IDENTIFICATION OF NOVEL NATURALLY OCCURRING GENETIC INTERACTIONS USING INTER-PERSON VARIATION AND VIRUS DYNAMICS IN HEPATITIS C VIRUS AND HUMAN IMMUNODEFICIENCY VIRUS INFECTIONS.

by
Ramy El-Diwany

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
August, 2015

Abstract

Hepatitis C virus (HCV) and human immunodeficiency virus (HIV) are two chronic viral infections of humans for which no vaccine exists. considerable research effort has focused on *in vitro, ex vivo,* and animal models due to inherent limitations in human research. The advent of highly quantitative and high-throughput measurement modalities from increasingly small amounts of human tissue have made it possible to generate unprecedented amounts of information from the material contained within a few drops of blood. This thesis presents three novel methods in which blood sampling of humans with HCV or HIV has generated large data sets that were then analyzed using the latest technologies and computational methods to describe previously unappreciated biological features of those infections.

This thesis explores two important aspects of HCV infection for which no surrogate for natural infection can be as informative: i) plasma microRNA signatures before, during, and after acute HCV infection using microRNA quantitative PCR array and ii) naturally occurring resistance polymorphisms in the HCV envelope to broadly neutralizing antibodies using a large panel of HCV pseudoparticles. This thesis also identifies novel restriction factors for HIV using RNA sequencing (RNAseq) technology to investigate naturally occurring inter-person variation in gene-induction in activated CD4+ T-cells of patients receiving an injection of pegylated interferon-alpha.

Thesis Advisors: Stuart Ray M.D. and Robert Siliciano M.D., Ph.D.

Thesis Readers: Ashwin Balagopal M.D. and Stuart Ray M.D.

Acknowledgements

The work presented here would not be possible without the generous support of the faculty, staff, and student members of the BCMB Graduate Training Program, Medical Scientist Training Program, and the Viral Hepatitis Research Center of the Johns Hopkins University School of Medicine. The author would like to thank Geraldine Seydoux, Robert F. Siliciano, Stuart C. Ray, David L. Thomas, and Ashwin Balagopal for formative discussion and genuine dedication to student training as members of his PhD thesis committee. The author would also like to acknowledge Justin R. Bailey for guidance with regards to the work performed on broadly neutralizing antibodies for HCV and continued willingness to marvel at the biology of infectious disease in humans.

Additionally, the author would like to thank the participants of the BBAASH, IFN PK, and UFO studies for contributing invaluable samples, William Osburn for expertise with regards to sampling BBAASH, Roxann Ashworth for assistance with execution of qPCR array profiling, Kenneth Witwer for assistance with analyzing microRNA array data, David Mohr of The Johns Hopkins Genetics Research Core Facility for expertise in library preparation and optimal sequencing, Hao Zhang for helpful discussion and assistance with designing FACS protocols, David Levine for perspective in all aspects of training, Joel Blankson for support with optimizing HIV-1 infection sampling and continuous mentorship through the MD-PhD training process, Sarah Wheelan for discussion and

encouragement with regards to computational methods, Supriya Munshaw for introduction to R, Candellaria Coggiano for perspective on large data sets, Paul Scherer for helpful discussion and reagents, Neil Neumann for helpful discussion and reagents, and the organizers and instructors of the Quantitative Systems Immunology Summer School at Boston University and the Computational Immunology seminar series at Johns Hopkins University School of Medicine for introduction to the computational methods implemented in this thesis. The author would like to thank the National Institutes of Health for their financial support through grants R01 DA 016078, U19 AI 088791, R37 DA 013806, 5R01 DA 031056, and 2R01 DA 016017.

Most importantly, the author would like to thank his wife Katherine and son Wesley for lovingly allowing him to chase answers to questions addressed inside and outside of this thesis.

Table of Contents

Background: Hepatitis C virus (HCV) and Human Immunodeficiency Virus-1 (HIV-1) are viruses with distinctly human pathogenesis. 1

CHAPTER 1: Acute HCV infection induces consistent changes in circulating miRNAs that are associated with non-lytic hepatocyte release 9

- **Abstract 9**
- **Introduction 11**
- **Materials and Methods 13**
- **Results 21**
- **Discussion 35**

CHAPTER 2: A polymorphism in hypervariable region 1 confers resistance to broadly neutralizing antibodies targeting opposite termini of the hepatitis C virus E2 protein. 42

- **Abstract 42**
- **Introduction 44**
- **Materials and Methods 46**
- **Results 51**
- **Discussion 63**

Chapter 3: Identification of novel interferon-induced restriction factors for HIV-1 from inter-human variation in activated T-cell transcription and virus dynamics *in vivo*. 68

- **Abstract 68**
- **Materials and Methods 72**
- **Discussion 99**

Conclusions 107

Appendix 110

- **References 110**
- **Abbreviations 129**
- **Curriculum Vitae 130**

This page is intentionally blank.

Background: Hepatitis C virus (HCV) and Human Immunodeficiency Virus-1 (HIV-1) are viruses with distinctly human pathogenesis.

Viruses are small infectious entities that replicate inside living cells of host organisms, including single-celled hosts such as amoeba and bacteria. The mechanisms host organisms employ to protect them from pathogenic infections are likely as ancient as the existence of cellular life. The endosymbiotic history of mitochondria and the more recent observation that ~8% of the human genome is retroviral DNA have challenged previous notions of what defines organism, flora, and pathogen. The genetic ecology of life facilitates evolution in ways that we have only been able to appreciate recently with the widespread availability of genome sequencing data. These analyses have revealed duplication and divergence of genes with antiviral properties among mammals (1) and more recent emerging evidence suggests that profound selection for genes related to antiviral immunity have occurred geographically among humans (2).

Many viruses that are pathogenic in humans also replicate in other organisms ranging from mosquitos to chimpanzees; however, infection in these other hosts is often less pathogenic and symptomatic. Often analogues of human viral infections exist in other species and useful comparisons can be made. In contrast to humans, these analogous infections may not cause disease yet important information can be learned from studying the presumed adaptations, either by the host or the virus,

that confer reduced pathogenesis. While these model systems have the potential to teach about important features of infections of humans, direct understanding of the pathogenesis of viral infections in humans remains challenging because it is unethical to perform the experiments used to describe the biology of infection and pathogenesis in other hosts. For this reason in particular, this thesis explores novel methods of data collection and analysis from naturally occurring human viral infections to fill in knowledge gaps regarding infection in this context.

The natural history of hepatitis C virus.

HCV is a hepatotropic single-stranded positive-sense RNA virus that infects over 170 million people worldwide. Acute HCV infection progresses to chronic infection in ~75% of cases (3–5). Of those who become infected, a large fraction will never experience any symptomatic manifestations of their disease; however, many will progress through a disease course characterized by chronic inflammation and the progressive deposition of fibrotic tissue within the liver, ultimately resulting in cirrhosis and, in some, end-stage liver disease and hepatocellular carcinoma.

Very little is understood about why certain individuals progress along a particular branch point in the natural history of infection. The most notable predictors are co-infection with HIV, *IFNL3* genotype, age, race, and gender; however, they collectively explain only a minority of the natural variation observed. Chimpanzees (*Pan troglodytes*) and northern

treeshrews (*Tupaia belangeri*) remain the only known organisms besides humans to support HCV replication naturally. While interesting surrogates for HCV infection in humans, their biological relevance and practicality have made them unpopular. Curiously, relatively high rates of spontaneous resolution and low levels of liver inflammation and disease occur with HCV infection of chimpanzees when compared to humans. Recently, national policy changes regarding the use of chimpanzees in research have disfavored their use in this setting. Overt pathogenesis has been observed with HCV in northern treeshrews (6); however, very rarely is sustained viremia observed and they are difficult to breed in captivity. Without a representative animal model for HCV, observation and molecular dissection of naturally occurring infections has the highest potential for illuminating biology that occurs at the system level such as pathogenesis.

The natural history of human immunodeficiency virus-1.

HIV-1 is a lentivirus that infects over 33 million people worldwide (7). If left untreated, infection with HIV leads to the acquired-immunodeficiency syndrome (AIDS) characterized by a dysfunctional immune system that renders the host more susceptible to life-threatening opportunistic infections and malignancy. Initial infection is asymptomatic in many cases but in some, manifests in a mononucleosis-like illness; during this acute stage high amounts of virus can be found in the blood. CD4+ T-

cells become infected at an exponential rate, and their resulting death causes a precipitous drop in their abundance in the blood (~50%). Though the host's immune system rarely controls infection to very low or undetectable amounts of virus, the host limits replication by ~10-100-fold within a few months and CD4+ T-cell abundance begins to rebound. In this asymptomatic phase, CD4+ T-cell counts remain depressed compared to pre-infection levels as replication continues. Eventually the destruction of CD4+ T-cells outpaces the host's ability to produce them and their numbers drop, leaving the host vulnerable to the infections and malignancies that define AIDS.

Lentiviruses similar to HIV exist for non-human primates and are referred to as simian immunodeficiency viruses (SIV). In their natural hosts, these viruses rarely cause immunodefficiency. Interestingly, when the variant of SIV that infects one primate species is introduced into another species of primate, disease features similar to AIDS in humans may emerge. Studying these model infections has revealed numerous insights into how immunodeficiency arises; however, extrapolation of these interactions to the biology of HIV infection in humans is difficult given the important differences that exist between the hosts and the viruses.

As we will discuss in chapter three, some of the most informative genetic interactions between humans and HIV-1 are interferon-induced immune genes that restrict viral production in host cells. Strong evidence

exists that genes that are associated with the immune system are some of the most rapidly changing genes in mammalian genomes(8). As such, many of the genetic interactions between host and virus may be not informative in the context of non-human primates. As was the case with HCV, limited experimental options and the distinct host-pathogen interactions in humans underscore the importance of studying molecular interactions between virus and host in the natural infection.

Studying humans

For the reasons discussed above and more, investigation of humans infected with HCV or HIV is necessary to understand the interactions (genetic or otherwise) that occur between host and pathogen. There are many approaches one can take to making these sorts of associations in humans with a viral infection. One approach is to profile many individuals currently infected, detail all of their relevant metrics, and survey these parameters for associations. This approach also referred to as cross-sectional, has the advantage of relatively simple data collection enabling the investigator to assemble large amounts of data at once to make important associations, especially when resources and time are very limited. This approach is not without limitations. Most importantly, because people are so different from one another prior to infection, large numbers of individuals must be observed so that changes associated with the

infection can be appreciated amidst these person-to-person differences unrelated to infection.

Another approach is to follow people longitudinally. Although resource intensive, cohort-based investigation allows the investigator to account for person-to-person differences in baseline gene expression by enabling the measurement of changes that occur through infection or treatment. As such, many fewer samples are needed to make significant associations. If for instance, gene A is induced by infection with HCV but its baseline expression varies dramatically from person to person, one can sample a person before and after infection and examine the *change* in the expression of gene A to determine if its expression is significantly associated with HCV infection. On the other hand, many more uninfected and infected people would need to be profiled cross-sectionally to appreciate a difference in levels of gene A between infected and uninfected individuals. Also, longitudinal data allows for more inference about causality, since temporal structure exists (a key postulate of causality). Another feature of longitudinal sampling is that it systematically adjusts for all of the variables that are not measured since those variables are largely constant over short durations in the same people. In chapters one and three we will use longitudinal sampling strategies to understand how infection- or treatment-induced changes can reveal potentially important genetic features of the host.

In chapter one, we will sample individuals throughout differing courses of acute infection with HCV and examine microRNA expression changes in plasma and uncover novel features of these regulatory molecules that arise in infection. In chapter two, we will examine naturally occurring variation in HCV envelope sequence and sensitivity to broadly neutralizing antibodies (key molecules for vaccine design) and identify an unexpected route of resistance. Finally in chapter three, we will compare HIV-1 virus dynamics with gene induction in response to interferon alpha treatment to identify novel restriction factors for HIV-1.

This page is intentionally blank.

CHAPTER 1: Acute HCV infection induces consistent changes in circulating miRNAs that are associated with non-lytic hepatocyte release

Abstract

Plasma microRNAs (miRNAs) change in abundance in response to disease, and have been associated with liver fibrosis severity in chronic hepatitis C virus (HCV) infection. However, the early dynamics of miRNA release during acute HCV infection are poorly understood. In addition, circulating miRNA signatures have been difficult to reproduce among separate populations. We studied plasma miRNA abundance during acute HCV infection to identify a miRNA signature of early infection. We measured 754 plasma miRNAs by quantitative PCR array in a discovery cohort of 22 individuals before and during acute HCV infection and after spontaneous resolution (n=11) or persistence (n=11) to identify a plasma miRNA signature. The discovery cohort derived from Baltimore Before and After Acute Study of Hepatitis. During acute HCV infection, increases in miR-122 ($P<.01$) and miR-885-5p ($P_{corrected}<.05$) and a decrease in miR-494 ($P_{corrected}<.05$) were observed at the earliest time-points after virus detection. Changes in miR-122 and miR-885-5p were sustained in persistent ($P<.001$) but not resolved HCV infection. The circulating miRNA signature of acute HCV infection was confirmed in a separate validation cohort that derived from the San Francisco-based UFO Study (n=28). As further confirmation, cellular changes of signature miRNAs were examined

in a tissue-culture model of HCV in hepatoma cells: HCV infection induced extracellular release of miR-122 and miR-885-5p despite unperturbed intracellular levels. In contrast, miR-494 accumulated intracellularly ($P<.05$). Collectively, these data are inconsistent with necrolytic release of hepatocyte miRNAs into the plasma during acute HCV infection of humans.

Introduction

Hepatitis C virus (HCV) is a single-stranded positive-sense hepatotropic RNA virus that infects over 170 million people worldwide. Acute HCV infection progresses to chronic infection in ~75% of cases (3–5). Little is understood about early host-pathogen interactions due to a lack of representative animal models and the difficulty in studying acute HCV, which is largely asymptomatic.

MicroRNAs (miRNAs) are 18-25 nucleotide non-coding RNA molecules with widespread gene-network regulatory capacity. miRNAs are readily detectable in cell-free components of blood, and circulating extracellular miRNAs have been associated with the severity and progression of numerous human diseases including hepatocellular carcinoma (HCC) (9, 10), liver injury (11), colorectal cancer (12), and cardiac injury (13). Little is understood about the early dynamics and release of circulating miRNAs in specific disease conditions. Importantly, circulaing miRNA signatures have been challenging to confirm in separate cohorts, raising questions as to whether consistent changes in circulating miRNAs occur (14, 15).

miRNAs have profound intracellular effects, and have been described to play a role in viral pathogenesis (16, 17). Indeed, HCV replication requires the hepatocyte-specific miRNA miR-122 for efficient replication (18), and a miR-122 antagonist was shown to have potent

antiviral effects in a human trial (19). Whereas conventional wisdom predicts that extracellular miRNAs found circulating in the plasma are the result of cell lysis, extracellular miRNAs have been shown to be selectively secreted by “donor” cells and can regulate genes in target cells in animal and tissue culture models (20–30). We hypothesized that acute HCV infection leads to regulated changes in the abundance of circulating microRNAs. We used multiple approaches to validate the circulating miRNA signature.

Materials and Methods

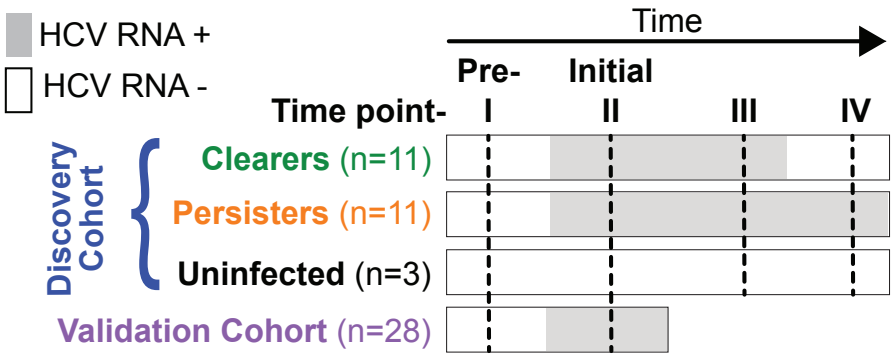

Figure 1-1. Study design. miRNAs were quantified in plasma samples from injection drug users in the discovery cohort who acquired acute HCV infection and either cleared infection (clearers; n=11) or developed persistent infection (persisters; n=11). miRNAs were also quantified in plasma samples from 3 participants in the discovery cohort who did not develop infection and who had available samples at time intervals that closely matched the intervals in the acutely infected participants. A separate validation cohort (n=28) was studied to confirm the findings in the discovery cohort.

Study population. The principal measurements in this study were performed using samples from a discovery cohort. Major findings were confirmed in a separate validation cohort. The discovery cohort was composed of 22 participants in the *Baltimore Before and After Acute Study of Hepatitis* (BBAASH) cohort, an ongoing study of injection drug users (IDU) who are at high risk for acquiring HCV infection (4). Eligible participants have a history of injection drug use and are negative for anti-HCV antibodies and HCV RNA at the time of enrollment. All enrollees are assessed for HBV and HIV status at enrollment, vaccinated for HBV if no evidence of previous vaccination, and monitored regularly for seroconversion at regular intervals as previously described. Written consent is obtained from each participant. Once enrolled, participants

receive counseling to reduce drug use and its complications. Phlebotomy is performed monthly, as previously described (4). Participant infection status is monitored by quantitative PCR (qPCR) for HCV RNA, and those with acute HCV infection are referred for evaluation for treatment. The validation cohort was composed of 28 participants in the San Francisco-based UFO Study (*You Find Out)*, a prospective study of young IDU at high-risk for acquiring HCV infection in whom incident acute HCV infection is identified and followed monthly (31). Written consent was obtained from each participant. The Institutional Review Board (IRB) of Johns Hopkins University School of Medicine approved the BBAASH study, and the IRB of the University of California at San Francisco approved the UFO Study.

Sample selection. Plasma samples were selected from participants for whom acute HCV infection and its outcome were fully documented. Four samples were examined over structured time intervals for each participant in the discovery cohort (**Figure 1-1**): time I) *pre-viremia,* defined by negative HCV RNA and anti-HCV antibody; time II) *initial viremia,* defined by the first positive HCV RNA test; time III) *late acute infection,* defined by the last available viremic plasma sample from persons who cleared infection (clearers) and a time-from-initial viremia-matched sample from persons with persistent infection (persisters); and time IV) *post-viremia,* defined by HCV RNA negative values ≥ 2 consecutive visits more than 4 weeks apart in clearers and a time-from-initial viremia-matched sample in

persisters. An additional 3 participants from the BBAASH study were included as negative controls. The negative control participants were injection drug users in whom HCV infection was never detected, and plasma samples were examined that corresponded to the same structured time intervals as in the discovery cohort to control for the natural variation in abundance of miRNAs over time. For confirmation of the principal findings, two samples were selected from each of the 28 individuals in the validation cohort: time I) *pre-viremia,* defined by negative HCV RNA and anti-HCV antibody; and time II) *initial viremia*, defined by the first HCV RNA positive sample within 8 months of incident HCV infection.

HCV RNA quantification. HCV RNA levels in plasma samples were previously quantified using a commercial RT-qPCR assay (TaqMan® HCV analyte-specific reagent, Roche Molecular Diagnostics, USA), which has a lower limit of quantification (LLOQ) of 50 IU/mL. Samples with detectable HCV RNA that were below the LLOQ were assigned a value of 25 IU/mL, and detection was confirmed using a nested PCR for the Core-E1 region of HCV as previously published (32).

miRNA isolation and quantification. RNA was isolated from 100 µL of plasma according to the miRvana Total RNA isolation protocol for plasma samples (Life Technologies). Prior to isolation, 5 ng of glycogen was added as a carrier polymer as previously described (33), and 0.5 pg of

Arabidopsis thaliana miRNA ath-miR159a (UUUGGAUUGAAGGGAGCUCUA, Integrated DNA Technologies) was added as an extraction control.

After extraction, RT/qPCR-based TaqMan® OpenArray® chips were used to quantify 754 miRNAs using Human MicroRNA Panel (Pool A v2.1 and Pool B v3.0) on the QuantStudio® 12K Flex machine (Life Technologies) at the Genetic Resources Core Facility, Johns Hopkins Institute of Genetic Medicine, Baltimore, MD. Ct values from the QuantStudio® machine that did not exceed 35 (the array-wide validated limit of quantitation) were exported using the ExpressionSuite® (Life Technologies) software. Array data were deposited on Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo; accession ID: GSE61187). Quantile normalization for arrayed data was performed, as has been previously described using the 'limma' package version 3.24.3 in R (www.r-project.org) version 3.1.2 (34–36).

The qPCR array was separately validated by performing individual RT/qPCR assays in duplicate for ath-miR-159a, miR-122, and miR-885-5p using the microRNA Taqman® assay. Fold change values between longitudinal samples were computed using the $2^{\Delta\Delta Ct}$ method against ath-miR-159a, and by using the $2^{\Delta Ct}$ method in the quantile-normalized approach with agreement between the two methods for all miRNAs for which results are presented. Quantile normalized data is presented.

In vitro HCV infection. A serum-free preparation of the J6/JFH-1 infectious clone of HCV (HCVcc) was produced by isolating serum-free supernatants from cells that were transfected with HCV RNA. Briefly, the pUC19-J6/JFH1 plasmid was linearized using Xba1 (New England Biolabs) and RNA was synthesized *in vitro* using the T7-polymerase based RiboMAX kit (Promega). Phenol-chloroform purified RNA was transfected into Huh 7.5.1 cells using DMRIE-C reagent (Life Technologies) according to the manufacturer's protocol. Cells were cultured in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum until 24 hours post-transfection. Serum-containing media was removed at 24 hours, cells were washed with PBS twice, and media was replaced with Ultraculture serum-free media (Biowhittaker). Serum-free supernatants from infected cells were collected 5-days post transfection, centrifuged at 1,500 x g for 15 minutes and passed through a 0.22 μm filter to remove cellular debris. HCV RNA was quantified using the protocol described above. Mock virus preparations were produced using the same protocol except no HCV RNA was added to the DMRIE-C transfection.

Serum-free cultured Huh 7.5.1 cells were infected with the serum-free HCVcc preparation at a multiplicity of infection of 5 IU/cell. Supernatants were centrifuged at 1,500 x g for 15 minutes and passed through a 0.22 μm filter to remove cellular debris. Supernatant and cellular miRNAs were isolated using the miRvana® PARIS and miRNA

isolation systems, respectively (Life Technologies), according to the manufacturer's protocol.

In vitro HCV infection with miRNA inhibition and over-expression. Seed cultures of Huh 7.5.1 cells were transfected with microRNA miRvana® mimics and inhibitors (Life Technologies) at a final concentration of 40nM using Oligofectamine (Life Technologies). Two days after transfection, cells were infected with HCVcc at a multiplicity of infection of 10 IU/cell. HCV RNA was quantitated 48 hours after infection in supernatants and after cell lysis. Immunohistochemistry for HCV was also performed at 48 hours after infection. Cells were fixed with 4% formaldehyde for 20 minutes then stained for HCV using the 9E10 anti-NS5A antibody (37) at 1:500 dilution in PBS + 3% bovine serum albumin + 0.3% Triton X-100 for one hour at room temperature. Cells were washed twice with PBS and stained using secondary antibody FITC-conjugated goat anti-mouse IgG (Life Technologies) at 1:1000 dilution in PBS + 3% bovine serum albumin + 0.3% Triton X-100 for one hour at room temperature. Cells were washed twice in PBS then imaged with a Zeiss inverted microscope with an Olympus camera.

Focus counting analysis. Cells were transfected, fixed, and stained as described above except that infection was performed using 4 IU/cell to facilitate focus counting. Images were acquired and foci were counted

using an AID iSpot Reader Spectrum operating AID ELISpot Reader version 7.0 at the Johns Hopkins Sidney Kimmel Comprehensive Cancer Center Immunology Core, Baltimore, MD.

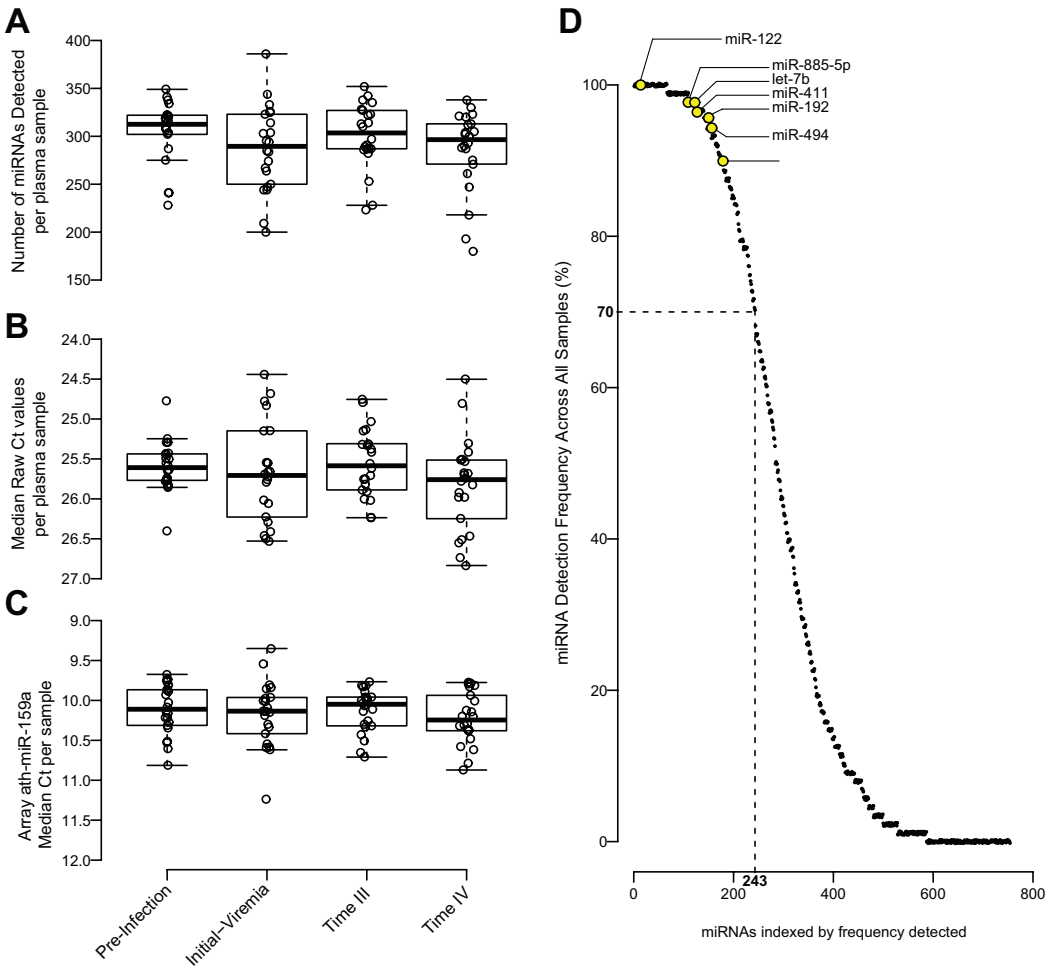

Figure 1-2. Quantitative PCR array performance characteristics indicate that miRNA quantification and detection was unbiased. (A) The number of microRNAs detected per sample and **(B)** the median qPCR crossing threshold (Ct) values among all measured miRNAs per sample are shown across all time points. **(C)** Median Ct values of the externally added control miRNA ath-miR-159a for all samples. Ath-miR-159a was separately included on each array > 10 times to improve the precision of its measurement. **(D)** miRNAs were ordered by their detection frequency (100*the number of samples in which a miRNA was detected divided by the total number of samples) in all samples from participants in the discovery cohort who developed acute HCV infection: 243 miRNAs were detected in > 70% of samples, and these were included in the final unbiased analysis.

Statistical analysis. To avoid false discovery due to incomplete coverage, only microRNAs that were detected at a frequency of greater than 70% across all samples were selected for analysis (**Figure 1-2D**). Two-tailed Wilcoxon rank-sum tests were performed when appropriate. Pearson's correlation was used to detect associations between continuous variables. Multiple comparisons correction was performed according to the Benjamini–Hochberg method (38). Ten microRNAs (including miR-122, miR-885-5p, and let-7b) were pre-specified before analysis for their relevance to HCV biology; correction for multiple comparisons was not employed for these microRNAs. Statistical analysis was performed using R version 3.1.2.

