockout mice spontaneously develop liver cancer, a tantalizing hypothesis emerges: HCV might effectively lower miR-122 levels in hepatocytes, which in turn may constitute an oncogenic stress in chronically infected cells. In all honesty, presenting this as an *a priori* hypothesis is a bit misleading, since it constitutes both the main conclusion and the main question raised by the work to be presented. Thus, only with a bit of hindsight can the following aims be proposed:

- Adapt and carry out AGO HiTS-CLIP analysis of HCV infected cells
- Characterize and validate AGO binding events on the HCV RNA genome.
- Characterize and validate AGO binding events on the host transcriptome.
- Interrogate the miR-122 target network in mouse and human livers.

Chapter 2: Materials and methods

Far from being a dense recitation, this section provides a holistic view of the materials and methods used in this thesis. This will include the reasoning behind the selections of reagents and protocols, the justifications for modifications, and key optimization experiments not discussed in the main text. The scope of this section will range from wet to dry, that is, from experiments at the bench or tissue culture hood, to finished bioinformatic analyses and visualization. It is my goal to provide the reader with a deep view of the thinking behind how experiments were planned and performed, how assumptions were managed, and in the case of seemingly arbitrary decision-points (particularly informatics ones), what course of action was decided upon. Most importantly, this section is written with practitioners in mind, in the hope that any graduate student or postdoc who is engaged in similar work may find something useful.

Selection, construction, and culture of cell lines

Huh-7.5 cells were clonally derived from Huh-7 cells that had been transduced with HCV subgenomic replicons (SGRs) and subsequently cured of HCV using IFN-alpha (Blight et al., 2002). These cells were found to exhibit enhanced HCV replication and virus production largely without inducing adaptive mutations, and resultantly have been a staple of HCV research in the Rice lab. Later work suggested that this enhancement in permissiveness was due to defective RIG-I signaling which abrogated an otherwise antiviral IFN response (Sumpter et al., 2005). With the observation that HCV NS3/4A specifically

antagonizes a key down stream component of RIG-I signaling, mitochondrial antiviral signaling protein (MAVS, a.k.a. IPS-1, VISA, and Cardif) (Meylan et al., 2005), we reasoned that the Huh-7.5 context would illuminate small RNA:viral RNA interactions in a setting where the innate immune response has failed, in other words, when the virus had effectively disabled the protein-based IFN response. Ultimately, however, these cells were chosen for the practical concern of enabling high infection frequency, unlike parental Huh-7 cells. Such a large and homogenously infected population of cells was critical for interpretable genomics results due to the high sensitivity, though still population based, measurements provided by CLIP and RNAseq.

Huh-7.5 cells and related variants were maintained in Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen, Life Technologies) supplemented with 5% heat inactivated fetal bovine serum (FBS, Hyclone, Thermo Scientific), and 0.1 mM nonessential amino acids (NEAA, Invitrogen, Life Technologies). Cells were maintained at 37°C in 5% CO₂. For single cells plated for clonal expansion, cell conditioned media with 10% FBS was used to enhance cell viability. To make conditioned media, cell were incubated with media containing 10% FBS and serially harvested every 24 hours. This conditioned media was filtered through 0.22µm Amicon filters (Millipore).

Huh-7.5 TetON cell construction

These cells were a kind gift from Nick Takacs, who made them as follows: Huh-7.5 cells were transduced with retroviral particles made in 293T cells

following co-transfection with pRetroX-Tet3G (Clontech), MLV GagPol and VSVg plasmids. Transduced cells were selected using 500 µg/mL G418, plated in 96-well format at 0.5-0.8 cells/well, and single cell clones were expanded. The clones were tested for permissiveness to HCV infection, growth rate, doxycycline dose-response of induction and timecourse of induction after transduction with a mCherry fluorescent protein reporter expressed from retroviral vector pRetroX TRE3G (Clontech). The clone deemed to have the optimal overall performance in these assays was re-named Huh-7.5 TetON and used in subsequent assays. Huh-7.5 TetON cells were maintained in DMEM with 10% FBS, 0.1mM NEAA, and 500µg/ml G418 (Gibco).

CRISPR mediated deletion of miR-122 in Huh-7.5 cells

Guide sequences 122.sgRNA1 and 122.sgRNA2 (to remove approximately 38 nucleotides including the miR-122 seed and partial stem-loop (Figure 2.1A) were cloned into pX330-U6-chimeric_BB-CBh-hSpCas9 (pX330, Addgene plasmid #42230) following published protocols (Ran et al., 2013). After sequence confirmation, 1.25µg of each pX330-122.sgRNA1 and pX330-122.sgRNA2 (2.5µg total) was used to electroporate 1×10^6 low passage Huh-7.5 cells with an Amaxa Nucleofector 2b machine (Lonza) using program T-28 and Nucleofector Kit V reagents, following the manufacturers instructions. The supplied GFP plasmid was used as an electroporation control. Two days after electroporation, cells were seeded in 96 well plates at 0.5 cells / well for single cell clonal expansion. To genotype single cell clones and to approximate editing

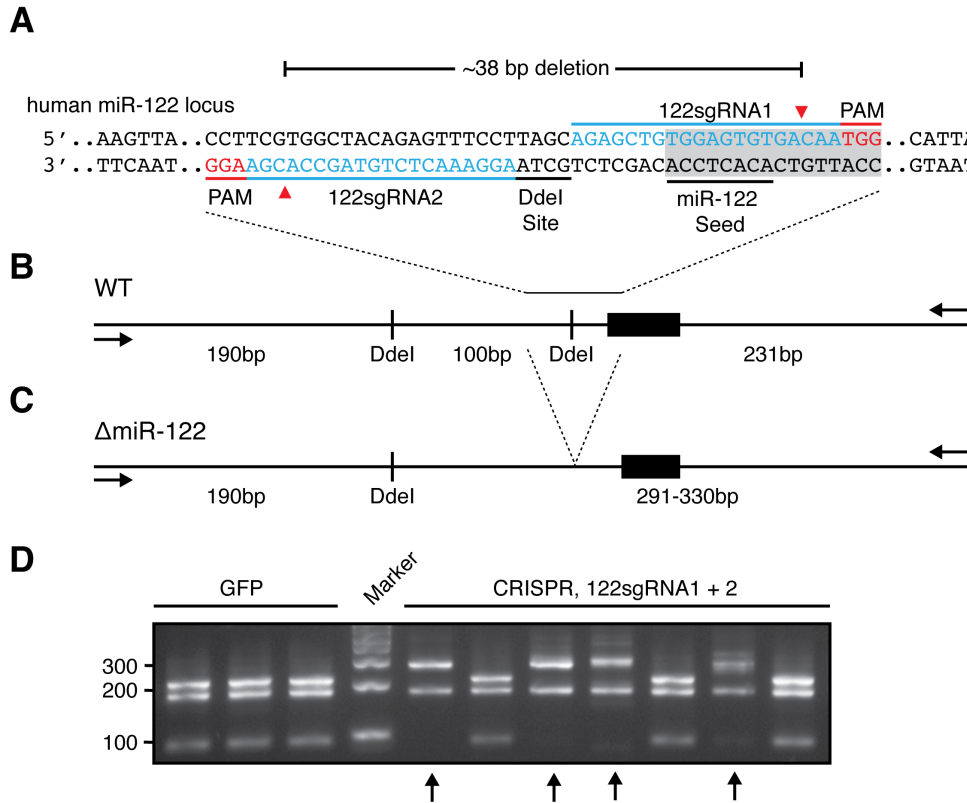


Figure 2.1. A CRISPR based targeting strategy to delete miR-122 from Huh-7.5 cells. **(A)** Guide RNAs were designed to delete a 38bp region encompassing the 5' end of the miR-122 stemloop, including the entire miR-122 seed, and an adjacent Ddel restriction site. Approximate cut sites indicated with red triangles. PAM: protospacer adjacent motif. **(B-C)** Location of Ddel restriction sites and fragment sizes relative to the miR-122 PCR amplicon for WT **(B)** and the Δ miR-122 **(C)** mutant, used for genotyping. Targeted deletion of one Ddel site is expected to yield two fragments (190 and ~300bp respectively), while unedited clones should produce three fragments (190, 100 and 231bp). **(D)** Genotyping of GFP or CRISPR transfected and expanded clonal cell lines after Ddel digestion of PCR amplicons. Suspected Δ miR-122 clones where one Ddel site has been deleted are indicated with black arrows.

efficiency in bulk cells, DNA was extracted using QuickExtract (Epicentre) and the miR-122 locus PCR amplified for 30 cycles. The resulting PCR products were purified, digested with DdeI (NEB), and underwent gel electrophoresis. As the deleted segment destroyed a DdeI restriction site, CRISPR deletions were identifiable on the basis of an altered restriction digest pattern (Figure 2.1B-D). Homozygous deletions, along with unedited and GFP only clones were expanded and used for subsequent studies. Guide sequences are in the Appendix.

CRISPR mediated deletion of Drosha in Huh-7.5 cells

For Drosha deletion, guide sequences Drosha.sgRNA1 and Drosha.sgRNA2 were designed to excise exon 9 and force an out of frame splicing between exons 8 and 10 as reported previously (Chong et al., 2008)(Figure 2.2). Guide RNAs were cloned into pX458-pSpCas9(BB)-2A-GFP (pX458, Addgene plasmid #48138). After sequence confirmation, transfection and single cell dilution cloning proceeded as above. To genotype single cell clones and to approximate editing efficiency in bulk cells, DNA was extracted using QuickExtract (Epicentre) and the Drosha exon 9 locus PCR amplified for 35 cycles. The resulting PCR products underwent gel electrophoresis, where deleted cells yielded shorter amplicons. One homozygous deletion clone (Δ Drosha) was isolated, along with unedited clones. Western and northern analysis of Δ Drosha cells are described in the next chapter. Guide sequences are listed in the Appendix.

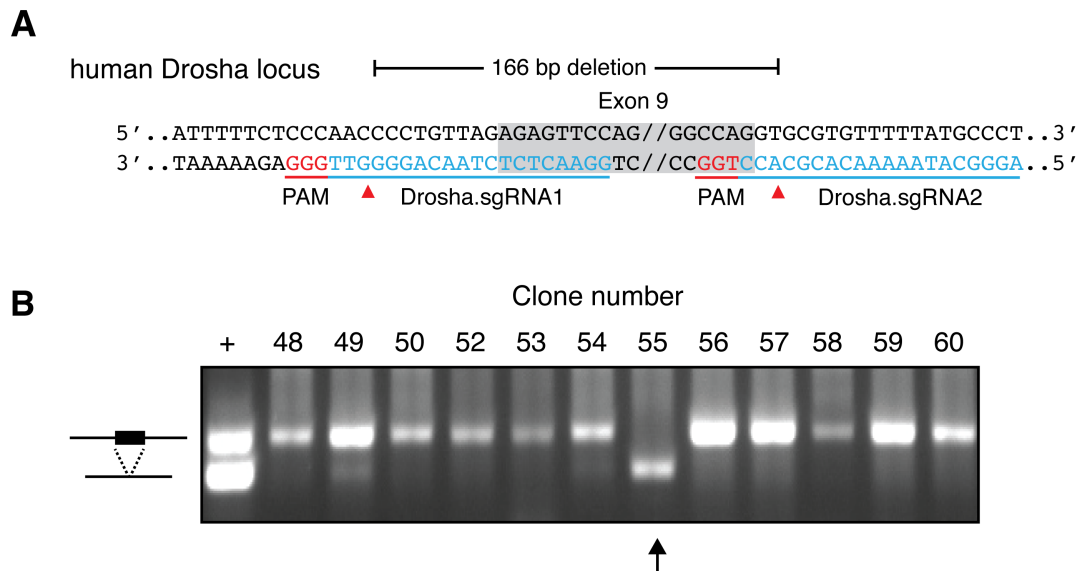


Figure 2.2. A CRISPR based targeting strategy to delete Drosha from Huh-7.5 cells. **(A)** Guide RNAs were designed to delete a 166bp region encompassing exon 9 of the Drosha gene. Approximate cut sites indicated with red triangles. PAM: protospacer adjacent motif. **(B)** Genotyping of CRISPR transfected and expanded clonal cell lines after PCR amplification surrounding exon 9. One clone (#55) exhibited a deletion consistent with exon 9 homozygous removal. Plus sign indicates PCR from bulk transfected cells prior to single cell dilution cloning.

CRISPR mediated deletion of Dicer in Huh-7.5 cells

Initial Dicer deletion was attempted by following a similar scheme as Drosha, to delete a coding exon with an out-of-frame splice to force a frameshift and premature stop. Exons 1 and 2 was successfully deleted independently in this manner, yet full Dicer activity was observed in exon 1 deleted cells, and residual activity in exon 2 deleted cells, likely due to a downstream start codon that resulted in a tolerable N-terminal deletion (Ma et al., 2008)(see Chapter 3). To make a complete knockout, I focused on deleting exon 19 in both WT and in exon 2 deleted cells, to induce a frameshift whereby any downstream translation would make a fragment lacking the helicase domain, the domain of unknown function (DUF), and the Piwi Argonaut and Zwillig (PAZ) domain, all known to be critical for dsRNA and miRNA processing by Dicer (Ma et al., 2012). Guide RNAs were cloned into pX458-pSpCas9(BB)-2A-GFP (pX458, Addgene plasmid #48138). After sequence confirmation, transfection and single cell dilution cloning proceeded as above. Genotyping was performed from QuickExtract (Epicentre) isolated DNA and the Dicer exon 19 locus was PCR amplified for 35 cycles. The resulting PCR products underwent gel electrophoresis, where deleted cells yielded shorter amplicons (Figure 2.3). Two homozygous deletion clones were isolated in WT cells (Δ DicerEx19) and one in exon 2 deleted cells (Δ DicerEx2.19), along with unedited clones. Western and northern analysis of Δ Dicer cells are described in the next chapter. Guide sequences are listed in the Appendix.

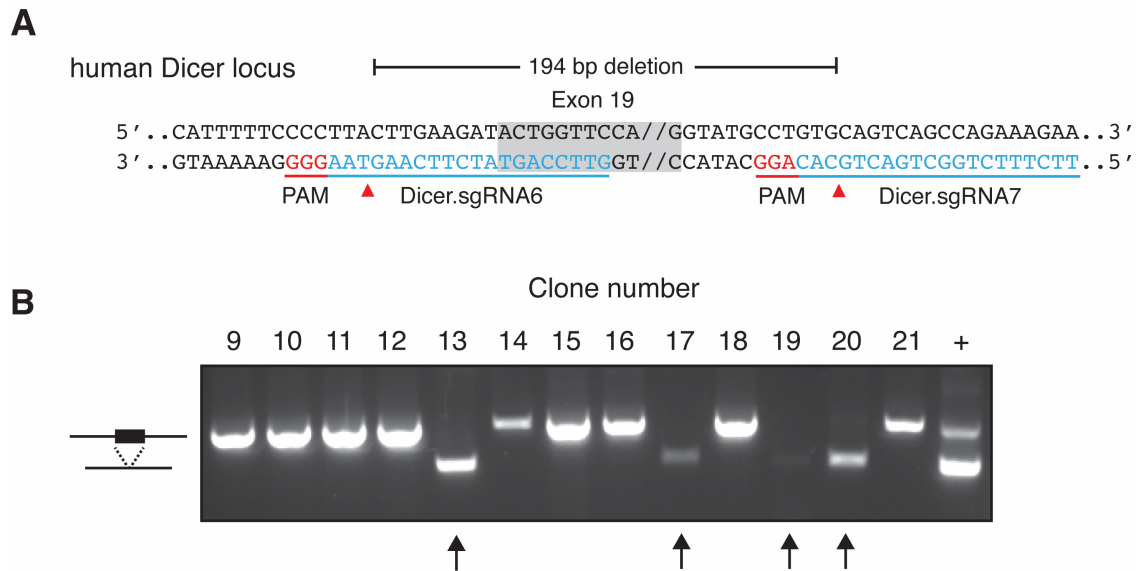


Figure 2.3. A CRISPR based targeting strategy to delete Dicer from Huh-7.5 cells. **(A)** Guide RNAs were designed to delete a 194bp region encompassing exon 19 of the Dicer gene. Approximate cut sites indicated with red triangles. PAM: protospacer adjacent motif. **(B)** Genotyping of CRISPR transfected and expanded clonal cell lines after PCR amplification surrounding exon 19. Four clones (12, 17, 19, 20) exhibited a deletion consistent with exon 19 homozygous removal. Plus sign indicates PCR from bulk transfected cells prior to single cell dilution cloning.

Generation of recombinant HCV plasmids

The selection of which HCV to use in these studies has largely centered on genotype 2 viruses, due to their ability to generate relatively high titer stocks (by historical HCV standards), as well as the ability to generate the large and homogeneously infected population of cells needed for CLIP. Moreover, as the 5'UTR miR-122 sites are completely conserved across all genotypes, results related to miR-122 studies with this virus are likely to be generalizable to other genotypes, though importantly, this has not been directly studied.

pJ6/JFH1-Clone2, pJ6/JFH-Clone2-5AB-Ypet, and pJc1FLAG(p7-nsGluc2A) are fully-infectious HCV non-reporter and reporter viruses respectively, that have been previously described (Catanese et al., 2013a; Horwitz et al., 2013; Marukian et al., 2008). Clone2 HCV is notable for efficient cell-to-cell spreading and is resistant to SR-BI antibody mediated inhibition, in this manner a high proportion of infected cells is thought to be achieved (Catanese et al., 2013a). To construct miR-15 dependent viruses in both backgrounds, we used overlap extension PCR. Briefly, pJ6/JFH1 Clone 2 plasmid DNA was digested with EcoRI and KpnI to yield a 1298nt fragment encompassing the start of T7 transcription, both miR-122 sites, and reading into the E1 coding sequence. A 1.2kb region (Fragment A), upstream of miR-122 sites, was PCR amplified from the parental 1298nt fragment with Accuprime Pfx Supermix (Invitrogen). Approximately 5ng of Fragment A was mixed with 5ng of a primer containing the EcoRI site, T7 start, 5' HCV sequence with miR-15 S1 and S2 sites and 20nt overlap with Fragment A, and the overlap product was amplified with 5μM of EcoRI and KpnI primers at either end. This PCR product was digested with EcoRI and KpnI, religated into the backbone, and sequence verified along the 1.3kb insert. The same strategy was used for miR-15 viruses in the Jc1FLAG2(p7-nsGluc2A) context but with EcoRI and BsiWI, resulting in a 1373nt fragment), 1st PCR primers (Pos43_F and Pos1358_R), and overlap PCR primers (Pos12343_F and Pos1358_R). Similarly, fusion PCR was used to generate pJ6/JFH1-Clone2-U3, -U3S2p3 and -U3S2p3,4 mutants using primers

m122toU3_F and F S2p34, R S2p34, F S2p3 and R S2p3. All cloning oligos are listed in the Appendix.

RNA transcription

In vitro transcripts were generated as previously described (Lindenbach et al., 2005). Briefly, plasmid DNA was linearized by XbaI and purified by using a Minelute column (Qiagen). RNA was transcribed from 1 µg of purified template by using the T7 RiboMAX Express RNA polymerase kit (Promega). Reaction mixtures were incubated at 37 °C for 30min to 1hr, followed by a 15-min digestion with 5U of RQ1 DNaseI (Promega). RNA was purified by using the RNeasy kit with DNaseI (Qiagen), and was quantified by absorbance at 260 nm and diluted to 0.5 µg/µl. Prior to storage at –80 °C, RNA integrity was determined by agarose gel electrophoresis and visualization by ethidium bromide staining.

RNA electroporation

Huh-7.5 cells were electroporated with HCV RNA as previously described (Lindenbach et al., 2005). Briefly, Huh-7.5 cells were treated with trypsin, washed twice with ice-cold phosphate-buffered saline (PBS) (Invitrogen) and resuspended at 1.75×10^7 cells/ml in PBS. Then, 5 µg of each RNA was combined with 0.4 ml of cell suspension and immediately pulsed using a BTX ElectroSquare Porator ECM 830 (820 V, 99 µs, five pulses). Electroporated cells were incubated at room temperature for 10 min prior to resuspension in 15 ml or 30 ml complete medium for non-reporter and reporter constructs, respectively. Resuspended cells were plated into 24-well, 6-well, and P150 tissue culture

dishes for reporter, RNA, and CLIP analysis respectively, at indicated timepoints. For U3 virus electroporations, cells were monitored until greater than 80% of cell were infected, then harvested (around 13-17 days post electroporation). Supernatants from electroporations meant for virus production were serially harvested and replaced starting at 48 hours, every 24 hours, for up to 5 days. Supernatants were filtered using 0.22 μ m filters to remove cell debris prior to concentration either under centrifugation in Amicon Ultracell 100kDa filters (Millipore), or concentration under N₂ pressure using a Stirred Cell ultrafiltration device and membranes (Millipore, Model 8400).

Two color miR-122 target fluorescent reporter construction

Construction of miR-122 fluorescent reporters largely mirrored previous work with miR-20 (Mukherji et al., 2011). In brief, a nuclear localization sequence (NLS: ATGGGCCCTAAAAAGAAGCGTAAAGTC) was appended to the N-terminus TagBFP and TagRFP (Evrogen) open reading frames via overlap PCR. The resulting nlsTagBFP was cloned into the pTre3G-BI vector (Clontech, #631337) between EcoRI and NdeI restriction sites. nlsTagRFP was PCR amplified to add 3'UTRs with N=1 bulged, mutant, or perfectly complementary miR-122 sites before insertion between BamHI and EcoRV sites in the nlsTagBFP containing plasmid. The artificial 3'UTRs for N=4 and N=6 constructs were chemically synthesized as GeneBlocks (IDT) before undergoing Gibson Assembly (NEB) with the BamHI/EcoRV digested pTre3G-nlsTagBFP plasmid, nlsTagRFP, and the artificial 3'UTR. Gibson assembly was also used to

seamlessly append full-length 3'UTRs of cellular targets to nlsTagRFP, after amplifying miR-122 target 3'UTRs from human genomic DNA. While Gibson assembly was initially chosen due to problems with overlapping restriction enzyme sites that could not be resolved with traditional cloning methods, the ease, flexibility and power to "scarlessly" modify a DNA sequence using Gibson assembly quickly made it the method of choice. All cloning oligos are listed in the Appendix.

HCV infection and LNA treatment assays

Infectious units were quantified by limiting dilution titration on naïve Huh-7.5 cells and counted using the median tissue culture infective dose [TCID₅₀] method (Lindenbach et al., 2005) or the focus-forming units (FFU) method (Gottwein et al., 2009). For infection experiments with WT or m15 Clone2 viruses, cells were seeded in 6 well or P150 plates, infected at an MOI of 1-2 the following day, inoculum removed at 6 hours post-infection (hip) and harvested for RNA-Seq or CLIP at 72 or 96hpi. Samples harvested for CLIP from cells electroporated with WT Clone2 RNA were harvested at 48 hours post-electroporation (hpe).

For miRNA and LNA experiments, cells were seeded the day before, and transfected with miRNA (Thermo Scientific) or LNA or miravirsin/SPC3649 (5'-CcAttGTcaCaCtCC-3'; LNA in capitals, DNA in lower case, Exiqon) at indicated concentrations using RNAi/Max (Invitrogen) for 48 hours. For cultures that were subsequently infected, this was done 24 hours later with WT, U3 or m15 viruses

at an MOI of 1-2. Cells were harvested at indicated time points for RNA and flow cytometry, while supernatants were harvested for titer measurements. Replication of Gaussia luciferase expressing HCV genomes was monitored by measurement of secreted Gluc in supernatants using the Renilla Luciferase Assay System (Promega) on a Berthold LB960 luminometer (Bad Wildbad, Germany) or an Omega Fluorostar reader (BMG Labtech). Media was replaced at each time point measured. No significant cytotoxicity was observed from the applied concentrations of LNA and miravirsin/SPC3649, as determined using CellTiter-Glo (Promega).

For two color fluorescent reporter measurements, Huh-7.5 TetON cells were plated at 1.5×10^5 cells/well in 6 well plates. Infection with WT-Clone2-5AB-Ypet virus was carried out 12 hours after plating at an MOI of 3. The inoculum was removed and replaced with fresh media after 6 hours. LNA and miR-122 mimics were transfected at the indicated concentrations 24 hours after infection as outlined above. Media was replaced 48 hours after infection to include $1 \mu\text{g/ml}$ doxycycline (Sigma) and reporter constructs were then transfected at $1.6 \mu\text{g/well}$ using Lipofectamine2000 (Invitrogen) reagent. To harvest, cells were trypsinized, fixed in FACS fixation buffer (0.5% PFA, 1% FBS, in 1xPBS) on ice for 10 minutes, and stored at 4°C in FACS buffer (1% FBS in 1xPBS). A small number of cells were harvested without fixation for RNA analysis.

RNA isolation and qPCR

Total RNA was prepared via Trizol extraction (Life Technologies, Carlsbad, CA), precipitation with ethanol, and yields determined by absorption spectroscopy using a NanoDrop (NanoDrop Products, Wilmington, DE). HCV genomes were quantified using the EraGen MultiCode-RTx method (EraGen Biosciences, Madison, WI) as described previously (Mulligan et al., 2009), and was run on the LightCycler 480 Real-Time PCR System (Roche Applied Sciences, Indianapolis, IN). MicroRNA qPCR was performed using the miScript RT II system (Qiagen) following the manufacturers instructions for absolute quantification. The resulting cDNA was used with the FastStart SYBR Green qPCR system (Roche Applied Sciences, Indianapolis, IN) following the manufacturers instructions and was run on a iCycler/IQ5 (BioRad). Synthetic miR-122, miR-21, miR-15a and miR-196 mimics (Dharmacon) were used to make a standard curves for each miRNA. The miscript RT II kit was also adapted to quantify HCV genomes (for Fig. 1G and Fig. S7) by using a sense HCV specific primer at the 3'end of the genome (JFH1-3UTR: CTGGTCTCTCTGCAGATCATGT). Previously determined amounts of in vitro transcribed HCV RNA were used to generate a standard curve. Per cell RNA amounts were determined by Trizol extracting RNA from 10^6 , 10^5 , and 10^4 Huh-7.5 cells at 60-80% confluence, in duplicate, summing the total amount of RNA recovered, and dividing by the number of cells to reach a measurement of 17.28 ± 4.83 pg RNA/cell. RNA measurements per nanogram were transformed to

per cell measurements with this number. For liver biopsy per cell estimates, we relied upon human hepatocellularity measurements of 139 ± 25 million cells/gram of liver from (Sohlenius-Sternbeck, 2006). Four liver specimens of various weights (2-40mg of tissue) underwent RNA extraction with Trizol. The summed total of extracted RNA per sample was divided by specimen weight to arrive at an RNA/mg of tissue estimate, and transformed by 139000 cells/mg of tissue to reach a measurement of 1.27 ± 0.39 pg RNA/cell. Per nanogram qPCR results were transformed accordingly.

Small RNA Northern analysis

Ten micrograms of Trizol extracted total RNA from Huh-7.5 cells was separated on a 15% acrylamide/7M urea gel. After transfer onto Hybond-N1 membrane (Amersham Pharmacia Biotech), small RNAs were detected with 32 P-end-labeled DNA probes complementary to human miR-122-5p, miR-16-5p, miR-21-5p, and U6 snRNA.

Western blotting

Cells were lysed at the indicated times using 1X PXL (1X PBS, tissue culture grade; no Mg^{2+} , no Ca^{2+} , 0.1% SDS, 0.5% deoxycholate, 0.5% NP-40, sterile filtered), triturated and placed on ice for 15 minutes. 15 μ g of protein lysate was separated on 4–12% Bis/Tris NuPage polyacrylamide gels (Invitrogen). Proteins were transferred to nitrocellulose membranes and proteins of interest were detected using antibodies against NS5A (9E10; 1:5000) (Lindenbach et al., 2005), AGO2 (ab32381, 1.3 μ g/mL, 1:1000) (Abcam), DICER (ab14601, 1:1000)

(Abcam), DROSHA (ab12286, 1:1000) (Abcam) or β -actin (AC15, 1:10000) (Sigma); secondary antibodies were AffiniPure Donkey-anti-Mouse IgG (H+L) (115-035-003, 1:10000) (Jackson Immuno Research), Goat-anti-Rabbit-HRP (31462, 1:10000) (Pierce), Goat-anti-rabbit IRDye 800CW (926-32211, 1:15000) (Licor Biosciences), or Goat-anti-rabbit IRDye 680RD (926-68070, 1:15000), (Licor Biosciences). Western blots were visualized using SuperSignal West Pico (Thermo Scientific) or the Odyssey CLx imaging system (Licor Biosciences).

Fluorescence microscopy

Images were captured on an Axioplan 2 imaging fluorescence microscope (Zeiss, Thornwood, NY) using Metavue Software (Molecular Devices, Sunnyvale, CA). Images were processed using ImageJ software (NIH, Bethesda, MD).

Luciferase reporter assays

Luciferase reporter vectors were cloned by inserting short oligonucleotides or PCR amplified target 3'UTRs into psiCHECK-2 (Promega) using XhoI and NotI. For luciferase reporter assays, 5×10^4 Huh-7.5 cells per well in 48-well plates were transfected over night with 2.56nM final concentration LNA122 (Exiqon) or miR-122 mimic (Thermo Fisher) using RNAi/MAX (Invitrogen). Alternatively, cells were infected with HCV (J6/JFH1-clone2), MOI=3 over night. 24 hrs later, cells were transfected with 1ng/well psiCHECK-2 reporter plasmid using Lipofectamine2000 (Invitrogen) and incubated over night before lysis in Passive Lysis Buffer and evaluation of luciferase levels using the Dual Luciferase Reporter Assay (Promega) on a Omega Fluorostar reader (BMG Labtech).

Flow cytometric analysis

For NS5A flow cytometry, cells were harvested using AccuMax (eBioscience) and fixed using Fixation/Permeabilization buffer (BD Biosciences) for 10 min at 4 °C. Fixed cells were washed with BD Perm/Wash buffer (BD Biosciences), incubated 30 min at RT with AlexaFluor-647-conjugated 9E10 antibody (1:4000 in BD Perm/Wash buffer), washed twice with BD Perm/Wash buffer and once with FACS buffer (PBS/3%FBS) prior to analysis using a BD FACS Calibur and BD FACSDiva software. Analysis was performed using FlowJo software.

Two-color fluorescent miRNA reporter flow cytometry

For fluorescent protein reporter measurements, cells were run on a MACSQuant VYB flow cytometer (Miltenyi Biotec) after fixation to detect TagBFP, TagRFP, and Ypet signals. The raw FACS data were analyzed with FlowJo to gate single, intact cells according to their forward (FSC-A) and side (SSC-A) scatter profiles. HCV positive cells were gated on the basis of Ypet signal above uninfected background. Untransfected cells were used to characterize the cellular autofluorescence in BFP and RFP channels, from which we subtracted the mean plus two standard deviations of the autofluorescent signal for each channel in transfected cells. Cells with BFP and RFP fluorescence levels less than 0 after background subtraction were excluded from further analyses. Data were log-transformed and binned according to BFP levels,

and the mean RFP signal was calculated for each BFP bin. See Chapter 4 for details.

High-throughput sequencing methods

The work presented in this thesis makes extensive use of high-throughput sequencing methods, many customized and invented in the Darnell lab. As such methods are constantly evolving, this section will outline the basic principles of experimental design using these techniques, describe the flow, aims and assumptions of these methods, as well as provide the finer details of their execution.

Experimental designs

The proper design of experiments involving highly sensitive genomics methods is critical for minimizing technical or biologically spurious sources of variation that at best mask the variation between the intended biological variables tested, or at worse mislead to false conclusions. Long known to be an issue in many microarray studies, batch effects represent a major and sometimes unacknowledged source of technical variation in next-generation sequencing (NGS) studies (Leek et al., 2010). To minimize the sometimes impossible task of correcting for a batch effect after-the-fact, CLIP and RNAseq experiments were designed to account for and mitigate batch effects in the following ways:

1. For all cell work, each well or dish from which a library was made is defined as a biologic replicate. One perturbed versus its un-perturbed sample (thus two biologic replicates) constituted one biologic replicate comparison.

2. For all animal and human work, all samples coming from the same donor are defined as a biologic replicate.
3. All cellular perturbations (virus infections, LNA treatments, etc) were paired with un-perturbed controls under identical conditions. Comparisons across controls from different experiments were not used beyond as a measure for cross batch variation.
4. Harvest of paired samples was carried out at the same time, in the same order (perturbed then control), and frozen in -80C. Subsequent CLIP and RNA library prep were initiated on paired samples on the same day.
5. For timecourse measurements, paired samples were frozen at the time of harvest. All subsequent CLIP or RNaseq library prep was initiated for all biological replicates for all time points on the same day.
6. Sample processing was done in separate tubes; where protein gels were concerned, all samples were separated by one empty lane. Where DNA gels were concerned, all samples were split by condition and run on separate gels. In other words, all perturbed biologic replicates were run on one gel, and all unperturbed samples on a separate gel. Samples were separated by one empty lane.
7. Pooling of samples was performed to include both perturbed and unperturbed controls in the same lane of the same sequencing run. Splitting samples across lanes was avoided, unless a measure of technical variation was sought.

Sample generation

As stated above, Huh-7.5 cells were used for all HTS work. Cells were infected with virus or electroporated with viral RNA, the media replaced at 6 hpi or 24hpe, respectively, and harvested at time points indicated in the figures. In all cases, wells with identical numbers of cells were mock infected or mock electroplated with media or PBS.

For miR-122 knockout mouse studies, the livers from five floxed control mice and four liver specific miR-122 knockout mice, were provided by Kalpana Ghoshal at the Ohio State University. Dissected livers were wrapped in foil and flash frozen in liquid nitrogen before being shipped to RU on dry ice.

For human normal versus liver tumor, five samples from patients with HCC (all HCV negative) were flash frozen in liquid nitrogen after autopsy, in accordance with IRB approved protocols at the Ohio State University. Specifically, per patient approximately one gram of tumor and histologically normal adjacent tissue were frozen simultaneously before being shipped to RU on dry ice.

Sample harvest

Cells growing in p150 plates were washed with cold PBS and irradiated over ice once for 400mJ/cm² and once again for 200mJ/cm² using a Spectrolinker XL-1500 (Spectronics Corporation). Cells were then trypsinized (Life Technologies), pelleted, and stored at -80°C until use. One p150 plate would typically yield between 3 and 10 million cells per pellet.

Mouse and human liver tissue was pulverized in a mortar and pestle with liquid nitrogen. A small portion (20-50mg) of the resulting powder was saved for RNA analysis. The remaining powder was crosslinked on a bed of dry ice with three irradiations at $400\text{mJ}/\text{cm}^2$ using a Stratlinker XL-1500 (Stratagene). All samples were then stored in -80°C until lysis.

Argonaute CLIP

Argonaute CLIP was performed generally following previous work (Chi et al., 2009) and is most comprehensively described in (Moore et al., 2014) with modifications listed here. The basic outline for AGO-CLIP consists two parts: from cells to autorad, that is, from cellular lysates to the isolation of AGO associated RNA fragments, and the DNA library preparation steps from small samples of isolated CLIPped RNA. For a schematic view of the standard AGO-CLIP protocol with variations, see Figure 2.4 and Figure 2.5.

Bead preparation

Protein A Dynabeads (Invitrogen) were washed 3x and resuspended in antibody binding buffer (AB: PBS, 0.02% Tween-20). Per p150 dish or liver sample, 100 μl of beads was used. Beads were rotated with 12.5 μl of bridging antibody (rabbit anti-mouse IgG, Jackson ImmunoResearch) per 100 μl of beads for 30 minutes at room temperature. Beads were then washed 3x with AB before adding 1 μl of the pan-AGO antibody (2A8, ascetic provided by Dr. Zissimos Mourelatos and described in (Nelson et al., 2007)) per 100 μl beads. For IgG only antibody controls, AGO antibody was omitted. Beads were rotated at for 30 min

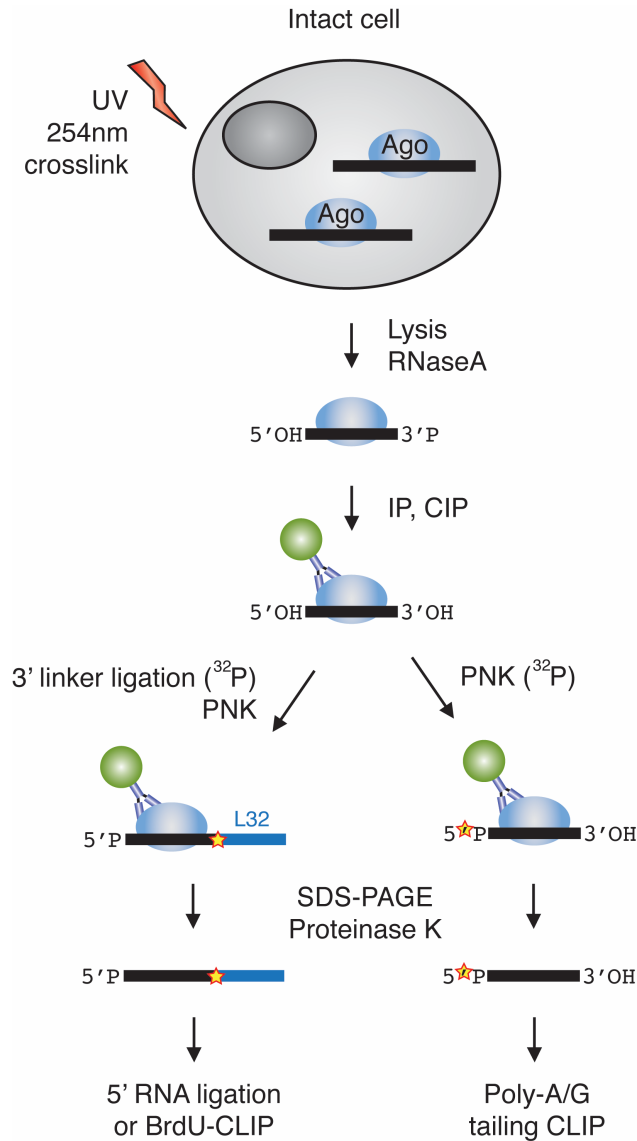


Figure 2.4. Ago-CLIP schematic from cells to Ago associated mRNA isolation. Cells or tissues are crosslinked, undergo lysis and partial RNase digestion before undergoing IP using anti-Ago (2A8) antibody coated magnetic beads. Following extensive wash steps, protein bound RNAs are enzymatically manipulated to enable down stream sequencing either by ligating a radiolabeled 3' RNA linker (left side) or direct radiolabeling with PNK (right). Radioactivity is indicated with a star. Following SDS-PAGE and proteinase K treatment, isolated Ago bound RNA is cloned. For RNA with 3' linkers, cloning proceeds with 5' linker ligation or BrdU-CLIP. For RNA without linkers, poly-A or -G tailing CLIP is implemented.

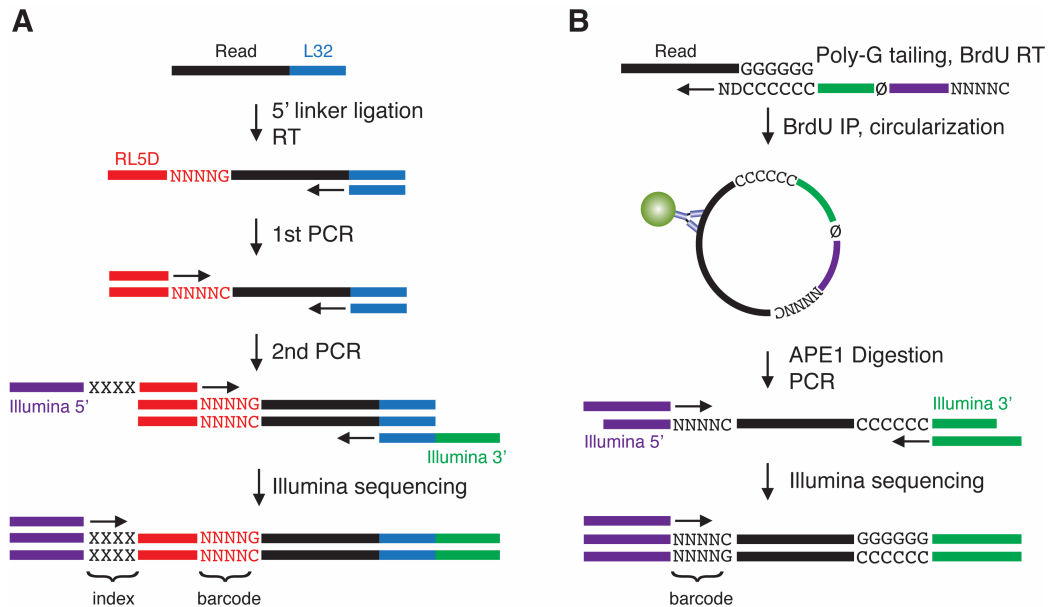


Figure 2.5. Two cloning strategies for Ago-CLIP isolated RNA fragments. **(A)** In standard CLIP, RNA reads undergo sequential ligation of 5' and 3' RNA adapters, followed by RT-PCR. Note that the RNA input contains the 3' adapter, as this is done on bead prior to SDS-PAGE where the radiolabeled 3' RNA linker is used for auto-radiography. The 5' adapter contains a barcode (NNNNG) such that each RNA tag is uniquely labeled, thus cataloging a single Ago binding event. An index specific for each library is added with a 2nd PCR step to enable multiplexing. Reads are sequenced starting with the index. **(B)** In Poly-G (or A) CLIP, completely independent of RNA ligation, reads are first tailed with a homopolymer polymerase and undergo RT using an oligo that recognizes the poly-G stretch, and using BrdUTP in place of UTP. This RT oligo contains Illumina sequencing adapters separated with an APE1 site. The cDNA undergoes BrdU IP, on bead circularization, APE1 digestion, then is PCR amplified to create the library. As in standard CLIP, each RNA read is tagged with a unique barcode.

at room temperature or overnight at 4°C. Just prior to IP, beads were washed in 1X PXL lysis buffer (1X PBS tissue culture grade without magnesium or calcium, 0.1% SDS, 0.5% Sodium-deoxycholate, 0.5% NP-40) with protease inhibitors (Roche, mini EDTA-free).

Cell lysate preparation and IP

Frozen lysates of crosslinked cells or powdered liver tissue were prepared by adding 1ml of 1X PXL (with protease inhibitors) and triturating to disrupt cells. Lysates treated with 10 μ l DNase (RQ1, Promega) for 5 minutes, thermomixing at 37°C at 1100rpm. Samples were then treated with RnaseA or RNaseI, first diluted to the indicated concentration by volume (e.g. 1:100, or 1:10,000) in lysis buffer and then added at 10 μ l per ml of lysate. Critical for interpreting CLIP autorads, the amount of RNase to be added was empirically determined in pilot experiments using an RNase titration. Range concentrations that yielded a smear 10-50kDa above the collapsed band were typically chosen, in some cases no RNase was required. Figure 2.6 shows an example of an RNase titration yielding a working concentration amenable for the partial RNase digest needed for CLIP. Lysates underwent thermomixing again for 5 minutes at 37°C at 1100rpm before being treated with 10 μ l RNasin RNase inhibitor (Promega) and spun at 4°C on max speed of a table-top micro centrifuge for 30 minutes. Supernatants, along with any lipid layer, were harvested and mixed with PXL equilibrated antibody bound beads for IP. A small amount of pre-IP lysate was kept to monitor Ip efficiency. Samples were rotated with beads at 4°C for 2-4 hours. Following IP, a small amount of post-IP lysate was kept. Beads were washed sequentially twice each with 1X PXL, 5x PXL (same as 1X but using 5X PBS), and 1X PNK buffer (50mM Tris-HCl, pH 7.5, 10mM MgCl₂, 0.5% NP-40).

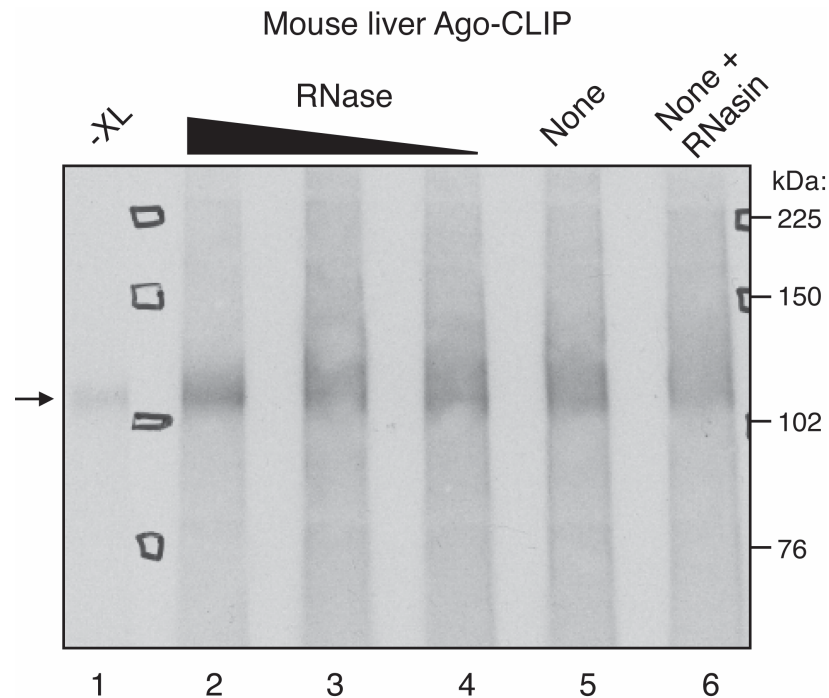


Figure 2.6. An RNaseA titration experiment for Ago-CLIP. Two mouse livers were processed for Ago-CLIP using 3' radiolabeled L32 linker under decreasing RNaseA concentrations (lanes 2-4), without RNaseA (lane 5), or without RNaseA plus RNase inhibitors (lane 6). Minus crosslink control from similar amount of tissue is shown (lane 1). Ago:miRNA complex at 110kDa is highlighted with an arrow. Note that as RNase concentration decreased the sharp 110 kDa band became more diffuse, indicating an RNase dependence "collapse" of the radioactive signal to the modal size of Ago plus miRNA. On the basis of this experiment for mouse liver, lane 4, corresponding to 1:10K dilution of RNase A was chosen for downstream processing, and became the RNaseA condition of choice for liver tissue studies in mouse and human.

Removal of 3' phosphate, 3' linker ligation, and re-phosphorylation of 5' ends of AGO bound RNA

To prevent RNA circularization of RNA tags, IPs were treated with alkaline phosphatase. Beads were resuspended in 80µl containing 1x dephosphorylation buffer, 3U of CIAP (Roche), RNasin inhibitor (Promega), and thermomixed for 20

minutes at 37°C, shaking at 1100rpm for 15s every 2 minutes. Samples were then washed as above sequentially in 1X PNK, 1X PNK plus 20mM EGTA, and twice with 1X PNK.

Radiolabeled linkers for 3' linker ligation per sample were set up by preparing a T4 phosphonucleotide kinase (PNK, NED) reaction according to the manufacturers instructions using 2.5µl ³²P-γ-ATP and an RNA linker, either 20pmol L32 or L34 RNA linker (Dharmacon or IDT) and incubating for 30 minutes at 37°C. To drive the reaction to completion, 0.5µl of 1mM ATP was added and incubated for an additional 5 minutes. Linkers were purified from free nucleotides using G-25 columns following the manufacturers instructions (GE Healthcare). Ligation reactions on bead were prepared using Rnl1 RNA ligase (Fermentas) and 12pmol radiolabeled linker following the manufacturers instructions. Samples were incubated for one hour at 16°C, shaking at 1100rpm for 15s every 4 minutes. After this hour, 60pmol cold linker was added and the reaction allowed to go overnight. The next morning, beads were washed twice each with 1XPXL, 5X PXL, and 1XPNK. A final PNK reaction was performed on bead to restore the 5' phosphate. To each sample an 80µl mix containing 1mM ATP, 1X T4 PNK buffer, 4µl T4 PNK and RNasin inhibitor. Samples were incubated for 20 minutes at 37°C, shaking at 1100rpm for 15s every 2 minutes.

SDS-PAGE resolution of AGO:RNA complexes

Beads were washed twice each with 1XPXL, 5X PXL, and 1XPNK. Protein was eluted off the beads by incubating with in 30µl of 1X LDS loading buffer

(Invitrogen) with reducing agent for 10 minutes at 70°C, shaking at 1100rpm. Supernatants were run on Novex NuPAGE 8% Bis-Tris gels (Invitrogen) in SDS-MOPS buffer (50mM MOPS, 50mM Tris Base, 0.1% SDS, 1mM EDTA, pH7.7) for 2.5 hours at 175V at 4°C. Radiolabeled protein RNA complexes were transferred to BA85 nitrocellulose (Whitman) using a Novex wet transfer apparatus for 1 hour at 30V. After transfer, the membrane was rinsed with RNase-free PBS, and exposed to Biomax MR film (Kodak) at -70°C typically from 3 hours to up to 5 days.

Recovery of AGO-RNA complexes

Nitrocellulose membranes were aligned with the exposed film and regions of the membrane from low RNase IP lanes were excised. Typically, binary AGO-miRNA complexes migrated at 110kDa, while AGO-mRNA and ternary AGO-miRNA-mRNA complexes migrated between 130-150kDa. These two regions (110kDa and 130-150kDa) were excised and processed separately, though I noted no qualitative differences between the ability to clone miRNAs or mRNA targets from either region. Put another way, mRNAs could often be cloned from the 110kDa band, and miRNAs from the 130-150kDa bands. Parallel processing of these regions was often used as technical replicates. In general, the regions excised did not extend beyond 150kDa, due to the large length of the RNA recovered. Short mRNA fragments between 30 and 60nts were preferred to facilitate read clustering and peak finding; I note this only because most NGS

RNA protocols call for longer (100-300nt) reads, to more effectively leverage paired end sequencing and more accurate splice site identification.

Proteinase K digestion, RNA recovery, and 5' linker ligation

RNA was liberated from membrane fragments using 200µl of a proteinase K solution (Roche) diluted in PK buffer (100mM Tris-HCl, pH 7.5, 50mM NaCl, 10mM EDTA) to 4mg/ml and incubated for 20 minutes at 37°C, shaking at 1000rpm. RNA was further denatured by adding 200µl of PK buffer with 7M urea and incubating for 20 minutes at 37°C, shaking at 1000rpm. RNA fragments underwent acid phenol:chloroform extraction and were precipitated overnight at -20°C in 1ml of 1:1 ethanol:isopropanol and 50µl of 3M NaOAc, pH 5.2.

RNA was pelleted by spinning at max speed (>13,000rpm) in a table top centrifuge at 4°C, and washed twice with 75% ethanol. Following speedvac drying of the RNA, the pellet was dissolved in 6µl RNase-free water. The 5' linker was ligated using Rnl1 (Fermantas) and 20 pmol of the 5' linker following the manufacturers instructions and in a total volume of 10µl. Samples were incubated at 16°C for 5 hours with intermittent shaking. Ligated RNA was DNase digested by adding 30µl RNase-free water, 5µl 10X RQ1 buffer, 2.5µl RQ1 DNase and 2.5µl RNasin inhibitor; and incubating at 37°C for 20 minutes. Samples then underwent phenol chloroform extraction as above.

Reverse transcription and 1st PCR

RNA was pelleted by spinning at max speed (>13,000rpm) in a table top centrifuge at 4°C, and washed twice with 75% ethanol, and dried. Following

speedvac drying of the RNA, the pellet was dissolved in 10µl RNase-free water. 8µl RNA was mixed with 10pmol DP3 primer, 3µl 3mM dNTPs, and incubated for 5 minutes at 65°C in a thermocycler; the remaining RNA from all samples was pooled and 8µl of this was used for the -RT control. Samples were ramped down to 50°C, after which RT was added: 1µl 0.1M DTT, 4µl 5x Superscript buffer, 1µl RNasin and 1µl Superscript III (Invitrogen) or 1µl water for -RT controls. Samples were incubated in a thermocycler at 50°C for 45 minutes, 55°C for 15 minutes, 90°C for 5 minutes then chilled to 4°C. PCR was performed immediately after using 27µl Accuprime Pfx (Invitrogen), 0.15µl DP5 primer (at 100 pmol/µl), 0.15µl DP3 primer (at 100 pmol/µl) , and 4µl of the RT reaction, for a total of 4 reactions per sample. PCR conditions were 95°C for 2 minutes, 20-35 cycles of 95°C for 20s denature, 58°C for 30s anneal, and 68°C for 30s extension. Typically, miRNA plus linker fixed products arose around 20-24 cycles, mRNA plus linkers from 22-30 cycles. Reactions were taken out the the thermocycler following extension starting at 22 cycles, every two or three cycles. RT negative samples were taken out 4-5 cycles after the last +RT sample.

The entire PCR reaction was loaded with an equal volume of 2x loading buffer (95% formamide, 5% 100mM EDTA pH 8.0, and a dash of bromophenol blue and xylene cyanol) and run on 10% denaturing PAGE gels in 1X TBE for 1 hour at 300V. Amplify molecular rulers (Biorad) were used as markers. To visualize DNA, gels were stained in 1X SYBR Gold (Molecular Probes) in 1X TBE for 10 minutes. PCR products corresponding to miRNAs flanked by adapters

(~55-60bp) or adapter flanked mRNAs (80-150bp) were cut from the gel. DNA was extracted by soaking gel slices in DNA diffusion buffer (0.5M ammonium acetate, 10mM Mg-acetate, 1mM EDTA, 0.1% SDS) and incubating at 50°C for 30 minutes, shaking at 1200rpm. Gel slurry was filtered through Whitman glass filters in Nanosep columns (VWR). DNA was recovered from the filtrate using Qiaquick gel purification (Qiagen) and resuspended in 30µl.

Multiplexed AGO-CLIP

To enable cost-effective multiplexing of standard CLIP libraries as well as compatibility with sequencing on MiSeq machines, we adapted the 2nd PCR step of the standard CLIP protocol to add sequencing adapters and 5' indices. This strategy uses the DP5 and DP3 sequences of the 1st PCR product as priming sites to add 5' indices and 3' adapters for a short (4-10 cycles) 2nd PCR step (Figure 2.5A). PCR was performed using 27µl, Accuprime Pfx (Invitrogen), 0.5µl MSFP5 5' primer (at 20 pmol/µl, each sample with different indexed primer), 0.5µl MSFP3 3' primer (at 20 pmol/µl) , and 3µl of the 1st PCR product, in triplicate. PCR conditions were 95°C for 2 minutes, 4, 7 or 10 cycles of 95°C for 20s denature, 58°C for 30s anneal, and 68°C for 30s extension. The 2nd multiplexed PCR product was separated on a 2% Metaphor agarose gel (Lonza) and the lowest cycle number visible purified Qiaquick gel purification (Qiagen) and resuspended in 30µl. Following DNA concentration and integrity analysis using TapeStation (Agilent), each individual sample was diluted down to 10nM, pooled

in equal parts with all other samples containing unique indices, and submitted for high throughput sequencing. All cloning oligos are listed in the Appendix.

Poly-G CLIP

Poly-G CLIP is a direct adaptation of the single linker ligation BrdU CLIP protocol (Weyn-Vanhentenryck et al., 2014) and was inspired by ribosomal profiling methods (Ingolia et al., 2009). In contrast to standard CLIP, which relies on two sequential RNA linker ligation steps, and BrdU CLIP, which relies on one RNA ligation step, the poly-G CLIP protocol relies on tailing CLIP'ed RNA, performing RT with a primer capable of circularization and subsequent cleavage to generate a cDNA library amenable for high-throughput sequencing (Figure 2.5B).

Regarding the choice of nucleotide for tailing, we chose G for a few practical but mostly empirical reasons. As poly-A tails can range in size from >100nt to >1kb, working out conditions for uniform poly-A tail lengths became impractical. Aldo Mele in the Darnell lab took an unbiased approach and tested two polyA polymerases, from *E. coli* or *S. cerevisiae*, and monitored their ability to generate homopolymeric tracts that could be used for priming. ATP yielded long tails for both enzymes as expected, but only yeast polyA polymerase created short, 12-15nt G tails on any input RNA (Figure 2.7). Moreover, the yeast enzyme most efficiently added tails to completion such that little input RNA was left untailied. On this basis of these experiments, I adapted Aldo's findings to aid in miRNA cloning efforts using polyG tailing.

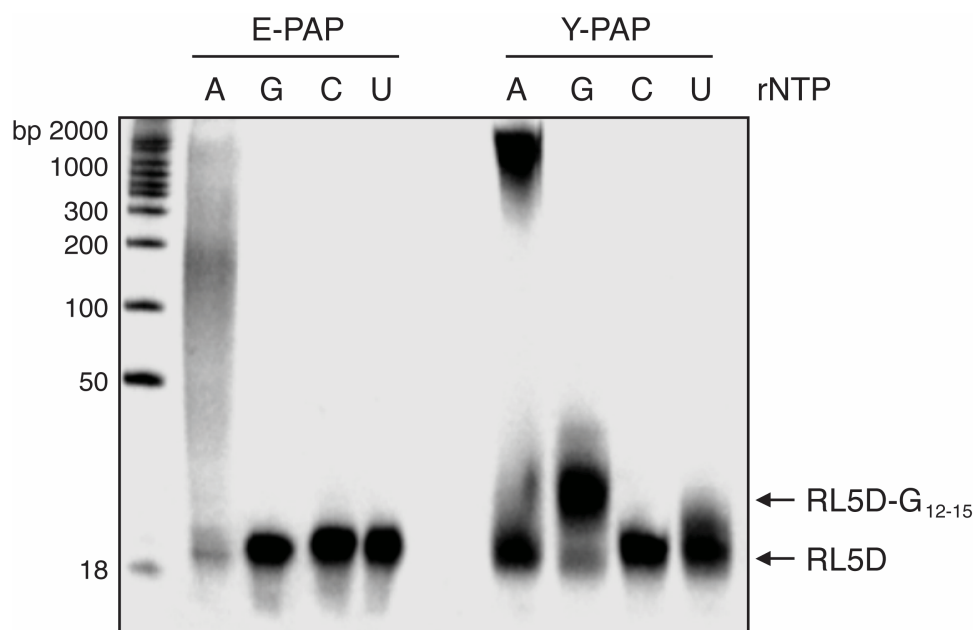


Figure 2.7. Ribonucleotide incorporation comparison between poly-A polymerase enzymes from *E. coli* (E-PAP) and *S. cerevisiae* (Y-PAP). Exactly 20 pmol RL5D RNA linker was radiolabeled with ^{32}P - γ -ATP in a PNK reaction and underwent poly A polymerase reactions with 5U E-PAP (left, NEB) or 300U Y-PAP (right, USB) in manufacturer recommended buffer conditions supplemented with 1mM of the indicated rNTP. Reactions proceeded for 20 minutes at 37°C, before stopping via heat denaturation at 65°C for 10 minutes. Samples were run on 20% denaturing Urea-PAGE gels and exposed to film. E-PAP readily incorporated ATP as expected forming products with median polyA lengths of 150-200nt, and was unable to incorporate an other rNTPs. Y-PAP synthesized polyA tails greater than 1kb in length and was largely incapable of incorporating UTP or CTP, but interestingly incorporated short well defined tails of polyG, between 12-15nt in length. Experiment performed by and data courtesy of Aldo Mele.

The standard CLIP protocol was followed to directly PNK label AGO bound RNA with ^{32}P - γ -ATP (Moore et al., 2014); the 3' RNA linker ligation steps were skipped. Following autorad exposure, 100-110kDa regions were excised from the membrane for RNA isolation as described above. The resulting RNA was tailed with yeast poly-A polymerase (USB Affymetrix 74225Y) using GTP with the final concentrations: 1X Y-PAP buffer, 1mM GTP, 7.5U RNAsin

(Promega N261), and 300U Y-PAP. The reaction was incubated at 37°C for 20 minutes, then 65°C for 10 minutes. After tailing, reverse transcription was carried out using 5µl polyG tailed RNA, 1µl 0.752M Tris, 1µl 8.2mM dATP, 1µl 8.2mM dCTP, 1µl 8.2mM dGTP (all Invitrogen), 1µl 8.2mM Br-dUTP (Sigma), 1µl 25µM RT Primer and 1µl H₂O. This mix was incubated for 3min at 75°C then ramped down to and held at 48°C. To this reaction, 1µl 82mM DTT, 1µl 10U/µl RNasin and 1µl SuperscriptIII (or H₂O for –RT) pre-warmed to 48°C was added to the original mix, then incubated for 45 minutes at 48°C, 15 minutes at 55°C, and 5 minutes at 85°C then held at 4°C. Following RT, 1µl of 2U/µl RNaseH (Invitrogen) was added to destroy RNA and incubated at 37°C for 20 minutes, after which 10µl RNase-free H₂O was added and the cDNA purified by spinning through a G-25 column.

cDNA purification and BrdU IP

The cDNA was then purified using Protein G dynabeads (Invitrogen) coupled with an anti-BrdU antibody (Santa Cruz). Briefly, 50µl of ProteinG dynabeads per sample were washed in AB buffer and resuspended in 25µl of AB buffer along with 25µl 50X Denhardt's Solution (Sigma) and rotated at room temperature for 1hr. Beads were then washed three times with 1X IP buffer (0.3X SSPE (Ambion), 1mM EDTA, 0.05% Tween-20), resuspended in 25µl IP buffer plus 25µl of anti-BrdU antibody (5µg, Santa Cruz, sc-32323) and rotated at room temperature a minimum of 45 minutes for antibody binding, then washed and equilibrated in 1X IP buffer. The equivalent of 25µl starting volume of beads were

used per BrdU IP. The volume of cDNA was brought up to 40µl with H₂O and 10µl 50X Denhardt's solution before adding 50µl of 2X IP buffer (2X IP buffer:0.6X SSPE (Ambion), 2mM EDTA, 0.1% Tween-20). The mix was incubated for 5 minutes 70°C, 2 minutes 25°C, spun down and added to the equivalent of 25µl starting volume of the prepared anti-BrdU beads. Tubes were rotated at room temperature for 30mins and washed once with 1X IP buffer plus 5X Denhardt's, twice in low salt buffer (15 mM Tris- HCl, pH 7.5, 5 mM EDTA), and twice with 1X IP buffer. cDNA was eluted from the beads via BrdU competitive elution by adding 50µl 100µM BrdU (Sigma) in 1X IP buffer, rotating for 30min and collecting the eluate. Eluted cDNA was purified by spinning through a G-25 column, the volume adjusted to 97.5µl with H₂O, after which 37.5µl 4X IP buffer and 15µl Denhardt's solution was added. The anti-BrdU IP was repeated as above with the remaining 25µl of prepared beads, and washed. The final two washes consisted of 1X CircLigase Wash Buffer (33mM Tris-acetate, 66mM KCl, pH7.8).

On-bead cDNA Circularization and Apel Linearization

On bead, cDNA was circularized by incubation for 1 hour, 60°C, 1300rpm thermomixer mixing for 15 seconds, every 30 seconds, with 2µl CircLigase 10X Reaction Buffer (Epicentre), 4µl 5M betane, 1µl 50mM MnCl₂, 0.5µl CircLigase ssDNA Ligase II (50U, Epicentre) and 12.5µl H₂O. Beads were washed by rotating for 5min, two times each with 1X IP buffer and Apel buffer (50mM potassium acetate, 20mM Tris-acetate, 10mM magnesium acetate, pH7.9).

cDNA was then linearized on bead by adding 2µl 10X NEB Reaction Buffer 4, 1.25µl ApeI (10U/µl NEB) and 16.75µl H₂O, and incubating for 1 hour, 37°C 1300rpm thermomixer mixing for 15 seconds, every 30 seconds. Beads were washed twice with low salt buffer, and twice with Phusion wash buffer (50mM Tris, pH8.0).

PCR amplification

To elute cDNA off of beads, 10µl 5X Phusion HF Buffer (NEB), 1µl 10mM dNTPs and 37.5µl H₂O were added, the mix transferred to a thin walled PCR tube and incubated at 98°C for 45 seconds. Supernatants from beads after agent capture were transferred to a fresh PCR tube. To this was added 0.5µl 20µM P5 primer, 0.5µl 20µM P3 primer, 0.5µl Phusion DNA polymerase and 0.5µl 50X SYBR Green I (Invitrogen).

PCR amplification was carried out on iQ5 or CFX real-time PCR machines (Biorad) in order to monitor amplification, with the samples being removed when the RFU signal reached ~800-1000. PCR cycle conditions were as follows: 1X 98°C 30s, 15-20 cycles as necessary 98°C 10s, 60°C 15s, 72°C 20s. PCR products were run on 2% metaphor agarose gels (Lonza), purified using Qiaquick gel purification (Qiagen) and resuspended in 30µl. Unlike standard CLIP, each RT primer contained its own index for multiplexing, and thus only one PCR reaction was ultimately performed to append Illumina sequencing linkers. Following DNA concentration and integrity analysis using TapeStation (Agilent), each individual sample was diluted down to 10nM, pooled in equal parts with all

other samples containing unique indices, and submitted for high throughput sequencing. All cloning oligos are listed in the Appendix.

mRNA-seq library construction

mRNA-seq libraries were prepared from Trizol extracted RNA following Illumina TruSeq protocols for poly-A selection, fragmentation, and adapter ligation. Alternatively, total RNA underwent RiboZero purification (Epicentre) to negatively select ribosomal RNA sequences. Multiplexed libraries were sequenced as 100nt single-end or paired-end runs on HiSeq-2000 sequencers.

Illumina sequencing

CLIP library next-gen sequencing was carried out on HiSeq2000 (RU genomics Core) or Hiseq2500 (NYGC) machines, set up for 100nt single-end (SE) sequencing and typically run between 2-5pM per lane, using v2 or v3 chemistry. Unless noted, the Illumina Read 1 primer was used. For RNAseq, samples were processed under 100nt paired-end sequencing, using standard Read1 and Read2 primers.

Bioinformatic analysis

Argonaute-CLIP

Analysis of AGO-CLIP data was carried out similar to previous work (Chi et al., 2009; Moore et al., 2014; Riley et al., 2012a). Data processing was carried out using the Galaxy suite of bioinformatics tools (Goecks et al., 2010) and the UCSC genome and table browsers (Kent et al., 2002), in addition to in-house tools developed by Chaolin Zhang. A schematic of data processing for CLIP is

presented in Figure 2.8. Using FASTX tools in Galaxy, FASTQ sequence files were first filtered for quality such that 80% of the read contained a mean PHRED score of 20. Reads were then collapsed on sequence to remove sequencing and PCR duplicates. The fold coverage, that is the raw reads divided by the collapsed reads, as a rough estimate of sequencing saturation, was typically as low as 5 but up to 100 fold depending on the presence of multiplexed samples and was used as a rough proxy for sample complexity.

In theory, a CLIP-seq library from an mRNA binding protein can be thought of as a sparsely sampled subset of a full mRNA-seq library since RNABPs by definition bind transcribed RNA. Thus it follows, especially for RNABPs with defined binding sites, and for which exon exon junctions play a minor role, that CLIP-seq libraries will be less complex than their parent RNA-seq libraries, due to the reduced sequence "space" that can be cloned. As a consequence, one should expect that fold coverages on average will be higher for CLIP than RNAseq. In practical terms this means that a far greater number of CLIP experiments can be sequenced per lane than RNAseq experiments. For instance, in singlet libraries, a Hiseq2000 lane will typically yield approximately 100M reads, for which AGO-CLIP will collapse to around 1-2M reads, a coverage of 50-100X; the parental RNAseq library would collapse to between 10-30M reads (3-10x coverage). Ten multiplexed AGO-CLIP libraries in the same scenario on average will yield around 1M reads per sample (10M total) for a per sample coverage of 10X, in line with one RNA-seq library.