Results

Table 1-1. Participant Characteristics.

	Discovery Cohort **n=22**	Validation Cohort **n=28**	Uninfected Controls **n=3**
Age,* median (min-max)	**25** (19-31)	**23** (16-28)	**29** (29-31)
White, n (%)	**20** (91)	**18** (64)	**2** (67)
Female, n (%)	**15** (68)	**10** (36)	**2** (67)
HCV genotype,* n (%)			
1a	**16** (73)	**11** (39)	**-**
1b	**2** (9)	**2** (7)	**-**
2	**1** (5)	**5** (18)	**-**
3a	**3** (14)	**10** (36)	**-**
***IFNL3* (rs12979860) genotype, n** (%)			
C/C	**11** (50)	**10** (36)	**NP**
C/T	**9** (41)	**15** (54)	**NP**
T/T	**2** (9)	**3** (11)	**NP**
ALT U/L,* median (IQR)	**41** (26-241)	**38** (19-83.25)	**NP**
HCV RNA (log_{10}IU/mL),*median (IQR)	**6.3** (5.3-6.7)	**5.8** (4.97-7.22)	**undetectable**
Days infected,*† median (IQR)	**26** (17-46)	**46** (39-96)	**-**

*At initial viremia
†Imputed infection date is the mid-point between the last known aviremic and first known viremic dates
"-" indicates that this was not applicable.
"NP" indicates that the measurement was not performed.
"undetectable" indicates that all samples over the sampling interval were below the lower limit of detection.

Study and array characteristics. The median (IQR) days-infected at initial viremia in the discovery cohort was 26 (17-46), the median (range) age was 25 years old (19-31), 20 (91%) were white, and 15 (68%) were female (**Table 1-1)**. The median (IQR) HCV RNA level upon initial viremia was 6.3 log_{10} IU/mL (5.3-6.7) and the median (IQR) ALT was 41 U/L (26-241). HCV genotype 1a was present in 16/22 (73%) participants, followed by genotype 1b in 2/22 (9%), genotype 2 in 1/22 (4.5%), and genotype 3 in 3/22 (13.6%). Although the discovery and validation cohorts both represent the natural history of acute HCV (39–42), the main differences

between them were race and *IFNL3* genotype: 18 (64%) of the validation cohort was white and only 10 (36%) were *IFNL3* (rs12979860) genotype C/C. In addition, the median (IQR) imputed number of days infected in the discovery cohort was 26 (17-46) days and 46 (39-96) days for the validation cohort (**Table 1-1**). None of the subjects included in either cohort had evidence of infection with HIV or HBV over the course of the sampling interval.

A median (range) of 303 miRNAs (180-386) was detected per sample: 108 miRNAs were detectable in > 98% of samples, and 468 miRNAs were detectable in ≤ 50% of samples. The analysis was limited to the 243 interpretable miRNAs that were detectable in greater than 70% of all samples. No significant differences were found in the number of miRNAs detected per sample, the un-normalized array median Ct values, or in the extraction efficiency of samples across the sampled time points (**Figure 1-2).**

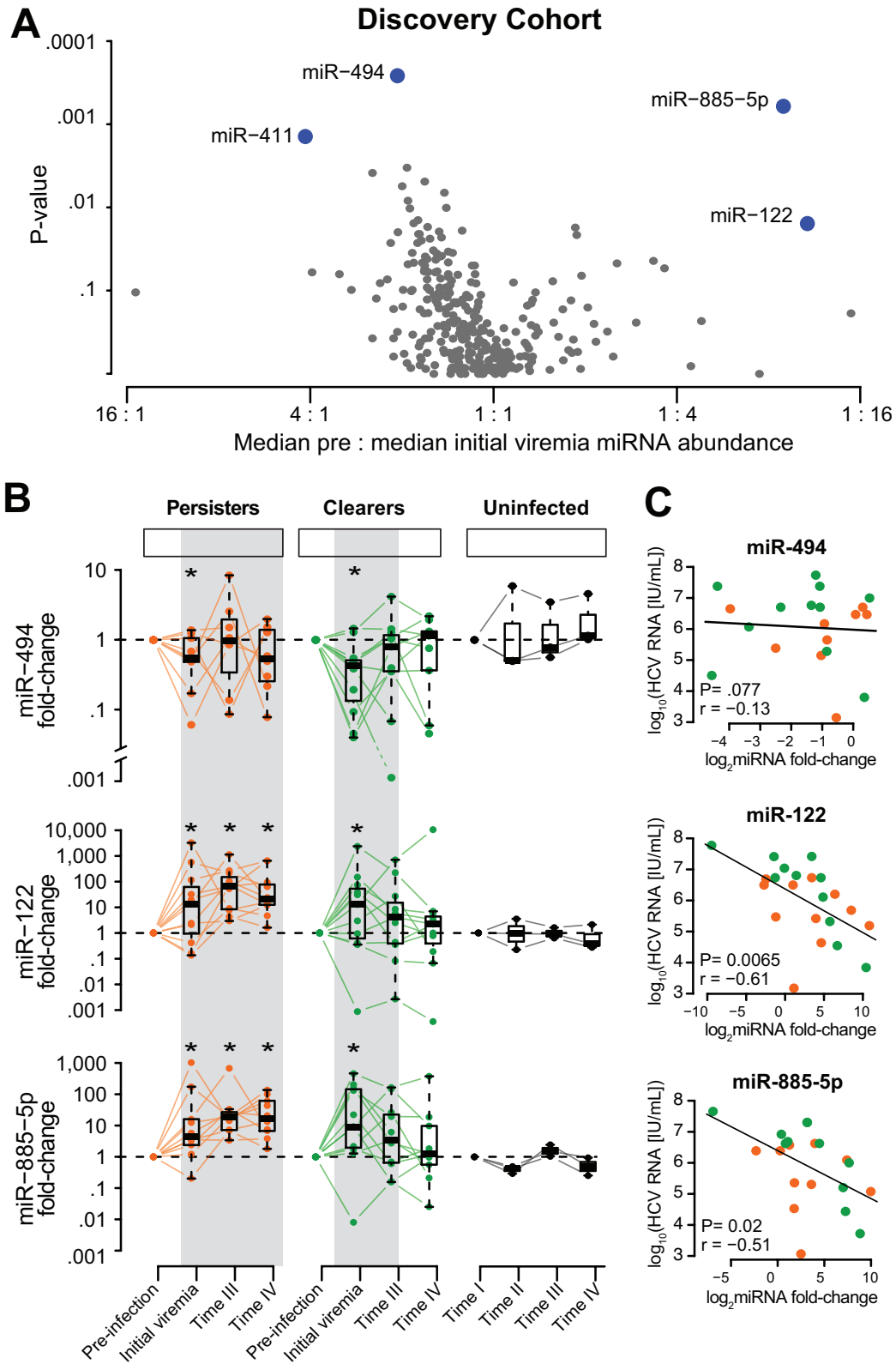
A
Discovery Cohort
miR−494
miR−885−5p
miR−411
miR−122
P-value
.0001
.001
.01
.1
16 : 1
4 : 1
1 : 1
1 : 4
1 : 16
Median pre : median initial viremia miRNA abundance
B
Persisters
Clearers
Uninfected
miR−494 fold-change
miR−122 fold-change
miR−885−5p fold-change
Pre-infection
Initial viremia
Time III
Time IV
Time I
Time II
C
miR-494
miR-122
miR-885-5p
log10(HCV RNA [IU/mL])
log2miRNA fold-change
P= .077
r = −0.13
P= 0.0065
r = −0.61
P= 0.02
r = −0.51

Figure 1-3. Acute HCV infection results in an increase in circulating miR-122 and miR-885-5p and a decrease in circulating miR-494. (A) A volcano plot depicts the median change in abundance of each miRNA across 22 participants between pre-viremia and initial viremia in the x-axis (miRNAs that are increased at initial viremia compared to pre-viremia are to the right of the graph). The y-axis shows the uncorrected p-value of the change in miRNA abundance. miRNAs in blue were significantly different between pre- and initial viremia in the discovery cohort after correction for multiple comparisons; no correction, however, was applied for miR-122, which was pre-specified in the study design for its relevance to HCV biology. **(B)** The change in circulating miRNA abundance versus pre-infection in matched time-intervals is displayed for BBAASH participants who were HCV negative (black) and who acquired acute HCV infection (green and orange lines). The selected time points were determined by participants who acquired HCV: before infection (pre-infection), during acute infection in the earliest (initial viremia) and latest (time III) available samples, and either after infection in clearers (green; n=11) or during the transition to chronic infection in persisters (orange; n=11) (time IV). Shown are changes in miRNA abundance from baseline in all 22 individuals in the discovery cohort for miR-411, miR-494, miR-122, and miR-885-5p. **(C)** In the x-axis is shown the change in abundance of miR-494, -122, and -885-5p between pre-infection and initial viremia. In the y-axis is shown the contemporary plasma $\log_{10}$ HCV RNA levels (IU/mL) at initial viremia. Individual points are further delineated for clearers (green; n=11) and persisters (orange; n=11).

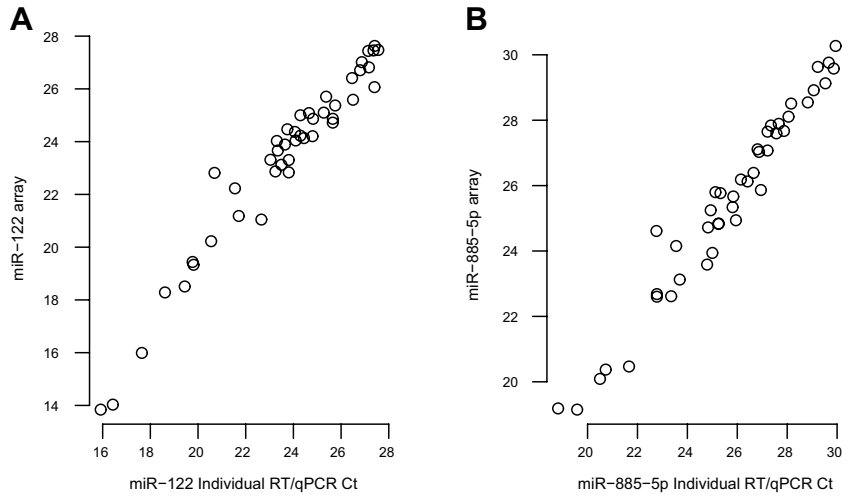

Figure 1-4. Validation of qPCR array results using individual RT/qPCR assays for select miRNAs. (A) miR-122 and **(B**) miR-885-5p abundance at pre-infection and initial-viremia time points for all infected individuals in the discovery cohort were confirmed using individual RT/qPCR Cts; qPCR array results compared favorably with individual RT/qPCR results ($r>0.97$, $p<2.2*e-16$ for both).

A circulating miRNA signature of acute HCV infection. With the onset of viremia, the abundance of four plasma miRNAs significantly changed (**Figure 1-3A**); the changes in abundance were not associated by their

constitutive expression in liver. Extracellular circulating miR-494, the second most abundant miRNA in the liver, and miRNA-411 decreased to a median (IQR) of 0.48-fold (0.19-0.78) and 0.24-fold (0.15-0.61), respectively ($P_{corrected}$<*.05* for each), compared to pre-infection levels; both miRNAs decreased in 15/22 (68%) participants. In contrast, circulating levels of miR-122, the most abundant miRNA in the liver, rose a median (IQR) of 5.3-fold (1.5-39.2; *P<.01*) over pre-infection levels, and increased levels were noted in 16/22 (73%) participants. miR-885-5p rose a median (IQR) of 9.0-fold (2.4-130.2; *P<.001*), and increased levels were noted in 19/22 (86%) participants. We confirmed that amplification scores for miR-494, -411, -122, and 885-5p were above 1.2, the accepted threshold. In addition, the changes in abundance of miR-122 and -885-5p were individually confirmed using separate RT/qPCR assays and correlated tightly with the results from the qPCR array (**Figure 1-4**).

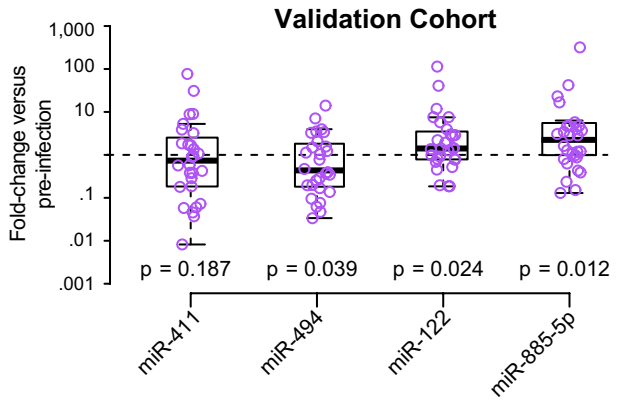

Figure 1-5. Validation of the plasma miRNA signature of acute HCV in a separate cohort. Boxplots with individual values show the changes in circulating miRNA abundance for miR-411, miR-494, miR-122, and miR-885-5p between pre-viremia and initial viremia in 28 participants of the UFO study. P-values generated by the non-parametric Wilcoxon rank-sum test are shown for the plotted miRNAs.

The extracellular miRNA signature was confirmed in a validation cohort of 28 participants in the San Francisco-based UFO study (**Table 1-1**). Changes in circulating miR-494, -122, and -885 with HCV infection were consistently observed in the validation cohort: miR-494 levels decreased a median (IQR) of 0.43 fold (0.19, 1.7; *P<.05*), miR-122 levels rose a median (IQR) of 2.3 fold (0.9-4.7; *P<.05*), and miR-885-5p rose a median (IQR) of 1.5 fold (0.9-3.25; *P<.05*). Changes in miR-411 were not consistently observed and so this miRNA was not studied further (**Figure 1-5**).

To determine whether the extracellular miRNA signature of acute HCV infection persisted during HCV viremia, miRNA abundance was quantified until the resolution of infection (clearers) or until early chronic infection (persisters)(**Figure 1-3B**). Four samples were taken from each person: i) *pre-infection* (negative HCV RNA and antibody tests); ii) *initial viremia* (first positive HCV RNA test); iii) *time III* (the last viremic visit in clearers and a time-from-initial viremia-matched sample from persisters); and iv) *time IV* (HCV RNA consistently negative on ≥ 2 consecutive visits more than 4 weeks apart in clearers and a time-from-initial viremia-matched sample in persisters). miR-494 levels decreased transiently at initial viremia, but were not significantly different from pre-infection levels at later time points (**Figure 1-3B**). In contrast, miR-122 and miR-885 levels were persistently elevated for the duration of viremia, returning to pre-infection levels at time IV in clearers, but remaining elevated in

persisters (*P<.001*) (**Figure 1-3B**). To verify that the extracellular miRNA signature was not the result of fluctuations over time that are unrelated to acute HCV infection, levels of miRNAs that constituted the signature were also examined in 3 IDUs in the BBAASH cohort who did not contract HCV over the course of their participation in the study. Plasma samples from uninfected IDUs were matched for intervals that were comparable with those observed in the discovery cohort (**Figure 1-3B**): no consistent changes in abundance were noted over time for miR-494, miR-122, and miR-885-5p among the 3 uninfected IDUs, further underscoring the findings in acute HCV infection.

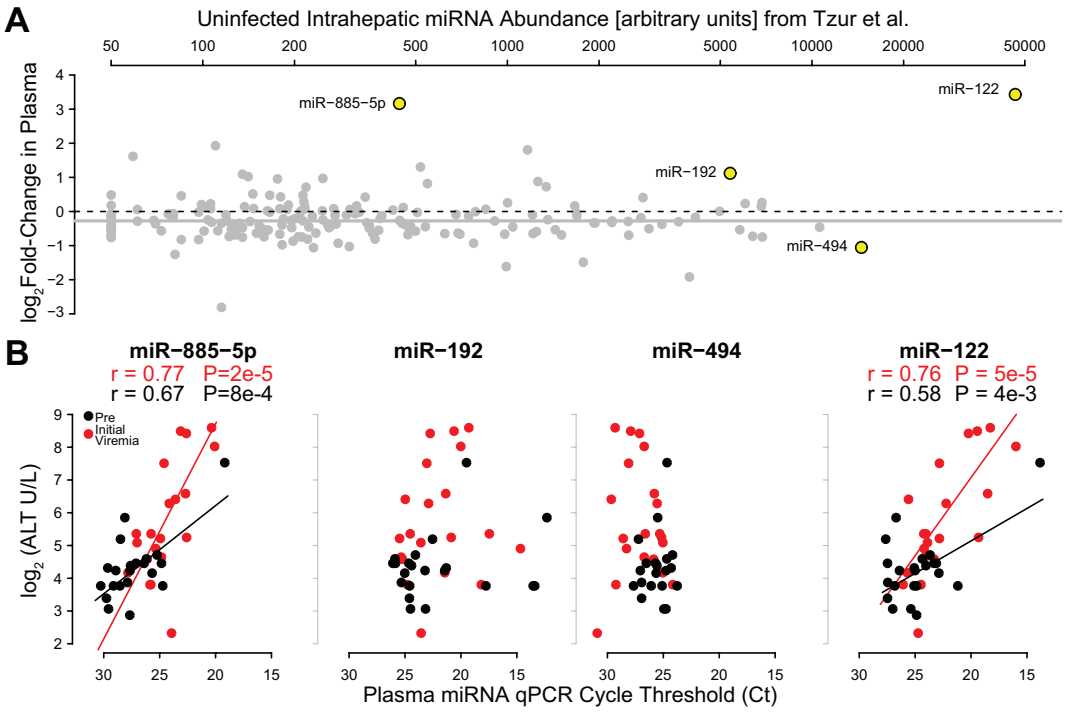

Figure 1-6. The circulating microRNA signature of acute HCV infection does not reflect intrahepatic miRNA abundance. (A) The change in circulating miRNA abundance from pre- to initial viremia for each of 185 miRNAs (y-axis) was compared to the intrahepatic abundance of the same miRNAs (x-axis), based on previously published data (43). When miR-122, the most abundant intrahepatic miRNA, was included, there was a statistically significant but weak correlation between intrahepatic abundance and

the change in circulating abundance of all studied miRNAs (r=0.19; P=.001). The relationship was lost when miR-122 was excluded from the analysis r=-.04,p=.46). **(B)** The relative circulating abundance of miR-885-5p, miR-192, miR-494, and miR-122 was compared to contemporaneous ALT levels at pre-(black) and initial (red) viremia and presented in order of their intrahepatic abundance. Regression lines and their corresponding Pearson's correlation coefficient are included for statistically significant associations.

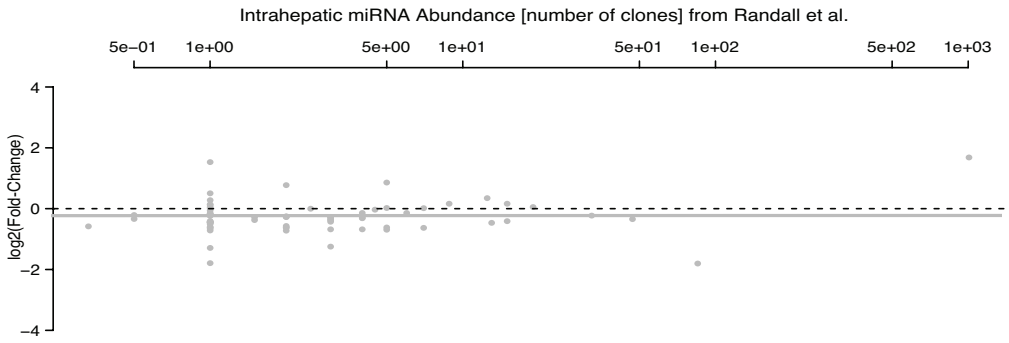

Figure 1-7. Plasma miRNA abundance is not related to intrahepatic miRNA abundance. The plasma miRNA abundance at initial viremia for each of 70 miRNAs (y-axes) was compared to the intrahepatic abundance of the same miRNAs (x-axes), based on previously published data.

We further investigated whether there were predictors of the circulating miRNA signature, first by testing whether its components were related to contemporary HCV RNA levels. While we observed no relationship between HCV RNA levels and the change in abundance of miR-494, we found inverse associations with miR-122 (r=-0.56, P=0.0065) and miR-885-5p (r=-0.5, p=0.02)(**Figure 1-3C**). These associations were only observed during the initial viremia time point. Next, we used two approaches to consider the role of hepatocyte damage in producing the extracellular miRNA signature. First, we adapted data from Tzur et al. describing the abundance of 185 miRNAs in human livers that were also profiled in our study (43). We found no relationship between intrahepatic abundance of miRNAs and their change in extracellular abundance upon

acute HCV infection (r=-0.04,*P=.46*; **Figure 1-6A**). Similarly, we applied data from 148 miRNAs examined by Randall et al. to our study and did not detect a relationship between liver abundance and change in plasma abundance after accounting for miR-122 (r=-0.09,*P=.28*; **Figure 1-7**) (44).

The second approach we used to study the role of hepatocyte damage in the extracellular miRNA signature was to examine associations with aminotransferases: while miR-122 and miR-885 were associated with contemporaneous ALT levels, miR-494 had no association with ALT despite its reported intrahepatic abundance (**Figures 1-6 and 1-7**). Conversely, miR-192, which has been reported as being increased and related to ALT levels in necrolytic drug-induced liver injury (45, 46), was not significantly elevated or associated with ALT levels in our study (**Figure 1-6B**). Taken together, these results suggest that changes in extracellular miRNA abundance are not solely the result of cell lysis.

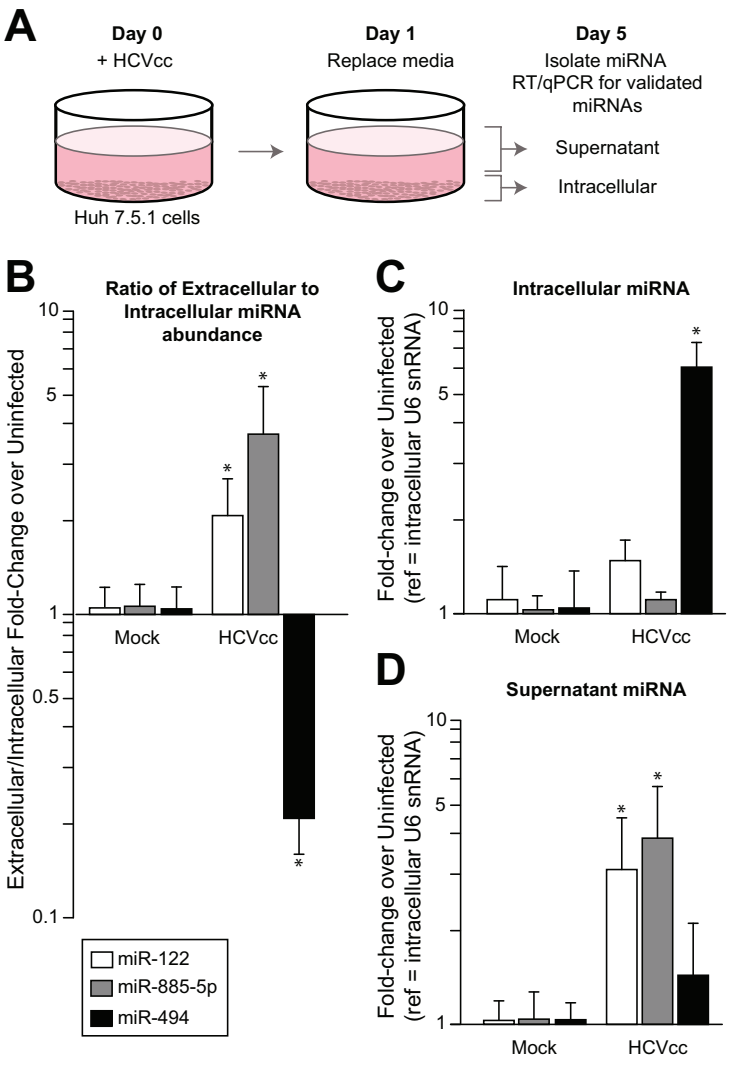

Figure 1-8. Infection *in vitro* recapitulates the observed signature of circulating miRNAs in acute HCV infection *in vivo*. (A) Huh 7.5.1 hepatoma cells were infected with the J6/JFH-1 infectious clone of HCV (HCVcc) and compared to uninfected cells (mock). All cell cultures and virus were prepared using serum-free to avoid confounding measurement of highly conserved miRNAs in fetal bovine serum. After 24 hours, cells were washed and the media replaced. Five days following infection, the ratio of the fold change of miRNAs in supernatants to fold-change within cells (**B**) shows the relative enrichment or depletion of miRNAs in supernatants relative to cells. Cellular (**C**) and supernatant (**D**) abundance of miRNAs were normalized to intracellular snRNA U6 abundance. Error bars indicate the standard deviation of four replicate infections and asterisks indicate significant ($P<0.05$) changes from mock controls.

The miRNA signature of acute HCV infection *in vitro*. To support that hepatocytes, and not other cells, contribute to the miRNA signature of acute HCV infection, we infected Huh 7.5.1 hepatoma cells with HCVcc

and quantified the expression of miRNAs in cells and supernatants, comparing these amounts to uninfected cells that were cultured for the same number of days (**Figure 1-8A**). Quantities of extracellular miRNAs were normalized to intracellular amounts of the same miRNAs to control for cell-death present in test and control wells. The extracellular signature that was found in circulating human plasma was confirmed *in vitro:* miR-122 and -885 increased outside of cells relative to inside, while miR-494 decreased (**Figure 1-8B**). Interestingly, intracellular miR-122 and miR-885-5p abundance did not vary with infection (**Figure 1-8C**) while extracellular levels significantly increased ($P<.05$ for both, **Figure 1-8D**). In contrast, the relative decrease in extracellular miR-494 abundance largely resulted from intracellular accumulation (**Figure 1-8C**). Collectively, these results suggest that the release or retention of the miRNAs that constitute the signature of acute HCV infection may be selective and specific.

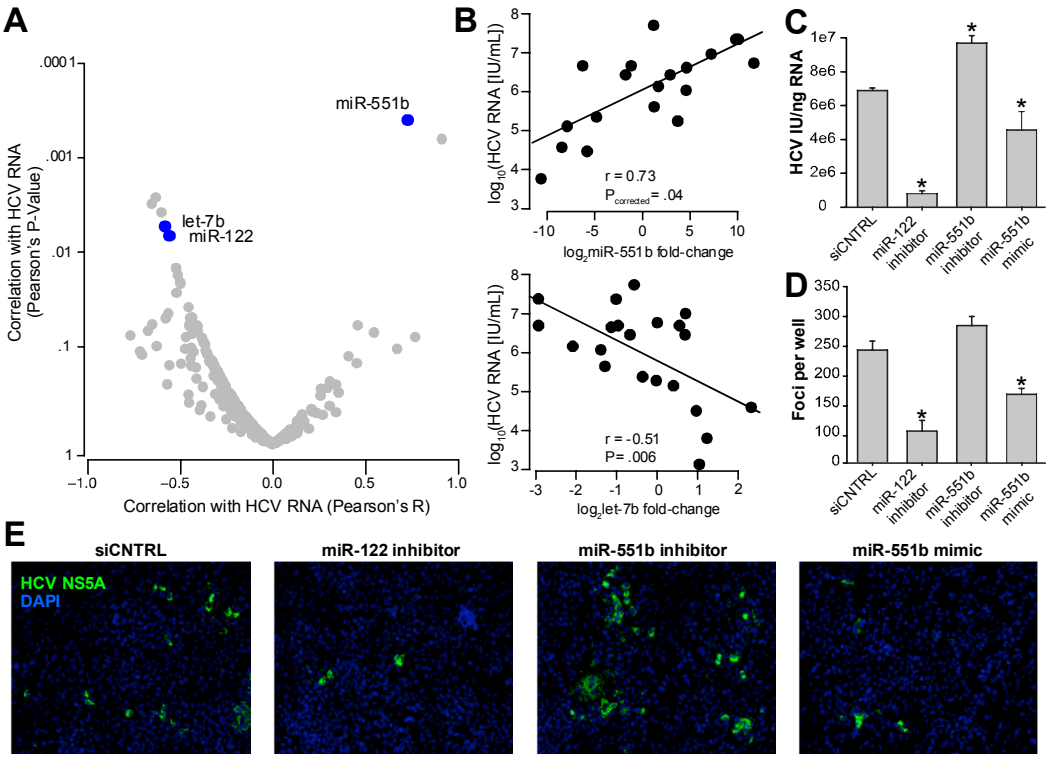

Figure 1-9. Circulating miR-122, let-7b, and miR-551b are associated with plasma HCV RNA abundance in acute infection. (A) Pearson's correlation P-values are plotted for fold change of each miRNA over pre-infection and HCV RNA abundance at acute viremia. Blue points indicate miRNAs with significant relationships with HCV RNA either after correction for multiple comparisons (miR-551b) or P<0.05 for miRNAs with known impacts on HCV replication (miR-122 and let-7b). **(B)** Scatterplots indicate the association between plasma HCV RNA levels and the relative abundance of let-7b (upper panel) and miR-551b (lower panel)**.** Regression line and corresponding Pearson's correlation coefficient is included. **(C)** Intracellular HCV RNA normalized to cellular RNA input from HCVcc infected Huh 7.5.1 cells transfected with indicated miRNA inhibitors or mimics is shown. Error bars indicate standard deviations between four repeat transfections and asterisks indicate significant changes in HCV RNA when compared to the negative control siRNA (siCNTRL). **(D)** Foci of infection, immunostained for HCV NS5A, were quantified from HCVcc infected Huh 7.5.1 cells under the labeled conditions. **(E)** Representative images from source wells used for **(D)** showing NS5A in green and nuclei (DAPI) in blue.

Extracellular miRNAs and plasma HCV RNA abundance. High HCV RNA levels during early acute infection have been found to be associated with spontaneous clearance (47). To understand potential determinants of HCV viremia in acute HCV infection, we examined changes in abundance of circulating miRNAs that were closely associated with HCV RNA levels

at initial viremia (**Figure 1-9A**), adjusting for multiple comparisons. The change in abundance of extracellular miR-551b was the most strongly associated with contemporaneous plasma HCV RNA levels (r=0.73; $P_{corrected}$=0.04; **Figures 1-9B**), although the median (IQR) amplification scores for miR-551b were 0.968 (0.899-1.024). As we found earlier, HCV RNA levels were inversely associated with plasma levels of miR-122 (**Figure 1-3C**). In addition, we found an inverse association between HCV RNA levels and let-7b (r=-0.59; *P=.005*), a miRNA that is induced by type 1 and 3 interferons and has anti-viral properties *in vitro* (48, 49).

To establish if miR-551b directly impacts HCV replication in hepatocytes, we transfected Huh 7.5.1 hepatoma cells with a miR-551b inhibitor and mature miR-551b mimic separately: miR-551b inhibition resulted in a significant increase in HCV RNA (*P<0.05*), while the miR-551b mimic led to a decline in HCV RNA abundance (*P<0.05*;**Figure 1-9C**). By comparison, miR-122 inhibition led to more potent suppression of HCV RNA levels. Transfection efficiency in these experiments was 81.5% ± 1.3%, and was measured using FITC-conjugated scrambled oligonucleotide.

We confirmed the HCV RNA findings by quantifying HCV protein expression. Cells were transfected with a miR-551b mimic, inhibitor, and a scrambled siRNA, then infected with HCV, and finally immunostained for the HCV NS5A protein. Cells that were treated with the miR-551b mimic showed a decrease in the number of infected cells when compared to the

scrambled siRNA control, while miR-551b antagonism led to an increase in the number of infected cells **(Figures 1-9D and 1-9E).** Focus counting revealed a significant decrease (*P<0.05*) in the number of foci in cells transfected with the miR-551b mimic, however, the increase in the number of foci observed in cells transfected with the miR-551b inhibitor did not reach significance (*P=0.1*) **(Figure 1-9D).**

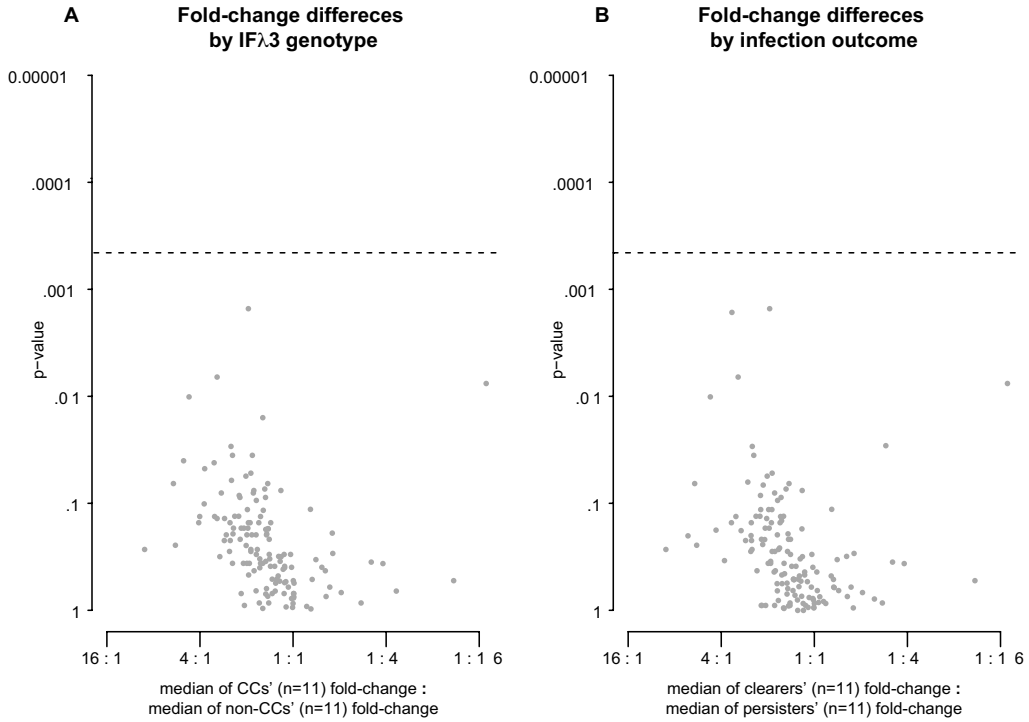

Figure 1-10. Changes in miRNA abundance did not significantly differ by *IFNL3* genotype or infection outcome. Fold-changes between pre-infection and initial viremia were separated into groups according to **(A)** *IFNL3* genotype or **(B)** infection outcome and each group was assigned a median fold change (n=11 for CC, nonCC [CT or TT], clearers, and persisters). The volcano plots depict the differences in the median fold-changes between each group in the x-axes in the form of a ratio (higher fold-change in the non-CTs or persisters to the left). In the y-axis is shown the uncorrected p-value of the change in miRNA abundance. The dashed lines indicate the threshold for significance.

Discussion

By comparing plasma miRNA levels in a well-characterized cohort before, during, and after acute HCV infection, we identified a consistent and specific extracellular miRNA signature of acute infection that was confirmed in a separate cohort and *in vitro*. Contrary to conventional wisdom, circulating miRNA abundance during acute infection did not simply represent cellular necrosis but a nonrandom process of retention and release that correlates with the onset and control of virus infection.

miR-122 is the most abundant miRNA in the liver; its increase in plasma levels during chronic infection has been presumed to be analogous to elevated serum aminotransferase levels, that are a consequence of cellular necrosis. However, miR-494, the second most abundant miRNA in the liver profiled, decreased in the plasma of acutely infected persons, and was found to increase intracellularly upon infection. We also confirmed that there was no association of acute HCV infection with plasma abundance of miR-192, a liver abundant microRNA that has been found to be increased in circulation during drug-induced liver injury (45, 46) and steatohepatitis (50). Taken together, these findings are inconsistent with necrolytic release of miRNAs, and implicate selectivity in the release of cellular miRNAs that determines their plasma abundance.

Our miRNA isolation technique does not distinguish between free, exosome-associated, or virion-associated miRNAs. The observed

increase in abundance of extracellular miR-122 upon HCV infection is consistent with the recent observations that miR-122 is packaged along with HCV RNA in exosomes (51). Moreover, the *inverse* relationship between circulating miR-122 and plasma HCV RNA levels suggest that miR-122 is not packaged directly into virions. Future research should focus on how miRNAs are differentially released and compartmentalized.

Circulating miRNA profiles have been associated with HCV infection in previous studies, although many of these studies were performed in persons with long-term chronic infection in whom liver fibrosis may have also contributed to miRNA abundance (52, 53). For example, changes in circulating miR-122 abundance have been reported in persons with chronic HCV infection and also in drug-induced liver injury (45) and non-alcoholic steatohepatitis (52). In addition, serum miR-122 levels in persons with chronic HCV infection who were treated with interferon and attained a sustained virologic response were higher than persons who had a null response (54, 55), and these results were largely explained by the genetic susceptibility marker near the *IFNL3* gene. In contrast, we did not find an association between the change in miR-122 levels upon acute infection and the outcome of infection or *IFNL3* genotype (**Figure 1-10**), although our sample size was not powered for this question. Given that there is an established positive association between HCV clearance and peak HCV RNA abundance (47), and that we found a negative relationship between circulating miR-122 and HCV RNA,

it is logical to speculate that the association between circulating miR-122, *IFNL3,* and HCV clearance may be consistent with what was observed during interferon treatment responses. However, we acknowledge that conclusive proof would require study of a larger cohort. We also note that we did not find changes in the abundance of miR-20a and miR-92a, in contrast to a previous report of miRNAs during acute HCV infection (17), although their study populations were composed of participants with symptomatic jaundice, which occurs during no more than 20% of acute HCV infections.

Circulating miR-885-5p has been less extensively characterized than miR-122. Gui et al. reported that serum miR-885-5p levels are increased in the setting of hepatocellular carcinoma, cirrhosis, and chronic hepatitis B (56). Our finding that miR-885-5p was elevated in acute infection is consistent with these previous findings and suggests that elevation in plasma miR-885-5p is likely part of a conserved response to liver stress.

To date, only two other publications exist regarding the role of miR-551b in any biological process (57, 58). Fu et al. describe a down-regulation of miR-551b in hepatoma cells expressing a mutant of the hepatitis B virus HBx protein that induces higher levels of cellular proliferation and Xu et al. describe a role for miR-551b in acquired apoptosis and chemotherapy resistance in human lung cancer cell-lines. Interestingly, we found that this miRNA reduced HCV replication even

after adjusting for cellular abundance (**Figure 1-9**). Further investigation into the mechanism of miR-551b is necessary and may reveal a novel antiviral response mediated by miR-551b with activity against viruses other than HCV.

Interestingly, both miR-122 (59–62) and miR-885-5p (63–65) have been identified as potent tumor suppressors with functional significance in cancer progression. In contrast, miR-494, which decreased in plasma abundance but increased within hepatocytes during acute HCV infection, has been reported to have oncogenic potential in hepatocytes (66) and anti-proliferative effects in breast cancer cells (67). As mentioned above, miR551b appears to have proliferative and anti-apoptotic properties in certain settings. Recent attention has focused on extracellular miRNAs as having the potential to regulate recipient cell gene expression (21–29); therefore, it is possible that the changes in extracellular miRNAs during acute HCV infection could compose part of an innate response to viral infection that modulates cellular proliferation and viral replication. Our finding that HCV RNA levels were inversely associated with changes in extracellular let-7b, a known anti-viral that is induced by interferon, support the hypothesis that the circulating miRNA signature may be part of an anti-viral response.

There were several limitations to this study. The large number of miRNAs considered increases the risk of type 1 error. Accordingly, we took the conservative approach by correcting for multiple comparisons.

More importantly, we confirmed our findings using multiple approaches: i) by identifying the same signature in a completely distinct group of individuals in the UFO cohort; ii) by finding that miRNA levels of the signature were invariant in uninfected IDUs over matching time intervals; and iii) by confirming the signature *in vitro*. We note that differences at baseline between the discovery and validation cohorts should have made it less likely that we would find the same circulating miRNA signature by chance, further supporting the validity of our findings.

On the other hand, it is possible that some small associations of miRNAs with acute HCV infection were missed by a study of this size. However, unlike larger cross-sectional studies in which disease associations are often confounded by person-to-person variability, we studied longitudinal samples in the same person before, during, and after acute HCV infection to establish consistent patterns of circulating miRNA abundance. Relative to cross sectional studies, this approach markedly improved our power by controlling for person-to-person variance.

Another limitation is that we did not sample liver directly. However, it would be unethical to use liver tissue in humans obtained for that purpose since liver biopsies are not routinely indicated during acute HCV infection and are very rarely performed before HCV infection. Therefore, we examined miRNA changes in hepatoma cells infected with HCV *in vitro* and used previously published data on miRNA expression in healthy liver.

There are remaining questions about the role of circulating miRNAs in acute HCV infection. Our results suggest that extracellular miRNA release is sequence-specific and correlates with the onset and outcome of HCV infection. Given the contribution of the implicated miRNAs in HCV replication and the prediction of clinical outcomes during chronic infection, these findings also suggest that the pathophysiology of chronic HCV infection may be established acutely as part of an innate response to viral infection. Future research is required to understand the mechanism of miRNA sorting and the role of plasma miRNAs in HCV infection.

This page is intentionally blank.

CHAPTER 2: A polymorphism in hypervariable region 1 confers resistance to broadly neutralizing antibodies targeting opposite termini of the hepatitis C virus E2 protein.

Abstract

Broadly neutralizing monoclonal antibodies (bnAbs) are informative for development of vaccines against highly variable viruses including hepatitis C virus (HCV), since they target relatively conserved viral epitopes that could serve as vaccine antigens. However, HCV resistance to bnAbs is poorly understood. HC33.4 and AR4A are two of the most potent bnAbs characterized to date, binding near the amino- and carboxy-terminus of HCV envelope (E2) protein, respectively. Given their distinct epitopes, it was surprising that these bnAbs clustered together in an analysis of neutralization profiles across a panel of natural HCV isolates, suggesting that some viral polymorphisms may confer resistance to both bnAbs; however, a competing hypothesis requiring mechanistic study was that the resistance polymorphisms co-evolved rather than sharing a mechanistic basis for resistance. To investigate this resistance, we developed a large, diverse panel of HCV variants and a novel computational method using amino acid sequence variation and neutralization sensitivity to identify bnAb resistance polymorphisms in E2. We identified polymorphisms conferring resistance to each bnAb individually, as well as a polymorphism outside of either binding epitope that confers resistance to both bnAbs.

This diverse viral panel in conjunction with a novel computational pipeline allows measurement of neutralizing antibody breadth, and prediction of naturally occurring bnAb resistance polymorphisms in E1E2. This analysis identified a polymorphism in hypervariable region 1 conferring resistance to two potent broadly neutralizing human mAbs targeting opposite termini of the E2 protein, and the method is broadly applicable to future studies to define mechanisms of HCV resistance to bnAbs.

Introduction

Hepatitis C virus (HCV) infects over 170 million people worldwide (3) and kills more people in the United States annually than HIV (68). Appalachian regions of the United States saw a >350% increase in the number of new HCV infections from 2009-2012 (69) and recent outbreaks in the United States have been attributed to the rapid increase in injection drug use (70). While direct-acting antiviral (DAA) therapy has revolutionized care for patients with HCV, control of the HCV pandemic remains challenging due frequent reinfection in high-risk individuals who have achieved a sustained virologic response from DAA therapy (71), transmission of NS5A inhibitor-resistant HCV variants from individuals failing DAA therapy (72), and the high proportion (~50%) of the infected individuals in the United States who are unaware asymptomatic carriers (73).

A major barrier to the development of a prophylactic vaccine for HCV is eliciting an immune response that is protective against a significant proportion of the naturally occurring viral variants (74, 75). Broadly neutralizing antibodies (bnAbs) against HCV prevent successful entry of a wide variety of naturally occurring HCV isolates. Globally, a majority of HCV infections are categorized into 6 genotypes based on genetic sequence that differ in their global localization with multiple subtypes occurring within some genotypes. Our group has recently reported

significant intragenotypic variation in sensitivity to a majority of the bnAbs characterized to date using an array of reporter viral particles displaying naturally occurring HCV envelopes (HCVpp) (76). Curiously, when the abilities of a panel of bnAbs to neutralize an array of 19 HCV genotype 1 envelopes were compared, distinct relationships between antibodies were observed, allowing grouping of all bnAbs into three distinct neutralization clusters, and suggesting that common E1E2 determinants of neutralization sensitivity are shared between bnAbs within each cluster.

One neutralization cluster includes the potent bnAbs HC33.4 and AR4A, although their described binding epitopes are near opposite termini of the E2 protein (**Figure 2-1**)(77, 78). We hypothesized that shared E1E2 resistance polymorphisms to these antibodies would explain the unexpected correlation between neutralization profiles of HC33.4 and AR4A.

Materials and Methods

Source of E1-E2 full-length envelope isolates. Plasma samples obtained from HCV infected subjects in the BBAASH cohort [15,16], Irish Anti-D cohort (79), and Swan Project (80) were used to construct a library of genotype 1 E1E2-expressing lentiviral pseudoparticles using a high-throughput production and screening approach.

HCV E1E2 expression plasmids. The E1E2 region was PCR amplified from cDNA reverse transcribed from viral RNA purified from subject plasma and cloned into the expression vector pcDNA3.2/V5/Dest (Invitrogen) using Gateway technology in a one-tube BP/LR reaction, as previously described (81).

Phylogenetic Analysis. Sanger sequencing of the entire length of the cloned E1E2 region was performed in both directions. Amino acid sequences from a nucleic acid MUSCLE alignment (82) were used to build a phylogenetic tree. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value(83). All trees are drawn to scale, with branch lengths measured in the number of substitutions per

site, and all positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA6 (84).

High-throughput HCVpp production and infectivity screening. HCVpp were produced from purified cDNA clones by lipofectamine-mediated transfection of HCV E1E2 and pNL4-3.Luc.R-E- plasmids into HEK293T cells in 96-well plates as previously described(81). Huh 7.5.1 cells were exposed to supernatants from 293T cell cultures in order to test for the presence of infectious HCVpp, as previously described (81). HCVpp were considered infectious if the luciferase activity in Huh 7.5.1 cell lysates was greater than 10,000 RLU. Isolates could have been from the same infected subject, however, must have differed by at least one amino acid from every other clone contained in the library. Envelopes that displayed enhanced infection (relative infection >1.2 with either bnAb) were discarded not included in the analysis or in the description of library meta-data.

Subject-adjusted Neutralizing Antibody Prediction of Resistance (SNAPR) algorithm. Amino acid alignments built as described in the phylogenetic analysis. To account for the uneven distribution of the infectious clones per subject, isolates from the library were selected at random from each person and added to the initial data set until all subjects

were represented by an equal number of isolates. This expanded alignment was then divided into groups at each amino acid position in the alignment based on amino acid present. Median associated relative infection values from each group were assigned and the relative infection values from the group with the lowest median relative infection value were compared to the relative infection values of all other groups using a Mann-Whitney U-test. All statistical analyses were performed using R version 3.1.2.

Site Directed Mutagenesis. The top ten positions were selected for further testing. The predicted ancestral clone was tested at each position, as was the clone with the highest sensitivity to HC33, the clone with the lowest sensitivity, the clone with median sensitivity out of all those containing the amino acid most closely associated with high sensitivity, and the clone with median sensitivity out of all those containing the amino acid associated with low sensitivity. Mutants were created using the QuickChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent) and ≥2-fold coverage by Sanger sequencing was performed to verify that all mutants differed from parent clones at only the desired locus.

Infectivity Measurements. Ninety-six-well flat bottom tissue culture plates (Falcon) were plated with 8,000 Hep3B cells per well and incubated overnight in a humidified CO_2 incubator at 37^{o}C. Media was removed from

the cells the following day and replaced with 50 µL of HCVpp. The plates were placed in a CO_2 incubator at 37^{o}C for 5 hours, after which the HCVpp were removed and replaced with 100 µL of phenol-free Hep3B media. After approximately 72 hours in the CO_2 incubator, media was removed from the cells and 50 µL of 1x Cell Culture Lysis Reagent (Promega) added and left to incubate for >5 minutes then 45 µL from each well were then transferred to a white, low-luminescence 96-well plate (Berthold) and read in a Berthold Luminometer (Berthold Technologies Centro LB960). Each sample was tested in duplicate. A mock pseudoparticle (no envelope) was used as a negative control.

Measurements of Neutralization. The same procedure used to measure infectivity was employed, but instead of pure pseudoparticle being added to the Hep3B cells, 50 µl of a mixture of 10% antibody (original concentration 10µg/ml) and 90% pseudoparticle (by volume) that had been incubated at 37^{o}C for 1 hour was added. HCVpp were tested in duplicate against both the antibody being tested and IgG, the negative control. Murine Leukemia Virus (MLV) was used as a positive control.

Calculation of Resistance. The mean fluorescence from the HCVpp tested against the antibody was calculated and divided by the mean RLU from the same pseudoparticles tested against IgG to determine the fraction of infection. The fraction of infection in presence of the antibody

was then divided by the fraction of infection in the presence of IgG to determine the relative resistance of each mutant.

Results

The expanded library of envelopes contains 117 distinct full-length naturally occurring E1E2 sequences. To construct a library of envelopes to predict relationships between sequence and neutralization sensitivity, we cloned 113 naturally occurring HCV genotype 1 envelopes that produced functional HCV pseudoparticles (HCVpp) containing a NL4.3Δenv-Luc reporter genome as previously described (76).

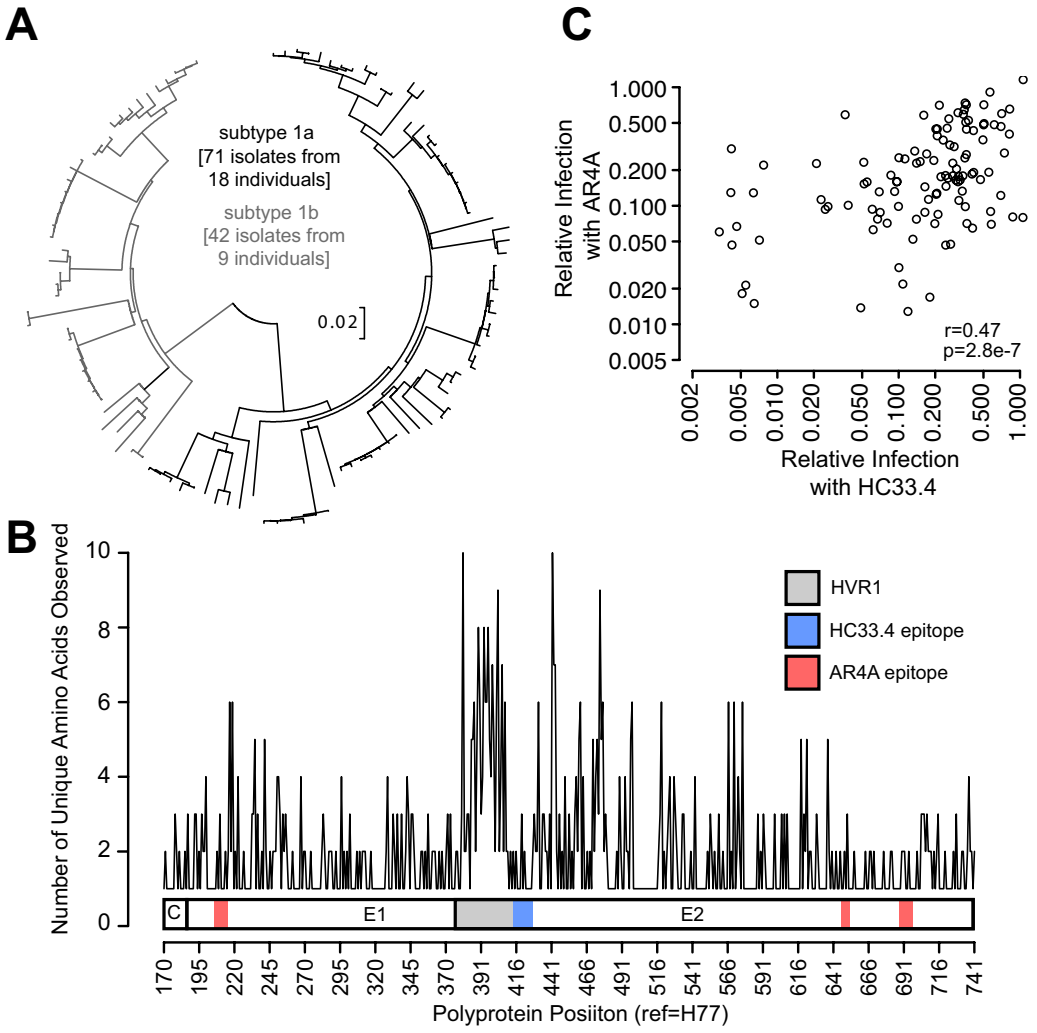

Figure 2-1. Construction of an E1E2 HCV envelope panel for neutralizing antibody breadth testing and sequence prediction of neutralizing antibody resistance polymorphisms. (**A**) Phylogenetic tree of E1E2 amino acid sequences determined by maximum likelihood (log likelihood=8668) shown with the distances drawn to scale. (**B**) Number of different amino acids present at each position in the full E1E2 panel alignment. Previously defined HC33.4 and AR4A binding epitopes are indicated (blue and pink), as is hypervariable region 1 (HVR1) (gray) and the portion of E2 included in the E2 core crystal structure determined by Kong *et al*, (ref) (green)(**C**) Spearman correlation between relative infection of 113 HCVpp in the presence of HC33.4 and AR4A. Each point indicates neutralization of a single HCvpp by HC33.4 on the x-axis and AR4A on the y-axis.

The resulting library includes 71 subtype 1a originating and 42 subtype 1b HCV envelopes (**Figure 2-1A**). This library contains 97% of amino acid polymorphisms present at greater than 5% frequency in a reference panel of 643 genotype 1 HCV isolates from GenBank (85) resulting in as many as 10 amino acid options in a single locus. The largest variety of sampling was observed in parts of the envelope not included in crystalized E2 (**Figure 2-3**).

Neutralization of the expanded panel of HCV envelope by AR4A and HC33.4 also displays large variation in bNAb sensitivity. By comparing entry efficiency of these envelopes in the presence of a HC33.4 or AR4A with entry in the presence of an isotype control, we quantified the relative infection of each envelope in the presence of each bnAb. AR4A and HC33.4 were associated with a median (IQR, min-max) relative infection of 0.17 (0.09-0.25,0.01-1.15) and 0.22 (0.09-0.27,0.003-1.1) respectively. Interestingly, HC33.4 more potently neutralized subtype 1a (relative infection median=0.17,IQR=0.06-0.35) than subtype 1b

(relative infection median=0.27,IQR=0.17-0.36) envelopes (P=0.002) while AR4A displayed no significant difference in subtype neutralization.

Shared sensitivity profiles for AR4A and HC33.4 is also observed in the expanded panel of HCV envelopes.

As was previously reported, the new library also displayed a significant correlation in neutralization sensitivity between HC33.4 and AR4A (r=0.44 $p=7x10^{-7}$) (Fig 1). Interestingly, this correlation was not as strong as we observed in our previous report that was based on just 19 isolates, suggesting that the expanded library now contains additional unshared resistance polymorphisms to each broadly neutralizing antibody.