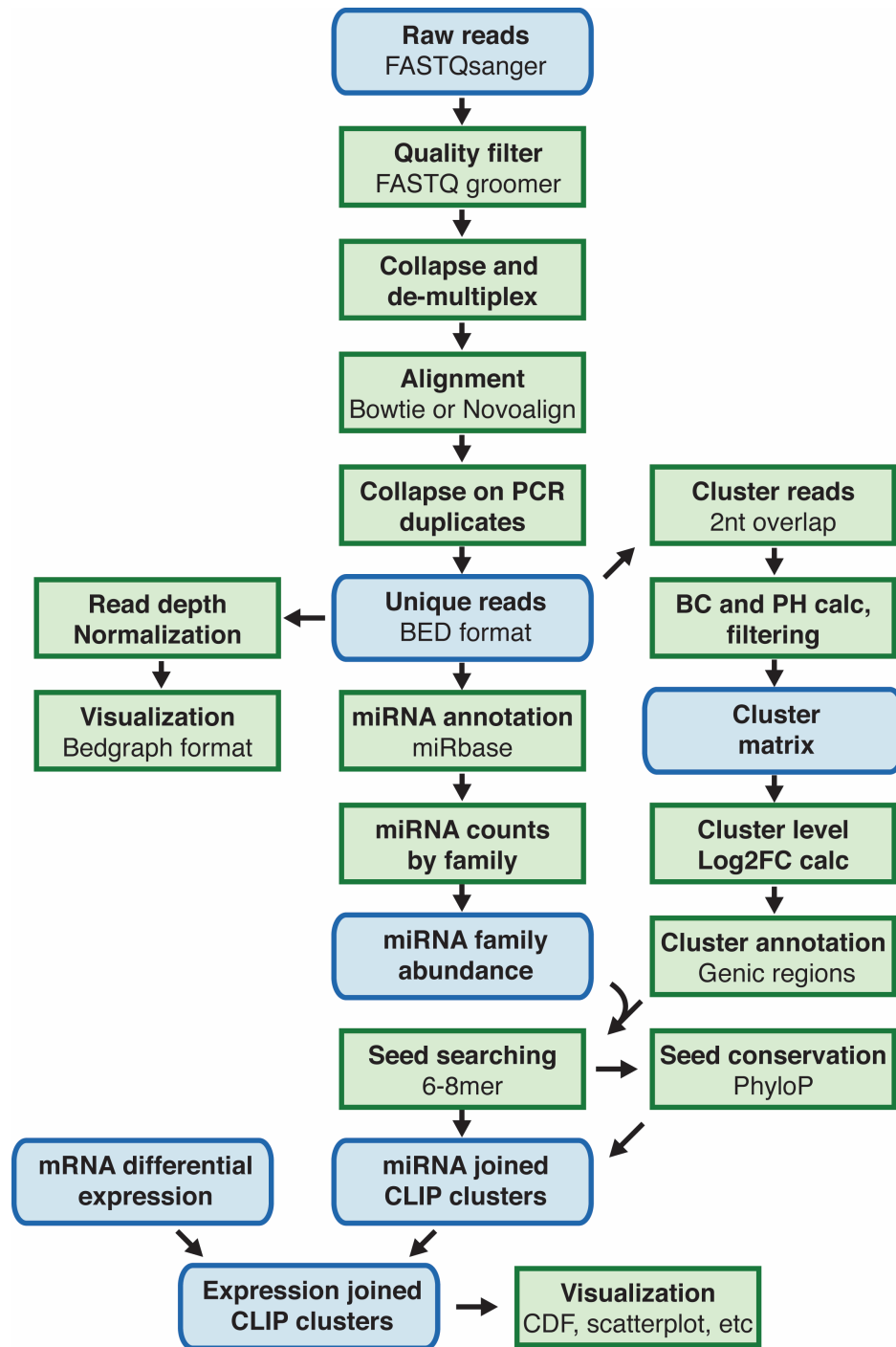


Figure 2.8. An analysis pipeline for CLIP data. Key raw and finished files in blue denoted by format where applicable. Data processing, calculation, and visualization steps in green boxes.

Barcoded samples, if any, were split after exact sequence collapse using the barcode splitter. 5'-degenerate linker sequences were removed for later estimation of unique binding events, where applicable (Licatalosi et al., 2012). The 3'-adapter sequence was also removed if present using the clip adapters tool while for poly-G tailed samples, an "adapter" sequence of GGGGGGGG was clipped from the 3' end. Reads shorter than 18nt were discarded for linker ligated samples, 15nt for poly-G samples. This final set of groomed reads represented the input for subsequent alignment.

Alignment of mRNA sample reads to either the human (hg18) or J6/JFH1-Clone2 (WT, U3 or m15) genome was carried out using the Bowtie (Langmead et al., 2009) or Novoalign (Novocraft) mapping programs allowing at most 2 mismatches (or indels) and discarding reads with multiple hits. Alignment of miRNA samples was performed similarly, but allowing for zero mismatches. After mapping, we collapsed reads on coordinates such that only those reads with sufficiently different degenerate linkers were kept; this distinguished unique binding events (tags) from PCR duplicates (Darnell et al., 2011; Licatalosi et al., 2012). Coordinates for mature miRNAs were constructed from mirBase (v18) for annotation and counting of miRNAs based on uncollapsed mapped reads.

Clustering of mRNA reads to genomic loci was carried out as in (Chi et al., 2009) where we typically specified a minimum biologic complexity (BC) of at least half of the libraries in the comparison (3 for 6 libraries, 5 for 10, etc). Reads at a particular cluster from different experiments were normalized to the read-depth of

their respective libraries for cross comparison. The resulting clusters were intersected with UCSC genes (a union of RefSeq, Uniprot, and Genbank gene definitions) to determine the genic identities and positions of AGO binding sites. Seed searches were carried out within robust AGO clusters (\pm 32nts) for all miRNA families identified from miRNA-CLIP specifying perfect base pairing for miRNA seeds to targets following 8mer, 7mer-A1, 7mer-m8, and 6mer(2-7) pairing rules (Bartel, 2009).

\log_2 fold changes comparing conditions were calculated as follows. A pseudo-count of 1 was added to all summed clusters per condition to allow for incorporation of conditions with zero clustered reads. After binning by BC, the sum total number of reads per cluster per condition was normalized to the read depth for that condition. Dividing these normalized CLIP abundance measurements per condition, we arrived at estimates over mock or controls for change in CLIP binding due to HCV, LNA and KO conditions.

Significant peaks on HCV RNA were called based on using scan statistics as described previously (Licatalosi et al., 2012). Gene wise p-values underwent Bonferroni multiple test correction where the significance cut-off on HCV RNA was set at 0.001. CIMS analysis on viral RNA was performed as described previously (Zhang and Darnell, 2011).

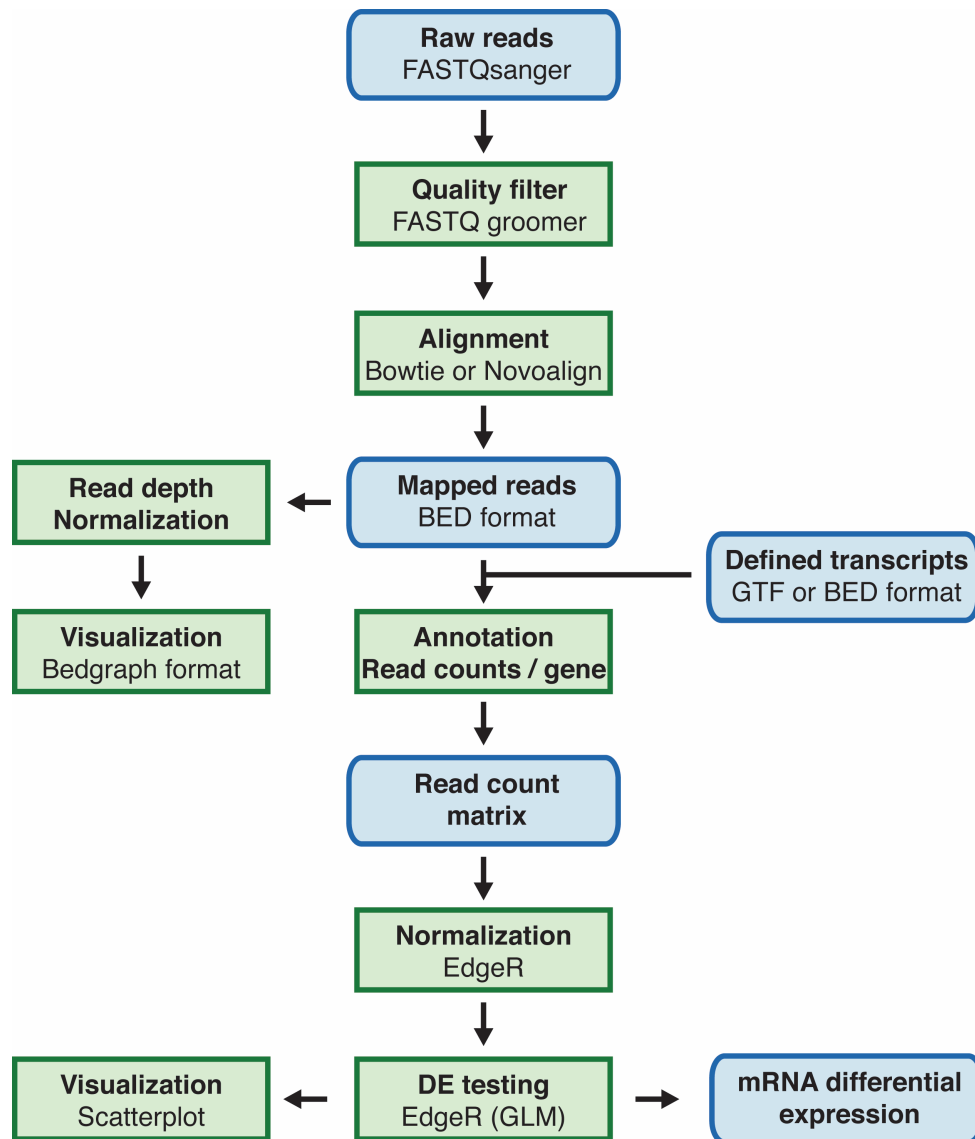


Figure 2.9. An analysis pipeline for RNAseq data. Key raw and finished files in blue denoted by format where applicable. Data processing, calculation, and visualization steps in green boxes.

RNA-Seq bioinformatics

Figure 2.9 outlines the pipeline for RNAseq analysis. Analysis of these data was carried out by first filtering FASTQ reads such that 80% of the read contained a mean score of 20. Reads were then directly mapped to hg18 using

Bowtie, allowing up to 2 mismatches and discarding reads with multiple hits. Mapped read coordinates were then intersected with a meta-transcript file of the longest isoform of every coding gene in UCSC genes and then counted. The resulting reads per gene count, for each RNA-Seq replicate then underwent statistical analysis and quantification using EdgeR (Robinson et al., 2009).

Conservation analysis

Interval coordinates of miRNA seed target sites within CLIP clusters were intersected with PhyloP scores (Pollard et al., 2010) for available mammalian nucleotides from the UCSC table browser. The resulting per-nucleotide scores were averaged across the core 6mer. The same procedure was performed for miRNA seed target coordinates from TargetScan6.2 (Lewis et al., 2005) predictions.

Statistical tests

Statistical analysis was carried out in Graphpad Prism or in R. Two-sided KS-Test were employed for CDFs to test both for goodness of fit and a difference in means. Mann Whitney U test (non parametric t-tests) were used for datasets for which no normality assumption could be made, Student's t-test were used otherwise. Where multiple ad-hoc or post-hoc comparisons were made, care was taken to account for multiple testing correction using one-way ANOVA and bonferroni correction.

Meta-analysis of published array data

Microarray datasets from the chimpanzee miravirsen study (Lanford et al., 2010) were processed by computing the \log_2 fold change across each gene for each paired chimp array (post- versus pre-treatment) and then averaged. Liver biopsy datasets were obtained from Gene Expression Omnibus (GSE15331, (Peng et al., 2009); GSE14323, (Mas et al., 2009)) and processed with GEO2R to remove non-median centered datasets. For GSE15331, a total of 96 HCV positive datasets representing 24 individuals in technical quadruplicate were compared to 20 HCV negative datasets representing 5 individuals in technical quadruplicate. For GSE14323, a total of 19 normal liver datasets were compared to 41 datasets from HCV infected patients with pre-malignant cirrhosis but no hepatocellular carcinoma. The above processed datasets from three independent studies were intersected with all conserved 8mer and 7m8 miR-122 or miR-15 predictions from TargetScan, or with CLIP derived miR-122 targets (7-8mers) from at least two miR-122 perturbation conditions.

Gene Ontology analysis

Gene ontology analysis was performed with unranked lists using the DAVID suite of bioinformatics tools (Huang et al., 2008) or the GOrilla suite for GO term enrichment (Eden et al., 2009). Gene list comparisons were carried out between the list of interest and a background set of AGO bound transcripts.

Conservation analysis

Interval coordinates of miRNA seed target sites within CLIP clusters were intersected with PhyloP scores (Pollard et al., 2010) for available mammalian nucleotides from the UCSC table browser. The resulting per-nucleotide scores were averaged across the core 6mer. The same procedure was performed for miRNA seed target coordinates from TargetScan6.2 (Lewis et al., 2005) predictions.

Statistical tests

Statistical analysis was carried out in Graphpad Prism or in R. Two-sided KS-Test were employed for CDFs to test both for goodness of fit and a difference in means. Mann Whitney U test (non parametric t-tests) were used for datasets for which no normality assumption could be made, Student's t-test were used otherwise. Where multiple ad-hoc or post-hoc comparisons were made, care was taken to account for multiple testing correction using one-way ANOVA and bonferroni correction.

Quantitative modeling of miR-122 sponging by HCV RNA

Our mathematical model expands upon a previously developed model for miRNA regulation (Mukherji et al., 2011) by adding the behavior of HCV and its interaction with miR-122. The mathematical model we developed describes the concentrations of a target mRNA species (r), an HCV mRNA species (h), and binding of miRNA (m) to form complexes with target or HCV mRNA species, respectively (r^* , h^*), see Figure 2.10.

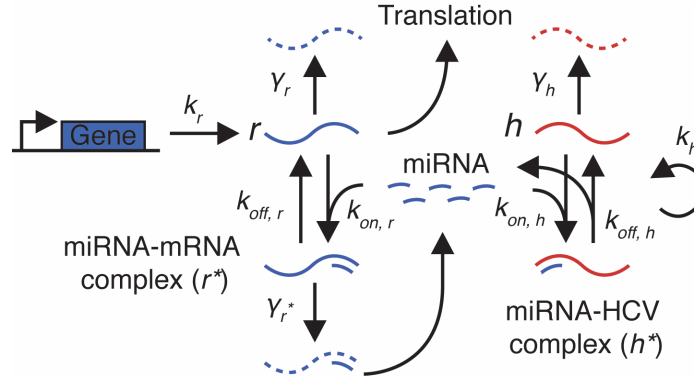


Figure 2.10. Illustration of model reactions for miR-122 dynamics, including transcription and translation of a target mRNA, binding to miR-122 and decay of mRNA species. HCV RNA can replicate, be degraded, or bind miR-122, functionally sequestering miR-122 and leading to de-repression of mRNA targets. See text for details.

Production of a target mRNA comes from transcription of a gene at a rate (k_r) with a corresponding degradation rate (γ_r). We assume the miRNA-mRNA complex (r^*) does not undergo translation and degrades at a rate allowing for recycling of the miRNA species into the pool. The total amount of miRNA is assumed to be constant and can bind to either target mRNA (r) or HCV mRNA (h). HCV RNA also decays at a particular rate (γ_h). We assumed a model where HCV RNA is only degraded in its unbound form, consistent with previous data (Li et al., 2013d; Shimakami et al., 2012).

From the schematic presented in Figure 2.10, from the mass-action equations for this system, we have the following set of equations:

$$\dot{r} = k_r - k_{+c}r \cdot m + k_{-c}r^* - \gamma_r r$$

$$\dot{h} = -k_{+h}h \cdot m + k_{-h}h^* + k_h h^* - \gamma_h h$$

$$\dot{r}^* = k_{+c}r \cdot m - k_{-c}r^* - \gamma_{r^*}r^*$$

$$\dot{h}^* = k_{+h}h \cdot m - k_{-h}h^*$$

where the total pool of miRNA is defined be:

$$m_T = m + r^* + h^*$$

Initially we performed numerical simulations of the full model, however to provide insight into the sponging effect of HCV we made simplifications to the model for analysis. Free miRNA (m) can be bound to either target mRNA (r) or HCV mRNA (h). At equilibrium, the amount of miRNA bound to the two species is proportional to the relative binding strengths as given by rearranging the equilibrium relations for these two species. Thus, the proportion of miRNA bound to each target is given by: $r^* / h^* = K_r / K_h$, where K_h (k_{+h}/k_{-h}) and K_r (k_{+r}/k_{-r}) are the equilibrium constants for the binding-unbinding reactions.

To simplify further, if we assume that the proportion of unbound miRNA is low in comparison to miRNA bound to HCV or mRNA target, then our conservation relation becomes $m_T = r^* + h^*$ and we can estimate the amount of miRNA bound HCV RNA combining the above relations to be:

$$h^* = \frac{m_T}{1 + \frac{K_r}{K_h}}$$

Solving for steady states in our full model as in Mukherji et al., the solution for r is given by the equation below:

$$r = \frac{1}{2} [r_0 - \lambda - \theta + \sqrt{(r_0 - \lambda - \theta)^2 + 4\lambda r_0}]$$

where,

$$r_0 = \frac{k_r}{\gamma_r}, \quad \lambda = \frac{\gamma_{R^*} + k_{off}}{k_{on}}, \quad \theta = \frac{\gamma_{R^*}}{\gamma_R} m_T$$

With HCV RNA present, the difference in this relation is a reduction in the total miRNA pool (m_T) such that the parameter theta is changed to:

$$\theta = \frac{\gamma_{R^*}}{\gamma_R} [m_T - h^*] = \frac{m_T}{1 + \frac{K_h}{K_r}}$$

Examining this relation, it is evident that as the binding strength of miR-122 to HCV RNA gets stronger, the available miR-122 pool is reduced. As the HCV binding strength approaches zero, then all of the original m_T pool is available. We can also approximate the effect of the number of binding sites and stoichiometry by assuming that N binding sites produces a similar effect to a single binding site with $N \cdot$ the binding strength. Simulating and fitting our data to values theta, we

used this formula to estimate the reduction in the pool of miRNA's caused by HCV mRNA.

Data fitting was initially performed on the $N = 4$ construct by simultaneously fitting the lambda and theta parameters to the steady-state equation for $r = f(r_0)$ above using a least-squares error fitting algorithm. In Figure 4.24C, lambda was scanned and fit to the different experimental N constructs' curves. In contrast to Mukherji et al., we found there to be inherent basal expression differences between the variable number of miR-122 sites as observed in the LNA data curves. We corrected the basal expression differences, which resulted in the alignment of the data sets at high BFP expression. In Figure 4.24 D-E, fitting was performed while changing only the value of theta representing a change in the available free miRNA but not in the binding strength parameter lambda. For $N = 4$ curves in both cases, we noted experiment-to-experiment variability in theta fits, though the relative changes within each experiment were consistent. For all fits, the inherent signal bias between RFP and BFP was corrected by using the difference in a group of approximately 10 data points at the high signal end for each curve, since they are expected to fall on a line of $x=y$.

Chapter 3: Argonaute HiTS-CLIP studies of HCV RNA

This chapter describes the overall scheme of studying HCV interactions with small RNAs using the AGO-CLIP method, and considers how these functional genomics techniques were used to address many questions in HCV RNA virus molecular biology in an unbiased manner. As CLIP permits the empirical observation of AGO binding sites with minimal technical perturbation, and is adaptable for *in vivo* work, we will define the interaction landscape between AGO and HCV RNA during infection in Huh-7.5 cells. Following this, we will demonstrate how AGO-CLIP profiles change in response to a variety of perturbations to include time, perturbing or removing miRNAs, utilizing miRNA independent viruses, and swapping the miRNA tropism of the virus. The impact on virus biology will be considered throughout.

The positive interaction of miR-122 with HCV RNA presents a unique outlier for miRNA function as it contradicts the canonical function of miRNAs as repressors of gene expression. A major goal of the current work is determining whether the miR-122 interaction with HCV RNA represents the predominant small RNA interaction with the viral genome, or is one of many. Moreover, with the HCV genome harboring multiple and conserved miR-122 sites beyond those in the 5'UTR, the timing and amount of binding at these and potentially other sites offers a window on the dynamics of small RNA usage by the viral genome.

Experimental setup

The first task for AGO-CLIP in virus infected cells is deciding the number of cells needed per CLIP sample. The initial AGO-CLIP paper obtained results from miR-124 transfected HeLa cells in 10cm dishes, which assuming 80% confluence at the time of harvest would correspond to $\sim 8 \times 10^6$ cells per sample (Chi et al., 2009). A similar cell number was selected for work done in Clone2 HCV infected Huh-7.5 cells, initially starting from electroporation, and later switching to infection. Importantly, independent electroporations or infections were performed, both in separate plates, but also on separate days and always with paired mock treated samples.

The experimental setup, as mentioned in the methods, was to electroporate HCV RNA or infect with virus approximately 5×10^6 cells per sample in a 15cm dish, allow for growth for between 48 and 72 hours, when most cell are infected and proceed with CLIP. Post electroporation, we measured infection frequency by flow cytometry of NS5A expression or by using the IPS-RFP reporter cells (Jones et al., 2010) and observed 80-90% infection frequency that was stable between 48 and 72 hours (Figure 3.1).

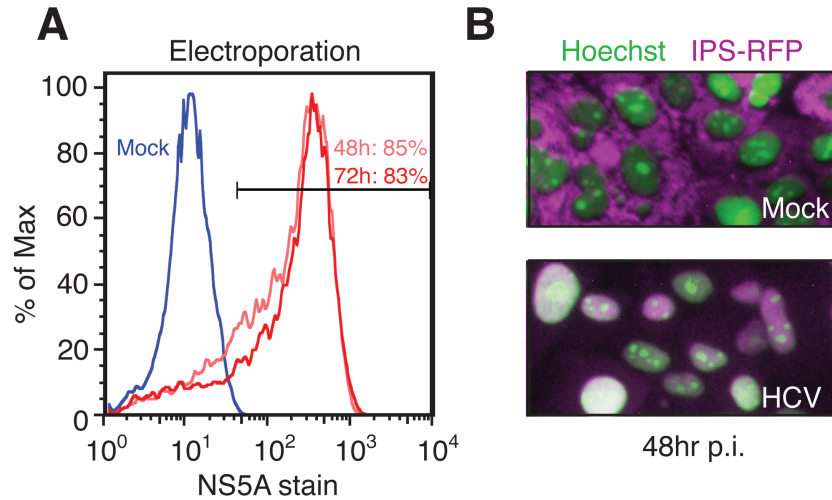


Figure 3.1. Infection frequency measurements for Huh-7.5 cells infected with Clone 2 HCV. **(A)** FACS analysis of NS5A stained cells at indicated time points post-electroporation, showing >80% infection frequency. **(B)** IPS-RFP cell reporter analysis of HCV infection, showing >90% infection frequency. In these cells an RFP with an NLS sequence is fused to the mitochondrial targeting domain of IPS1 (aka MAVS). HCV NS3/4A protease expression cleaves IPS-RFP, thus relocalizing fluorescence to the nucleus (Jones et al., 2010). In general, this assay for measuring HCV replication is slightly more sensitive than NS5A staining, due to the enzymatic nature of NS3/4A. Consequently, slightly higher infection frequency (or at the very least protein translation) can be measured with this method compared to NS5A FACS. Overall, both methods approximately agree.

Following crosslinking and harvest, cells were processed for CLIP as described in the methods, using a pan-AGO antibody that recognizes all four known mammalian AGOs (Nelson et al., 2007). A typical autoradiogram is shown in Figure 3.2A. Here, ^{32}P -labelled RNA crosslinked to AGO was run on an 8% Bis-Tris gel, transferred onto nitrocellulose and exposed for 2 days. An IgG antibody was used as a non-specific antibody control. Additionally, a no crosslink control is shown where only a faint band around 105-110kDa is apparent. This would correspond to tightly associated AGO:miRNA complexes (Lima et al., 2009). UV crosslinking results in a dramatic increase in signal, where importantly, 130-150kDa "smears" corresponding to AGO:mRNA or AGO:miRNA:mRNA complexes appear alongside a more apparent 110kDa band. High RNase conditions tend to collapse the signal to the model size of the protein plus a protected RNA fragment, though this is not always the case. In this experiment, regions from mock or HCV infected lanes were cut corresponding to 110, 130 or 150kDa as a first size selection step for the isolation of AGO bound miRNAs and mRNAs.

As a measure of IP efficiency, pre- and post-IP lysates were probed via western blot to monitor the extent of AGO depletion (Figure 3.2B). In both uninfected and infected cells, approximately 60% of total AGO protein was depleted in lysates, indicating that the majority of AGO:RNA complexes were recovered for CLIP.

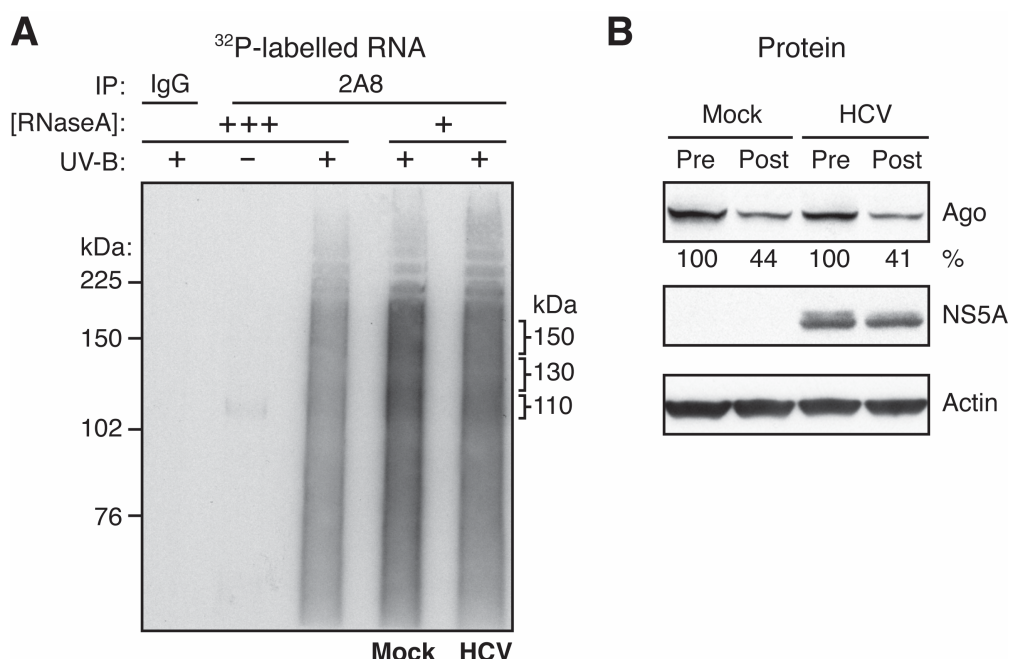


Figure 3.2. Autoradiogram of AGO bound RNA and measure of AGO depletion. **(A)** Autoradiogram of ³²P-labelled RNA bound to AGO after IP using the pan-AGO 2A8 antibody. AGOs migrate at 97kDa, which shifts by RNA species bound by approximately 1kDa/3nt, thus AGO bound to ~22nt miRNA should be around 105kDa. IgG used as a non-specific control. Minus crosslink control shows the stringency of washes in that only a faint Ago:miRNA complex remains. High RNase and low RNase conditions indicated. Regions excised corresponding to 110, 130 or 150kDa are highlighted. **(B)** Pre and post IP lysates were probed for Ago, NS5A and Actin. Percent of Ago depletion from mock or HCV infected cells is indicated and are based on normalizing to actin.

Excised membrane regions underwent proteinase K treatment to liberate the RNA, which then underwent 5' adapter ligation and RT-PCR for DNA library generation (see Methods). The "first" PCR provides the next size selection step after the autorad, and yields a key measure of success or failure for library cloning (Figure 3.3). Here, the cDNA derived from each membrane fragment (110,130 or 150kDa) underwent PCR for various cycle numbers. Importantly,

care was taken to catch the earliest cycle number for which an amplified signal could be detected on a 6M Urea-PAGE gel, to preserve the maximum library complexity and minimize the long known non-uniformity of PCR amplification of cDNA libraries (Patanjali et al., 1991). Adapter flanked miRNAs of ~60nt could be observed with as few as 20 cycles of PCR. Target mRNAs typically migrated as "smears" owing to the partial RNase digest, ran between ~80-180nt, and were observed in as few as 24 cycles. In general, PCR from the 110kDa membrane fragments tended to yield miRNA+adapter sized products, though this was not always the case, as some mRNA smears could be observed. 130-150kDa regions almost always yielded mRNA "smears."

After isolation of miRNA or mRNA "smears," a second PCR step was performed to append 4mer indices (one per sample) to enable multiplexing, and to add Illumina sequences for NGS. This PCR step was typically very short (between 4-10 cycles). In the end, libraries with miRNA sized fragments were ~150nt; mRNA between ~175-250nt. These libraries were concentration normalized, pooled such that each index was represented once, and submitted for high-throughput sequencing.

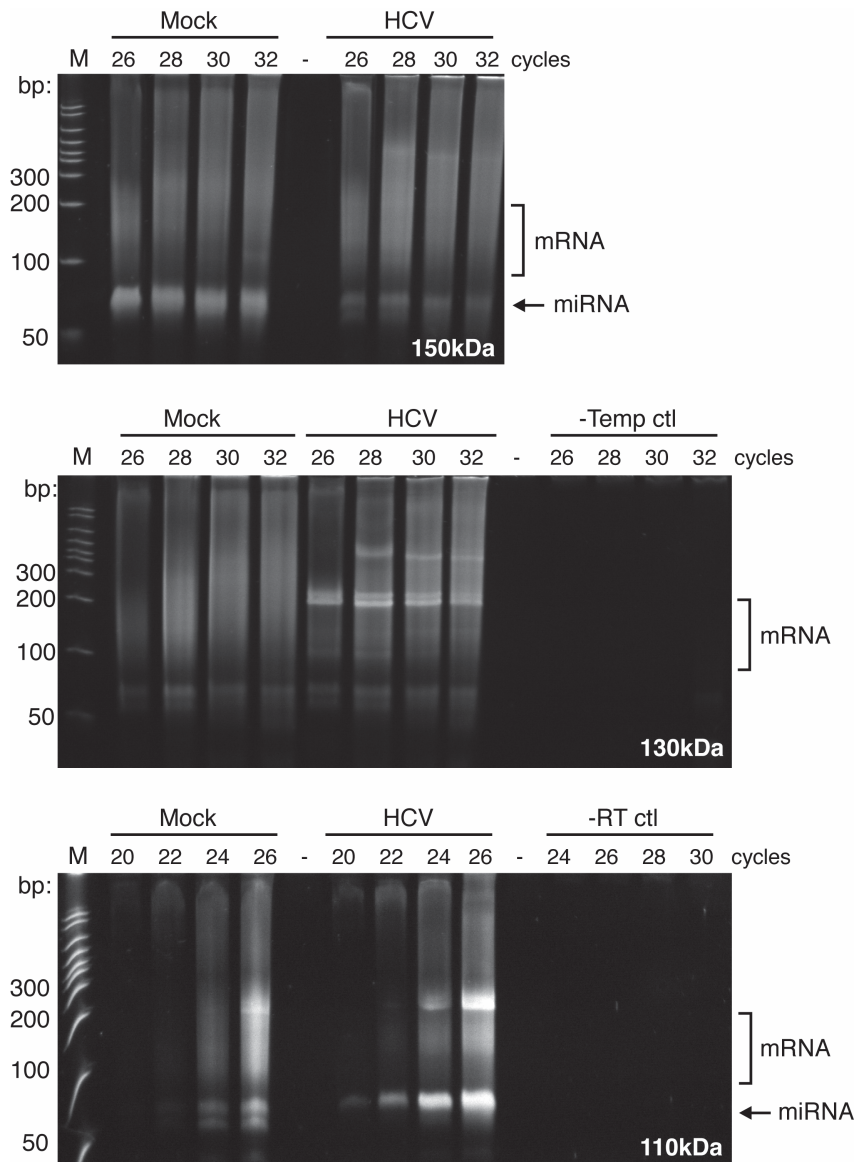


Figure 3.3. 1st PCR results from cloning AGO bound RNAs. Each panel depicts the results for excised membrane fragments (110, 130 or 150 kDa) after PCR amplification for indicated number of cycles. Minus reverse transcriptase (-RT) and minus template (-Temp) PCR controls are shown. Bands corresponding to miRNA + adapters (~60nt) and mRNA targets + adapters" smears" (~80-180nt) are indicated. Note that while 110kDa bands predominantly yield miRNA sized products, this is not exclusive. mRNA smears are more common in 130 and 150kDa bands, where miRNAs can also be seen. Regions selected for 2nd PCR multiplexing steps were chosen such that no "bandiness" was evident for mRNA smears.

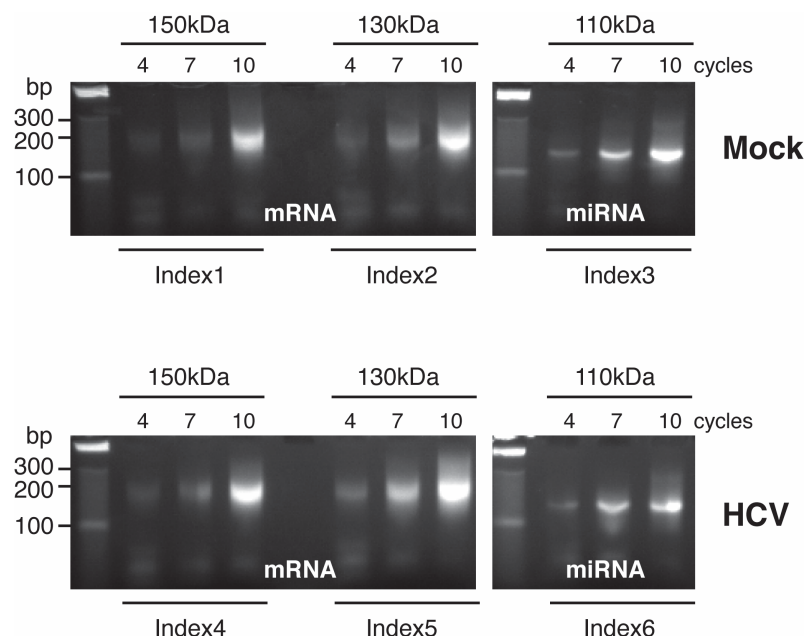


Figure 3.4. Multiplex adapter 2nd PCR results from cloning AGO bound RNAs. After amplification with multiplexed primers at indicated cycle number, mRNAs from 150 and 130kDa regions yielded products between 180 and 200bp as expected. Cloned miRNAs plus adapters from 100kDa regions yielded a sharp band at ~150bp. Libraries from Mock or HCV infected libraries were qualitatively identical at this stage.

An AGO binding map of HCV RNA confirms extensive miR-122 engagement

Three matched experiments were prepared and sequenced in the above manner, resulting in a total of six libraries. After processing reads for quality and trimming adapters, reads were mapped on to the input Clone2 HCV RNA genome, and collapsed on the basis of a degenerate linker sequence introduced during RNA linker ligation (see Methods). Thus, each resulting unique read ("tag") can be considered an individual AGO binding event. Approximately 1-2% of input reads mapped to the HCV genome. The resulting AGO binding map is presented in Figure 3.5, where tags from independent experiments are

delineated by color and are displayed across a linear representation of the HCV genome.

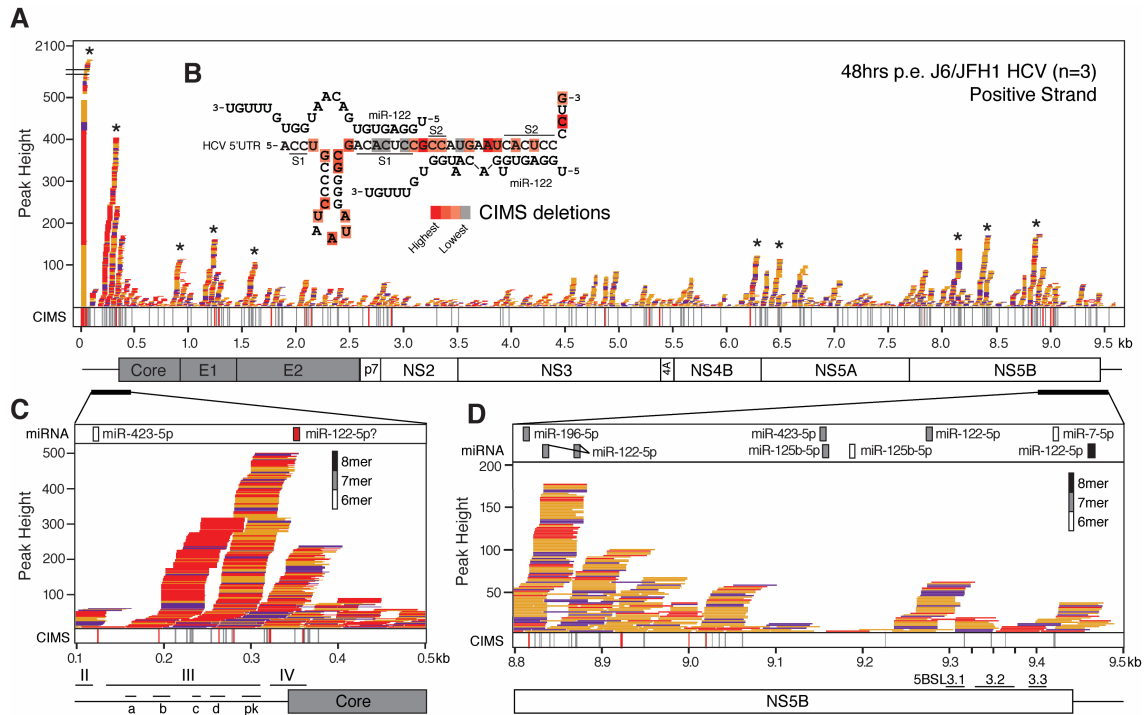


Figure 3.5. An Argonaute binding map on HCV RNA. **(A)** Mock-subtracted binding map of Ago-CLIP reads across HCV genomic RNA, from three biological replicates, denoted by read color. Statistically significant peaks above a uniform background are indicated by asterisks. Bottom CIMS track shows location of all deletions (gray) and statistically significant CIMS deletions (red). **(B)** Schematic of a miR-122 binding model to S1 and S2 highlighting locations of CIMS deletions. **(C)** Zoom in view of Ago binding from (A) across the viral IRES into the coding sequence. IRES domains (II-IV), associated stemloops (a-d) and the pseudoknot (pk) region are indicated. Upper track displays seeds for the top 50 miRNA seeds from CLIP data, seed types denoted by colors. A proposed non-canonical miR-122 site is highlighted in red (Pang et al., 2012). **(D)** Zoom in view of Ago binding from (A) across the NS5B peaks. Known stem loops in the NS5B coding sequence are indicated.

Not displayed is the negative strand or alignments to mock infected cells, where in either case a vastly low percentage of reads mapped to HCV (not shown). This

not only presents a relief in terms of experimental success (no HCV mapping reads in the mock), but it also gives a sense of the high strand asymmetry exhibited by AGO binding. After removing reads mapping to the polyU tract (which is polyA on the negative strand and upon which mapped reads are indistinguishable from Illumina sequencing artifacts), >98% of reads mapped to the positive strand, and no peaks for the remaining low number of reads were observed on the negative strand. This would be consistent with AGO exclusively binding the positive strand, and playing little role, if any, on anti-sense genomes. Moreover, it also likely indicates that no RNAi-like response is occurring, given that this would imply Dicer activity on dsRNA in a manner that would generate siRISCs that target both strands of HCV.

By far, AGO binding to positive strands at the extreme end of the 5'UTR, overlapping with S1/S2 miR-122 sites, and is the most frequent binding event on the HCV genome, accounting for 50% of all HCV mapped tags (Figure 3.5A). The high sequencing depth achieved at S1/S2 sites enabled nucleotide resolution of AGO crosslink sites, based on the observation that RT tends to skip crosslinked amino-acid-RNA adducts and thus results in deletions at the site of crosslinking (Zhang and Darnell, 2011). The resulting crosslink induced mutation analysis (CIMS) revealed a statistically significant enrichment of deletions in areas of close proximity to, but not within, the miR-122 S1 and S2 binding sites (Figure 3.5B). Interestingly, the most significant crosslink site was at the G in position 28 of the HCV genome. A recent report described G28A mutations as natural HCV

variants resistant to miR-122 antagonism: the G28A mutant was attenuated in WT cells, but could replicate in cells with reduced miR-122 levels (Israelow et al., 2014). While the significance of this crosslink site remains to be determined, one possibility is that AGO binding to HCV RNA is tuned by contacts outside of the seed regions which structurally remain flexible in the AGO binding pocket in the absence of base-pairing with the miRNA (Schirle et al., 2014). A G28A mutation might favor tighter miR-122 binding given that the opposed base on miR-122 is a U. Moreover, its position, between S1 and S2 auxiliary pairing may be relevant for the simultaneous AGO:miR-122 occupancy, though the precise contribution is unknown.

Returning to the CLIP map, additional peaks of AGO binding were observed. To identify which peaks were "significant", we employed scan statistics to test if a peak was above a hypothetical background where all tags were uniformly distributed across the HCV genome (Darnell et al., 2011; Licatalosi et al., 2012). In this manner, the S1/S2 peak was significant as were peaks in the IRES, E1, E2, NS5A and NS5B regions of the genome.

The second largest peak in the viral IRES specifically overlapped with the pseudoknot and the coding start site (Figure 3.5C). We noted no canonical 7mer or 8mer binding site for the top 50 miRNA families expressed in these cells, opening the possibility that some type of non-canonical or miRNA independent AGO binding is occurring. Interestingly, a non-canonical miR-122 site in the IRES was proposed based on SHAPE analysis of the HCV 5'UTR with miR-122 *in vitro*

(Pang et al., 2012). Our results with AGO could be consistent with miR-122 engagement. Alternatively, as the 5'UTR is highly structured, the rather broad IRES peak could be due to crosslinking with AGO:miR-122 bound at S1/S2, but on the outer surface of the AGO protein. This would imply that the IRES peak is dependent on AGO binding to S1/S2, which is an idea that we will return to. Additionally, a role for AGO in regulating IRES translation, however minor, cannot be ruled out (see below).

Along with S1 and S2 binding, four additional miR-122 sites are known on the HCV genome, notably in the NS5B coding sequence and in the 3'UTR (Jopling et al., 2005; Nasheri et al., 2011). Consistent with previous data, no significant AGO binding was observed on the 8mer miR-122 site in the 3'UTR (Figure 3.5D). The largest peak outside of the 5'UTR, in NS5B, overlapped with two miR-122 7mer seeds, one of which has been shown to be very slightly repressive (Nasheri et al., 2011). While reports have suggested numerous miRNAs that interact with HCV RNA (reviewed in (Gupta et al., 2014)), among these we only observed expression of let-7 and miR-196 families in our miRNA-CLIP data (see next chapter), and no seeds from either of these families were observed within significant Ago binding peaks. Taken together, these data suggest that miR-122 constitutes the predominant miRNA interaction with HCV RNA in these cells, and is largely confined to the 5'UTR.

We can only speculate as to the nature of the low-level amount of AGO binding across the HCV coding sequence. Given the high abundance of HCV

RNA in Huh-7.5 cells, it's possible that the low level binding reflects AGO-CLIP "noise" of rare binding events that are likely to survive filtering because they uniquely map. Still, it remains possible that some biology is at play. AGO was initially cloned as eukaryotic initiation factor 2C (EIF2C), which remains the root name for AGOs 1-4 (Zou et al., 1998). As the name suggests, the initial phenotype for AGO suggested a role in translation, specifically in the ternary complex formation between eIF2, Met-tRNA, and GTP, known to be the first step in translation initiation (Roy et al., 1988). In this work, AGO (referred to in this study as Co-eIF-2A) was found to be important for stabilizing ternary complex formation on an mRNA prior to its transfer to the 40S ribosomal subunit (Roy et al., 1988). This literature has been largely superseded by emphasis on post-transcriptional gene silencing by AGO, but it raises the possibility that AGO may be "moonlighting" with a role in HCV translation. The HCV IRES is unique in that it can bind to the 40S subunit in the absence of any initiation factors; it appears that only after 40S subunit engagement do eIF3 and the ternary complex bind to initiate translation (Ji et al., 2004; Pestova et al., 1998). If AGO plays a role in stabilizing the ternary complex and is already in the neighborhood by binding at miR-122 sites, then perhaps its primary function extends beyond protecting the uncapped genome, and is required to get translation initiation going efficiently. The observed CLIP map on the IRES may be consistent with this.

In a general sense, the overall view of the AGO-CLIP map, with a large peak at the 5' UTR and AUG start, broad coating along the CDS and basically

zero signal from the stop codon into the 3'UTR, looks very similar to ribosomal profiling on a highly expressed gene. A prototypical example is FMRP, an RNABP known to be associated with translating ribosome such that CLIP with this protein reveals broad coating of coding exons (Darnell et al., 2011). Could this be evidence of a role for AGO ahead of the translocating ribosome or in association with it? Clearly, more work is needed to resolve this.

The above are but a few interpretations drawn from the AGO-CLIP map on the HCV RNA genome and I hope they raise more questions than answers. It bears keeping in mind that a map constructed from cells infected for two days, undoubtedly reflects the AGO binding of a mixture of viral genomes at various stages of the viral life cycle. Thus we cannot conclude based on this static map which binding events are linked and therefore dependent on one another. Nor can we assign a weight to each binding event, assuming that they are independent: does the 50% engagement at the 5'UTR mean that 50% of viral RNAs are engaged with miR-122? This is perhaps the wrong question. But we can begin to address the dynamics of AGO binding on HCV RNA by turning our attention toward a timecourse.

AGO binding to miR-122 sites on HCV RNA occurs early and is replication independent

The motivation behind this experiment was to manipulate as little as possible. Unlike work aimed at overexpressing or knocking down AGO or miR-

122, we aimed to use CLIP to measure how endogenous AGO:miR-122 complexes would react to a flood of incoming HCV genomes. We tested three genomes: the WT Clone 2 HCV genome which could launch an infection; a viral polymerase defective mutant (GNN) in the same background which could be translated, but could not undergo RNA replication; and a miR-122 null binder with S1 and S2 sites mutated at position 3 (p3). We reasoned that replication dependent and independent AGO binding could be distinguished in addition miR-122 dependent binding with these mutants. We performed CLIP at 6, 12, 24, and 48 hours post-electroporation in these contexts. We observed comparable AGO binding at the 5'UTR miR-122 sites of WT and GNN mutants as early as 6 hours post electroporation and not at all for the p3 mutant (Figure 3.6A-C). AGO binding to miR-122 sites in particular remained stable throughout the WT timecourse, but decreased steadily for the GNN mutant. Interestingly, in all three contexts we observed some AGO association with the IRES, suggesting miR-122 and replication independence. AGO binding to regions outside the 5'UTR emerged after 24 hours of replication, and were not observed in the GNN or p3 mutants. The overall amount of Ago binding for all three genomes' RNAs correlated with RNA abundance as measured by qPCR (Figure 3.6D-E), suggesting replication or abundance dependent Ago targeting of the viral ORF. This would not be inconsistent with sequencing noise. The persistence of IRES binding can be interpreted in at least two ways. As previously hypothesized, a non-canonical miR-122 site may be responsible (Pang et al., 2012).

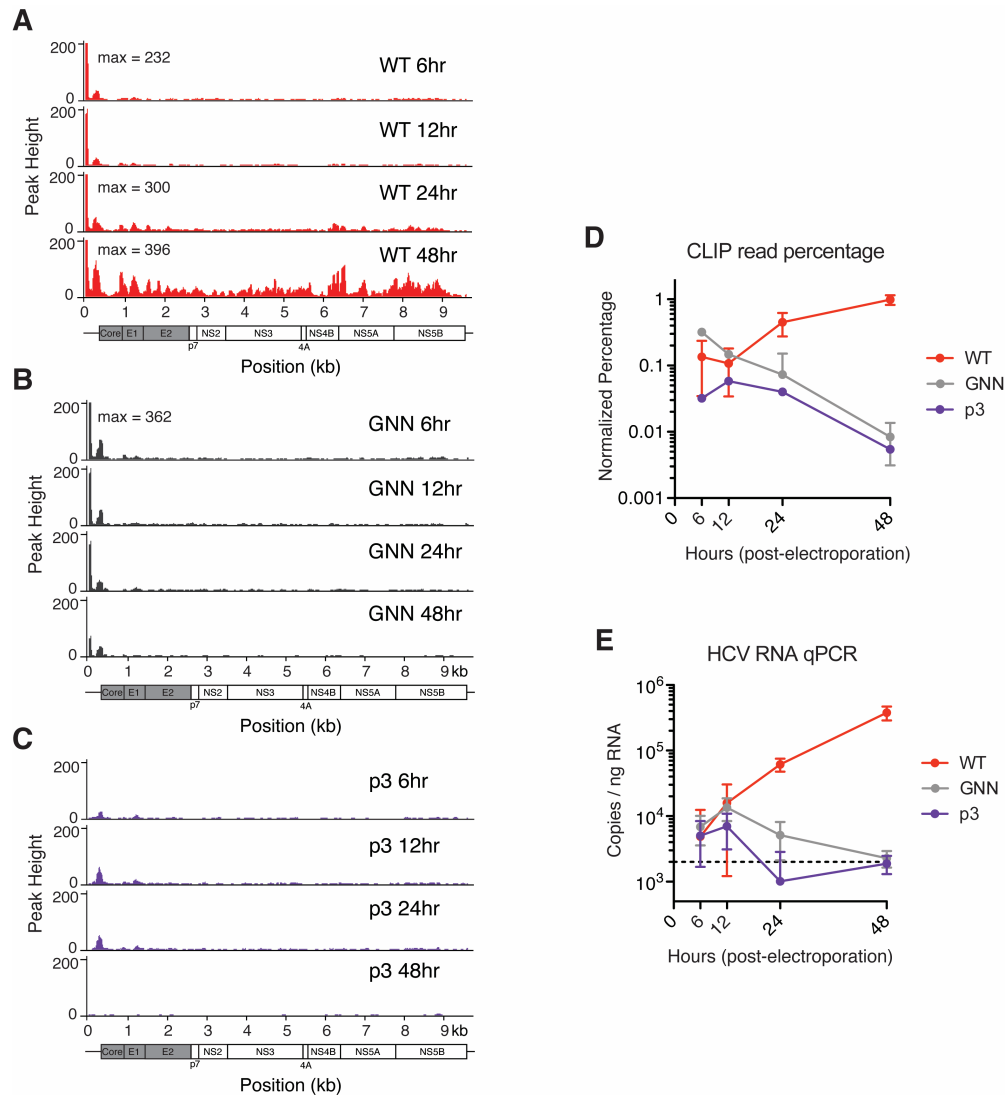


Figure 3.6. Ago binding timecourse of WT and replication deficient (GNN) HCV RNA genomes. **(A-C)** Cells were electroporated with WT, and replication deficient GNN, or p3 genomes and harvested for CLIP at indicated timepoints. The p3 mutant contains a C→G mutation at the third position for both S1 and S2 binding sites and is considered a miR-122 null binder. Data are displayed across the HCV genome as a density plot of raw peak height. The maximum peak height for S1/S2 binding is indicated if beyond scale. **(D)** Normalized percentages of HCV mapping unique reads for WT, GNN and p3 mutants across the timecourse. **(E)** qPCR measurements of HCV RNA copies / ng of total RNA across the timecourse for WT, GNN and p3 electroporations. Dashed line indicates the limit of detection. All data are mean \pm SD.

Alternatively, AGO may be playing a miRNA-independent role in regulating the IRES. Future work using replication defective genomes that cannot undergo translation due to poor eIF2 ternary complex recruitment might help shed light on this. Overall, these data indicate that while HCV RNA is bound by endogenous AGO at S1/S2 miR-122 sites early upon infection, additional replication or abundance dependent Ago binding of the viral ORF can occur at a lower level.

An HCV resistant to miR-122 antagonism engages miR-122

In 2011, Li, Bukh, and colleagues described an interesting HCV that was resistant to locked nucleic acid (LNA) based miR-122 antagonism (Li et al., 2011). After passaging HCV genomes lacking stem-loop I (SL1) of the 5'UTR, a replication competent recombinant virus was isolated where SL1 was replaced with part of cellular U3 snoRNA (Figure 3.7A). Interestingly, this "U3" virus lost the S1 miR-122 site but retained the S2 site. Based on U3 virus' ability to replicate despite miR-122 inhibition, we reasoned that CLIP with the U3 virus could yield insight on miR-122 independent binding with a replicating virus. We additionally hypothesized that this virus would generally behave as a miR-122 null binder.

Surprisingly, CLIP revealed extensive AGO binding to the 5'UTR, specifically overlapping with the S2 miR-122 site (Figure 3.7B). CIMS analysis corroborated the likely engagement of miR-122 at S2 as significant crosslinking was observed 1nt upstream of S2 auxiliary pairing, mirroring results with WT

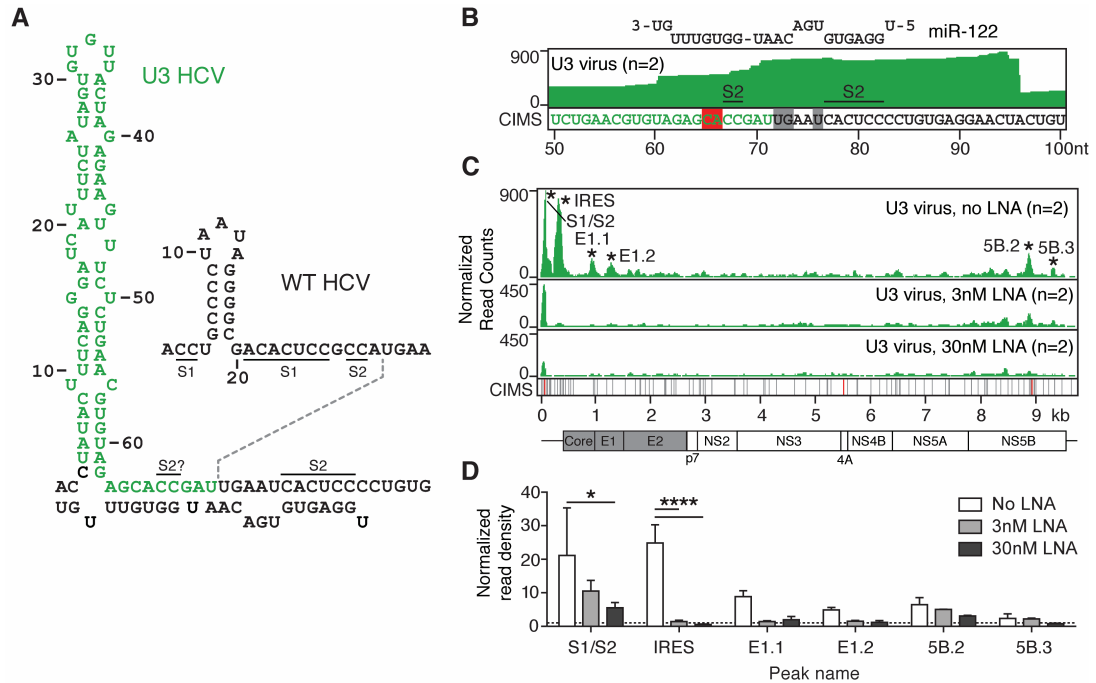


Figure 3.7. An HCV mutant resistant to miR-122 antagonism engages Ago and miR-122. **(A)** Schematic of the U3 virus 5'UTR with proposed S2 miR-122 binding. U3 snoRNA sequence is shown in green, adapted from (Li et al., 2011). For reference, the WT HCV end with stem loop I is shown. **(B)** Ago binding map (top track) and CIMS locations (bottom track) across the U3 virus 5'UTR corresponding to miR-122 binding at S2. Relevant CIMS deletions sites are shown in gray (not significant) and red (significant). U3 snoRNA sequence is shown in green. **(C)** Ago binding map across the U3 virus genome after treatment with increasing doses of LNA122. Significant peaks are named by location and are indicated by asterisks. Bottom CIMS track shows location of all deletions (gray) and statistically significant CIMS deletions (red) for the untreated dataset. **(D)** Ago binding in significant peaks from untreated U3 datasets in (C) shown as normalized read densities calculated per dataset. **** $P < 0.0001$, * $P < 0.05$, one-way ANOVA with bonferroni correction. Error bars, \pm SD. This CLIP experiment was performed with Troels Scheel.

virus presented earlier. To verify the LNA resistance of this virus, we performed CLIP in U3 virus infected cells after treatment with 3nM or 30nM LNA-122. Here we observed a sharp dose dependent decrease in AGO engagement of the 5'UTR, and a less extensive decrease across the genome, consistent with a pro-

viral effect for miR-122 binding to S2 (Figure 3.7C-D). Thus, it appears that miR-122 is capable of residually engaging the U3 virus genome by binding S2, despite its relative miR-122 independence. One possible reason for this might be that miR-122 makes many more base contacts with the U3 virus, as modeled in Figure 3.7A, and as a result miR-122 bound to the U3 viral genome may be less susceptible to LNA based antagonism. Attempts to make S2 p3 mutants of the U3 virus were all inviable (not shown), which at first glance argues that miR-122 binding to S2 remains required for viability. And yet, it's also possible that the S2 sequence has overlapping functions, such as PCBP2 binding (Masaki et al., 2015). Thus a bit of caution is warranted. Despite the preponderance of evidence that the U3 virus still binds miR-122, we cannot firmly conclude that this binding plays the same critical role that it does with WT virus. Most importantly, the striking miR-122 independence of this virus points to a potential avenue of resistance to LNA based therapeutics in the form of recombinant viruses with similarly large 5'UTR stem-loops.

CRISPR disruption of small RNA biogenesis, part I

The results so far have primarily drawn upon the introduction of RNA genomes into the cell and rely upon CLIP of endogenous AGO:miRNA to reveal the binding landscape of miRNAs on the HCV genome. As a followup, we wondered what would happen upon perturbing the miRNA content of the cell as a complementary means of exploring miRNA:HCV RNA interactions. The approach

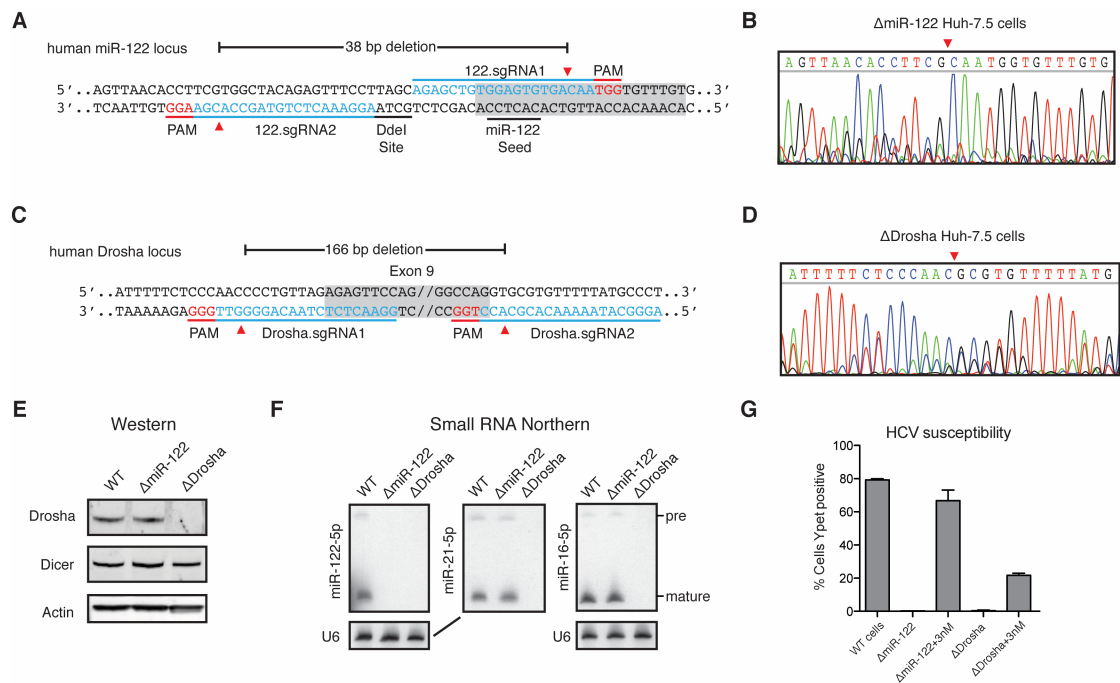


Figure 3.8. A CRISPR based targeting strategy to delete miR-122 and Drosha from Huh-7.5 cells. **(A)** Guide RNAs were designed to delete a 38bp region encompassing the 5' end of the miR-122 stemloop, including the entire miR-122 seed, and an adjacent *Ddel* restriction site. DNA cleavage sites indicated with red triangles. miR-122-5p highlighted in gray. PAM: protospacer adjacent motif. **(B)** DNA chromatogram indicating the precise genomic location (red triangle) for non-homologous end joining following CRISPR cleavage, resulting in miR-122 deletion. **(C)** Guide RNAs were designed to delete a 166bp region encompassing exon 9 of the Drosha gene, highlighted in gray. DNA cleavage sites indicated with red triangles. **(D)** DNA chromatogram indicating the precise genomic location (red triangle) for non-homologous end joining following CRISPR cleavage, resulting in Drosha exon 9 deletion. **(E)** Western blot analysis of ΔmiR-122 and ΔDrosha Huh-7.5 cells for Drosha, Dicer, and Actin. **(F)** Small RNA northern blot analysis of ΔmiR-122 and ΔDrosha Huh-7.5 cells for miR-122, miR-21, miR-16, and U6 RNA as a control. Mature and pre-miRNA bands are indicated. **(G)** HCV susceptibility of ΔmiR-122 and ΔDrosha Huh-7.5 cells supplemented with 3nM miR-122 mimic, assayed via flow cytometry 72hrs after Clone2-Ypet infection.

using LNAs, as in the U3 virus experiments, has the main benefit of being transient and simple to carry out, yet is not without confounding factors. Quite simply, no one really knows how LNAs inhibit miRNAs, beyond the basic mechanism of behaving like a very stable and tightly binding nucleic acid analog that can hybridize to a target sequence. Do LNAs inhibit pri-, pre-, mature, and/or AGO loaded mature miRNAs? Can we assume that results obtained with LNAs preclude ANY activity for the miRNA being inhibited? The same concerns surround the alternative approach of inhibiting small RNA biogenesis by siRNA or shRNA knockdown of Drosha or Dicer, with the added theoretical wrinkle of relying on circular logic that is self-limiting. The very enzymes required to generate silencing RNAs are themselves the targets of silencing.

To avoid these issues, we opted for an approach using CRISPR genome engineering to delete the miR-122, Drosha, and Dicer genes from Huh-7.5 cells. The scope of small RNA disruption enabled by these CRISPR mutants ranges from the minimal to the drastic: from the targeted deletion of one miRNA to the removal of all miRNAs. Schematized in Figure 3.8A, we targeted the miR-122 locus with transfected Cas9 to delete a 38bp stretch encompassing the miR-122 seed and a DdeI site used for genotyping (see methods). For Drosha deletion, we targeted exon 9 in a manner that would result in an out-of-frame splice between exons 8 and 10 resulting in a nonsense mediated decay substrate (Figure 3.8B). This was the same strategy used to generate the Drosha knockout mouse (Chong et al., 2008). Single cell dilution and expansion yielded Δ miR-122

and Δ Drosha clones harbored the intended deletions (Figure 3.8C-D). We observed the specific loss of Drosha protein only in Δ Drosha cells via western blot (Figure 3.8E). Functional characterization of miRNA processing via Northern blot showed that both pre- and mature miR-122 was specifically lost in Δ miR-122 cells, but not other abundant miRNAs such as miR-21 or miR-16. In Δ Drosha cells, all three of these Drosha-dependent miRNAs were lost (Figure 3.8F). The successful creation of both of these cell lines, despite altered miRNAs, suggests that miR-122 and miRNAs in general are not critically required for cell viability. However, some differences in proliferation rate were noted: while Δ miR-122 cells grew no differently than WT cells, Δ Drosha cells proliferated at half the rate (not shown). This is consistent with known role for miRNAs in regulating the cell cycle (Bueno and Malumbres, 2011).

We next tested HCV susceptibility of Δ miR-122 and Δ Drosha cells. As expected, HCV could not replicate in either cell line, but importantly, replication could be rescued with miR-122 mimic transfection (Figure 3.8G). In Δ miR-122 cells, this rescue was complete at 3nM, and not significantly different from unedited cells. In contrast, HCV replication could not be fully rescued in Δ Drosha cells. In this context, Δ Drosha cells lack all miRNAs with the exception of miR-122, thus it's possible that either additional miRNAs are directly or indirectly required for HCV replication, or that global miRNA suppression is toxic to virus for non-specific and pleiotropic reasons. Nevertheless, the ability to launch HCV

infection in these cells after miR-122 mimic transfection opens the possibility of exploring the role of all miR-122 independent binding on the HCV genome.

Probing non miR-122 dependent AGO binding on HCV RNA

With HCV permissive cell lines lacking miR-122 or all canonical mature miRNAs, we next decided to pursue CLIP in these contexts after HCV infection. Δ Drosha cells were first transfected with 3nM miR-122 mimic and then infected with WT HCV for three days before CLIP. Importantly, no HCV replication was observed in these cells without miR-122 supplementation (Figure 3.8G). We reasoned that compared to WT cells, CLIP in HCV infected Δ Drosha cells supplemented with miR-122 would only retain or enhance miR-122 dependent peaks, and that peaks lost in Δ Drosha might be indicative of other miRNA binding. We observed that the S1/S2, IRES, E1.2 and NS5B peaks remained significant in Δ Drosha+miR-122 cells, but the low level background generally persisted (Figure 3.9A). Ago binding to E1 and E2 peaks was reduced in Δ Drosha cells, suggesting that a minor proportion Ago binding on HCV RNA is due to other miRNAs, however no canonical seeds were found for these peaks. As a means of quantifying the relative amounts of Ago binding across peak regions between cell types with different infection frequencies, the density of reads under each peak region was normalized to total read depth and plotted per dataset.