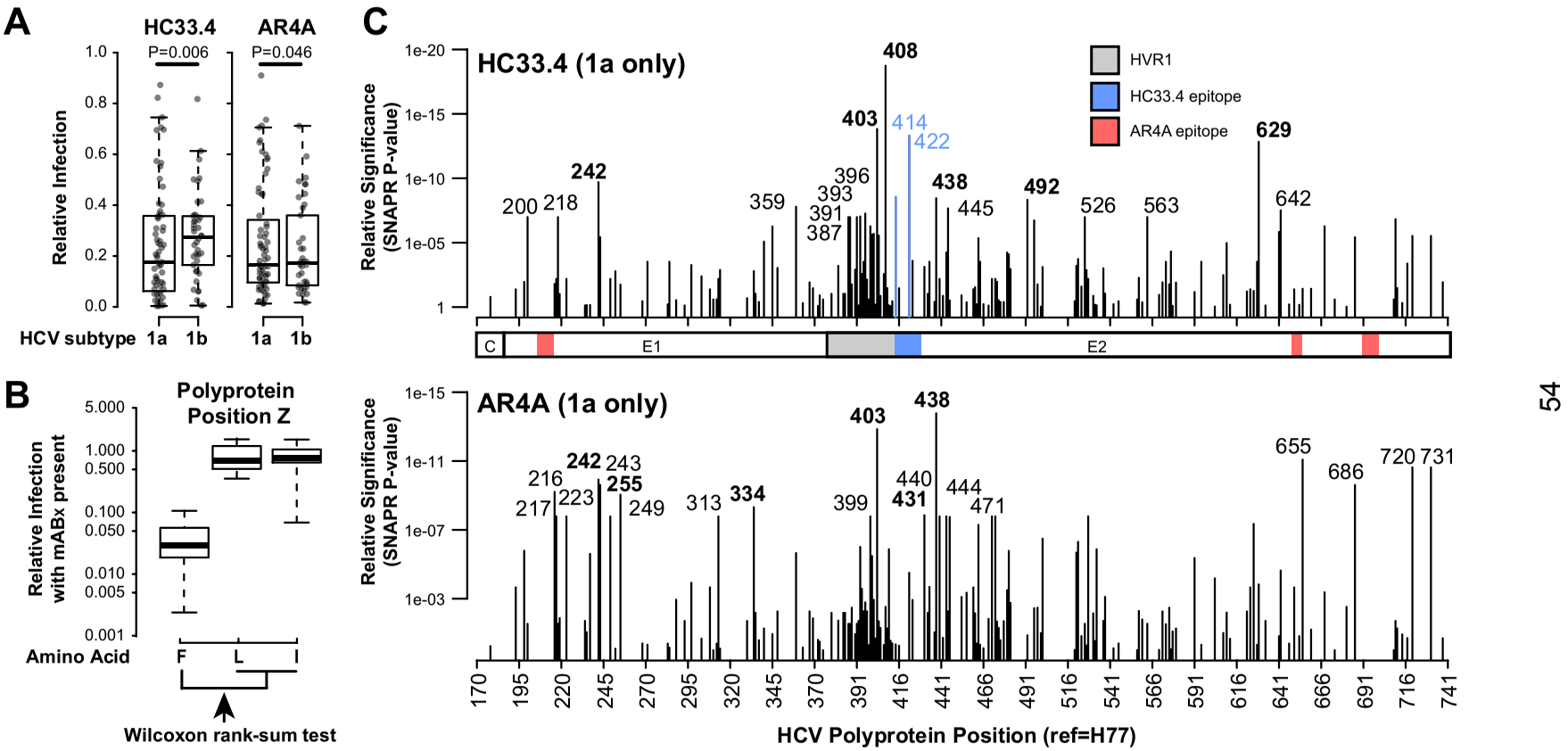
A
HC33.4
AR4A
P=0.006
P=0.046
Relative Infection
1.0
0.8
0.6
0.4
0.2
0.0
HCV subtype 1a 1b 1a 1b
B
Polyprotein Position Z
Relative Infection with mABx present
5.000
1.000
0.500
0.100
0.050
0.010
0.005
0.001
Amino Acid F L I
Wilcoxon rank-sum test
C
HC33.4 (1a only)
AR4A (1a only)
Relative Significance (SNAPR P-value)
1e-20
1e-15
1e-10
1e-05
1
1e-11
1e-07
1e-03
HVR1
HC33.4 epitope
AR4A epitope
C
E1
E2
HCV Polyprotein Position (ref=H77)

Figure 2-2. Identification of position-specific sequence variation that significantly relates to neutralization sensitivity. (**A**) Subtype differences in relative infection. (**B**) At each envelope alignment position, isolates were divided into group according to amino acid present. Relative infection (infection in the presence of mAb/infection in the presence of nonspecific IgG) values from the most sensitive group (as determine by the medians) was compared to the relative infection associated with all other amino acid choices as a single group. (**C**) The resulting significance values (after adjustment for isolate subject origin) using only subtype 1a clones against HC33.4 and AR4A. Previously defined HC33.4 and AR4A binding epitopes are indicated (blue and red), as is hypervariable region 1 (HVR1) (gray).

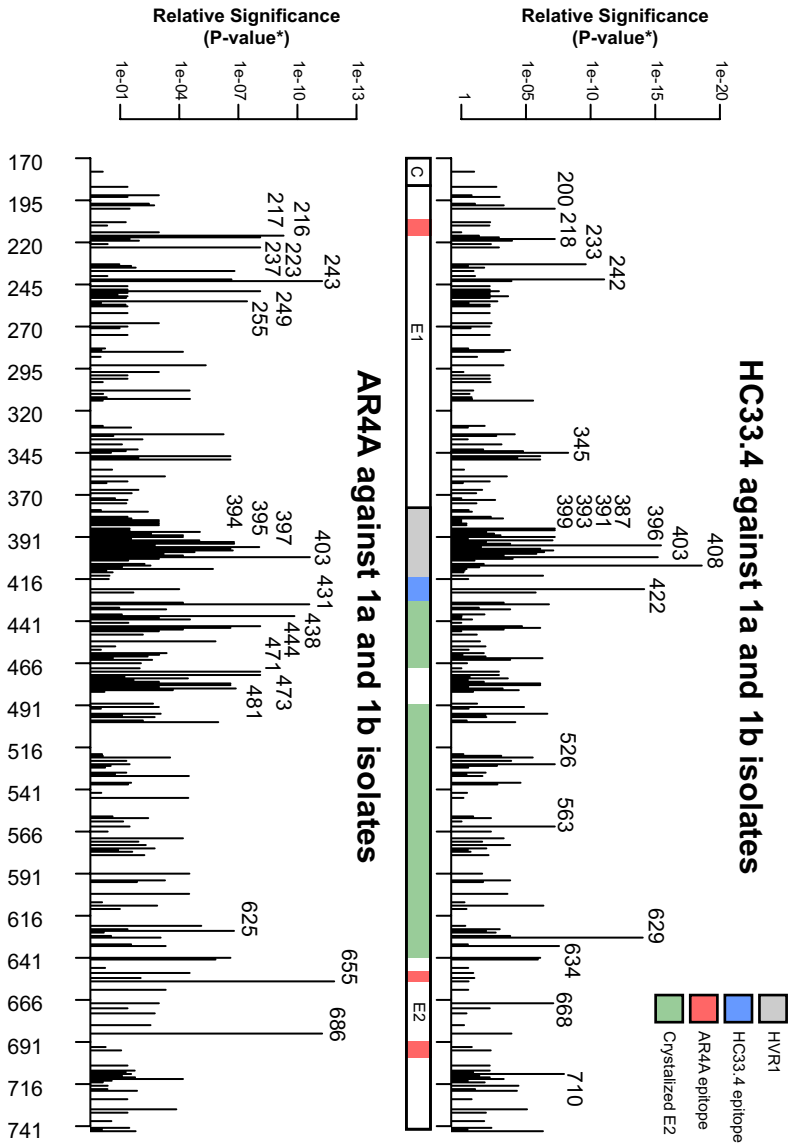

Figure 2-3. SNAPR analysis of all genotype 1 isolates. The significance values (after adjustment for isolate subject origin) using all genotype 1 clones against HC33.4 and AR4A. Shown epitopes are based on data from previous reports. Previously defined HC33.4 and AR4A binding epitopes are indicated (blue and red), as is hypervariable

region 1 (HVR1) (gray), and the portion of E2 included in the E2 core crystal structure determined by Kong *et al* (86) (green).

Subject-adjusted Neutralizing Antibody Prediction of Resistance-polymorphisms (SNAPR) and (LASSO): using variation in sequence and neutralization sensitivity to predict resistance polymorphisms in the HCV envelope. Nearest-neighbor joining alignments were built and at each amino acid position all envelopes were divided into groups according to the residue occupying that position. The neutralization values at those sites were compared and a Wilcoxon-rank sum test was performed comparing the neutralization values of all clones carrying the amino acid associated with the highest neutralization sensitivity with the neutralization values of all of the other clones (**Figure 2-2B**). Due to the higher degree of similarity among isolates originating from the same subject, neutralization data from isolates from underrepresented subjects were randomly selected for repeating until the number of data points in the library representing each individual was identical. Because of the potential for subtype differences to dominate findings, grouped genotype (**Figure 2-3**) and subtype 1a only analyses were performed separately. Grouped and subtype 1a only p-values spanned approximately 10 and 20 orders of magnitude for HC33.4 and AR4A respectively (**Figure 2-2**).

HC33.4		
Polymorphism	P-value SNAPR	Coefficient LASSO
X408M	**1.82E-19**	**3.30E-02**
X403L	**1.48E-14**	**-4.34E-02**
X629V	**1.47E-13**	**2.79E-02**
X242M	**2.00E-10**	**-3.76E-02**
X438L	**3.55E-09**	**-2.94E-02**
X492R	**4.74E-09**	**-2.91E-02**
X445Y	2.17E-08	2.84E-02
X396T	5.45E-08	4.04E-02
X393G	9.21E-08	-4.42E-02
X496I	1.86E-07	-3.28E-02
X399L	5.31E-07	4.29E-02
X641E	1.47E-06	-3.67E-02
X401N	2.04E-06	4.05E-02
X404S	2.77E-06	3.38E-02
X610N	1.07E-05	-2.91E-02
X431E	7.47E-04	-2.80E-02
X303G	4.08E-03	3.71E-02
X388S	1.72E-02	4.27E-02
X446K	3.06E-01	3.06E-02
X402L	6.45E-01	2.86E-02

AR4A		
Polymorphism	P-value SNAPR	Coefficient LASSO
X438L	**1.68E-14**	**-4.24E-02**
X438V	**NA**	**3.37E-02**
X403L	**1.40E-13**	**-4.74E-02**
X242M	1.20E-10	-3.34E-02
X255T	**9.04E-10**	**3.39E-02**
X334V	**4.80E-09**	**3.14E-02**
X431E	**1.41E-08**	**-3.96E-02**
X444Y	1.65E-08	2.95E-02
X501Q	3.25E-07	-2.91E-02
X410N	1.34E-06	3.35E-02
X359M	2.23E-06	2.91E-02
X401G	1.16E-03	-3.74E-02
X396T	**1.63E-03**	**3.13E-02**
X397S	5.46E-03	3.70E-02
X397R	5.46E-03	-3.00E-02
X384T	1.88E-02	4.01E-02
X384S	NA	-3.10E-02
X387Y	2.88E-02	3.09E-02
X405Q	4.98E-02	-3.87E-02
X402L	2.03E-01	4.13E-02

Table 2-1. Agreement between Subject-adjusted (SNAPR) and phylogenetic-adjusted (LASSO) methods resistance polymorphism predictions based on the same data set. The 20 most significant resistance polymorphism predictions by LASSO for HC33.4 and AR4A, cross referenced with SNAPR P-values. Predictions that were identified in the top 20 both predictions of both algorithms are shown in bold.

Because variation in the distance between any given two clones in the library, we also implemented a method that adjusts specifically for phylogeny, and not just by subject. The same data set was analyzed using this modality. The 20 most significant predictions from LASSO for HC33.4 and AR4A (**Table 2-1**) were compared to the top 20 most significant SNAPR predictions yielding 6 and 7 resistance-polymorphism predictions respectively.

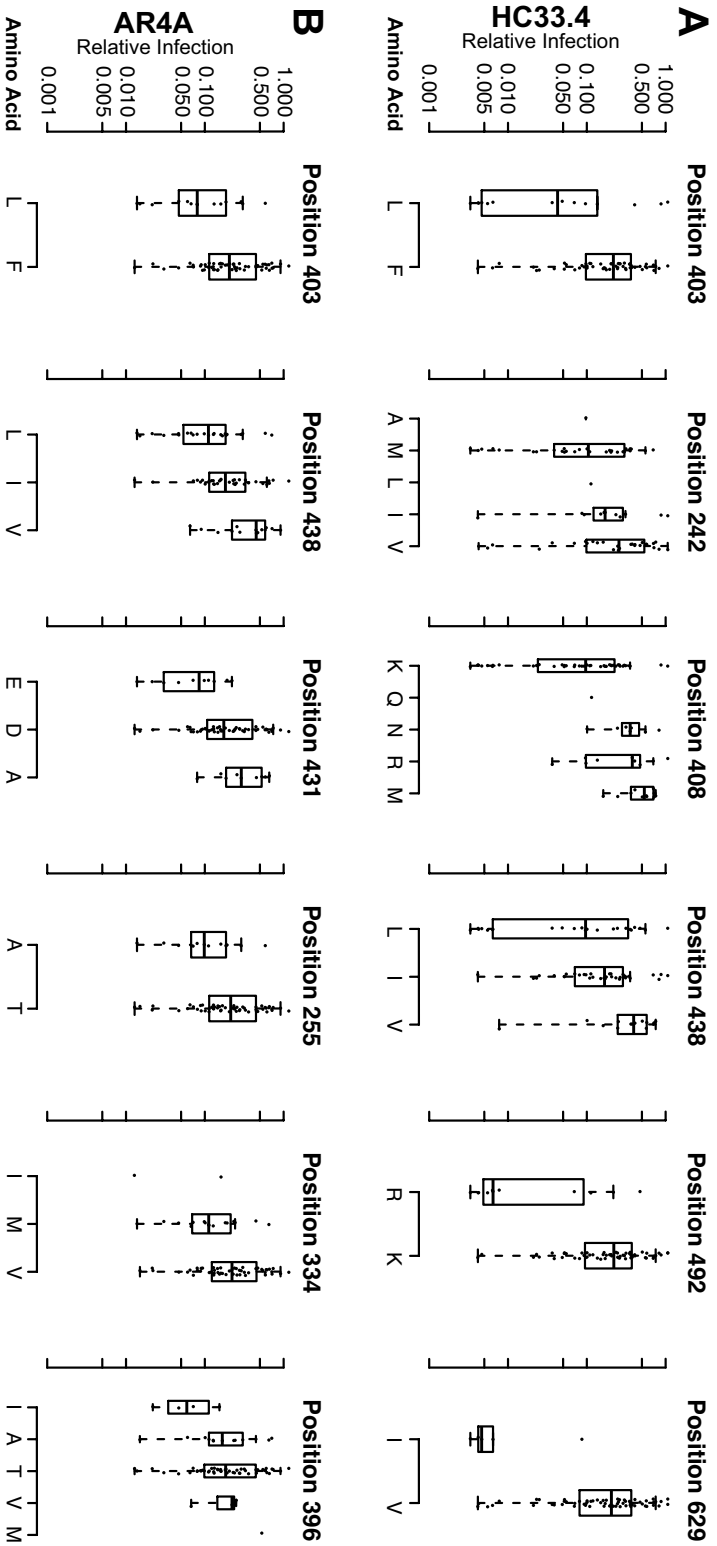
A
HC33.4
Relative Infection
1.000
0.500
0.100
0.050
0.010
0.005
0.001
Amino Acid
Position 403
L
F
Position 242
A
M
L
I
V
Position 408
K
Q
N
R
M
Position 438
L
I
V
Position 492
R
K
Position 629
I
V
B
AR4A
Relative Infection
1.000
0.500
0.100
0.050
0.010
0.005
0.001
Amino Acid
Position 403
L
F
Position 438
L
I
V
Position 431
E
D
A
Position 255
A
T
Position 334
I
M
V
Position 396
I
A
T
V
M

Figure 2-4. Comparison of relative infection of HCVpp produced using E1E2 clones with each amino acid residue at positions predicted to influence bnAb sensitivity. Boxplots showing the relative infection values for all isolates grouped by amino acid present at the indicated position in the presence of HC33.4 (**A**) and AR4A (**B**) in order of significance.

HC33.4 and AR4A resistance polymorphisms are distributed outside of described epitopes for each. Both SNAPR and LASSO predicted that X403L, X242M, X408M, X438L, X492R, and X629V were determinants of neutralization sensitivity for HC33.4 although the described binding epitope is between polyprotein positions 410-423. Similarly, X403L, X438L, X431E, X244T, X438T, X334V, and X396T were predicted to be determinants of neutralization sensitivity to AR4A by both algorithms though the described epitope is split between the amino (polyprotein positions 200-210) and carboxy (polyprotein positions 650-710) termini of E2 (**Figure 2-3**).

Polyprotein Position 403 is significantly associated with HC33.4 and AR4A neutralization efficiency. Given the imperfect correlation in neutralization sensitivity between HC33.4 and AR4A, it is not surprising that many of loci predicted to be determinants of sensitivity to HC33.4 and AR4A are not shared. Interestingly, multiple analytic modalities revealed that 403 as a shared determinant of neutralization sensitivity for HC33.4 and AR4A.

For HC33.4, a leucine occupation at polyprotein position 403 was associated with a median (IQR) 0.042 (0.0048 to 0.13) fraction infection in

the unadjusted data and a median (IQR) 0.0064 (0.0047 to 0.064) fraction infection in the SNAPR adjusted data. In contrast, phenylalanine was associated with a median (IQR) 0.22 (0.097 to 0.36) fraction infection in the unadjusted data and a median (IQR) 0.22 (0.10 to 0.38) fraction infection in the SNAPR adjusted data (**Figure 2-4A**). A Wilcoxon rank-sum comparison of the HC33.4 fraction infection of isolates containing leucine with values of those containing phenylalanine resulted in a P-value of $9.9x10^{-4}$ in the unadjusted data, and a P-value of 1.5e-14 in the subject adjusted data. LASSO analysis assigned a coefficient of -0.044 for L403F with HC33.4, the strongest coefficient produced (**Table 2-1**).

For AR4A, a leucine occupation at polyprotein position 403 was associated with a median (IQR) 0.080 (0.050 to 0.18) relative infection in the unadjusted data and a median (IQR) 0.067 (0.017 to 0.13) fraction infection in the SNAPR adjusted data. In contrast, phenylalanine was associated with a median (IQR) 0.20 (0.11 to 0.44) relative infection in the unadjusted data and a median (IQR) 0.22 (0.14 to 0.45) fraction infection in the SNAPR adjusted data (**Figure 2-4B**). A Wilcoxon rank-sum comparison of the AR4A fraction infection of isolates containing leucine with values of those containing phenylalanine resulted in a P-value of 0.012 in the unadjusted data, and a P-value of 1.4e-13 in the subject-adjusted data. Analysis with lasso coefficient of -0.043 for L403F with AR4A, the second strongest coefficient produced (**Table 2-1**).

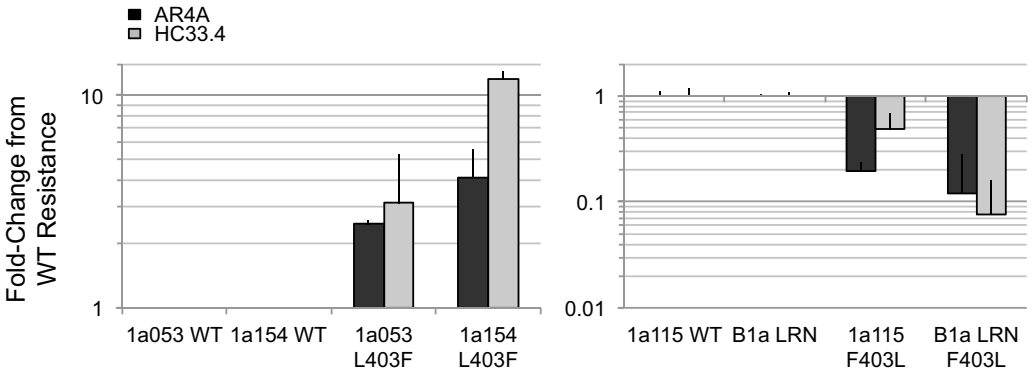

Figure 2-5. Fold change in neutralization resistance after site-directed mutagenesis at polyprotein position403 in multiple unrelated E1E2 clones. (**A**) Fold change in resistance of HCVpp with two unrelated wild type (WT) E1E2 proteins or HCVpp with the same E1E2 proteins after introduction of L403F. (**B**) Fold change in neutralization of HCVpp with two unrelated wild type (WT) E1E2 proteins or HCVpp with the same E1E2 proteins after introduction of F403L.

Mutations introduced into sensitive and resistant clones validate 403 as a shared resistance polymorphism. To test the prediction that amino acid occupation at position 403 could be a shared determinant of sensitivity to HC33.4 and AR4A, we introduced a leucine (the amino acid associated with sensitivity) into multiple resistant HCV envelopes bearing a phenylalanine (the amino acid associated with resistance) using site-directed mutagenesis (SDM). SDM of leucine at position 403 to phenylalanine (L403F), in the E1E2 background isolates 1a53 and 1a154 (H77), which differ at 46 amino acids out of the 581 amino acid alignment, conferred a 3.0-12.3-fold increase in resistance to HC33.4 ($P<0.01$ for each background). Interestingly, the same mutations conferred 2.5-5.3-fold increase in resistance to AR4A ($P<0.01$ for each background)(**Figure 2-5A**).

Conversely, site directed mutagenesis of phenylalanine at position 403 to leucine (F403L), in the E1E2 backgrounds 1a115 and Bole1a (subtype 1a ancestral sequence), which differ at 38 amino acids out of the 581 amino acid alignment, conferred a 3.0-12.3-fold increase in resistance to HC33.4 ($P<0.01$ for each background). The same mutations also conferred 2.5-5.3-fold increase in resistance to AR4A ($P<0.01$ for each background)(**Figure 2-5B**).

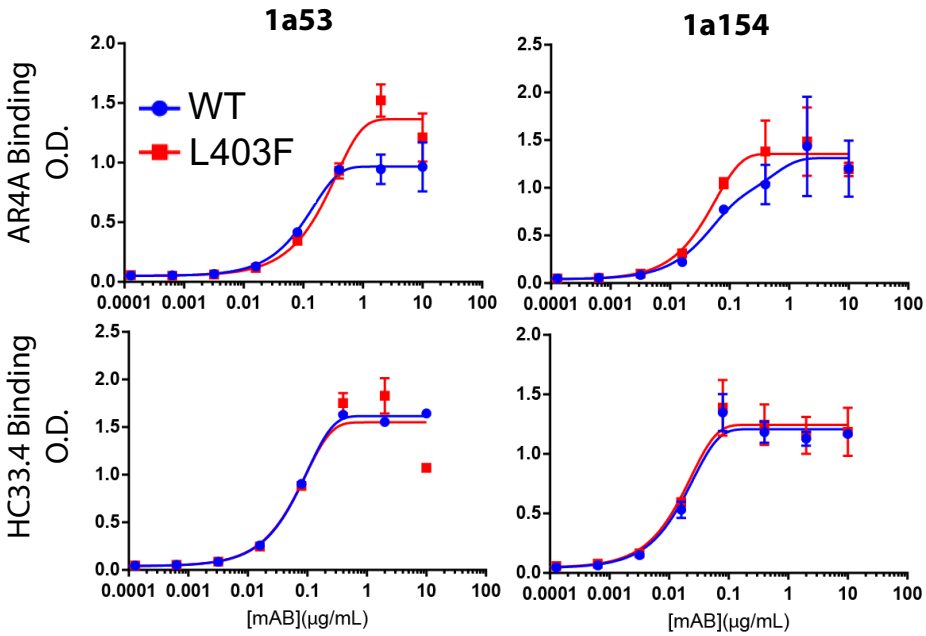

Figure 2-6. Binding of HC33.4 and AR4A is not impacted by SDM at position 403. Binding of HC33.4 or AR4A to wildtype (WT) E1E2 proteins or E1E2 with leucine mutated to phenylalanine at position 403.

Discussion

We have developed a high-throughput platform for prediction of HCV bnAb resistance polymorphisms and identified an unexpected shared resistance polymorphism for two highly potent bnAbs for HCV not contained within the described epitopes for either.

As our ability grows to query larger and larger sets of naturally occurring HCV isolates for their sensitivity to broadly neutralizing antibodies, so does our understanding of variation in sensitivity—thereby addressing a key barrier to developing an effective prophylactic vaccine to HCV. In a previous report, we found that antibodies tend to cluster into groups with respect to the isolates that they neutralize the most potently. Conventional wisdom would predict that these clusters are defined by their shared epitopes. The observation that a shared and unshared resistance polymorphisms to two broadly neutralizing antibodies that fall into the same neutralization cluster are located at positions in neither antibody's described epitope suggests that these clusters are defined by something far more complex. One theory is that HCV varies greatly in its receptor dependence (as it has >4 described entry factors) and that these clusters are defined by a group of isolates with a specific preference in receptor tropism. In this theory, all the antibodies that fall into the cluster interrupt a specific entry preference. Interestingly, we profiled the binding affinity of SDM envelopes at position 403 and found no difference in binding affinity

(**Figure 2-6**) further suggesting that the mechanism of resistance confirmed by this polymorphism and the features that define neutralization clusters are more complex than simple antibody-epitope interactions.

Recent visualization of the E2 protein in complex with a bnAb has been informative (86), however, large deletions in E2 to facilitate crystallization preclude analysis of many putative antibody/epitope interactions. To simulate global diversity, broad neutralization has typically been assessed through evaluation of neutralization potency in several representative clones per genotype in the context of a replication competent genotype 2a HCV genome (77). While HCV E1E2 full-length chimeras allow for observation of the extended life-cycle of HCV infection, there limited availability means that they encapsulate only a minority of intragenotypic variation.

While we were able to identify an unexpected resistance polymorphism to multiple bnAbs, there were limitations to the study design and approach. We only sampled 97% of the naturally occurring polymorphisms that exist at a ≥5% threshold in a large set of Genbank HCV genotype 1 sequences. When the minimum frequency threshold for polymorphism incidence is reduced to ≥1%, the coverage is reduced to 78%. As has been observed in the emergence of DAA resistance, even resistant clones with <1% pretreatment frequency within a single host is sufficient for treatment failure (87). More extensive panels are required to probe rarely occurring natural polymorphisms in HCV. Though all

envelopes were isolated from individuals with ongoing viremia only ~50% of the isolates produced functional envelopes in the HCVpp system. This may reflect the distribution of functional viruses that circulate in the blood of infected patients but likely represents a bias with unknown consequence—a limitation existent at a greater frequency in all *in vitro* systems for HCV to date. Due to the lack of robust animal models for HCV infection that represent similar inoculum requirements, frequency of spontaneous resolution, and viral replication/diversity, it is likely that the only validation for broad neutralization and resistance will be in the form of vaccine efficacy studies. While these studies are ongoing, we provide another line of reasoning that supports vaccination strategies that promote the generation of antibody responses spanning more than one neutralization cluster.

In conclusion, we sampled 113 naturally occurring HCV envelopes for their sensitivity to two of the most potent broadly neutralizing antibodies characterized to date and found significant variation in isolate sensitivity and identified a shared resistance polymorphism located outside the predicted binding epitope of both antibodies. We have previously identified that neutralizing antibodies fall into distinct clusters with respect to which isolates they neutralize most potently. Here we have presented evidence that one feature of this clustering is far more complex than simple epitope recognition. More extensive panels are required to

identify more rare but potentially significant resistance polymorphisms for broadly neutralizing antibodies.

This page is intentionally blank.

Chapter 3: Identification of novel interferon-induced restriction factors for HIV-1 from inter-human variation in activated T-cell transcription and virus dynamics *in vivo*.

Abstract

The production of type 1 interferons is critical to host inhibition of Simian and Human Immunodeficiency viruses (SIV and HIV respectively)(88–90). Though there are hundreds of interferon-stimulated genes (ISGs), which of these restrict HIV-1 replication is largely unknown (with several notable exceptions(91–95)). To identify ISGs that restrict HIV-1 replication, we selected HIV-1 permissive cells (CD38+/HLA-DR+/CD4+/CD3+ lymphocytes, or activated CD4+ T-Cells) from 19 humans with untreated HIV-1 infection before and after administration of pegylated-interferon alpha 2a (IFN) and studied RNA abundance using RNA sequencing (RNAseq). By comparing the target cell RNA abundance induced by IFN by the resultant reduction in plasma HIV-1 RNA, we identified 13 ISGs that might explain how interferon restricts HIV-1 replication. Included were recognized (e.g.s, MX2, APOBEC3A) and mostly novel HIV-1 (e.g.s, BCL2L14, CMPK2, and LAMP3) candidate HIV-1 restriction factors. Using principal component analyses to identify ISGs that independently restrict HIV-1 replication, we selected CMPK2 and BCL2L14 for further study novel candidate restriction factors. Each of these and MX2 (as control) was confirmed ($P<0.05$) by using RNA interference to attenuate the effect of interferon on HIV-1 replication in cell

culture. These studies provide new insights into how type-1 interferons restrict in humans the replication of HIV-1 and possibly other chronic viral infections.