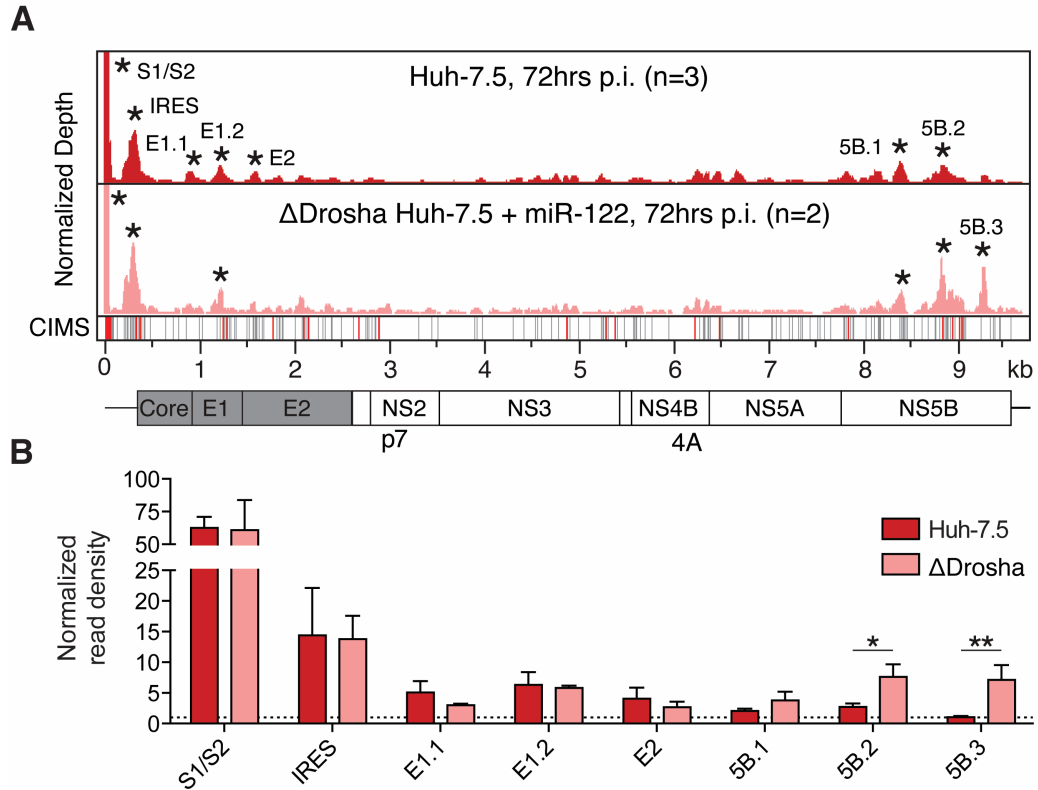


Figure 3.9. Argonaute binding maps on HCV RNA in the absence of miRNAs. **(A)** Mock-subtracted binding map of Ago-CLIP reads across HCV genomic RNA in WT or Δ Drosha Huh-7.5 cells. Data were normalized to total cellular and virus read depth for cross track comparison. Statistically significant peaks per track are named by location and are indicated by asterisks. Bottom CIMS track shows location of all deletions (gray) and statistically significant CIMS deletions (red) from the WT track. **(B)** Ago binding in significant peaks from WT Huh-7.5 cells in (A) shown as normalized read densities calculated per dataset. Data were normalized to background read density of non-peak regions and set to one (dashed line). Asterisks, ** $P < 0.01$, * $P < 0.05$, Student's t-test. Error bars, \pm SD.

We can appreciate the relative amount of AGO binding per peak region and determine the extent of signal loss or gain upon removal of miRNAs (Figure 3.9B). Clearly, S1/S2 and IRES binding are unaffected, and peaks in NS5B (mainly 5B.2 and the emergence of 5B.3) were enhanced in Δ Drosha cells, all suggesting miR-122 dependence. The 5B.3 peak is truly novel, and is perfectly

conserved among consensus HCV genotypes (Figure 3.10). The proximity of the remaining peaks to the read density of regions not in peaks (aka, the background) suggests AGO-CLIP noise as mentioned previously. Based on these results, we can be more confident that miRNAs other than miR-122 play little, if any, role in binding HCV RNA, at least in the context of an Huh-7.5 cell.

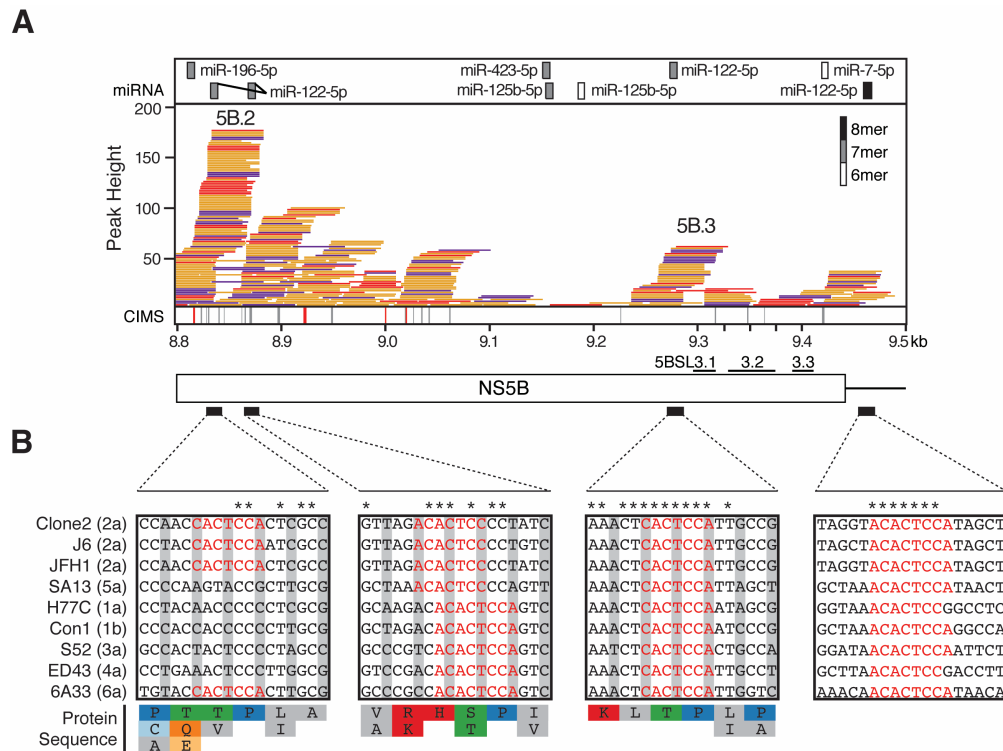


Figure 3.10. miR-122 binding sites in the NS5B coding sequence and 3'UTR of HCV RNA. **(A)** Ago binding map in HCV infected WT cells across the NS5B coding sequence, from three experiments. **(B)** Sequence alignments for indicated major HCV genotypes (1-6) of bound miR-122 sites presented in (A). miR-122 seed highlighted in red. Third codon positions shaded in gray; amino acids are colored according to physiochemical properties and are presented as the Clone2 sequence with deviations in other genotypes below. Conserved nucleotides are marked with asterisks.

The U3 virus replicates in the absence of miR-122 and miRNAs

Given the seeming miR-122 independence of U3 virus replication, we decided to test U3 replication in Δ miR-122 and Δ Drosha cells, and if possible, perform AGO-CLIP to resolve the miRNA dependence of the unique AGO binding pattern on this viral genome. We hoped to address the paradox observed with this virus: the U3 virus is resistant to miR-122 antagonism, and yet retains the S2 site that is clearly and specifically bound by AGO.

Starting with Δ miR-122 cells, we compared WT versus U3 virus infection in these cells measuring HCV RNA over a timecourse. As expected, WT virus was completely abolished in Δ miR-122 cells, but could be rescued with 3nM miR-122 supplied in trans (Figure 3.11A). Notably, no difference was observed in U3 virus replication in WT versus Δ miR-122 cells, and miR-122 supplementation had a mild enhancing effects for the U3 virus, that was only significant in the last timepoint (Figure 3.11B). These results strongly suggest that U3 virus replication is independent of miR-122. This was further corroborated by AGO-CLIP in U3 infected Δ miR-122 and Δ Drosha cells. U3 virus replication was slower in Δ Drosha cells compared to Δ miR-122 cells (Figure 3.11C), however, we achieved sufficient percentages (>30%) to enable AGO-CLIP. The U3 virus displayed no notable AGO binding in both Δ miR-122 and Δ Drosha contexts (Figure 3.11D). These data alongside our LNA results confirm that the S2 site engages miR-122, but that this binding is somewhat dispensable. Furthermore, this suggests that S1 and S2 might have distinct non-overlapping functions. In

the case of the U3 virus, the S1 site appears to have been replaced with a large stem loop. The recent observation that the S2 site also binds to PCBP2 and may compete with miR-122 binding points to a different role beyond RNA stabilization (Masaki et al., 2015).

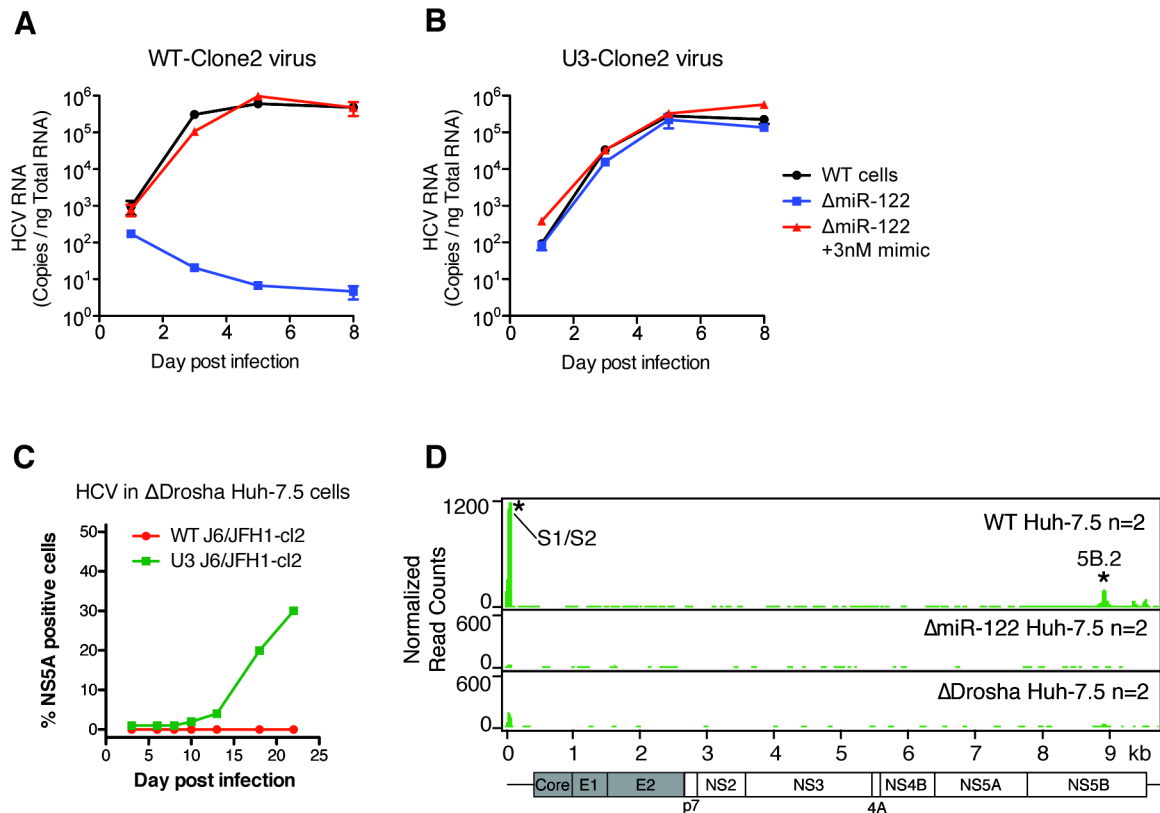


Figure 3.11. Timecourse qPCR measurements of WT-Clone2 virus **(A)** or U3-Clone2 virus **(B)** in WT cells or in ΔmiR-122 cells with or without 3nM miR-122 supplementation. Error bars, \pm SD. AGO-CLIP with U3 virus in ΔmiR-122 and ΔDrosha cells. **(C)** Percent cells ΔDrosha cells infected with U3 virus over time. **(D)** U3 virus AGO-CLIP in WT Huh-7.5 cells (top), ΔmiR-122 (middle), and ΔDrosha Huh-7.5 cells (bottom). These experiments performed by Troels Scheel.

Swapping miRNA tropism of HCV: an attempt with miR-21

The results thus far have focused on the WT virus and the U3 virus, which no longer critically requires miR-122, but can still bind it. To get around the confounding factors of the U3 virus, such as its large SL1 and its loss of the S1 miR-122 site, we opted to swap the miRNA tropism of WT virus as a means of minimally creating a miR-122 independent, but not miRNA independent, virus. This idea was first explored with miR-21 as a miRNA expressed in Huh-7 cell derivatives, but with a sufficiently altered seed compared to miR-122 (Figure 3.12A-B) (Jopling et al., 2008). The added benefit of starting with miR-21 was that there was some interesting negative data worth revisiting. In an early report, the S1 miR-122 site in an HCV replicon system was mutated to a miR-21 site and found to be inviable in Huh-7 cells (Jopling et al., 2008). Based on this result, it was argued that the AGO:miR-122 complex might be unique in its engagement with HCV RNA, and that mutation of S1 may be interfering with an overlapping function. With the discovery of critical auxiliary pairing at positions 2-3 and 30-31 of the HCV genome with miR-122, the issue was revisited using a chimeric miRNA that consisted of the miR-21 seed (positions 1-8) and the remainder as miR-122 (positions 9-22). This miRNA was not capable of rescuing replicons bearing tandem miR-21 S1 and S2 sites, despite the fact that this chimeric miRNA should have been able to bind the auxiliary sites (Roberts et al., 2011).

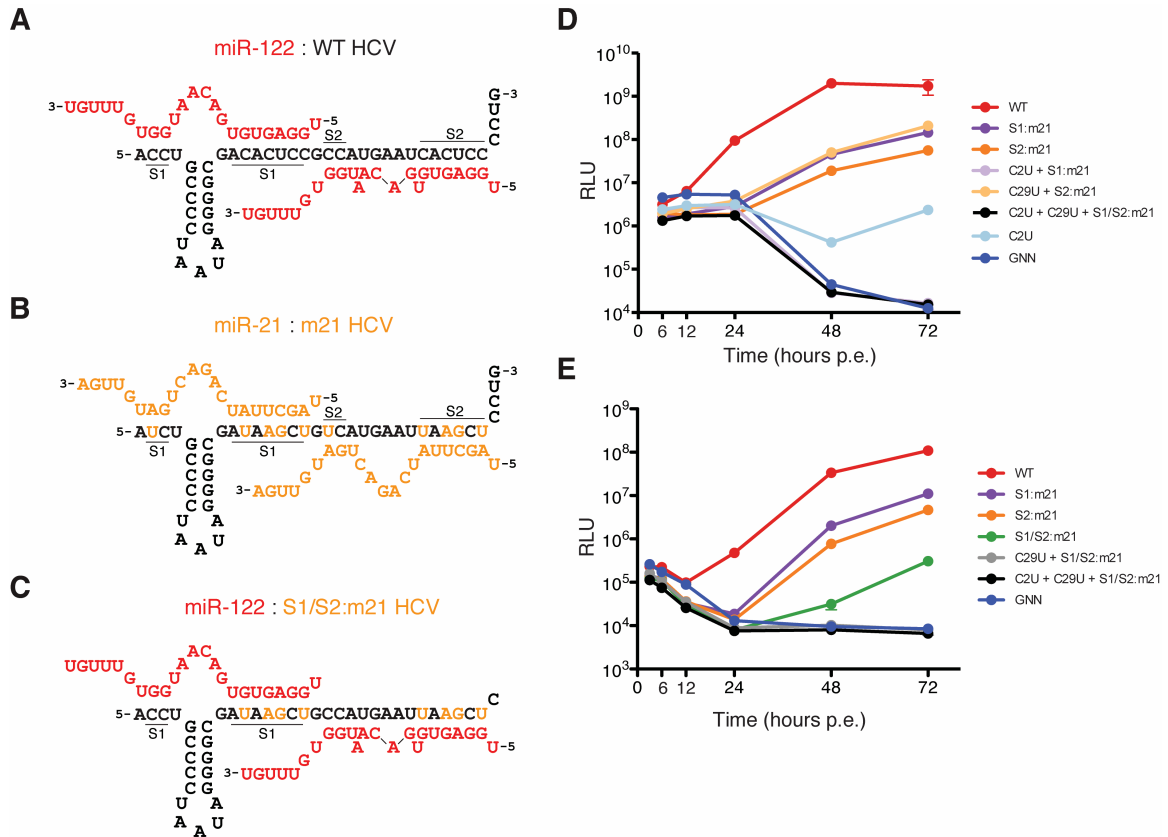


Figure 3.12. Exchanging miRNA tropism for HCV to use miR-21. **(A)** WT HCV RNA in complex with miR-122 based on extensive genetic and SHAPE data (Jopling et al., 2008; Machlin et al., 2011; Mortimer and Doudna, 2013). **(B)** A proposed model for miR-21a engagement with a miR-21 family dependent HCV ("m21"). **(C)** A proposed model for miR-122 engagement with a S1/S2 miR-21 HCV. Note that G:U wobbles are capable at positions 2 and 6. **(D)** Full length mono-cistronic secreted luciferase reporter virus (Jc1 background) timecourse of various mutants on the way to a miR-21 dependent virus. WT and GNN mutants were contained miR-122 S1 and S2. Variations in swapping S1 or S2 seeds, or in mutating positions 2 or 29 to accomodate miR-21 are noted. Samples were washed at 24 and 48 hours after supernatant harvest. **(E)** Luciferase reporter virus measurements as in (D) with additional mutants. Samples were washed after all timepoints measured. All samples plotted as means \pm SD.

Additionally, an important HCV seemed to be missing: a miR-21 dependent virus that incorporated the appropriate base changes at positions 2 and 30 to enable endogenous miR-21 recognition in a manner analogous to miR-

122 (Figure 3.12B). We set out to make this virus, reasoning that individual S1 and S2 miR-21 swaps would not be viable, and that the full miR-21 swap mutant with additional C to U mutations at positions 2 and 29 would replicate.

To our surprise, both S1 and S2 seed mutants were individually viable when swapped for miR-21, each exhibiting one log attenuation compared to WT genomes (Figure 3.12D). The S1/S2 double mutant was viable though attenuated by closer to three logs compared to WT (Figure 3.12E). These results with full-length virus contradict earlier reports and showcase a tolerance for seed mutations. Interestingly, the full miR-21 binding mutant shown in Figure 3.12B was not viable (in black, Figure 3.12D-E). Working backwards from this mutant, we uncovered a high sensitivity for mutations at nucleotide 2 of the HCV genome. In all cases, mutations to U at this position were either inviable or very highly attenuated. A complete S2 swap (including a C29U mutation) was as viable as an S1 swap, however combining these mutations did not result in any viral replication (in gray, Figure 3.12E). That miR-21 S1/S2 swaps were viable, but when combined with C2U or C29T mutations had lost this viability, suggested that miR-21 was not rescuing this virus as thought and that perhaps, miR-122 was still engaging this genome. Following the observation that S1/S2 virus replication could be boosted with miR-122 mimic (not shown), we settled on a model whereby this virus represented a maximal deviation from canonical miR-122 binding in that many G:C base pairs were replaced by G:U wobbles (Figure 3.12C). Furthermore, the drastic inhibition of the C2U mutation hinted at an

overlapping function beyond miRNA pairing that was perturbed upon mutagenesis.

Swapping miRNA tropism of HCV to use miR-15

Perhaps the biggest lesson of our experience with miR-21 was that the C at position 2 of the HCV genome likely has an overlapping function beyond binding a miRNA. A likely candidate for this additional function is a role in RNA replication. Learning from this, we reasoned that a viable replacement for miR-122 would need to maintain a GG at analogous positions to pair at auxiliary sites, should not allow for G:U wobbles in the seed, and should be expressed at similar levels to miR-122. Taking this candidate approach over a literature based approach, we settled on the miR-15 family. In particular, miR-15a/b satisfied the above criteria, with five non-G:U changes in the seed compared to miR-122, maintenance of GG at positions 15-16, and present as ~4.8% of miRNAs in the cell compared to ~8.3% for miR-122 (Figure 3.13).

While miR-16 dependent viruses were inviable (not shown), an electroporated miR-15a/b dependent HCV luciferase reporter virus (m15) was viable, and replicated to within one log of the WT virus after 72hrs (Figure 3.14A). Notably, the miR-15 virus was resistant to increasing concentrations of LNA122 ($IC_{50} > 50nM$) but susceptible to LNA15a/b ($IC_{50} = 10nM$) (Figure 3.14B). Unlike WT virus, the m15 virus was also viable in Δ miR-122 cells (Figure 3.14C). Taken together, these data strongly support miR-122 independent virus replication by exchanging miRNA tropism to use miR-15a/b.

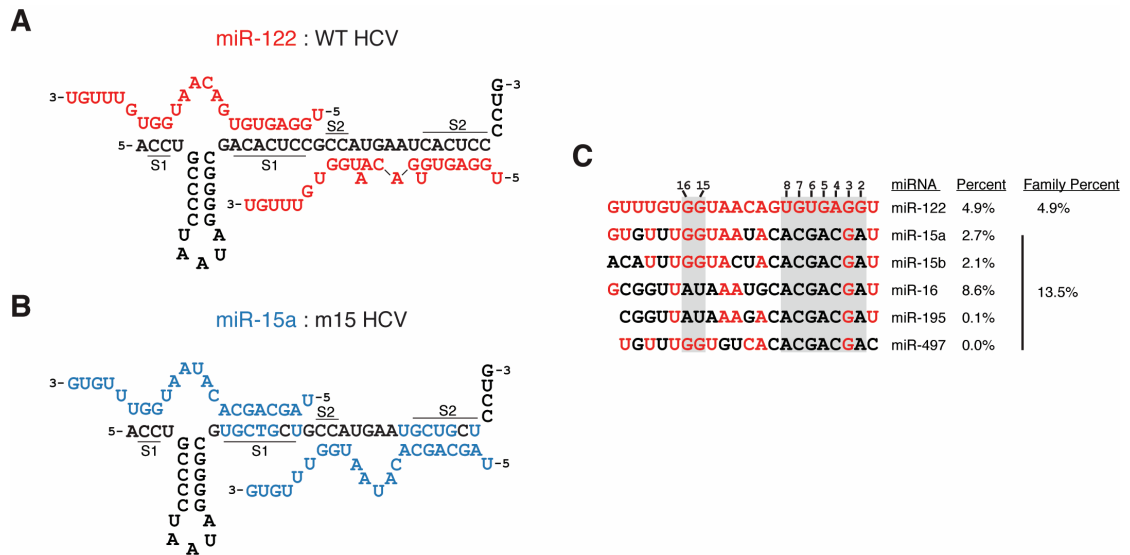


Figure 3.13. Exchanging miRNA tropism for HCV. **(A)** WT HCV RNA in complex with miR-122 based on extensive genetic and SHAPE data (Jopling et al., 2008; Machlin et al., 2011; Mortimer and Doudna, 2013). **(B)** A proposed model for miR-15a engagement with a miR-15 family dependent HCV ("m15"). **(D)** Conservation of miR-122 (the sole family member) with the miR-15 family reveals extensive differences in the seed (6 changes with no possibility for G:U wobble pairing), maintenance of seed GC content, and importantly, the preservation of the GG at positions 15-16, critical for the 3' supplementary pairing shown in (C). Note that GG is not preserved in miR-161. Percent abundances of these miRNAs on AGO as measured by polyG-CLIP (see next chapter) are indicated.

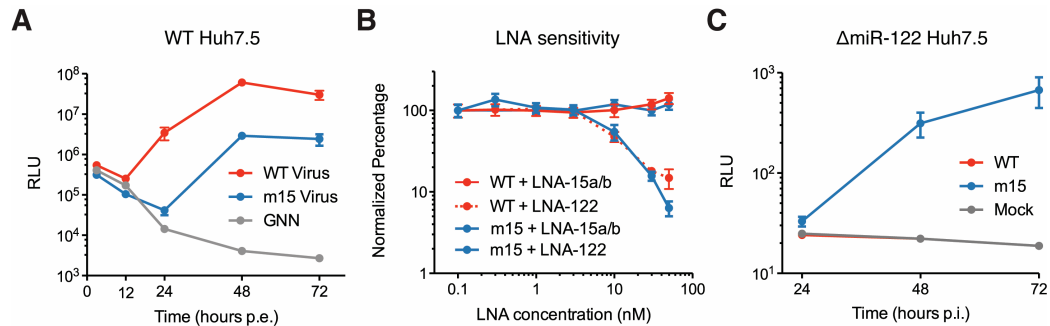


Figure 3.14. Functional characterization of a miR-15 dependent HCV. **(A)** Luciferase measurements of supernatants from WT and m15 HCV reporter virus electroporations (\pm SD). Non-replicating GNN control is shown. Media was replaced at each time point measured. **(B)** Dose response of WT and m15 reporter viruses following pre-treatment with LNA inhibitors of miR-122 or miR-15a/b at indicated concentrations, measured at 96 hours post-infection (\pm SD). **(C)** Time course post infection of Δ miR-122 Huh-7.5 cells with indicated viruses (\pm SD). These experiments were performed with Katharina Shaw.

We expanded these analyses with the non-reporter viruses in the Clone2 background and measured sensitivities to 50nM LNA122 or LNA15 prior to virus infection. As expected WT HCV RNA levels were reduced 20-fold with LNA122 treatment but were resistant to LNA15, the converse was true for the m15 virus, though with less attenuation (Figure 3.15A-B). We noted that the expected inhibitory LNA treatment for either WT or m15 virus reduced infected cell frequency (Figure 3.15C-D). Lastly, as expected, WT virus infectivity was severely reduced in LNA122 treated cells but was unaffected by LNA15 (Figure 3.15E). The infectivity for the m15 virus presented an interesting result in that it was reduced in the presence of LNA15 as expected, but also in the presence of LNA122 (Figure 3.15F).

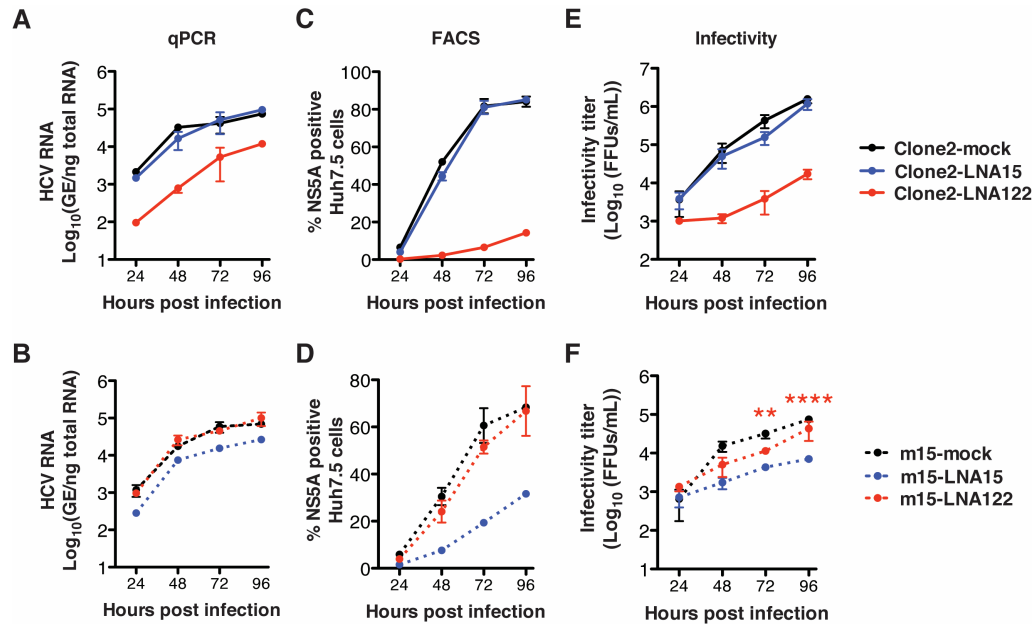


Figure 3.15. LNA sensitivity assays comparing WT HCV and m15 HCV. **(A-B)** qPCR measurements for WT (A) or m15 (B) viruses were performed at indicated timepoints post infection, after mock treating, or treating with LNA inhibitors to miR-15a/b (blue) or miR-122 (red) at 50nM, from duplicate experiments. Data presented as Genome Equivalents (GE) / ng of total cellular RNA (\pm SD). **(C-D)** Percent NS5A positive cells (\pm SD) in WT (C) and m15 (D) virus infections following LNA treatment as above, measured by flow cytometry. **(E-F)** Infectivity titer measurements of virus released from LNA or mock treated cells as above for WT (E) and m15 (F) virus infections, as FFU/mL (\pm SD). These experiments were performed with Troels Scheel and Eiko Nishiuchi.

At least two possibilities might explain the apparent reduction in m15 virus titer upon treatment with LNA122. Conceivably, miR-122 may have direct pro-viral roles on HCV RNA beyond S1/S2 sites, such as the putative IRES binding site, or sites in NS5B. Alternatively, and not mutually exclusively, the miR-122 regulated transcriptome if globally de-repressed may alter the intracellular environment such that it is "less pro-viral." I hesitate to use the term "anti-viral" for LNA122 treatment in this indirect context, since this would imply some level of

cellular intent. For example, consider a miR-122 target that if de-repressed would signal the cell to mount an anti-viral response or initiate cell-cycle arrest. Such scenarios have been recently proposed with limited experimental support (Xiong et al., 2014; Zhao et al., 2013). Arguably, such a target could constitute a sensor for miR-122 levels. However, if de-repressing miR-122 targets inhibits m15 virus due to a passive indirect effect, such as altering the lipid composition or trafficking, but playing no selected for role in detecting or combatting virus, this would be evidence of a non-specific "anti-viral" effect. In any case, as the above two scenarios are simultaneously possible and are completely dependent on the underlying host transcriptome, we will revisit them in the next chapter. What is common to both perhaps, is that the virus faces a miR-122 defined ceiling that limits its replication.

Refocusing on the virus, the ability of the m15 virus to replicate in Δ miR-122 cells offers an alternative means to explore miR-122 dependent binding as an obverse experiment to that presented in Figure 3.9. Recall how that experiment was performed with WT virus in Δ Drosha cells supplemented with miR-122, thus only miR-122 binding was expected to be observed. In the experiment presented in Figure 3.16, we performed CLIP in m15 virus infected WT and Δ miR-122 cells. Here we expect a specific loss of miR-122 dependent binding where no change should be observed at S1/S2 sites since these depend on the miR-15 family.

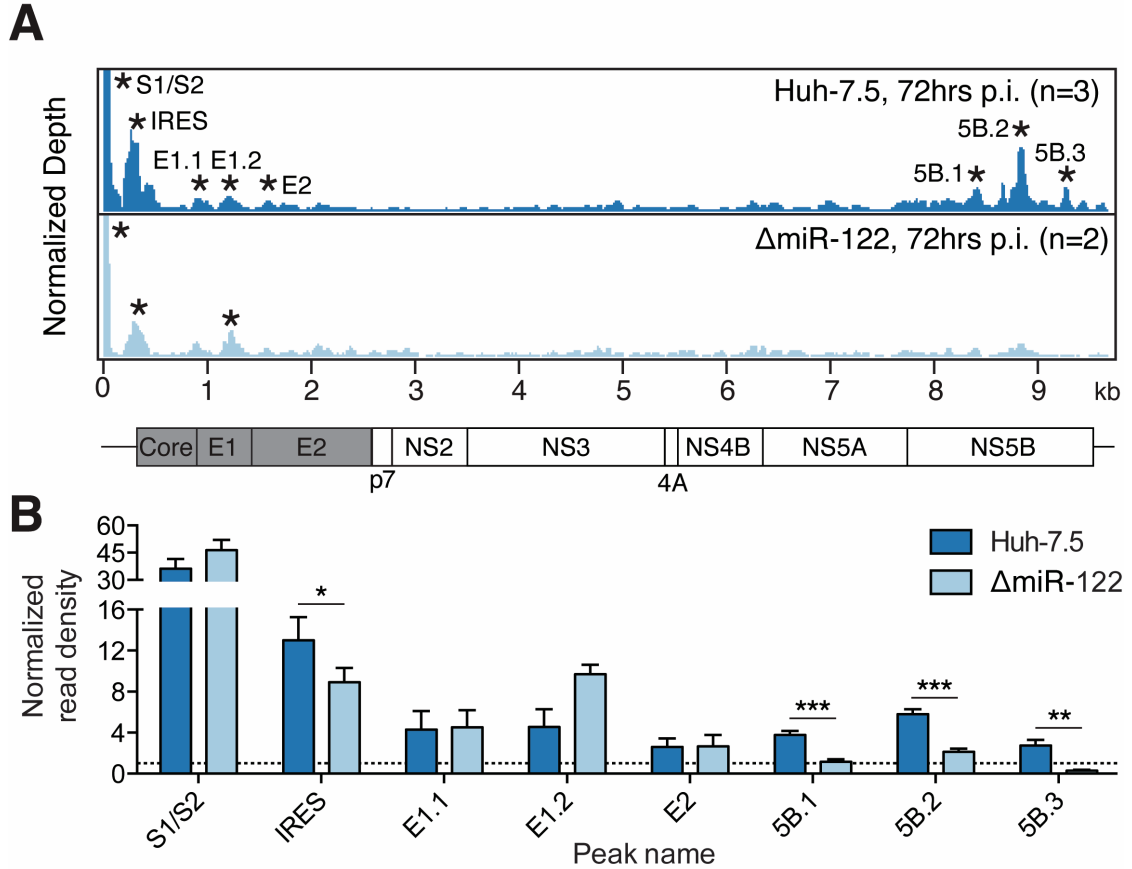


Figure 3.16. AGO binding maps on m15 HCV reveal extent of miR-122 dependent targeting. **(A)** Ago binding map of m15 virus infection in WT Huh-7.5 (top panel) or Δ miR-122 Huh-7.5 cells (bottom panel). Data were normalized to total cellular read depth for cross track comparison. Statistically significant peaks per track are named by location and are indicated by asterisks. **(B)** Ago binding in significant peaks from (A) shown as normalized read densities calculated per dataset. Two-sided Student's t-test used. Error bars, \pm SD.

The AGO-CLIP pattern on m15 virus was similar to that observed with WT virus, indicating that the S1/S2 swap can recapitulate the previously observed binding pattern (Figure 3.16A, top panel). In Δ miR-122 cells we noted the reductions in AGO binding in the viral IRES and in the NS5B peaks, strongly supporting miR-122 dependent binding at these sites (Figure 3.16A, bottom panel). Quantification

of read density within peaks revealed statistically significant reductions in putative miR-122 dependent sites (Figure 3.16B). These data offer confirmation for endogenous miR-122 dependent binding outside of S1/S2.

The above discussion on the apparent reduction of m15 virus to LNA122 suggests a miR-122 defined ceiling for viral replication. While it remains to be determined if miR-122 binding on the viral RNA, outside of S1/S2, plays a direct role in augmenting HCV RNA levels, the Δ miR-122 cell context offers a platform to address the general ceiling question. In this experiment, we re-introduced miR-122 at various concentrations into Δ miR-122 cells, and infected with m15 virus. If the ceiling hypothesis holds, then we could expect to see a miR-122 dependent enhancement of m15 virus. Comparing untreated Δ miR-122 cells or cells treated with 3nM or 30nM of miR-122, we noted a two-fold enhancement of viral RNA and infectious titer levels in the presence of miR-122 (Figure 3.17A and C). We also detected a slight increase in the percentage of infected cells (Figure 3.17B). These data, combined with the observations of LNA122 inhibition of the m15 virus, lend support for a replication ceiling that is in some measure dictated by miR-122 levels. Moreover, as we'll see in the next chapter, this ceiling appears modified by the virus to functionally reduce the miR-122 levels in the cell.

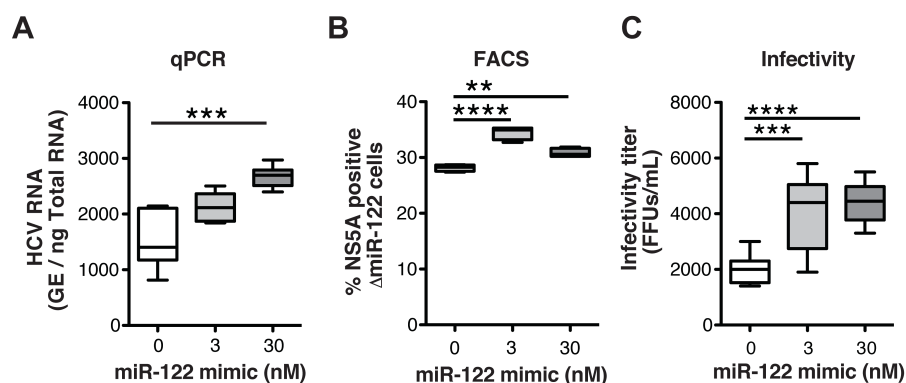


Figure 3.17. Effect of miR-122 reintroduction to m15 virus infected Δ miR-122 cells. **(A)** qPCR measurements for m15 virus RNA in Δ miR-122 cells supplemented with miR-122 mimic at indicated concentrations, from quadruplicate experiments. **(B-C)** FACS analysis of infection frequency (B) and infectivity measurements (C) as in (A). Asterisks: **** $P < 0.0001$, *** $P < 0.001$, ** $P < 0.01$, one-way ANOVA with Bonferroni multiple comparison correction.

CRISPR disruption of small RNA biogenesis, part II

In contrast to Drosha, which is required for miRNA biogenesis but plays no role in RNAi, Dicer represents a critical node for both miRNA and siRNA biogenesis. Successful deletion of Dicer in Huh-7.5 cells would thus enable an unbiased exploration of any role for RNAi in HCV infected cells. In combination with Δ Drosha cells, we also reason that these reagents may also prove useful in revealing the respective contributions of miRNAs or siRNA in other virus infections.

More proximally, Δ Dicer cells have the potential to resolve seemingly contradictory data regarding miR-122 usage by HCV RNA. One study found that Dicer knockdown reduced HCV RNA accumulation and could be rescued by adding back mature miR-122 duplexes (Zhang et al., 2012). Based on these

results, the authors concluded that mature miR-122 produced by Dicer activity on pre-miR-122 was key for maintaining HCV RNA abundance. And yet, mounting evidence suggests that pre-miR-122 may play an active regulatory role, especially in liver, where pre-miR-122 oscillates in a circadian manner and correlates with oscillating miR-122 target repression despite invariant (and high) levels of mature miR-122 (Gatfield et al., 2009). This work suggests an interesting hypothesis whereby mature miR-122 levels may not necessarily explain the full regulatory potential of this miRNA. Conceivably, newly synthesized miR-122 or the pre-miR-122 intermediate itself, may play a regulatory role distinct from "old" miR-122 (Gatfield et al., 2009). Starting from this premise, Cox, Doudna, Sarnow and colleagues tested the abilities of pre-miR-122 molecules to maintain HCV RNA abundance. By using pre-miR-122 hairpins with DNA bases at known positions of Dicer processing, they were able to make hairpins refractory to Dicer cleavage *in vitro* (Cox et al., 2013). These Dicer resistant hairpin molecules were observed to function in miRNA and siRNA assays, and in addition appeared to maintain HCV RNA abundance. Further testing of a bulged pre-miR-122 that exhibited poor Dicer processing and not able to function in miRNA and siRNA assays was also able to maintain HCV RNA abundance (Cox et al., 2013). Overall, these results suggest a role for pre-miR-122 in binding and stabilizing HCV RNA.

The above two studies present somewhat contradictory findings. The first essentially argues that mature miR-122 is required for HCV RNA maintenance,

and thus implies a critical need for Dicer enzymatic activity. The second argues that pre-miR-122 can function to maintain HCV RNA, and therefore could imply that Dicer is largely dispensable. So which is it? Note that the results of the first study do not entirely rule out pre-miR-122 function so much as provide clear evidence that mature miR-122 functions on HCV RNA.

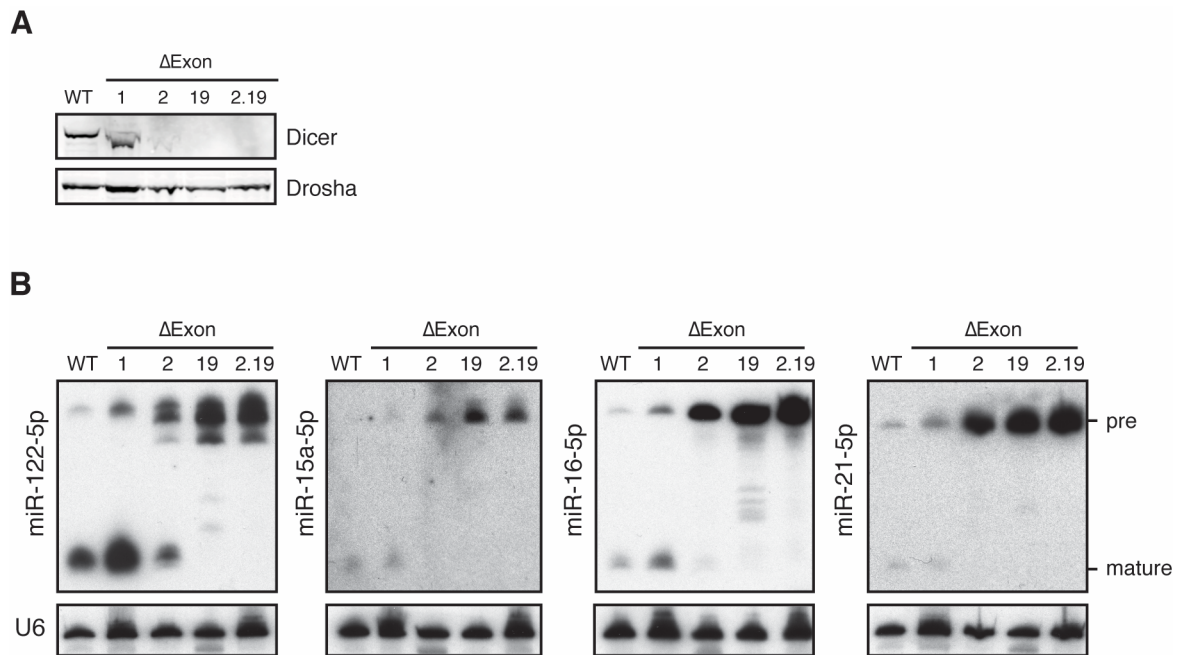


Figure 3.18. Functional characterization of CRISPR engineered Dicer mutants in Huh-7.5 cells. **(A)** Western blot of Dicer in WT, Δ EX1, Δ EX2, Δ EX19 and Δ EX2.19 cells. Drosha served as a loading control. **(B)** Small RNA northern analysis of each of the above Dicer mutants probing for miR-122, miR-15a, miR-16 and miR-21. U6 RNA served as a loading control.

To address this unambiguously, we decided to delete Dicer in Huh-7.5 cells and test their susceptibility to HCV. Our initial attempt focused on deleting the first coding exon, where we reasoned that removing the start codon would prevent Dicer translation. Exon 1 was successfully deleted, but to our surprise,

Dicer protein was made at near WT levels and miRNA processing proceeded normally (Figure 3.18). In some experiments we also noticed that the loading control U6 RNA appeared degraded specifically in these cells (not shown). The nearest in-frame methionine is in exon 4, which would result in a protein missing 149 amino acids from the N-terminus, corresponding to approximately a 17kDa reduction in size (from ~220kDa to 203kDa). In terms of protein structure, this would result in a Dicer mutant with a non-functional DEXDc helicase domain. Such mutants have been extensively characterized *in vitro* and are known to exhibit WT reaction velocities for miRNA processing but intriguingly, were found to be dramatically enhanced in siRNA processing (Ma et al., 2008; 2012). In this manner, the Dicer helicase domain is thought to be an inhibitor of Dicer siRNA but not miRNA processing. More recent work has revealed a curious form of Dicer active in mouse oocytes that essentially lacks the helicase domain (Flemr et al., 2013). This oocyte specific Dicer, called Dicer^o, arose from a retro-transposon inserted alternative first exon between exons 6 and 7 of the Dicer gene. Found only in mice and rats, this Dicer also exhibits enhanced siRNA biogenesis and is thought to promote endogenous RNAi in the mouse germline for reasons that are still not yet clear (Flemr et al., 2013). Could our Dicer Δ EX1 Huh-7.5 cells be used to illustrate an antiviral RNAi response in somatic cells? Experiments are currently underway to address this.

Taking a cue from Δ Drosha cells, where an exon was deleted to promote an out-of-frame splice to generate an NMD substrate, we next deleted exon 2 of

Dicer, thinking that by preserving translation in exon 1, no truncated Dicer would be made. Dicer Δ EX2 cells exhibited a near complete absence of Dicer protein levels; however a very faint band could still be detected in long exposures and resembled Dicer Δ EX1 in size (Figure 3.18A). These cells also displayed a marked increase in pre-miRNAs, suggesting a major defect in Dicer processing. Unlike pre-miR-15a, pre-miR-16 and pre-miR-21, the precursor for miR-122 yielded additional unprocessed bands (Figure 3.18B). And yet, a small amount of mature miRNA could clearly still be made from miR-122 and miR-16, likely due to residual truncated Dicer activity (Figure 3.18A). Total Dicer knockouts these cells were not.

As a last attempt, I deleted exon 19, which is immediately upstream of the RNase III domain coding exons, and whose deletion would induce an out-of-frame splice (Dicer Δ EX19). Protein made any upstream start codon would encounter a stop before RNaseIII domains and thus would not function, if it happened to survive NMD. Additionally, any start downstream would make protein consisting only of the RNaseIII and the dsRNA binding domains, and also would not function (Ma et al., 2012). Exon 19 was also deleted Dicer Δ EX2 cells as an additional redundant strategy (making Dicer Δ EX2.19 cells). This third targeting strategy finally worked. Dicer protein levels were completely abolished, as was pre-miRNA processing in cell lines lacking exon 19 (Figure 3.18).

CRISPR mediated genome engineering is rightfully lauded as a superior alternative to traditional knockdown and knockout approaches for its ease and

flexibility. As the above exercise shows, targeting a few regions in the same gene largely bears this out, with the important caveat that not all knockout engineering attempts will yield the same results. In this instance, we fortuitously achieved a few very interesting mutants: a putative "super" Dicer expressing cell, a cell expressing very little Dicer, and true Dicer knockouts.

Testing HCV dependence on Dicer: a role for pre-miR-122?

Now with a panel of endogenous Dicer mutant Huh-7.5 cells, we next tested their permissiveness to HCV. If pre-miR-122 can function to stabilize HCV RNA as posited by (Cox et al., 2013), we would expect to see HCV replication in Dicer Δ EX19 and Dicer Δ EX2.19 cells, were only pre-miR-122 is made. On the other hand, these cells should be refractory to HCV replication if only mature miR-122 is required.

We observed no replication in both Dicer Δ EX19 and Dicer Δ EX2.19 cells after infection with Jc1 luciferase reporter virus (Figure 3.19B). Virus replication could be rescued in these cells by transfecting 30nM miR-122 mimic, largely bypassing the Dicer requirement. We also tested Dicer Δ EX1 and Dicer Δ EX2 cells, and found that virus replication was attenuated by one log in both contexts. This was notable in light of the fact that Dicer Δ EX1 cells expressed WT levels of protein, while Dicer Δ EX2 cells expressed 95% less Dicer (Figure 3.19A). This may reflect more promiscuous Dicer activity in Δ EX1 cells, though this has yet to be fully confirmed. Overall, these data preliminarily suggest that endogenous pre-miR-122 cannot promote HCV replication. As this conclusion disagrees with

previous work, there are alternative hypotheses to consider in reconciling these disparate findings. Conceivably, Dicer may play a role in loading pre-miR-122 molecules apart from its RNase activity. Restoration of Dicer mutants that cannot cleave miRNA stem loops will need to be tested to see if HCV replication can be rescued in this context. Additionally, we've yet to test if transfected pre-miR-122 can rescue HCV replication in Dicer Δ EX19 and Dicer Δ EX2.19 cells. This is important considering that the increase in pre-miR-122 yielded multiple RNA species, in contrast to miRs -15a, -16, and -21 in which a single pre-miRNA was unregulated (Figure 3.18B). Taken together, these data bring up many fascinating questions to explore using genome engineered manipulations of small RNA activity in the cell.

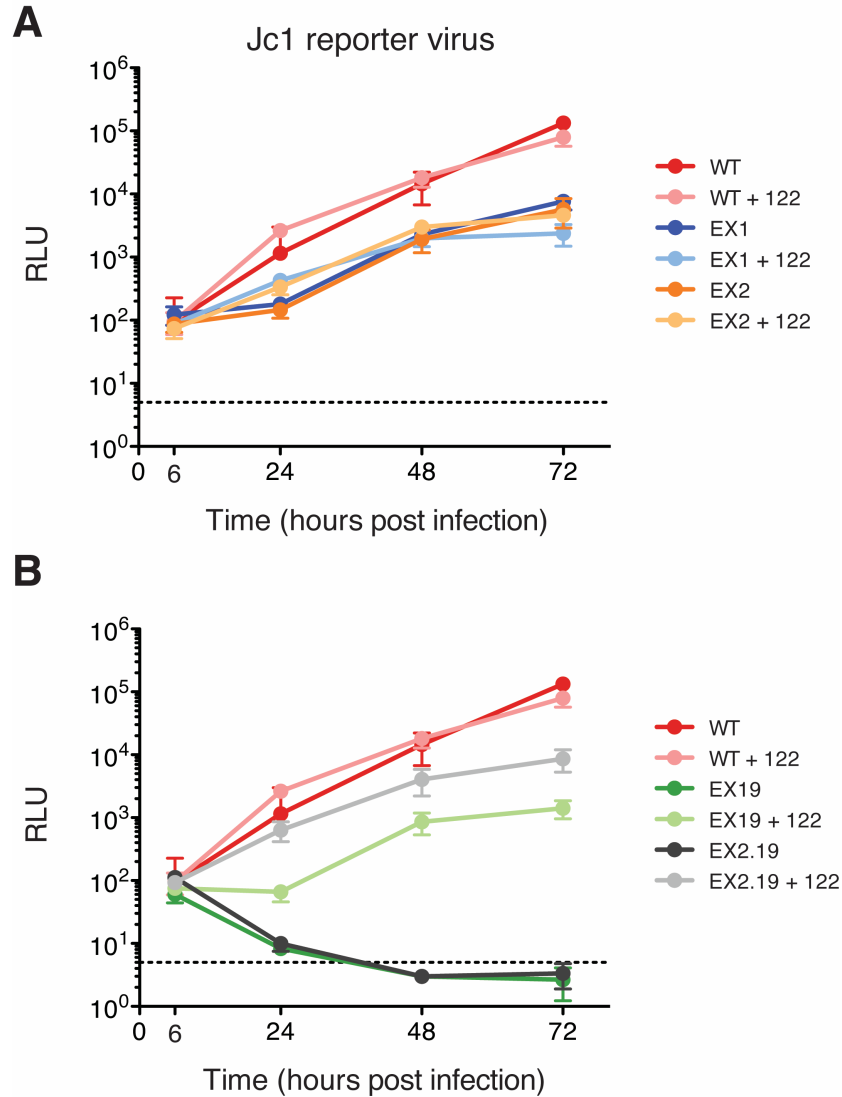


Figure 3.19. HCV replication timecourse measurements in various Dicer CRISPR mutants. Jc1 luciferase reporter virus was used at an MOI of 1 to infect WT Huh-7.5 cells, or Huh-7.5 cells harboring genomic deletions of the indicated coding exons of Dicer. Experiments were performed in the absence or presence (+122) of 30nM transfected miR-122 mimic. Cells were washed at each timepoint measured. Panels A and B present the same WT data for comparison of the remaining mutants. Dashed line indicates mean signal from mock-infected cells. Error bars, \pm SD.

Chapter 3 summary and discussion

By applying AGO-CLIP to a virus that appears to require a miRNA for its replication, we've generated conclusive biochemical evidence of the miR-122 interaction with HCV RNA that overwhelmingly agrees with the genetic evidence. Additional and likely miRNA dependent AGO binding was also observed and may reflect additional roles for small RNAs in HCV biology, though by and large, miR-122 appears to be the most important for this virus. We've also explored multiple host and viral contexts for HCV replication. By deleting miR-122, Drosha, or Dicer genes, we've begun to probe endogenous small RNA function on HCV at the genomic level. Establishing a miR-122 independent virus (U3) in conjunction with swapping HCV miRNA tropism to use miR-15, we've also added new viral contexts to address HCV small RNA use. Driven by CLIP, these results provide answers to many questions that can now be addressed.

One major question still to be resolved in HCV biology is how positive strand genomes "decide" between multiple fates: to undergo translation, to package into virions, serve as templates for negative strand synthesis, or get degraded. miR-122 has been proposed as a regulator in these processes, specifically favoring RNA synthesis, by competing at the S2 site with PCBP2 induced RNA circularization, which favors translation (Masaki et al., 2015). Its aforementioned roles in stabilizing viral RNA and perhaps promoting translation make it seem as though miR-122 indeed plays multiple parts. Thus, our results with CLIP likely capture a mixture of RNAs at various stages of the virus lifecycle.

We can consider future means of leveraging CLIP to address the central issue: at which stage(s) of the viral lifecycle is miR-122 important?

One way to frame this question is to consider whether S1 and S2 sites possess distinct functions. Pioneering work established that miR-122 engagement at both S1 and S2 sites was critically required for HCV replication (Jopling et al., 2008). CLIP data agrees with this, though both S1 and S2 are too close to distinguish individual AGO binding events. One way around this would be to use chimeric viruses, where S1 binds miR-122 and S2 binds miR-15, or vice versa. Preliminary work with Yingpu Yu and Troels Scheel in the Rice lab has shown the viability of such swap mutants. Combined with detailed CIMS analysis and recent AGO-CLIP advances to identify miRNA:target ligation events, the stage is set to functionally unravel S1 and S2 activities in the future (Grosswendt et al., 2014; Helwak et al., 2013)(Moore, et al. Submitted). AGO and miR-122 relocalization to lipid droplets upon HCV infection provides yet another clue to interpret CLIP maps (Ariumi et al., 2011; Berezhna et al., 2011). It's conceivable that an AGO-CLIP experiment can be designed to specifically interrogate AGO association in different subcellular compartments. We can draw inspiration from the recent development of proximity specific ribosome profiling on this front (Jan et al., 2014).

Perhaps a more compelling approach with potential inroads to *in vivo* interrogation of miR-122 usage by a virus is to consider the recently discovered hepaciviruses of horses, rats, and mice (Drexler et al., 2013; Firth et al., 2014;

Kapoor et al., 2013). Non-primate hepacivirus (NPHV), commonly found in horse sera, contains at least one miR-122 site in its 5'UTR and is thought to be engaged with AGO (Scheel et al., 2015). Rodent hepacivirus (RHV) has also been shown to contain a miR-122 site in the 5'UTR (Kapoor et al., 2013). If these viruses represent hepatotropic ancestors of HCV, they may help explain how HCV acquired the usage of miR-122 as a key determinant in the virus life-cycle. Clearly, much work remains to be done in this area.

Of course, given the success of AGO-CLIP in HCV studies, one also gets the impression that even more compelling applications of these methods lie with other RNABPs, in other viruses, and certainly other host contexts per the above discussion. One compelling example are arthropod borne (arbo) viruses that cycle through insect and mammalian hosts and uniquely face two very different immune systems. As mentioned in the introduction, RNAi plays a key role in innate anti-viral defense in insects, while the protein based interferon response appears to dominate in mammalian hosts. AGO-CLIP is thoroughly poised to give an unbiased picture of endogenous small RNA function in both contexts, and may uncover host: viral interactions at the RNA level as targets for therapy. Arboviruses such as chikungunya virus (CHIKV, of the *Togaviridae*), yellow fever and West Nile viruses (YFV and WNV, both of the *Flaviviridae*) are good starting candidates, and current work in the Rice lab is aimed at elucidating small RNAi interaction maps for these viruses in insect and mammalian hosts. That all of the above viruses are emerging threats driven by human induced climate

change lends a sense of urgency to the mission (Fischer et al., 2014; Gould and Higgs, 2009).

The general application of AGO-CLIP as a screening and discovery tool warrants some discussion as we consider how viruses may exploit small RNAs to their benefit, and how to properly interpret CLIP maps. Obviously, viruses encoding their own miRNAs present an intuitive case for AGO-CLIP whereby reads perfectly aligning to a 22-23nt stretch are likely to be miRNA genes, as has been extensively shown for many herpesviruses (Haecker et al., 2012; Riley et al., 2012a; Skalsky et al., 2012). These primarily constitute evidence of AGO loading, while AGO association with mRNA, which yields a broader peak, can be reasonably assumed to be repressive. Clearly HCV requiring miR-122 is an outlier in this regard, and we propose that the differences in the proportional amount of AGO binding (its "peakiness") across the viral genome may indicate whether a small RNA is required by the virus. HCV, where over 50% of all AGO binding events map to the first 50nt of the viral genome would largely bear this out. AGO-CLIP work in *Pestiviruses* by Troels Scheel has identified miRNA dependence among certain members of this virus family, where again, a single peak predominates (unpublished observations). In the case of requiring a small RNA, we would expect the amount of AGO binding to be both peak-like and to positively correlate with RNA abundance, an observation we were able to confirm with HCV comparing WT and GNN timecourses (Figure 3.6).

Repressive events are somewhat harder to interpret on the basis of AGO-CLIP results alone. Given that AGO binding typically results in RNA degradation, CLIP results that "catch" these moments should at steady-state be an underestimation of the amount of AGO binding at these loci. As a result, AGO-CLIP alone is not sufficient to conclude which binding events on viral genomes are repressive and which may be informatics noise. This is especially true in light of CLIP results on viruses that do not yield peaks, but instead are broadly "coated" with AGO binding (not shown). However, by pairing AGO-CLIP with RNAseq (or for a single virus, qPCR), and then manipulating small RNA levels with inhibitors or as genomic deletions, one can overcome this limitation and focus on the *change* in AGO association in these contexts. This is essentially what was done with the U3 virus in Δ miR-122 and Δ Drosha cells. Establishing what a true AGO-CLIP signature of viral RNA repression can also be done by miRNA overexpression or by forcing an RNAi response (with siRNA) and watching how AGO binding changes. These experiments have not yet been systematically performed and no doubt provide a compelling jumping off point for future studies.

Chapter 4: Argonaute HiTS-CLIP studies of HCV infected host cells

In the previous chapter, we focused on the insights from AGO-CLIP on the HCV genome in a variety of infection contexts. This arguably constitutes the main goal of AGO-CLIP in HCV infected cells. But part of the utility of CLIP is that it provides a wealth of data on the host transcriptome in addition to the virus. With this data, we can begin to piece together the cellular miRNA regulatory response to a viral infection in an unbiased manner. To do this, we'll begin by analyzing the small RNA populations within AGO and how they are similar and different from total RNA based measurements. We'll then move on to a novel and globally observed phenomena of miR-122 sponging by HCV RNA. This sponge effect will be used to establish the miR-122 regulated target network in Huh-7.5 cells and will be the subject of rigorous experimental examination and validation. It is remarkable, but to some degree unsurprising in hindsight, that the critical miR-122:HCV interaction could have a specific and de-repressive effect on the host miR-122 targetome. The results to be presented here simultaneously offer a global yet transcript specific view of miR-122 sponging by HCV, for which many possible functions and consequences will be discussed.

AGO-CLIP miRNA abundance measurements - correcting ligation bias

Perhaps the most basic analysis to perform with AGO-CLIP data is the determination of the AGO associated miRNA profile. This can either be done by aligning groomed reads to a database of mature miRNA sequences (such as

miRbase) or by aligning to the genome and counting reads overlapping with annotated miRNA coordinates. The former approach is the easiest to carry out but is constrained by exact sequence; the latter is constrained mainly by mapping coordinates and has the potential to pick up novel miRNAs.

Applying either approach we noted significant differences between previous data on miRNA abundance in Huh-7.5 cells and results with AGO-CLIP. In previous small RNA cloning measurements from the lab, miR-122 was measured to be the most abundant miRNA in Huh-7.5 cells, at around 20% of total miRNA (Randall et al., 2007). In contrast, AGO-CLIP miRNA abundance measurement placed miR-122 as the 65th most abundant miRNA, closer to 0.16% of the total (Figure 4.1A). Combined with miRNAs that displayed enrichment in AGO-CLIP, very little correlation was observed with previous small RNA cloning results (Figure 4.1A).

We can extend this analysis further by inferring the miRNA profile based on mRNA clusters. This was based on the intuitive idea, first brought up by (Chi et al., 2009), that counting miRNA seeds within mRNA clusters, and weighting by cluster read density could be used to infer the underlying miRNA profile. In other words, the AGO abundance on mRNA targets likely reflects which miRNAs are actually used. The resulting target normalized total frequency showed that miR-122 was among the top ten most frequent seeds within mRNA clusters, indicating its functional abundance on AGO. And yet, CLIP derived miRNA abundance measurements displayed no correlation with the inferred profile (Figure 4.1B).

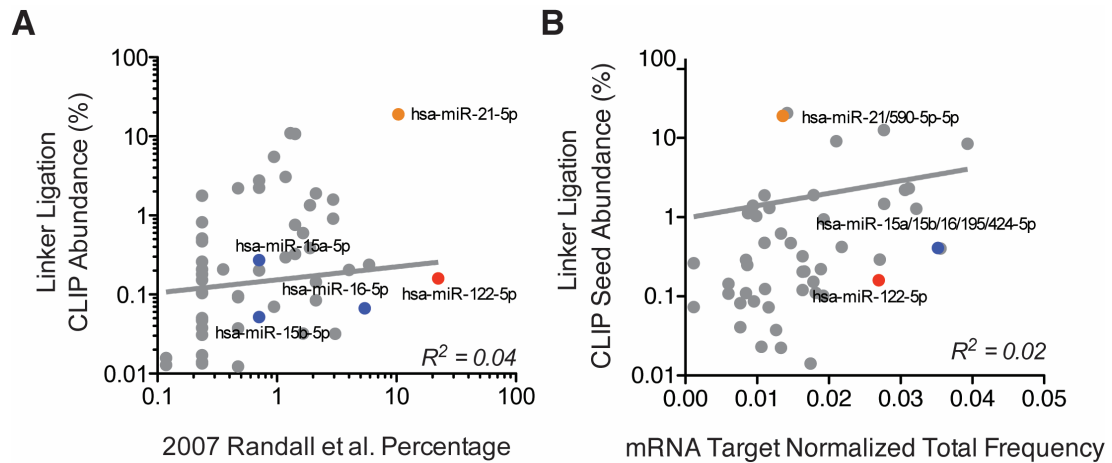


Figure 4.1. miRNA bias in linker- ligation dependent AGO-CLIP. **(A)** Scatterplot comparing published miRNA percent abundance measurements from total RNA (Randall et al., 2007) to linker ligation CLIP. Squared correlation coefficient of the linear regression best-fit line is shown. miR-122, miR-15/16 and miR-21 families are highlighted. **(B)** Scatterplot comparing linker ligation CLIP miRNA percent abundance to the normalized abundance of miRNA seeds on mRNA clusters, from linker ligation CLIP, as in (Chi et al., 2009). Squared correlation coefficient of the best-fit line is shown.

There are a few possible reasons for this discrepancy, at least one biological and two technical. As most miRNA measurements are made from biochemically purified total RNA where levels are assumed to be functional (that is, AGO associated), one possibility is that the assumption of functionality does not hold. Conceivably, the miRNA profile from total RNA could be categorically different from the miRNA profile that is actually AGO associated. Such a bold biologic hypothesis would challenge most miRNA measurements made from total RNA as they are devoid of cellular context and RNABP interactions key to their function. Perhaps AGO-CLIP provides the real answer?

As attractive as the above hypothesis may be at first, there are two technical concerns related to cloning bias that warrant attention: the effect of UV crosslinking and linker ligation. As UV crosslinking is thought to work on the order of ångstroms, perhaps different miRNAs sit in the AGO binding pocket each in their own way and as a result are not uniformly cross-linked. Added to this is the concern that crosslinking, while amenable to stringent purification, may differentially impact a miRNA's ability to be reverse transcribed into cDNA due to the crosslinked amino-acid:RNA adduct that may interfere with RT. We can use this last point by considering CIMS deletions on miRNAs to see if the pattern and amount of CIMS deletions varies as a function of miRNA abundance. Shown in Figure 4.2, we can appreciate the uniqueness of CIMS deletions per miRNA in line with previous analyses in brain (Zhang and Darnell, 2011). Notably, CIMS to miRNA seeds is largely absent, which agrees with structural studies of miRNA occupancy in AGO (Schirle et al., 2014). The most crosslinked nucleotides are between positions 8-12, in regions not generally observed to base pair with targets. Additionally, among this representative cohort of miRNA examined, all four RNA bases are somewhat equally crosslinked (7U, 3G, 6A, 5C). Perhaps the most striking feature is the percent deletions across miRNAs. The most abundant miRNA measured by CLIP is miR-21 (19% of total reads), less than 0.5% of these reads harbored CIMS deletions. In contrast, over 6% of miR-122 reads (0.16% of total) exhibited CIMS. By and large, it appears as if the most abundant miRNAs in CLIP are among the least cross-linked. However, whether this is a

cause or a symptom of the altered miRNA profile on AGO compared to total RNA is unclear. Cloning miRNAs from non-crosslinked AGO IPs revealed a similar miRNA profile to CLIP in these cells, with drastically lower depth (not shown). Taken together, these results imply, but as yet do not confirm, a cross-link bias in miRNA abundance measurements with AGO-CLIP.

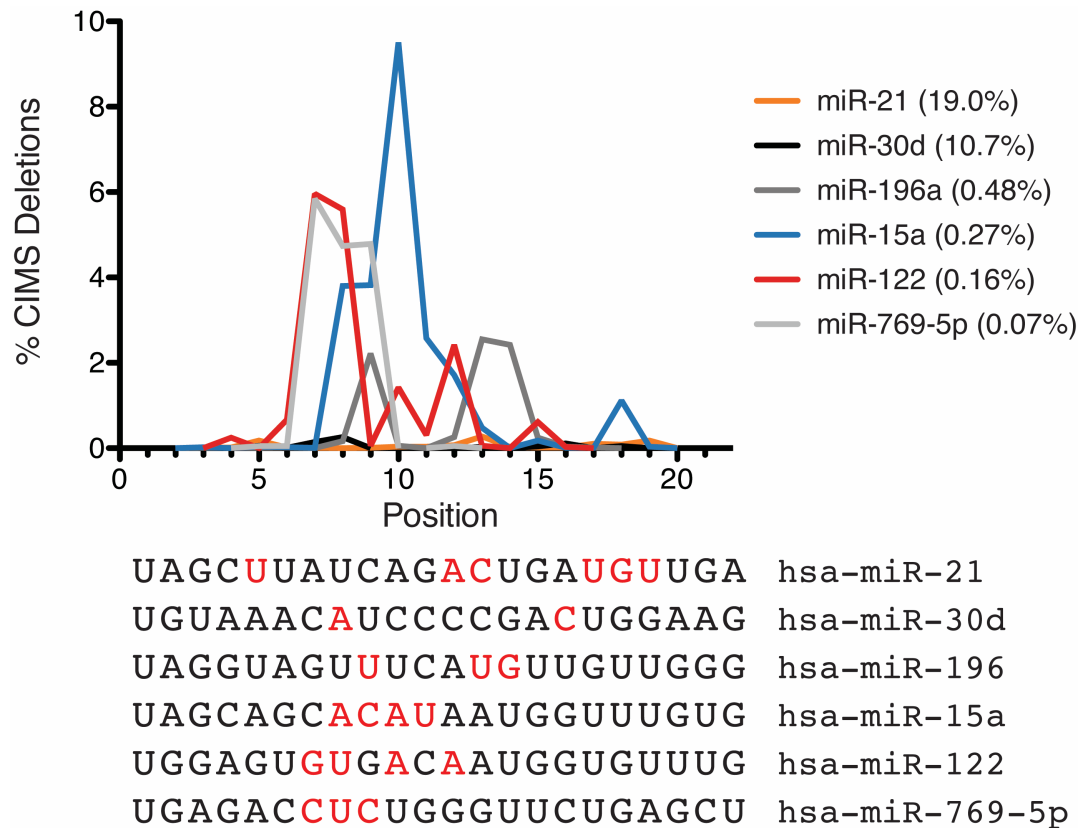


Figure 4.2. CIMS analysis of select miRNAs from AGO-CLIP. The percent of reads mapping to each respective miRNA with CIMS deletions is plotted by 5' to 3' position. miRNAs are ordered by decreasing CLIP abundance. miRNA sequences are shown below with significant CIMS nucleotides highlighted in red.