Introduction

Interferons are host defense cytokines that are found in all vertebrates (96). Their release coordinates the expression of hundreds of cell-autonomous defense genes (interferon-stimulated genes; ISGs) to control intracellular pathogens, especially viruses. Interferon-alpha (IFN), a quintessential type 1 interferon, has been used to treat chronic virus infections, and has activity against HIV-1: HIV-1 infected individuals treated with IFN experience a decline in the abundance of plasma viral RNA (97) and the number of cells harboring viral genomes (98), presumably by inducing ISGs that restrict HIV-1 replication. The interferon-induced cytidine deaminase APOBEC3G introduces stop codons into retroviral genome intermediates preventing their ability to produce infectious progeny (4). These inactivated viruses shape the functional diversity of HIV-1 genomes within the latent reservoir, an identified barrier to cure (99). Several other HIV-1 restriction factors that are ISGs have been described, such as SAMHD1, BST2, and MX2, but they do not fully account for the antiviral effects of IFN. Hundreds of ISGs are induced upon IFN administration; however, only a select group of them are likely to be responsible for restricting HIV-1 replication. We searched for novel ISGs that are also HIV-1 restriction factors in infected people who received IFN.

The natural history of HIV-1 infection, which includes declining CD4+ T-cell abundance without a decrease in circulating virus, suggests

that viral replication is relatively insensitive to changes in the availability of uninfected cells. Indeed, established models also suggest that viremia is robust to changes in the availability of uninfected cells and much more sensitive to changes in the rates of infection and virus production in permissive cells Thus, we reasoned that variation in the induction of ISGs with restrictive capacity for HIV-1 in infected and otherwise permissive cells determines the variation in the decline in HIV-1 viremia in response to IFN *in vivo*. We sought to identify novel restriction factors for HIV-1 by comparing person-to-person differences in HIV-1 restriction and transcription dynamics in activated CD4 T-cells from HIV-1 infected patients through IFN treatment using a novel analytical approach to human cohort RNA sequencing (RNAseq).

Materials and Methods

Modeling the contributions of various features of HIV-1 infection. The three differential equations $dx/dt = \lambda - dx - \beta xv$, $dy/dt = \beta xv - ay$, and $dv/dt = ky - uv$ (λ =production rate of uninfected cells, x=number of uninfected cells, d=decay rate of uninfected cells, β=first order rate constant of infection, v=plasma viral load, y=number of infected cells, a=decay rate of infected cells, and k=first order rate constant of virus production in infected cells) were integrated using "integrateODE" function in the 'mosiac' package version 0.9.2-2 in R (www.r-project.org) version 3.1.2 for 1000 days using standard starting values (x=1e6,y=1,v=100000, λ=1e5,d=0.1,a=0.5,B=2e-7,k=100,u=5) (100–102). The solution values were then used as input into another integration but varying either x, β, k, or a by the indicated factor and run to simulate an additional 3 weeks (300 days for a) with these new imposed values.

Human subjects. The primary outcome was to compare gene expression in activated T-cell with HIV RNA changes following administration of IFN (ΔHIV_{IFN}): 20 persons with HIV-HCV coinfection were enrolled in a prospective study of HIV and HCV viral kinetics pre-antiretroviral therapy (ART). Persons with chronic HIV and HCV infections were recruited from the Johns Hopkins HIV Clinic, the Baltimore City Sexually Transmitted Diseases Clinic, and other area clinics. HIV infection was established by

detection of HIV-1 antibodies and an HIV RNA level >400 c/mL; chronic HCV infection was determined by detection of HCV antibodies and HCV RNA >100,000 IU/mL for >6 months. Subjects had received <24 months of ART over their entire lives and none within 6 months. Subjects also were excluded if HBsAg was detected in plasma; they were pregnant; there was a history of severe depression or any uncontrolled disease; platelet count was <50,000/mm^3; there was a contraindication to use of raltegravir, tenofovir DF, or emtricitabine. Because therapy was judged too urgent to wait for study procedures as per current treatment guidelines for HIV and HCV, persons at screening whose CD4+ T lymphocyte counts were <200/mm^3 or who had cirrhosis were excluded. A total of 32 patients were screened to identify 20 study subjects. All subjects gave written informed consent to the protocol as approved by the Johns Hopkins Institutional Review Board.

Subjects were admitted to the Johns Hopkins Hospital Clinical Research Unit. A single 1.5 μg/kg of peginterferon alpha 2b (IFN) was administered subcutaneously. Thereafter, blood was collected every 6 hours and the patient was discharged 24 hours after the IFN dose. The patient returned 48 hours, 72 hours, 7 days, and 14 days after the IFN dose. At day 14, ART consisting of raltegravir, tenofovir, and emtricitabine was given by the Johns Hopkins Research Pharmacy Investigational Drug Service. The Johns Hopkins University School of Medicine Institutional Review Board approves this study.

Laboratory Testing. Unless otherwise indicated, all laboratory testing was performed in the clinical laboratory of the Johns Hopkins Hospital. Single-nucleotide polymorphisms (SNPs) at position rs12979860, which lies upstream of *IFNL3*, have been strongly associated with HCV RNA changes to IFN. To assess *IFNL3* status, DNA was extracted from peripheral blood mononuclear cells using the QIAamp DNA Blood Mini Kit (Qiagen), and SNP genotype at position rs12979860 was performed with TaqMan custom SNP genotyping assays (Life Technologies) and using a Roche LightCycler 480 Real-Time System (Roche Applied Science, Indianapolis, IN).

Viral RNA Testing. To reduce interassay variance, HCV and HIV RNA testing for a given subject was done at the same time on plasma centrifuged within 30 minutes of collection and stored at −20°C for up to 25 hours and then at −80°C until testing. HCV RNA testing was done using the RealTime HCV Amplification Reagent Kit Abbot). To provide information in “real time,” such as for screening into the study, additional HCV RNA tests were done by the commercial laboratory of the Johns Hopkins Hospital using the Roche Cobas AmpliPrep/Cobas TaqMan HCV Test, v. 1.0. Although both were reported in international units, analyses were primarily done on results from one or the other laboratory. HIV RNA testing was done using the RealTime HIV Assay (Abbott). HCV genotype

was determined in the Johns Hopkins Hospital clinical laboratory by direct sequencing of the Core-E1 regions of the HCV genome. CD4+ T-lymphocyte count was measured by flow cytometry of whole blood that was delivered to the Johns Hopkins Hospital clinical laboratory.

Isolation of activated CD4 T-cells. Peripheral blood mononuclear cells (PBMCs) were separated from whole blood and frozen as previously described (103). Freshly thawed cells were washed and incubated with CD3-FITC (Biolegend, San Diego, CA), CD4-PECy7 (Biolegend), CD8-APC (BD Biosciences), HLA-DR-PE (Biolegend), and CD38-BV421 (Biolegend) 40 minutes at 4ºC per the manufacturer's recommendation. Immediately before sorting, plasma membrane compromised cells were labeled with propidium iodide (Sigma). Flourescence-activated cell sorting (FACS) was performed on a MoFlo Legacy Sorter (Beckman-Coulter) at the Johns Hopkins School of Public Health Flow Cytometery Core Facility. The population of interest was sorted directly into ≥4 volumes of Quick-RNA™ MicroPrep lysis buffer (Zymo Research) per the manufacturer's recommendation. Sorting was stopped when the number of sorted cells reached 125,000 cells although many samples did not reach this number. Flow-cytometry analysis on two randomly selected post-sort samples revealed >95% purity. Sorted samples were vortexed, incubated for 10 minutes at room temperature, vortexed again, and frozen at -80ºC until isolation.

Activated T-Cell RNA isolation. Isolation was performed using the Quick-RNA™ MicroPrep kit (Zymo Research, Irvine, CA) according the manufacturer's protocol without the on-column DNAase treatment. The eluate was treated with DNAse-I (Qiagen, Germantown, MD) according to the manufacturer's protocol then purified and concentrated using the RNA Clean-up and Concentrator kit (Zymo Research, Irvine, CA) according the manufacturer. The high-sensitivity assay for RNA or DNA was performed on RNA isolations and cDNA libraries respectively using a 2100 Biolanalyzer (Agilent).

Library preparation and sequencing. Complementary DNA (cDNA) libraries were produced using the Ovation® Single-cell RNAseq kit (NuGEN) according to the manufacturer's specifications. Briefly, reverse transcription was carried out using a random hexamer to oligodT ratio of 50:1 and unique barcodes for each individual's samples were ligated to ~250bp enzymatically fragmented molecules. All samples were linearly amplified using 19 cycles of PCR using primer annealing sites contained within the adapters. All sequencing was performed on a HiSeq2500 (Illumina) at the Johns Hopkins Genetics Research Core Facility. Each individual's samples were pooled three to each lane (one other sample from a third time point was included from each individual for an unrelated

project) in an equimolar fashion and read in rapid-run mode to produce 100bp paired-end reads.

Sequence Mapping. To avoid resampling of fragments due to PCR duplication, identical sequences were removed from data as is standard practice. Reads were assembled with RSEM (104) using hg19 as the reference human genome. Transcript variant analysis was performed using all annotated UCSC isoforms as of 7/15/2014.

Differential Expression Calculation. Differential expression of genes and isoforms was calculated using EBseq, an empirical Bayes hierarchical model for expression analysis of RNAseq data (105).

Statistical Analyses. Comparison values, including post-correction fold-change, post-correction probability of differential expression (PPDE), and post-correction probability of equal expression (PPEE) values, were further analyzed in using ‘stats' in R version 3.1.2. Measurements for which neither the PPDE or PPEE was ≥0.95 were discarded. In the determination of significance of the gene’s change across the cohort, all fold-change calculations with PPEE ≥ 0.95 were set to 0. If there were remaining measurements for a gene in >10 of the individuals, a two-sided one-sample T-test was performed on $\log_2$ transformed fold-change values. The resulting P-values were adjusted for multiple comparisons using the

Benajami-Hochberg method (38). Genes with adjusted P-values ≤0.05 were considered ISGs. Spearman rank-correlations were performed between fold-changes between the observed ISGs and viral load decline with fold-changes with PPEE ≥ 0.95 retained as their original values. Pairwise correlation plots were constructed using 'corrplot' version 0.73. K-means clustering was performed using 'cluster' version 1.15.3. Wilcoxon rank-sum tests were performed on untransformed data in discovery analyses unless indicated otherwise. Two-sample unpaired one-way t-tests were performed in cell-culture studies as these experiments were performed to validate observations previously made *in vivo*.

***In vitro* IFN treatments.** Six cell lines THP-1, MT4, MOLT4, A3.01, PM1, and Jurkat cells (NIH AIDS Reagent Bank) were treated with 1,000 U/mL IFN for 0,6,12, and 24 hours and RNA was isolated using the QuickRNA microprep kit (Zymo Research). Reverse transcription was performed using SuperScript III (Invitrogen) according to the manufacturer's protocol using oligo-dT primers only. Quantitative PCR was performed using LightCycler SybrGreen master mix (BioRad) according to the manufacturer and run on the LightCycler 480 (BioRad) according to the manufacturer's protocol. Fold-changes are shown in reference to untreated samples and normalized to the geometric mean of Ct values from RPL13, RPL37, and β-actin (Integrated DNA technologies, assay IDs Hs.PT.58.47294843, Hs.PT.58.1213197, and Hs.PT.56a.40703009.g

respectively) measured using the LightCycler 480 Probes master mix (BioRad).

Virus Preparation. The full-length HIV-IIIB, pNL4.3 delta nef GFP or pNL4.3 delta env (mock) plasmids were transfected into 293T cells and supernatants were spinnoculated onto PM1 cells (NIH AIDS Reagent Bank) for 2 hours at 1,200g. Cells were then incubated at 37ºC for 10 days. Supernatants were passed through 0.22 μm filter and virus was pelleted through a 20% sucrose layer for 2 hours at 150,000g. Pellets were resuspended in RPMI containing 10% FBS and stored at -80ºC. Virus concentration was determined using the 96-well format Alliance HIV-1 P24 Antigen Elisa kit (Perkin-Elmer, Waltham, MA) according to the manufacturer's protocol.

Testing cell lines for their IFN-mediated restriction potential for HIV-1. The indicated five cell lines (NIH AIDS Reagent Bank) were centrifuge inoculated with varying amounts of HIV-IIIB. For each inoculum concentration, 200,000 cells were centrifuge inoculated with 50μL of RPMI containing 10% FBS at 1,200g for 2 hours. Cells were allowed to rest for 24 hours at 37ºC and then each inoculum was split into two wells, one that received no treatment while the other received 1,000 U/mL IFN. Both wells were incubated for an additional 72 hours at 37ºC before supernatants were collected and p24 measurements were performed.

***In vitro* RNAi, infection, IFN treatment, and measurement of HIV production.** MT4 cells (NIH AIDS Reagent Bank) were transfected with SMARTpool siGENOME siRNAs (Dharmacon) negative controls or siRNAs at 100nM against the indicated gene of interest using the Amaxa Cell Line Nucleofector V kit (Lonza) according to the manufacturer's protocol. Cells were rested in recovery media (20% FBS) at 37ºC for 24 hours and 200,000 cells were centrifuge-innoculated with purified replication-competent HIV-IIIB (10pg p24) in 50uL of 10%FBS RPMI at 1,200g for 2 hours. After washing twice with media, cells were rested for 24 hours at 37ºC and split into two wells: one containing 1,000 Units IFN/mL and one containing an equal volume of the same 10% FBS RPMI used to resuspend the IFN (untreated). After 48 hours of IFN treatment, supernatants were removed, diluted 1:1,000 in DMEM containing 10% FBS. One hundred microliters of this dilution was transferred to 10,000 TZM-bl cells (NIH AIDS Reagent Bank) in 100uL DMEM containing 10% FBS and 20μg/mL diethylaminoethanol and incubated at 37ºC for 48 hours. Cells were washed with PBS and assayed for luciferase activity using the Luciferase assay kit (Promega). Fraction infection was calculated by dividing the luciferase reading from the TZM-bl cells exposed to the IFN-treated supernatants by the luciferase from the TZM-bl cells exposed to the untreated supernatants.

Results

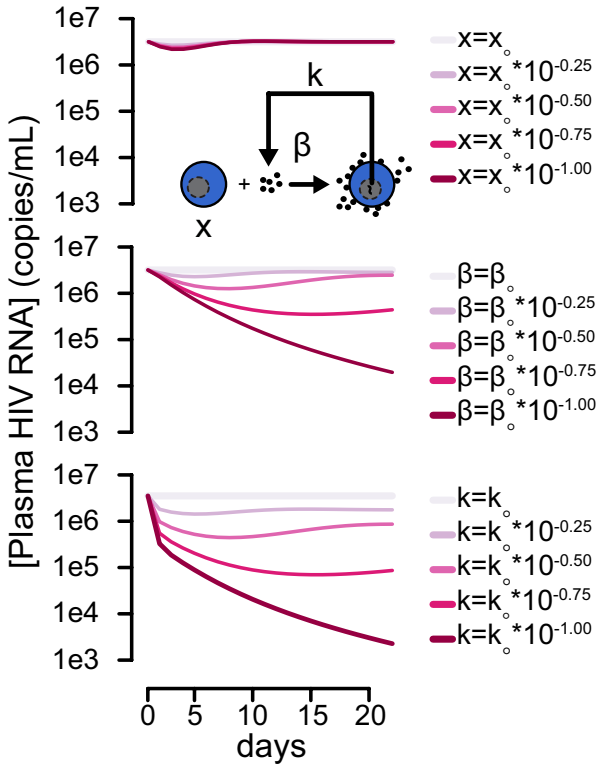

Figure 3-1. HIV-1 viremia is more sensitive to changes in the rate of infection and the rate of virus production than changes in the availability of uninfected target cells. Infections were simulated using three standard differential equations *dx/dt* = λ-dx-βxv, *dy/dt* = βxv-ay, and *dv/dt* = ky-uv (*λ* =production rate of uninfected cells, x=number of uninfected cells, d=decay rate of uninfected cells, β=first order rate constant of infection, v=plasma viral load, y=number of infected cells, a=decay rate of infected cells, and k=first order rate constant of virus production in infected cells) using the starting values (x=1e6,y=1,v=100000, λ=1e5,d=0.1,a=0.5,B=2e-7,k=100,u=5). Resulting values were then used as input into the same equations separately varying the indicated variable by the indicated factor and run for 21 days.

Modeling plasma HIV-1 RNA kinetics after IFN. We adapted standard mathematical models of HIV-1 kinetics after antiviral treatment to consider an array of ISGs as putative antivirals, accounting for how IFN might affect HIV-1 viremia (100, 102, 106). Because IFN administration is known to cause a reduction in CD4+ T-cell abundance in HIV-1 infected individuals(107), we modeled whether the effect of IFN on HIV-1 was via

ISG induction or by decreasing the availability of target cells. We hypothesized that ISGs could either affect the rate of infection of susceptible cells (β) or the rate of virus production in infected cells (k) (**Figure 3-1**). Modeling revealed that HIV-1 kinetics after IFN administration was more consistent with ISG induction affecting β or k, rather than a change in the total number of target cells.

Table 3-1. Participant Characteristics.

Characteristic	**N = 19 (100%)**
Age,*	
median (range)	**49.2** (20.8-60.6)
Sex	
Male, n (%)	**15** (78.9)
Race	
Black, n (%)	**12** (63.2)
BMI, kg/m^2*	
median (range)	**23.3** (17.7-39.0)
HIV RNA, $\log_{10}$ copies/mL*	
median (range)	**4.27** (2.91-5.44)
CD4+ T-cell count, cells/mm^3*	
median (range)	**376** (203-759)
HCV RNA ($\log_{10}$IU/mL),*	
median (IQR)	**26** (17-46)
HCV genotype,* n (%)	
1a	**15** (79)
1b	**1** (5.3)
2b	**1** (5.3)
3a	**2** (10.5)
***IFNL3* (rs12979860) genotype, n** (%)	
C/C	**1** (5.3)
C/T	**11** (57.9)
T/T	**7** (36.6)

*at baseline

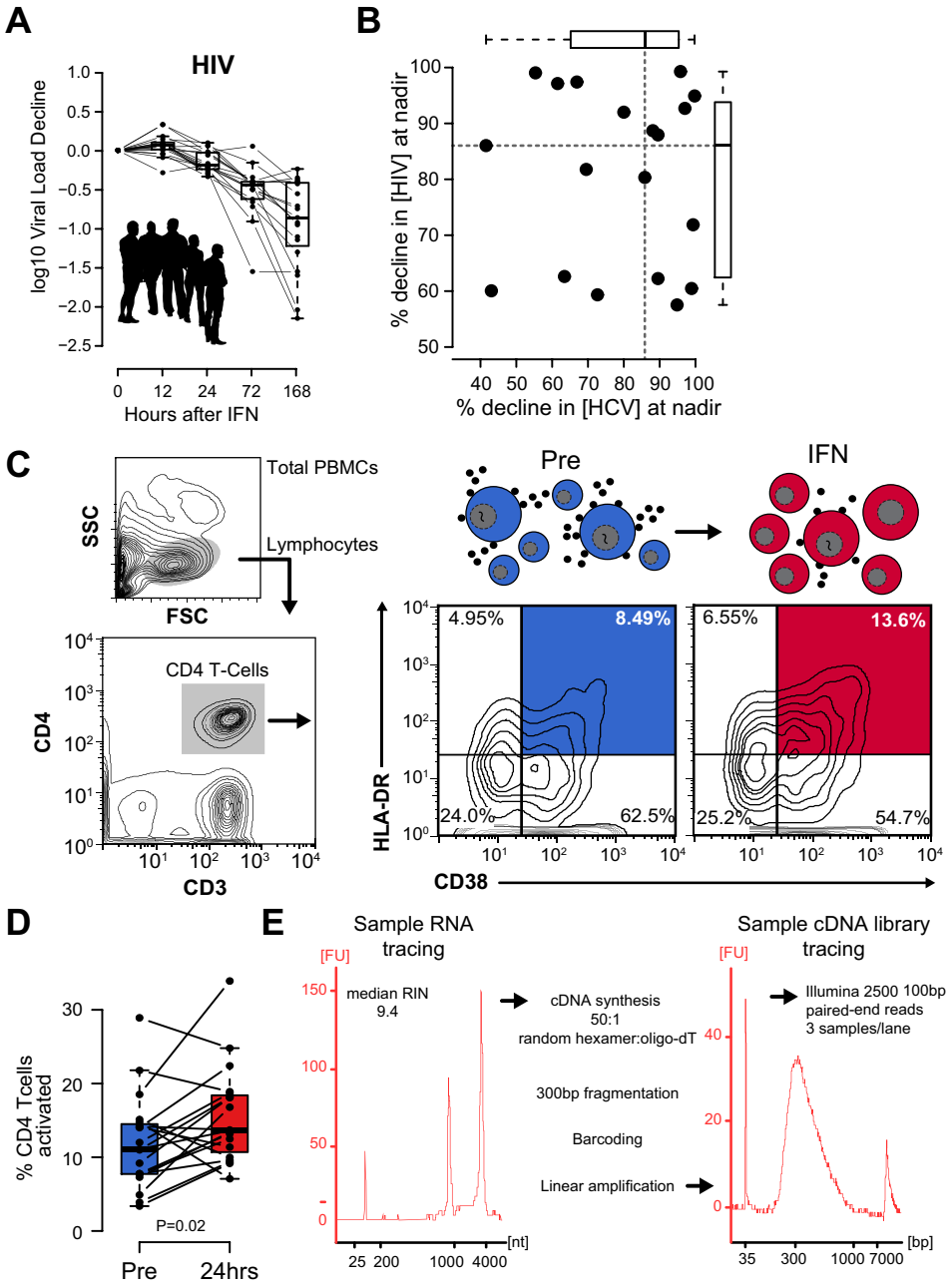

Figure 3-2. *In vivo* IFN responsiveness and activated CD4 T-cell transcriptome library preparation. (**A**) Boxplots with overlaid values of $\log_{10}$ changes in HIV and HCV RNA from baseline are shown with lines connecting values from each subject over the indicated time-points. Lines indicate longitudinal measurements from the same individual. (**B**) Scatterplot and boxplots comparing the percent decline in viremia from baseline for HIV and HCV at each subject's corresponding nadir virus measurement. (**C**) Flow cytometery data and sorting algorithm for isolation of activated CD4 T-cells from total PBMCs at baseline (Pre, blue) and after 24 hours after injection with IFN (IFN, red). (**D**) Boxplots with overlaid points showing the percent of CD4 T-cells bearing the activation markers HLA-DR and CD38 with lines connecting participants across the two measured timepoints. (**E**) Sample bioanalyzer tracing of RNA, overview of library preparation, and sample bioanalyzer tracing of prepared cDNA library. Fluorescence units, FU; RNA or cDNA length in nucleotides (nt).

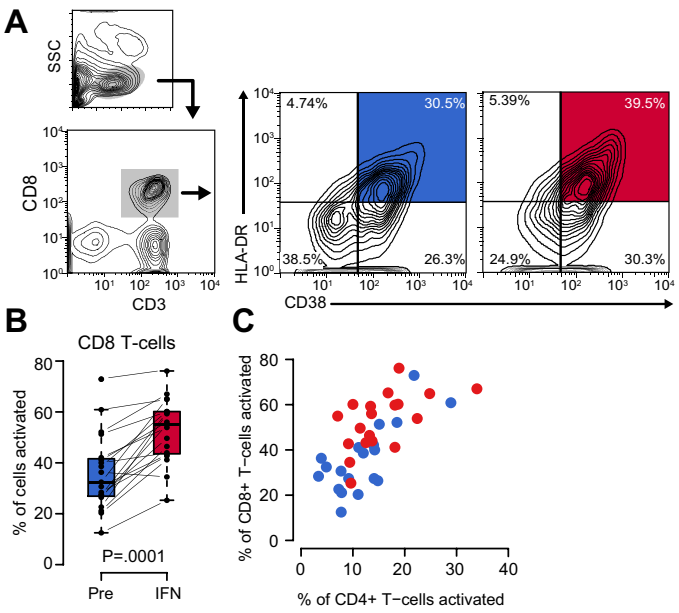

Figure 3-3. Activation status changes in CD8 T-cells reveal that IFN activated both CD4 and CD8 T-cells. (**A**) Flow cytometery profiling of activated CD8 T-cells from total PBMCs at baseline (Pre, blue) and after 24 hours after injection with IFN (IFN, red). (**B**) Boxplots with overlaid points showing the percent of CD8 T-cells bearing the activation markers HLA-DR and CD38 with lines connecting participants across the two measured timepoints.

Person-to-person variability in HIV-1 virus kinetics. Administration of a single weight-based injection of IFN to 19 untreated HIV-1 infected persons resulted in a cohort-wide median (range) 0.85 (0.24 to 2.14) $\log_{10}$(decline in HIV RNA copies/mL) at one week (P=0.003) (**Figure 2A**). Participants were co-infected with HCV, and HCV viral kinetics were performed in parallel over the same intervals (13). Strikingly, HIV-1 and HCV viral kinetics did not correspond within the same people (**Figure 2B**), suggesting that the ISGs that restrict each virus are distinct.

RNA sequencing from activated CD4+ T cells before and after IFN. Peripheral blood mononuclear cells (PBMCs) were collected before and

24 hours after IFN. Activated CD4+ and CD8+ T cells, defined by co-expression of CD38 and HLA-DR, were separately quantified and sorted before and after IFN (**Figure 3-2C**). Compared to pre-IFN, the proportion of CD4+/CD38+/HLA-DR+ and CD8+/CD38+/HLA-DR+ T cells increased by a median (IQR) 5.4 (-1.6 to 6.9; p=0.02) and 19.1 (14.2 to 22.5; P=0.0001)**,** respectively, after IFN (**Figure 3-2D** and **Figure 3-3**) suggesting that IFN impacts global T-cell activation. We focused on activated CD4+ T cells, rather than total or resting CD4+ T-cells (108–110), pre- and post-IFN to avoid confounding due to their varying proportions. Thus activated CD4 T-cells were sorted using fluorescence-activated cell-sorting (FACS) from all subjects immediately before and 24 hours after IFN administration. A median (range) of 8.21e4 (1.66e4 to 1.25e5) cells were sorted from which a median (range) of 36.61 (4.86-69.59) nanograms of RNA was recovered with a median (range) RNA integrity number (RIN) of 9.4 (8.2-9.9). Complimentary DNAs (cDNA) were enzymatically fragmented to approximately 250bp and barcoded adapters were ligated followed by 19 cycles of linear amplification. Paired-end RNA sequencing (RNAseq) was performed (**Figure 3-2E**) and a median (range) of 87e6 (41e6 to 180e6) 100bp sequences were recovered from each sample. Duplicate sequences (presumably due to PCR) were removed prior to genomic alignment and accounted for a median 8.84% of the total reads.