The second technical concern, and one that compounds the potential cross-link bias raised above, is that of cloning bias. Stemming from sometimes

conflicting results across labs, over the past decade it has been increasingly clear that the dual RNA linker ligation steps using T4 RNA ligase to clone small RNAs can dramatically influence the resulting miRNA profile, and must be taken into account. The key enzyme involved, RNA ligase 1 (Rnl1) from T4 DNA bacteriophage, evolved primarily to counter a host-mediated defense mechanism whereby the anti-codon loop of lysine charged tRNAs is cleaved in phage infected cells (Amitsur et al., 1987; Silber et al., 1972). By repairing cleaved tRNA with Rnl1 and T4 poly-nucleotide kinase (T4 PNK), T4 DNA phage is able to ensure availability of tRNA-Lys (Amitsur et al., 1987). Not surprisingly, more recent work has found that RNA ligation of adapters to miRNAs using Rnl1 prefers structural features that most closely mimic anti-codon loops, and as a result introduce cloning bias (Zhuang et al., 2012). Solutions to this problem range from using RNA ligase mutants with altered ligation efficiencies to modifying RNA adapters to a random pool (Brown et al., 2013; Hafner et al., 2012; Zhuang et al., 2012).

As an alternative, we opted to try an RNA ligation independent approach (Figure 4.3A). Inspired by ribosomal profiling cloning methods (Ingolia et al., 2009), we pursued an approach where miRNAs are tailed with G nucleotides (see methods). This tail then serves as a handle for reverse transcription using a primer containing linker sequences separated with an APE1 ssDNA digestion site. Following RT, the cDNA is circularized with an ssDNA ligase, and cleaved with APE1 to generate a cDNA product amenable to PCR. This cloning

technique, called polyG-CLIP, thus provides an RNA ligation independent means to quantify miRNAs from crosslinked AGO IPs. Note however, that polyG-CLIP isn't totally ligation free: an ssDNA ligase, modified from an RNA ligase is used.

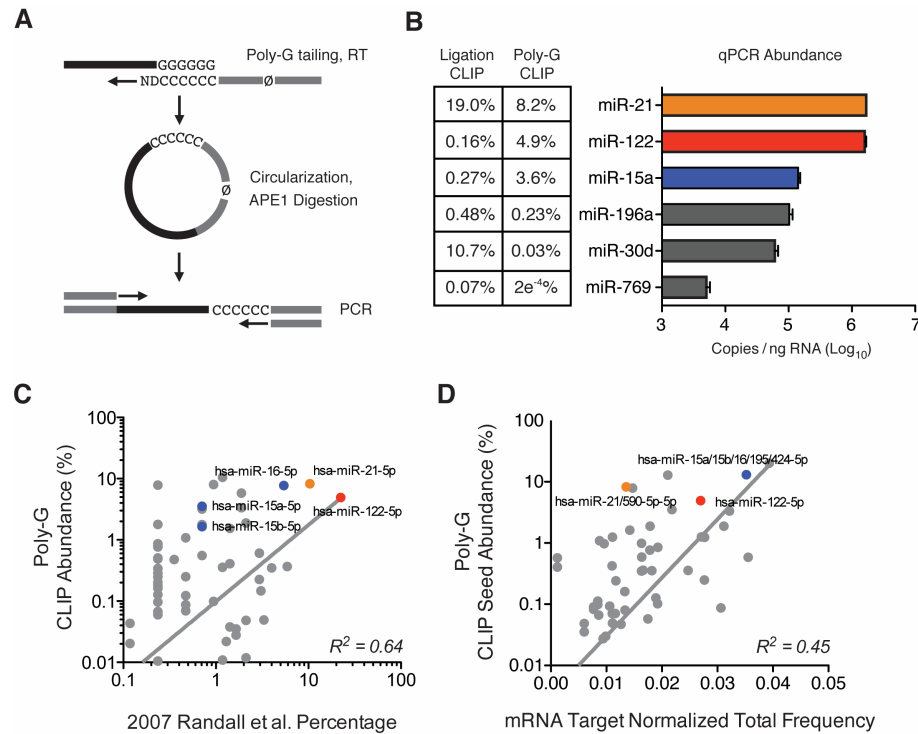


Figure 4.3. miRNA bias correction with polyG-CLIP. **(A)** . A 3' adapter ligation independent miRNA cloning strategy consists of poly-G tailing small RNA (in black) followed by reverse transcript (RT) using an anchored oligo (in gray) containing an APE1 cut site. Resulting cDNA products are circularized and re-linearized with APE1 nuclease, prior to PCR. **(B)** Absolute miRNA quantification of select miRNAs measured by qPCR, ordered by decreasing abundance. The table highlights the abundance of these miRNAs via linker ligation or poly-G CLIP. **(C)** Scatterplot comparing published miRNA percent abundance measurements from total RNA (Randall et al., 2007) to polyG CLIP. Squared correlation coefficient of the linear regression best-fit line is shown. miR-122, miR-15/16 and miR-21 families are highlighted. **(D)** Scatterplot comparing polyG CLIP miRNA percent abundance to the normalized abundance of miRNA seeds on mRNA clusters, from linker ligation CLIP, as in (Chi et al., 2009). Squared correlation coefficient of the best-fit line is shown.

Applying polyG-CLIP and comparing to qPCR results of select miRNAs, we noted that polyG-CLIP agreed with qPCR abundance order, unlike standard ligation based CLIP (Figure 4.3B). Moreover, comparisons to previous miRNA data yielded a much higher correlation (Figure 4.3C). Poly-CLIP miRNA measurements were also well correlated to the inferred miRNA profile from standard ligation based CLIP (Figure 4.3D). Taken together, these results strongly suggest a miRNA cloning bias as a result RNA ligation that can be corrected with alternate approaches.

And yet one question lingers regarding the standard RNA ligation approach: why would bias be observed on the miRNA profile, but not on mRNA targets? I reason that this is due to the nature of the RNA cloned: unlike an individual miRNA with a fixed nucleotide sequence (and not thought to be accessible by RNase), mRNA targets undergo a partial RNase digestion such that a variety of 5' and 3' ends are generated for each target. As a consequence, the "cloneability" of an mRNA target is likely to be much more uniform than a miRNA, and thus mRNAs should not exhibit such a drastic bias. This would explain why the inferred miRNA from mRNA clusters in standard CLIP agrees with polyG-CLIP miRNA measurements (Figure 4.3D). While polyG-CLIP appeared to be superior to quantifying miRNAs, the overall read depth for both miRNAs and mRNAs was lower compared to standard CLIP (not shown). For miRNAs, few in number, short and relatively abundant, the lower depth did not present a major hurdle for quantification, as showcased above. The same cannot

be said for mRNAs, whose much lower read depth in polyG-CLIP (compared to standard CLIP), limited the utility of this technique for studying mRNA targets.

As a compromise, polyG-CLIP was used to determine the AGO bound miRNA profile, while the standard RNA ligation based approach was used to determine mRNA targets. For subsequent analysis on mRNA-CLIP clusters, we focused on searching the top 50 seed families derived from poly-G CLIP studies, which constituted over 97% of miRNA identified in Huh-7.5 cells.

As a last point, it should be noted that addressing the above concerns does not preclude the initial biologic interpretation, that the AGO associated miRNA profile may be categorically different from the profile measured from total RNA. Indeed, despite correcting for cloning bias, the correlation between polyG-CLIP and either total small RNA cloning or the inferred miRNA profile is not perfect (Figure 4.3C-D). It remains plausible, perhaps even likely, that miRNA measurements from total RNA can miss the real picture of AGO:miRNA function precisely because they fail to capture the functional RNABP complex in action. Future work based on AGO-CLIP provides a fantastic means to address this largely overlooked and somewhat mysterious assumption (Chen, 2013).

A tour of AGO binding in Huh-7.5 cells

Before we go into global and somewhat abstract bioinformatic analyses of AGO-CLIP targeting during HCV infection, it's perhaps useful to get a feel for the data by simply looking at it in a genome browser. In this manner, we can observe what AGO binding an mRNA transcript looks like "in the wild" so to speak. This

provides the most accessible introduction to viewing these data, and it's also where I began historically, not knowing what else to do with a 2GB sequencing file. Exploring it with a genome browser seemed as good a place to start as any, and importantly, was the jumping off point for many hypotheses to be discussed.

But before we launch our guided safari, a few particulars. mRNA CLIP reads were aligned to the human genome and clustered (defined here as any two reads overlapping by at least 1nt). As a quick sanity check, we can perceive the importance of biologic replicates by considering where reads annotate from any one or more independent libraries. The idea here is that the more something is independently and repeatedly confirmed, the more likely it is to reflect the true global picture of RNABP activity. Such an analysis was first performed with the 3'UTR binding nElavl protein CLIP in mouse brain, where increasing biologic complexity increased the 3'UTR signal (Ince-Dunn et al., 2012). These results showed that repetition can dramatically lower spurious clusters from individual libraries. Performing a similar analysis on AGO-CLIP data from Huh-7.5 cells, we observed something very similar (Figure 4.4). Here, clusters in at least one library are great in number (>250K) but mostly map to introns and deep intergenic regions, and are thought to represent one-off spurious noise. As we increase the stringency for clusters, where constituent reads must come from two, three and so on up to six libraries, we can appreciate how the percentage of clusters mapping to 3'UTRs and CDS exons, places of well known AGO association, increases with BC. Note however, that the total number of clusters decreases

with increased BC stringency. This tradeoff introduces a decision point in defining a cutoff for what is significant. While the lowest BC contains all true positives by definition, it is plagued by a high number of false positives. On the other hand, the highest BC is likely too stringent as it selects for highly abundant (or highly clonable) mRNAs at the expense of lower abundant, but true, mRNA targets. In other words, it suffers from a false negative problem. Thus an arbitrary decision must be made. For the purposes of these studies, we split it down the middle: I stipulated a minimum BC of the number of libraries divided by two, in the case of these six libraries (3 Mock, 3 +HCV), we consider clusters composed of reads from any three libraries at a minimum.

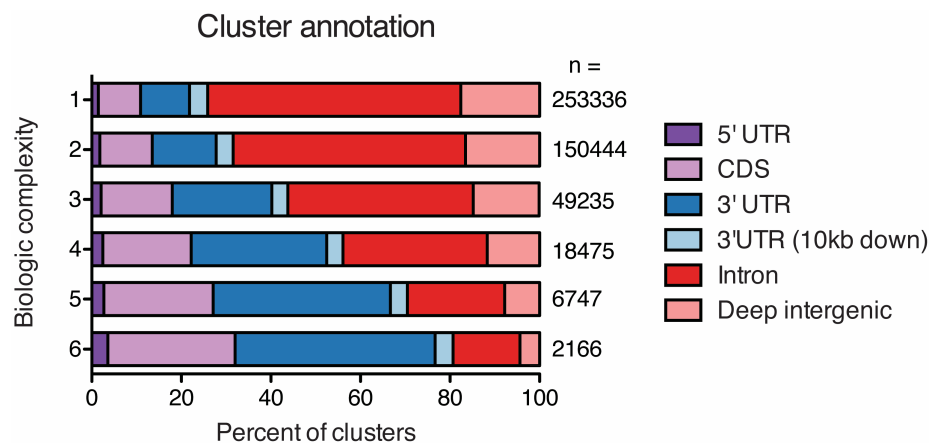


Figure 4.4. Cluster annotation results by increasing biologic complexity. AGO-CLIP reads from up to six libraries were clustered with a minimum overlap of 2nt. Clusters with reads from any one, two, three and so on up to six libraries the were annotated. The resulting proportions of clusters overlapping known genic elements is plotted above, alongside the total number of clusters with the respective biologic complexity (BC).

Now, to begin our tour of AGO bound mRNAs, we'll start by focusing on the hepatic nature of Huh-7.5 cells. The parental line for these cells, Huh-7, was originally isolated in 1982 from a well differentiated hepatocellular carcinoma in a 57-year-old Japanese male (Nakabayashi et al., 1982). As such, numerous hepatic transcripts are maintained in these cells befitting their liver origin. One notable transcript 3'UTR bound by AGO is 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR) (Figure 4.5A), the rate limiting enzyme for cholesterol biosynthesis and famously, the main target of cholesterol reducing statin drugs. Interestingly, miR-122 inhibition in mice and in chimpanzees reduced serum cholesterol, and in mice, reduced HMGCR expression was observed (Elmén et al., 2008; Esau et al., 2006; Krützfeldt et al., 2005; Lanford et al., 2010). Given the importance of miR-122 to HCV replication and to host lipid biosynthesis, later work explored a possible connection between the two and found essentially that there was none: miR-122's activity on HCV and on regulating cholesterol biosynthesis occurred via separate and distinct mechanisms (Norman and Sarnow, 2010). With no miR-122 sites and so far no confirmed miRNA sites on HMGCR, the above data point to some role for miR-122 in regulating an inhibitor to cholesterol biosynthesis upstream of HMGCR (Norman and Sarnow, 2010). This possibility remains to be fully explored. Data presented in Figure 4.5A open the possibility of direct miRNA regulation of HMGCR where interestingly, miR-21 was recently shown to regulate HMGCR mRNA in a cellular model nonalcoholic fatty liver disease (NAFLD) (Sun et al., 2015).

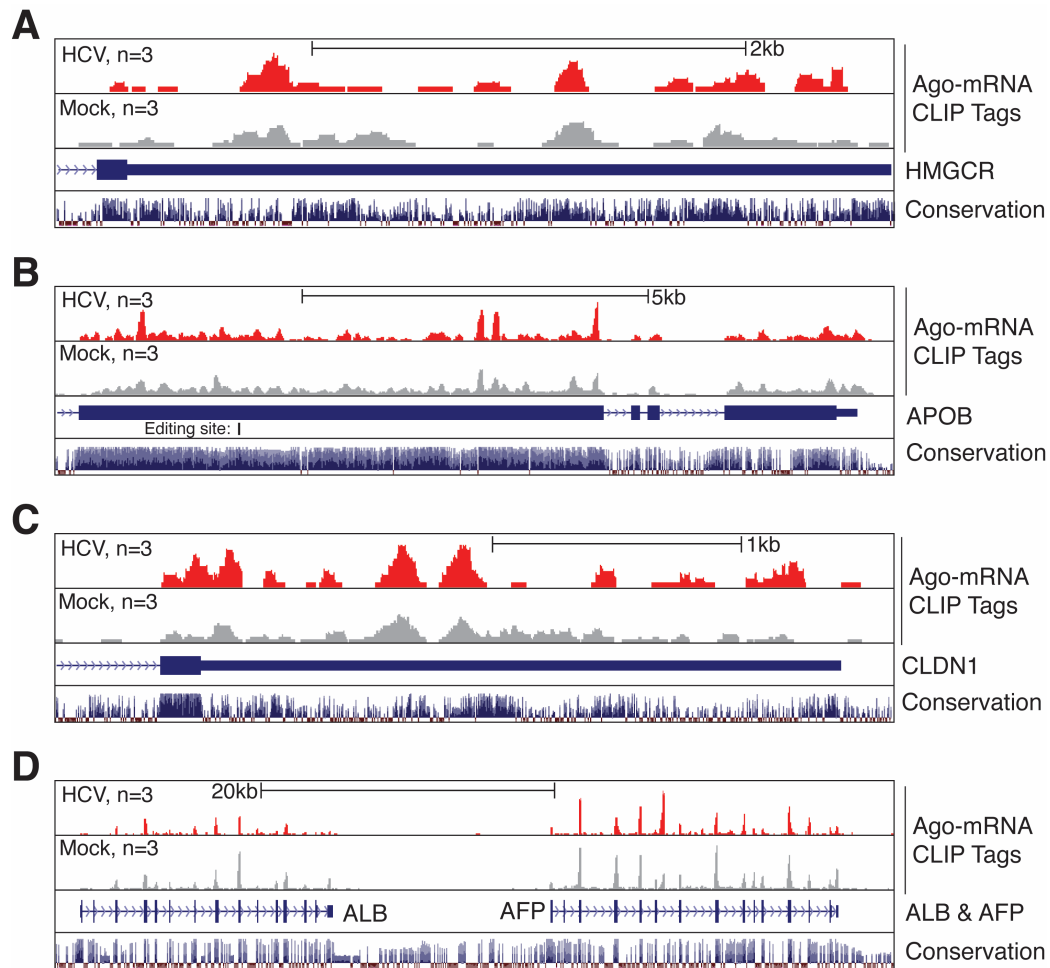


Figure 4.5. Genome browser examples of AGO-CLIP binding to notable hepatic transcripts. AGO binding in the presence of WT HCV or uninfected cells in the upper red and gray tracks respectively, from triplicate experiments. Reads are plotted as normalized densities to enable cross track comparison. miR-122 family seeds noted by red boxes, peaks changing between conditions denoted by asterisks. PhyloP conservation track shown at the bottom. Genes: **(A)** 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), **(B)** apolipoprotein B (APOB) CAA to UAA editing site indicated, **(C)** claudin-1 (CLDN1), and the **(D)** albumin and alpha-fetoprotein (ALB and AFP) locus.

Consistent with a hepatic origin of Huh-7.5 cells is robust expression (and surprising AGO targeting) of apolipoprotein B (APOB) (Figure 4.5B). Long known to be a key player in very low density lipoprotein (VLDL) secretion, APOB is also

associated with HCV lipoviroparticles (LVPs) (Felmlee et al., 2013). The APOB transcript features an unusually long coding exon that in the liver results in a 100kDa protein (ApoB100), whereas in the small intestine this exon is edited by APOBEC1 to produce an in-frame stop codon, resulting in a truncated 48kDa protein (ApoB48) critical for chylomicron function. The long APOB exon is coated with AGO binding, much of it downstream of the editing site (Figure 4.5B). Could this be a surveillance mechanism to ensure that any errantly edited APOB transcript does not result in production of the 48kDa protein in liver? This hypothesis remains to be tested.

Claudin-1 (CLDN1) mRNA, encoding a tight junction protein and co-receptor for HCV, is also bound by AGO via putative miR-155 sites (Evans et al., 2007; Qin et al., 2013) (Figure 4.5C). Interestingly, Claudin-1 is upregulated in cirrhosis and HCC, suggesting a tumor suppressive role for miR-155 (Holczbauer et al., 2014). This parallels observations with occludin (OCLN) another HCV entry factor and tight junction protein, whose transcript contains miR-122 sites (Ploss et al., 2009; Sendi et al., 2014). Recent work has shown that miR-122 expression reduces OCLN mRNA in hepatocytes, and in this manner prevents HCV entry (Sendi et al., 2014). Whether this represents a bonafide super-infection exclusion mechanism remains to be determined.

As a fulcrum that captures both the hepatocyte and HCC derived nature of Huh-7.5 cells, it is notable that both albumin (ALB) and alpha-fetoprotein (AFP) are robustly expressed and bound by AGO (Figure 4.5D). The former protein is

the quintessential liver product and is the most abundant blood protein in mammals, where it's required to maintain osmotic pressure of body fluids (Theodore Peters, 1995). AFP is thought to be the fetal form of serum albumin and is normally not present in the adult liver, but is a common marker for tumors, especially HCC (Bialecki and Di Bisceglie, 2005). Whether AGO binding to the coding exons of these genes plays any role in their regulation is completely unknown.

Of course, AFP is not the only indicator of the HCC origin of these cells. Both MYC and MYCN oncogenes are expressed, the latter while particularly associated with poor progression in neuroblastoma, is highly bound by AGO in Huh-7.5 cells (Figure 4.6A). The myc family of transcription factors are well known drivers of HCC and recent work has implicated tumor suppressive miRNAs (including miR-122) participating in feedback loops to inhibit myc expression (Han et al., 2013; Wang et al., 2013a). PEG10, also highly bound by AGO, is a paternally expressed imprinted gene that, like myc, is upregulated in many HCC cases and uniquely in the regenerating liver (Okabe et al., 2003; Tsou et al., 2003) (Figure 4.6B). At least two non-coding RNAs involved in cancer are also bound by AGO, one aptly named highly upregulated in liver cancer (HULC) (Matouk et al., 2009) (Figure 4.6C-D). The other XIST, the key non-coding effector of X-inactivation, is not normally expressed in males but is known to be deregulated human cancers, regardless of gender (Weakley et al., 2011). As the function of a miRNA is typically cast as a repressor of protein expression, finding

non-coding RNAs bound by AGO presents a bit of a conundrum. What is AGO doing? Long noncoding RNA biology is at an interesting frontier in exploring this, and perhaps other RNABP interactions in various disease states (Shi et al., 2013).

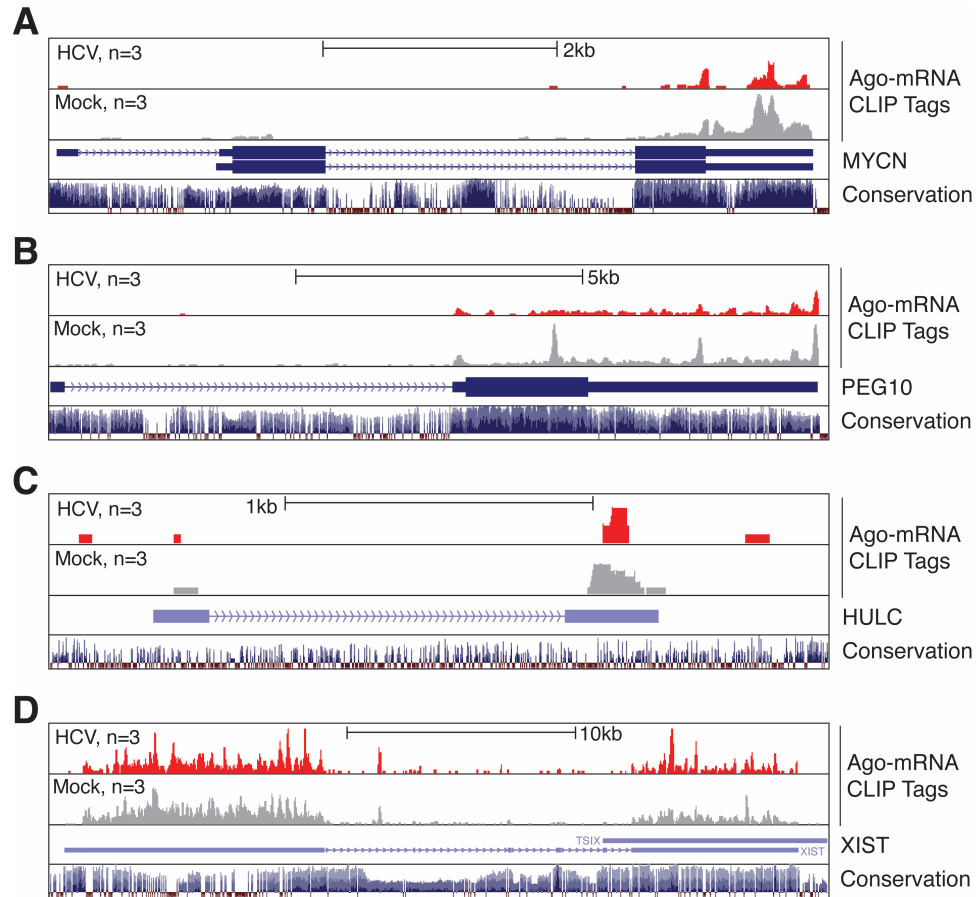


Figure 4.6. Genome browser examples of AGO-CLIP binding to notable coding and non-coding transcripts associated with HCC. AGO binding in the presence of WT HCV or uninfected cells in the upper red and gray tracks respectively, from triplicate experiments. Reads are plotted as normalized densities to enable cross track comparison. miR-122 family seeds noted by red boxes, peaks changing between conditions denoted by asterisks. PhyloP conservation track shown at the bottom. Genes: **(A)** v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog (MYCN), **(B)** paternally expressed 10 (PEG10), **(C)** hepatocellular carcinoma up-regulated long non-coding RNA (HULC), and **(D)** X inactive specific transcript (non-protein coding).

As the above examples indicate, far more questions and testable hypotheses than answers are raised by observing AGO-CLIP data with "boots on the ground" level views. The descriptive nature of these data imply a role for miRNAs in regulating a variety of liver and cancer specific pathways, some not having anything to do with HCV. Indeed, all of the above examples, with the exception of PEG10, did not show appreciable change upon HCV infection compared to mock controls. While certainly a rich dataset for understanding liver or cancer biology, more comprehensive and miRNA focused means of analysis are clearly needed to make sense of these miRNA:mRNA interaction maps and their shifting role during HCV infection.

As the molecular link between AGO and HCV RNA, this is where miR-122 becomes important. Consider three known miR-122 targets (Figure 4.7). The first, initially called cationic amino acid transporter 1 (CAT-1) was historically the first miR-122 target validated (Chang et al., 2004). This transcript contains three 7mer sites of which AGO binding is detected in two of them and interestingly, appears reduced upon HCV infection (Figure 4.7A). A similar observation was made on citrate synthase (CS) and pyruvate kinase muscle isoform 2 (PKM2) mRNAs, both confirmed miR-122 targets (Jung et al., 2011; Krützfeldt et al., 2005) (Figure 4.7B-C). Searching within AGO mRNA peaks for miR-122 seeds to putatively define novel miR-122 targets, we observed a similar pattern. For miR-122 seeds in the 3'UTRs of SLC1A5, SFT2D1 and CANX mRNAs, AGO binding was again reduced upon virus infection (Figure 4.8). Taken together, these

observations suggest a link between the HCV dependence on miR-122, and de-repression of endogenous miR-122 targets. To fully explore this, we must consider miR-122 targets as a group and compare how targets of other miRNAs as a group respond to HCV infection.

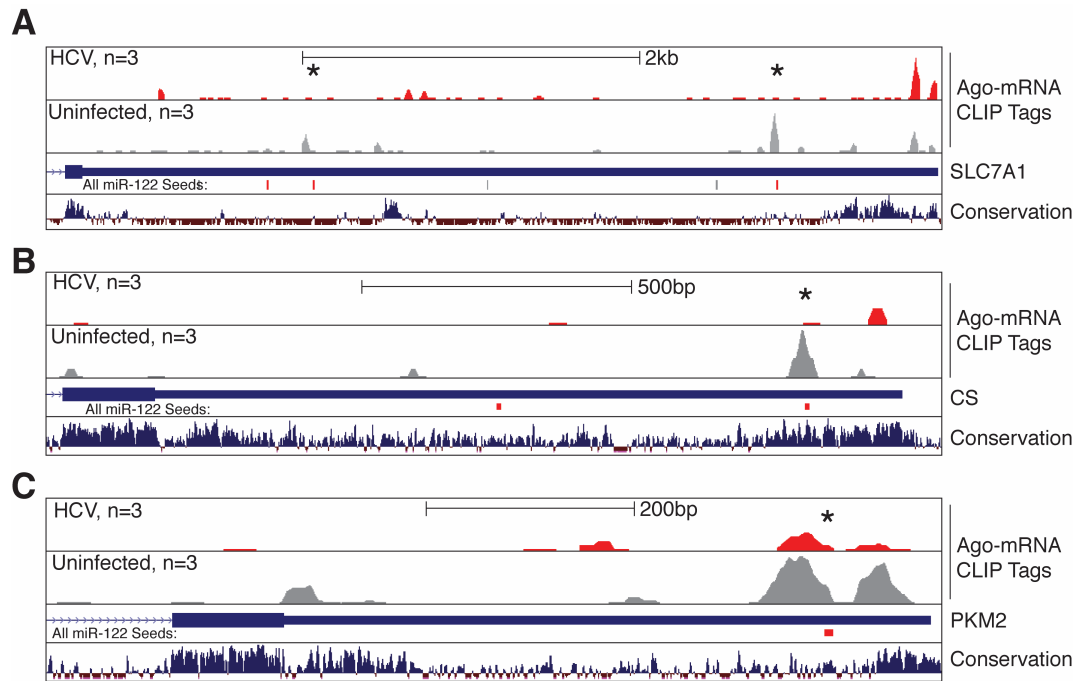


Figure 4.7. Genome browser examples of AGO-CLIP binding to previously known miR-122 targets. AGO binding in the presence of WT HCV or uninfected cells in the upper red and gray tracks respectively, from triplicate experiments. Reads are plotted as normalized densities to enable cross track comparison. 7mer miR-122 family seeds noted by red boxes (6mers in gray), peaks changing between conditions denoted by asterisks. PhyloP conservation track shown at the bottom. Genes: (A) solute carrier family 7 (cationic amino acid transporter, y+ system), member 1 (SLC7A1, aka CAT-1), (B) citrate synthase (CS), (C) pyruvate kinase muscle isoform 2 (PKM2).

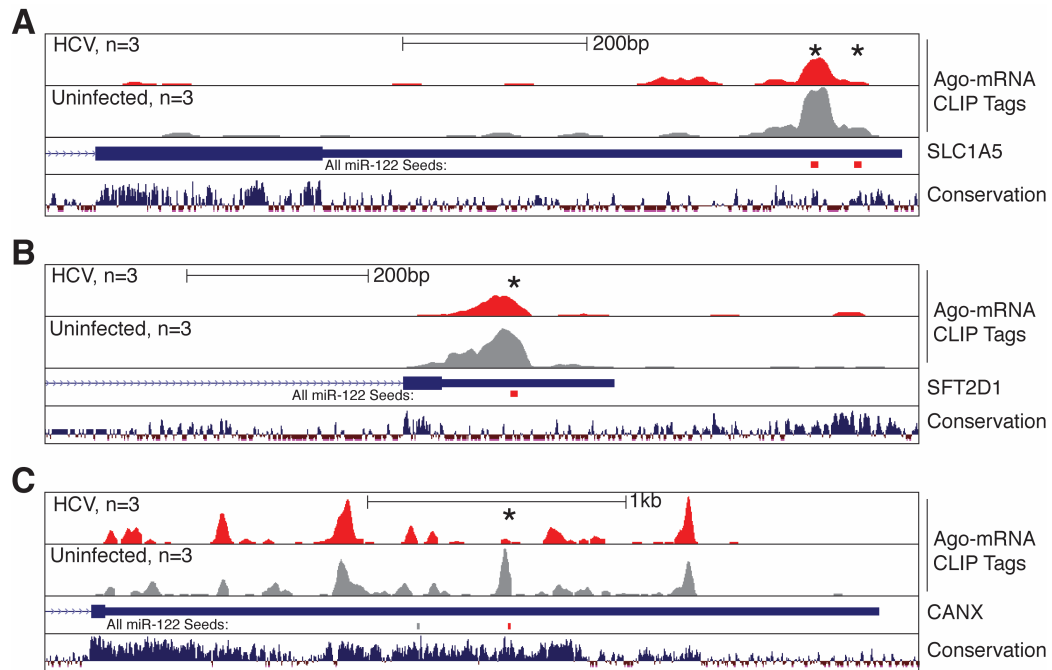


Figure 4.8. Genome browser examples of novel AGO-CLIP identified miR-122 targets. AGO binding in the presence of WT HCV or uninfected cells in the upper red and gray tracks respectively, from triplicate experiments. Reads are plotted as normalized densities to enable cross track comparison. miR-122 family seeds noted by red boxes, peaks changing between conditions denoted by asterisks. PhyloP conservation track shown at the bottom. Genes: **(A)** solute carrier family 1 (neutral amino acid transporter), member 5 (SLC1A5), **(B)** vesicle transport protein SFT2A (SFT2D1), **(C)** calnexin (CANX).

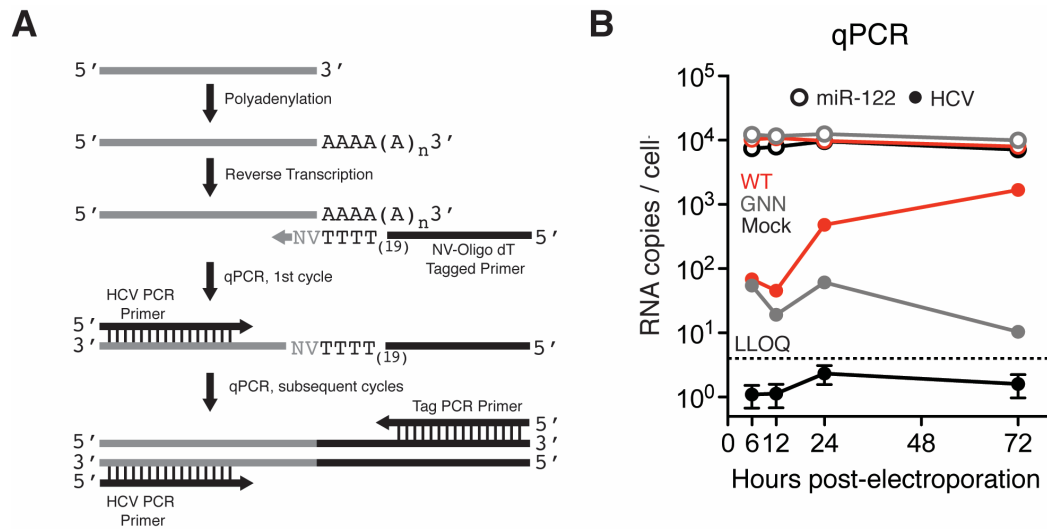


Figure 4.9. An common assay to detect miRNA and HCV RNA levels. **(A)** Non poly-adenylate RNAs (such as HCV RNA or miRNAs) are poly-adenylated and reverse transcribed with a tagged oligo-dT primer capable of recognizing the junction with non-polyA. The resulting cDNA uses an HCV 3'UTR primer or a sense miRNA primer and the common tag primer for qPCR. **(B)** Absolute qPCR measurements of miR-122 and HCV RNA levels at indicated time points post electroporation (n=3). Replication-deficient J6/JFH1-GNN and mock controls are shown. Dashed line indicates lower limit of quantitation. Error bars, \pm SD.

HCV infection functionally reduces AGO binding on host miR-122 targets

Given the crucial requirement of miR-122 for HCV replication, and in light of the result that HCV RNA levels accumulate to within one log of miR-122 levels (Figure 4.9), we hypothesized that the HCV genome may act as a "sponge" for cellular miR-122, where viral replication may exert a broadly de-repressive effect on host miR-122 targets. This would be consistent with CLIP results on miR-122 targets presented thus far. Two key predictions of the data can be made based on this hypothesis: (1) AGO binding to miR-122 targets as a population should be reduced upon HCV infection, and consequently, (2) the aggregate expression of

this population should increase. We can globally explore this bioinformatically by assigning a miRNA to all CLIP derived mRNA clusters with a miRNA seed, computing the fold change at a given cluster between HCV and mock infected cells, and plotting these values as a cumulative distribution. I assigned a miRNA family to clusters bearing seeds based on polyG-CLIP abundance of the top 50 expressed miRNA families. In cases where a cluster could be bound by more than one miRNA family due to the presence of multiple seeds, both miRNA families were assigned to that cluster. The median number of miRNA seeds per cluster in these data was 1.4, in line with previous estimates (Farh et al., 2005). A log₂ fold-change between HCV or mock infected cells score was calculated per miRNA-assigned cluster, and the cumulative distribution of these score was plotted, binned by miRNA family. In these plots, if there was no difference in AGO binding between HCV and mock infected cells, the score for that cluster was near 0. Changes showing a decrease of AGO binding upon virus infection would be <0, while increases in AGO binding would be >0 (Figure 4.10A).

Viewing the entirety of the data in this manner and focusing on 7-8mer seeds, we found that miR-122 3'UTR targets as a population exhibited reduced AGO binding upon virus infection (Figure 4.10B). This was highly specific to miR-122 targets, as the targets of the miR-15 family, or the top ten miRNA families exclusive of miR-122, exhibited no such change. Significant changes in miR-122 binding were observed for all canonical seed types (as defined in (Bartel, 2009)).

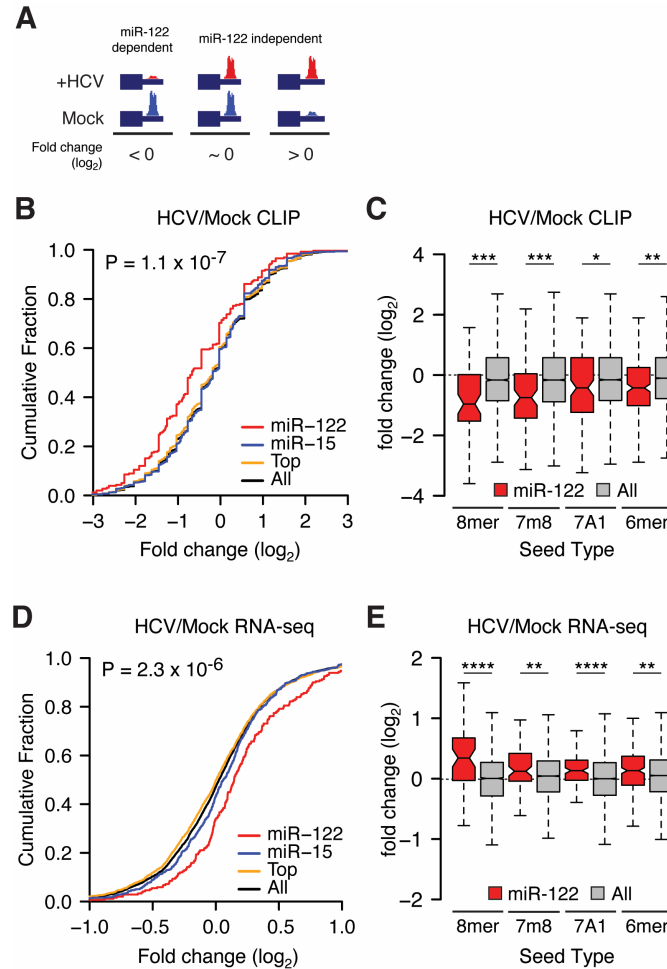


Figure 4.10. HCV infection de-represses endogenous miR-122 targets. **(A)** miR-122 dependence definition by Ago binding profiles. **(B)** Cumulative density function (CDF) of the log₂ fold change in CLIP binding between infected and uninfected cells for all 3'UTR clusters containing indicated 7-8mer seeds by family, from triplicate experiments. "Top" refers to the top 10 miRNA families, exclusive of miR-122. "All" refers to the top 50 miRNA families, inclusive of miR-122. Two-sided K-S test p-value between miR-122 and all targets shown. **(C)** The mean log₂ fold change (\pm ranges) in CLIP binding on miR-122 3'UTR targets versus all targets during HCV infection broken down by seed type. **(D)** A CDF plot during HCV infection as in (B) but measuring target mRNA expression via RNA-Seq, from duplicate experiments at 72hrs post-infection. Targets with more than one miRNA binding site were collapsed such that no gene is represented more than once per category. **(E)** The mean log₂ fold change (\pm ranges) in mRNA expression of miR-122 3'UTR targets versus all targets during HCV infection broken down by seed type.

However the greatest change was observed with 7-8mers likely reflecting the higher intrinsic false-positivity associated with 6mers (Figure 4.10C). Through RNA-seq measurements we observed functional de-repression of CLIP derived miR-122 3'UTR targets after virus infection such that greater RNA abundance was evident when compared to all miRNA targets (Figure 4.10D). Likewise, significant expression changes for all miR-122 target seed types were observed in the reciprocal direction as CLIP (Figure 4.10E). Taken together, these results establish a correlative role for HCV replication in functionally reducing miR-122 activity, both at the level of decreased AGO binding and subsequently increased mRNA levels.

While emphasis of miRNA activity largely centers on 3'UTRs, we can peer objectively into CLIP data to look for any effects on other genic elements: 5'UTRs, coding exons and introns. If miR-122 inhibition by HCV is truly general, we would expect to see AGO CLIP changes similar to 3'UTRs on other genic regions, where these changes should likely not have much of an effect on mRNA levels. miRNA activity is well known to be weaker if AGO binding occurs on coding regions, presumably due to competition with ribosomes (Bartel, 2009). Moreover, there is very little general evidence that miRNA binding to 5'UTRs or introns influences mRNA levels, however there are exceptions.

Unlike HCV, AGO rarely binds cellular 5'UTR elements, of those that appear to be bound, there is no miR-122 specific difference in AGO binding due to HCV infection, nor is there a significant change in mRNA abundance (Figure

4.11A). The low number of 5'UTRs bound is likely the reason for this, as the results are quite noisy. The story is different with miR-122 targeted coding exons,

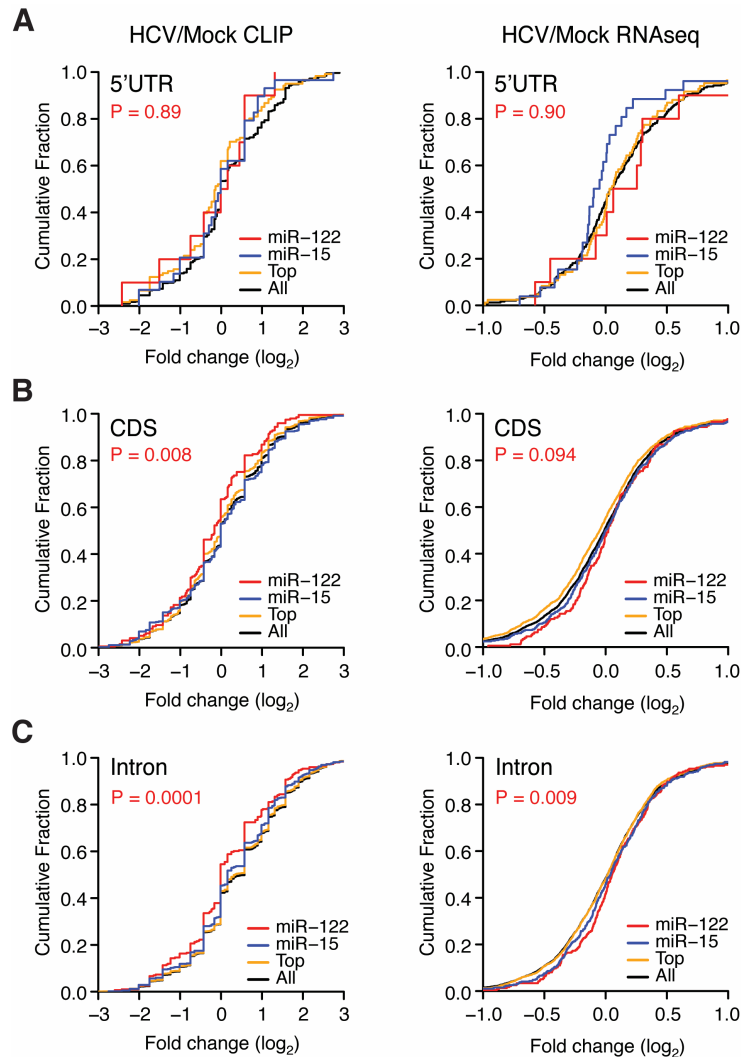


Figure 4.11. HCV infection effects on non-3'UTR miR-122 targets. CDF plots of the \log_2 fold change in CLIP target binding (left) and mRNA expression (right) between infected and uninfected cells for 5'UTR (**A**), CDS (**B**), and intronic (**C**) targets containing indicated 7-8mer seeds by family, from three experiments. "Top" refers to the top 10 miRNA families, exclusive of miR-122. "All" refers to the top 50 miRNA families, inclusive of miR-122. Two-sided K-S test p-value between miR-122 and all targets shown. For mRNA-seq, targets with more than one miRNA binding site were collapsed such that no gene is represented more than once per category.

there is a slight reduction of AGO binding on exons upon HCV infection, however this difference does not result in a significant increase in mRNA (Figure 4.11B). This is consistent with previous observations on exonic miRNA targeting as expected.

miR-122 binding to introns presents an interesting observation as there is a slight decrease in intronic binding upon HCV infection, and also a very slight increase in mRNA expression (Figure 4.11D). Whether this points to a role in miRNA regulation of transcript levels by binding exons is highly speculative. Recall that the population sizes between 3'UTRs, exons, introns and 5'UTRs are uneven at any biologic complexity (Figure 4.4). Thus any significance calculations are likely to be spuriously magnified with a high enough sample size as is very common in genomics and other data science research. This problem, known as the "p-value problem", makes it very difficult to interpret if the slight mRNA change due to miR-122 intronic binding is meaningful (Lin et al., 2013). It's perhaps more intuitive to consider that the effect of mRNA abundance is very marginal on introns and coding exons (in Figure 4.11, the red curves are only slightly different than the black curves), while for 3'UTRs there is a drastic difference (Figure 4.10). The most conservative interpretation is that AGO binding changes on miR-122 targets due to HCV can be observed throughout genic elements targeted by miR-122, but they are functionally most pronounced on 3'UTRs.

As, mentioned in chapter 3, numerous miRNAs beyond miR-122 have been proposed to bind HCV RNA, even in Huh-7 cell derivatives (reviewed in (Gupta et al., 2014)). Of these miRNA families, only the let-7 and miR-196 families were identified as AGO associated in polyG-CLIP. Conceivably, binding of these miRNAs to HCV could reduce their functional activity for host targets in a similar stoichiometric sponge manner as miR-122. We did not observe this Figure 4.12. AGO binding and mRNA expression of let-7 and miR-196 targets was not altered by HCV infection. This agrees with the CLIP map on HCV, as neither of these miRNA seeds were observed in AGO peaks on the viral genome. Overall, this points to miR-122 as being very unique among miRNAs in that its criticality for HCV replication results in specific and observable changes to host miR-122 targets.

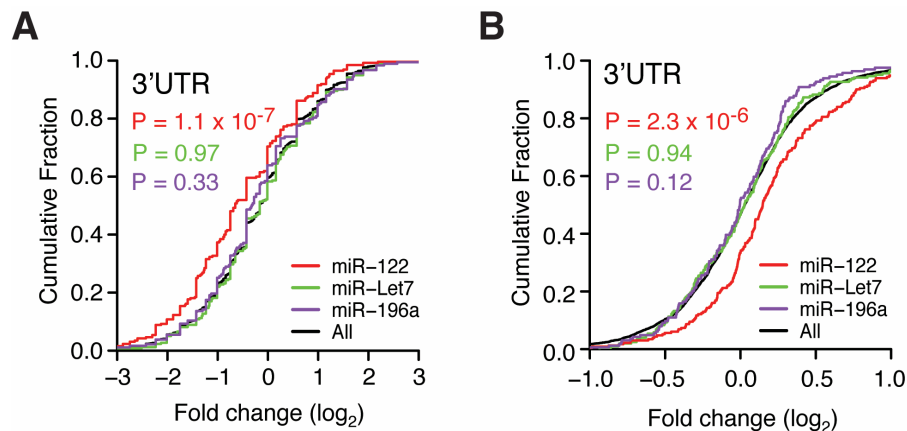


Figure 4.12. HCV infection effects on miRNA targets for miRNAs that putatively bind HCV RNA. CDF plots of the log₂ fold change in CLIP target binding (**A**) and mRNA expression (**B**) between infected and uninfected cells highlighting let-7 and miR-196 family 3'UTR targets. Two-sided K-S test p-value between let-7 (green), miR-196 (purple), or miR-122 (red) and all targets shown.

As CLIP permits the biochemical identification of miRNA targets, it is worthwhile to consider how it compares to bioinformatic prediction. Compared to Targetscan6.2 (TS) (Lewis et al., 2005), we found that CLIP largely complemented and expanded upon predicted miR-122 targets (Figure 4.13). 3'UTR targets identified via CLIP and predicted by TS exhibited the greatest change in Ago binding (Figure 4.13A) and mRNA de-repression (Figure 4.13B-C) compared to expressed targets unique to either search modality. Of the expressed 731 miR-122 CLIP targets of all seed types identified via CLIP, 48% and 9% overlapped with non-conserved and conserved TS predictions, respectively (Figure 4.13D). Focusing on a more stringent set of 7mer and 8mer seeds for CLIP data yielded even greater overlap, such that only 5% of CLIP derived targets were not represented in either TS conservation category (Figure 4.13E). These results highlight a broad convergence between CLIP and bioinformatic prediction to outline a set of miR-122 targets specifically derepressed upon virus infection. Given the bias of bioinformatic prediction to favor 3'UTRs and conserved sites, in addition to being unable factor in the expressed transcriptome, the empirical nature of CLIP methods make it uniquely poised to given an unbiased account of AGO binding, as the above results suggest.

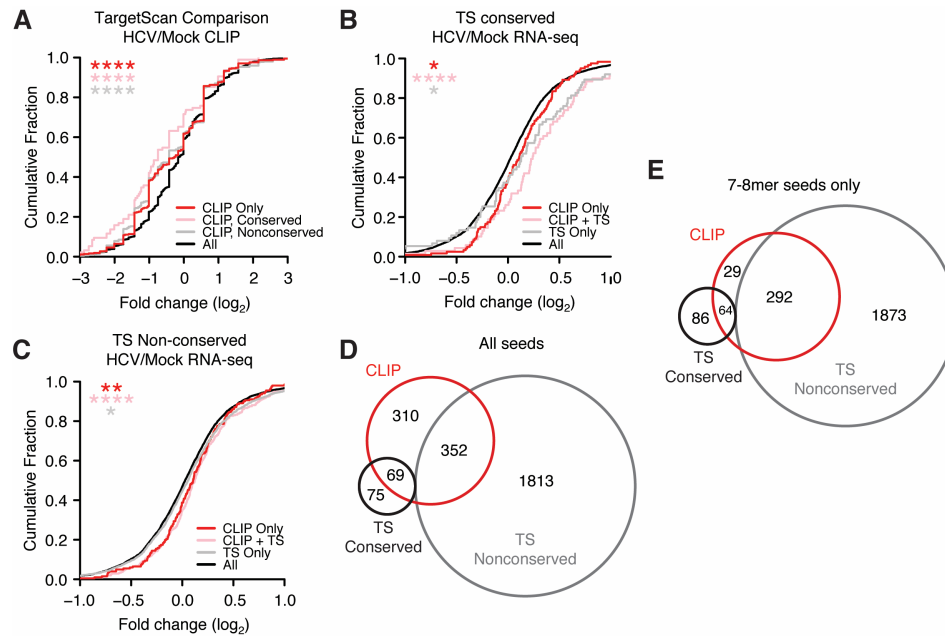


Figure 4.13. CLIP comparison to the miRNA prediction tool Targetscan6.2. **(A)** CDF plot of Ago binding comparing 3'UTR miR-122 targets unique to CLIP ("CLIP Only"), CLIP targets predicted to be conserved ("CLIP, Conserved") and non-conserved ("CLIP, Nonconserved") according to Targetscan (TS) bioinformatic prediction. **(B)** CDF plot of RNAseq data comparing TS conserved miR-122 targets confirmed in CLIP data ("CLIP + TS"), expressed miR-122 targets without CLIP evidence ("TS Only"), or CLIP targets lacking TS prediction ("CLIP only"). Conservation for CLIP only targets was defined by an averaged PhyloP score across the miR-122 seed greater than 0.5. **(C)** CDF plot of RNAseq data as in (B) comparing non-conserved TS targets. Conservation for CLIP only targets was defined as in (B) **(D-E)** Proportional Venn diagram showing the overlap of conserved and non-conserved TS predictions to CLIP derived miR-122 targets of all seed types (D) or 7mer and 8mer seeds only (E). Asterisks: **** $P < 0.0001$, *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, two-sided K-S test.

Thus far, our results suggest that HCV functionally reduces miR-122 levels. If so, then CLIP results obtained from artificially inhibiting or deleting miR-122 in the absence of virus should be similar. We decided to corroborate our observations with HCV by performing CLIP after pharmacologic inhibition of miR-122 and in Δ miR-122 Huh-7.5 cells. The reduced Ago binding on miR-122 3'UTR

targets during HCV infection was similar to 30nM LNA122 or miravirsin treatment (Figure 4.14A-B) and to Δ miR-122 cells compared to unedited controls (Figure 4.14C).

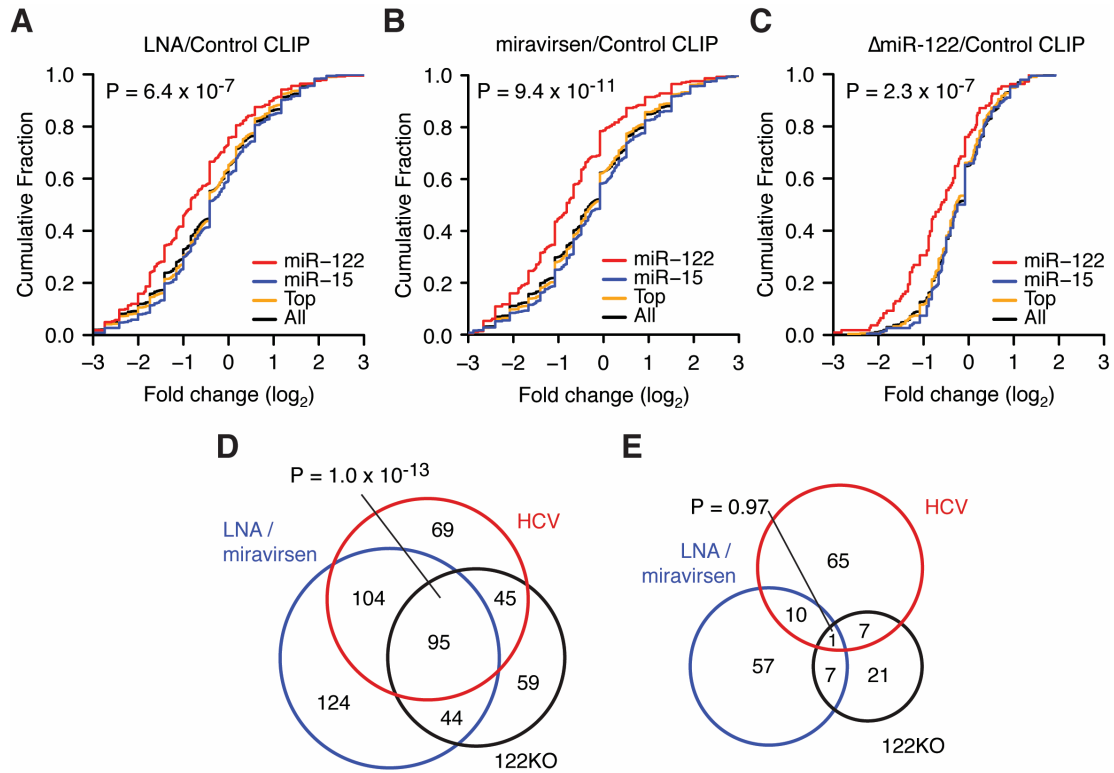


Figure 4.14. MiR-122 target de-repression in LNA, miravirsin and Δ miR-122 cells. Cumulative density function (CDF) of the log₂ fold change in CLIP binding between treatment over control cells with LNA122 (**A**) or miravirsin (**B**) at 30nM or genetic deletion of miR-122 (Δ miR-122) (**C**), for all 3'UTR clusters containing indicated 7-8mer seeds by family, from triplicate experiments. "Top" refers to the top 10 miRNA families, exclusive of miR-122. "All" refers to the top 50 miRNA families, inclusive of miR-122. Two-sided K-S test p-value between miR-122 and all targets shown. (**D**) Proportional Venn diagram showing the overlap of miR-122 targets with reduced (<0) CLIP binding across Δ miR-122, LNA or miravirsin treatment, and HCV infection conditions. Hypergeometric p-value of overlap shown. (**E**) Proportional Venn diagram showing the overlap of miR-122 targets with greater (>0) CLIP binding across Δ miR-122, LNA or miravirsin treatment, and HCV infection conditions. Hypergeometric p-value of overlap shown.

Interrogating the list of 3'UTR targets exhibiting reduced Ago binding across these three conditions revealed highly significant overlap (Figure 4.14D) suggesting that the effect of HCV replication on lowering functional miR-122 levels is as general as antagonizing miR-122. Interestingly, we can also observe the noise in this system, by looking the overlap of miR-122 targets that exhibited enhanced AGO binding in the presence of virus, LNA or miR-122 deletion. Indeed, there is no significant overlap (Figure 4.14E).

Overall, while the above results compellingly suggest that HCV functionally de-represses miR-122 targets, it is still only correlative. Establishing a causative role was in truth the main motivation to construct the miR-15 dependent virus: by swapping the miRNA tropism of HCV, we should expect de-repression of miR-122 targets to be relieved and redirected to miR-15 targets. Parsing RNAseq from m15 versus mock infected cells, we indeed observed that miR-122 targets were no longer de-repressed, and that miR-15 family targets were (Figure 4.15A). These results highlight the causal nature of an HCV induced miRNA sponge as both functional and somewhat modular. We note that the m15 virus sponge effect was generally weaker than for the WT virus, likely due to the lower replication level observed and possibly to binding of the miR-15 family member, miR-16, which shares the seed site but may not be able to engage the m15 genome due to lack of auxiliary pairing (see previous chapter). On a transcript level, we could also observe via CLIP that binding to miR-15 targets was reduced in m15 virus but not WT virus infected cells; likewise, miR-122

targets with reduced binding in WT virus infection showed no reduction in m15 infected cells (Figure 4.15B-C). Thus at the level of AGO binding and mRNA expression, we can observe HCV RNA as a causal miRNA sponge in action.

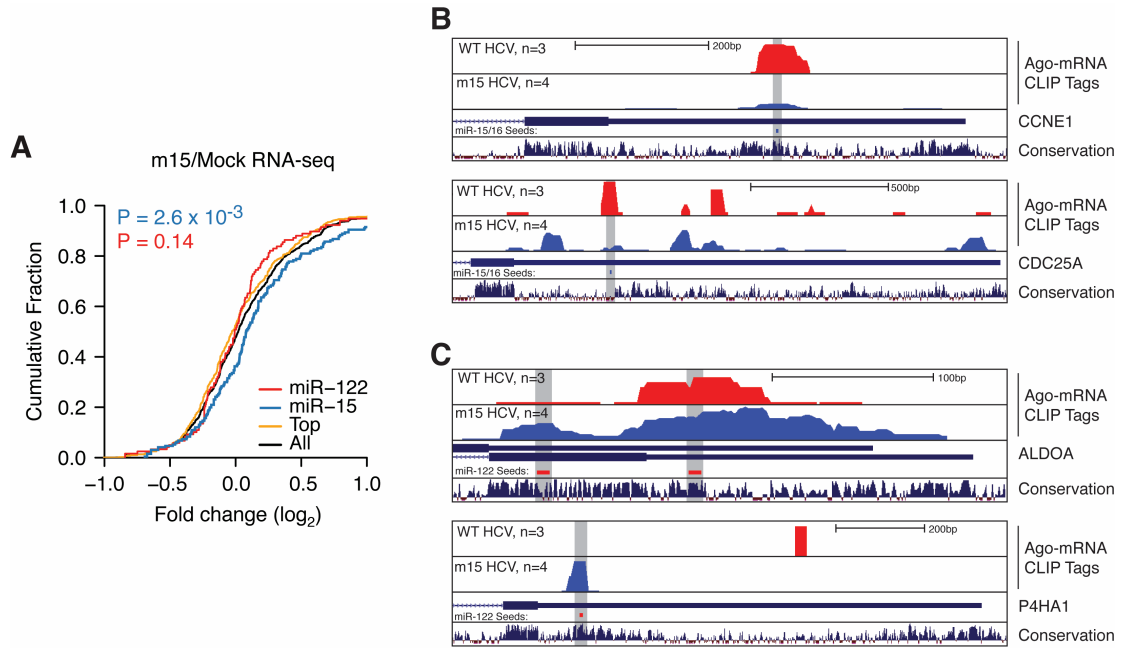


Figure 4.15. Exchanging HCV miRNA tropism redirects functional miRNA sequestration. **(A)** CDF plot of the \log_2 fold change in mRNA expression between HCV m15 infected and uninfected cells for all 3'UTR clusters containing indicated 7-8mer seeds by family, from duplicate experiments at 96hrs post infection. "Top" refers to the top 10 miRNA families, exclusive of miR-122 and miR-15. "All" refers to the top 50 miRNA families, inclusive of miR-122 and miR-15. Two-sided K-S test P-value comparing miR-15 (blue) or miR-122 (red) clusters to "All" is shown. **(B)** Genome browser tracks of two known miR-15 family targets, CCNE1 and CDC25A. AGO binding in the presence of WT HCV (3 libraries) or m15 HCV (4 libraries) in the upper red and blue tracks, respectively. Reads are plotted as normalized densities to enable cross track comparison. miR-15/16 family seeds noted by gray bar. PhyloP conservation track shown at the bottom. **(C)** Genome browser tracks as in (B) of two known miR-122 family targets, ALDOA and P4HA1. MiR-122 family seeds noted by gray bar.

Validation of HCV induced miR-122 sequestration in bulk and single cells

Based on the results so far, it's arguable that HCV RNA can be said to "validate" miR-122 targets, given the effect of WT virus on host transcripts containing miR-122 seeds. The same can apply to the m15 virus and miR-15 family targets. It bears stressing that all of the above HCV work was done by observing the effects of virus on the endogenous transcriptome, without manipulating miRNA levels. Thus, any further attempts to validate the sponge effect are likely to yield fruit and may provide additional mechanistic insight on how HCV usage of miR-122 impacts host targets.

As a first attempt, we decided to observe in practical terms, what mRNA de-repression due to sponging of a miRNA looks like via qPCR to validate changes observed via RNAseq, and luciferase reporter assays, to measure the magnitude and seed dependence of HCV induced miR-122 sponging. In general, qPCR of 3'UTR miR-122 targets were increased between 1.1 and 1.8 fold in the presence of virus, in line with the median RNA-seq measurement of a 50% change, and in line with estimates of miRNA function (Mukherji et al., 2011) (Figure 4.16). As expected, only 3'UTR targets appeared de-repressed (8 of 12) and not coding exon targets (0 of 3) Figure 4.16. These results provide an independent validation of CLIP and RNAseq measurements.

Next, we used luciferase reporters of miRNA activity to validate the HCV miR-122 sponge on individual 3'UTRs. Using miRNA mimics or LNAs as positive controls for repression and de-repression respectively, we observed that HCV

infection broadly resembled LNA inhibition with reporters bearing miR-122 seeds, or with cloned full length miR-122 target 3'UTRs (Figure 4.17). As expected, this effect was specific to miR-122 containing seeds, as no de-repression by HCV was observed on a miR-17 reporter.

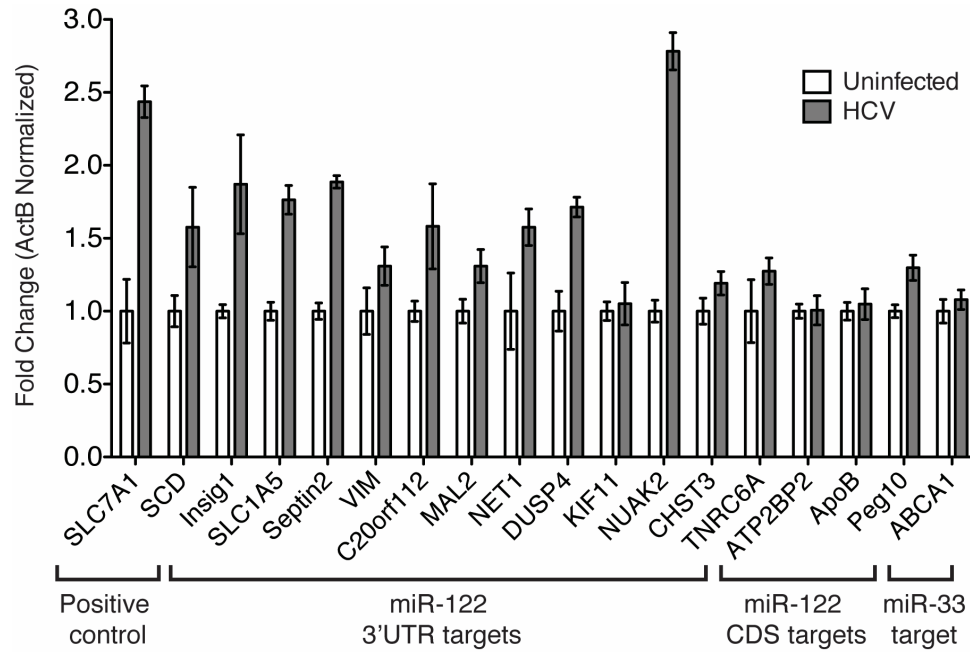


Figure 4.16. QPCR of AGO-CLIP derived miR-122 targets after HCV infection, from triplicate experiments. Two previously known miR-122 targets (SLC7A1 and SCD) served as positive controls. 3'UTR and CDS miR-122 targets tested are indicated. A miR-33 target (ABCA1) served as a negative control. Error bars, \pm SD.

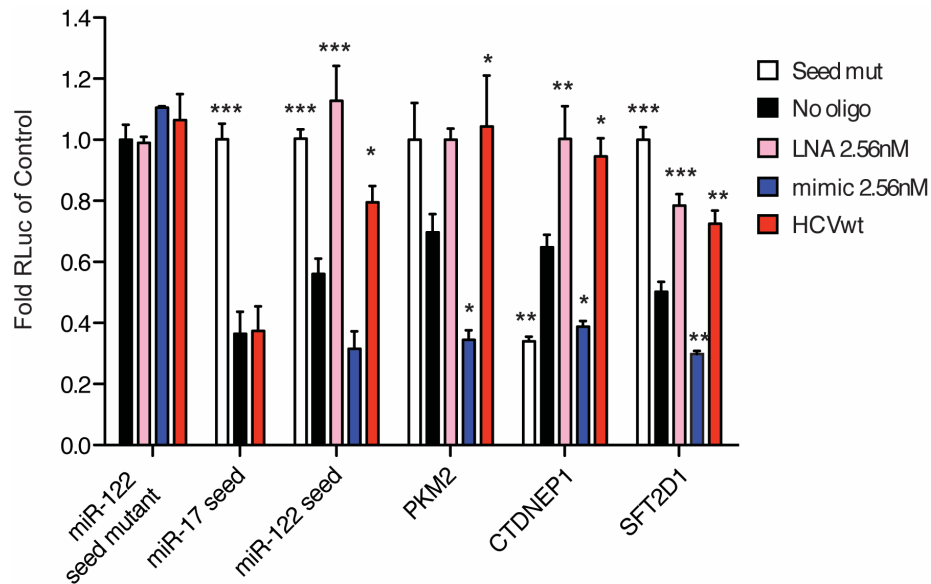


Figure 4.17. Validation of HCV induced de-repression of miR-122 targets with luciferase reporters. Luciferase reporter measurements for synthetic miR-122, miR-17, or cellular 3'UTR target constructs. Data were normalized to "no oligo" p3,4 mutant conditions. Significance testing was performed relative to endogenous "no oligo" repression for each tested construct. Asterisks: *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, ANOVA with bonferroni correction.

Perhaps an under appreciated but salient point of the results presented thus far is that they stem from measurements on populations of cells. As cellular mRNA and HCV RNA expression levels are known to vary widely between individual cells (Kandathil et al., 2013; Sheahan et al., 2014), we sought to achieve a more thorough understanding of the HCV miRNA sponge on host miRNA targets at a quantitative single-cell level. Previous work by Mukherji, van Oudenaarden, Sharp, and colleagues demonstrated that miRNAs generate thresholds of gene expression such that miRNA repression can be highest on low abundance targets, and can be virtually non-existent on high abundance targets

(Mukherji et al., 2011). Furthermore, these thresholds can be altered upon manipulating miRNA levels. Perhaps the most intriguing feature of this work is that it provides a testable mathematical model of miRNA function across a range of mRNA expression patterns. How might the HCV:miR-122 sponge work in this context?

To test this, we adapted the strategy used by Mukherji et al. to construct two-color tet-inducible fluorescent reporters of miRNA activity amenable to flow cytometry (Figure 4.18) (Mukherji et al., 2011). In this assay, a bidirectional Tet promoter drives expression of blue and red fluorescent proteins (TagBFP and TagRFP). Each fluorescent protein is tagged with a nuclear localization sequence (NLS) to aid in flow cytometric analysis. The 3'UTR of TagRFP is engineered to contain *N* binding sites for miR-122, or full 3'UTRs of selected miR-122 targets, while TagBFP serves as an expression control. When expressed in Huh-7.5 TetON cells (see methods), a TagRFP with no miR-122 sites will be expressed at equal per-cell levels as TagBFP. However, upon introducing miR-122 sites to the 3'UTR of TagRFP, its expression will be repressed by endogenous miR-122 (Figure 4.18A). In this manner, manipulation of miRNA levels with mimic, LNA or virus can be used to determine miRNA activity on TagRFP across a range of TagBFP expression values.