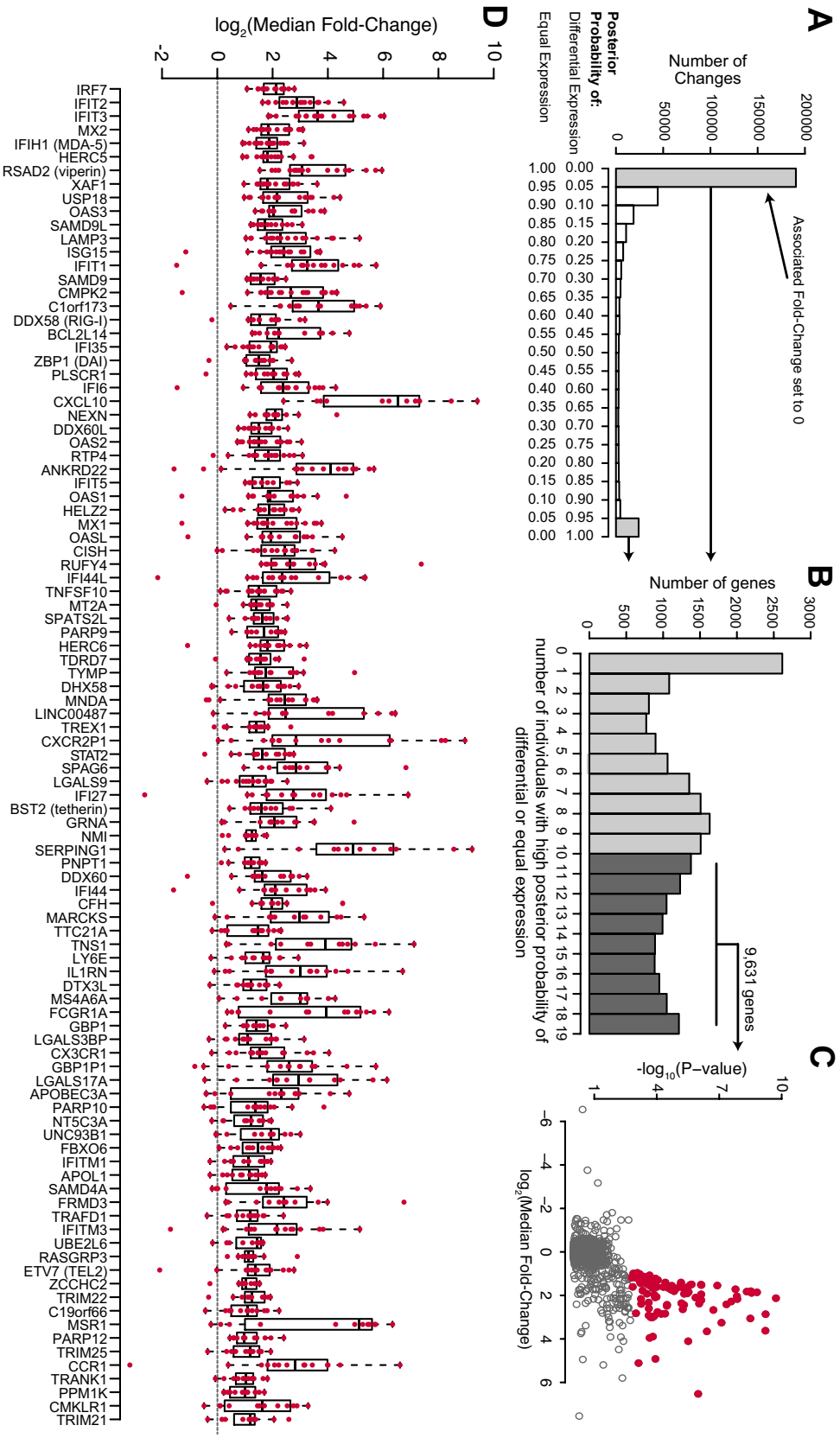
A
Number of Changes
0
50000
100000
150000
200000
Posterior Probability of:
Differential Expression
Equal Expression
1.00 0.00
0.95 0.05
0.90 0.10
0.85 0.15
0.80 0.20
0.75 0.25
0.70 0.30
0.65 0.35
0.60 0.40
0.55 0.45
0.50 0.50
0.45 0.55
0.40 0.60
0.35 0.65
0.30 0.70
0.25 0.75
0.20 0.80
0.15 0.85
0.10 0.90
0.05 0.95
0.00 1.00
Associated Fold-Change set to 0
B
Number of genes
0
500
1000
1500
2000
2500
3000
number of individuals with high posterior probability of differential or equal expression
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
9,631 genes
C
-log10(P−value)
1
4
7
10
log2(Median Fold-Change)
−6 −4 −2 0 2 4 6
D
log2(Median Fold-Change)
-2 0 2 4 6 8 10
IRF7
IFIT2
IFIT3
MX2
IFIH1 (MDA-5)
HERC5
RSAD2 (viperin)
XAF1
USP18
OAS3
SAMD9L
LAMP3
ISG15
IFIT1
SAMD9
CMPK2
C1orf173
DDX58 (RIG-I)
BCL2L14
IFI35
ZBP1 (DAI)
PLSCR1
IFI6
CXCL10
NEXN
DDX60L
OAS2
RTP4
ANKRD22
IFIT5
OAS1
HELZ2
MX1
OASL
CISH
RUFY4
IFI44L
TNFSF10
MT2A
SPATS2L
PARP9
HERC6
TDRD7
TYMP
DHX58
MNDA
LINC00487
TREX1
CXCR2P1
STAT2
SPAG6
LGALS9
IFI27
BST2 (tetherin)
GRNA
NMI
SERPING1
PNPT1
DDX60
IFI44
CFH
MARCKS
TTC21A
TNS1
LY6E
IL1RN
DTX3L
MS4A6A
FCGR1A
GBP1
LGALS3BP
CX3CR1
GBP1P1
LGALS17A
APOBEC3A
PARP10
NT5C3A
UNC93B1
FBXO6
IFITM1
APOL1
SAMD4A
FRMD3
TRAFD1
IFITM3
UBE2L6
RASGRP3
ETV7 (TEL2)
ZCCHC2
TRIM22
C19orf66
MSR1
PARP12
TRIM25
CCR1
TRANK1
PPM1K
CMKLR1
TRIM21

Figure 3-4. An integrated algorithm for *de novo* characterization of ISGs in activated CD4 T-cells from HIV-1 humans receiving IFN. (**A**) Histogram showing the distribution of PPDE and PPEE for all genes in which a determination could be made over the two sampled time-points. Meta-data from a merged data set for all individuals is shown. (B) Histogram showing the distribution of the number of genes with the indicated number of individuals detecting that gene with PPDE or PPEE≥0.95. Volcano plot comparing cohort-wide median post-correction fold-change values (with values corresponding to PPEE ≥0.95 set to 0) and p-values for the 9,631 genes with high-confidence quantitation in at least 50% of the cohort. (**D**) Boxplots with overlaid points showing the individual fold-changes for the 99 genes that significantly changed with IFN across the cohort after adjustment for multiple comparisons with no fold-change adjustment imposed for PPEE values ≥0.95.

Defining an empirical set of IFN-regulated genes in activated CD4+ T cells from cohort RNAseq data. Since we did not know which genes would be induced by IFN in activated CD4+ T cells, we first determined a set of ISGs in activated CD4+ T cell that were consistently induced in the cohort. An empirical Bayes hierarchical clustering model ('EBseq') was implemented to determine the posterior probability of differential vs. equal expression (PPDE and PPEE respectively) and post-correction fold-change for each gene after IFN treatment in each individual (**Figure 3-4A**). In total, 21,930 genes were equally or differentially expressed with a posterior probability ≥0.95 in at least one individual; 9,631 were equally or differentially expressed with a posterior probability ≥0.95 in ≥11 members of the cohort, which was the pre-determined threshold for further study (**Figure 3-4B**). The resulting $\log_2$ fold-changes for all individuals (with all fold-changes associated with PPEE ≥0.95 set to 0) were tested and significance was adjusted for each gene tested (**Figure 3-4C**). This yielded a *de novo* set of 99 genes that were *bona fide* ISGs in the activated CD4+ T-cells from the 19 people (**Figure 3-4D**). The list of ISGs

included multiple known HIV-1 restriction factors such as MX2 and BST2 (tetherin), several ISGs with previously described activity against other viruses (RSAD2 [Viperin], DDX58 [RIG-I], and IFITM3), and genes that have not been previously described as having antiviral roles (including but not limited to BCL2L14 [BCL-G], CMPK2, and LAMP3).

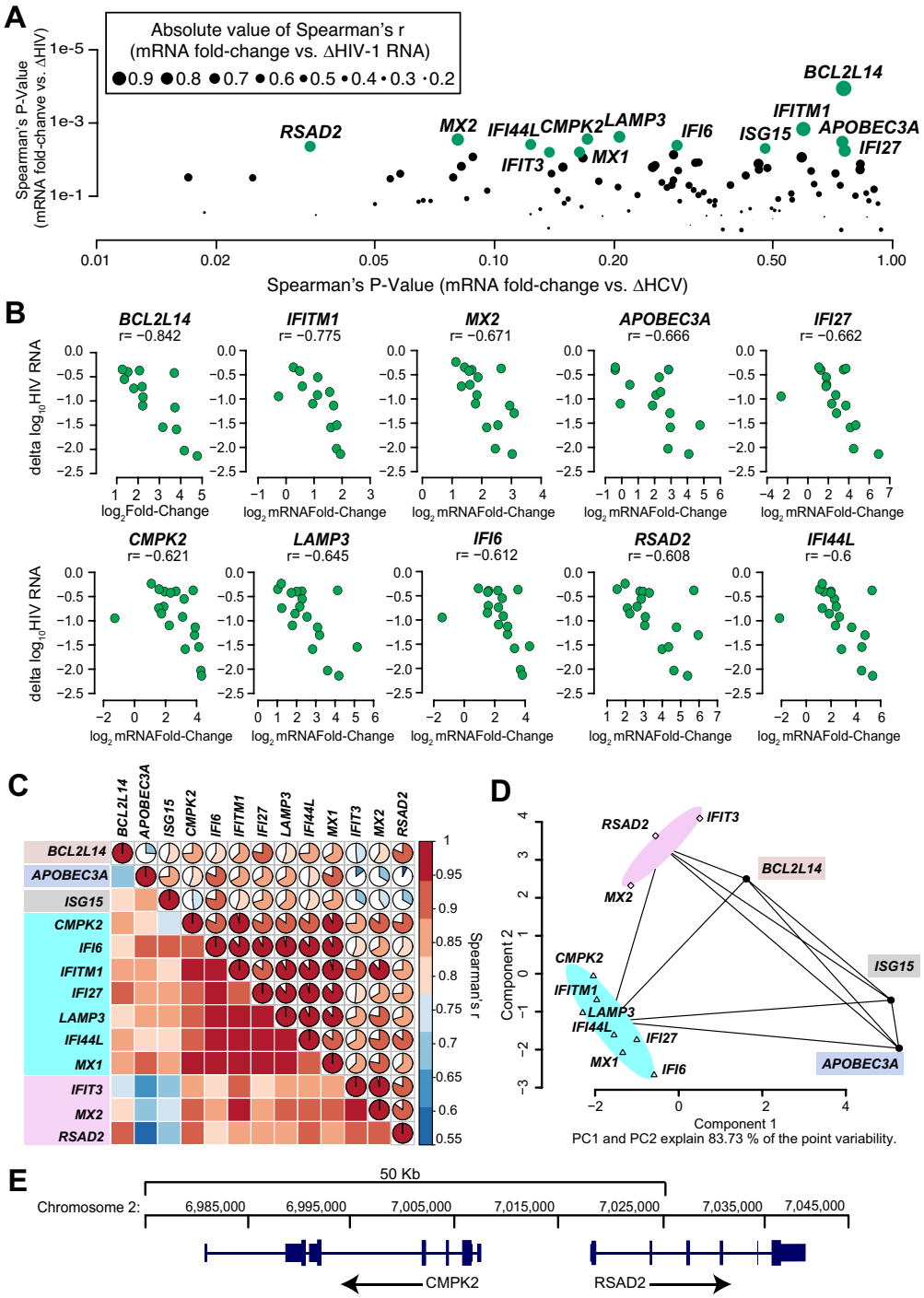

Figure 3-5. Relating ISG induction with viral RNA decline. (A) Scatter plot comparing P-values of Spearman correlations of each ISG's fold-change with plasma HIV-1 RNA decline at one week (y-axis) and plasma HCV RNA decline at one-week nadir (x-axis). (B) Individual scatterplots of 10 genes that significantly related to HIV and not HCV viral load decline after adjustment for multiple comparisons (scatterplots of MX1, ISG15, and IFIT3 are shown in Supplemental Figure). (C) Pairwise correlation plots of fold-changes for the genes shown in (B). (D) K-means clustering of principal component analysis grouping all points into 5 clusters. The same color scheme used to demark clusters in (D) is used in (C).

An algorithm for identifying ISGs that are putative HIV-1 restriction factors. As expected, persons-to-person variability in ISG expression was substantial for some genes **(Figure 3-4D**). We developed an algorithm that ranked ISGs as putative HIV-1 restriction factors by comparing the magnitude of increase of each ISG (retaining EBseq post-correction fold-changes when PPEE≥0.95) with the magnitude of plasma HIV-1 RNA decline across the cohort. The algorithm revealed 13 genes whose upregulation in activated CD4+ T cells were closely correlated with plasma HIV-1 RNA decline. To validate against false detection, we performed the same algorithm using HCV kinetics in the same people, and as expected found no significantly correlated ISGs. As expected, no gene-induction events in activated CD4 T-cells related to HCV RNA decline significantly after adjustment for multiple comparisons, although RSAD2, which related to HIV-1 RNA decline did relate to HCV RNA decline with no adjustment (P=0.04) (**Figure 3-5A**). The 13 putative HIV-1 restriction ISGs included MX2 and IFI6, two genes that are known to restrict HIV-1 *in vitro* (8, 91). In addition, several genes were identified for which no evaluation of antiviral activity has been performed, including BCL2L14, CMPK2, LAMP3 (**Figure 3-5B, Figure 3-6**).

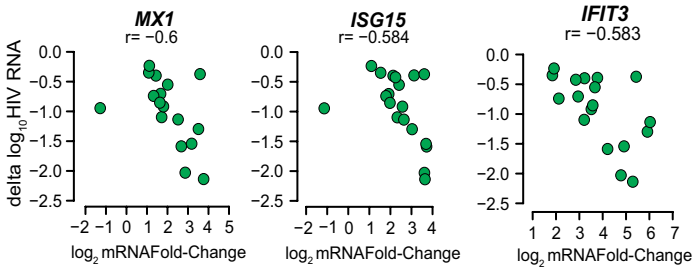

Figure 3-6. Scatterplots of additional genes that significantly related to HIV-1 restriction *in vivo.* Scatterplots labeled with the corresponding Spearman rank correlation estimate comparing the post-correction fold-changes of the indicated gene with plasma HIV-1 RNA decline at one-week.

We next asked whether induction of these 13 putative restriction factors for HIV-1 were induced by IFN in clusters to test whether their relative effect on HIV-1 could parsed independently of each other. Fortuitously, pairwise-wise correlation analysis of post-correction fold-change values revealed a hierarchy in the relatedness of gene induction (**Figure 3-5C**). We characterized this hierarchy by performing k-means clustering on principal component analysis of the fold-change values and revealed five distinct clusters of genes with BCL2L14, APOBEC3A, and ISG15 each in their own cluster and the rest of the genes contained in two additional clusters (**Figure 3-5D**). Interestingly, although RSAD2 and CMPK2 are directly adjacent to each other on chromosome 2 (**Figure 3-5E**), they segregated into two distinct clusters with other members distributed across different chromosomes, suggesting differential transcriptional regulation at the CMPK2 locus.

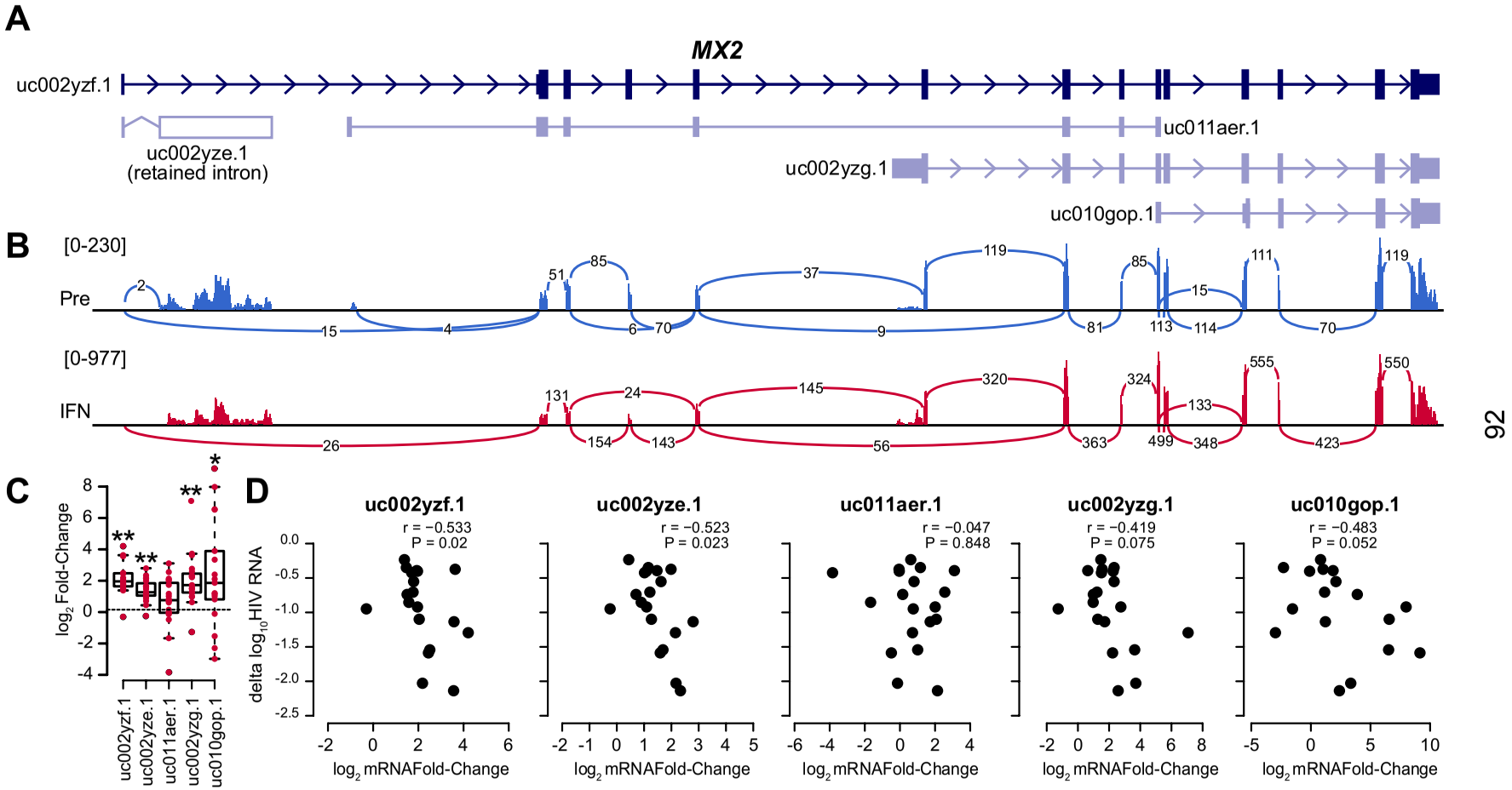
A
MX2
uc002yzf.1
uc002yze.1
(retained intron)
uc011aer.1
uc002yzg.1
uc010gop.1
B
[0-230]
Pre
[0-977]
IFN
C
log2 Fold-Change
uc002yzf.1
uc002yze.1
uc011aer.1
uc002yzg.1
uc010gop.1
D
delta log10 HIV RNA
uc002yzf.1
r = −0.533
P = 0.02
uc002yze.1
r = −0.523
P = 0.023
uc011aer.1
r = −0.047
P = 0.848
uc002yzg.1
r = −0.419
P = 0.075
uc010gop.1
r = −0.483
P = 0.052
log2 mRNAFold-Change

Figure 3-7. Multiple MX2 isoforms are induced by IFN and relate to viral restriction. (**A**) Relative genomic positions of MX2 isoform exons. The sole Refseq annotated transcript is shown in dark blue and the non-Refseq UCSC variants are shown in light blue. Arrow-studded lines indicate 5' to 3' directionality of transcripts and connectivity between exons (filled bars) and increased girth of bars indicates coding regions. Unfilled bars indicate retained introns. (**B**) Sashimi plots showing the mapped read density [range] within regions of MX2 immediately before (Pre, blue), and 24 hours after the administration of IFN (IFN, red). Curved lines are labeled with the number of reads spanning the indicated exon-exon junction. (**C**) Modified boxplots with points overlaid of post-correction fold-changes in each isoform with IFN treatment for each individual in which the measurement was possible. * $P<0.01$, ** $P\leq0.0001$. (**D**) Scatterplots with Spearman rank-correlation values showing the relationship between the post-correction fold-change values for each gene and virus RNA decline.

Specific ISG isoforms are associated with HIV-1 restriction.

Differential splicing of transcripts is evident in their sequence and thus RNAseq data can be used to examine and quantitate unique splice variants within a sample. We subsetted genes by their transcriptional variants for MX2, CMPK2, and BCL-G to test the hypothesis that individual isoforms were responsible for HIV-1 restriction (LAMP3 has only one annotated variant). We examined 5 distinct transcript variants of MX2 (**Figure 3-7A**) for their induction by IFN and their relationship with viral restriction (**Figure 3-7B-C**). Four out of the five (3 coding and 1 non-coding) variants were significantly induced by IFN ($P<0.05$ for all). Only one of the three coding variants (uc002yzf.1), also known as the full-length MX2, was associated with viral restriction ($P=0.02$) (**Figure 3-7C**). In addition, the non-coding variant of MX2, which contains the first exon and a partially-spliced retained intron (uc002yzf.1), was highly expressed and associated with both to viral restriction ($P=0.023$) (**Figure 3-S3C**) and full-length MX2 expression ($P=3.4\times10^{-7}$), although their splicing is mutually exclusive (**Figure 3-7A**).

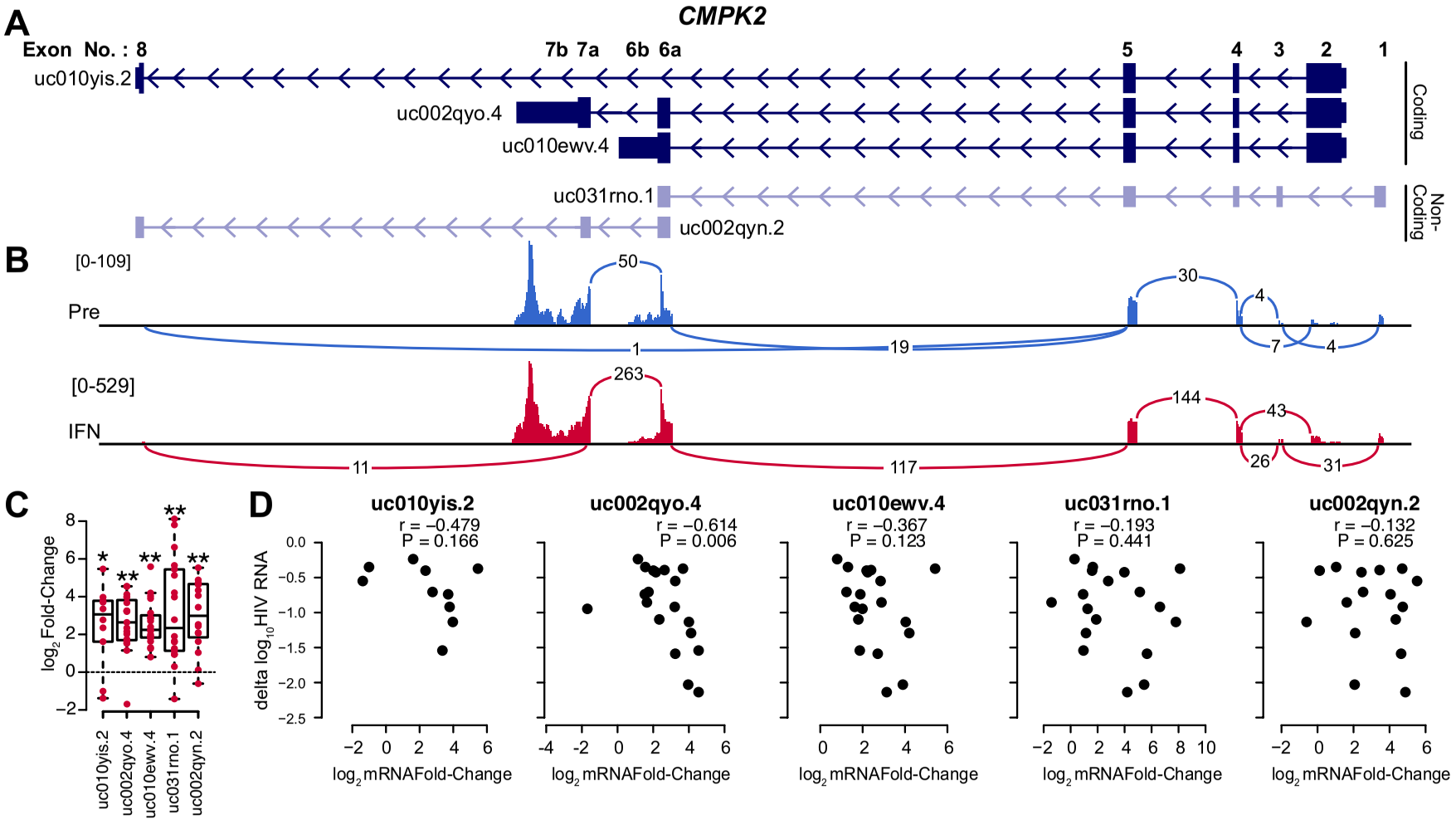
A
CMPK2
Exon No. : 8
7b 7a
6b 6a
5
4
3
2
1
uc010yis.2
uc002qyo.4
uc010ewv.4
uc031rno.1
uc002qyn.2
Coding
Non-Coding
B
[0-109]
Pre
50
30
4
1
19
7
4
[0-529]
IFN
263
144
43
11
117
26
31
C
log2 Fold-Change
uc010yis.2
uc002qyo.4
uc010ewv.4
uc031rno.1
uc002qyn.2
D
uc010yis.2
r = –0.479
P = 0.166
uc002qyo.4
r = –0.614
P = 0.006
uc010ewv.4
r = –0.367
P = 0.123
uc031rno.1
r = –0.193
P = 0.441
uc002qyn.2
r = –0.132
P = 0.625
delta log10 HIV RNA
log2 mRNAFold-Change

Figure 3-8. CMPK2 isoform analysis (**A**) Relative genomic positions of all detected UCSC annotated CMPK2 isoform exons. Detected variants that are Refseq annotated are shown in dark blue and non-Refseq UCSC annotated variants are shown in light blue. Arrow-studded lines indicate 5' to 3' directionality of transcripts and connectivity between exons (filled bars). Increased girth of bars indicates coding regions. (**B**) Sashimi plots showing the mapped read density [range] within regions of CMPK2 immediately before (Pre, blue), and 24 hours after the administration of IFN (IFN, red). Curved lines are labeled with the number of reads spanning the indicated exon-exon junction. (**C**) Modified boxplots with points overlaid of post-correction fold-changes in each isoform with IFN treatment for each individual in which the measurement was possible. * $P<0.01$, ** $P\leq0.0001$. (**D**) Scatterplots with Spearman rank-correlation values showing the relationship between the post-correction fold-change values for each gene and virus RNA decline.

We performed the same analysis for CMPK2, parsing out upregulation of each of five detected isoforms (3 coding and 2 noncoding) (**Figure 3-8A**) after IFN ($P<0.05$ for all)(**Figure 3-8B**). All isoforms were induced after IFN. However, similar to MX2, only one coding variant (uc002qyo.4) was associated with a decline in plasma HIV-1 RNA level after IFN ($P=0.006$)(**Figure 3-8C**).

Table 3-2. Quantitative PCR primer sequences and amplicon location by exon.

Gene	Exons Spanned*	Forward Primer	Reverse Primer
MX2	E1 to E2	TGATTTCTCCATCCTGAACGTG	CATGTGCTGTCTCCCTGTC
LAMP3	E2 to E3	ATGTTGAAGTATCTCCGAGGTG	GTCAAGACTGGAATTTATCAGGTTC
CMPK2 uc002qyo.4	E7b**	AATTTGGGACTGAGGGAGATG	CTACACTGGCATGCTGATGA
CMPK2 uc010ewv.4	E6a to E6b**	CCCACAGCAGCTACCATTAT	CAGTTCTACACCACTGCCTTTA
CMPK2 uc010yis.2	E3 to E4	AATCTGCCAAATCTCCTGTGA	CCA AGATTC CAGTCC CAC AA
CMPK2 uc031rno.1	E1 to E2	CGCTCTTAACTGCTCCACAA	GAGAGGGCAATTCCAAGAAGA
CMPK2 uc002qyn.2	E2 to E3	AGGTGGAGAGAAGGATGGTATAA	AAATGTCCTACCAGCGGATG
BCL2L14 uc001rac.3	E1 to E2	AATATCCTTACTGCCACCTGAC	GGGATTTCTTCCAGGTCACA
BCL2L14 uc001rae.3	E4 to E5	GCCAAAATTGTTGAGCTGCT	GCCATCAAACCATCCTGTG

*If no isoform specified, then Refseq exon annotation was used.
**Refer to Figure 3-8 for exon numbering.