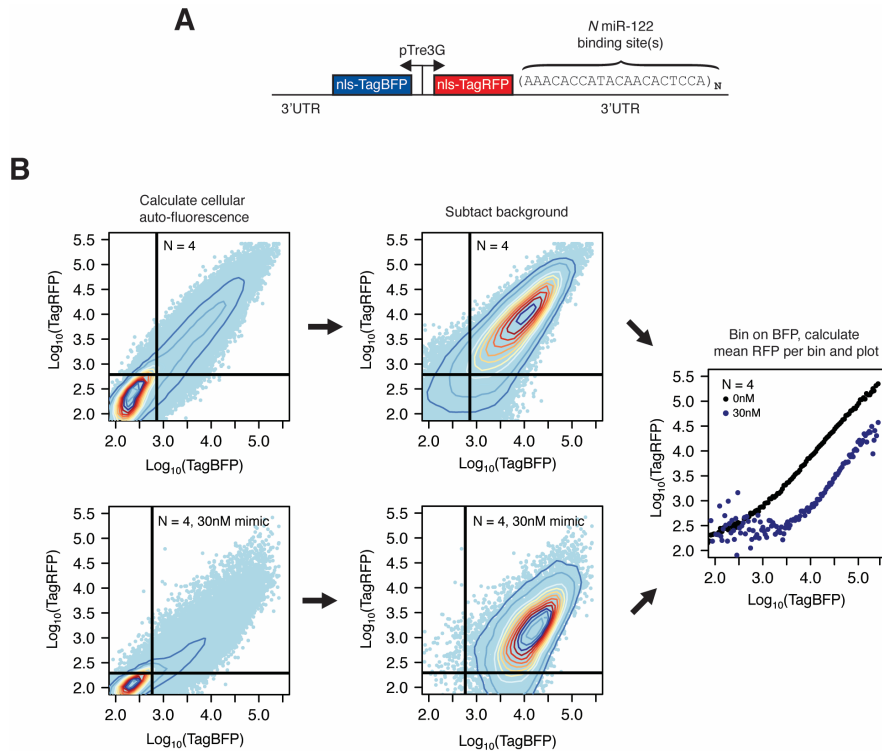


Figure 4.18. A single-cell assay for miRNA activity. **(A)** Two-color fluorescent reporter containing a bidirectional Tet promoter that drives expression of blue and red fluorescent proteins (TagBFP and TagRFP). Each fluorescent protein is tagged with a nuclear localization sequence (NLS) to aid in flow cytometric analysis. The 3'UTR of TagRFP is engineered to contain N binding sites for miR-122, or full 3'UTRs of miR-122 targets. **(B)** Sample processing starts by calculating the mean cellular BFP and RFP auto-fluorescence from untransfected cells (mean denoted by lines). The auto-fluorescence of each channel plus two standard deviations is subtracted from every point. The resulting distribution is binned across the BFP signal, and the RFP mean calculated for each bin and plotted. In this example, a construct with four miR-122 sites was transfected alone or in the presence of 30nM miR-122 mimic. Contour lines for each FACS plot denote 10% increments of the total density.

Testing reporters with $N = 1$ and 6 miR-122 binding sites in the presence of miR-122 mimic, we observed dose dependent miR-122 repression that increased with N , as expected, whereas adding LNA122 decreased repression (Figure 4.19A-D). HCV infection, using a Ypet expressing clone2 reporter virus, in both contexts resembled LNA inhibition where de-repression was notably more

pronounced in cells expressing low amounts of reporter, extending previously reported miRNA thresholding effects for miR-20 to miR-122 (Mukherji et al., 2011)(Figure 4.19A and C). Importantly, no such changes were observed for a reporter with a p3,4 miR-122 seed ("N1m") (Figure 4.19E-F). Additionally, we tested a reporter with a perfectly complementary miR-122 site, thus making the miRNA behave as a siRNA. As highlighted in the introduction, this distinction is useful to separate the modes of action of small RNAs on AGO: miRNAs appear to work stoichiometrically to suppress gene expression, and as a consequence, their repressive activity can be diluted out by expressing more target mRNA. This is why miRNA repression appears to fail at high target expression levels. In contrast, siRNAs work catalytically as AGO:siRNA complexes can move on to a new target after mRNA cleavage. By using miR-122 as an siRNA, we expect that no such "threshold" should be observed as it would for miRNAs. Indeed, miR-122 as siRNA reporter exhibited no thresholding such that mimic repression, or LNA and HCV de-repression was observed at all expression levels (Figure 4.19G-H). Taken together, these data suggest that HCV infection modulates functional miR-122 levels to relieve endogenous repression on host targets in a stoichiometric manner, and is governed by target expression level and the number of miRNA binding sites.

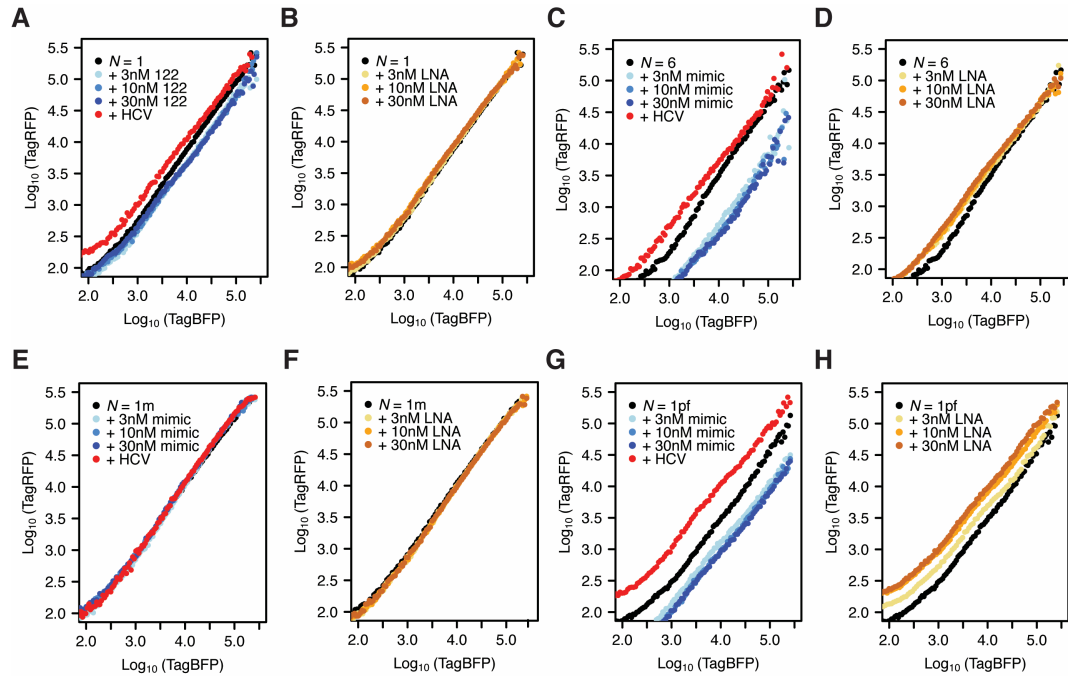


Figure 4.19. Single cell fluorescent reporters of miRNA activity measurements after modulating miR-122 levels. **(A)** Log-log transfer functions for $N=1$ miR-122 site in the presence of increasing concentrations of miR-122 mimic or HCV infection. **(B)** Transfer functions for $N = 1$ miR-122 site in the presence of increasing concentrations of LNA inhibitors to miR-122. **(C-H)** Transfer functions as in (A) and (B) for $N=6$ **(C-D)**, a mutated miR-122 seed ("1m") **(E-F)**, and $N = 1$ perfectly complementary ("1pf") miR-122 sites **(G-H)**.

We extended this analysis to measure actual cellular miR-122 target 3'UTRs using this system. We observed modest de-repression upon HCV infection for previously known targets with one miR-122 site, AldoA, PKM2, and P4HA1, but not for CS 3'UTRs (Figure 4.20A-D). We also tested novel CLIP-identified targets CTDNEP1, SFT2D1, MASP1 and MAL2 and obtained similar results, with all four tested being reduced upon miR-122 mimic addition and all except MAL2 de-repressed upon adding virus (Figure 4.20A-D).

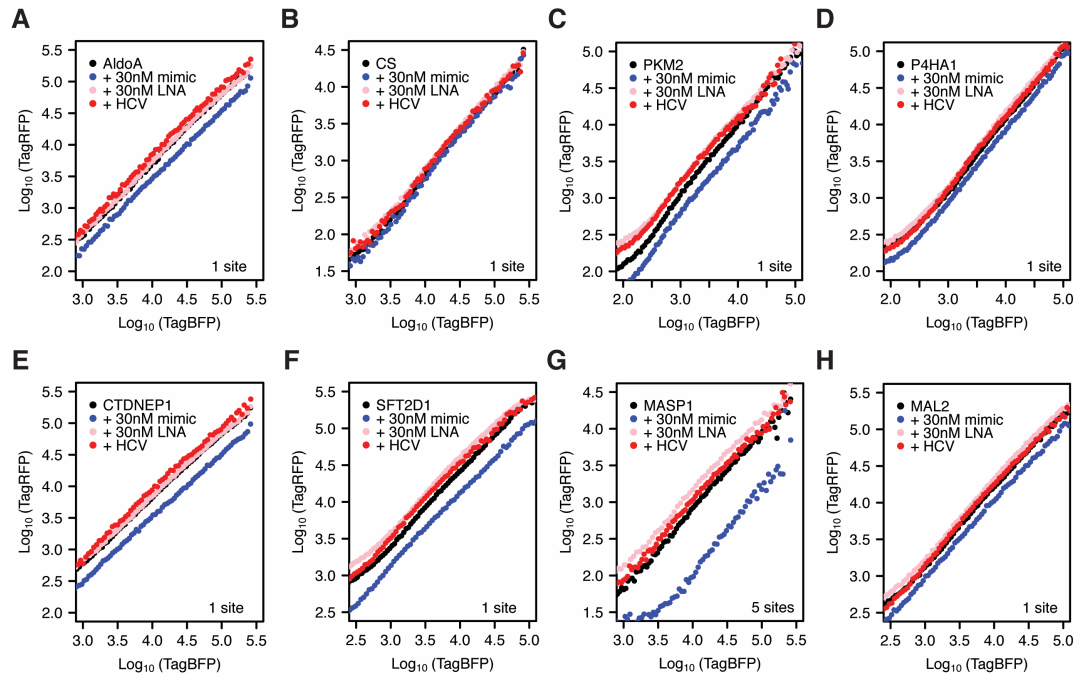


Figure 4.20. Fluorescent reporter measurements for select cellular 3'UTRs. Log-log transfer functions of full-length cellular 3'UTRs in the presence of miR-122 mimic, LNA inhibitor, or HCV infection. The effective number of miR-122 sites (N) is indicated for each construct. **(A-D)**: previously reported miR-122 targets confirmed via CLIP. **(E-H)**: Novel CLIP derived miR-122 target 3'UTRs.

In broad terms, HCV infection resembled LNA inhibition of miR-122 for these reporter measurements. One caveat that must be addressed is the apparent discrepancy between miR-122 levels in liver versus those in Huh-7.5 cells. As miR-122 levels in Huh7 derived cells are estimated to be 10-fold lower than primary adult liver tissue (Chang et al., 2004), we next explored the HCV sponge effect in the presence of excess miR-122. Exogenous miR-122 addition increased intracellular miR-122 in a dose dependent manner by up to 10-fold in Huh-7.5 cells, within the range of miR-122 levels measured from patient liver biopsies (Figure 4.21A-B). As no changes in HCV RNA levels were observed

upon adding miR-122, the resulting miR-122:HCV ratio went from ~15:1 fold at the lowest, to over 100:1 with 30nM of miR-122 mimic added (Figure 4.21C). Testing $N = 1$ or AldoA 3'UTR reporter constructs in this *in vivo*-like context, we observed that HCV infection was able to relieve 30nM of mimic repression to untreated levels for low but not high abundance targets (Figure 4.22A-B). The ability for HCV to rescue excess miR-122 repression was not as pronounced for the $N = 6$ construct (Figure 4.22C) whereas a reporter containing a perfectly complementary miR-122 site was particularly sensitive to rescue by HCV replication (Figure 4.22D). Similar, dose-dependent results were obtained under 0.3 or 3nM mimic treatment for all constructs (Figure 4.23). Taken together, these results suggest that miR-122 sponging by HCV RNA can exist in more physiologic miR-122 concentration settings.

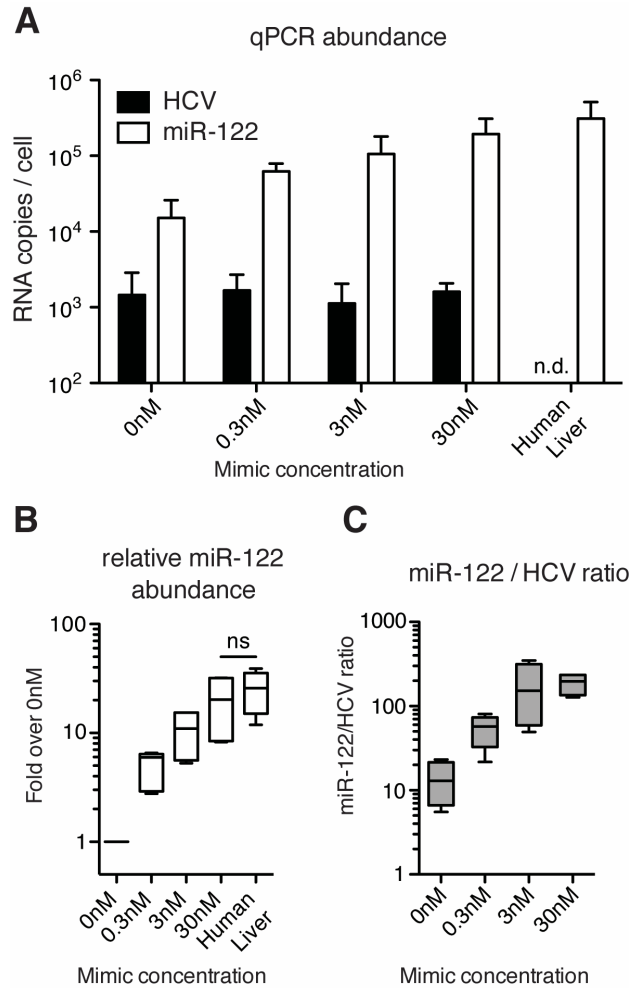


Figure 4.21. Boosting miR-122 levels in Huh-7.5 cells to mimic human liver. **(A)** qPCR measurements of miR-122 and HCV RNA copy numbers per cell in infected Huh-7.5 cells after addition of increasing amount of miR-122 mimic. miR-122 copy number from human liver biopsy specimens is also shown. n.d.: not done. **(B)** The relative miR-122 abundance over mock Huh-7.5 cells, from data in (A). Significance testing was performed via Student's two-tailed t-test. **(C)** The miR-122:HCV ratio from data in (A).

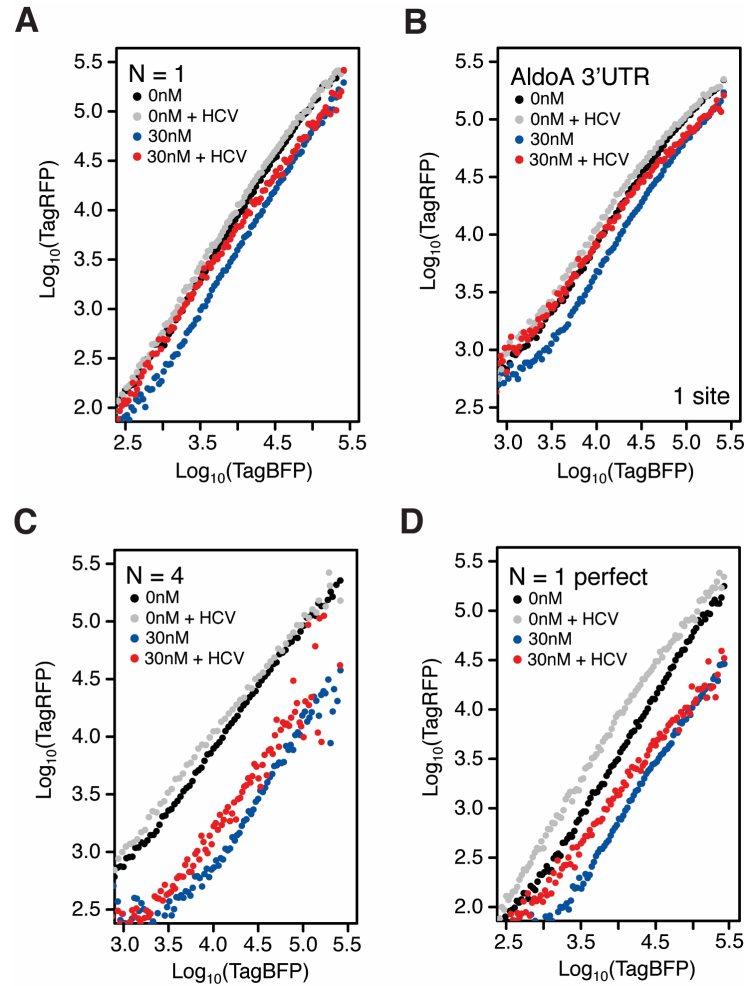


Figure 4.22. Validation of miR-122 sponging by HCV under *in-vivo*-like levels of miR-122. Log-log transfer functions for $N = 1$ (**A**), ALDOA 3'UTR (**B**), $N = 4$ (**C**) or one perfectly complementary (**D**) miR-122 site single cell fluorescent reporter constructs in the presence or absence of 30nM miRNA mimic and/or HCV infection.

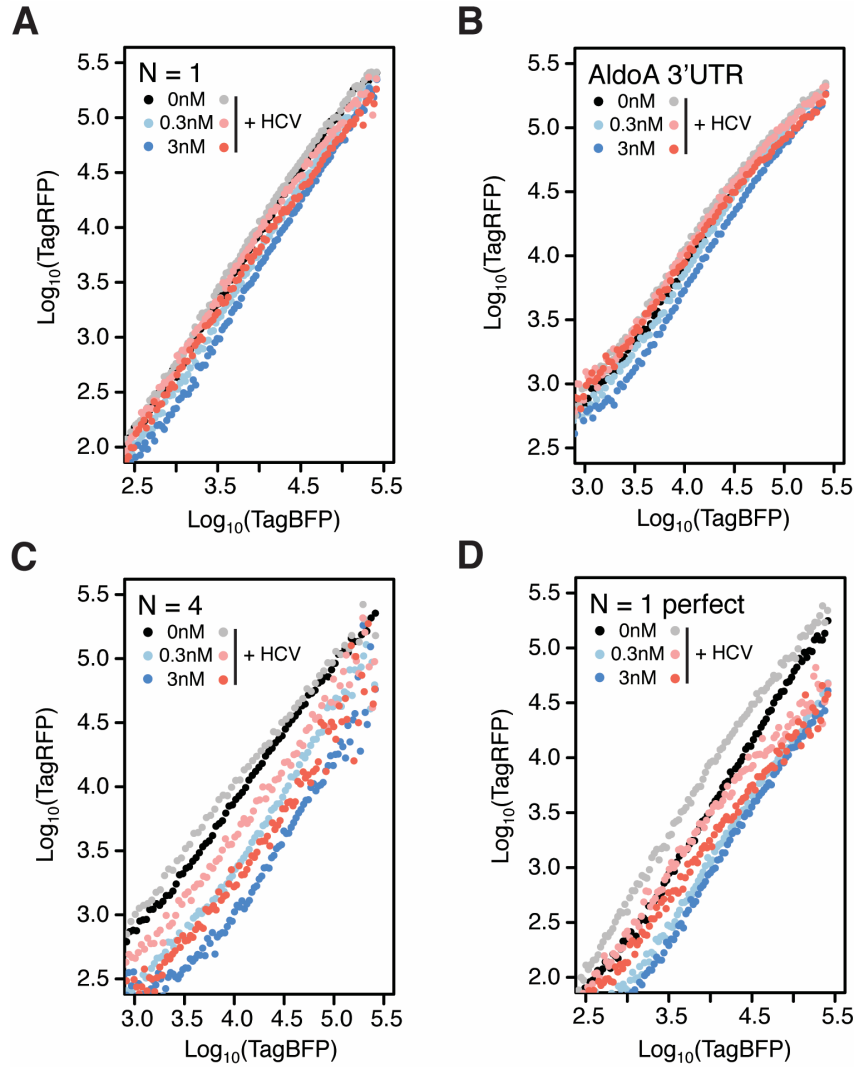


Figure 4.23. Validation of miR-122 sponging by HCV in 0.3nM or 3nM of miR-122. Log-log transfer functions for $N = 1$ (A), ALDOA 3'UTR (B), $N = 4$ (C) or one perfectly complementary (D) miR-122 site single cell fluorescent reporter constructs in the presence or absence of 0.3nM or 3nM miRNA mimic and/or HCV infection.

A quantitative model of miR-122 sponging by HCV RNA

Based on these dose-dependent mimic and LNA reporter system measurements, we expanded the miRNA model of gene regulation presented by Mukherji et al. to incorporate a competing self-replicating viral target (Figure 4.24A). Here, if HCV RNA is present at sufficiently high numbers or has relatively high binding strengths compared to other miR-122 targets, it acts to reduce the available miR-122 pool, and de-represses miR-122 targets (r , measured as TagRFP fluorescence) relative to non-targets (r_o , measured as TagBFP fluorescence) (Figure 4.24B). We developed a quantitative model for HCV induced reduction of the miR-122 pool in this scenario (see methods). This model describes the concentrations of a target mRNA species (r), an HCV mRNA species (h), and binding of miRNA (m) to form complexes with target or HCV mRNA species, respectively (r^* , h^*). Production of a target mRNA comes from transcription of a gene at a rate (k_r) with a corresponding degradation rate (γ_r). We assume the miRNA-mRNA complex (r^*) does not undergo translation and degrades at a rate allowing for recycling of the miRNA species into the pool. The total amount of miRNA is assumed to be constant and can bind to either target mRNA (r) or HCV mRNA (h). HCV RNA also decays at a particular rate (γ_h). As previously demonstrated, HCV RNA degradation is slower when bound to miR-122 (Li et al., 2013d; Shimakami et al., 2012), and we therefore assumed a model of HCV degradation primarily from its unbound form. Assuming steady state levels of HCV RNA at the time of measurements resulted in a decrease of

the model parameter θ , which governs the amount of free miRNA in the system. The number of miR-122 sites is estimated by the model parameter λ that is related to the total binding strength of miR-122 to a particular site. By tuning these parameters, we accurately fitted experimental data of endogenous miR-122 repression of reporters with increasing numbers of miR-122 sites (Figure 4.24C).

To explore the effect of HCV on the miR-122 pool, we fitted the model to experimental data with four miR-122 sites during infection (Figure 4.24D), and estimated the change to parameter θ to correspond to an approximate 50% reduction in available miR-122. A similar result was obtained for the $N = 4$ construct in the presence of HCV and 30nM miR-122 mimic (Figure 4.24E). The model estimated that the highest theoretical HCV levels reducing the miR-122 pool by 90% could de-repress mRNA targets by up to 4.5-fold for low-expressed mRNAs (Figure 4.24F). Synthetic reporter and cellular measurements agreed with model predictions for 50% reductions in miR-122 levels, where de-repression was most drastic for low expressed targets harboring multiple miR-122 sites or an siRNA-like context (Figure 4.24G-H). Taken as a whole, the quantitative model outlines several factors controlling HCV induced de-repression of host mRNA targets given steady-state levels of HCV RNA: the expression level of the target mRNA, mRNA-miR-122 binding strength, and the number of sites on the target mRNA.

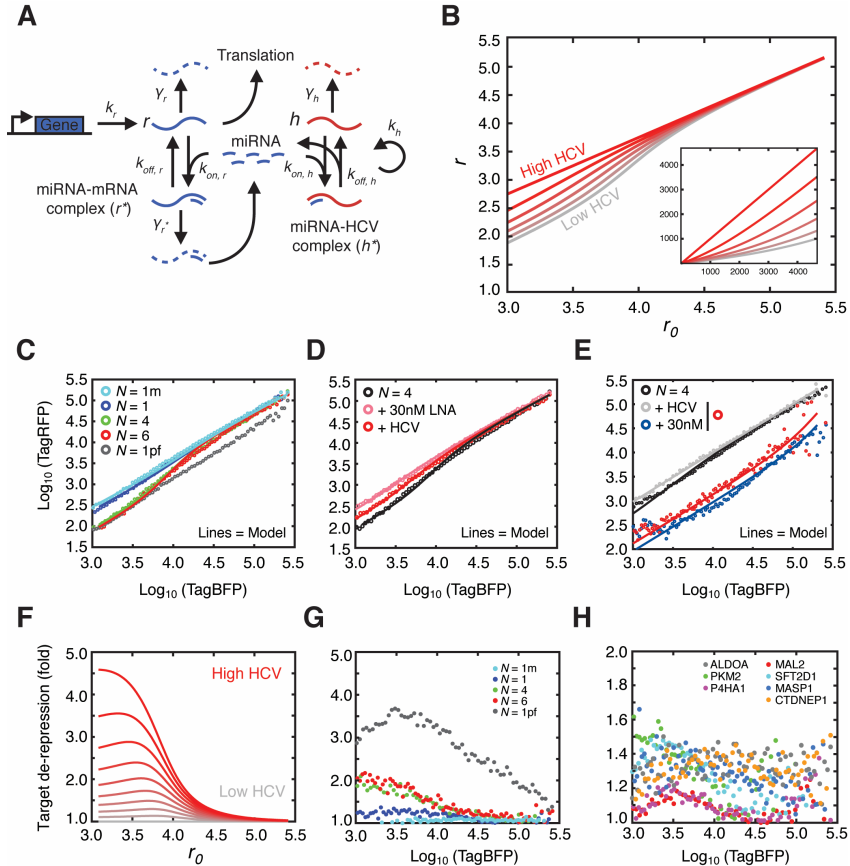


Figure 4.24. Quantitative modeling of miR-122 sequestration by HCV. **(A)** Illustration of model reactions for miR-122 dynamics, including transcription and translation of a target mRNA, binding to miR-122 and decay of mRNA species. HCV RNA can replicate, be degraded, or bind miR-122, functionally sequestering miR-122 and leading to de-repression of mRNA targets. **(B)** Increasing amounts of HCV (or a relative increase in binding strength at miR-122 sites) leads to changes in single-cell gene expression as compared to unregulated targets, with stronger effects at the low mRNA expression levels. Parameters used are fitted from data in (C). Each curve, from top to bottom, represents a reduction in the miRNA pool by 20%. Inset displays model on a linear scale. **(C)** Model fitting of the steady state approximation to experimental data while increasing the number of binding sites corresponding to changes in total binding strength. **(D)** Model fitting for the $N = 4$ case showing a 50% reduction in the miRNA pool by HCV modeled by a proportional change in the theta parameter. **(E)** Model fitting for the $N = 4$ construct under 30nM miR-122 mimic addition \pm HCV infection. **(F)** Increasing HCV:miR-122 binding strength or HCV RNA abundance results in functional de-repression of miR-122 targets. The curves (top to bottom) represent 10 percentage-points increases in the available miR-122 pool (10% to 100% availability). **(G-H)** Experimental HCV induced derepression of synthetic miR-122 binding site constructs (G) or endogenous 3' UTRs with miR-122 binding sites (H). Modeling done with Tal Danino.

As the model largely agrees with previous work, applying this framework to other more robust miRNA sponge systems (viral or otherwise) or testing other HCVs with different replication capacities than the robust clone 2 virus, will be useful to refine our understanding of miRNA function in the cell.

Chapter 4 summary and discussion

In this chapter, we've explored the converse side of AGO-CLIP data in HCV infected cells by focusing on the host. Starting from genome browser gazing, continuing with sponge hypothesis formation, global validation using RNAseq, and further characterization in bulk and single cells, our results chiefly establish that HCV RNA may act as a competitive inhibitor of miR-122 activity. Natural viral derived miRNA sponges have been described previously (Cazalla et al., 2010; Lee et al., 2013) though no examples are currently known from RNA viruses. Whether this systems-level phenomenon occurs with other, more robust RNA virus infections remains to be explored, and presents a novel frontier worthy of followup. For HCV biology, our elucidation of a miR-122 sponge raises a number of questions that revolve around two key issues, namely, whether the HCV:miR-122 sponge exists *in vivo* and whether such a miRNA sponge has a functional or pathogenic role in HCV disease biology.

On the likelihood of an HCV:miR-122 sponge in liver

The idea that HCV RNA may act as a competitive inhibitor of miR-122 is closely aligned with proposed roles for competing endogenous RNAs (ceRNAs)

(Salmena et al., 2011). While both HCV RNA and ceRNAs share the theme of de-repressing a miRNA regulated network by increasing the pool of available targets through RNA expression, they importantly differ in their mode of interaction with miRNAs. HCV genomic RNA critically requires miR-122 interaction to stabilize the viral genome and stimulate translation and replication, while most cellular transcripts are repressed upon encountering a miRNA. Moreover, unlike cellular mRNA targets, HCV genomic RNA is its own substrate for replication, and thus constitutes a direct and positive feedback loop to sequester additional miR-122. These distinguishing features suggest different parameters for HCV versus cellular ceRNA based sponge effects on a miRNA target network.

This is particularly relevant in light of recent findings showing that endogenous miR-122 repression is only relieved when ceRNAs are forcibly expressed at super-physiological levels in mouse liver (Denzler et al., 2014). In this work, the authors over-expressed the AldoA 3'UTR to constitute up to 40% of RNA in the cell; only at these extremes was de-repression of endogenous miR-122 targets observed. Notwithstanding the fact that no attempt was made to confirm that the over-expressed miR-122 target was at all regulated by endogenous miR-122, the authors generally concluded that ceRNA effects are unlikely under the normal paradigm whereby an individual ceRNA is induced, even if highly so. While perhaps over-generalized, this work raises some key

issues that are worth exploring as we consider HCV RNA as a functional miR-122 inhibitor.

The first key issue relates to cellular infection frequency and per cell HCV RNA abundance. Unlike the >90% infection frequencies in Huh-7.5 cells, the percentage of infected hepatocytes in chronically infected patients based on in situ hybridization of liver biopsy tissue, ranges from as low as 0.07% to as high as 100%, with medians in the 20-40% range (Liang et al., 2009; Pal et al., 2006). Combined with HCV genotype, dynamic replication variation within the liver, circadian miR-122 expression and regulation, and host variability in innate immune responses (Sheahan et al., 2014), a complex picture of HCV infection emerges that would largely mask observations of HCV sponge effects in bulk cell or tissue Ago-CLIP or RNAseq measurements. HCV levels per cell are estimated to range from 1 to 10^2 copies per hepatocyte in contrast to Huh-7 derivatives with in excess of 10^3 copies per cell (Kandathil et al., 2013) (Figure 4.21A). However, extrapolating from liver biopsies or back-calculations of the number of cells required to account for HCV levels in serum are all confounded such that it is very unclear what the HCV RNA level truly is in a "highly infected" human hepatocyte. So, on this point, we can only ask a speculative question: assuming a distribution centered on a few hundred RNA copies per cell, how likely are these viral RNAs capable of inducing a sponge effect?

This leads to the second main concern: miR-122 abundance. Our bulk cell measurements estimate ten-fold higher miR-122 levels in primary liver tissue

versus hepatoma cell line derivatives (10^5 vs 10^4 estimate ranges per cell, Figure 4.21A) and are in line with previous estimates (Chang et al., 2004; Denzler et al., 2014). Taken at face value, this suggests that the HCV:miR-122 sponge might be a feature of replication unique to Huh-7.5 cells and likely plays little if any role in the liver, where viral RNAs are drastically outnumbered by miR-122.

However, there are a number of assumptions with this interpretation that we must consider. Chief among them is the assumption that miR-122 binding to HCV RNA is functionally and mechanistically the same as binding to a cellular 3'UTR. Considering the basic premise of miRNA function, this is demonstrably false: while miRNAs function to repress mRNAs, no target transcript "needs" to encounter miRNA for its function to undergo translation. HCV RNA on the other hand, critically requires a miRNA for its function to make viral protein (by stabilizing the viral RNA and/or directly aiding in translation). In the context of these opposing outcomes of miRNA binding, ceRNAs are argued to behave as ordinary mRNA targets (Salmena et al., 2011), as they are neither stabilized by miRNA binding nor is there as direct a mechanism of propagating ceRNA activity as cytoplasmic viral RNA replication. Our results that HCV can exert sponge effects in the presence of excess miR-122 would support the view that HCV RNA as a miRNA sponge does not abide by simple miRNA stoichiometry assumptions made for ceRNAs or mRNA targets (Figure 4.22).

But why might this be? This idea of "needing" to encounter a miRNA might imply a more active mechanism of miR-122:HCV RNA binding, that goes beyond

passive miR-122 abundance. Conceivably, the virus could have evolved a specific means to recruit miR-122 as an essential host factor to sites of viral replication or assembly. Indeed, mounting evidence supports such a hypothesis. HCV replication is known to disperse and prevent formation of processing bodies (P-bodies), the sites of miRNA silencing (Ariumi et al., 2011; Berezhna et al., 2011; Pager et al., 2013). Additionally, AGO2, miR-122 and Dicer, but not RISC components GW182 and DCP2, selectively partition to sites of HCV replication / assembly at lipid droplets (Ariumi et al., 2011; Berezhna et al., 2011). These cell biological findings highlight the possibility that HCV RNA coercion of miR-122 might engender a sponge effect, even when miR-122 is highly abundant.

As mentioned in the introduction but refined here in full, the mouse liver provides additional evidence that total miR-122 levels aren't the whole story. Despite invariantly high levels of mature miR-122 in the liver, miR-122 transcription is under circadian control and a number of targets oscillate in a circadian manner corresponding to repression by miR-122 (Gatfield et al., 2009). This puzzling finding could be explained in at least two ways (Figure 4.25A-B). MiR-122 could function to amplify target protein oscillation by repressing basal translation at an expression trough; thus miRNA function cycles between switch-like and neutral modes (recall Figure 1.7B)(Figure 4.25A).



Figure 4.25. Potential mechanisms of miR-122 activity despite invariant mature levels. **(A)** Constant miR-122 (dark gray) may shape circadian rhythm of targets (light gray) by repressing basal levels of translation. In this case, only targets represented by the dotted area is available for translation. This context would impart higher protein amplitudes throughout the circadian cycle. **(B)** Conceivably, a newly synthesized or a chemically distinct miR-122 sub-population may exist. If this circadian expressed population was functionally distinct from bulk miR-122, target mRNAs could be subject to circadian oscillation (dotted area again depicts fraction of mRNA available for translation). **(C)** In one model of how HCV RNA might depress miR-122 targets, we can consider miR-122 bound to HCV as effectively reducing the total miR-122 pool, thus enabling more target to be translated (dotted area). **(D)** In another model, HCV RNA may hijack the more "active" newly synthesized or precursor that is perhaps a limiting miR-122 subpopulation, such that mature levels are reduced through this upstream targeting. Panels A and B modified from (Gatfield et al., 2009).

Should a sub-population of miR-122 exist that is either chemically altered and/or newly synthesized, and assuming this sub-population is functionally distinct from total levels, then it's possible that circadian expression of this sub-population influences target mRNA translational ability to oscillate (Figure 4.25B). These in fact are the main mechanisms proposed by Gatfield, et al. and are bolstered by observations of cytoplasmic modifications to miR-122, particularly adenylation and uridylation (Burns et al., 2011; Katoh et al., 2009). Relating changing miR-122 activity due to HCV RNA can be similarly considered, whereby viral RNA replication has a direct and negative impact on miR-122 levels available to repress targets (Figure 4.25C). Along the lines of newly synthesized miR-122, perhaps HCV actually targets this lower abundance and limiting precursor of mature miR-122, and in so doing reduces mature miRNA by preventing biogenesis (Figure 4.25D).

All of the above suggests additional nuanced factors and raises questions that are largely obscured from bulk miRNA abundance measurements of high miR-122 levels in liver, and the assumption that all this miRNA is functional. Are distinct AGO:miR-122 complexes recruited to HCV RNA in special compartments? Are these complexes limiting? Might coercion of newly synthesized or chemically distinct miR-122 be the mechanism of miRNA subversion? What happens to miR-122 after it engages the viral genome?

Numerous experimental paths forward spring to mind to explore these ideas, a few of which are summarized here. The miR-15 dependent virus

presents an excellent test case to see if miR-15 family members can be redirected to sites of HCV replication and assembly in place of miR-122. A detailed analysis of S1 and S2 swap mutants where one site is miR-122 and the other site miR-15 could be used to test if the 2x stoichiometry of binding sites is needed to render an effective sponge. Deletion of a cytoplasmic polymerase GLD2, that is known to stabilize miR-122 could be used to evaluate the possibility of mature but largely non-functional "old" miR-122 that is distinct from newly synthesized miR-122 (D'Ambrogio et al., 2012; Katoh et al., 2009). Combined with the development of cell compartment specific AGO-CLIP, these experiments may continue to unravel the unique interplay between HCV RNA and miR-122.

On the functional role of the HCV:miR-122 sponge

The establishment of a miR-15 dependent HCV suggests that the miR-122 sponge effect is largely dispensable for the virus in the Huh-7.5 cell context. From the previous chapter, recall that LNA-122 slightly reduced m15 virus titers and that restoring miR-122 in Δ miR-122 cells increased titers, suggesting that the miR-122 sponge may reflect a trade-off for the large, positive and direct impact of miR-122 on WT HCV replication. Conceivably, HCV replication may exert enough pressure on miR-122 levels to de-repress targets such that the cellular environment is passively altered to negatively impact viral replication. Or more actively, some targets may act as sensors for low miR-122 levels, and by extension, the health of the hepatocyte. While future work will be needed to shed

light on specific players involved in this process, our data suggest that viral replication faces a ceiling by reducing levels of an otherwise pro-viral miRNA.

Note that throughout this chapter I have avoided discussing specific targets that are de-repressed upon miR-122 sponging, even those targets extensively tested in reporter assays. This has been deliberate. Due to the hepatoma nature of Huh-7.5 cells, extreme caution is needed when attempting to extrapolate results to liver biology, no matter how compelling. The transcriptomes between these cell types, notwithstanding hepatic examples above, are so drastically different so as to be nearly incomparable (not shown, but median correlation was $\sim 0.2-0.5$). So we're likely to suffer from both a false-positive and false negative problem if we start to seriously parse Huh-7.5 cell data without informed knowledge of actual miRNA targeting in the liver. This will be the stated goal of the next chapter.

But we can begin consideration of an important question: how might the HCV miR-122 sponge impact a hepatocyte? Work with miR-122 knockout mice, which develop progressive liver disease that spontaneously results in HCC, suggests that miR-122 tumor suppressor activity is essential for long-term liver homeostasis (Hsu et al., 2012; Tsai et al., 2012). It is tantalizing to speculate that miR-122 sequestration in a chronic HCV infection may be a molecular link to the heterogeneous liver dysfunction that characterizes HCV induced disease(s). Many miR-122 targets found in Huh-7.5 cells would corroborate this with no shortage of potential mechanisms. For instance, miR-122 targets that we confirm

or establish via CLIP and reporter measurements, such as P4HA1, PKM2, and MASP1 are known to be upregulated in fibrosis or HCC (Jung et al., 2011; Li et al., 2013b; Liu et al., 2014a), with MASP1 notable for being specifically linked with HCV-associated HCC (Saeed et al., 2013). P4HA1, or prolyl 4-hydroxylase, is a key enzyme in collagen maturation; miR-122 suppression of this enzyme is hypothesized as a mechanism to control fibrogenic responses in the liver and the associated action of hepatic stellate cells (Li et al., 2013b). Pyruvate kinase muscle isoform 2, a key enzyme in glycolysis, is thought to be a primary driver in cancer metabolism. That miR-122 targets PKM2 provides an intuitive hypothesis: loss of miR-122 promotes PKM2 expression, and alters cellular metabolism to rely on glycolysis instead of oxidative phosphorylation based cellular energetics. The enzyme, as a miR-122 target specifically, has been linked to poor progression of HCC and in breast cancer (Fong et al., 2015; Liu et al., 2014a). Lastly, mannan binding lectin-associated serine protease 1 (MASP1), known to be unregulated by HCV via unknown mechanisms (Saeed et al., 2013), according to our data appears to be a miR-122 target. This secreted protease is thought to promote the proliferation of hepatic stellate cells which in turn adopt a pro-fibrotic course. In this manner, levels of this protein are thought to be a specific indicator of HCV induced liver disease, though the putative role of MASP1 in the virus lifecycle is unknown. The above examples are but the tip of the iceberg. As we begin to consider miR-122 targets in the liver, the stage is set for HCV induced sponging to play a role in defining virus specific pathology.

Overall, as it remains possible that functional effects of such a sponge may primarily impact highly infected cells, our data highlight the possibility of searching for transcriptome level changes to the miR-122 target network in response to HCV infection in individual hepatocytes. The extension of CLIP and RNA-seq in single-cell and primary contexts provides a compelling platform to address these and other long-term disease driven changes to a miRNA target network.

Chapter 5: AGO-CLIP studies of mouse and human liver

Thus far, we have focused on a so called HCV driven "sponge effect" whereby viral RNA uses miR-122 to an extent that host miR-122 targets are depressed. What we haven't delved into very much are the identities and functions of these actual targets and their roles in HCV or liver biology. Due to the Huh-7.5 context, this has been appropriate, given that we cannot assume a similar transcriptome in these cells compared to hepatocytes. The work presented in this chapter aims to overcome this limitation by exploring miR-122 targeting in its native *in vivo* setting in both mouse and human liver. The availability of a miR-122 knockout mouse, and its associated cancer related phenotype, offers an excellent platform to unambiguously map the entire miR-122 regulated network *in situ* and begin to define the molecular consequences of miR-122 dysregulation. Combined with RNAseq, we are further likely able to distinguish between primary effects due to loss of direct miR-122 binding, and secondary effects of altered signaling. In addition, AGO-CLIP studies of human liver add direct and clinically relevant depth to results obtained in mice, by providing points of comparison across species and across tissue types (liver versus Huh-7.5 for instance) in the hunt for targets most likely to impact HCV or liver disease biology.

Throughout, while this work was conducted squarely with HCV in mind, it must be noted that no virus will be present in the ensuing work. As we'll see,

miR-122 use by HCV is but one of the many surprises to be gleaned from this miRNA.

Experimental setup for AGO-CLIP in miR-122 KO mice

This work was initiated jointly with Kalpana Ghoshal at the Ohio State University. Shortly after publishing the initial description of a miR-122 knockout mouse in 2012 (discussed below), Kalpana contacted us about performing AGO-CLIP in these mice. Sensing an overlap with the HCV story then forming, I was brought in to lead the effort. Or more accurately, I enthusiastically volunteered to do AGO-CLIP in this *in vivo* context.

My enthusiasm was multi-fold. At this stage, I was convinced that the miR-122 sponge effect was probably real, but validating targets became an issue in Huh-7.5 cells given the high false positive (transcripts expressed in Huh-7.5 but not in liver) and false negative (liver transcripts lost in Huh-7.5 cells) likelihoods of RNA expression compared to liver. The literature comparing Huh-7.5 and liver transcriptomes is thin, but analyzing RNAseq data generated by William Schneider of human fetal liver cultures (HFLCs), a presumed closer substitute for liver, the transcriptome correlation of these data with Huh-7.5 was ~0.4 (not shown). This low correlation for ostensibly the same tissue type did not bode well for direct extrapolation of relevant liver targets in Huh-7.5 cells.

The miR-122 knockout mouse is the ideal context to solve this problem and possess an unusual phenotype. While miR-122 has been proposed to influence liver development (Laudadio et al., 2012; Xu et al., 2010), the first main

surprise of the knockout mouse is that livers develop normally, with no observable defects in anatomy and physiology (Hsu et al., 2012). A liver phenotype does emerge at 10 weeks of age when microsteatosis and liver inflammation begin to occur. Hepatic triglyceride synthesis increases at this stage, and yet serum cholesterol, LDL and HDL, all decrease. The result is an accumulation of fat in hepatocytes (steatosis) that progressively gets worse as the animals age. Combined with an infiltration of IL-6 producing immune cells, this steatosis progresses to a well defined fibrosis and hepatitis by six months of age. Between 12-17 months of age, miR-122 knockout mice then spontaneously develop HCC where interestingly, the penetrance differs by sex: 50% of males develop HCC whereas only 10% of females are affected (Hsu et al., 2012). This difference phenocopies the known male skew of HCC in humans (El-Serag, 2011).

The the slowly unfolding HCC observed by Hsu, Ghoshal and colleagues in miR-122 knockout mice appears consistent with the human trajectory of disease as correlating with age, sex and hepatic injury. It also points to a role for miR-122 as a key factor in maintaining long-term liver homeostasis, as a putative tumor suppressor. This last point was addressed in the initial report, as miR-122 overexpression was able to rescue liver tumorigenesis in a model driven by overexpressed c-MYC, drastically reducing tumor burden (Hsu et al., 2012).

Along the lines of the main hypothesis of the HCV sponge, that chronic repression of miR-122 may influence HCV associated disease progression, the

miR-122 knockout mouse permits a refinement of this hypothesis by providing the extreme case. Since long-term deletion of miR-122 results in spontaneous HCC, perhaps by investigating the miR-122 targets network in these mice, we might gain insight as to how HCV may do the same in humans. This was the main motivation to pursue AGO-CLIP in these mice.

The second point of enthusiasm involves using miRNA knockout systems for AGO-CLIP studies. Historically, miRNA activity in mammalian cells has typically been characterized by miRNA overexpression and monitoring of a target reporter, usually in a contrived cell type or in a context where the miRNA is usually not expressed. This basic idea was extended in the pioneering AGO-CLIP study, where miR-124 (a brain specific miRNA) was overexpressed in HeLa cells to look for novel miR-124 dependent peaks (Chi et al., 2009). While no doubt a useful proof-of-principle for "validating" a miRNA target with CLIP, this idea is plagued with problems: from overexpression artifacts that cannot be confirmed with knockouts (Vidigal and Ventura, 2015), to the aforementioned context problem where the miRNA encounters targets it usually never observes *in vivo*. Performing CLIP in a miRNA knockout context addresses these limitations succinctly, and as we've seen with Δ miR-122 Huh-7.5 cells, allows for the unambiguous determination of miRNA target by looking for the specific *loss* of AGO-CLIP peaks. This approach has been validated *in vivo* with great accuracy using miR-155 knockout mice, probing effects of this miRNA on immune cell function (Loeb et al., 2012)

To begin, we decided that to avoid any confounding events due to emerging liver pathology at 10 weeks, we would perform AGO-CLIP in miR-122 knockout mouse livers and floxed littermate controls from 6-week old mice, before pathology emerges. Both control and knockout mice were on the same feeding schedule, and importantly, livers were flash frozen at the same time of day (mornings). AGO-CLIP was performed on four knockout livers (1 female, 3 males) and five floxed controls (2 female, 3 male), for a total of nine mice (Figure 5.1). The autorads gave clear 110kDa miRNA bands in addition to mRNA smears between 130-150kDa. mRNA sized regions were excised and used to make DNA libraries compatible with NGS and sequenced.

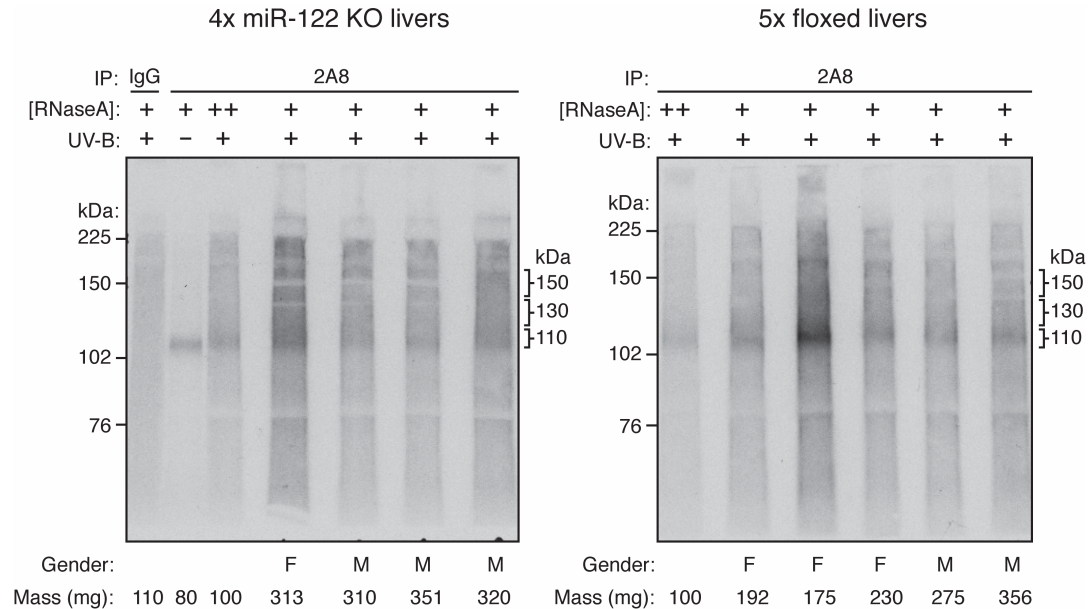


Figure 5.1. Autoradiogram of AGO bound RNA in miR-122 KO mice and floxed controls. Autoradiogram of ^{32}P -labelled RNA bound to AGO after IP using the pan-AGO 2A8 antibody. AGOs migrate at 97kDa, which shifts by RNA species bound by approximately 1kDa/3nt, thus AGO bound to ~22nt miRNA should be around 105kDa. IgG used as a non-specific control. Minus crosslink control shows the stringency of washes in that only a faint Ago:miRNA complex remains. High RNase and low RNase conditions indicated. Regions excised corresponding to 110, 130 or 150kDa are highlighted. Left panel depicts signal from four miR-122 knockout animals, where the gender and the amount of liver tissue used per sample is indicated. The right panel depicts the same but for five floxed control mice, treated as "WT."

As a confirmation of miR-122 loss, I first analyzed the miRNA profiles between WT and KO mice. While standard CLIP was performed with its attendant miRNA cloning bias issues, this analysis focuses only on comparing the same miRNA across conditions. Thus, assuming the cloning bias for a particular miRNA is constant between conditions, this issue is circumvented. We observed that mature miR-122 (-5p) and its passenger strand (-3p) were the

most significantly reduced miRNAs in knockout livers, as expected (Figure 5.2). An additional miRNA, miR-31-5p was also reduced, while a number of miRNAs were upregulated in knockout livers (Figure 5.2). These results suggest additional roles for miR-122 in affecting the abundance of other miRNAs.

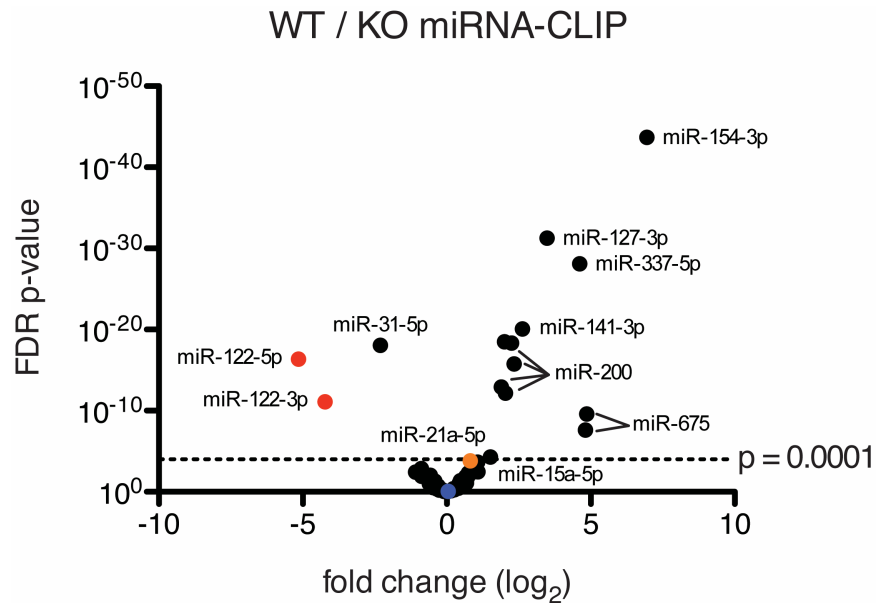


Figure 5.2. Volcano plot showing the log 2 fold change between miR-122 knockout versus wild-type mouse liver tissue of miRNA-CLIP results. Y-axis denotes FDR, where $p = 0.0001$ line is indicated. Notable miRNAs are highlighted by color and/or label.

Next, groomed reads were aligned to the mouse genome (mm9) and clustered mapped reads that overlapped by 1nt. Clusters were then grouped by BC and plotted as before to get a sense of the distribution of clusters across various genomic regions (see Chapter 4). In contrast to data in Huh-7.5 cells, where up to 50% of clusters mapped to 3'UTRs, most clusters in mice mapped to coding exons (>50%) with at best 30% in 3'UTRs Figure 5.3. For downstream

analysis of miRNA seeds, a cutoff of five libraries with a minimum peak height of 10 reads was used. Unlike Huh-7.5 data where only a BC threshold was used, the much higher depth of AGO-CLIP achieved in mouse liver made peak height filter appropriate, if only to enable a more "stringent" definition of clusters, though in truth, just as specified with Huh-7.5 cells, this call was arbitrary.

With these cluster definitions, a more detailed annotation was performed comparing all WT clusters to KO clusters, including those commonly shared. We observed that clusters in 3'UTRs were only mildly reduced in KO libraries (21.4% to 19.3%) (Figure 5.4). Given that a typically miRNA target contains binding sites for more than one miRNA, this is perhaps somewhat expected. Less expected was the decrease in coding exon clusters in knockout virus WT livers (40.2% to 31.5%), suggesting that a substantial fraction of unique miR-122 binding occurs in coding exons. Other genic regions, such as introns, transposable elements (TEs), 5'UTRs, and intergenic regions, were not appreciably different between KO and WT livers.

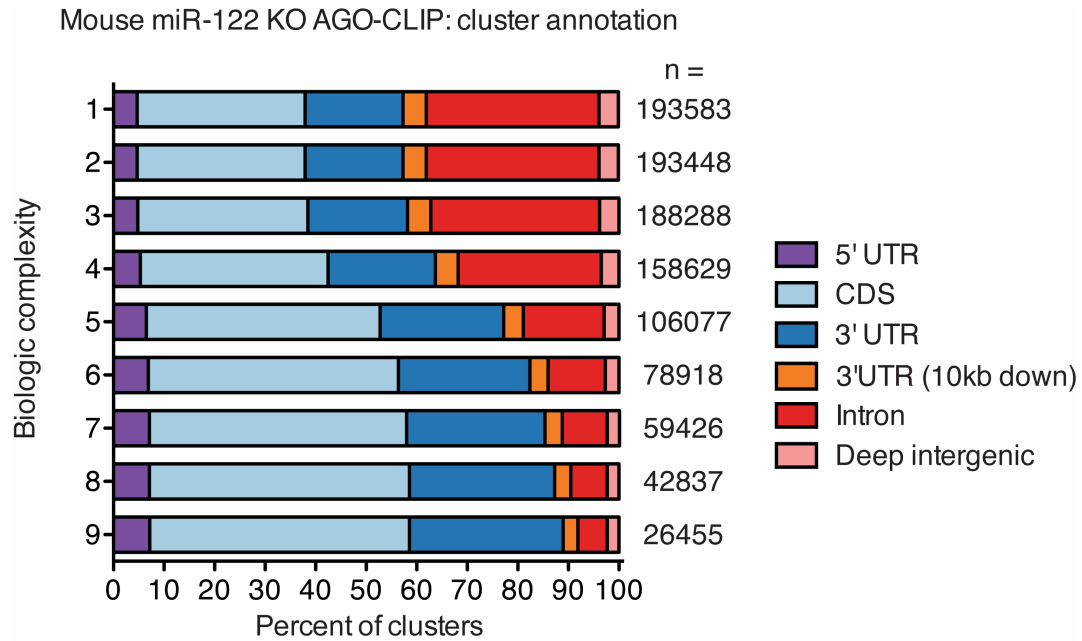


Figure 5.3. Cluster annotation results by increasing biologic complexity. AGO-CLIP reads from up to nine libraries from mouse liver were clustered with a minimum overlap of 2nt. Clusters with reads from any one, two, three and so on up to nine libraries the were annotated. The resulting proportions of clusters overlapping known genic elements is plotted above, alongside the total number of clusters with the respective biologic complexity (BC).

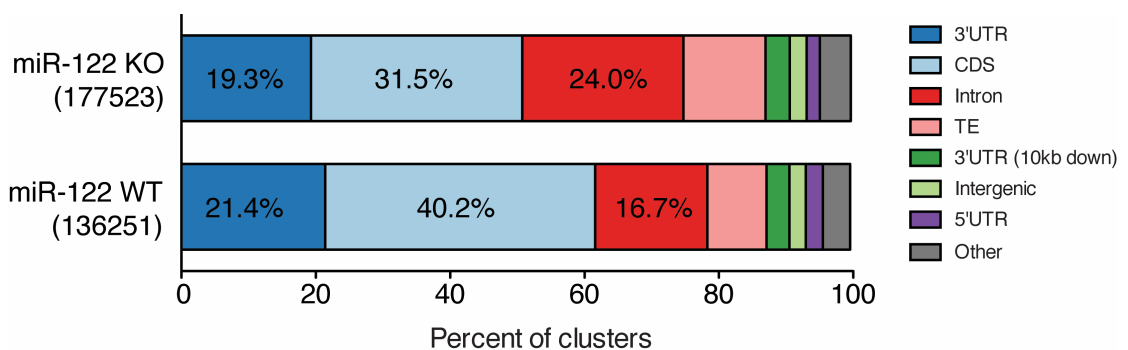


Figure 5.4. Proportion of clusters after robust clustering of overlapping reads from miR-122-deficient (KO) or floxed control (WT) libraries mapping to various regions of the genome. Clusters unique to each case plus those in common are indicated. TE, transposable elements; Other, includes non-coding RNA, satellite repeats, low complexity elements, etc.

miR-122 dependent peaks in mouse liver: some unexpected findings

Having confirmed miR-122 loss in knockout livers with CLIP, and after specifying BC and size cutoffs for clusters, we next approached defining miR-122 peaks. To do so in a less biased fashion than searching for the canonical miR-122 seed, we opted instead to search for instances where a cluster was specifically lost in KO libraries, extract the underlying sequence for the AGO footprint (+/- 32nt from the peak position), and search for enriched motifs (Figure 5.5A-B). For putative miR-122 dependent clusters, this analysis revealed significant enrichment for 8mer, 7merM8, 7merA1 and 6mer miR-122 seeds (Figure 5.5B). Unexpectedly, it also revealed a non-canonical miR-122 site motif, characterized by a G-bulge between the 6th and 7th seed positions (Figure 5.5B and D). Such "nucleation" bulges were observed previously with miR-124 between positions 5 and 6, and were demonstrated to be functional (Chi et al., 2012). We confirmed the presence of both canonical and G-bulged (hereafter also called miR-122-B) seeds by searching for motifs relative to AGO cluster peak coordinates. As expected we found significant enrichment for canonical miR-122 seeds within 20 nucleotides of AGO peaks (Figure 5.5C). We observed a similar enrichment, though fewer target sites, with the bulged motif (Figure 5.5D). Taken together, these results point to an alternative mode of miR-122 target recognition that occurs *in vivo*.

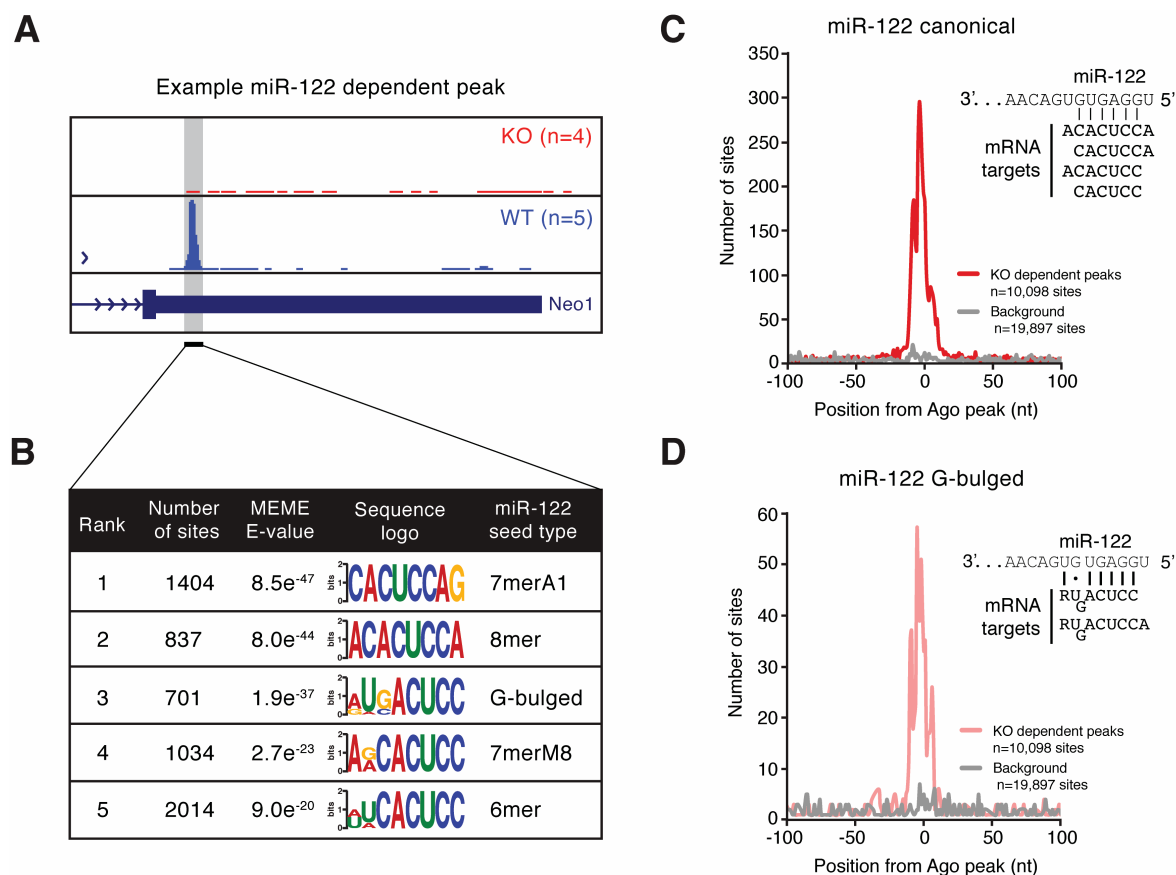


Figure 5.5. Identification of a novel G-bulged miR-122 motif. **(A)** Example of a miR-122 dependent peak. Robust WT signal is absent in KO libraries. **(B)** MEME analysis table of top motifs within clusters displaying miR-122 dependence as in (A). **(C)** Canonical miR-122 motif enrichment around Ago peaks. miR-122 KO dependent peaks (red) were compared to background clusters (gray). Inset shows miR-122 motif binding profiles. **(D)** Bulged miR-122 motif enrichment around Ago peaks. miR-122 KO dependent peaks (pink) were compared to background clusters (gray). Inset shows bulged miR-122 motifs binding profiles.

After assigning miRNAs (the top 100 expressed) to clusters by seed presence, we annotated the miR-122 and miR-122-B targetomes in detail (Figure 5.6). About 8% of CLIP clusters contained no discernable canonical miRNA seed, consistent with previous reports for "orphan" clusters (Chi et al., 2009; 2012). Roughly 5% of AGO-CLIP clusters contained a miR-122 seed, an additional 1%

of clusters contained a miR-122-B seed, for a total of ~4500 miR-122 targets spread across a variety of genomic regions (Figure 5.6A). We noted that 3'UTRs were targeted in similar proportion between canonical and G-bulged miR-122 sites, however, coding exon targeting predominated for both, and was considerably expanded for miR-122-B targets (Figure 5.6B). Additionally, between 20-40% of miR-122 targeting appears to occur in regions beyond 3'UTRs and coding exons. miR-122 dependent AGO targeting of introns, intergenic regions, 5'UTRs and TEs is totally novel and functionally unexplored, but clearly expands the AGO:miR-122 bound transcriptome. Combined with a binding preference for coding exons, and the identification of a G-bulged miR-122 seed, the classic assumption by virtue of bioinformatic target prediction that most miRNAs regulate a few hundred targets, is perhaps in need of revision.

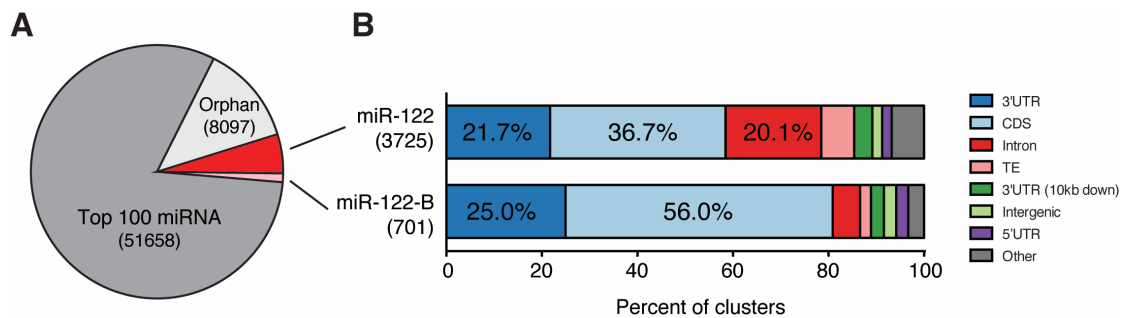


Figure 5.6. Annotation of miR-122 dependent clusters. **(A)** Pie chart highlighting miR-122 or bulged miR-122 targets as a proportion of all miRNA targets in mouse. Of the ~64,000 miRNA matched clusters roughly 5% are canonical miR-122 bound targets, and 1% G-bulged sites. Total number of clusters per category indicated. **(B)** Genomic annotation breakdown of canonical or bulged miR-122 sites.

Global analysis of miR-122 AGO peaks in WT versus miR-122 KO mice

With an expanded set of RNAs containing miR-122 or miR-122-B seeds, we next sought to quantify the fold change between WT and KO libraries. Similar to our analysis comparing HCV versus mock infected Huh-7.5 cells, we first calculated from read depth normalized data the log₂ fold-change between WT and KO samples per miRNA-assigned cluster. The cumulative distribution of these score was plotted, binned by miRNA family. Again, for these plots, if there was no difference in AGO binding at a particular cluster between WT and KO libraries, the score for that cluster was near 0. Changes showing a decrease of AGO binding upon miR-122 loss would be >0 , while increases in AGO binding would be <0 (Figure 5.7A).

Focusing on 3'UTRs, we observed a highly significant reduction in AGO binding in miR-122 seed containing clusters in KO libraries. This was true for both canonical and G-bulged seeds (Figure 5.7B). Other miRNA families, such as miR-21 and miR-15/16 were not significantly altered in their AGO-CLIP distributions. Two examples of canonical and G-bulged miR-122 targeting are shown in Figure 5.7D-E. In both cases, a miR-122 seed can clearly be discerned as underlying peaks lost in knockout livers.

To corroborate our CLIP results with effects on RNA abundance, we performed RNAseq in the same livers used for CLIP. Here we expect that loss of miR-122 should result in a measurable increase of target mRNA levels presumably for both seed types.