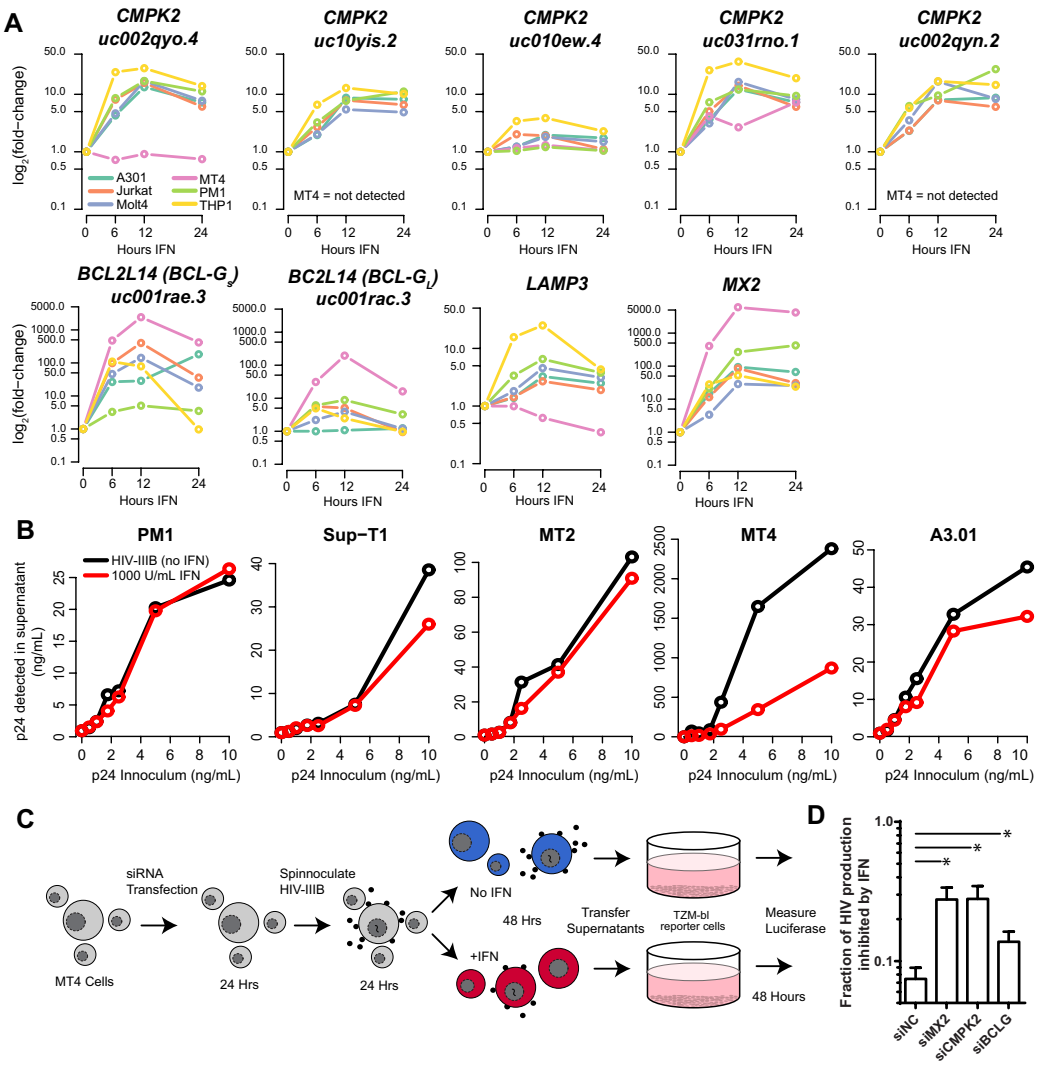

Figure 3-9. *In vitro* validation of CMPK2 and BCL2L14 as IFN-induced restriction factors for HIV-1. (**A**) Real-time qPCR data showing individual isoforms for CMPK2 and BCL2L14 as well LAMP3 and MX2 regulation in response to IFN in the 6 indicated cell lines at 0, 6, 12, and 24 hours. (**B**) IFN-mediated inhibition of HIV-1 in the indicated 5 cell lines known to support HIV replication. (**C**) Experimental scheme for measuring IFN-mediated inhibition of HIV-1 in cells in which CMPK2 and BCL-G have been down-modulated by RNAi. (**D**) Barplots showing inhibition measurements with error bars indicating standard error over the mean from 3 replicate transfections. * indicates P<0.05.

IFN inducibility and capacity for HIV-1 restriction of CMPK2, BCL2L14, and LAMP3. We tested whether IFN induced the putative HIV-1 restriction factors BCL2L14, LAMP3, and CMPK2 by treating multiple cell lines with IFN over a time-course of 24-hours and quantifying

intracellular mRNA levels of each gene using RT/qPCR with qPCR primer sets spanning unique exons or exon/exon junctions to distinguish the individual isoforms (**Figure 3-9A**). All three putative restriction factors were increased upon IFN treatment in most of the cell lines that were tested, although to different degrees. Notably, some isoforms of each gene were more robustly induced than others: e.g., BCL-G_L isoform uc001rac.3 exhibited <10-fold IFN-induction in 5/6 cell lines, whereas BCL-G_S isoform uc001rae.3 exhibited >10-fold induction in 5/6 cell lines. CMPK2 isoform uc002qyo.4 was induced in 5/6 cell lines, although not in MT4 cells (**Figure 3-9A**).

We next established which cell line(s) supported robust HIV-1 replication, and in which did IFN demonstrate the most antiviral activity. Six cell lines were inoculated with infectious HIV-1 at titrated concentrations to establish replication capacity of the virus in the cells (**Figure 3-9B**). Separate well were treated with IFN to inhibit viral replication. MT4 cells appeared to support the most robust HIV-1 replication and were also the most potent inhibitors of HIV-1 replication in response to IFN.

RNAi was used to determine if CMPK2 and BCL2L14 expression contributed to IFN-mediated restriction of HIV-1 *in vitro*. We transfected MT4 cells with siRNAs directed against CMPK2, BCL2L14, and MX2 as a positive control and then treated cells with IFN (**Figure 3-9C**). IFN resulted in a 92.6% inhibition of virus production in cells transfected with negative

control siRNAs (**Figure 3-9C**). In cells where siRNAs against MX2, our positive control were introduced, IFN resulted in a 72.3% inhibition of virus production, a significantly less amount than the negative control siRNA (P=0.015). When siRNAs against CMPK2 were introduced into these cells, IFN resulted in a 72.1% inhibition of virus production, a significantly less amount than the negative control siRNA (P=0.019) (**Figure 3-9C**). IFN treatment of cells in which siRNAs against BCL2L14 had been introduced resulted in an 86.2% inhibition of virus production, a significantly less amount than the negative control (P=0.049) (**Figure 3-9C**).

Discussion

By sequencing mRNAs in the principal cells in which HIV-1 replicates, we have identified IFN-induced HIV-1 restriction factors that directly contribute to the effect of IFN on HIV-1 among infected people with uncontrolled viremia. Our findings underscore the remarkable diversity of the IFN transcriptional program that is in place to defend against scores of intracellular pathogens: of the 99 ISGs that we identified in activated CD4+ T cells, only a fraction were relevant for HIV-1. The complexity of IFN signaling and ISG regulation is also evident, and was seen from the level of the individual to the level of each unique gene isoform. Understanding how our novel panel of ISGs restricts HIV-1 replication may illuminate vulnerabilities in the life cycle of an otherwise efficient pathogen that can be exploited as therapeutic targets.

There have been several genome-wide explorations of HIV-1 restriction factors, and these have yielded a plethora of insights about HIV-1 replication (111, 112). However previous reports have either hinged on the use of cell lines for discovery, or have utilized cell types that are not the natural target of HIV-1 infection. In addition, by using an open-ended sequencing approach, we have developed a rigorously-defined panel of ISGs in activated CD4+ T cells, expanding the list of ISGs that have been previously described as relevant in HIV-1 (113).

CMPK2 has been annotated as an ISG before, but not as an antiviral effector. Its genomic location on Chromosome 2 is nestled between RSAD2 and the newly described NRIR, an interferon-induced long non-coding RNA (linc-RNA) that is a negative regulator of IFN responses. Hence, CMPK2 sits in a transcriptionally active hotspot for IFN regulation. Our data collectively support that CMPK2, and not RSAD2 or NRIR, restricts HIV-1 replication in activated CD4 T-cells. CMPK2 catalyzes the phosphorylation of dUMP to dUDP, an intermediate step in the overall conversion of dUMP to dUTP. It is possible that increasing intracellular stores of dUTP may contribute to the lethal uracylation of retroviral genomes that has been described elsewhere (114). All CMPK2 coding isoforms contain the active site of the translated protein, and all were induced by IFN, so it is not clear why only the uc002.qyo.4 isoform was associated with HIV-1 restriction *in vivo.* The CMPK2-uc002.qyo.4 isoform differs from the other isoforms only in an additional 41 amino acids at the N-terminus, and in that it has a unique 3'UTR. These differences suggest that the CMPK2- uc002.qyo.4 isoform may possess modified activity conferred by the additional domain and/or that the stability of the transcript is improved by this alternate 3' UTR.

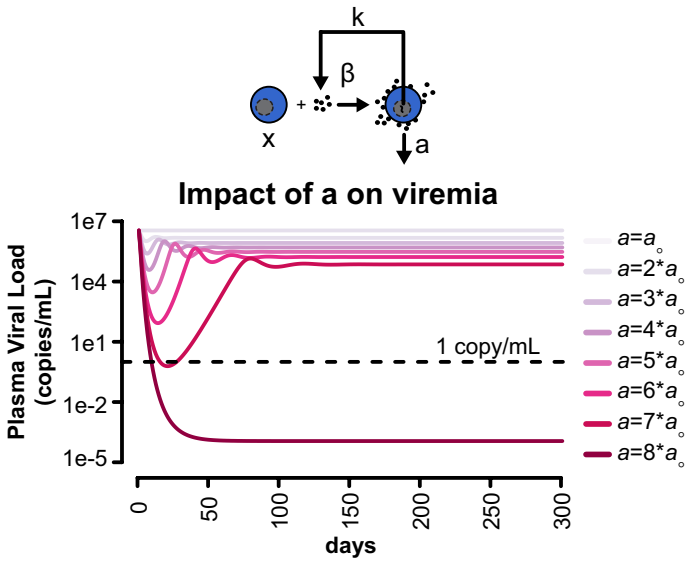

Figure 3-10. The impact of infected cell decay rate on HIV-1 virmeia. As in **Figure 1,** infections were simulated using three differential equations *dx/dt* = λ-dx-βxv, *dy/dt* = βxv-ay, and *dv/dt* = ky-uv (*λ* =production rate of uninfected cells, x=number of uninfected cells, d=decay rate of uninfected cells, β=first order rate constant of infection, v=plasma viral load, y=number of infected cells, a=decay rate of infected cells, and k=first order rate constant of virus production in infected cells) using the starting values (x=1e6,y=1,v=100000, λ=1e5,d=0.1,a=0.5,B=2e-7,k=100,u=5). Resulting values were then used as input into the same equations varying a by the indicated variable by the indicated factor and run for 300 days.

The function of BCLG remains poorly understood. There is controversy as to whether BCL-G is involved in apoptosis (115, 116). The BCLG nomenclature derives from its BH2 and BH3 domains, although the extent of its similarity to canonical BH-containing proteins such as BCL2 is unknown. Although IFN is known to reduce CD4+ T cell counts, this is believed to occur through bone-marrow production suppression and not increased apoptosis of terminally differentiated cells. Without a massive shift to apoptosis in IFN exposed cells, it is tempting to theorize that if BCL2L14 increases the tendency for apoptosis, it does so selectively (i.e. in infected cells). This would represent a novel form of viral restriction as

modeling suggests that an increased decay rate of infected cells could enduringly suppressing viremia to below 1 copy/liter when the decay rate increases by ≥8-fold (**Figure 3-10**). While knockdown of BCL2L14 isoforms significantly limited IFN-mediated HIV-1 restriction, it was to a lower extent than CMPK2 or MX2. This may have been due to the observation that MT4 cells, the cell type in which the RNAi experiment was performed, was the most potent inducer of BCL2L14 and significant residual induction through IFN treatment still occurred. More potent suppression or oblation of the locus is needed to confirm this hypothesis.

We encountered several challenges in this investigation. Biases in reverse transcription and the library preparation may have favored the most abundant transcripts that were then targeted for sequencing, omitting less abundant but still relevant transcripts. This may have resulted in patchy quantification of some transcripts across the cohort. To account for this possibility, we pre-specified that we would only analyze transcripts that were detectable in ≥11/19 individuals to ensure that only consistently detected genes were considered. A second challenge was in selecting CD4+ T cell markers that corresponded well with HIV-1 infected or HIV-1 susceptible cells. We used CD38 and HLA-DR, which are well established for identifying activated CD4+ T cells. However, It is possible that other markers would have more specifically identified HIV-1 susceptible cells. Focusing on only HIV-1 infected cells, however, would have omitted the effects of IFN on preventing infection of new cells, which

translates mathematically to which ISGs impact β, the first order rate constant of infection (**Figure 3-1**). An important point is that we did not have sufficient power to do a complete isoform analysis on all of the 21,930 unique genes that were sequenced. However, we were able to analyze the isoforms of individual genes of interest after employing the computational algorithm to identify putative ISG HIV-1 restriction factors. Further research is needed to understand how the functional activity of these isoforms differ, especially with regard to CMPK2 and MX2. Furthermore, while our *in vitro* RNAi experiments against CMPK2 demonstrated that it plays a role in IFN-mediated HIV-1 restriction, this was evident in the MT4 cell line, which did not demonstrate a substantial increase in CMPK2 expression upon IFN stimulation. These results suggest that CMPK2 catalysis may be rate-limiting for IFN-mediated restriction of HIV-1.

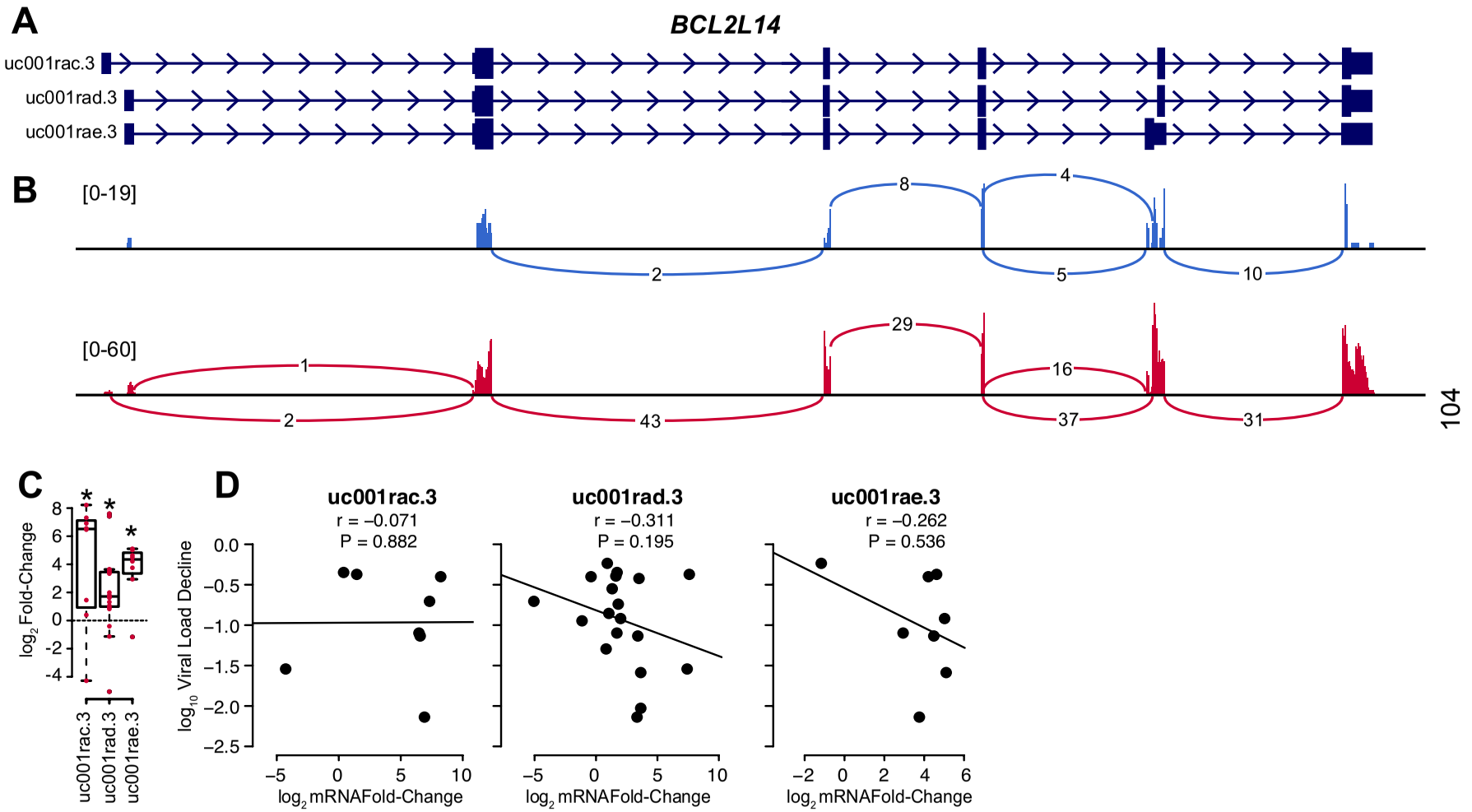
A
BCL2L14
uc001rac.3
uc001rad.3
uc001rae.3
B
[0-19]
8
4
2
5
10
[0-60]
29
1
16
2
43
37
31
104
C
log2 Fold-Change
8
6
4
2
0
-2
-4
*
*
*
uc001rac.3
uc001rad.3
uc001rae.3
D
uc001rac.3
r = –0.071
P = 0.882
uc001rad.3
r = –0.311
P = 0.195
uc001rae.3
r = –0.262
P = 0.536
log10 Viral Load Decline
0.0
–0.5
–1.0
–1.5
–2.0
–2.5
–5
0
5
10
–2
2
4
6
log2 mRNAFold-Change

Figure 3-11. Multiple isoforms for BCL2L14 are induced by IFN. (**A**) Relative genomic positions of BCL2L14 isoform exons. Refseq annotated transcripts are shown in dark blue and the non-Refseq UCSC variants are shown in light blue. Arrow-studded lines indicate 5' to 3' directionality of transcripts and connectivity between exons (filled bars) and increased girth of bars indicates coding regions. Unfilled bars indicate retained introns. (**B**) Sashimi plots showing the mapped read density [range] within regions of BCL2L14 immediately before (Pre, blue), and 24 hours after the administration of IFN (IFN, red). Curved lines are labeled with the number of reads spanning the indicated exon-exon junction. (**C**) Modified boxplots with points overlaid of post-correction fold-changes in each isoform with IFN treatment for each individual in which the measurement was possible. * $P<0.01$, ** $P\leq0.0001$. (**D**) Scatterplots with Spearman rank-correlation values showing the relationship between the post-correction fold-change values for each gene and virus RNA decline. Missing points are from individuals in which the isoform could not be identified.

The likelihood that specific exons and junctions within a gene will have sufficient coverage for quantitation depends on many features including (but not limited to) exon size, location (terminal sequence is less likely to be captured), expression relative to all other genes, and number of sequences acquired for each sample. There were three detected UCSC isoforms of BCL2L14 defined by two differential splice sites in our samples. One of these sites is between the 4^{th} and 5^{th} exon, however, it only distinguishes uc001rad.3 from uc001rac.3 or uc001rae.3. Detection of the splice site that distinguishes uc001rac.3 from uc001rae.3 (short non-coding exon 1) was sufficiently low to preclude quantitation for 11/19 individuals (**Figure 3-11**). With these data we were able to determine that all isoforms were induced by IFN treatment, however, we were unable to establish a significant rank-correlation relationship with any particular isoform. Due to the priceless nature of RNAs acquired from these specimens, high likelihood of low expression of these isoforms in pre-treated samples, and importance of quantitating many reference genes for

accurate normalization, we elected not to employ qPCR to determine for changes of individual isoforms.

We provide evidence for two novel interferon-induced HIV-1 restriction factors. HIV-1 restriction factors provide insights into the ongoing battle between pathogen and host, and may provide novel therapeutic targets for HIV-1 control. Unlike many other identified host-restriction factors for HIV-1, the genes we identified are highly variable in their induction from host-to-host providing an inconsistent landscape of restriction for HIV-1 to realize fitness advantage. Whether or not this is a protective mechanism for host species whose generation time and mutation rates are miniscule in comparison to viruses remains to be determined.

Conclusions

This thesis presents three novel methods for sampling natural HCV and HIV infections to describe novel features of the molecular biology of infection.

By comparing plasma miRNA levels in a well-characterized cohort before, during, and after acute HCV infection in chapter one, we identified a consistent and specific extracellular miRNA signature of acute infection that validated in two separate methods; in another cohort and in a tissue culture model of infection. Contrary to predictions made by conventional wisdom, circulating miRNA abundance during acute infection was not simply the result of cellular necrosis but indeed a nonrandom process of retention and release that tracked with the onset and control of virus infection. This suggests that extracellular miRNA release is sequence-specific and is a result of the onset and outcome of HCV infection. Given the contribution of miRNAs in HCV replication, their prediction of clinical outcomes during chronic infection, and their potential to regulate networks of gene expression, these findings also suggest that the pathophysiology of chronic HCV infection may be established acutely as part of an innate response to viral infection.

In chapter two, we developed a high-throughput platform for prediction of HCV bnAb resistance polymorphisms and identified an unexpected shared resistance polymorphism for two highly-potent bnAbs for HCV not contained within the described epitopes for either. By

sampling 113 naturally occurring HCV envelopes for their sensitivity to two of the most potent broadly neutralizing antibodies characterized to date, we found significant variation in isolate sensitivity and identified a shared resistance polymorphism located outside the predicted contact residues of both antibodies. Previous research from our lab has identified that neutralizing antibodies fall into distinct clusters with respect to which isolates they neutralize most potently. We presented evidence the unifying features of clustering are far more complex than simple epitope recognition.

In chapter three, we sequenced mRNAs in the principal cells in which HIV-1 replicates and subsequently identified IFN-induced HIV-1 restriction factors that contribute to the effect of IFN on HIV-1 among infected people with uncontrolled viremia. These IFN-induced restriction factors for HIV-1 provide insights into the ongoing battle between pathogen and host to replicate and limit the repercussions of infection. Since host factors are relatively static in their sequence, their proteins and pathways may provide novel therapeutic targets for controlling HIV-1 infection. Unlike many other identified host-restriction factors for HIV-1, the genes we identified are highly variable in their induction from host-to-host providing an inconsistent landscape of restriction for HIV-1 to realize fitness advantage heightening the possibility that viral proteins that target their destruction may not exist.

The findings in chapters one and three not only highlight the importance of characterizing relationships in humans, but that significant and profound person-to-person differences are evident even in early and short sampling intervals. The findings in chapter two suggest viral evasion from neutralization is far more complex than sequence changes in binding epitopes and perhaps a far more complex process involving significant structural and receptor tropism changes through mutation in other parts of the envelope. These findings too underscore the importance of studying natural infections in humans, and particularly the person-to-person differences that may be a previously unappreciated host species-specific defense mechanism against pathogens.

In conclusion, careful and limited sampling of humans with natural viral infections can provide insights for which no model can provide surrogacy. In particular, genetic variation between hosts within a species and viral infections that they harbor are particularly useful in dissecting the molecular interactions that define the battles between host and pathogen.

Appendix

References

1. **Henderson S**, **Fenton T**. 2015. APOBEC3 genes: retroviral restriction factors to cancer drivers. Trends Mol Med **21**:274–284.

2. **Key FM**, **Peter B**, **Dennis MY**, **Huerta-Sánchez E**, **Tang W**, **Prokunina-Olsson L**, **Nielsen R**, **Andrés AM**. PLOS Genetics: Selection on a Variant Associated with Improved Viral Clearance Drives Local, Adaptive Pseudogenization of Interferon Lambda 4 (IFNL4).

3. **Thomas DL**, **Astemborski J**, **Rai RM**, **et al**. 2000. The natural history of hepatitis c virus infection: Host, viral, and environmental factors. JAMA **284**:450–456.

4. **Cox AL**, **Netski DM**, **Mosbruger T**, **Sherman SG**, **Strathdee S**, **Ompad D**, **Vlahov D**, **Chien D**, **Shyamala V**, **Ray SC**, **Thomas DL**. 2005. Prospective evaluation of community-acquired acute-phase hepatitis C virus infection. Clin Infect Dis Off Publ Infect Dis Soc Am **40**:951–958.

5. **Maheshwari A**, **Ray S**, **Thuluvath PJ**. 2008. Acute hepatitis C. Lancet **372**:321–332.

6. **Amako Y**, **Tsukiyama-Kohara K**, **Katsume A**, **Hirata Y**, **Sekiguchi S**, **Tobita Y**, **Hayashi Y**, **Hishima T**, **Funata N**, **Yonekawa H**, **Kohara M**. 2010. Pathogenesis of hepatitis C virus infection in Tupaia belangeri. J Virol **84**:303–311.

7. **Fettig J**, **Swaminathan M**, **Murrill CS**, **Kaplan JE**. 2014. Global epidemiology of HIV. Infect Dis Clin North Am **28**:323–337.

8. **McLaren PJ**, **Gawanbacht A**, **Pyndiah N**, **Krapp C**, **Hotter D**, **Kluge SF**, **Götz N**, **Heilmann J**, **Mack K**, **Sauter D**, **Thompson D**, **Perreaud J**, **Rausell A**, **Munoz M**, **Ciuffi A**, **Kirchhoff F**, **Telenti A**. 2015. Identification of potential HIV restriction factors by combining evolutionary genomic signatures with functional analyses. Retrovirology **12**:41.

9. **Kong YW**, **Ferland-McCollough D**, **Jackson TJ**, **Bushell M**. 2012. microRNAs in cancer management. Lancet Oncol **13**:e249–e258.

10. **Köberle V**, **Kronenberger B**, **Pleli T**, **Trojan J**, **Imelmann E**, **Peveling-Oberhag J**, **Welker M-W**, **Elhendawy M**, **Zeuzem S**, **Piiper A**, **Waidmann O**. 2013. Serum microRNA-1 and microRNA-122 are prognostic markers in patients with hepatocellular carcinoma. Eur J Cancer **49**:3442–3449.

11. **Bala S**, **Petrasek J**, **Mundkur S**, **Catalano D**, **Levin I**, **Ward J**, **Alao H**, **Kodys K**, **Szabo G**. 2012. Circulating microRNAs in exosomes indicate hepatocyte injury and inflammation in alcoholic, drug-induced, and inflammatory liver diseases. HEPATOLOGY **56**:1946–1957.

12. **Menéndez P**, **Padilla D**, **Villarejo P**, **Palomino T**, **Nieto P**, **Menéndez JM**, **Rodríguez-Montes JA**. 2013. Prognostic implications of serum microRNA-21 in colorectal cancer. J Surg Oncol **108**:369–373.

13. **Cheng Y**, **Tan N**, **Yang J**, **Liu X**, **Cao X**, **He P**, **Dong X**, **Qin S**, **Zhang C**. 2010. A translational study of circulating cell-free microRNA-1 in acute myocardial infarction. Clin Sci Lond Engl 1979 **119**:87–95.

14. **Witwer KW**. 2015. Circulating microRNA biomarker studies: pitfalls and potential solutions. Clin Chem **61**:56–63.

15. **Leidner RS**, **Li L**, **Thompson CL**. 2013. Dampening enthusiasm for circulating microRNA in breast cancer. PloS One **8**:e57841.

16. **Bellare P**, **Ganem D**. 2009. Regulation of KSHV lytic switch protein expression by a virus-encoded microRNA: an evolutionary adaptation that fine-tunes lytic reactivation. Cell Host Microbe **6**:570–575.

17. **Shrivastava S**, **Petrone J**, **Steele R**, **Lauer GM**, **Di Bisceglie AM**, **Ray RB**. 2013. Up-regulation of circulating miR-20a is correlated with hepatitis C virus-mediated liver disease progression. HEPATOLOGY **58**:863–871.

18. **Jopling CL**, **Yi M**, **Lancaster AM**, **Lemon SM**, **Sarnow P**. 2005. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. Science **309**:1577–1581.

19. **Janssen HLA**, **Reesink HW**, **Lawitz EJ**, **Zeuzem S**, **Rodriguez-Torres M**, **Patel K**, **van der Meer AJ**, **Patick AK**, **Chen A**, **Zhou Y**, **Persson R**, **King BD**, **Kauppinen S**, **Levin AA**, **Hodges MR**. 2013. Treatment of HCV infection by targeting microRNA. N Engl J Med **368**:1685–1694.