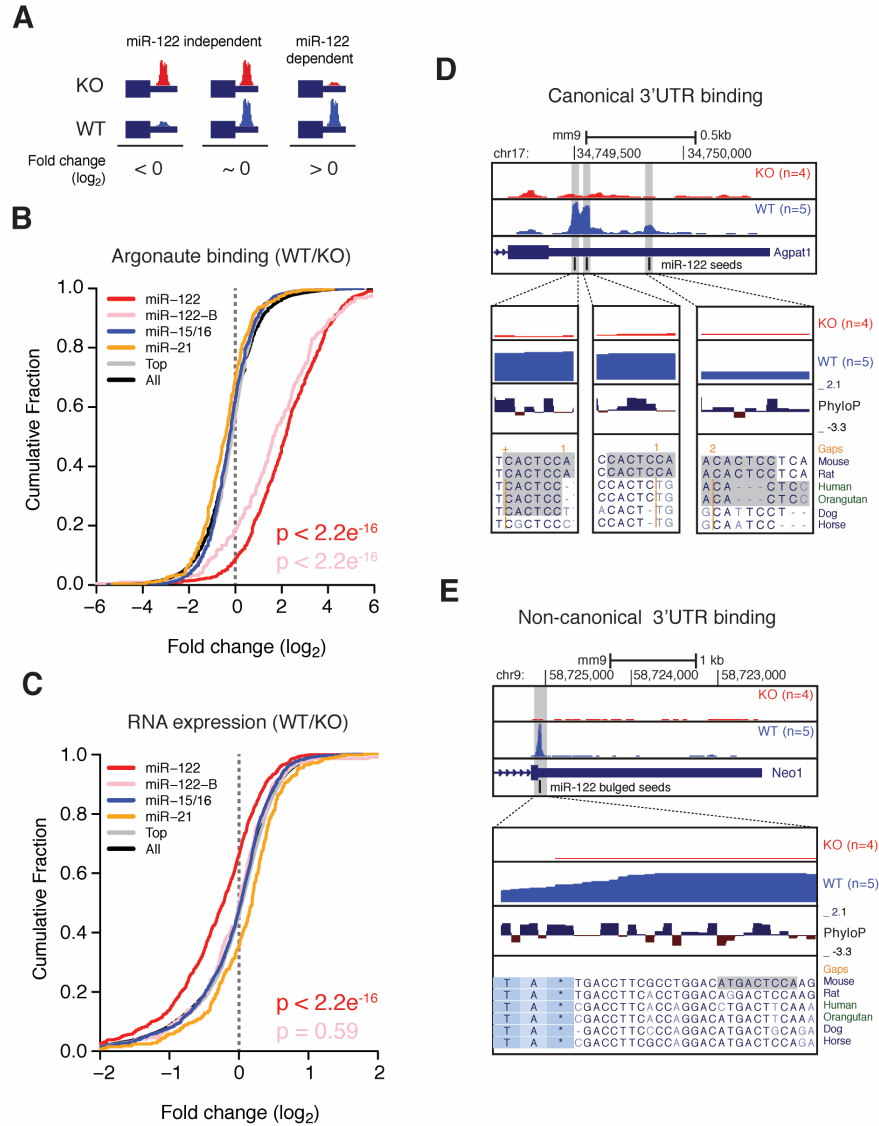


Figure 5.7. Global AGO binding changes on 3'UTRs affect mRNA abundance specifically for miR-122 targets. **(A)** miR-122 dependence definition by Ago binding profiles. **(B)** Cumulative distribution fraction (CDF) plot of the log2 fold change in Ago binding within cluster harboring indicated miRNA seeds. KS Test p-values indicated for miR-122 distributions vs all targets. **(C)** CDF plot of the log2 fold change in RNA expression of transcripts harboring indicated miRNA seeds, as in (B). Clusters were collapsed by gene so that no gene is represented twice. KS Test p-values indicated for miR-122 distributions vs all. **(D)** Canonical 3'UTR binding example of the *Agpat1* 3'UTR. Inset nucleotide conservation scores (PhyloP) and multi-alignments. **(E)** Non-canonical 3'UTR binding example of the *Neo1* 3'UTR. Inset nucleotide conservation scores (PhyloP) and multi-alignments.

While we observed robust de-repression for canonical targets, we noted that 3'UTR targets with G-bulged seeds were not significantly de-repressed (Figure 5.7C). Thus it seems that while the non-canonical seed is clearly used in the sense that AGO:miR-122 complex binds, it does not appear to confer any repression.

There are a few plausible explanations for this. Unlike the results obtained with miR-124, where G-bulges occur at between positions 5 and 6 of the mRNA and maintain full Watson-Crick base pairing (Chi et al., 2012), the miR-122 G-bulge occurs between positions 6 and 7 and induces a G:U base pair at position 7. Conceivably, this may confer a difference in the ability for miR-122 to nucleate on a target where the end result is a less stable seed with a G:U base pair. Structural evidence is consistent with this hypothesis, where notably base pairing at positions 6 and 7 requires that the target mRNA displace an alpha-helix that normally kinks the miRNA at the end of the seed region and permits binding beyond position 6 (Schirle et al., 2014). A bulge in the corresponding position in the mRNA target may alter the conformational dynamics of AGO, to say nothing of potentially modulating repression. For now, the most conservative interpretation for the lack of repression for G-bulged miR-122 targets is to take them at face value, and posit that G-bulged targets may be competitive inhibitors of canonical miR-122 activity, may serve as reservoirs for AGO:miR-122 complexes, and/or possess additional functions beyond mediating mRNA repression.

The same conclusion appears to hold for miR-122 binding in regions beyond the 3'UTR. Summarized in Figure 5.8, we noted that for both canonical and G-bulged miR-122 targets, significant loss of AGO binding was observed in KO samples in coding exons, introns, transposable elements (LTRs, LINEs and SINEs), intergenic regions, and 5'UTRs (Figure 5.8A). However, for transcript regions that can be measured by RNAseq, only canonical miR-122 seed targets in 3'UTRs and to a lesser extent coding exons were de-repressed as a class in KO livers (Figure 5.8B). No significant effects in RNA levels were observed for intronic and 5'UTR targets ((Figure 5.8B). Taken together, these data provide robust confirmation for functional miR-122 regulation of 3'UTR targets as a prevailing mechanism for regulating target mRNA levels, while at the same time showcasing the limits of miRNA targeting function as strictly impacting mRNA levels. What all the extra AGO:miR-122 binding is functionally doing can only be guessed at with these data, though on a transcript level some targets are worth emphasizing.

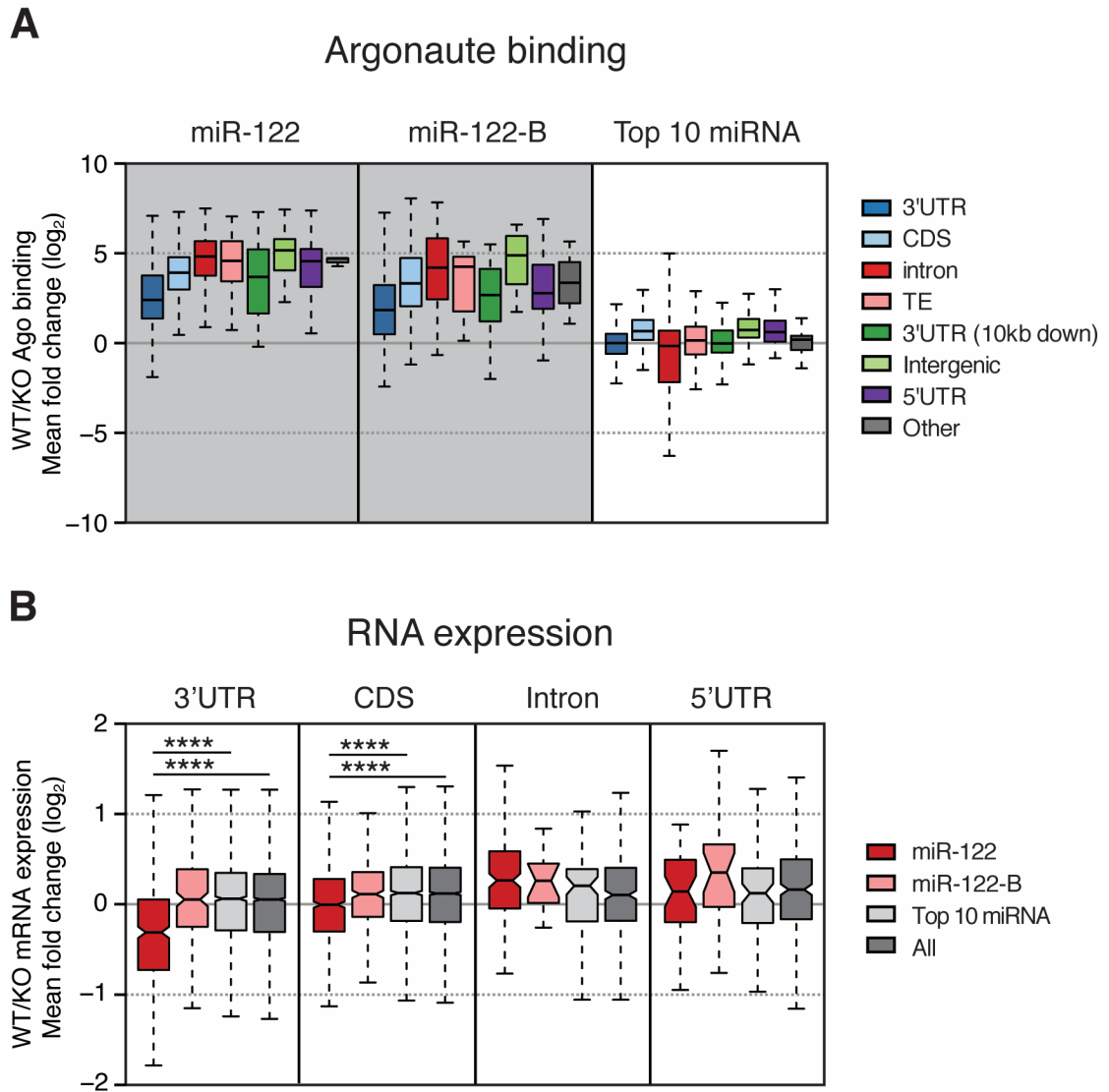


Figure 5.8. Widespread canonical and non-canonical miR-122 targeting throughout the transcriptome. **(A)** AGO binding profiles for miR-122, bulged miR-122, and the top 10 miRNA family targets (exclusive of miR-122), broken down by annotated region. Gray shading indicate statistical significance ($P < 10e-10$, KS-test) for each annotated category compared to the same category among the top 10 miRNA families. **(B)** RNA expression profiles (as in a) for miR-122, bulged miR-122, the top 10 miRNA, and all miRNA family targets, by expressed annotated region. Asterisks indicate statistical significance ($P < 10e-10$ KS-test) for each annotated category comparison. All other comparisons were not significant.

A brief survey of miR-122 dependent AGO binding beyond 3'UTRs

In the following sections, I will focus primarily on 3'UTR targets derepressed in miR-122 KO livers as the key players whose dysregulation probably contributes most to the cancer phenotype of these mice. Still, some discussion is appropriate for non-3'UTR genic targets, especially those for which a measurable change in mRNA abundance can be assessed from paired RNAseq data. While repressive 3'UTR targeting may be the most straightforward to interpret globally across a transcriptome, it is by no means definitive at the transcript level, as the case of HCV clearly demonstrates. So we must suspend our general concept of "miRNAs as repressors of mRNA at 3'UTRs" when we zoom in, and consider alternatives.

For instance, perhaps miR-122 binding stabilizes cellular RNA, as in HCV, or activates translation, as in AU rich elements in the 3'UTR of TNF mRNA under stress (Vasudevan et al., 2007). An enhancement of translation has also been shown for AGO binding the 5'UTR of ribosomal protein transcripts with miR-10a (Ørom et al., 2008). Intronic targeting has in general been only obliquely addressed functionally, but given the plurality of evidence of AGO shuttling to the nucleus, where localization is often dictated by miRNA isoforms, roles in regulating mRNA stability, alternative splicing, mRNA processing, and perhaps RNA transcription cannot be ruled out (Hwang et al., 2007; Schraivogel and Meister, 2014). Lastly, recent evidence that non-linear splicing of exons to form

circular RNAs (circRNAs) may act as miRNA sponges offers yet another functional mechanism to consider for non-3'UTR binding (Memczak et al., 2013). In all of the above cases, it is not necessary that mRNA levels be impacted, so our inability to see a change in RNA abundance does not preclude additional and as yet unknown functions. Just the same however, it is important to stress that the null hypothesis, that these non-3'UTR AGO binding events "aren't doing much" must be taken seriously, particularly when considering the lack of purifying selection observed for the majority of non-3'UTR sites (discussed in the next section). Still, some outliers appear very interesting.

Carboxylesterase1 (Ces1) proteins are a class of enzyme that participate in drug, lipid and xenobiotic metabolism of small molecules. Primarily expressed in the liver, the Ces1 gene has undergone many duplication events with at least 5 members clustered in a 0.2Mb stretch of chromosome 8 (Holmes et al., 2010). Interestingly, Ces family members 1d, 1e, and 1f all exhibit miR-122 dependent binding in coding exons, and all are significantly down-regulated (~2-3 fold) in KO livers (Figure 5.9A). Notably, Ces1 expression also goes down in mouse livers with age (Fu et al., 2012). Thus, it's possible that miR-122 may act to stabilize or stimulate translation of this protein family, and that over time, lower miR-122 results in lower expression of these detoxifying enzymes. Thinking of miR-122 KO livers from an aging perspective might be an interesting new frame for interrogating the altered liver biology of these mice.

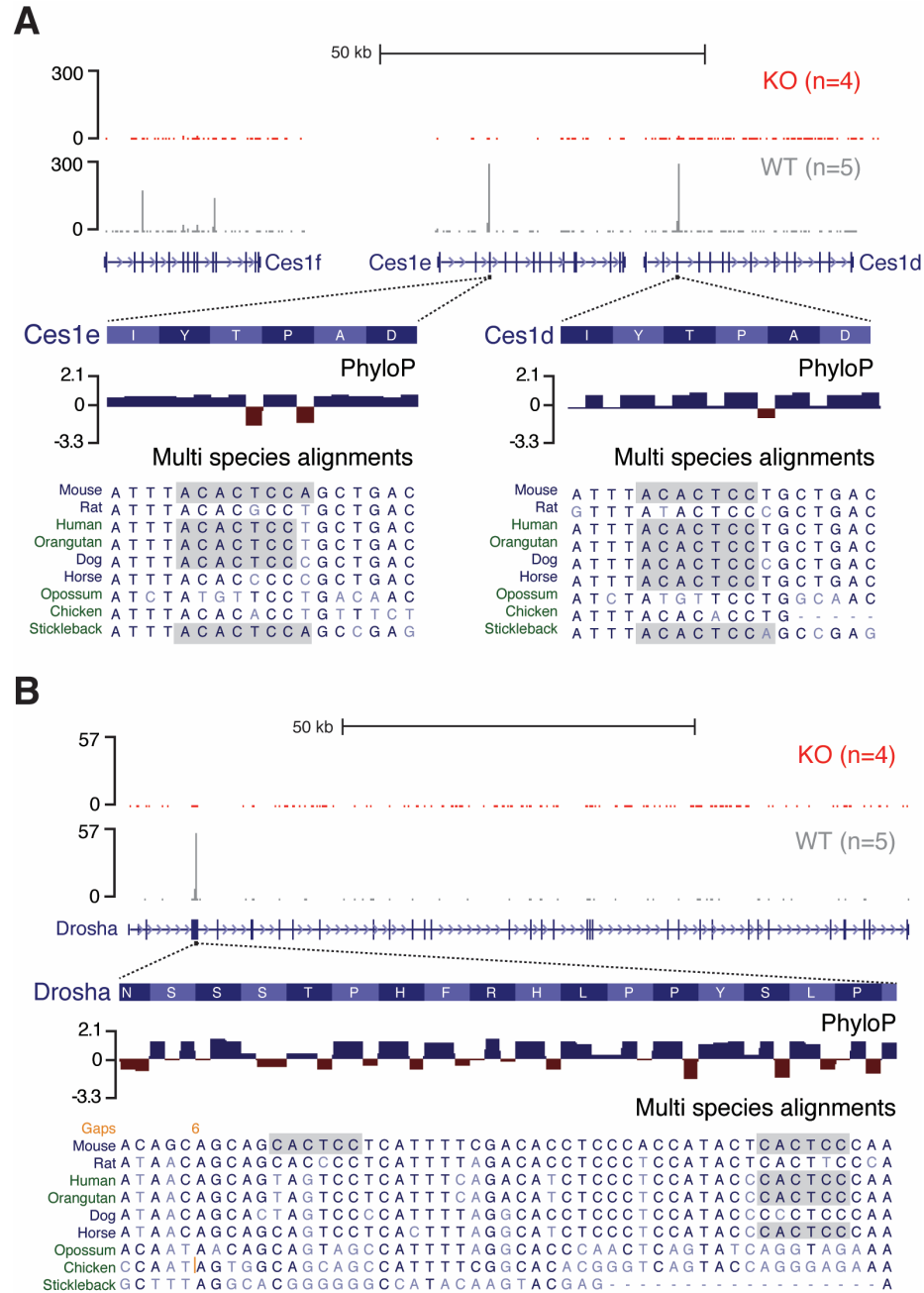


Figure 5.9. Two examples of miR-122 regulation on cellular coding exons. **(A)** Ces1 or **(B)** Drosha genes are shown below WT and KO CLIP tracks after read depth normalization. Zoom in view displays nucleotide conservation scores (PhyloP) and multi species alignments for indicated regions. miR-122 seeds highlighted in gray. In both cases, transcript levels were significantly altered in KO livers. See text for details.

Another example of coding exon targeting is Drosha, a key nuclear RNase involved in miRNA biogenesis. The second exon of Drosha has two miR-122 sites in mice, one of which is reasonably well-conserved (Figure 5.9B). As Drosha expression goes up in KO livers, it's likely that miR-122 is normally playing a repressive role on Drosha gene expression. This could constitute a feedback loop to regulate Drosha expression levels as a means of controlling global miRNA biogenesis. In this manner, the cell might be able to "sense" altered miR-122 (or other miRNA) abundance and regulate miRNA processing accordingly. Along these lines, the Dicer 3'UTR possesses a non-canonical miR-122 site, found only in mice (not shown), however, little change in RNA levels was noted.

Additional interesting results can be observed in 5'UTRs that possess miR-122 sites. One example is CXC chemokine-ligand-1 (Cxcl1). Previously called Gro-alpha, this inflammatory cytokine is produced by a variety of cells and binds the Cxcr2 chemokine receptor to promote fibrogenesis and angiogenesis. Shown in Figure 5.10A, it possesses clear miR-122 binding sites in its 5'UTR. In KO livers, Cxcl1 expression is significantly increased (~3.2 fold), thus it's likely that miR-122 normally plays a repressive role in binding the 5'UTR. Cxcl1 expression has been recognized as an inflammatory hallmark in a number of cancers, including HCC (Bandapalli et al., 2012; Tang et al., 2012). Moreover, increased Cxcl1 expression has been identified as a predisposition marker for HCV and alcohol induced cirrhosis, where interestingly a SNP in the Cxcl1 intron

appears to be involved (Nischalke et al., 2012; 2013). In mice, two miR-122 sites are present in the 5'UTR, however, and this is a theme which we'll return to, neither of them are conserved in humans (Figure 5.10A).

Another example of 5'UTR binding is Mavs, a well known anti-viral signaling molecule and as discussed in the introduction, a target for the HCV NS3/4A protease to blunt the innate immune response. The Mavs 5'UTR robustly binds miR-122 and expression is elevated ~2 fold in KO livers, thus pointing to a normally repressive role for miR-122 binding (Figure 5.10B). It's certainly tempting to hypothesize that Mavs regulation by miRNA at the 5'UTR may act as a miR-122 sensor, where in the case of the HCV:miR-122 sponge, an up regulation of Mavs may help dampen the effects of NS3. But alas, this makes zero sense considering that the miR-122 site is not conserved in humans (Figure 5.10B).

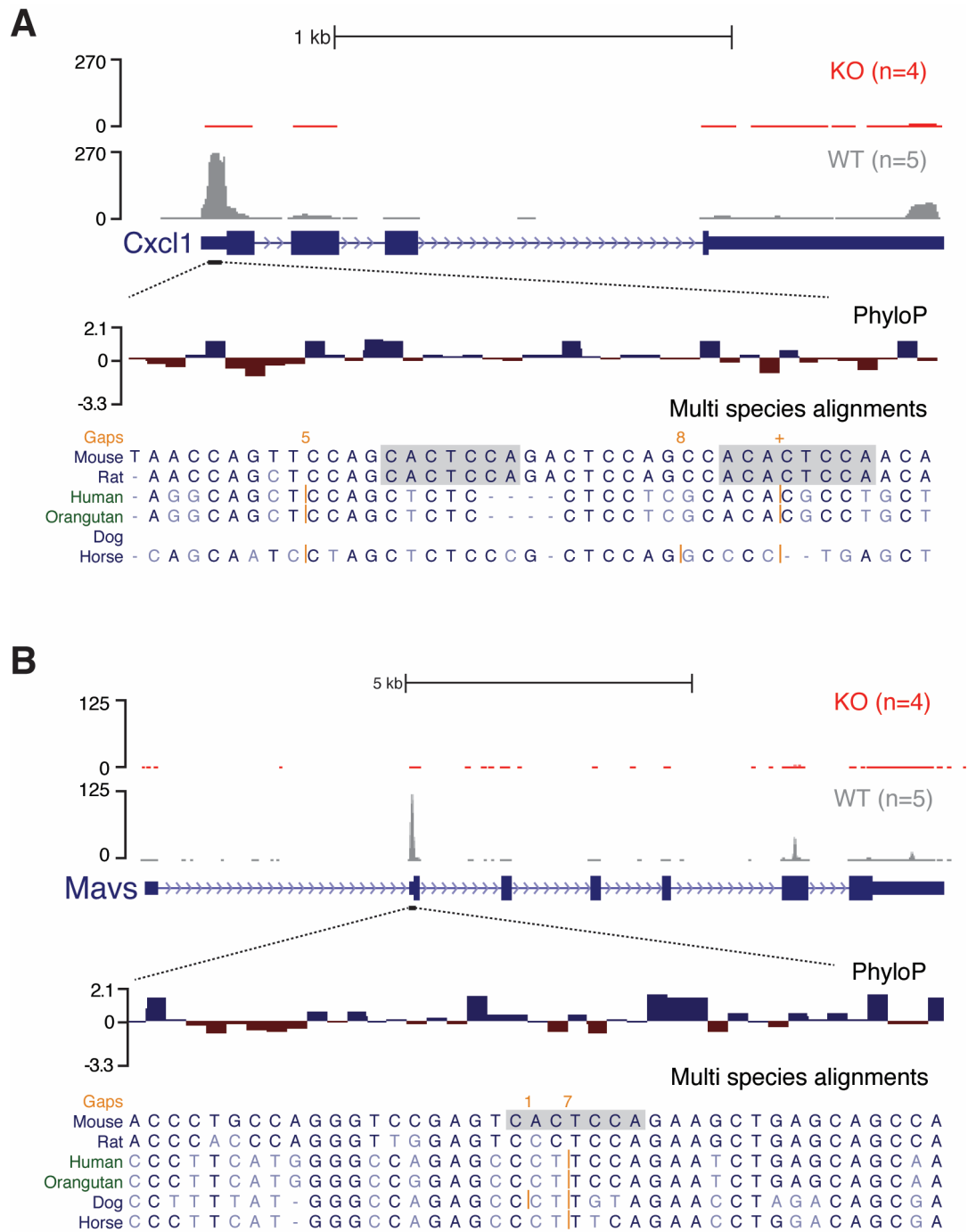


Figure 5.10. Two examples of miR-122 regulation on cellular 5'UTRs. **(A)** Cxcl1 or **(B)** Mavs genes are shown below WT and KO CLIP tracks after read depth normalization. Zoom in view displays nucleotide conservation scores (PhyloP) and multi species alignments for indicated regions. miR-122 seed highlighted in gray. In both cases, transcript levels were significantly upregulated in KO livers.

miR-122 3'UTR targets are poorly conserved

Despite clear and robust miR-122 binding along with a correlated change in expression level, the preceding examples appear in large part to be species specific, given that the miR-122 seed sites do not appear conserved in mammalian lineages, especially between mouse and human. We can explore this in a more general fashion by considering the nucleotide conservation of the core 6mer of a miRNA binding site, using PhyloP conservation scores across sequenced mammalian genomes (Pollard et al., 2010). Importantly, PhyloP scores are calculated per nucleotide and assume a neutral model of evolution. Positive scores indicate high conservation, while negative scores indicate rapidly diverging sequences; scores near zero are undergoing neutral drift. By plotting the mean PhyloP score across a miRNA 6mer, and binning by miRNA and genomic region, we can discern how well conserved miR-122 target sites are relative to other highly expressed miRNAs. The expectation, largely afforded by miRNA target prediction efforts, is that highly conserved miRNAs (such as miR-122) usually regulate a conserved set of targets.

The results are shown in Figure 5.11. As expected exonic regions are more conserved (>0) than intronic regions (~ 0) which are largely undergoing neutral drift. For the top ten expressed miRNAs, conservation on 3'UTRs approaches that of coding exons. 3'UTR targets for miR-122 however, are significantly less conserved compared to the top 10 or all miRNAs (Figure 5.11A). Such a pattern is repeated among coding exon miR-122 targets generally, and is

magnified by focusing on 3rd codon positions, which are under less selective pressure to maintain protein coding sequences at the nucleotide level (Figure 5.11B). Bulged miR-122 sites generally were not statistically different than the top ten miRNA or all miRNAs. These results are somewhat surprising in that they imply that most canonical miR-122 targeting in mice is species specific. Moreover, it also predicts that only a select subset of targets are shared between mouse and human, and that humans likely have a completely different set of miR-122 regulated targets. The causes, consequences, and potential mechanisms for miR-122 species specificity despite the fact that this miRNA is highly conserved down to fish will be addressed in the discussion. For now, it suffices to bring this point up as we begin to consider what classes of targets are derepressed in KO livers.

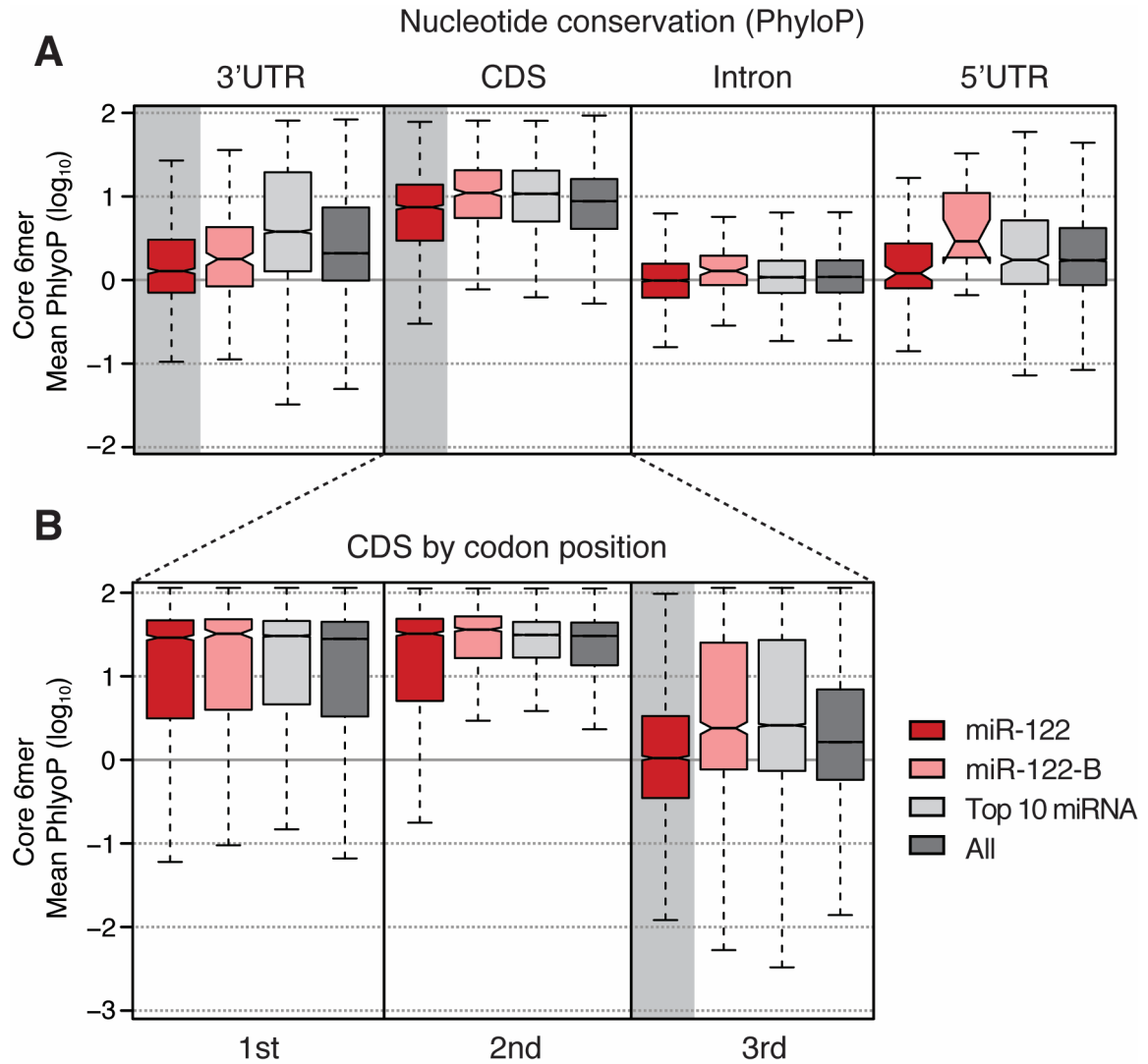


Figure 5.11. miR-122 CLIP targets are not well conserved in mouse. **(A)** Mean nucleotide conservation scores for mouse miR-122, bulged miR-122, the top 10 miRNAs and all miRNA targets, denoted by genic region. PhyloP scores from placental mammals were averaged across the core miRNA 6mer and plotted. Shaded areas indicate statistical significance compared to the top 10 or all miRNA targets ($P < 10e-10$, KS-test). **(B)** Per nucleotide PhyloP scores calculated as in (A) were further subdivided by codon position for coding exons. Shaded areas indicate statistical significance ($P < 10e-10$, KS-test).

miR-122 regulates a diverse set of metabolic proteins

Pioneering studies in mice following miR-122 antagonism with anti-sense oligonucleotides (ASOs) or locked nucleic acids (LNAs) revealed altered lipid profiles characterized by reduced plasma cholesterol levels, increased hepatic fatty-acid oxidation, and a decrease in hepatic cholesterol and fatty-acid synthesis rates (Elmén et al., 2008; Esau et al., 2006; Krützfeldt et al., 2005). These studies identified a number of direct miR-122 targets using microarrays, and further defined secondary targets whose expression was altered but did not contain predicted miR-122 sites. Indeed several genes with crucial to lipid metabolism but lacking miR-122 sites (*Acc1*, *Acly*, *Scd1*, *Srebp2* and *Hmgcr*) were down-regulated (Esau et al., 2006; Krützfeldt et al., 2005). Consistent with these findings, miR-122 antagonism decreased serum cholesterol levels in chimpanzees (Lanford et al., 2010).

As the miR-122 KO mouse provides a more detailed identification of all miR-122 regulated genes (via RNAseq) and enables subsetting of direct targets (via AGO-CLIP), we revisited proposed groupings of miR-122 targets using gene ontology (GO) analysis. We first focused on all de-repressed mRNAs in miR-122 livers using all expressed genes as the background control. We found that transcripts encoding small molecule and lipid metabolism were highly enriched biologic processes (Figure 5.12A). Significant enrichment for proteins in membrane bound organelles and secreted vesicles was also observed (Figure 5.12B). Focusing on direct miR-122 targets (CDS and 3'UTR), we observed a

similar enrichment for biologic process and cellular component (Figure 5.12C-D). No qualitative difference was observed between GO results of miR-122 versus bulged miR-122 targets. Taken together, these data largely confirm and expand upon previous studies with the identification of a non-canonical miR-122 binding site as contributing to the overall theme of miR-122 as a key regulator of liver metabolism. And yet, as we consider the apparent species specificity of a substantial fraction of miR-122 regulation, we must ask: does the same hold for humans?

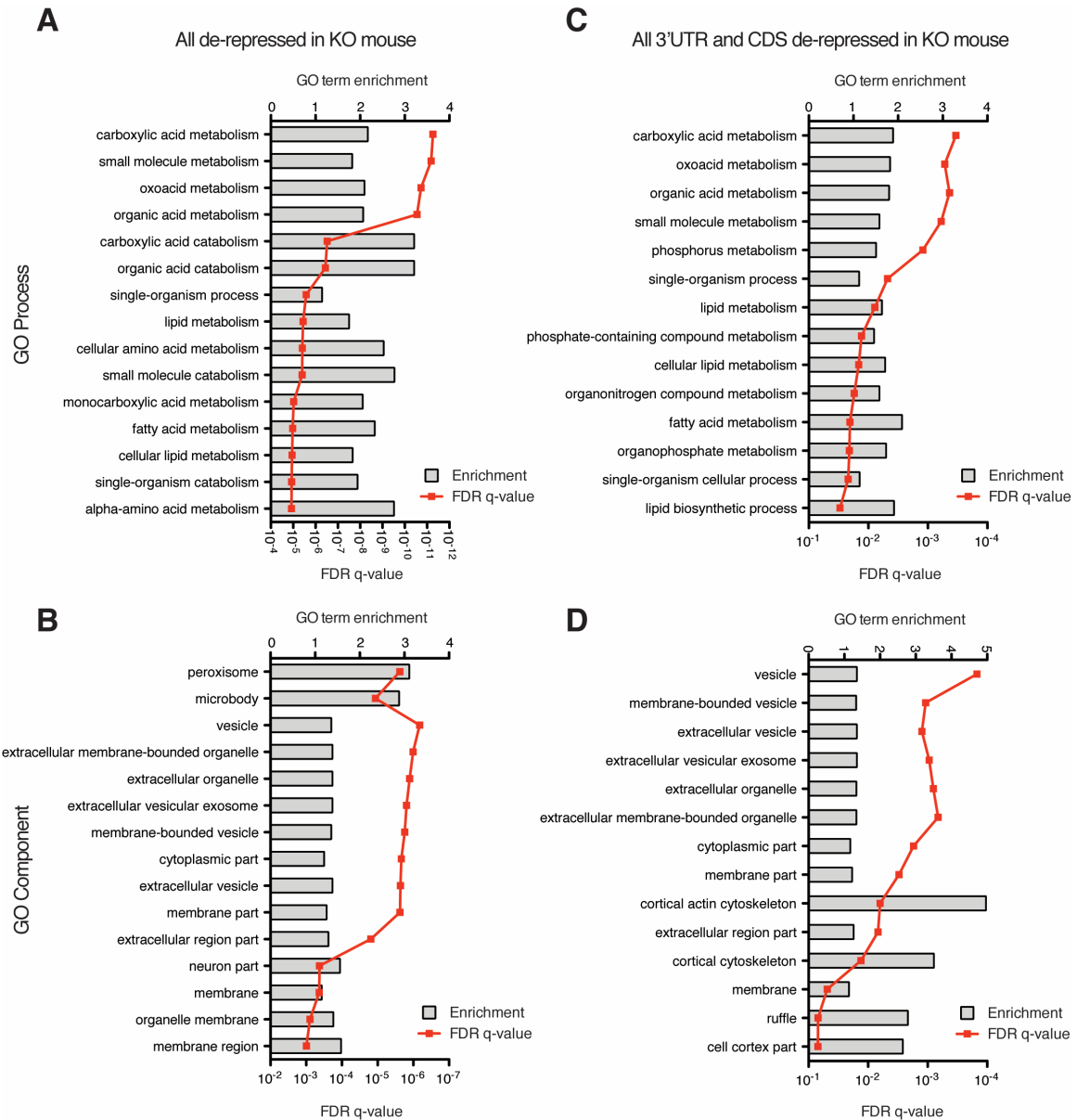


Figure 5.12. Gene ontology (GO) analysis of de-repressed miR-122 targets in mice. **(A)** GO term enrichment and FDR significance of the top GO Process terms for all miR-122 targets de-repressed in miR-122 KO mice. **(B)** As in (A), but for the top GO Component terms. **(C)** GO term enrichment and FDR significance of the top GO Process terms for all 3'UTR and CDS miR-122 targets de-repressed in miR-122 KO mice. **(D)** As in (C), but for the top GO Component terms.

Establishing a miR-122 target network in human liver

The results presented thus far with miR-122 knockout mice livers have been illuminating to define the complex web of targets directly and indirectly regulated by miR-122. Given the high conservation and expression of miR-122 in mammalian livers, it has so far been reasonable to assume that miR-122 likely plays similar roles in these varied organisms. Surprisingly, the high frequency of mouse specific miR-122 targeting suggests otherwise.

To address this and to enable a direct comparison between human and mouse, we decided to perform AGO-CLIP in human liver samples procured by Kalpana Ghoshal at OSU. Specifically, we chose liver resections from HCC patients to compare tumor to histologically "normal" adjacent tissue. Pathologically, these samples exhibited moderate to poorly differentiated HCC with adjacent fibrosis and cirrhosis (Table 5.1). Importantly, we used lower miR-122 expression in HCC tumors compared to adjacent tissue as an additional selection criteria (Figure 5.13).

We reasoned that the matched nature of these samples (tumor versus normal from the same patient) combined with differences in measured miR-122 abundance would increase our chances in identifying not just static miR-122 targets (i.e. transcripts with seeds) but also to observe AGO-CLIP binding changes analogous to the miR-122 KO mouse. AGO-CLIP was performed on five human HCC samples and five matched controls for a total of ten libraries.

Table 5.1 HCC subject characteristics

Sample	Age/sex	Pathological evaluation
1	58/F	HCC, moderately to poorly differentiated surrounding liver shows mild steatosis and chronic portal inflammation.
2	33/M	Multi-focal HCC with intrahepatic cholangiocarcinoma, minimal portal inflammation, scant steatosis.
3	59/M	Moderately differentiated HCC, with vascular invasion, established cirrhosis in surrounding liver.
4	63/M	HCC, moderately to poorly differentiated cirrhosis with moderate to marked microvesicular steatosis.
5	60/M	HCC, moderately to poorly differentiated Areas of necrosis present, Surrounding liver shows minimal inflammation without cirrhosis

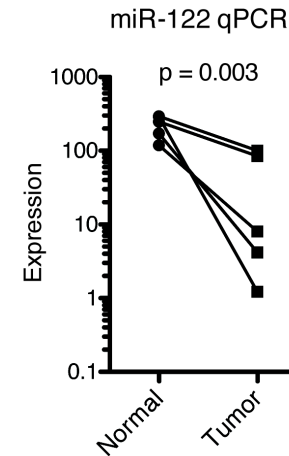


Table 5.1 and Figure 5.13. HCC subject characteristics and concurrent miR-122 qPCR in human liver samples. (Left) Subject characteristics table for five HCC patients. (Right) miR-122 qPCR in arbitrary expression values for paired tumor versus normal adjacent tissue for five HCC patient liver samples. Paired t-test p-value between normal versus tumor samples shown.

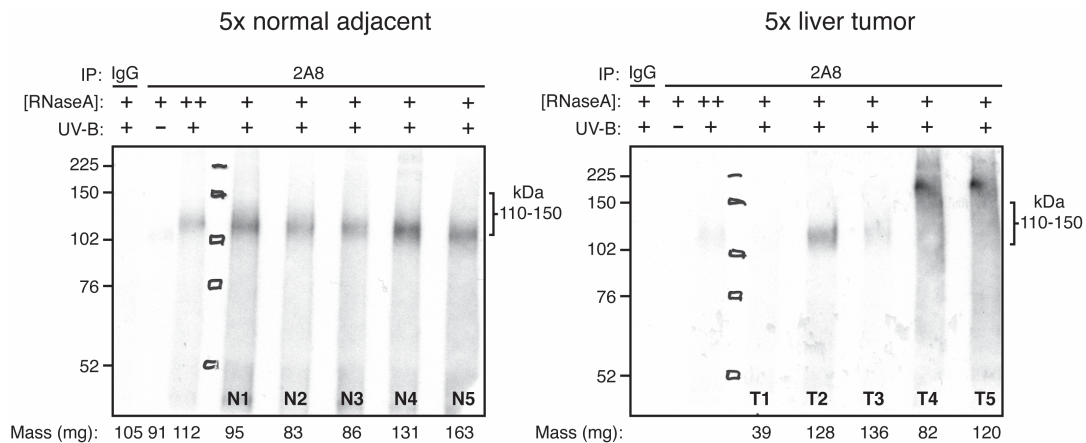


Figure 5.14. Autoradiogram of AGO bound RNA in human HCC tumors or matched normal adjacent tissue. Autoradiogram of ^{32}P -labelled RNA bound to AGO after IP using the pan-AGO 2A8 antibody. AGOs migrate at 97kDa, which shifts by RNA species bound by approximately 1kDa/3nt, thus AGO bound to ~22nt miRNA should be around 105kDa. IgG used as a non-specific control. Minus crosslink control shows the stringency of washes in that only a faint Ago:miRNA complex remains. High RNase and low RNase conditions indicated. Region excised corresponding to 110-150kDa complexes is highlighted. Left panel depicts signal from five patient samples of normal adjacent tissue, where the amount of liver tissue used per sample is indicated. The right panel depicts five liver tumors matched by number to normal adjacent tissue.

The autorads gave clear 110kDa miRNA bands in normal tissue, however some differences in signal intensity were apparent in matched tumor samples (Figure 5.14). One general reason for this may be the lower tissue volume used compared to mouse liver. Nevertheless, library construction after excising mRNA sized regions was successful (not shown) and so we moved forward with sequencing.

As before, the first analysis we performed was on the miRNA profiles between tumor and normal samples. We confirmed our qPCR results and observed that miR-122 was indeed reduced on AGO in tumor samples, in addition to other miRNAs, notably miR-194 and miR-144 Figure 5.15. Loss of miR-194 and miR-144 in hepatocytes have been previously implicated in HCC progression and liver inflammation (Krützfeldt et al., 2012; Li et al., 2015). A number of miRNAs, such as miR-21 and miR-221/222 exhibited increased AGO association in tumors, and largely confirms previous and recent observations (Dong et al., 2015; Karakatsanis et al., 2013; Ogawa et al., 2012; Sun et al., 2015). Taken together, these largely congruous results set the stage for exploring miRNA target regulation where clearly miR-122 is one of many miRNAs involved in cancer progression.

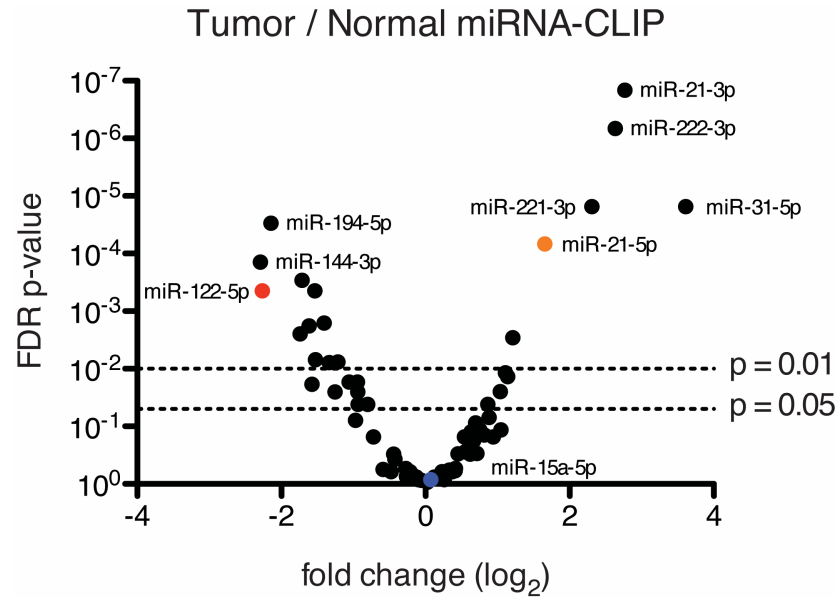


Figure 5.15. Volcano plot showing the log 2 fold change between tumor versus normal human liver tissue of miRNA-CLIP results. Y-axis denotes FDR, where $p = 0.01$ and 0.05 lines are indicated. Notable miRNAs are highlighted by color and/or label.

Next, groomed reads were aligned to the human genome (hg18) and clustered such that mapped reads that overlapped by 1nt. Clusters were then grouped by BC and plotted as before to get a sense of the distribution of clusters across various genomic regions (see Chapter 4). A greater proportion of coding exon targeting was evident across all libraries and this frequency did not appear to vary much with increasing BC (Figure 5.16). For downstream analysis of miRNA seeds, a cutoff of five libraries was used.

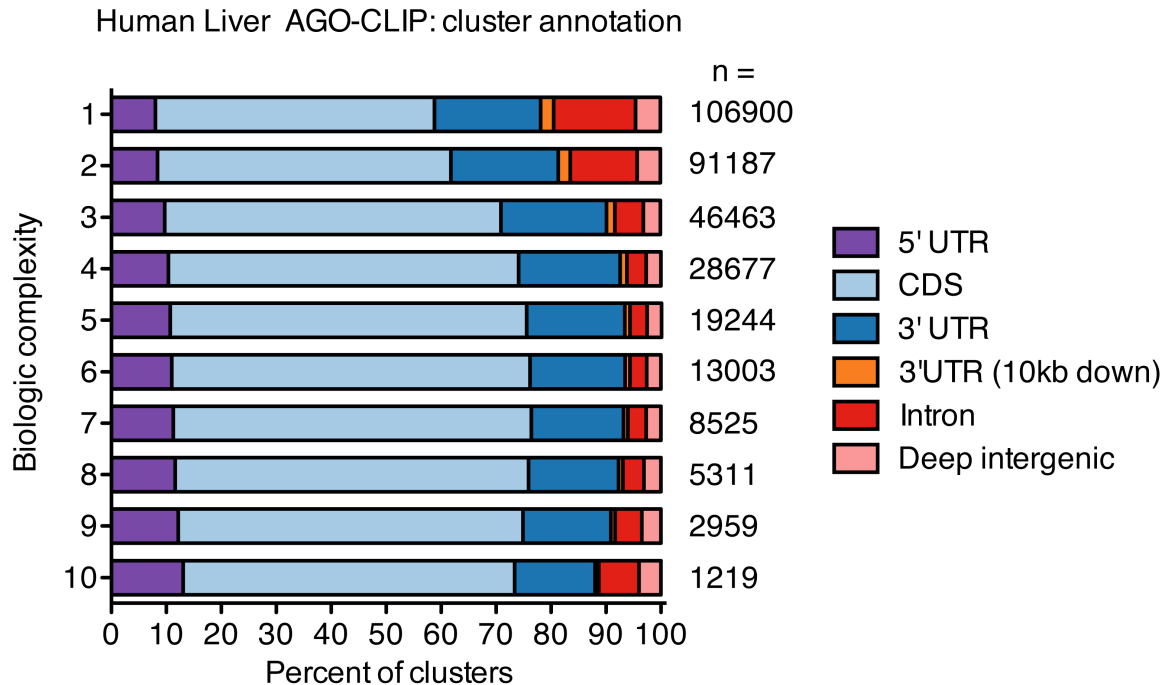


Figure 5.16. Cluster annotation results by increasing biologic complexity. AGO-CLIP reads from up to 10 libraries from human liver were clustered with a minimum overlap of 2nt. Clusters with reads from any one, two, three and so on up to ten libraries the were annotated. The resulting proportions of clusters overlapping known genic elements is plotted above, alongside the total number of clusters with the respective biologic complexity (BC).

With these cluster definitions, a more detailed annotation was performed comparing all normal tissue versus tumor clusters, including those commonly shared (Figure 5.17). Of the ~19000 clusters at this stringency, close to 20% were in 3'UTRs and an even greater proportion (~70%) mapped to coding regions than in mice. Other genic regions, such as introns, transposable elements (TEs), 5'UTRs, and intergenic regions, were not appreciably different between made up the remaining 10%. Comparing the same annotation in mouse livers (Figure 5.4), 3'UTR binding was consistent at 20% of all clusters. The

marked increase in coding exon binding between human and mouse (or conversely, the decrease in intronic and TE binding) is a bit puzzling and probably reflects distinct transcriptome differences between human and mouse liver. Some amount of bioinformatic noise can also not be ruled out.

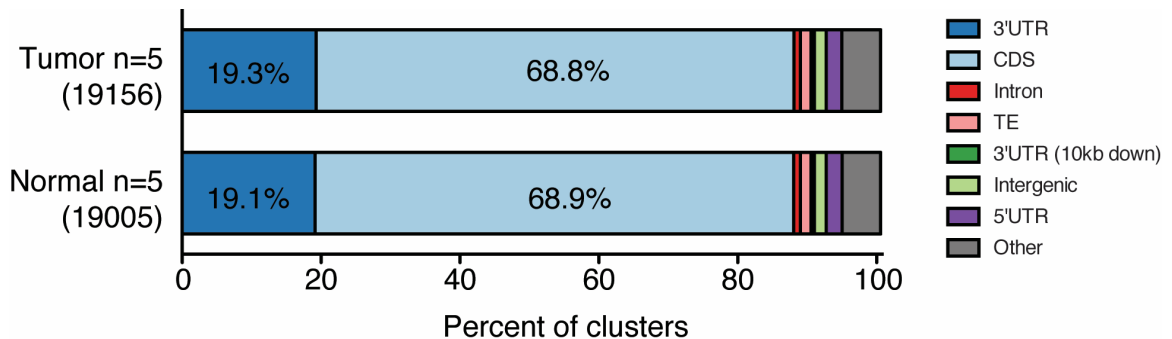


Figure 5.17. Cluster annotation of normal adjacent liver tissue (Normal) or matched tumor (Tumor) CLIP libraries from human liver mapping to various regions of the genome. Clusters unique to each case plus those in common are indicated. A BC filter of 5 of 10 libraries was specified. TE, transposable elements; Other, includes non-coding RNA, satellite repeats, low complexity elements, etc.

Just the same, we went forward with the same analysis performed in mouse liver, where clusters were grouped by the presence of miRNA seed and genomic region, and the log2 fold change between normal and tumor plotted as a CDF. In this plot, as in the mouse, positive values indicate greater AGO binding in the normal tissue, while negative values indicate greater binding in tumors (Figure 5.18A). For miR-122 3'UTR targets, we observed significant reductions in AGO binding in tumor tissue for both canonical and G-bulged targets, consistent with miR-122 loss in this context (Figure 5.18B).

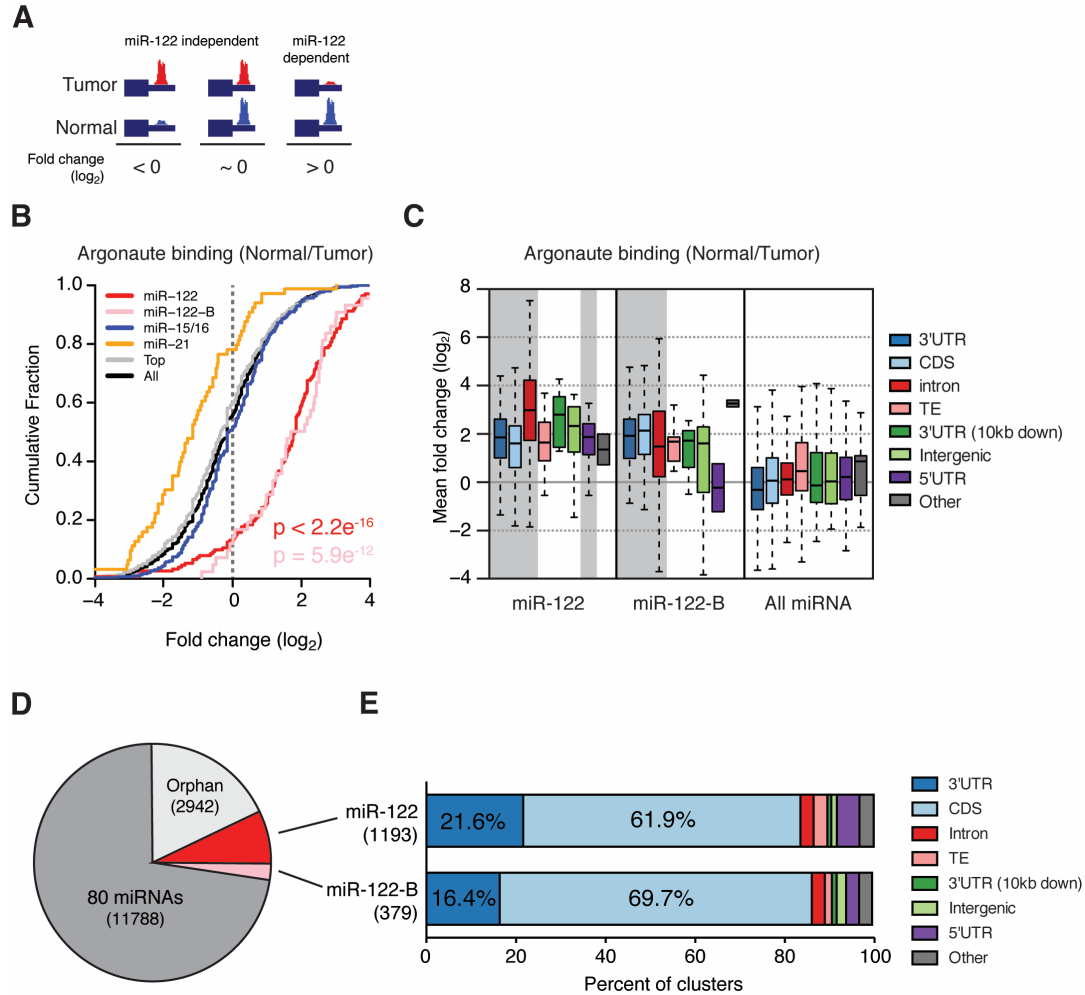


Figure 5.18. Global reduction of AGO binding miR-122 targets observed in liver tumors. **(A)** miR-122 dependence definition by Ago binding profiles. **(B)** cumulative distribution fraction (CDF) plot of the \log_2 fold change in Ago binding within cluster harboring indicated miRNA seeds. KS Test p-values indicated for miR-122 distributions vs all targets. **(C)** AGO binding profiles for miR-122, bulged miR-122, and all miRNA targets (exclusive of miR-122), broken down by annotated region. Gray shading indicate statistical significance ($P < 10e^{-5}$, KS-test) for each annotated category compared to the same category among all miRNA targets. **(D)** Pie chart highlighting miR-122 or bulged miR-122 targets as a proportion of all miRNA targets in mouse. Total number of clusters per category indicated. **(E)** Genomic annotation breakdown of canonical or bulged miR-122 sites.

Intriguingly, we also observed the opposite for some miRNAs, most notably miR-21, whose increased expression in tumors also results in increased AGO binding of miR-21 targets.

Focusing on miR-122 binding, we observed as in mice, that reduced AGO binding was not specific to 3'UTRs and was observed for both canonical and G-bulged sites significantly in genic regions (Figure 5.18C). The miR-122 targetome in human liver constituted around 9% of all AGO clusters, slightly higher than the ~6% in mice (Figure 5.18D). The miR-122 target network also skewed heavily to coding exons (>60%) and was generally similar to the global annotation of miRNA targeting in liver.

In the last analysis similar to that performed in mice, we investigated the conservation of miR-122 seeds using PhyloP scores and again observed that 3'UTR and to a lesser extent coding exon canonical miR-122 targets were generally less conserved than the top 10 or all expressed miRNA (Figure 5.19).

Taken together, these results extend conclusions obtained in miR-122 KO mice with the human observation of a G-bulged miR-122 seed, extensive non-3-UTR AGO binding, and importantly, poor conservation of miR-122 seeds that may point to high species specificity of miR-122 target regulation.

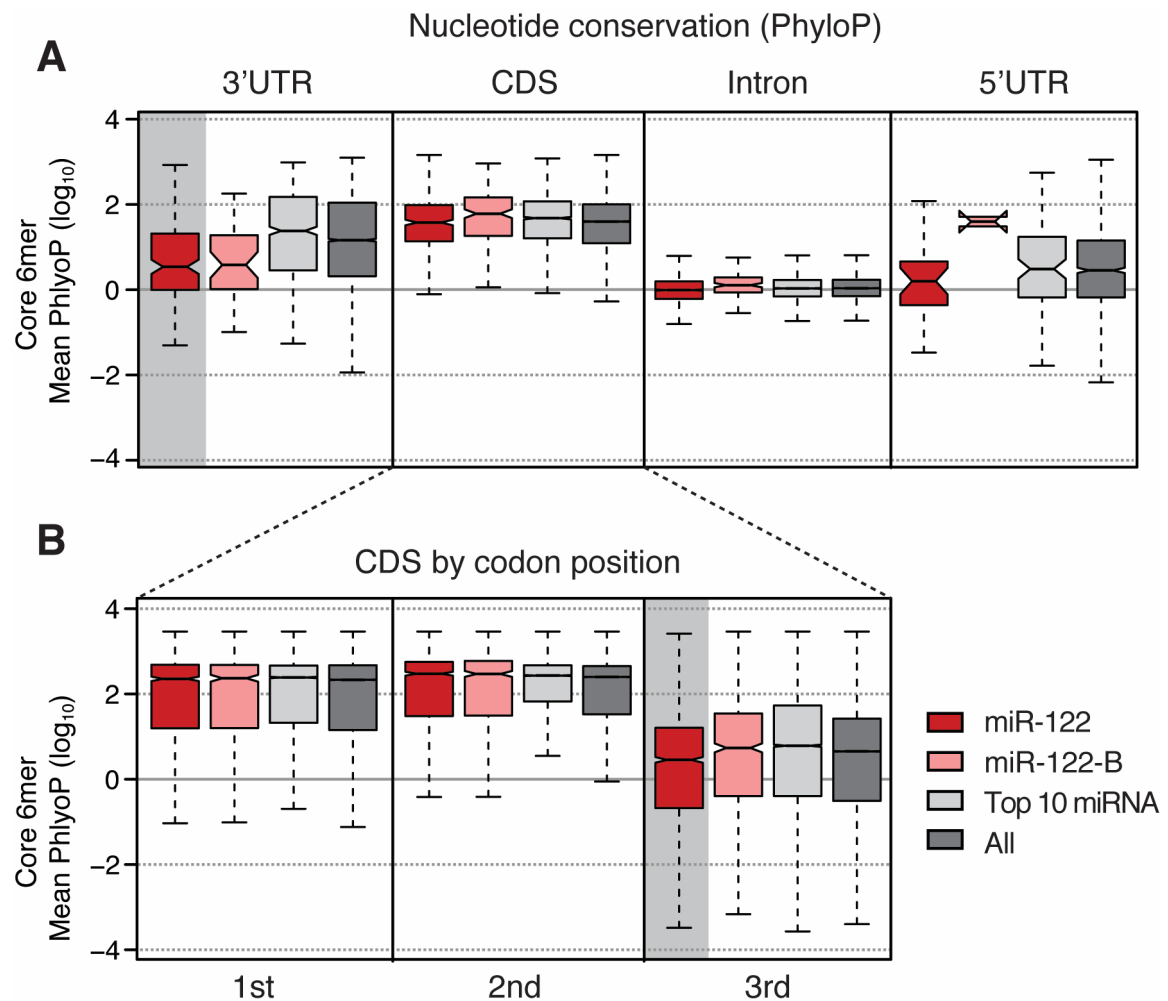


Figure 5.19. miR-122 CLIP targets are not well conserved in human. **(A)** Mean nucleotide conservation scores for human miR-122, bulged miR-122, the top 10 miRNAs and all miRNA targets, denoted by genic region. PhyloP scores from placental mammals were averaged across the core miRNA 6mer and plotted. Shaded areas indicate statistical significance ($P < 10e-10$, KS-test). **(B)** Per nucleotide PhyloP scores calculated as in (A) were further subdivided by codon position for coding exons. Shaded areas indicate statistical significance ($P < 10e-10$, KS-test).

Intersecting miR-122 target networks between mouse and human

Considering how well the miR-122 KO mouse appears to phenocopy the slow and chronic progression to HCC in humans, its reasonable to propose that miR-122 targets shared between mouse and human might be key drivers of the HCC phenotype. In this context, the high species specificity exhibited by miR-122 target regulation between mouse and human can be seen as more of a feature than a bug: by intersecting these disparate datasets and focusing on common targets, we can filter out species specific noise to settle on a core set of conserved factors that presumably play the biggest roles in HCC. Of course, the alternative hypothesis, that much of the species specificity itself may drive pathogenesis is not without weight. This is especially true given the HCC cancer phenotype which strikes mice well past the age of sexual maturity, thus envisioning how natural selection acts to maintain miR-122 targets in this scenario is highly problematic, absent additional factors that affect individual fitness. This point will be addressed further in the discussion. For our purposes, we should simply consider the species specificity of miR-122 targets in the broader context of the multi-factorial nature of HCC. Or to use an analogy, let's say there are 1000 ways to get liver cancer. Far from miR-122 deregulation making that total 1001, the species specificity might make it more like 2001.

As of the writing of this chapter (mid-March 2015), RNAseq in matched normal and tumor samples has not yet been performed. So for the ensuing analysis, miR-122 target CLIP comparisons have been made with all transcripts

de-repressed in miR-122 KO livers. Irrespective of genic region, of the ~3900 genomic loci harboring canonical and G-bulged miR-122 sites in mice roughly ~15% (613) possessed any overlap with human targets in liver (Figure 5.20A). The human targetome, smaller than the mouse due to lower sequencing depth, overlapped considerably with all mouse targets. Focusing on the set of targets de-repressed in mice, only 76 targets (2% of mouse and 7% of human) were common among these datasets.

This theme is repeated when we zoom in on 3'UTR targets. Of the ~1200 miR-122 3'UTR targets in mice, only 29 (or ~2%) were both de-repressed in mouse livers and found in human (Figure 5.20B). Focusing on miR-122 targets identified as HCV de-repressed in Huh-7.5 cells further reduces the overlap to 14 targets with an additional 18 in common with de-repressed mouse targets (likely what remains to be found if human livers were sequenced deeper) (Figure 5.20C).

It must be noted that even of the 29 targets in common in Figure 5.20B, nearly half can be considered co-incidental: miR-122 sites are not conserved between human and mouse paralogs. Thus in these targets, miR-122 regulation appears to have evolved independently. Such an example is shown in Figure 5.21. This transcript encodes an alkaline phosphatase of unknown function that is expressed in liver, bone and kidney. In both mouse and human it clearly possesses AGO bound miR-122 sites, however the primary sequence and precise location relative to the stop codon are not conserved.

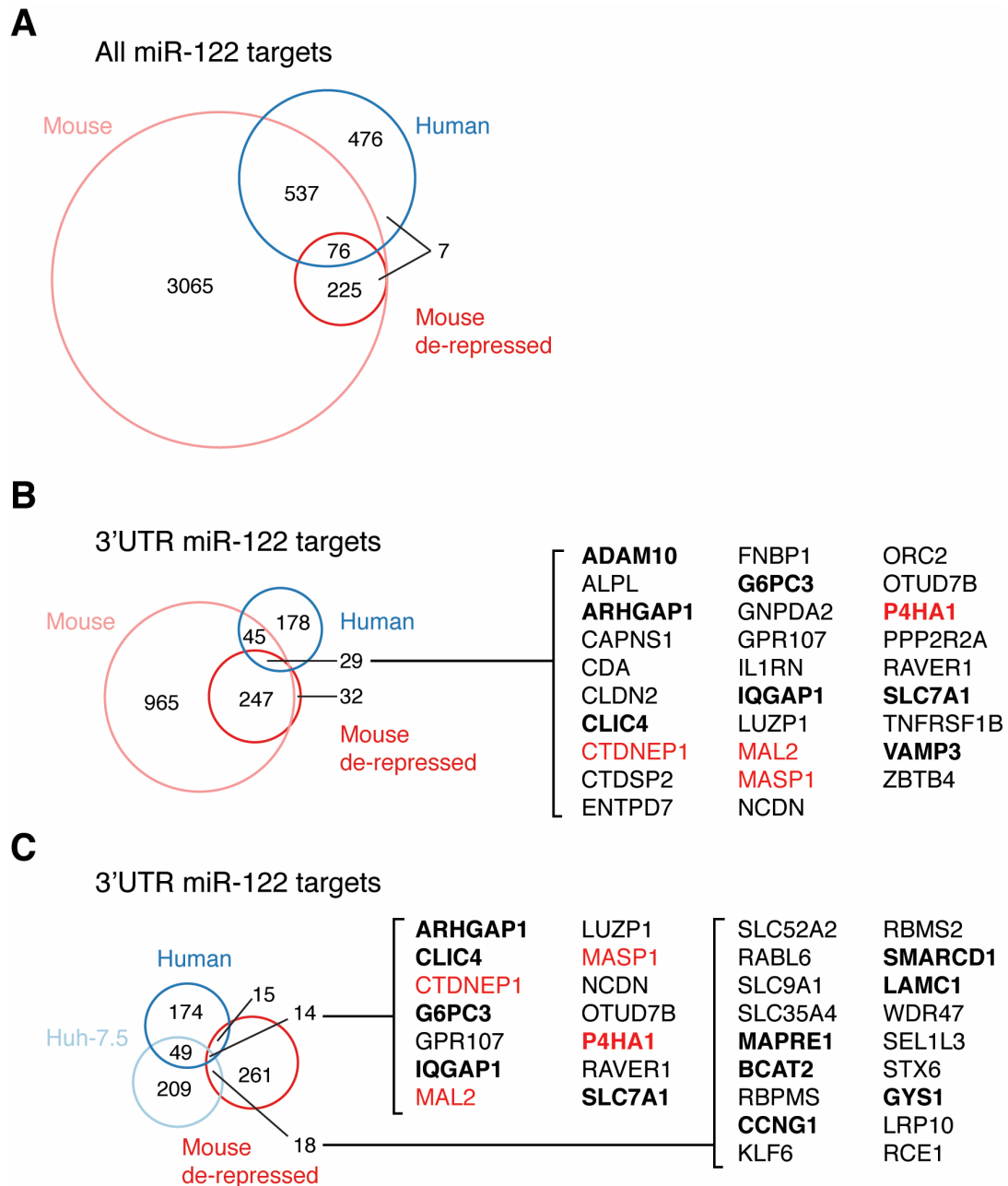


Figure 5.20. miR-122 target overlaps between human and mouse. **(A)** Proportional Venn diagram showing overlap of all CLIP identified canonical and bulged miR-122 targets in mouse or human liver, compared to all de-repressed miR-122 targets measured by RNAseq. **(B)** Diagram as in (A) focusing only on 3'UTR targets. Genes found in all three categories are listed at right, in bold are any targets with an previous confirmatory data, in red are targets validated using either luciferase or single cell reporters from this thesis. **(C)** Diagram as in (B) displaying overlap with all CLIP derived miR-122 targets from Huh-7.5 cells.

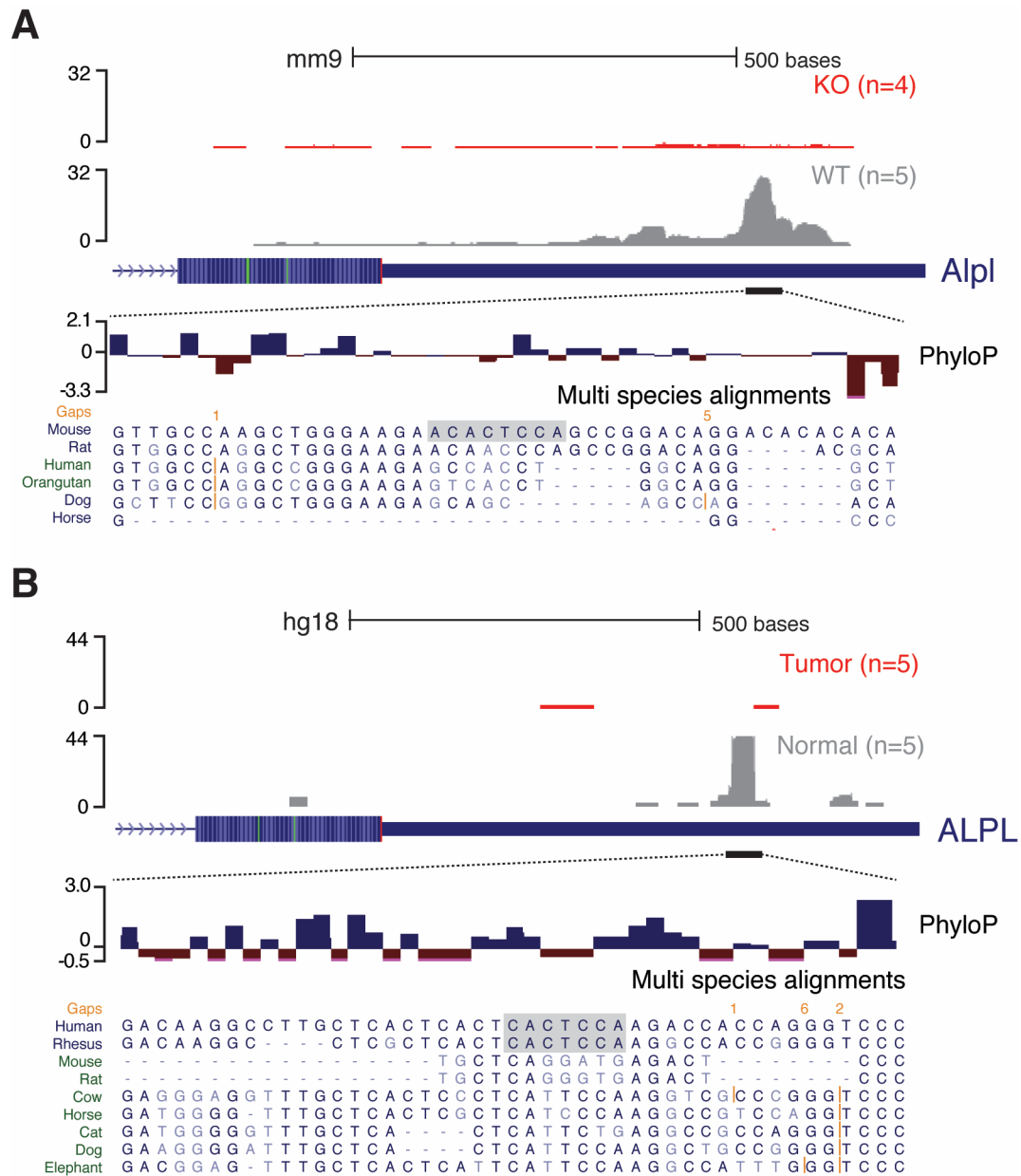


Figure 5.21. miR-122 targeting ALPL 3'UTR is conserved while the binding site is not. AGO-CLIP data in mouse miR-122 or WT mouse liver (**A**) or human normal and tumor tissue (**B**). Zoom in view displays nucleotide conservation scores (PhyloP) and multi species alignments for indicated regions. miR-122 seed highlighted in gray.

Of the number of genes that are conserved between human and mouse and that are derepressed in mouse KO livers, one novel miR-122 target stands out. Interleukin-1 receptor antagonist protein (IL1RN) is a well studied and potent negative regulator of IL-1 β , a key pro-inflammatory signaling molecule (Gabay et al., 2010). Deficiency in IL1RN results excessive IL-1 β signaling to produce acute and chronic inflammation in a variety of animal models and human disease contexts (Aksentijevich et al., 2009; Hirsch et al., 1996; Nicklin et al., 2000; Reddy et al., 2009). IL-1 β has also been implicated as an important mediator in liver pathology, where increased IL-1 β activity has been observed from patients with chronic liver diseases to fulminant hepatic failure (Sekiyama et al., 1994; Tilg et al., 1992).

IL1RN is a clear miR-122 target in both mouse and humans, with additionally conserved supplementary pairing (Figure 5.22). This identification of IL1RN as a miR-122 target *in vivo* raises numerous hypotheses. Conceivably, miR-122 could normally play a role in permitting IL-1 β signaling by regulating IL1RN. Should an IL-1 β driven inflammatory response overshoot, and reduce miR-122 levels as is typically seen in many liver diseases (see Introduction), derepressed IL1RN may work to dampen the IL-1 β cascade. This feedback loop would make sense in the context of an acute HCV infection as a key axis to silence by potentially sequestering miR-122. As HCV replicates and induces an inflammatory response, reducing miR-122 would derepress IL1RN to suppress this exact immune response.

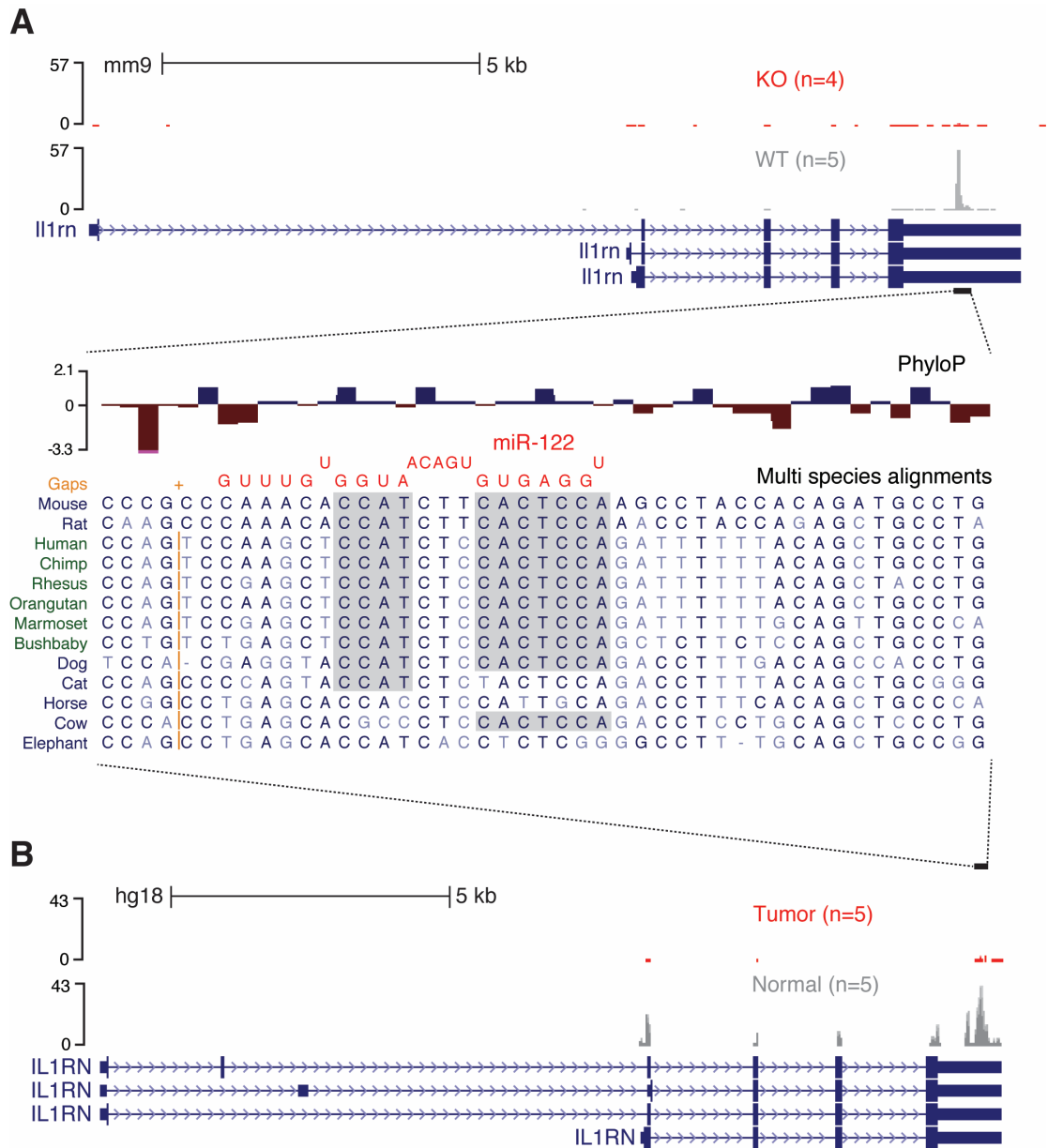


Figure 5.22. Conserved miR-122 targeting of the IL1RN 3'UTR in murine and human liver. AGO-CLIP data in mouse miR-122 or WT mouse liver (**A**) or human normal and tumor tissue (**B**). Zoom in view displays nucleotide conservation scores (PhyloP) and multi species alignments for indicated regions. miR-122 is highlighted in red, extensively conserved seed waiting highlighted in gray.

In accordance with this, increased IL1RN (and IL-1B) expression has been observed in patients with acute and chronic active HCV infections (Gramantieri et

al., 1999; Libra et al., 2006). As an added wrinkle, IL1RN has also been identified as interferon induced in the context of HCV therapy (Cotler et al., 2002). And so we conclude with the following testable scenario, should HCV prevent the activation of an endogenous IFN response via NS3/4A inactivation of RIG-I signaling, derepression of a negative regulator of inflammation (IL1RN) by miRNA sequestration might be means to bypass IFN induction as suppress an inflammatory response.

The above hypothesis is one of many that can be made for the common targets shared between humans and mice as we begin to consider additional roles for miR-122 as a positive regulator for HCV proliferation *in vivo*. Clearly, much work remains to be done on this front.

Chapter 5 summary and discussion

In this chapter, we've explored efforts to elucidate the miR-122 targetome *in vivo* using a miR-122 knockout mouse model and human liver tissue. We've uncovered both non-canonical G-bulged miR-122 sites and revealed widespread miR-122 dependent loci, well beyond the normal 3'UTR sites of miRNA function. Furthermore, we've revealed surprisingly low conservation of targets for such a highly conserved miRNA. In this section, we'll discuss how this might have come to be, and what we might expect from the miR-122 targetome if its deregulation contributes to a similar liver pathology in mice and humans despite regulating different target sets.

Redefining miRNA function: the Seitz hypothesis and target conservation

In 2009, Hervé Seitz pointed out that miRNAs present at least three apparent observational paradoxes (Seitz, 2009). The first paradox is how exactly a miRNA can have hundreds of targets, and yet many *in vivo* phenotypes of specific miRNA loss can be rescued by repressing a few, and often only single targets. Indeed, the developmental phenotypes for the pioneering miRNAs in *C. elegans*, *lin-4* and *let-7*, can each be rescued by introducing complementary seed mutations in their respective targets *lin-14* and *lin-41* (Ambros, 1989; Slack et al., 2000). One well-regarded explanation for this discrepancy is that many *in vivo* experiments tend to overlook, or not follow up, on subtle phenotypic changes associated with other targets. This dovetails with another view that many miRNA loss-of-function phenotypes are context specific, and may not exist basally but require something to induce them. Indeed, a current evaluation of *in vivo* miRNA knockout studies agrees with these views, particularly in light of the fact that many miRNA knockouts possess no obvious phenotypes, often despite ample overexpression data (Vidigal and Ventura, 2015). On its face, this should strike fear in the heart of any miRNA researcher who relies on overexpressed miRNAs to repress luciferase reporters as a surrogate for miRNA target validation. But this is perhaps a separate discussion.

Compounding the miRNA:target discrepancy is the surprisingly modest repression exerted by miRNAs. This is the second major paradox. From a virologist's perspective, miRNA repression of target protein expression by less

than 2-fold is a pittance compared to the orders of magnitude of RNA replication and transcriptional regulation observed in viral infections and their counteracting innate immune responses. Seitz draws a more reasonable parallel by considering that 2-fold effects are well within intraindividual variations in gene expression and transcriptional noise (Seitz, 2009). Moreover, that only a fraction of human genes display haplo-insufficient phenotypes would only further support this argument (Huang et al., 2010). Fitting weakly repressive miRNAs within a complex and robust transcriptional control system that is largely buffered from fluctuations in gene expression is thus at odds with the clear importance of miRNA function.

To resolve these contradictions, Seitz proposed a redefinition of terms. Instead of positing that the primary function of miRNA is to repress mRNA by binding to seed sites, we should consider the obverse, that the function of mRNA seed targeting is to regulate miRNA activity. In other words, most mRNAs primarily function as miRNA sponges where the 2-fold repression is inconsequential for most targets except for the select few that are contextually very sensitive to gene dosage. This explanation provides a framework to resolve the both of the above paradoxes by predicting that most miRNA targets aren't functional in the sense that their repression has downstream consequences. Rather, they function more to set and continuously alter the effective miRNA concentration where, and this is the key, each targeted transcript responds *uniquely* but not independently of other targets. This interpretation is remarkably consistent at all scales of miRNA function: from thresholding effects on individual

transcripts thoroughly described and tested earlier (Mukherji et al., 2011), to *in vivo* phenotypes that appear only in specific contexts, such as genetic or environmental stress (Vidigal and Ventura, 2015).

For Seitz, the miRNA sponge or "pseudotarget hypothesis" also solves a third paradox: while many miRNA targets are conserved among closely related species, they differ considerably between distant species, even for highly conserved miRNAs (Chen and Rajewsky, 2006). This is essentially what we've confirmed for miR-122 targets between humans and mice, in that such targets are poorly conserved across these species, but often are retained in close neighbors such as macaques and rats, respectively. The pseudotarget hypothesis predicts such poor conservation among distantly related organisms precisely because repression is irrelevant for most targets. During evolution, the precise targetomes of a miRNA are free to diverge so long as pseudotargets are replaced with mRNAs with similar expression patterns (Seitz, 2009), or alternatively, by evolving exquisite control of miRNA transcription and degradation. As a consequence, this hypothesis strongly argues that target site conservation is a key feature of "real" targets.

Our results with miR-122 in liver would agree with this view, given that the vast majority of targets, while AGO associated, do not result in substantive repression. Still, it's possible that for many putative pseudotargets the right context for a visible phenotype remains elusive, but conceptually exists. Note also that the above discussion primarily revolves around conserved targeting in

3'UTRs. Our CLIP data alongside other work would contest this narrow framework, given that AGO binding occurs well beyond 3'UTRs, can be conserved in these regions, and does not often behave according to strict seed pairing rules (Helwak et al., 2013; Loeb et al., 2012). Moreover, the demonstration of many species specific miRNA phenotypes in a variety of systems suggests that the pseudotarget hypothesis can be a source of neo-functionalization for miRNA binding (Vidigal and Ventura, 2015).

The broadened framework originally proposed by Seitz has in large part underwritten the competing endogenous RNA (ceRNA) hypothesis as its more formalized and highly debated extension (Denzler et al., 2014; Ebert and Sharp, 2010; 2012; Figliuzzi et al., 2013; Jens and Rajewsky, 2014; Salmena et al., 2011). Critics rightfully point out that relatively minor fluctuations of individual target transcripts are unlikely to perturb global miRNA function, assuming that miRNA levels are in vast excess to the overall target abundance (Jens and Rajewsky, 2014). For highly abundant miRNAs such as miR-122, only when target abundance approaches or exceeds the miRNA expression level are ceRNA effects observed (Denzler et al., 2014). Nevertheless, as researchers are only beginning to appreciate the extent of AGO targeting outside of 3'UTRs and the prevalence of non-canonical binding, the general assumption that miRNA levels are in vast excess to target abundance will require revisiting.

The miR-122 knockout molecular phenotype and HCV: a proposed evolutionary connection

So where does the above discussion leave our results with miR-122 and HCV? I propose that the expansion of miR-122 bound loci defined by AGO-CLIP, combined with the relatively poor conservation observed for miR-122 target seed sites (as pseudo targets) may bear directly on two primary and unique features of this miRNA: its high abundance and its coercion by HCV.

At the outset, miR-122 is both an outlier for being highly liver specific and also the most abundant miRNA in any tissue measured, especially in humans (Ladewig et al., 2012; Lagos-Quintana et al., 2002). A sufficient explanation for these observations has yet to emerge. As mentioned earlier, the miR-122 cancer phenotype occurs after sexual maturity and KO mice display no apparent problems in breeding (K. Ghoshal, personal communication), unlike other miRNAs with critical roles in development or cellular differentiation (de Pontual et al., 2011; Rodriguez et al., 2007; Vidigal and Ventura, 2015). Assuming that the reproductive fitness of miR-122 KO mice is only mildly affected by miR-122 loss, it is difficult to fathom how a miR-122 target network could *not* undergo genetic drift (Lynch and Hagner, 2015). Consider a small core set of targets that are sensitive to ~2-fold repression, among a pool of pseudo-targets which are insensitive. Absent strongly deleterious acquisition of miR-122 binding which is cleared through selective avoidance (Bartel and Chen, 2004), most pseudo-targets and even non-targets are free to acquire miR-122 sites with no

deleterious effects. This would explain the the divergence of miR-122 targetomes between humans and mice and their relative evolutionary proximity to closely related species. Furthermore, this could also serve as the engine for neofunctionalization which would be expected from such "evolutionary meandering" of intermolecular interactions (Lynch and Hagner, 2015). Even among highly conserved miRNAs, such a mechanism could also account for the emergence of species specific acquisition of sites that may then undergo positive selection as part of the neofunctionalization process, as has been observed in humans (Li et al., 2012; Saunders et al., 2007).

Now, one could imagine a scenario where at some point the expanded miR-122 targetome could negatively impact those core set of targets that must be repressed, due to the overall target abundance becoming inhibitory to repression. As a result, the hepatocyte may evolve a means to increase miR-122 abundance, and the cycle repeats. Notably, absolute miR-122 abundance per cell is well known in mouse (~66,000 copies / cell) and human, which contains nearly twice the amount of miR-122 (~135,000 copies per cell) (Chang et al., 2004; Denzler et al., 2014).

Overall, this hypothesis makes at least four testable predictions: that miR-122 level should scale with its absolute target abundance across various organisms (measured by CLIP and not by informatic prediction), that targetomes should diverge between distant organisms (as is predicted in vertebrates (Xu et

al., 2013)), and that miR-122 promoter complexity as well as levels per hepatocyte should scale with organismal complexity.

If we consider HCV infection in this context, one reasonable argument for its acquisition of miR-122 dependence as a protective shield is due to the simple fact that miR-122 is both highly abundant in liver, but mostly irrelevant for most targets. In other words, HCV has evolved to use miR-122 precisely because it provides an unexploited niche relative to other miRNAs. At first glance, this appears to resemble a null hypothesis for any pro-viral role for the miR-122. However this need not be the case, especially if the substantial and species specific targeting in humans represents the right mix of targets that would make sense to de-repress. For conserved targets such as IL1RN (Figure 5.22), and non-conserved targets such as Drosha (Figure 5.9B), it is straightforward to envision pro-viral paths for the HCV sponge in humans.

All of this is to say that many surprises are likely to continue to be in store for the myriad of roles that miR-122 plays in the livers of vertebrates. Expanding upon our results to test specific predictions in other organisms and model systems, using novel HCV variants (such as the U3 and miR-15 dependent viruses), putative murine and equine cousins to HCV, and empowered by unbiased genomic methods such as CLIP, we are poised to unravel the complex host-viral interaction of information that is centered upon one 22 nucleotide snippet of RNA, life's indispensable molecule.

Chapter 6: Concluding remarks

In this very short chapter, I aim less to provide a comprehensive model of the work presented in this thesis than to highlight the unanswered questions this work raises and point out viable paths forward.

In succinct terms, the following model is proposed that links HCV replication in the infected hepatocyte to outward manifestations of HCV induced disease. Mechanistically, miR-122 is posited as a central molecular player in this process. In normal hepatocytes, miR-122 target repression is required to maintain long-term liver homeostasis. Upon HCV infection, viral RNA coercion of miR-122 results in de-repression of host miR-122 targets (Figure 6.1). Due to this miRNA sponging by HCV RNA, liver homeostasis is perturbed and can result in a constellation of liver complications to include HCC. Indirect pro- or anti-viral benefits from the miR-122 sponge are possible, but currently unknown.

At the outset, the miR-122:HCV sponge relies on a two part assumption: constant miR-122 levels and sufficiently high enough viral replication to impact the effective miR-122 concentration in the cell. As shown in this thesis and addressed in previous chapter discussions, mature miR-122 levels appear relatively constant *in vitro* and *in vivo*. Capturing the dynamic nature of cytoplasmic HCV replication that underlies long-term chronicity in the liver is the next major challenge. While the sponge effect described in this work can be thoroughly tested in cell culture systems, its ultimate validation *in vivo* will require appropriate model systems and technologies aimed at interrogating single cells.

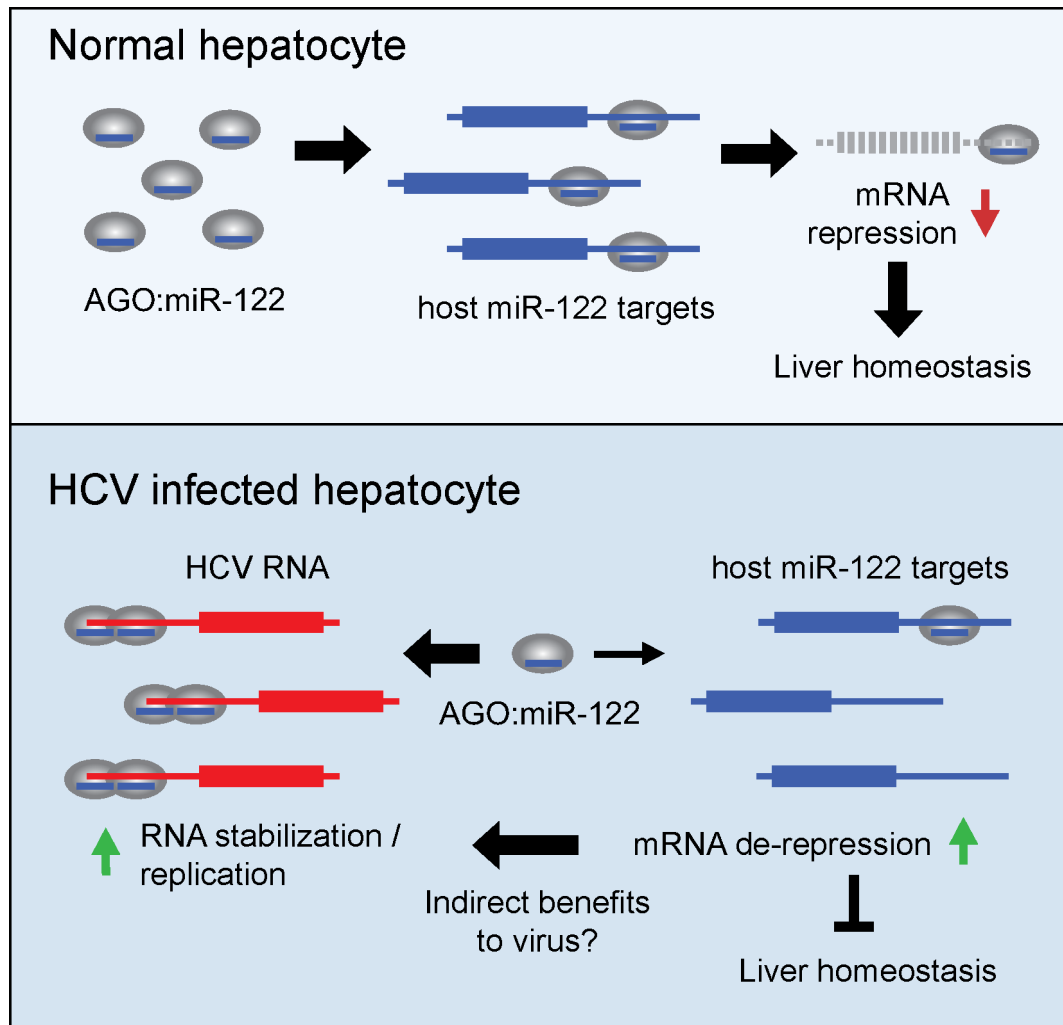


Figure 6.1. A concise model for HCV:miR-122 sponge induced liver disease. In normal hepatocytes, miR-122 target repression is required to maintain long-term liver homeostasis. Upon HCV infection, viral RNA coercion of miR-122 results in de-repression of host miR-122 targets, which in turn de-stabilizes liver homeostasis. Indirect benefits (or road-blocks) to the virus as a result of miR-122 target de-repression are possible but currently unknown. This figure adapted from (Luna et al., 2015).

To this end, recently described hepaciviruses of rodents, as close homologs to HCV, might offer long sought after disease-causing surrogates for HCV infection. That these viruses have miR-122 sites in their 5'UTRs is highly

suggestive of their need for this miRNA and may point to potential sponge effects.

But given that our data also pinpoint high species specificity to miR-122 targeting between mice and humans, its possible that distinct human specific miR-122 targets, when de-repressed, might be key drivers for human disease. To this end, the application of single cell transcriptomic analysis to HCV infected human liver cells represents an alternative path forward, should murine hepaciviruses fail to phenocopy human disease. Additionally, the *in vitro* Huh-7.5 cell context, where the sponge effect was discovered, can continue to play a role by exploring the stoichiometric requirements for a viral RNA to sponge miRNA. Testing other HCV genotypes with different replicative potentials, testing heterozygous miR-122 knockout mutant cells, and probing the sponge effect induced by the m15 virus offer proximal means to cement our observations while providing a roadmap for investigations of other RNA viruses.

In sum, the genomic portrait I've outlined for HCV and miR-122 in this thesis is exactly that: a single perspective on the multifaceted relationship between host and virus. As biology matures in the post-genomic era, and static reductionist pictures yield to data-driven dynamism, the stage is set for further systems level views of viruses. Today as ever, Pasteur's words remain as relevant as they did in the 19th century: "Messieurs, c'est les microbes qui auront le dernier mot." (Gentlemen, it is the microbes who will have the last word).

APPENDIX of OLIGOS

Oligo name	Sequence (5' to 3')
miR-15 dependent HCV construct cloning	
Pos43_F	CCTGTGAGGAACTACTGTCTTC
Pos12343_F	CGGCCAGTGAATTCTAATACGACTCA
Pos1293_R	CGGTGTCCAGTGATGGTACCAG
Pos1358_R	GCATCGCGTACGCCAAGATC
m122to15_F	CAGTGAATTCTAATACGACTCACTATAGACCTGCCC CTAATAGGGGCGTGCTGCTGCCATGAATGCTGCTCC TGTGAGGAACTACTGTC
m122toU3_F	CAGTGAATTCTAATACGACTCACTATAGACCTATACT TTCAGGGATCATTTCTATAGTGTGTTACTAGAGAAGT TTCTCTGAACGTGTAGAGCACCGATTGAATCACTCC CCTGTGAGGAACTACTGTCTTC
F S2p34	GATTGAATCACAGCCCTGTGAG
R S2p34	CTCACAGGGCTGTGATTCAATC
F S2p3	GATTGAATCACTGCCCTGTGAG
R S2p3	CTCACAGGGCAGTGATTCAATC
Luciferase reporter vector cloning	
miR-122_seed_F	ATATCTCGAGTCTAGCCACATGACACTCCATATGCG GCCGCATTA
miR-122_seed_R	TAATGCGGCCGCATATGGAGTGTCATGTGGCTAGAC TCGAGATAT
miR-122_seedmut_F	TCGAGTCTAGCCACATGACACAGCATATGC
miR-122_seedmut_R	GGCCGCATATGCTGTGTCATGTGGCTAGAC
miR-17_4xseed_F	/5Phos/TCGAGTCAAGGTCAGCTGCACTTTAGTACAG TCTGTCATGCACTTTATAACGGCCCCGCTTGCACTT TACAATCAGCAACCTTGCACTTTATATGC
miR-17_4xseed_R	/5Phos/GGCCGCATATAAAGTGCAAGGTTGCTGATTG TAAAGTGCAAGCGGGGCCGTTATAAAGTGCAATGACA GACTGTACTAAAGTGCAAGCTGACCTTGAC
miR-17_4xseed_seedmut_F	/5Phos/TCGAGTCAAGGTCAGCTGCACaaTAGTACAG TCTGTCATGCACaaTATAACGGCCCCGCTTGCAACaaT ACAATCAGCAACCTTGCAACaaTATATGC
miR-17_4xseed_seedmut_R	/5Phos/GGCCGCATATATTGTGCAAGGTTGCTGATTG TATTGTGCAAGCGGGGCCGTTATATTGTGCATGACA GACTGTACTATTGTGCAGCTGACCTTGAC
PKM2_3UTR_F1881	ATATCTCGAGTGGACCCCAGAGCCCCTCC

PKM2_3UTR_R2 501	TAAT <u>GCGGCCGCG</u> GTGCTTCAGCTGTTGTTTATTG
PKM2_seedmut_ F	GCAGCAAACACAGCACCCCTCCACC
PKM2_seedmut_ R	GGTGGAGGGTGCTGTGTTTGCTGC
CTDNEP1_3UT R_F1183	ATATCTCGAGCAGCTGCTCCCCCTCCACCTG
CTDNEP1_3UT R_R1528	TAAT <u>GCGGCCGCT</u> GACTTGGCTTCTCTTTGAGC
CTDNEP1_seed mut_F	CGTGTTACACAGCATGGAAACC
CTDNEP1_seed mut_R	GGTTTCCATGCTGTGTGAACACG
SFT2D1_3UTR_ F511	ATATCTCGAGAAATCAGAACTTGTGGAAAAGAG
SFT2D1_3UTR_ R699	TAAT <u>GCGGCCGCT</u> TTAAAGTTACAAGCATTTAATGG
SFT2D1_seedm ut_F	CCATAAAACACAGCAGGAACAAC
SFT2D1_seedm ut_R	GTTGTTCCCTGCTGTGTTTTATGG
pTre3G fluorescent reporter vector cloning	
EcoRI- nlsTagBFP_F	CGGAATTCACCATGGGCCCTAAAAAGAAGCGTAAAG TCATGAGCGAGCTGATTAAGGAGA
NdeI- nlsTagBFP_R	GGAATTCCATATGTTATCCGGAATTAAGCTTGTGCC C
nlsTagRFP_F	ATGGGCCCTAAAAAGAAGCGTAAAGTCATGGTGTCT AAGGGCGAAGAGC
pT3G-MCS2- BamHI-nls_F	GGGGCCCAGATCTCCGCGGGGATCCACCATGGGCC CTAAAAAGAAGCGTAAAGTC
RFP-ClaI-N=1m- EcoRV-pT3G_R	CTGATTATGATCCTCTGGAGATATCCGGCTCCAGTC TTGTATCCTGTTTCCGGCATCGATTTATCCGGAATTA AGTTTGTGCCCC
RFP-ClaI-N=1- EcoRV-pT3G_R	CTGATTATGATCCTCTGGAGATATCCGGCTGGAGTG TTGTATGGTGTTCGCGCATCGATTTATCCGGAATTA AGTTTGTGCCCC
RFP-ClaI-N=1pf- EcoRV-pT3G_R	CTGATTATGATCCTCTGGAGATATCCGGCTGGAGTG TGACAATGGTGTTCGCGCATCGATTTATCCGGAAT TAAGTTTGTGCCCC
RFP-end_R	TTATCCGGAATTAAGTTTGTGCCCC
RFP-ClaI-N=4- EcoRV-pT3G	GGGGCACAACTTAATTCGGATAAATCGATCGGCA AACACCATACAACACTCCACTAGAAACACCATAAAA

	CACTCCAGATCAAACACCATACAACACTCCAGCATA AACACCATAAAACACTCCACGGCGATATCTCCAGAG GATCATAATCAG
RFP-ClaI-N=6- EcoRV-pT3G	GGGGCACAACTTAATTCCGGATAAATCGATCGGCA AACACCATAACAACACTCCACTAGAAACACCATAAAA CACTCCAGATCAAACACCATACAACACTCCAGCATA AACACCATAAAACACTCCAGTGCAAACACCATACAA CACTCCACGAAAAACACCATAAAACACTCCACGGCG ATATCTCCAGAGGATCATAATCAG
pT3G-CS- 3UTR_F	GGGGCACAACTTAATTCCGGATAAAACTGGAGACT GGGTGAAAGTG
CS-3UTR- BamHI-pT3G_R	CTGATTATGATCCTCTGGAGATATCGGCTCAAGCAT TGTTAATAAAAACATT
pT3G-MASP1- 3UTR_F	GGGGCACAACTTAATTCCGGATAAGCTGACTTACT TCCTCGGGGC
MASP-3UTR- BamHI-pT3G_R	CTGATTATGATCCTCTGGAGATATCTTGGTTATCCAC GAGGGTTTATTTCCAC
pT3G-MAL2- 3UTR_F	GGGGCACAACTTAATTCCGGATAAACTCCTTAGA AACTGGCAG
MAL2-3UTR- BamHI-pT3G_R	CTGATTATGATCCTCTGGAGATATCATTATCAAATGA ATACTTTATTAGAGACA
pT3G-SFT2D1- 3UTR_F	GGGGCACAACTTAATTCCGGATAAAAATCAGAAAC TTGTGGAAAAGAGC
SFT2D1-3UTR- BamHI-pT3G_R	CTGATTATGATCCTCTGGAGATATCTTTAAAGTTACA AGCATTTAATGGTTTAATC
pT3G-PKM2- 3UTR_F	GGGGCACAACTTAATTCCGGATAATGGACCCCAGA GCCCCTCC
PKM2-3UTR- BamHI-pT3G_R	CTGATTATGATCCTCTGGAGATATCGGTGCTTCAGC TGTTGTTTATTGAC
pT3G- CTDNEP1- 3UTR_F	GGGGCACAACTTAATTCCGGATAACAGCTGCTCCC CCTCCACC
CTDNEP1- 3UTR-BamHI- pT3G_R	CTGATTATGATCCTCTGGAGATATCTGACTTGGCTTC TCTTTGAGCCTC
pT3G-P4HA1- 3UTR_F	GGGGCACAACTTAATTCCGGATAACAAACAGGCTT CCCTTTTTCTCCT
P4HA1-3UTR- BamHI-pT3G_R	CTGATTATGATCCTCTGGAGATATCTTAAAAAAGATT TAAGATCATAAATAGGTCATTG
pT3G-AldoA- 3UTR_F	GGGGCACAACTTAATTCCGGATAAGCGGAGGTGTT CCCAGGCTGC
AldoA-3UTR- BamHI-pT3G_R	CTGATTATGATCCTCTGGAGATATCGCCGACTCCCC CTTAAATAGCTGTTTA

CRISPR mutant cloning and sequencing	
122.sgRNA1a	CACCGAGAGCTGTGGAGTGTGACAA
122.sgRNA1b	AAACTTGTCACTCCACAGCTCTC
122.sgRNA2a	CACCGAGGAACTCTGTAGCCACGA
122.sgRNA2b	AAACTCGTGGCTACAGAGTTTCCTC
122Genomic_F	GTTGCAAACAGAGTTCCTGTCC
122Genomic_R	GAGAGGCAGGGTTCAGCTAACC
Drosha.sgRNA1a	CACCGGAACCTCTCTAACAGGGGTT
Drosha.sgRNA1b	AAACAACCCCTGTTAGAGAGTTCC
Drosha.sgRNA2a	CACCGAGGGCATAAAACACGCACC
Drosha.sgRNA2b	AAACGGTGCGTGTTTTTATGCCCT
DroshaGenomic_F	TCATGCCTATAATCCCAGCACTTTAGG
DroshaGenomic_R	CAGTTTCAATACTAGGCAACATGTATCTTCTGAG
Dicer.sgRNAEX1.1a	CACCGAACCCTGCATGATTGTGTAA
Dicer.sgRNAEX1.1b	AAACTTACACAATCATGCAGGGTTC
Dicer.sgRNAEX1.2a	CACCGTTATACGCCAAGAAAATATC
Dicer.sgRNAEX1.2b	AAACGATATTTTCTTGGCGTATAAC
DicerEX1_GenomicF	CACTGGGTAAGGTACAGAATGCTTGAC
DicerEX1_GenomicR	TTAATCAGAAGTGGGAGGCCTGAAAG
Dicer.sgRNAEX2.1a	CACCGCGTATAAAGACGTAGACTGT
Dicer.sgRNAEX2.1b	AAACACAGTCTACGTCTTTATACGC
Dicer.sgRNAEX2.2a	CACCGTATTTTATCAAGTTGTGCA
Dicer.sgRNAEX2.2b	AAACTGCACAACCTTGATAAAATAC
DicerEX2_GenomicF	CATTTCAACACGTATTTACACAATACAGCTG
DicerEX2_GenomicR	CAGAAGCAGAACATTTAAGAGAAACAGC
Dicer.sgRNAEX19.1a	CACCGAACCAGTATCTTCAAGTAA
Dicer.sgRNAEX19.1b	AAACTTACTTGAAGATACTGGTTC

Dicer.sgRNAEX1 9.2a	CACCGTTCTTTCTGGCTGACTGCAC
Dicer.sgRNAEX1 9.2b	AAACGTGCAGTCAGCCAGAAAGAAC
DicerEX19_Geno micF	GAGAGTTTTGTTTTTATACTTGGCCCATT
DicerEX19_Geno micR	AGACTGAAAATGCTGATGCAGTAAC
pX330 and pX458 insert sequencing primer	ACTATCATATGCTTACCGTAAC
Small RNA northern probes	
miR-122-5p_AS	AAACACCATTGTCACACTCCA
miR-21-5p_AS	TCAACATCAGTCTGATAAGCTA
miR-16-5p_AS	CGCCAATATTTACGTGCTGCTA
U6_snRNA_AS	GAATTTGCGTGTATCCTTGCGCAGGGGCCATGCTA A
Poly-G CLIP	
ND7C-IndexGCAT	/5phos/GNNNNNNNNTATGCGATCGTCGGACTGTAGA ACTCT/idSp/CAAGCAGAAGACGGCATAACGACCCCCC CDN
ND7C-IndexGTCA	/5phos/GNNNNNNNNTTGACGATCGTCGGACTGTAGA ACTCT/idSp/CAAGCAGAAGACGGCATAACGACCCCCC CDN
ND7C-IndexACTG	/5phos/GNNNNNNNNTCAGTGATCGTCGGACTGTAGA ACTCT/idSp/CAAGCAGAAGACGGCATAACGACCCCCC CDN
ND7C-IndexAGCT	/5phos/GNNNNNNNNTAGCTGATCGTCGGACTGTAGA ACTCT/idSp/CAAGCAGAAGACGGCATAACGACCCCCC CDN
DSFP5a	AATGATACGGCGACCAACGACAGGTTTCAGAGTTCTA CAGTCCGACG
DSFP3a	CAAGCAGAAGACGGCATA
SSP2	GACAGGTTTCAGAGTTCTACAGTCCGACGATC
Multiplexed "Super-CLIP" primers	
Anchor_MiSeq_CT AG_DP5_1	AATGATACGGCGACCAACGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTCTAGAGGGAGGAC GATGCGG
Anchor_MiSeq_GATC _DP5_2	AATGATACGGCGACCAACGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTGATCAGGGAGGAC GATGCGG
Anchor_MiSeq_CGTA _DP5_3	AATGATACGGCGACCAACGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTCGTAAGGGAGGAC

	GATGCGG
Anchor_MiSeq_ GCAT_DP5_4	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTGCATAGGGAGGAC GATGCGG
Anchor_MiSeq_ GTGC_DP5_5	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTGTGCAGGGAGGAC GATGCGG
Anchor_MiSeq_ ACCG_DP5_6	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTACCGAGGGAGGAC GATGCGG
Anchor_MiSeq_ CATG_DP5_7	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTCATGAGGGAGGAC GATGCGG
Anchor_MiSeq_T GCA_DP5_8	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTTGCAAGGGAGGAC GATGCGG
Anchor_MiSeq_T CAC_DP5_9	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTTCACAGGGAGGAC GATGCGG
Anchor_MiSeq_ AGTG_DP5_10	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTAGTGAGGGAGGAC GATGCGG
Anchor_MiSeq_T ACG_DP5_11	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTTACGAGGGAGGAC GATGCGG
Anchor_MiSeq_ ATGC_DP5_12	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTATGCAGGGAGGAC GATGCGG
Anchor_MiSeq_ CCGA_DP5_13	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTCCGAAGGGAGGAC GATGCGG
Anchor_MiSeq_ CTCC_DP5_14	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTCTCCAGGGAGGAC GATGCGG
Anchor_MiSeq_ GGCT_DP5_15	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTGGCTAGGGAGGAC GATGCGG
Anchor_MiSeq_ GAGG_DP5_16	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTGAGGAGGGAGGAC GATGCGG
Anchor_MiSeq_ AACC_DP5_17	AATGATACGGCGACCACCGAGATCTACACTCTTTCC CTACACGACGCTCTTCCGATCTAACCAGGGAGGAC GATGCGG

MSFP3	CAAGCAGAAGACGGCATACGAGATCCGCTGGAAGT GACTGACAC
-------	--

REFERENCES

- Adinolfi, L.E., Gambardella, M., Andreana, A., Tripodi, M.F., Utili, R., and Ruggiero, G. (2001). Steatosis accelerates the progression of liver damage of chronic hepatitis C patients and correlates with specific HCV genotype and visceral obesity. *Hepatology* 33, 1358–1364.
- Aksentijevich, I., Masters, S.L., Ferguson, P.J., Dancey, P., Frenkel, J., van Royen-Kerkhoff, A., Laxer, R., Tedgård, U., Cowen, E.W., Pham, T.-H., et al. (2009). An autoinflammatory disease with deficiency of the interleukin-1-receptor antagonist. *New England Journal of Medicine* 360, 2426–2437.
- Alter, H.J., Purcell, R.H., Shih, J.W., Melpolder, J.C., Houghton, M., Choo, Q.-L., and Kuo, G. (1989). Detection of Antibody to Hepatitis C Virus in Prospectively Followed Transfusion Recipients with Acute and Chronic Non-A, Non-B Hepatitis. *New England Journal of Medicine* 321, 1494–1500.
- Altuvia, Y., Landgraf, P., Lithwick, G., Elefant, N., Pfeffer, S., Aravin, A., Brownstein, M.J., Tuschl, T., and Margalit, H. (2005). Clustering and conservation patterns of human microRNAs. *Nucleic Acids Res.* 33, 2697–2706.
- Ambros, V. (1989). A hierarchy of regulatory genes controls a larva-to-adult developmental switch in *C. elegans*. *Cell* 57, 49–57.
- Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in *C-elegans*. *Curr Biol* 13, 807–818.
- Amitsur, M., Levitz, R., and Kaufmann, G. (1987). Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine tRNA. *Embo J* 6, 2499–2503.
- Andre, P., Komurian-Pradel, F., Deforges, S., Perret, M., Berland, J.L., Sodoyer, M., Pol, S., Bréchet, C., Paranhos-Baccala, G., and Lotteau, V. (2002). Characterization of low- and very-low-density hepatitis C virus RNA-containing particles. *Journal of Virology* 76, 6919–6928.
- Aravin, A.A., Hannon, G.J., and Brennecke, J. (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 318, 761–764.
- Ariumi, Y., Kuroki, M., Kushima, Y., Osugi, K., Hijikata, M., Maki, M., Ikeda, M., and Kato, N. (2011). Hepatitis C virus hijacks P-body and stress granule components around lipid droplets. *Journal of Virology* 85, 6882–6892.

- Bai, S., Nasser, M.W., Wang, B., Hsu, S.-H., Datta, J., Kutay, H., Yadav, A., Nuovo, G., Kumar, P., and Ghoshal, K. (2009). MicroRNA-122 inhibits tumorigenic properties of hepatocellular carcinoma cells and sensitizes these cells to sorafenib. *J. Biol. Chem.* *284*, 32015–32027.
- Bala, S., Petrasek, J., Mundkur, S., Catalano, D., Levin, I., Ward, J., Alao, H., Kodys, K., and Szabo, G. (2012). Circulating microRNAs in exosomes indicate hepatocyte injury and inflammation in alcoholic, drug-induced, and inflammatory liver diseases. *Hepatology* *56*, 1946–1957.
- Bandapalli, O.R., Ehrmann, F., Ehemann, V., Gaida, M., Macher-Goeppinger, S., Wente, M., Schirmacher, P., and Brand, K. (2012). Down-regulation of CXCL1 inhibits tumor growth in colorectal liver metastasis. *Cytokine* *57*, 46–53.
- Barba, G., Harper, F., Harada, T., Kohara, M., Goulinet, S., Matsuura, Y., Eder, G., Schaff, Z., Chapman, M.J., MIYAMURA, T., et al. (1997). Hepatitis C virus core protein shows a cytoplasmic localization and associates to cellular lipid storage droplets. *Proc Natl Acad Sci USA* *94*, 1200–1205.
- Bartel, D.P., and Chen, C.Z. (2004). Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat Rev Genet* *5*, 396–400.
- Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* *116*, 281–297.
- Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* *136*, 215–233.
- Bartenschlager, R., Penin, F., Lohmann, V., and Andre, P. (2011). Assembly of infectious hepatitis C virus particles. *Trends in Microbiology* *19*, 95–103.
- Baulcombe, D. (2004). RNA silencing in plants. *Nature* *431*, 356.
- Beard, M.R., and Helbig, K.J. (2008). Control of HCV replication: when size does not matter. *Hepatology* *47*, 1092–1094.
- Bellecave, P., Sarasin-Filipowicz, M., Donze, O., Kennel, A., Gouttenoire, J., Meylan, E., Terracciano, L., Tschopp, J., Sarrazin, C., Berg, T., et al. (2010). Cleavage of Mitochondrial Antiviral Signaling Protein in the Liver of Patients with Chronic Hepatitis C Correlates with a Reduced Activation of the Endogenous Interferon System. *Hepatology* *51*, 1127–1136.
- Bennasser, Y., Le, S.-Y., Benkirane, M., and Jeang, K.-T. (2005). Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing. *Immunity* *22*, 607–619.

- Berezhna, S.Y.S., Supekova, L.L., Sever, M.J.M., Schultz, P.G.P., and Deniz, A.A.A. (2011). Dual regulation of hepatitis C viral RNA by cellular RNAi requires partitioning of Ago2 to lipid droplets and P-bodies. *RNA* *17*, 1831–1845.
- Berger, K.L., Cooper, J.D., Heaton, N.S., Yoon, R., Oakland, T.E., Jordan, T.X., Mateu, G., Grakoui, A., and Randall, G. (2009). Roles for endocytic trafficking and phosphatidylinositol 4-kinase III alpha in hepatitis C virus replication. *Proceedings of the National Academy of Sciences* *106*, 7582.
- Beutler, B., Eidenschenk, C., Crozat, K., Imler, J.-L., Takeuchi, O., Hoffmann, J.A., and Akira, S. (2007). Genetic analysis of resistance to viral infection. *Nat Rev Immunol* *7*, 753–766.
- Bialecki, E., and Di Bisceglie, A. (2005). Diagnosis of hepatocellular carcinoma. *HPB: Official Journal of the International Hepato Pancreato Biliary Association* *7*, 26–34.
- Biggins, S.W., Bambha, K.M., Terrault, N.A., Inadomi, J., Shiboski, S., Dodge, J.L., Gralla, J., Rosen, H.R., and Roberts, J.P. (2012). Projected future increase in aging hepatitis C virus-infected liver transplant candidates: A potential effect of hepatocellular carcinoma. *Liver Transpl* *18*, 1471–1478.
- Blight, K.J., McKeating, J.A., and Rice, C.M. (2002). Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. *Journal of Virology* *76*, 13001–13014.
- Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* *10*, 185–191.
- Bondy-Denomy, J., and Davidson, A.R. (2014). To acquire or resist: the complex biological effects of CRISPR-Cas systems. *Trends in Microbiology* *22*, 218–225.
- Brimacombe, R., Stiege, W., Kyriatsoulis, A., and Maly, P. (1988). Intra-RNA and RNA-protein cross-linking techniques in *Escherichia coli* ribosomes. *Meth Enzymol* *164*, 287–309.
- Brown, M., Suryawanshi, H., Hafner, M., Farazi, T.A., and Tuschl, T. (2013). Mammalian miRNA curation through next-generation sequencing. *Front Genet* *4*, 145.
- Brown, V., Jin, P., Ceman, S., Darnell, J.C., O'Donnell, W.T., Tenenbaum, S.A., Jin, X., Feng, Y., Wilkinson, K.D., Keene, J.D., et al. (2001). Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* *107*, 477–487.

Bruno, S., Silini, E., Crosignani, A., Borzio, F., Leandro, G., Bono, F., Asti, M., Rossi, S., Larghi, A., Cerino, A., et al. (1997). Hepatitis C virus genotypes and risk of hepatocellular carcinoma in cirrhosis: A prospective study. *Hepatology* 25, 754–758.

Bueno, M.J., and Malumbres, M. (2011). MicroRNAs and the cell cycle. *Biochim Biophys Acta* 1812, 592–601.

Bukh, J., Purcell, R.H., and Miller, R.H. (1993). At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. *Proc Natl Acad Sci USA* 90, 8234–8238.

Burbelo, P.D., Dubovi, E.J., Simmonds, P., Medina, J.L., Henriquez, J.A., Mishra, N., Wagner, J., Tokarz, R., Cullen, J.M., Iadarola, M.J., et al. (2012). Serology-enabled discovery of genetically diverse hepaciviruses in a new host. *Journal of Virology* 86, 6171–6178.

Burgess, D.J. (2013). Small RNAs: antiviral RNAi in mammals. *Nat Rev Genet* 14, 821–821.

Burns, D.M., D'Ambrogio, A., Nottrott, S., and Richter, J.D. (2011). CPEB and two poly(A) polymerases control miR-122 stability and p53 mRNA translation. *Nature* 473, 105–U125.

Camus, G., Herker, E., Modi, A.A., Haas, J.T., Ramage, H.R., Farese, R.V., and Ott, M. (2013). Diacylglycerol acyltransferase-1 localizes hepatitis C virus NS5A protein to lipid droplets and enhances NS5A interaction with the viral capsid core. *Journal of Biological Chemistry* 288, 9915–9923.

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

Castoldi, M., and Muckenthaler, M.U. (2012). Regulation of iron homeostasis by microRNAs. *Cell Mol Life Sci* 69, 3945–3952.

Castoldi, M.M., Spasic, M.M.V., Altamura, S.S., Elmén, J.J., Lindow, M.M., Kiss, J.J., Stolte, J.J., Sparla, R.R., D'Alessandro, L.A.L., Klingmüller, U.U., et al. (2011). The liver-specific microRNA miR-122 controls systemic iron homeostasis in mice. *J. Clin. Invest.* 121, 1386–1396.

Catanese, M.T., Loureiro, J., Jones, C.T., Dorner, M., Hahn, von, T., and Rice, C.M. (2013a). Different Requirements for Scavenger Receptor Class B Type I in Hepatitis C Virus Cell-Free versus Cell-to-Cell Transmission. *Journal of Virology* 87, 8282–8293.

- Catanese, M.T., Uryu, K., Kopp, M., Edwards, T.J., Andrus, L., Rice, W.J., Silvestry, M., Kuhn, R.J., and Rice, C.M. (2013b). Ultrastructural analysis of hepatitis C virus particles. *Proceedings of the National Academy of Sciences*.
- Cazalla, D., Yario, T., Steitz, J.A., and Steitz, J. (2010). Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA. *Science* *328*, 1563–1566.
- Cermelli, S., Ruggieri, A., Marrero, J.A., Ioannou, G.N., and Beretta, L. (2011). Circulating microRNAs in patients with chronic hepatitis C and non-alcoholic fatty liver disease. *PLoS ONE* *6*, e23937.
- Chak, E., Talal, A.H., Sherman, K.E., Schiff, E.R., and Saab, S. (2011). Hepatitis C virus infection in USA: an estimate of true prevalence. *Liver Int* *31*, 1090–1101.
- Chang, J., Provost, P., and Taylor, J.M. (2003). Resistance of Human Hepatitis Delta Virus RNAs to Dicer Activity. *Journal of Virology* *77*, 11910–11917.
- Chang, J., Nicolas, E., Marks, D., Sander, C., Lerro, A., Buendia, M.A., Xu, C., Mason, W.S., Moloshok, T., Bort, R., et al. (2004). miR-122, a mammalian liver-specific microRNA, is processed from hcr mRNA and may downregulate the high affinity cationic amino acid transporter CAT-1. *RNA Biol* *1*, 106–113.
- Chen, C.-Z. (2013). An unsolved mystery: the target-recognizing RNA species of microRNA genes. *Biochimie* *95*, 1663–1676.
- Chen, K., and Rajewsky, N. (2006). Deep conservation of microRNA-target relationships and 3'UTR motifs in vertebrates, flies, and nematodes. *Cold Spring Harb. Symp. Quant. Biol.* *71*, 149–156.
- Chen, Y., Shen, A., Rider, P.J., Yu, Y., Wu, K., Mu, Y., Hao, Q., Liu, Y., Gong, H., Zhu, Y., et al. (2011). A liver-specific microRNA binds to a highly conserved RNA sequence of hepatitis B virus and negatively regulates viral gene expression and replication. *Faseb J* *25*, 4511–4521.
- Chen, Y., Li, L., Zhou, Z., Wang, N., Zhang, C.-Y., and Zen, K. (2012). A pilot study of serum microRNA signatures as a novel biomarker for occult hepatitis B virus infection. *Med Microbiol Immunol* *201*, 389–395.
- Cheung, V.G., Conlin, L.K., Weber, T.M., Arcaro, M., Jen, K.-Y., Morley, M., and Spielman, R.S. (2003). Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat. Genet.* *33*, 422–425.
- Chi, S.W., Hannon, G.J., and Darnell, R.B. (2012). An alternative mode of microRNA target recognition. *Nat. Struct. Mol. Biol.* *19*, 321–327.

- Chi, S.W., Zang, J.B., Mele, A., and Darnell, R.B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* *460*, 479–486.
- Chong, M.M.W., Rasmussen, J.P., Rudensky, A.Y., and Littman, D.R. (2008). The RNaseIII enzyme Drosha is critical in T cells for preventing lethal inflammatory disease. *J Exp Med* *205*, 2005–2017.
- Choo, Q.L., Kuo, G., Weiner, A.J., Overby, L.R., Bradley, D.W., and Houghton, M. (1989). Isolation of a Cdna Clone Derived From a Blood-Borne Non-a, Non-B Viral-Hepatitis Genome. *Science* *244*, 359–362.
- Cotler, S.J., Craft, T., Ferris, M., Morrissey, M., McCone, J., Reddy, K.R., Conrad, A., Jensen, D.M., Albrecht, J., and Taylor, M.W. (2002). Induction of IL-1Ra in resistant and responsive hepatitis C patients following treatment with IFN-con1. *J. Interferon Cytokine Res.* *22*, 549–554.
- Coulouarn, C., Factor, V.M., Andersen, J.B., Durkin, M.E., and Thorgeirsson, S.S. (2009). Loss of miR-122 expression in liver cancer correlates with suppression of the hepatic phenotype and gain of metastatic properties. *Oncogene* *28*, 3526–3536.
- Cox, E.M., Sagan, S.M., Mortimer, S.A.W., Doudna, J.A., and Sarnow, P. (2013). Enhancement of hepatitis C viral RNA abundance by precursor miR-122 molecules. *RNA* *19*, 1825–1832.
- Cullen, B.R. (2009). Viral and cellular messenger RNA targets of viral microRNAs. *Nature* *457*, 421–425.
- Cullen, B.R. (2010). Five Questions about Viruses and MicroRNAs. *PLoS Pathog* *6*, e1000787.
- Cullen, B.R. (2013). MicroRNAs as mediators of viral evasion of the immune system. *Nat Immunol* *14*, 205–210.
- D'Ambrogio, A., Gu, W., Udagawa, T., Mello, C.C., and Richter, J.D. (2012). Specific miRNA Stabilization by Gld2-Catalyzed Monoadenylation. *CellReports* *2*, 1537–1545.
- Da Costa, D., Turek, M., Felmlee, D.J., Girardi, E., Pfeffer, S., Long, G., Bartenschlager, R., Zeisel, M.B., and Baumert, T.F. (2012). Reconstitution of the Entire Hepatitis C Virus Life Cycle in Nonhepatic Cells. *Journal of Virology* *86*, 11919–11925.
- Darnell, J.C., Jensen, K.B., Jin, P., Brown, V., Warren, S.T., and Darnell, R.B. (2001). Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* *107*, 489–499.

Darnell, J.C., Mostovetsky, O., and Darnell, R.B. (2005). FMRP RNA targets: identification and validation. *Genes Brain Behav.* *4*, 341–349.

Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y.S., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W., et al. (2011). FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism. *Cell* *146*, 247–261.

Darnell, R.B. (2010). HITS-CLIP: panoramic views of protein-RNA regulation in living cells. *Wiley Interdiscip Rev RNA* *1*, 266–286.

Davis, G.L., Alter, M.J., Serag, El, H., Poynard, T., and Jennings, L.W. (2010). Aging of Hepatitis C Virus (HCV)-Infected Persons in the United States: A Multiple Cohort Model of HCV Prevalence and Disease Progression. *Yagast* *138*, 513–521.e516.

de Pontual, L., Yao, E., Callier, P., Faivre, L., Drouin, V., Cariou, S., Van Haeringen, A., Geneviève, D., Goldenberg, A., Oufadem, M., et al. (2011). Germline deletion of the miR-17~92 cluster causes skeletal and growth defects in humans. *Nature Publishing Group* *43*, 1026–1030.

Dedon, P.C., Soultis, J.A., Allis, C.D., and Gorovsky, M.A. (1991). A simplified formaldehyde fixation and immunoprecipitation technique for studying protein-DNA interactions. *Anal Biochem* *197*, 83–90.

Denzler, R., Agarwal, V., Stefano, J., Bartel, D.P., and Stoffel, M. (2014). Assessing the ceRNA Hypothesis with Quantitative Measurements of miRNA and Target Abundance. *Molecular Cell* 1–11.

Diamond, M.S., and Schoggins, J.W. (2013). Host restriction factor screening: let the virus do the work. *Cell Host Microbe* *14*, 229–231.

Ding, X., Ding, J., Ning, J., Yi, F., Chen, J., Zhao, D., Zheng, J., Liang, Z., Hu, Z., and Du, Q. (2012). Circulating microRNA-122 as a potential biomarker for liver injury. *Mol Med Rep* *5*, 1428–1432.

Dittmann, M., Hoffmann, H.-H., Scull, M.A., Gilmore, R.H., Bell, K.L., Ciancanelli, M., Wilson, S.J., Crotta, S., Yu, Y., Flatley, B., et al. (2015). A serpin shapes the extracellular environment to prevent influenza A virus maturation. *Cell* *160*, 631–643.

Dong, R., Zheng, Y., Chen, G., Zhao, R., Zhou, Z., and Zheng, S. (2015). miR-222 overexpression may contribute to liver fibrosis in biliary atresia by targeting PPP2R2A. *J. Pediatr. Gastroenterol. Nutr.* *60*, 84–90.

Drexler, J.F., Corman, V.M., Müller, M.A., Lukashev, A.N., Gmyl, A., Coutard, B., Adam, A., Ritz, D., Leijten, L.M., van Riel, D., et al. (2013). Evidence for novel hepaciviruses in rodents. *PLoS Pathog* 9, e1003438.

Du, T., Zhou, G., and Roizman, B. (2011). HSV-1 gene expression from reactivated ganglia is disordered and concurrent with suppression of latency-associated transcript and miRNAs. *Proc Natl Acad Sci USA* 108, 18820–18824.

Ebert, M.S., and Sharp, P.A. (2010). Emerging Roles for Natural MicroRNA Sponges. *Curr Biol* 20, 0–0.

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

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10, 48–7.

Egger, D., Wölk, B., Gosert, R., Bianchi, L., Blum, H.E., Moradpour, D., and Bienz, K. (2002). Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *Journal of Virology* 76, 5974–5984.

Eis, P.S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M.F., Lund, E., and Dahlberg, J.E. (2005). Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc Natl Acad Sci USA* 102, 3627–3632.

El-Serag, H.B. (2011). Hepatocellular Carcinoma. *New England Journal of Medicine* 365, 1118–1127.

Elkayam, E., Kuhn, C.-D., Tocilj, A., Haase, A.D., Greene, E.M., Hannon, G.J., and Joshua-Tor, L. (2012). The Structure of Human Argonaute-2 in Complex with miR-20a. *Cell* 150, 100–110.

Elmén, J., Lindow, M., Silahtaroglu, A., Bak, M., Christensen, M., Lind-Thomsen, A., Hedtjörn, M., Hansen, J.B., Hansen, H.F., Straarup, E.M., et al. (2008). Antagonism of microRNA-122 in mice by systemically administered LNA-antimiR leads to up-regulation of a large set of predicted target mRNAs in the liver. *Nucleic Acids Res.* 36, 1153–1162.

Esau, C., Davis, S., Murray, S.F., Yu, X.X., Pandey, S.K., Pear, M., Watts, L., Booten, S.L., Graham, M., McKay, R., et al. (2006). miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting. *Cell Metabolism* 3, 87–98.

Etiemble, J., Moroy, T., Jacquemin, E., Tiollais, P., and Buendia, M.A. (1989). Fused transcripts of c-myc and a new cellular locus, hcr in a primary liver tumor. *Oncogene* 4, 51–57.

Evans, M.J., Hahn, von, T., Tscherne, D.M., Syder, A.J., Panis, M., Wölk, B., Hatzioannou, T., McKeating, J.A., Bieniasz, P.D., and Rice, C.M. (2007). Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature* 446, 801–805.

Farh, K.K.-H.K., Grimson, A.A., Jan, C.C., Lewis, B.P.B., Johnston, W.K.W., Lim, L.P.L., Burge, C.B.C., and Bartel, D.P.D. (2005). The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* 310, 1817–1821.

Farid, W.R.R., Pan, Q., van der Meer, A.J.P., de Ruiter, P.E., Ramakrishnaiah, V., de Jonge, J., Kwekkeboom, J., Janssen, H.L.A., Metselaar, H.J., Tilanus, H.W., et al. (2012). Hepatocyte-derived microRNAs as serum biomarkers of hepatic injury and rejection after liver transplantation. *Liver Transpl* 18, 290–297.

Fecko, C.J., Munson, K.M., Saunders, A., Sun, G., Begley, T.P., Lis, J.T., and Webb, W.W. (2007). Comparison of femtosecond laser and continuous wave UV sources for protein-nucleic acid crosslinking. *Photochem Photobiol* 83, 1394–1404.

Felmlee, D., Hafirassou, M., Lefevre, M., Baumert, T., and Schuster, C. (2013). Hepatitis C Virus, Cholesterol and Lipoproteins — Impact for the Viral Life Cycle and Pathogenesis of Liver Disease. *Viruses* 5, 1292–1324.

Figliuzzi, M., Marinari, E., and De Martino, A. (2013). MicroRNAs as a Selective Channel of Communication between Competing RNAs: a Steady-State Theory. *Biophysj* 104, 1203–1213.

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* 9, 102–114.

Firth, C., Bhat, M., Firth, M.A., Williams, S.H., Frye, M.J., Simmonds, P., Conte, J.M., Ng, J., Garcia, J., Bhuva, N.P., et al. (2014). Detection of Zoonotic Pathogens and Characterization of Novel Viruses Carried by Commensal *Rattus norvegicus* in New York City. *mBio* 5, e01933–14–e01933–14.

Fischer, M., Staples, J.E., Arboviral Diseases Branch, National Center for Emerging and Zoonotic Infectious Diseases, CDC (2014). Notes from the field: chikungunya virus spreads in the Americas - Caribbean and South America, 2013-2014. *MMWR Morb. Mortal. Wkly. Rep.* 63, 500–501.

Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahovicek, K., and Svoboda, P. (2013). A Retrotransposon-Driven Dicer Isoform Directs Endogenous Small Interfering RNA Production in Mouse Oocytes. *Cell* 155, 807–816.

Flisiak, R., Jaroszewicz, J., Flisiak, I., and Łapiński, T. (2012). Update on alisporivir in treatment of viral hepatitis C. *Expert Opin Investig Drugs* 21, 375–382.

Fong, M.Y., Zhou, W., Liu, L., Alontaga, A.Y., Chandra, M., Ashby, J., Chow, A., O'Connor, S.T.F., Li, S., Chin, A.R., et al. (2015). Breast-cancer-secreted miR-122 reprograms glucose metabolism in premetastatic niche to promote metastasis. *Nat Cell Biol* 17, 183–194.

Fornari, F., Gramantieri, L., Giovannini, C., Veronese, A., Ferracin, M., Sabbioni, S., Calin, G.A., Grazi, G.L., Croce, C.M., Tavorari, S., et al. (2009). MiR-122/cyclin G1 interaction modulates p53 activity and affects doxorubicin sensitivity of human hepatocarcinoma cells. *Cancer Research* 69, 5761–5767.

Foy, E., Li, K., Wang, C.F., Sumpter, R., Ikeda, M., Lemon, S.M., and Gale, M. (2003). Regulation of interferon regulatory factor-3 by the hepatitis C virus serine protease. *Science* 300, 1145–1148.

Frankel, L.B., Di Malta, C., Wen, J., Eskelinen, E.-L., Ballabio, A., and Lund, A.H. (2014). A non-conserved miRNA regulates lysosomal function and impacts on a human lysosomal storage disorder. *Nature Communications* 5, 5840–11.

Fraser, C.S., and Doudna, J.A. (2007). Structural and mechanistic insights into hepatitis C viral translation initiation. *Nat Rev Microbiol* 5, 29–38.

Friedman, R.C.R., Farh, K.K.-H.K., Burge, C.B.C., and Bartel, D.P.D. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genes Dev* 19, 92–105.

Fu, Z.D., Csanaky, I.L., and Klaassen, C.D. (2012). Effects of Aging on mRNA Profiles for Drug-Metabolizing Enzymes and Transporters in Livers of Male and Female Mice. *Drug Metabolism and Disposition* 40, 1216–1225.

Fukuhara, T., Kambara, H., Shiokawa, M., Ono, C., Katoh, H., Morita, E., Okuzaki, D., Maehara, Y., Koike, K., and Matsuura, Y. (2012). Expression of MicroRNA miR-122 Facilitates an Efficient Replication in Nonhepatic Cells upon Infection with Hepatitis C Virus. *Journal of Virology* 86, 7918–7933.

Gabay, C., Lamacchia, C., and Palmer, G. (2010). IL-1 pathways in inflammation and human diseases. *Nat Rev Rheumatol* 6, 232–241.

Gangaraju, V.K., and Lin, H. (2009). MicroRNAs: key regulators of stem cells. *Nat Rev Mol Cell Biol* *10*, 116–125.

Gao, Y., He, Y., Ding, J., Wu, K., Hu, B., Liu, Y., Wu, Y., Guo, B., Shen, Y., Landi, D., et al. (2009). An insertion/deletion polymorphism at miRNA-122-binding site in the interleukin-1alpha 3' untranslated region confers risk for hepatocellular carcinoma. *Carcinogenesis* *30*, 2064–2069.

Garcia, D.M., Baek, D., Shin, C., Bell, G.W., Grimson, A., and Bartel, D.P. (2011). Weak seed-pairing stability and high target-site abundance decrease the proficiency of Isy-6 and other microRNAs. *Nat. Struct. Mol. Biol.* *18*, 1139–1146.

Gastaminza, P., Cheng, G., Wieland, S., Zhong, J., Liao, W., and Chisari, F.V. (2008). Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. *Journal of Virology* *82*, 2129.

Gatfield, D., Le Martelot, G., Vejnar, C.E., Gerlach, D., Schaad, O., Fleury-Olela, F., Ruskeepaa, A.-L., Oresic, M., Esau, C.C., Zdobnov, E.M., et al. (2009). Integration of microRNA miR-122 in hepatic circadian gene expression. *Genes Dev* *23*, 1313–1326.

Gilmour, D.S., and Lis, J.T. (1984). Detecting Protein-Dna Interactions In vivo - Distribution of Rna-Polymerase on Specific Bacterial Genes. *Proc Natl Acad Sci USA* *81*, 4275–4279.

Goecks, J., Nekrutenko, A., Taylor, J., Galaxy Team (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* *11*, R86.

Gottwein, J.M., Scheel, T.K.H., Hoegh, A.M., Lademann, J.B., Eugen-Olsen, J., Lisby, G., and Bukh, J. (2007). Robust Hepatitis C Genotype 3a Cell Culture Releasing Adapted Intergenotypic 3a/2a (S52/JFH1) Viruses. *Gastroenterology* *133*, 1614-1626.

Gottwein, J.M., Scheel, T.K.H., Jensen, T.B., Lademann, J.B., Prentoe, J.C., Knudsen, M.L., Hoegh, A.M., and Bukh, J. (2009). Development and characterization of hepatitis C virus genotype 1-7 cell culture systems: Role of CD81 and scavenger receptor class B type I and effect of antiviral drugs. *Hepatology* *49*, 364–377.

Gould, E.A., and Higgs, S. (2009). Impact of climate change and other factors on emerging arbovirus diseases. *Transactions of the Royal Society of Tropical Medicine and Hygiene* *103*, 109–121.

Gramantieri, L., Casali, A., Trerè, D., Gaiani, S., Piscaglia, F., Chieco, P., Cola, B., and Bolondi, L. (1999). Imbalance of IL-1 beta and IL-1 receptor antagonist mRNA in liver tissue from hepatitis C virus (HCV)-related chronic hepatitis. *Clin Exp Immunol* 115, 515–520.

Griffiths-Jones, S., Saini, H.K., van Dongen, S., and Enright, A.J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Res.* 36, D154–D158.

Grimson, A., Farh, K.K.-H., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Molecular Cell* 27, 15–15.

Grosswendt, S., Filipchuk, A., Manzano, M., Klironomos, F., Schilling, M., Herzog, M., Gottwein, E., and Rajewsky, N. (2014). Unambiguous Identification of miRNA:Target Site Interactions by Different Types of Ligation Reactions. *Molecular Cell* 54, 1042–1054.

Gupta, P., Cairns, M.J., and Saksena, N.K. (2014). Regulation of gene expression by microRNA in HCV infection and HCV-mediated hepatocellular carcinoma. *Virology Journal* 11, 64–64.

Haecker, I.I., Gay, L.A.L., Yang, Y.Y., Hu, J.J., Morse, A.M.A., McIntyre, L.M.L., and Renne, R.R. (2012). Ago HITS-CLIP expands understanding of Kaposi's sarcoma-associated herpesvirus miRNA function in primary effusion lymphomas. *PLoS Pathog* 8, e1002884–e1002884.

Hafner, M.M., Landthaler, M.M., Burger, L.L., Khorshid, M.M., Hausser, J.J., Berninger, P.P., Rothballer, A.A., Ascano, M.M., Jungkamp, A.-C.A., Munschauer, M.M., et al. (2010). Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. *Cell* 141, 13–13.

Hafner, M., Renwick, N., Farazi, T.A., Mihailović, A., Pena, J.T.G., and Tuschl, T. (2012). Barcoded cDNA library preparation for small RNA profiling by next-generation sequencing. *Methods* 58, 164–170.

Han, H., Sun, D., Li, W., Shen, H., Zhu, Y., Li, C., Chen, Y., Lu, L., Li, W., Zhang, J., et al. (2013). A c-Myc-MicroRNA functional feedback loop affects hepatocarcinogenesis. *Hepatology* 57, 2378–2389.

Han, J.-Q., and Barton, D.J. (2002). Activation and evasion of the antiviral 2'–5' oligoadenylate synthetase/ribonuclease L pathway by hepatitis C virus mRNA. *RNA* 8, 512–525.

Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* 495, 384–388.

Helwak, A., Kudla, G., Dudnakova, T., and Tollervey, D. (2013). Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding. *Cell* 153, 654–665.

Henke, J.I., Goergen, D., Zheng, J., Song, Y., Schüttler, C.G., Fehr, C., Jünemann, C., and Niepmann, M. (2008). microRNA-122 stimulates translation of hepatitis C virus RNA. *Embo J* 27, 3300–3310.

Herbst, D.A., and Reddy, K.R. (2013). Risk factors for hepatocellular carcinoma. *Clinical Liver Disease* 1, 180–182.

Herker, E., Harris, C., Hernandez, C., Carpentier, A., Kaehlcke, K., Rosenberg, A.R., Farese, R.V.J., and Ott, M. (2010). Efficient hepatitis C virus particle formation requires diacylglycerol acyltransferase-1. *Nat Med* 16, 1295–1298.

Hirsch, E., Irikura, V.M., Paul, S.M., and Hirsh, D. (1996). Functions of interleukin 1 receptor antagonist in gene knockout and overproducing mice. *Proc Natl Acad Sci USA* 93, 11008–11013.

Holczbauer, Á., Gyöngyösi, B., Lotz, G., Törzsök, P., Kaposi-Novák, P., Szijártó, A., Tátrai, P., Kupcsulik, P., Schaff, Z., and Kiss, A. (2014). Increased Expression of Claudin-1 and Claudin-7 in Liver Cirrhosis and Hepatocellular Carcinoma. *Pathol. Oncol. Res.* 20, 493–502.

Holmes, R.S., Wright, M.W., Lalederkind, S.J.F., Cox, L.A., Hosokawa, M., Imai, T., Ishibashi, S., Lehner, R., Miyazaki, M., Perkins, E.J., et al. (2010). Recommended nomenclature for five mammalian carboxylesterase gene families: human, mouse, and rat genes and proteins. *Mamm Genome* 21, 427–441.

Hoofnagle, J.H. (1997). Hepatitis C: the clinical spectrum of disease. (Wiley Subscription Services, Inc., A Wiley Company), pp. 15S–20S.

Hoofnagle, J.H., and Seeff, L.B. (2006). Peginterferon and ribavirin for chronic hepatitis C. *New England Journal of Medicine* 355, 2444–2451.

Horner, S.M., and Gale, M. (2013). Regulation of hepatic innate immunity by hepatitis C virus. *Nat Med* 19, 879–888.

Horwitz, J.A., Dorner, M., Friling, T., Donovan, B.M., Vogt, A., Loureiro, J., Oh, T., Rice, C.M., and Ploss, A. (2013). Expression of heterologous proteins flanked by NS3-4A cleavage sites within the hepatitis C virus polyprotein. *Virology* 439, 23–33.

Hsu, S.-H., Wang, B., Kota, J., Yu, J., Costinean, S., Kutay, H., Yu, L., Bai, S., La Perle, K., Chivukula, R.R., et al. (2012). Essential metabolic, anti-inflammatory, and anti-tumorigenic functions of miR-122 in liver. *J. Clin. Invest.* *122*, 2871–2883.

Hu, J., Wang, Z., Tan, C.-J., Liao, B.-Y., Zhang, X., Xu, M., Dai, Z., Qiu, S.-J., Huang, X.-W., Sun, J., et al. (2013). Plasma microRNA, a potential biomarker for acute rejection after liver transplantation. *Transplantation* *95*, 991–999.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* *4*, 44–57.

Huang, N., Lee, I., Marcotte, E.M., and Hurles, M.E. (2010). Characterising and Predicting Haploinsufficiency in the Human Genome. *PLoS Genet* *6*, e1001154–11.

Hussain, M., and Asgari, S. (2014). MicroRNA-like viral small RNA from Dengue virus 2 autoregulates its replication in mosquito cells. *Proc Natl Acad Sci USA* *111*, 2746–2751.

Hutvagner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science* *297*, 2056–2060.

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

Ince-Dunn, G., Okano, H.J., Jensen, K.B., Park, W.-Y., Zhong, R., Ule, J., Mele, A., Fak, J.J., Yang, C., Zhang, C., et al. (2012). Neuronal Elav-like (Hu) Proteins Regulate RNA Splicing and Abundance to Control Glutamate Levels and Neuronal Excitability. *Neuron* *75*, 1067–1080.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* *324*, 218–223.

Israelow, B., Mullokandov, G., Agudo, J., Sourisseau, M., Bashir, A., Maldonado, A.Y., Dar, A.C., Brown, B.D., and Evans, M.J. (2014). Hepatitis C virus genetics affects miR-122 requirements and response to miR-122 inhibitors. *Nature Communications* *5*, 1–11.

Ito, T., and Lai, M.M. (1999). An internal polypyrimidine-tract-binding protein-binding site in the hepatitis C virus RNA attenuates translation, which is relieved by the 3'-untranslated sequence. *Virology* *254*, 288–296.

- Jan, C.H., Williams, C.C., and Weissman, J.S. (2014). Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. *Science* *346*, 1257521–1257521.
- Jangra, R.K., Yi, M., and Lemon, S.M. (2010). Regulation of Hepatitis C Virus Translation and Infectious Virus Production by the MicroRNA miR-122. *Journal of Virology* *84*, 6615–6625.
- Janssen, H.L.A., Reesink, H.W., Lawitz, E.J., Zeuzem, S., Rodriguez-Torres, M., Patel, K., van der Meer, A.J., Patick, A.K., Chen, A., Zhou, Y., et al. (2013). Treatment of HCV Infection by Targeting MicroRNA. *New England Journal of Medicine* *386*(18), 1685-1694.
- Jens, M., and Rajewsky, N. (2014). Competition between target sites of regulators shapes post-transcriptional gene regulation. *Nat Rev Genet* 1–14.
- Ji, H., Fraser, C.S., Yu, Y., Leary, J., and Doudna, J.A. (2004). Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc Natl Acad Sci USA* *101*, 16990–16995.
- Jones, C.T., Catanese, M.T., Law, L.M.J., Khetani, S.R., Syder, A.J., Ploss, A., Oh, T.S., Schoggins, J.W., MacDonald, M.R., Bhatia, S.N., et al. (2010). Real-time imaging of hepatitis C virus infection using a fluorescent cell-based reporter system. *Nat. Biotechnol.* *28*, 167–171.
- Jones, C.T., Murray, C.L., Eastman, D.K., Tassello, J., and Rice, C.M. (2007). Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus. *Journal of Virology* *81*, 8374–8383.
- Jopling, C.L., Norman, K.L., and Sarnow, P. (2006). Positive and negative modulation of viral and cellular mRNAs by liver-specific microRNA miR-122. *Cold Spring Harb. Symp. Quant. Biol.* *71*, 369–376.
- Jopling, C.L., Schütz, S., and Sarnow, P. (2008). Position-dependent function for a tandem microRNA miR-122-binding site located in the hepatitis C virus RNA genome. *Cell Host Microbe* *4*, 77–85.
- Jopling, C.L., Yi, M., Lancaster, A.M., Lemon, S.M., and Sarnow, P. (2005). Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* *309*, 1577–1581.
- Jung, C.J.C., Iyengar, S.S., Blahnik, K.R.K., Ajuha, T.P.T., Jiang, J.X.J., Farnham, P.J.P., and Zern, M.M. (2011). Epigenetic modulation of miR-122 facilitates human embryonic stem cell self-renewal and hepatocellular carcinoma proliferation. *PLoS ONE* *6*, e27740–e27740.

Kandathil, A.J., Graw, F., Quinn, J., Hwang, H.S., Torbenson, M., Perelson, A.S., Ray, S.C., Thomas, D.L., Ribeiro, R.M., and Balagopal, A. (2013). *Gastroenterology* 145, 1404–1413.e1410.

Kapoor, A., Simmonds, P., Gerold, G., Qaisar, N., Jain, K., Henriquez, J.A., Firth, C., Hirschberg, D.L., Rice, C.M., Shields, S., et al. (2011). Characterization of a canine homolog of hepatitis C virus. *Proceedings of the National Academy of Sciences* 108, 11608–11613.

Kapoor, A., Simmonds, P., Scheel, T.K.H., Hjelle, B., Cullen, J.M., Burbelo, P.D., Chauhan, L.V., Duraisamy, R., Sanchez Leon, M., Jain, K., et al. (2013). Identification of rodent homologs of hepatitis C virus and pegiviruses. *mBio* 4, e00216–13.

Karakatsanis, A., Papaconstantinou, I., Gazouli, M., Lyberopoulou, A., Polymeneas, G., and Voros, D. (2013). Expression of microRNAs, miR-21, miR-31, miR-122, miR-145, miR-146a, miR-200c, miR-221, miR-222, and miR-223 in patients with hepatocellular carcinoma or intrahepatic cholangiocarcinoma and its prognostic significance. *Mol. Carcinog.* 52, 297–303.

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

Keller, P., Petrie, J.T., De Rose, P., Gerin, I., Wright, W.S., Chiang, S.-H., Nielsen, A.R., Fischer, C.P., Pedersen, B.K., and MacDougald, O.A. (2008). Fat-specific protein 27 regulates storage of triacylglycerol. *J. Biol. Chem.* 283, 14355–14365.

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

Khan, A.G., Whidby, J., Miller, M.T., Scarborough, H., Zatorski, A.V., Cygan, A., Price, A.A., Yost, S.A., Bohannon, C.D., Jacob, J., et al. (2014). Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2. *Nature* 509, 381–384.

Kincaid, R.P., Burke, J.M., and Sullivan, C.S. (2012). RNA virus microRNA that mimics a B-cell oncomiR. *Proc Natl Acad Sci USA* 109, 3077–3082.

Kincaid, R.P., Chen, Y., Cox, J.E., Rethwilm, A., and Sullivan, C.S. (2014). Noncanonical MicroRNA (miRNA) Biogenesis Gives Rise to Retroviral Mimics of Lymphoproliferative and Immunosuppressive Host miRNAs. *mBio* 5, e00074–14–e00074–14.

- Kincaid, R.P., and Sullivan, C.S. (2012). Virus-Encoded microRNAs: An Overview and a Look to the Future. *PLoS Pathog* 8.
- King, H.A., Cobbold, L.C., and Willis, A.E. (2010). The role of IRES trans-acting factors in regulating translation initiation. *Biochem. Soc. Trans* 38, 1581–1586.
- Kiyosawa, K., Sodeyama, T., Tanaka, E., Gibo, Y., Yoshizawa, K., Nakano, Y., Furuta, S., Akahane, Y., Nishioka, K., Purcell, R.H., et al. (1990). Interrelationship of blood transfusion, non-A, non-B hepatitis and hepatocellular carcinoma: Analysis by detection of antibody to hepatitis C virus. *Hepatology* 12, 671–675.
- Kojima, S.S., Gatfield, D.D., Esau, C.C.C., and Green, C.B.C. (2010). MicroRNA-122 modulates the rhythmic expression profile of the circadian deadenylase Nocturnin in mouse liver. *PLoS ONE* 5, e11264–e11264.
- Krützfeldt, J., Rajewsky, N., Braich, R., Rajeev, K.G., Tuschl, T., Manoharan, M., and Stoffel, M. (2005). Silencing of microRNAs in vivo with “antagomirs.” *Nature* 438, 685–689.
- Krützfeldt, J., Rösch, N., Hausser, J., Manoharan, M., Zavolan, M., and Stoffel, M. (2012). MicroRNA-194 is a target of transcription factor 1 (Tcf1, HNF1 α) in adult liver and controls expression of frizzled-6. *Hepatology* 55, 98–107.
- Kunte, D.P., Delacruz, M., Wali, R.K., Menon, A., Du, H., Stypula, Y., Patel, A., Backman, V., and Roy, H.K. (2012). Dysregulation of microRNAs in colonic field carcinogenesis: implications for screening. *PLoS ONE* 7, e45591.
- Kuo, G., Choo, Q.L., Alter, H.J., Gitnick, G.L., Redeker, A.G., Purcell, R.H., Miyamura, T., Dienstag, J.L., Alter, M.J., Stevens, C.E., et al. (1989). An Assay for Circulating Antibodies to a Major Etiologic Virus of Human Non-a, Non-B-Hepatitis. *Science* 244, 362–364.
- Kuo, M.H., and Allis, C.D. (1999). In vivo cross-linking and immunoprecipitation for studying dynamic protein: DNA associations in chromatin environment. *Methods* 19, 425–433.
- Kutay, H., Bai, S., Datta, J., Motiwala, T., Pogribny, I., Frankel, W., Jacob, S.T., and Ghoshal, K. (2006). Downregulation of miR-122 in the rodent and human hepatocellular carcinomas. *J Cell Biochem* 99, 671–678.
- Kyrmizi, I., Hatzis, P., Katakili, N., Tronche, F., Gonzalez, F.J., and Talianidis, I. (2006). Plasticity and expanding complexity of the hepatic transcription factor network during liver development. *Genes Dev* 20, 2293–2305.

- Ladewig, E., Okamura, K., Flynt, A.S., Westholm, J.O., and Lai, E.C. (2012). Discovery of hundreds of mirtrons in mouse and human small RNA data. *Genome Research* 22, 1634–1645.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science* 294, 853–858.
- Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. (2002). Identification of tissue-specific microRNAs from mouse. *Curr Biol* 12, 735–739.
- Lanford, R.E., Hildebrandt-Eriksen, E.S., Petri, A., Persson, R., Lindow, M., Munk, M.E., Kauppinen, S., and Orum, H. (2010). Therapeutic Silencing of MicroRNA-122 in Primates with Chronic Hepatitis C Virus Infection. *Science* 327, 198–201.
- Laudadio, I., Manfredi, I., Achouri, Y., Schmidt, D., Wilson, M.D., Cordi, S., Thorrez, L., Knoops, L., Jacquemin, P., Schuit, F., et al. (2012). A Feedback Loop Between the Liver-Enriched Transcription Factor Network and Mir-122 Controls Hepatocyte Differentiation. *Gastroenterology* 142, 119–129.
- Lavie, M., Goffard, A., and Dubuisson, J. (2007). Assembly of a functional HCV glycoprotein heterodimer. *Curr Issues Mol Biol* 9, 71–86.
- Lee, C.-T., Risom, T., and Strauss, W.M. (2007). Evolutionary Conservation of MicroRNA Regulatory Circuits: An Examination of MicroRNA Gene Complexity and Conserved MicroRNA-Target Interactions through Metazoan Phylogeny. *DNA and Cell Biology* 26, 209–218.
- Lee, R.C., and Ambros, V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862–864.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C-Elegans Heterochronic Gene Lin-4 Encodes Small RNAs with Antisense Complementarity to Lin-14. *Cell* 75, 843–854.
- Lee, S., Song, J., Kim, S., Kim, J., Hong, Y., Kim, Y., Kim, D., Baek, D., and Ahn, K. (2013). Selective Degradation of Host MicroRNAs by an Intergenic HCMV Noncoding RNA Accelerates Virus Production. *Cell Host Microbe* 13, 678–690.
- Leek, J.T., Scharpf, R.B., Bravo, H.C., Simcha, D., Langmead, B., Johnson, W.E., Geman, D., Baggerly, K., and Irizarry, R.A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat Rev Genet* 11, 733–739.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20.

Li, C.C., Wang, Y.Y., Wang, S.S., Wu, B.B., Hao, J.J., Fan, H.H., Ju, Y.Y., Ding, Y.Y., Chen, L.L., Chu, X.X., et al. (2013a). Hepatitis B virus mRNA-mediated miR-122 inhibition upregulates PTTG1-binding protein, which promotes hepatocellular carcinoma tumor growth and cell invasion. *Journal of Virology* **87**, 2193–2205.

Li, D., Wang, X., Lan, X., Li, Y., Liu, L., Yi, J., Li, J., Sun, Q., Wang, Y., Li, H., et al. (2015). Down-regulation of miR-144 elicits proinflammatory cytokine production by targeting toll-like receptor 2 in nonalcoholic steatohepatitis of high-fat-diet-induced metabolic syndrome E3 rats. *Mol. Cell. Endocrinol.* **402**, 1–12.

Li, J., Ghazwani, M., Zhang, Y., Lu, J., Li, J., Fan, J., Gandhi, C.R., and Li, S. (2013b). miR-122 regulates collagen production via targeting hepatic stellate cells and suppressing P4HA1 expression. *J Hepatol* **58**, 522–528.

Li, J., Liu, Y., Xin, X., Kim, T.S., Cabeza, E.A., Ren, J., Nielsen, R., Wrana, J.L., and Zhang, Z. (2012). Evidence for Positive Selection on a Number of MicroRNA Regulatory Interactions during Recent Human Evolution. *PLoS Genet* **8**, e1002578–12.

Li, Y., Lu, J., Han, Y., Fan, X., and Ding, S.W. (2013c). RNA Interference Functions as an Antiviral Immunity Mechanism in Mammals. *Science* **342**, 231–234.

Li, Y.-P., Gottwein, J.M., Scheel, T.K., Jensen, T.B., and Bukh, J. (2011). MicroRNA-122 antagonism against hepatitis C virus genotypes 1-6 and reduced efficacy by host RNA insertion or mutations in the HCV 5' UTR. *Proceedings of the National Academy of Sciences* **108**, 4991–4996.

Li, Y., Masaki, T., Yamane, D., McGivern, D.R., and Lemon, S.M. (2013d). Competing and noncompeting activities of miR-122 and the 5' exonuclease Xrn1 in regulation of hepatitis C virus replication. *Proc Natl Acad Sci USA* **110**, 1881–1886.

Li, Y., Wang, J., Kanai, R., and Modis, Y. (2013e). Crystal structure of glycoprotein E2 from bovine viral diarrhea virus. *Proceedings of the National Academy of Sciences* **110**, 6810.

Liang, Y., Shilagard, T., Xiao, S.Y., Snyder, N., Lau, D., Cicalese, L., Weiss, H., Vargas, G., and Lemon, S.M. (2009). Visualizing Hepatitis C Virus Infections in Human Liver by Two-Photon Microscopy. *Gastroenterology* **137**, 1448–1458.

Libra, M.M., Mangano, K.K., Anzaldi, M.M., Quattrocchi, C.C., Donia, M.M., di Marco, R.R., Signorelli, S.S., Scalia, G.G., Zignego, A.L.A., de Re, V.V., et al. (2006). Analysis of interleukin (IL)-1beta IL-1 receptor antagonist, soluble IL-1 receptor type II and IL-1 accessory protein in HCV-associated lymphoproliferative disorders. *Oncol Rep* 15, 1305–1308.

Licatalosi, D.D., Yano, M., Fak, J.J., Mele, A., Grabinski, S.E., Zhang, C., and Darnell, R.B. (2012). Ptbp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. *Genes Dev* 26, 1626–1642.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., et al. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456, 464–469.

Lima, W.F., Wu, H., Nichols, J.G., Sun, H., Murray, H.M., and Crooke, S.T. (2009). Binding and cleavage specificities of human Argonaute2. *Journal of Biological Chemistry* 284, 26017–26028.

Lin, J., and Cullen, B.R. (2007). Analysis of the interaction of primate retroviruses with the human RNA interference machinery. *Journal of Virology* 81, 12218–12226.

Lin, L.-T., Noyce, R.S., Pham, T.N.Q., Wilson, J.A., Sisson, G.R., Michalak, T.I., Mossman, K.L., and Richardson, C.D. (2010). Replication of subgenomic hepatitis C virus replicons in mouse fibroblasts is facilitated by deletion of interferon regulatory factor 3 and expression of liver-specific microRNA 122. *Journal of Virology* 84, 9170–9180.

Lin, M., Lucas, H.C., Jr, and Shmueli, G. (2013). Research Commentary—Too Big to Fail: Large Samples and the p-Value Problem. *Information Systems Research* 24, 906–917.

Lindenbach, B.D., and Rice, C.M. (2005). Unravelling hepatitis C virus replication from genome to function. *Nature* 436, 933–938.

Lindenbach, B.D., and Rice, C.M. (2013). The ins and outs of hepatitis C virus entry and assembly. *Nat Rev Microbiol* 11, 688–700.

Lindenbach, B.D., Evans, M.J., Syder, A.J., Wölk, B., Tellinghuisen, T.L., Liu, C.C., Maruyama, T., Hynes, R.O., Burton, D.R., McKeating, J.A., et al. (2005). Complete replication of hepatitis C virus in cell culture. *Science* 309, 623–626.

Liu, A.M., Xu, Z., Shek, F.H., Wong, K.-F., Lee, N.P., Poon, R.T., Chen, J., and Luk, J.M. (2014a). miR-122 targets pyruvate kinase M2 and affects metabolism of hepatocellular carcinoma. *PLoS ONE* 9, e86872.

Liu, X.-H., Sun, M., Nie, F.-Q., Ge, Y.-B., Zhang, E.-B., Yin, D.-D., Kong, R., Xia, R., Lu, K.-H., Li, J.-H., et al. (2014b). Lnc RNA HOTAIR functions as a competing endogenous RNA to regulate HER2 expression by sponging miR-331-3p in gastric cancer. *Mol Cancer* 13, 92.

Loeb, G.B., Khan, A.A., Canner, D., Hiatt, J.B., Shendure, J., Darnell, R.B., Leslie, C.S., and Rudensky, A.Y. (2012). Transcriptome-wide miR-155 binding map reveals widespread noncanonical microRNA targeting. *Molecular Cell* 48, 760–770.

Lohmann, V., Körner, F., Koch, J., Herian, U., Theilmann, L., and Bartenschlager, R. (1999). Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* 285, 110–113.

Lohmann, V. (2013). Hepatitis C Virus RNA Replication. *Curr. Top. Microbiol. Immunol.* 369, 167–198.

Luna, J.M., Scheel, T.K.H., Danino, T., Shaw, K.S., Mele, A., Fak, J.J., Nishiuchi, E., Takacs, C.N., Catanese, M.T., de Jong, Y.P., et al. (2015). Hepatitis C Virus RNA Functionally Sequesters miR-122. *Cell* 160, 1099–1110.

Luo, G. (1999). Cellular proteins bind to the poly(U) tract of the 3' untranslated region of hepatitis C virus RNA genome. *Virology* 256, 118.

Lynch, M., and Hagner, K. (2015). Evolutionary meandering of intermolecular interactions along the drift barrier. *Proceedings of the National Academy of Sciences* 112, E30–E38.

Ma, E., MacRae, I.J., Kirsch, J.F., and Doudna, J.A. (2008). Autoinhibition of Human Dicer by Its Internal Helicase Domain. *Journal of Molecular Biology* 380, 237–243.

Ma, E., Zhou, K., Kidwell, M.A., and Doudna, J.A. (2012). Coordinated Activities of Human Dicer Domains in Regulatory RNA Processing. *Journal of Molecular Biology* 422, 466–476.

Machlin, E.S., Sarnow, P., and Sagan, S.M. (2011). Masking the 5' terminal nucleotides of the hepatitis C virus genome by an unconventional microRNA-target RNA complex. *Proceedings of the National Academy of Sciences* 108, 3193–3198.

MacKay, C.R., Wang, J.P., and Kurt-Jones, E.A. (2014). Dicer's role as an antiviral: still an enigma. *Curr. Opin. Immunol.* 26, 49–55.

- Maillard, P.V., Ciaudo, C., Marchais, A., Li, Y., Jay, F., Ding, S.W., and Voinnet, O. (2013). Antiviral RNA Interference in Mammalian Cells. *Science* *342*, 235–238.
- Malim, M.H., and Bieniasz, P.D. (2012). HIV Restriction Factors and Mechanisms of Evasion. *Cold Spring Harb Perspect Med* *2*, a006940.
- Marukian, S., Jones, C.T., Andrus, L., Evans, M.J., Ritola, K.D., Charles, E.D., Rice, C.M., and Dustin, L.B. (2008). Cell culture-produced hepatitis C virus does not infect peripheral blood mononuclear cells. *Hepatology* *48*, 1843–1850.
- Mas, V.R., Maluf, D.G., Archer, K.J., Yanek, K., Kong, X., Kulik, L., Freise, C.E., Olthoff, K.M., Ghobrial, R.M., McIver, P., et al. (2009). Genes involved in viral carcinogenesis and tumor initiation in hepatitis C virus-induced hepatocellular carcinoma. *Mol. Med.* *15*, 85–94.
- Masaki, T., Arend, K.C., Li, Y., Yamane, D., McGivern, D.R., Kato, T., Wakita, T., Moorman, N.J., and Lemon, S.M. (2015). miR-122 Stimulates Hepatitis C Virus RNA Synthesis by Altering the Balance of Viral RNAs Engaged in Replication versus Translation. *Cell Host Microbe* *17*, 217–228.
- Matouk, I.J., Abbasi, I., Hochberg, A., Galun, E., Dweik, H., and Akkawi, M. (2009). Highly upregulated in liver cancer noncoding RNA is overexpressed in hepatic colorectal metastasis. *Eur J Gastroenterol Hepatol* *21*, 688–692.
- Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* *495*, 333–338.
- Merz, A., Long, G., Hiet, M.-S., Bruegger, B., Chlanda, P., Andre, P., Wieland, F., Krijnse-Locker, J., and Bartenschlager, R. (2011). Biochemical and Morphological Properties of Hepatitis C Virus Particles and Determination of Their Lipidome. *J. Biol. Chem.* *286*, 3018–3032.
- Messina, J.P., Humphreys, I., Flaxman, A., Brown, A., Cooke, G.S., Pybus, O.G., and Barnes, E. (2015). Global distribution and prevalence of hepatitis C virus genotypes. *Hepatology* *61*, 77–87.
- Meylan, E., Curran, J., Hofmann, K., Moradpour, D., Binder, M., Bartenschlager, R., and Tschopp, J. (2005). Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature* *437*, 1167–1172.
- Mili, S., and Steitz, J.A. (2004). Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* *10*, 1692–1694.

Miyanari, Y., Atsuzawa, K., Usuda, N., Watashi, K., Hishiki, T., Zayas, M., Bartenschlager, R., Wakita, T., Hijikata, M., and Shimotohno, K. (2007). The lipid droplet is an important organelle for hepatitis C virus production. *Nat Cell Biol* 9, 1089–1097.

Mohd Hanafiah, K., Groeger, J., Flaxman, A.D., and Wiersma, S.T. (2013). Global epidemiology of hepatitis C virus infection: New estimates of age-specific antibody to HCV seroprevalence. *Hepatology* 57, 1333–1342.

Moon, S.L., Anderson, J.R., Kumagai, Y., Wilusz, C.J., Akira, S., Khromykh, A.A., and Wilusz, J. (2012). A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular exoribonuclease XRN1 and alters host mRNA stability. *RNA* 18, 2029–2040.

Moore, M.J., Zhang, C., Gantman, E.C., Mele, A., Darnell, J.C., and Darnell, R.B. (2014). Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. *Nat Protoc* 9, 263–293.

Moroy, T., Etiemble, J., Bougueleret, L., Hadchouel, M., Tiollais, P., and Buendia, M.A. (1989). Structure and expression of hcr, a locus rearranged with c-myc in a woodchuck hepatocellular carcinoma. *Oncogene* 4, 59–65.

Moroy, T., Marchio, A., Etiemble, J., Trépo, C., Tiollais, P., and Buendia, M.A. (1986). Rearrangement and enhanced expression of c-myc in hepatocellular carcinoma of hepatitis virus infected woodchucks. *Nature* 324, 276–279.

Mortimer, S.A., and Doudna, J.A. (2013). Unconventional miR-122 binding stabilizes the HCV genome by forming a trimolecular RNA structure. *Nucleic Acids Res.*

Mukherji, S., Ebert, M.S., Zheng, G.X.Y., Tsang, J.S., Sharp, P.A., and van Oudenaarden, A. (2011). MicroRNAs can generate thresholds in target gene expression. *Nat. Genet.* 43, 854–859.

Murphy, D.G., Sablon, E., Chamberland, J., Fournier, E., Dandavino, R., and Tremblay, C.L. (2014). Hepatitis C Virus Genotype 7: A New Genotype Originating from Central Africa. *J Clin Microbiol.*

Nakabayashi, H., Taketa, K., Miyano, K., Yamane, T., and Sato, J. (1982). Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. *Cancer Research* 42, 3858–3863.

Nakaya, T., Alexiou, P., Maragkakis, M., Chang, A., and Mourelatos, Z. (2013). FUS regulates genes coding for RNA-binding proteins in neurons by binding to their highly conserved introns. *RNA* 19, 498–509.

Narbus, C.M., Israelow, B., Sourisseau, M., Michta, M.L., Hopcraft, S.E., Zeiner, G.M., and Evans, M.J. (2011). HepG2 Cells Expressing MicroRNA miR-122 Support the Entire Hepatitis C Virus Life Cycle. *Journal of Virology* *85*, 12087–12092.

Nasheri, N., Singaravelu, R., Goodmurphy, M., Lyn, R.K., and Pezacki, J.P. (2011). Competing roles of microRNA-122 recognition elements in hepatitis C virus RNA. *Virology* *410*, 336–344.

Nelson, P.T., De Planell-Saguer, M., Lamprinaki, S., Kiriakidou, M., Zhang, P., O'Doherty, U., and Mourelatos, Z. (2007). A novel monoclonal antibody against human Argonaute proteins reveals unexpected characteristics of miRNAs in human blood cells. *RNA* *13*, 1787–1792.

Nemeth, E., Tuttle, M.S., Powelson, J., Vaughn, M.B., Donovan, A., Ward, D.M., Ganz, T., and Kaplan, J. (2004). Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. *Science* *306*, 2090–2093.

Nicklin, M.J., Hughes, D.E., Barton, J.L., Ure, J.M., and Duff, G.W. (2000). Arterial inflammation in mice lacking the interleukin 1 receptor antagonist gene. *J Exp Med* *191*, 303–312.

Nischalke, H.D., Berger, C., Luda, C., Müller, T., Berg, T., Coenen, M., Krämer, B., Körner, C., Trebicka, J., Grünhage, F., et al. (2012). The CXCL1 rs4074 A allele is associated with enhanced CXCL1 responses to TLR2 ligands and predisposes to cirrhosis in HCV genotype 1-infected Caucasian patients. *J Hepatol* *56*, 758–764.

Nischalke, H.D., Berger, C., Lutz, P., Langhans, B., Wolter, F., Eisenhardt, M., Krämer, B., Kokordelis, P., Glässner, A., Müller, T., et al. (2013). Influence of the CXCL1 rs4074 A allele on alcohol induced cirrhosis and HCC in patients of European descent. *PLoS ONE* *8*, e80848.

Nkontchou, G., Ziol, M., Aout, M., Lhabadie, M., Baazia, Y., Mahmoudi, A., Roulot, D., Ganne-Carrie, N., Grando-Lemaire, V., Trinchet, J.-C., et al. (2011). HCV genotype 3 is associated with a higher hepatocellular carcinoma incidence in patients with ongoing viral C cirrhosis. *J. Viral Hepat.* *18*, e516–e522.

Norman, K.L., and Sarnow, P. (2010). Modulation of hepatitis C virus RNA abundance and the isoprenoid biosynthesis pathway by microRNA miR-122 involves distinct mechanisms. *84*, 666–670.

Ogawa, T., Enomoto, M., Fujii, H., Sekiya, Y., Yoshizato, K., Ikeda, K., and Kawada, N. (2012). MicroRNA-221/222 upregulation indicates the activation of stellate cells and the progression of liver fibrosis. *Gut* *61*, 1600–1609.

Okabe, H., Satoh, S., Furukawa, Y., Kato, T., Hasegawa, S., Nakajima, Y., Yamaoka, Y., and Nakamura, Y. (2003). Involvement of PEG10 in human hepatocellular carcinogenesis through interaction with SIAH1. *Cancer Research* 63, 3043–3048.

Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007). The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in *Drosophila*. *Cell* 130, 89–100.

Omari, El, K., Iourin, O., Harlos, K., Grimes, J.M., and Stuart, D.I. (2013). Structure of a Pestivirus Envelope Glycoprotein E2 Clarifies Its Role in Cell Entry. *Cell Reports* 3, 30–35.

Pager, C.T., Schütz, S., Abraham, T.M., Luo, G., and Sarnow, P. (2013). Modulation of hepatitis C virus RNA abundance and virus release by dispersion of processing bodies and enrichment of stress granules. *Virology* 435, 472–484.

Pal, S., Shuhart, M.C., Thomassen, L., Emerson, S.S., Su, T., Feuerborn, N., Kae, J., and Gretch, D.R. (2006). Intrahepatic hepatitis C virus replication correlates with chronic hepatitis C disease severity in vivo. *Journal of Virology* 80, 2280–2290.

Pang, P.S., Pham, E.A., Elazar, M., Patel, S.G., Eckart, M.R., and Glenn, J.S. (2012). Structural Map of a MicroRNA-122:Hepatitis C Virus Complex. *Journal of Virology* 86, 1250–1254.

Parameswaran, P., Sklan, E., Wilkins, C., Burgon, T., Samuel, M.A., Lu, R., Ansel, K.M., Heissmeyer, V., Einav, S., Jackson, W., et al. (2010). Six RNA viruses and forty-one hosts: viral small RNAs and modulation of small RNA repertoires in vertebrate and invertebrate systems. *PLoS Pathog* 6, e1000764.

Patanjali, S.R., Parimoo, S., and Weissman, S.M. (1991). Construction of a Uniform-Abundance (Normalized) Cdna Library. *Proc Natl Acad Sci USA* 88, 1943–1947.

Pedersen, I.M., Cheng, G., Wieland, S., Volinia, S., Croce, C.M., Chisari, F.V., and David, M. (2007). Interferon modulation of cellular microRNAs as an antiviral mechanism. *Nature* 449, 919–922.

Peng, X., Li, Y., Walters, K.-A., Rosenzweig, E.R., Lederer, S.L., Aicher, L.D., Prohl, S., and Katze, M.G. (2009). Computational identification of hepatitis C virus associated microRNA-mRNA regulatory modules in human livers. *BMC Genomics* 10, 373.

Pestova, T.V., Shatsky, I.N., Fletcher, S.P., Jackson, R.J., and Hellen, C.U. (1998). A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev* 12, 67–83.

Petrik, J., Parker, H., and Alexander, G.J. (1999). Human hepatic glyceraldehyde-3-phosphate dehydrogenase binds to the poly(U) tract of the 3' non-coding region of hepatitis C virus genomic RNA. *J. Gen. Virol.* 80 (Pt 12), 3109–3113.

Pfeffer, S., Zavolan, M., Grässer, F.A., Chien, M., Russo, J.J., Ju, J., John, B., Enright, A.J., Marks, D., Sander, C., et al. (2004). Identification of virus-encoded microRNAs. *Science* 304, 734–736.

Phan, T., Beran, R.K.F., Peters, C., Lorenz, I.C., and Lindenbach, B.D. (2009). Hepatitis C Virus NS2 Protein Contributes to Virus Particle Assembly via Opposing Epistatic Interactions with the E1-E2 Glycoprotein and NS3-NS4A Enzyme Complexes. *Journal of Virology* 83, 8379–8395.

Pijlman, G.P., Funk, A., Kondratieva, N., Leung, J., Torres, S., van der Aa, L., Liu, W.J., Palmenberg, A.C., Shi, P.-Y., Hall, R.A., et al. (2008). A Highly Structured, Nuclease-Resistant, Noncoding RNA Produced by Flaviviruses Is Required for Pathogenicity. *Cell Host Microbe* 4, 579–591.

Ploss, A., Evans, M.J., Gaysinskaya, V.A., Panis, M., You, H., de Jong, Y.P., and Rice, C.M. (2009). Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature* 457, 882–886.

Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Research* 20, 110–121.

Prokunina-Olsson, L., Muchmore, B., Tang, W., Pfeiffer, R.M., Park, H., Dickensheets, H., Hergott, D., Porter-Gill, P., Mumy, A., Kohaar, I., et al. (2013). A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus. *Nat. Genet.* 45, 164–171.

Qi, P., Cheng, S.-Q., Wang, H., Li, N., Chen, Y.-F., and Gao, C.-F. (2011). Serum microRNAs as biomarkers for hepatocellular carcinoma in Chinese patients with chronic hepatitis B virus infection. *PLoS ONE* 6, e28486.

Qin, W., Ren, Q., Liu, Te, Huang, Y., and Wang, J. (2013). MicroRNA-155 is a novel suppressor of ovarian cancer-initiating cells that targets CLDN1. *FEBS Letters* 587, 1434–1439.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8, 2281–2308.

Randall, G., Panis, M., Cooper, J.D., Tellinghuisen, T.L., Sukhodolets, K.E., Pfeffer, S., Landthaler, M., Landgraf, P., Kan, S., Lindenbach, B.D., et al. (2007). Cellular cofactors affecting hepatitis C virus infection and replication. *Proc Natl Acad Sci USA* 104, 12884–12889.

Ray, U., and Das, S. (2011). Interplay between NS3 protease and human La protein regulates translation-replication switch of Hepatitis C virus. *Sci. Rep.* 1.

Reddi, H.V., Driscoll, C.B., Madde, P., Milosevic, D., Hurley, R.M., McDonough, S.J., Hallanger-Johnson, J., McIver, B., and Eberhardt, N.L. (2013). Redifferentiation and induction of tumor suppressors miR-122 and miR-375 by the PAX8/PPARgamma fusion protein inhibits anaplastic thyroid cancer: a novel therapeutic strategy. *Cancer Gene Ther* 20, 267–275.

Reddy, S., Jia, S., Geoffrey, R., Lorier, R., Suchi, M., Broeckel, U., Hessner, M.J., and Verbsky, J. (2009). An autoinflammatory disease due to homozygous deletion of the IL1RN locus. *New England Journal of Medicine* 360, 2438–2444.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and RUVKUN, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901–906.

Reiss, S., Rebhan, I., Backes, P., Romero-Brey, I., Erfle, H., Matula, P., Kaderali, L., Poenisch, M., Blankenburg, H., Hiet, M.-S., et al. (2011). Recruitment and Activation of a Lipid Kinase by Hepatitis C Virus NS5A Is Essential for Integrity of the Membranous Replication Compartment. *Cell Host Microbe* 9, 32–45.

Reynolds, J.E., Kaminski, A., Kettinen, H.J., Grace, K., Clarke, B.E., Carroll, A.R., Rowlands, D.J., and Jackson, R.J. (1995). Unique features of internal initiation of hepatitis C virus RNA translation. *Embo J* 14, 6010–6020.

Rijnbrand, R.C., Abbink, T.E., Haasnoot, P.C., Spaan, W.J., and Bredenbeek, P.J. (1996). The influence of AUG codons in the hepatitis C virus 5' nontranslated region on translation and mapping of the translation initiation window. *Virology* 226, 47–56.

Riley, K.J., Rabinowitz, G.S., Yario, T.A., Luna, J.M., Darnell, R.B., and Steitz, J.A. (2012a). EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. *Embo J* 31, 2207–2221.

Riley, K.J., Yario, T.A., and Steitz, J.A. (2012b). Association of Argonaute proteins and microRNAs can occur after cell lysis. *RNA* 18, 1581–1585.

Roberts, A.P.E., Lewis, A.P., and Jopling, C.L. (2011). miR-122 activates hepatitis C virus translation by a specialized mechanism requiring particular RNA components. *Nucleic Acids Res.* *39*, 7716–7729.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* *26*, 139–140.

Rodriguez, A., Vigorito, E., Clare, S., Warren, M.V., Couttet, P., Soond, D.R., van Dongen, S., Grocock, R.J., Das, P.P., Miska, E.A., et al. (2007). Requirement of bic/microRNA-155 for normal immune function. *Science* *316*, 608–611.

Romero-Brey, I., Merz, A., Chiramel, A., Lee, J.-Y., Chlanda, P., Haselman, U., Santarella-Mellwig, R., Habermann, A., Hoppe, S., Kallis, S., et al. (2012). *PLoS Pathog.*

Rosenfeld, A.B., and Racaniello, V.R. (2005). Hepatitis C Virus Internal Ribosome Entry Site-Dependent Translation in *Saccharomyces cerevisiae* Is Independent of Polypyrimidine Tract-Binding Protein, Poly(rC)-Binding Protein 2, and La Protein. *Journal of Virology* *79*, 10126–10137.

Roy, A.L., Chakrabarti, D., Datta, B., Hileman, R.E., and Gupta, N.K. (1988). Natural Messenger-Rna Is Required for Directing Met-Transfer Rna Binding to 40s Ribosomal-Subunits in Animal-Cells - Involvement of Co-Eif-2a in Natural Messenger Rna-Directed Initiation Complex-Formation. *Biochemistry* *27*, 8203–8209.

Ruan, H., and Pownall, H.J. (2001). Overexpression of 1-acyl-glycerol-3-phosphate acyltransferase- α enhances lipid storage in cellular models of adipose tissue and skeletal muscle. *Diabetes* *50*, 233–240.

Rubbia-Brandt, L., Quadri, R., Abid, K., Giostra, E., Malé, P.J., Mentha, G., Spahr, L., Zarski, J.P., Borisch, B., Hadengue, A., et al. (2000). Hepatocyte steatosis is a cytopathic effect of hepatitis C virus genotype 3. *J Hepatol* *33*, 106–115.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature* *448*, 83–86.

Sadler, M.D., and Lee, S.S. (2015). Revolution in hepatitis C antiviral therapy. *Br. Med. Bull.* *113*, 31–44.

Saeed, A., Baloch, K., Brown, R.J.P., Wallis, R., Chen, L., Dexter, L., McClure, C.P., Shakesheff, K., and Thomson, B.J. (2013). Mannan binding lectin-associated serine protease 1 is induced by hepatitis C virus infection and activates human hepatic stellate cells. *Clin Exp Immunol* *174*, 265–273.

Saetrom, P., Heale, B.S.E., Snøve, O., Aagaard, L., Alluin, J., and Rossi, J.J. (2007). Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic Acids Res.* **35**, 2342.

Sagan, S.M., and Sarnow, P. (2013). Molecular biology. RNAi, Antiviral after all. *Science* **342**, 207–208.

Sagan, S.M., Sarnow, P., and Wilson, J.A. (2013). Modulation of GB virus B (GBV-B) RNA abundance by microRNA-122: Dependence on and escape from microRNA-122 restriction. *Journal of Virology*.

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P.P. (2011). A ceRNA Hypothesis: The Rosetta Stone of a Hidden RNA Language? *Cell* **146**, 353–358.

Sarasin-Filipowicz, M., Krol, J., Markiewicz, I., Heim, M.H., Filipowicz, W. (2009). Decreased levels of microRNA miR-122 in individuals with hepatitis C responding poorly to interferon therapy. *Nature Medicine*. **15**, 31–33.

Saunders, M.A., Liang, H., and Li, W.-H. (2007). Human polymorphism at microRNAs and microRNA target sites. *Proc Natl Acad Sci USA* **104**, 3300–3305.

Scheel, T.K.H., and Rice, C.M. (2013). Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. *Nat Med* **19**, 837–849.

Scheel, T.K.H., Kapoor, A., Nishiuchi, E., Brock, K.V., Yu, Y., Andrus, L., Gu, M., Renshaw, R.W., Dubovi, E.J., McDonough, S.P., et al. (2015). Characterization of nonprimate hepacivirus and construction of a functional molecular clone. *Proceedings of the National Academy of Sciences* **112**, 2192–2197.

Scheller, N., Mina, L.B., Galao, R.P., Chari, A., Gimenez-Barcons, M., Noueiry, A., Fischer, U., Meyerhans, A., and Diez, J. (2009). Translation and replication of hepatitis C virus genomic RNA depends on ancient cellular proteins that control mRNA fates. *Proc Natl Acad Sci USA* **106**, 13517–13522.

Schirle, N.T., Sheu-Gruttadauria, J., and MacRae, I.J. (2014). Structural basis for microRNA targeting. *Science* **346**, 608–613.

Schnettler, E., Sterken, M.G., Leung, J.Y., Metz, S.W., Geertsema, C., Goldbach, R.W., Vlak, J.M., Kohl, A., Khromykh, A.A., and Pijlman, G.P. (2012). Noncoding Flavivirus RNA Displays RNA Interference Suppressor Activity in Insect and Mammalian Cells. *Journal of Virology* **86**, 13486–13500.

Schoggins, J.W., and Rice, C.M. (2011). Interferon-stimulated genes and their antiviral effector functions. *Current Opinion in Virology* **1**, 519.

- Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., and Rice, C.M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* **472**, 481–485.
- Schraivogel, D., and Meister, G. (2014). Import routes and nuclear functions of Argonaute and other small RNA-silencing proteins. *Trends Biochem Sci* **39**, 420–431.
- Sedano, C.D., and Sarnow, P. (2014). Hepatitis C virus subverts liver-specific miR-122 to protect the viral genome from exoribonuclease Xrn2. *Cell Host Microbe* **16**, 257–264.
- Seitz, H. (2009). Redefining MicroRNA Targets. *Curr Biol* **19**, 870-873.
- Sekiyama, K.D., Yoshiba, M., and Thomson, A.W. (1994). Circulating proinflammatory cytokines (IL-1 beta, TNF-alpha, and IL-6) and IL-1 receptor antagonist (IL-1Ra) in fulminant hepatic failure and acute hepatitis. *Clin Exp Immunol* **98**, 71–77.
- Sendi, H., Mehrab-Mohseni, M., Foureau, D.M., Ghosh, S., Walling, T.L., Steuerwald, N., Zamor, P.J., Kaplan, K.J., Jacobs, C., Ahrens, W.A., et al. (2014). miR-122 decreases HCV entry into hepatocytes through binding to the 3' UTR of OCLN mRNA. *Liver Int* **35**, 1315-1323.
- Sharma, N.R., Mateu, G., Dreux, M., Grakoui, A., Cosset, F.-L., and Melikyan, G.B. (2011). Hepatitis C Virus Is Primed by CD81 Protein for Low pH-dependent Fusion. *J. Biol. Chem.* **286**, 30361–30376.
- Sheahan, T., Imanaka, N., Marukian, S., Dorner, M., Liu, P., Ploss, A., and Rice, C.M. (2014). Interferon Lambda Alleles Predict Innate Antiviral Immune Responses and Hepatitis C Virus Permissiveness. *Cell Host Microbe* **15**, 190–202.
- Shi, X., Sun, M., Liu, H., Yao, Y., and Song, Y. (2013). Long non-coding RNAs: A new frontier in the study of human diseases. *Cancer Lett* **339**, 159–166.
- Shimakami, T., Yamane, D., Jangra, R.K., Kempf, B.J., Spaniel, C., Barton, D.J., and Lemon, S.M. (2012). Stabilization of hepatitis C virus RNA by an Ago2-miR-122 complex. *Proc Natl Acad Sci USA* **109**, 941–946.
- Shulla, A., and Randall, G. (2012). Hepatitis C virus-host interactions, replication, and viral assembly. *Current Opinion in Virology* **2**, 725–732.
- Silber, R., Malathi, V.G., and Hurwitz, J. (1972). Purification and properties of bacteriophage T4-induced RNA ligase. *Proc Natl Acad Sci USA* **69**, 3009–3013.

Simmonds, P., Bukh, J., Combet, C., Deléage, G., Enomoto, N., Feinstone, S., Halfon, P., Inchauspé, G., Kuiken, C., Maertens, G., et al. (2005). Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology* 42, 962–973.

Singaravelu, R., Russell, R.S., Tyrrell, D.L., and Pezacki, J.P. (2014). ScienceDirectHepatitis C virus and microRNAs: miRed in a host of possibilities. *Current Opinion in Virology* 7, 1–10.

Skalsky, R.L.R., Corcoran, D.L.D., Gottwein, E.E., Frank, C.L.C., Kang, D.D., Hafner, M.M., Nusbaum, J.D.J., Feederle, R.R., Delecluse, H.-J.H., Luftig, M.A.M., et al. (2012). The viral and cellular microRNA targetome in lymphoblastoid cell lines. *PLoS Pathog* 8, e1002484–e1002484.

Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and RUVKUN, G. (2000). The lin-41 RBCC gene acts in the *C. elegans* heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Molecular Cell* 5, 659–669.

Smith, D.B., Bukh, J., Kuiken, C., Muerhoff, A.S., Rice, C.M., Stapleton, J.T., and Simmonds, P. (2014). Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. *Hepatology* 59, 318–327.

Sohlenius-Sternbeck, A.-K. (2006). Determination of the hepatocellularity number for human, dog, rabbit, rat and mouse livers from protein concentration measurements. *Toxicol in Vitro* 20, 1582–1586.

Solomon, M.J., and Varshavsky, A. (1985). Formaldehyde-Mediated Dna Protein Crosslinking - a Probe for Invivo Chromatin Structures. *Proc Natl Acad Sci USA* 82, 6470–6474.

Song, K., Han, C., Zhang, J., Lu, D., Dash, S., Feitelson, M., Lim, K., and Wu, T. (2013). Epigenetic regulation of MicroRNA-122 by peroxisome proliferator activated receptor-gamma and hepatitis b virus X protein in hepatocellular carcinoma cells. *Hepatology* 58, 1681–1692.

Spångberg, K., Wiklund, L., and Schwartz, S. (2000). HuR, a Protein Implicated in Oncogene and Growth Factor mRNA Decay, Binds to the 3' Ends of Hepatitis C Virus RNA of Both Polarities. *Virology* 274, 378–390.

Stark, A., Brennecke, J., Bushati, N., Russell, R.B., and Cohen, S.M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* 123, 1133–1146.

Starkey Lewis, P.J., Dear, J., Platt, V., Simpson, K.J., Craig, D.G.N., Antoine, D.J., French, N.S., Dhaun, N., Webb, D.J., Costello, E.M., et al. (2011). Circulating microRNAs as potential markers of human drug-induced liver injury. *Hepatology* 54, 1767–1776.

Stevens, J.G., Wagner, E.K., Devi-Rao, G.B., Cook, M.L., and Feldman, L.T. (1987). RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. *Science* 235, 1056–1059.

Sumpter, R., Loo, Y.-M., Foy, E., Li, K., Yoneyama, M., Fujita, T., Lemon, S.M., and Gale, M. (2005). Regulating intracellular antiviral defense and permissiveness to hepatitis C virus RNA replication through a cellular RNA helicase, RIG-I. *Journal of Virology* 79, 2689–2699.

Sun, C., Huang, F., Liu, X., Xiao, X., Yang, M., Hu, G., Liu, H., and Liao, L. (2015). miR-21 regulates triglyceride and cholesterol metabolism in non-alcoholic fatty liver disease by targeting HMGCR. *Int J Mol Med* 35, 847–853.

Takeuchi, K., and Reue, K. (2009). Biochemistry, physiology, and genetics of GPAT, AGPAT, and lipin enzymes in triglyceride synthesis. *Am J Physiol Endocrinol Metab* 296, E1195–E1209.

Tang, K.H., Ma, S., Lee, T.K., Chan, Y.P., Kwan, P.S., Tong, C.M., Ng, I.O., Man, K., To, K.-F., Lai, P.B., et al. (2012). CD133(+) liver tumor-initiating cells promote tumor angiogenesis, growth, and self-renewal through neurotensin/interleukin-8/CXCL1 signaling. *Hepatology* 55, 807–820.

Tanguy, M., and Miska, E.A. (2013). Antiviral RNA interference in animals: piecing together the evidence. *Nat. Struct. Mol. Biol.* 20, 1239–1241.

Taylor, D.R., Shi, S.T., Romano, P.R., Barber, G.N., and Lai, M.M. (1999). Inhibition of the interferon-inducible protein kinase PKR by HCV E2 protein. *Science* 285, 107–110.

Tennant, B.C., and Gerin, J.L. (2001). The woodchuck model of hepatitis B virus infection. *Hepatology* 33, 89–102.

Theodore Peters, J. (1995). *All About Albumin* (Academic Press).

Thomas, D.L., Astemborski, J., Rai, R.M., Anania, F.A., Schaeffer, M., Galai, N., Nolt, K., Nelson, K.E., Strathdee, S.A., Johnson, L., et al. (2000). The natural history of hepatitis C virus infection - Host, viral, and environmental factors. *Jama* 284, 450–456.

Thomas, D.L. (2013). Global control of hepatitis C: where challenge meets opportunity. *Nat Med* 19, 850–858.

Tilg, H., Wilmer, A., Vogel, W., Herold, M., Nölchen, B., Judmaier, G., and Huber, C. (1992). Serum levels of cytokines in chronic liver diseases. *Ygast* 103, 264–274.

Triboulet, R., Mari, B., Lin, Y.-L., Chable-Bessia, C., Bennasser, Y., Lebrigand, K., Cardinaud, B., Maurin, T., Barbry, P., Baillat, V., et al. (2007). Suppression of microRNA-silencing pathway by HIV-1 during virus replication. *Science* 315, 1579–1582.

Trobaugh, D.W., Gardner, C.L., Sun, C., Haddow, A.D., Wang, E., Chapnik, E., Mildner, A., Weaver, S.C., Ryman, K.D., and Klimstra, W.B. (2013). RNA viruses can hijack vertebrate microRNAs to suppress innate immunity. *Nature* 1–15.

Tsai, W.-C., Hsu, P.W.-C., Lai, T.-C., Chau, G.-Y., Lin, C.-W., Chen, C.-M., Lin, C.-D., Liao, Y.-L., Wang, J.-L., Chau, Y.-P., et al. (2009). MicroRNA-122, a tumor suppressor microRNA that regulates intrahepatic metastasis of hepatocellular carcinoma. *Hepatology* 49, 1571–1582.

Tsai, W.-C., Hsu, S.-D., Hsu, C.-S., Lai, T.-C., Chen, S.-J., Shen, R., Huang, Y., Chen, H.-C., Lee, C.-H., Tsai, T.-F., et al. (2012). MicroRNA-122 plays a critical role in liver homeostasis and hepatocarcinogenesis. *J. Clin. Invest.* 122, 2884–2897.

Tsou, A.-P., Chuang, Y.-C., Su, J.-Y., Yang, C.-W., Liao, Y.-L., Liu, W.-K., Chiu, J.-H., and Chou, C.-K. (2003). Overexpression of a novel imprinted gene, PEG10, in human hepatocellular carcinoma and in regenerating mouse livers. *J Biomed Sci* 10, 625–635.

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M., and Nomoto, A. (1992). Internal ribosome entry site within hepatitis C virus RNA. *Journal of Virology* 66, 1476–1483.

Ueda, H.R., Chen, W., Adachi, A., Wakamatsu, H., Hayashi, S., Takasugi, T., Nagano, M., Nakahama, K.-I., Suzuki, Y., Sugano, S., et al. (2002). A transcription factor response element for gene expression during circadian night. *Nature* 418, 534–539.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302, 1212–1215.

Umbach, J.L., Kramer, M.F., Jurak, I., Karnowski, H.W., Coen, D.M., and Cullen, B.R. (2008). MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature* 454, 780–783.

- Umbach, J.L., and Cullen, B.R. (2009). The role of RNAi and microRNAs in animal virus replication and antiviral immunity. *Genes Dev* 23, 1151–1164.
- Vaillancourt, F.H., Brault, M., Pilote, L., Uyttersprot, N., Gaillard, E.T., Stoltz, J.H., Knight, B.L., Pantages, L., McFarland, M., Breitfelder, S., et al. (2012). Evaluation of phosphatidylinositol-4-kinase III α as a hepatitis C virus drug target. *Journal of Virology* 86, 11607.
- van der Ree, M.H., van der Meer, A.J., de Bruijne, J., Maan, R., van Vliet, A., Welzel, T.M., Zeuzem, S., Lawitz, E.J., Rodriguez-Torres, M., Kupcova, V., et al. (2014). Long-term safety and efficacy of microRNA-targeted therapy in chronic hepatitis C patients. *Antiviral Res* 111, 53–59.
- Vanharanta, S., Marney, C.B., Shu, W., Valiente, M., Zou, Y., Mele, A., Darnell, R.B., and Massague, J. (2014). Loss of the multifunctional RNA-binding protein RBM47 as a source of selectable metastatic traits in breast cancer. *eLife* 3.
- Vasudevan, S., Tong, Y., and Steitz, J.A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. *Science* 318, 1931–1934.
- Vereide, D.T., Seto, E., Chiu, Y.-F., Hayes, M., Tagawa, T., Grundhoff, A., Hammerschmidt, W., and Sugden, B. (2013). Epstein–Barr virus maintains lymphomas via its miRNAs. *Oncogene* 33, 1258–1264.
- Vidigal, J.A., and Ventura, A. (2015). The biological functions of miRNAs: lessons from in vivo studies. *Trends in Cell Biology* 25, 137–147.
- Vila-Brau, A., De Sousa-Coelho, A.L., Goncalves, J.F., Haro, D., and Marrero, P.F. (2013). Fsp27/CIDEA is a CREB target gene induced during early fasting in liver and regulated by FA oxidation rate. *J Lipid Res* 54, 592–601.
- Villano, S.A., Vlahov, D., Nelson, K.E., Cohn, S., and Thomas, D.L. (1999). Persistence of viremia and the importance of long-term follow-up after acute hepatitis C infection. *Hepatology* 29, 908–914.
- Waidmann, O., Bihrer, V., Pleli, T., Farnik, H., Berger, A., Zeuzem, S., Kronenberger, B., and Piiper, A. (2012). Serum microRNA-122 levels in different groups of patients with chronic hepatitis B virus infection. *J. Viral Hepat.* 19, e58–e65.
- Wakita, T., Pietschmann, T., Kato, T., Date, T., Miyamoto, M., Zhao, Z., Murthy, K., Habermann, A., Kräusslich, H.-G., Mizokami, M., et al. (2005). Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat Med* 11, 791–796.

Wang, B., Hsu, S.-H., Wang, X., Kutay, H., Bid, H.K., Yu, J., Ganju, R.K., Jacob, S.T., Yuneva, M., and Ghoshal, K. (2013a). Reciprocal regulation of microRNA-122 and c-Myc in hepatocellular cancer: Role of E2F1 and transcription factor dimerization partner 2. *Hepatology* 59, 555–566.

Wang, C., Sarnow, P., and Siddiqui, A. (1993). Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *Journal of Virology* 67, 3338–3344.

Wang, E.T., Cody, N.A.L., Jog, S., Biancolella, M., Wang, T.T., Treacy, D.J., Luo, S., Schroth, G.P., Housman, D.E., Reddy, S., et al. (2012a). Transcriptome-wide Regulation of Pre-mRNA Splicing and mRNA Localization by Muscleblind Proteins. *Cell* 150, 710–724.

Wang, F., Yang, F., Zhang, L., and Sun, S. (2010). Is increasing the copy number of the microRNA precursor in a vector really helpful for further improving the level of mature microRNA in stable cells? *Hepatology* 52, 1517–authorreply1518.

Wang, L., Jeng, K.-S., and Lai, M.M.C. (2011). Poly(C)-binding protein 2 interacts with sequences required for viral replication in the hepatitis C virus (HCV) 5' untranslated region and directs HCV RNA replication through circularizing the viral genome. *Journal of Virology* 85, 7964.

Wang, S., Qiu, L., Yan, X., Jin, W., Wang, Y., Chen, L., Wu, E., Ye, X., Gao, G.F., Wang, F., et al. (2012b). Loss of microRNA 122 expression in patients with hepatitis B enhances hepatitis B virus replication through cyclin G(1) -modulated P53 activity. *Hepatology* 55, 730–741.

Wang, X., Lam, E.K.Y., Zhang, J., Jin, H., and Sung, J.J.Y. (2009). MicroRNA-122a functions as a novel tumor suppressor downstream of adenomatous polyposis coli in gastrointestinal cancers. *Biochem Biophys Res Commun* 387, 376–380.

Wang, Y., Xu, Z., Jiang, J., Xu, C., Kang, J., Xiao, L., Wu, M., Xiong, J., Guo, X., and Liu, H. (2013b). Endogenous miRNA Sponge lincRNA-RoR Regulates Oct4, Nanog, and Sox2 in Human Embryonic Stem Cell Self-Renewal. *Developmental Cell* 25, 69–80.

Watashi, K. (2003). Cyclosporin A suppresses replication of hepatitis C virus genome in cultured hepatocytes. *Hepatology* 38, 1282–1288.

Weakley, S.M., Wang, H., Yao, Q., and Chen, C. (2011). Expression and Function of a Large Non-coding RNA Gene XIST in Human Cancer. *World J Surg* 35, 1751–1756.

Weinlich, S., Hüttelmaier, S., Schierhorn, A., Behrens, S.-E., Ostareck-Lederer, A., and Ostareck, D.H. (2009). IGF2BP1 enhances HCV IRES-mediated translation initiation via the 3'UTR. *RNA* 15, 1542.

Weyn-Vanhentenryck, S.M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., Xue, C., Herre, M., Silver, P.A., Zhang, M.Q., et al. (2014). HITS-CLIP and Integrative Modeling Define the Rbfox Splicing-Regulatory Network Linked to Brain Development and Autism. *CellReports* 6, 1139–1152.

Whisnant, A.W., Kehl, T., Bao, Q., Materniak, M., Kuzmak, J., Lochelt, M., and Cullen, B.R. (2014). Identification of Novel, Highly Expressed Retroviral MicroRNAs in Cells Infected by Bovine Foamy Virus. *Journal of Virology* 88, 4679–4686.

Wilson, J.A., Zhang, C., Huys, A., and Richardson, C.D. (2011). Human Ago2 is required for efficient microRNA 122 regulation of hepatitis C virus RNA accumulation and translation. *Journal of Virology* 85, 2342–2350.

Wilson, S.J., Schoggins, J.W., Zang, T., Kutluay, S.B., Jouvenet, N., Alim, M.A., Bitzegeio, J., Rice, C.M., and Bieniasz, P.D. (2012). Inhibition of HIV-1 particle assembly by 2',3'-cyclic-nucleotide 3'-phosphodiesterase. *Cell Host Microbe* 12, 585–597.

Wozniak, A.L., Griffin, S., Rowlands, D., Harris, M., Yi, M., Lemon, S.M., and Weinman, S.A. (2010). Intracellular Proton Conductance of the Hepatitis C Virus p7 Protein and Its Contribution to Infectious Virus Production. *PLoS Pathog* 6, e1001087.

Wu, X., Somlo, G., Yu, Y., Palomares, M.R., Li, A.X., Zhou, W., Chow, A., Yen, Y., Rossi, J.J., Gao, H., et al. (2012). De novo sequencing of circulating miRNAs identifies novel markers predicting clinical outcome of locally advanced breast cancer. *J Transl Med* 10, 42.

Xiong, Y., Zhang, C., Yuan, J., Zhu, Y., Tan, Z., Kuang, X., and Wang, X. (2014). Hepatitis C virus represses the cellular antiviral response by upregulating the expression of signal transducer and activator of transcription 3 through sponging microRNA-122. *Mol Med Rep*.

Xu, H., He, J.-H., Xiao, Z.-D., Zhang, Q.-Q., Chen, Y.-Q., Zhou, H., and Qu, L.-H. (2010). Liver-enriched transcription factors regulate microRNA-122 that targets CUTL1 during liver development. *Hepatology* 52, 1431–1442.

Xu, J., Zhang, R., Shen, Y., Liu, G., Lu, X., and Wu, C.I. (2013). The evolution of evolvability in microRNA target sites in vertebrates. *Genome Research* 23, 1810–1816.

- Xu, J., Zhu, X., Wu, L., Yang, R., Yang, Z., Wang, Q., and Wu, F. (2012). MicroRNA-122 suppresses cell proliferation and induces cell apoptosis in hepatocellular carcinoma by directly targeting Wnt/beta-catenin pathway. *Liver Int* 32, 752–760.
- Zhang, C., Huys, A., Thibault, P.A., and Wilson, J.A. (2012). Requirements for human Dicer and TRBP in microRNA-122 regulation of HCV translation and RNA abundance. *Virology* 433, 479–488.
- Zhang, C., and Darnell, R.B. (2011). Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat. Biotechnol.* 29, 607–614.
- Zhang, Y., Jia, Y., Zheng, R., Guo, Y., Wang, Y., Guo, H., Fei, M., and Sun, S. (2010). Plasma microRNA-122 as a biomarker for viral-, alcohol-, and chemical-related hepatic diseases. *Clin Chem* 56, 1830–1838.
- Zhao, L., Li, F., and Taylor, E.W. (2013). Can tobacco use promote HCV-induced miR-122 hijacking and hepatocarcinogenesis? *Med Hypotheses* 80, 131–133.
- Zhuang, F., Fuchs, R.T., Sun, Z., Zheng, Y., and Robb, G.B. (2012). Structural bias in T4 DNA ligase-mediated 3'-adapter ligation. *Nucleic Acids Res.* 40, e54–e54.
- Zou, C., Zhang, Z., Wu, S., and Osterman, J.C. (1998). Molecular cloning and characterization of a rabbit eIF2C protein. *Gene* 211, 187–194.
- Ørom, U.A., Nielsen, F.C., and Lund, A.H. (2008). MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Molecular Cell* 30, 460–471.