20. **Li J**, **Liu K**, **Liu Y**, **Xu Y**, **Zhang F**, **Yang H**, **Liu J**, **Pan T**, **Chen J**, **Wu M**, **Zhou X**, **Yuan Z**. 2013. Exosomes mediate the cell-to-cell transmission of IFN-α-induced antiviral activity. Nat Immunol **14**:793–803.

21. **Roccaro AM**, **Sacco A**, **Maiso P**, **Azab AK**, **Tai Y-T**, **Reagan M**, **Azab F**, **Flores LM**, **Campigotto F**, **Weller E**, **Anderson KC**,

Scadden DT, **Ghobrial IM**. 2013. BM mesenchymal stromal cell-derived exosomes facilitate multiple myeloma progression. J Clin Invest **123**:1542–1555.

22. **Tadokoro H**, **Umezu T**, **Ohyashiki K**, **Hirano T**, **Ohyashiki JH**. 2013. Exosomes derived from hypoxic leukemia cells enhance tube formation in endothelial cells. J Biol Chem.

23. **Morello M**, **Minciacchi VR**, **de Candia P**, **Yang J**, **Posadas E**, **Kim H**, **Griffiths D**, **Bhowmick N**, **Chung LW**, **Gandellini P**, **Freeman MR**, **Demichelis F**, **Di Vizio D**. 2013. Large oncosomes mediate intercellular transfer of functional microRNA. Cell Cycle Georget Tex **12**:3526–3536.

24. **Delorme-Axford E**, **Donker RB**, **Mouillet J-F**, **Chu T**, **Bayer A**, **Ouyang Y**, **Wang T**, **Stolz DB**, **Sarkar SN**, **Morelli AE**, **Sadovsky Y**, **Coyne CB**. 2013. Human placental trophoblasts confer viral resistance to recipient cells. Proc Natl Acad Sci U S A **110**:12048–12053.

25. **van Balkom BWM**, **de Jong OG**, **Smits M**, **Brummelman J**, **Ouden K den**, **de Bree PM**, **van Eijndhoven MAJ**, **Pegtel DM**, **Stoorvogel W**, **Würdinger T**, **Verhaar MC**. 2013. Endothelial cells require miR-214 to secrete exosomes that suppress senescence and induce angiogenesis in human and mouse endothelial cells. Blood **121**:3997–4006, S1–15.

26. **Xin H**, **Li Y**, **Buller B**, **Katakowski M**, **Zhang Y**, **Wang X**, **Shang X**, **Zhang ZG**, **Chopp M**. 2012. Exosome-mediated transfer of miR-133b from multipotent mesenchymal stromal cells to neural cells contributes to neurite outgrowth. Stem Cells Dayt Ohio **30**:1556–1564.

27. **Montecalvo A**, **Larregina AT**, **Shufesky WJ**, **Stolz DB**, **Sullivan MLG**, **Karlsson JM**, **Baty CJ**, **Gibson GA**, **Erdos G**, **Wang Z**, **Milosevic J**, **Tkacheva OA**, **Divito SJ**, **Jordan R**, **Lyons-Weiler J**, **Watkins SC**, **Morelli AE**. 2012. Mechanism of transfer of functional microRNAs between mouse dendritic cells via exosomes. Blood **119**:756–766.

28. **Valadi H**, **Ekström K**, **Bossios A**, **Sjöstrand M**, **Lee JJ**, **Lötvall JO**. 2007. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol **9**:654–659.

29. **Mittelbrunn M**, **Gutiérrez-Vázquez C**, **Villarroya-Beltri C**, **González S**, **Sánchez-Cabo F**, **González MÁ**, **Bernad A**, **Sánchez-Madrid F**. 2011. Unidirectional transfer of microRNA-loaded exosomes from T cells to antigen-presenting cells. Nat Commun **2**:282.

30. **Okoye IS**, **Coomes SM**, **Pelly VS**, **Czieso S**, **Papayannopoulos V**, **Tolmachova T**, **Seabra MC**, **Wilson MS**. 2014. MicroRNA-Containing T-Regulatory-Cell-Derived Exosomes Suppress Pathogenic T Helper 1 Cells. Immunity **41**:89–103.

31. **Page K**, **Hahn JA**, **Evans J**, **Shiboski S**, **Lum P**, **Delwart E**, **Tobler L**, **Andrews W**, **Avanesyan L**, **Cooper S**, **Busch MP**. 2009. Acute hepatitis C virus infection in young adult injection drug users: a prospective study of incident infection, resolution, and reinfection. J Infect Dis **200**:1216–1226.

32. **Ray SC**, **Arthur RR**, **Carella A**, **Bukh J**, **Thomas DL**. 2000. Genetic epidemiology of hepatitis C virus throughout egypt. J Infect Dis **182**:698–707.

33. **McAlexander MA**, **Phillips MJ**, **Witwer KW**. 2013. Comparison of Methods for miRNA Extraction from Plasma and Quantitative Recovery of RNA from Cerebrospinal Fluid. Front Genet **4**:83.

34. **Bolstad BM**, **Irizarry RA**, **Astrand M**, **Speed TP**. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinforma Oxf Engl **19**:185–193.

35. **Witwer KW**, **Sarbanes SL**, **Liu J**, **Clements JE**. 2011. A plasma microRNA signature of acute lentiviral infection: biomarkers of central nervous system disease. AIDS Lond Engl **25**:2057–2067.

36. **Ritchie ME**, **Phipson B**, **Wu D**, **Hu Y**, **Law CW**, **Shi W**, **Smyth GK**. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res **43**:e47.

37. **Lindenbach BD**, **Evans MJ**, **Syder AJ**, **Wölk B**, **Tellinghuisen TL**, **Liu CC**, **Maruyama T**, **Hynes RO**, **Burton DR**, **McKeating JA**, **Rice CM**. 2005. Complete replication of hepatitis C virus in cell culture. Science **309**:623–626.

38. **Hochberg Y**, **Benjamini Y**. 1990. More powerful procedures for multiple significance testing. Stat Med **9**:811–818.

39. **Hajarizadeh B**, **Grady B**, **Page K**, **Kim AY**, **McGovern BH**, **Cox AL**, **Rice TM**, **Sacks-Davis R**, **Bruneau J**, **Morris M**, **Amin J**, **Schinkel J**, **Applegate T**, **Maher L**, **Hellard M**, **Lloyd AR**, **Prins M**, **Dore GJ**, **Grebely J**, **InC3 Study Group**. 2015. Patterns of hepatitis C virus RNA levels during acute infection: the InC3 study. PloS One **10**:e0122232.

40. **Hajarizadeh B**, **Grady B**, **Page K**, **Kim AY**, **McGovern BH**, **Cox AL**, **Rice TM**, **Sacks-Davis R**, **Bruneau J**, **Morris M**, **Amin J**, **Schinkel J**, **Applegate T**, **Maher L**, **Hellard M**, **Lloyd AR**, **Prins M**, **Geskus RB**, **Dore GJ**, **Grebely J**, **the InC3 Study Group**. 2015. Factors associated with hepatitis C virus RNA levels in early chronic infection: the InC(3) study. J Viral Hepat.

41. **Hajarizadeh B**, **Grady B**, **Page K**, **Kim AY**, **McGovern BH**, **Cox AL**, **Rice TM**, **Sacks-Davis R**, **Bruneau J**, **Morris M**, **Amin J**, **Schinkel J**, **Applegate T**, **Maher L**, **Hellard M**, **Lloyd AR**, **Prins M**, **Geskus RB**, **Dore GJ**, **Grebely J**, **InC(3)Study Group**. 2014. Interferon lambda 3 genotype predicts hepatitis C virus RNA levels in early acute infection among people who inject drugs: the InC(3) study. J Clin Virol Off Publ Pan Am Soc Clin Virol **61**:430–434.

42. **Sacks-Davis R**, **Grebely J**, **Dore GJ**, **Osburn W**, **Cox AL**, **Rice TM**, **Spelman T**, **Bruneau J**, **Prins M**, **Kim AY**, **McGovern BH**, **Shoukry NH**, **Schinkel J**, **Allen TM**, **Morris M**, **Hajarizadeh B**, **Maher L**, **Lloyd AR**, **Page K**, **Hellard M**, **InC3 study group**. 2015. Hepatitis C Virus Reinfection and Spontaneous Clearance of Reinfection-the InC3 Study. J Infect Dis.

43. **Tzur G**, **Israel A**, **Levy A**, **Benjamin H**, **Meiri E**, **Shufaro Y**, **Meir K**, **Khvalevsky E**, **Spector Y**, **Rojansky N**, **Bentwich Z**, **Reubinoff BE**, **Galun E**. 2009. Comprehensive gene and microRNA expression profiling reveals a role for microRNAs in human liver development. PloS One **4**:e7511.

44. **Randall G**, **Panis M**, **Cooper JD**, **Tellinghuisen TL**, **Sukhodolets KE**, **Pfeffer S**, **Landthaler M**, **Landgraf P**, **Kan S**, **Lindenbach BD**, **Chien M**, **Weir DB**, **Russo JJ**, **Ju J**, **Brownstein MJ**, **Sheridan R**, **Sander C**, **Zavolan M**, **Tuschl T**, **Rice CM**. 2007. Cellular cofactors

affecting hepatitis C virus infection and replication. Proc Natl Acad Sci U S A **104**:12884–12889.

45. **Starkey Lewis PJ**, **Dear J**, **Platt V**, **Simpson KJ**, **Craig DGN**, **Antoine DJ**, **French NS**, **Dhaun N**, **Webb DJ**, **Costello EM**, **Neoptolemos JP**, **Moggs J**, **Goldring CE**, **Park BK**. 2011. Circulating microRNAs as potential markers of human drug-induced liver injury. HEPATOLOGY **54**:1767–1776.

46. **Ward J**, **Kanchagar C**, **Veksler-Lublinsky I**, **Lee RC**, **McGill MR**, **Jaeschke H**, **Curry SC**, **Ambros VR**. 2014. Circulating microRNA profiles in human patients with acetaminophen hepatotoxicity or ischemic hepatitis. Proc Natl Acad Sci U S A **111**:12169–12174.

47. **Liu L**, **Fisher BE**, **Thomas DL**, **Cox AL**, **Ray SC**. 2012. Spontaneous clearance of primary acute hepatitis C virus infection correlated with high initial viral RNA level and rapid HVR1 evolution. HEPATOLOGY **55**:1684–1691.

48. **Witwer KW**, **Sisk JM**, **Gama L**, **Clements JE**. 2010. MicroRNA regulation of IFN-beta protein expression: rapid and sensitive modulation of the innate immune response. J Immunol Baltim Md 1950 **184**:2369–2376.

49. **Cheng J-C**, **Yeh Y-J**, **Tseng C-P**, **Hsu S-D**, **Chang Y-L**, **Sakamoto N**, **Huang H-D**. 2012. Let-7b is a novel regulator of hepatitis C virus replication. Cell Mol Life Sci CMLS **69**:2621–2633.

50. **Tryndyak V**, **de Conti A**, **Kobets T**, **Kutanzi K**, **Koturbash I**, **Han T**, **Fuscoe JC**, **Latendresse JR**, **Melnyk S**, **Shymonyak S**, **Collins L**, **Ross SA**, **Rusyn I**, **Beland FA**, **Pogribny IP**. 2012. Interstrain differences in the severity of liver injury induced by a choline- and

folate-deficient diet in mice are associated with dysregulation of genes involved in lipid metabolism. FASEB J **26**:4592–4602.

51. **Bukong TN**, **Momen-Heravi F**, **Kodys K**, **Bala S**, **Szabo G**. 2014. Exosomes from hepatitis C infected patients transmit HCV infection and contain replication competent viral RNA in complex with Ago2-miR122-HSP90. PLoS Pathog **10**:e1004424.

52. **Cermelli S**, **Ruggieri A**, **Marrero JA**, **Ioannou GN**, **Beretta L**. 2011. Circulating microRNAs in patients with chronic hepatitis C and non-alcoholic fatty liver disease. PloS One **6**:e23937.

53. **Trebicka J**, **Anadol E**, **Elfimova N**, **Strack I**, **Roggendorf M**, **Viazov S**, **Wedemeyer I**, **Drebber U**, **Rockstroh J**, **Sauerbruch T**, **Dienes H-P**, **Odenthal M**. 2013. Hepatic and serum levels of miR-122 after chronic HCV-induced fibrosis. J Hepatol **58**:234–239.

54. **Su T-H**, **Liu C-H**, **Liu C-J**, **Chen C-L**, **Ting T-T**, **Tseng T-C**, **Chen P-J**, **Kao J-H**, **Chen D-S**. 2013. Serum microRNA-122 level correlates with virologic responses to pegylated interferon therapy in chronic hepatitis C. Proc Natl Acad Sci U S A **110**:7844–7849.

55. **Estrabaud E**, **Lapalus M**, **Broët P**, **Appourchaux K**, **De Muynck S**, **Lada O**, **Martinot-Peignoux M**, **Bièche I**, **Valla D**, **Bedossa P**, **Marcellin P**, **Vidaud M**, **Asselah T**. 2014. Reduction of microRNA 122 expression in IFNL3 CT/TT carriers and during progression of fibrosis in patients with chronic hepatitis C. J Virol **88**:6394–6402.

56. **Gui J**, **Tian Y**, **Wen X**, **Zhang W**, **Zhang P**, **Gao J**, **Run W**, **Tian L**, **Jia X**, **Gao Y**. 2011. Serum microRNA characterization identifies miR-885-5p as a potential marker for detecting liver pathologies. Clin Sci Lond Engl 1979 **120**:183–193.

57. **Fu X**, **Tan D**, **Hou Z**, **Hu Z**, **Liu G**, **Ouyang Y**, **Liu F**. 2012. [Effect of microRNA on proliferation caused by mutant HBx in human hepatocytes]. Zhonghua Gan Zang Bing Za Zhi Zhonghua Ganzangbing Zazhi Chin J Hepatol **20**:598–604.

58. **Xu X**, **Wells A**, **Padilla MT**, **Kato K**, **Kim KC**, **Lin Y**. 2014. A signaling pathway consisting of miR-551b, catalase and MUC1 contributes to acquired apoptosis resistance and chemoresistance. Carcinogenesis **35**:2457–2466.

59. **Wang B**, **Hsu S**, **Wang X**, **Kutay H**, **Bid HK**, **Yu J**, **Ganju RK**, **Jacob ST**, **Yuneva M**, **Ghoshal K**. 2014. Reciprocal regulation of microRNA-122 and c-Myc in hepatocellular cancer: role of E2F1 and transcription factor dimerization partner 2. HEPATOLOGY **59**:555–566.

60. **Bai S**, **Nasser MW**, **Wang B**, **Hsu S-H**, **Datta J**, **Kutay H**, **Yadav A**, **Nuovo G**, **Kumar P**, **Ghoshal K**. 2009. MicroRNA-122 inhibits tumorigenic properties of hepatocellular carcinoma cells and sensitizes these cells to sorafenib. J Biol Chem **284**:32015–32027.

61. **Hsu S-H**, **Ghoshal K**. 2013. MicroRNAs in Liver Health and Disease. Curr Pathobiol Rep **1**:53–62.

62. **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**, **Hsu M-T**, **Wu J-C**, **Huang H-D**, **Shiao M-S**, **Hsiao M**, **Tsou A-P**. 2012. MicroRNA-122 plays a critical role in liver homeostasis and hepatocarcinogenesis. J Clin Invest **122**:2884–2897.

63. **Afanasyeva EA**, **Mestdagh P**, **Kumps C**, **Vandesompele J**, **Ehemann V**, **Theissen J**, **Fischer M**, **Zapatka M**, **Brors B**, **Savelyeva L**, **Sagulenko V**, **Speleman F**, **Schwab M**,

Westermann F. 2011. MicroRNA miR-885-5p targets CDK2 and MCM5, activates p53 and inhibits proliferation and survival. Cell Death Differ **18**:974–984.

64. **Yan W**, **Zhang W**, **Sun L**, **Liu Y**, **You G**, **Wang Y**, **Kang C**, **You Y**, **Jiang T**. 2011. Identification of MMP-9 specific microRNA expression profile as potential targets of anti-invasion therapy in glioblastoma multiforme. Brain Res **1411**:108–115.

65. **Guan X**, **Liu Z**, **Liu H**, **Yu H**, **Wang L-E**, **Sturgis EM**, **Li G**, **Wei Q**. 2013. A functional variant at the miR-885-5p binding site of CASP3 confers risk of both index and second primary malignancies in patients with head and neck cancer. FASEB J Off Publ Fed Am Soc Exp Biol **27**:1404–1412.

66. **Lim L**, **Balakrishnan A**, **Huskey N**, **Jones KD**, **Jodari M**, **Ng R**, **Song G**, **Riordan J**, **Anderton B**, **Cheung S-T**, **Willenbring H**, **Dupuy A**, **Chen X**, **Brown D**, **Chang AN**, **Goga A**. 2014. MicroRNA-494 within an oncogenic microRNA megacluster regulates G1/S transition in liver tumorigenesis through suppression of mutated in colorectal cancer. HEPATOLOGY **59**:202–215.

67. **Song L**, **Liu D**, **Wang B**, **He J**, **Zhang S**, **Dai Z**, **Ma X**, **Wang X**. 2015. miR-494 suppresses the progression of breast cancer in vitro by targeting CXCR4 through the Wnt/β-catenin signaling pathway. Oncol Rep **34**:525–531.

68. **Holmberg SD**, **Spradling PR**, **Moorman AC**, **Denniston MM**. 2013. Hepatitis C in the United States. N Engl J Med **368**:1859–1861.

69. **Zibbell JE**, **Iqbal K**, **Patel RC**, **Suryaprasad A**, **Sanders KJ**, **Moore-Moravian L**, **Serrecchia J**, **Blankenship S**, **Ward JW**, **Holtzman D**.

2015. Increases in hepatitis C virus infection related to injection drug use among persons aged ≤30 years - kentucky, tennessee, virginia, and west virginia, 2006-2012. MMWR Morb Mortal Wkly Rep **64**:453–458.

70. **Conrad C**, **Bradley HM**, **Broz D**, **Buddha S**, **Chapman EL**, **Galang RR**, **Hillman D**, **Hon J**, **Hoover KW**, **Patel MR**, **Perez A**, **Peters PJ**, **Pontones P**, **Roseberry JC**, **Sandoval M**, **Shields J**, **Walthall J**, **Waterhouse D**, **Weidle PJ**, **Wu H**, **Duwve JM**. 2015. Community Outbreak of HIV Infection Linked to Injection Drug Use of Oxymorphone - Indiana, 2015. MMWR Morb Mortal Wkly Rep **64**:443–444.

71. Five–Year Risk of Late Relapse or Reinfection With Hepatitis C After Sustained Virological Response: Meta-analysis of 49 Studies in 8534 Patients.

72. **Franco S**, **Tural C**, **Nevot M**, **Moltó J**, **Rockstroh JK**, **Clotet B**, **Martinez MA**. 2014. Detection of a sexually transmitted hepatitis C virus protease inhibitor-resistance variant in a human immunodeficiency virus-infected homosexual man. Gastroenterology **147**:599–601.e1.

73. **Armstrong GL**, **Wasley A**, **Simard EP**, **McQuillan GM**, **Kuhnert WL**, **Alter MJ**. 2006. The Prevalence of Hepatitis C Virus Infection in the United States, 1999 through 2002. Ann Intern Med **144**:705–714.

74. **Liang TJ**. 2013. Current progress in development of hepatitis C virus vaccines. Nat Med **19**:869–878.

75. **Mohd Hanafiah K**, **Groeger J**, **Flaxman AD**, **Wiersma ST**. 2013. Global epidemiology of hepatitis C virus infection: new estimates of

age-specific antibody to HCV seroprevalence. HEPATOLOGY **57**:1333–1342.

76. **Bailey JR**, **Wasilewski LN**, **Snider AE**, **El-Diwany R**, **Osburn WO**, **Keck Z**, **Foung SKH**, **Ray SC**. 2015. Naturally selected hepatitis C virus polymorphisms confer broad neutralizing antibody resistance. J Clin Invest **125**:437–447.

77. **Giang E**, **Dorner M**, **Prentoe JC**, **Dreux M**, **Evans MJ**, **Bukh J**, **Rice CM**, **Ploss A**, **Burton DR**, **Law M**. 2012. Human broadly neutralizing antibodies to the envelope glycoprotein complex of hepatitis C virus. Proc Natl Acad Sci U S A **109**:6205–6210.

78. **Wong JAJ-X**, **Bhat R**, **Hockman D**, **Logan M**, **Chen C**, **Levin A**, **Frey SE**, **Belshe RB**, **Tyrrell DL**, **Law JLM**, **Houghton M**. 2014. Recombinant hepatitis C virus envelope glycoprotein vaccine elicits antibodies targeting multiple epitopes on the envelope glycoproteins associated with broad cross-neutralization. J Virol **88**:14278–14288.

79. **Kenny-Walsh E**. 1999. Clinical outcomes after hepatitis C infection from contaminated anti-D immune globulin. Irish Hepatology Research Group. N Engl J Med **340**:1228–1233.

80. **Raghuraman S**, **Park H**, **Osburn WO**, **Winkelstein E**, **Edlin BR**, **Rehermann B**. 2012. Spontaneous clearance of chronic hepatitis C virus infection is associated with appearance of neutralizing antibodies and reversal of T-cell exhaustion. J Infect Dis **205**:763–771.

81. **Osburn WO**, **Snider AE**, **Wells BL**, **Latanich R**, **Bailey JR**, **Thomas DL**, **Cox AL**, **Ray SC**. 2014. Clearance of hepatitis C infection is associated with the early appearance of broad neutralizing antibody responses. HEPATOLOGY **59**:2140–2151.

82. **Edgar RC**. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32**:1792–1797.

83. **Jones DT**, **Taylor WR**, **Thornton JM**. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci CABIOS **8**:275–282.

84. **Tamura K**, **Stecher G**, **Peterson D**, **Filipski A**, **Kumar S**. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol **30**:2725–2729.

85. **Munshaw S**, **Bailey JR**, **Liu L**, **Osburn WO**, **Burke KP**, **Cox AL**, **Ray SC**. 2012. Computational reconstruction of Bole1a, a representative synthetic hepatitis C virus subtype 1a genome. J Virol **86**:5915–5921.

86. **Kong L**, **Giang E**, **Nieusma T**, **Kadam RU**, **Cogburn KE**, **Hua Y**, **Dai X**, **Stanfield RL**, **Burton DR**, **Ward AB**, **Wilson IA**, **Law M**. 2013. Hepatitis C virus E2 envelope glycoprotein core structure. Science **342**:1090–1094.

87. **Bartolini B**, **Lionetti R**, **Giombini E**, **Sias C**, **Taibi C**, **Montalbano M**, **Offizi G D'**, **McPhee F**, **Hughes EA**, **Zhou N**, **Ippolito G**, **Garbuglia AR**, **Capobianchi MR**. 2015. Dynamics of HCV genotype 4 resistance-associated variants during virologic escape with pIFN/RBV + daclatasvir: A case study using ultra deep pyrosequencing. J Clin Virol **66**:38–43.

88. **Sandler NG**, **Bosinger SE**, **Estes JD**, **Zhu RTR**, **Tharp GK**, **Boritz E**, **Levin D**, **Wijeyesinghe S**, **Makamdop KN**, **del Prete GQ**, **Hill BJ**, **Timmer JK**, **Reiss E**, **Yarden G**, **Darko S**, **Contijoch E**, **Todd JP**, **Silvestri G**, **Nason M**, **Norgren RB**, **Keele BF**, **Rao S**, **Langer JA**, **Lifson JD**, **Schreiber G**, **Douek DC**. 2014. Type I interferon

responses in rhesus macaques prevent SIV infection and slow disease progression. Nature **511**:601–605.

89. **Goujon C**, **Malim MH**. 2010. Characterization of the alpha interferon-induced postentry block to HIV-1 infection in primary human macrophages and T cells. J Virol **84**:9254–9266.

90. **Bosinger SE**, **Li Q**, **Gordon SN**, **Klatt NR**, **Duan L**, **Xu L**, **Francella N**, **Sidahmed A**, **Smith AJ**, **Cramer EM**, **Zeng M**, **Masopust D**, **Carlis JV**, **Ran L**, **Vanderford TH**, **Paiardini M**, **Isett RB**, **Baldwin DA**, **Else JG**, **Staprans SI**, **Silvestri G**, **Haase AT**, **Kelvin DJ**. 2009. Global genomic analysis reveals rapid control of a robust innate response in SIV-infected sooty mangabeys. J Clin Invest **119**:3556–3572.

91. **Goujon C**, **Moncorgé O**, **Bauby H**, **Doyle T**, **Ward CC**, **Schaller T**, **Hué S**, **Barclay WS**, **Schulz R**, **Malim MH**. 2013. Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. Nature **502**:559–562.

92. **Soros VB**, **Yonemoto W**, **Greene WC**. 2007. Newly synthesized APOBEC3G is incorporated into HIV virions, inhibited by HIV RNA, and subsequently activated by RNase H. PLoS Pathog **3**:e15.

93. **Sheehy AM**, **Gaddis NC**, **Choi JD**, **Malim MH**. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature **418**:646–650.

94. **Neil SJD**, **Zang T**, **Bieniasz PD**. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature **451**:425–430.

95. **Laguette N**, **Sobhian B**, **Casartelli N**, **Ringeard M**, **Chable-Bessia C**, **Ségéral E**, **Yatim A**, **Emiliani S**, **Schwartz O**, **Benkirane M**.

2011. SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. Nature **474**:654–657.

96. **Langevin C**, **Aleksejeva E**, **Passoni G**, **Palha N**, **Levraud J-P**, **Boudinot P**. 2013. The antiviral innate immune response in fish: evolution and conservation of the IFN system. J Mol Biol **425**:4904–4920.

97. **Asmuth DM**, **Murphy RL**, **Rosenkranz SL**, **Lertora JJL**, **Kottilil S**, **Cramer Y**, **Chan ES**, **Schooley RT**, **Rinaldo CR**, **Thielman N**, **Li X-D**, **Wahl SM**, **Shore J**, **Janik J**, **Lempicki RA**, **Simpson Y**, **Pollard RB**, **AIDS Clinical Trials Group A5192 Team**. 2010. Safety, tolerability, and mechanisms of antiretroviral activity of pegylated interferon Alfa-2a in HIV-1-monoinfected participants: a phase II clinical trial. J Infect Dis **201**:1686–1696.

98. **Azzoni L**, **Foulkes AS**, **Papasavvas E**, **Mexas AM**, **Lynn KM**, **Mounzer K**, **Tebas P**, **Jacobson JM**, **Frank I**, **Busch MP**, **Deeks SG**, **Carrington M**, **O'Doherty U**, **Kostman J**, **Montaner LJ**. 2013. Pegylated Interferon alfa-2a monotherapy results in suppression of HIV type 1 replication and decreased cell-associated HIV DNA integration. J Infect Dis **207**:213–222.

99. **Ireton RC**, **Gale M**. 2013. Systems biology analyses to define host responses to HCV infection and therapy. Curr Top Microbiol Immunol **363**:143–167.

100. **Wodarz D**, **Nowak MA**. 2002. Mathematical models of HIV pathogenesis and treatment. BioEssays News Rev Mol Cell Dev Biol **24**:1178–1187.

101. **Perelson AS**, **Neumann AU**, **Markowitz M**, **Leonard JM**, **Ho DD**. 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science **271**:1582–1586.

102. **Finzi D**, **Siliciano RF**. 1998. Viral Dynamics in HIV-1 Infection. Cell **93**:665–671.

103. **Balagopal A**, **Kandathil AJ**, **Higgins YH**, **Wood J**, **Richer J**, **Quinn J**, **Eldred L**, **Li Z**, **Ray SC**, **Sulkowski MS**, **Thomas DL**. 2014. Antiretroviral therapy, interferon sensitivity, and virologic setpoint in human immunodeficiency virus/hepatitis C virus coinfected patients. HEPATOLOGY **60**:477–486.

104. **Li B**, **Dewey CN**. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics **12**:323.

105. **Leng N**, **Dawson JA**, **Thomson JA**, **Ruotti V**, **Rissman AI**, **Smits BMG**, **Haag JD**, **Gould MN**, **Stewart RM**, **Kendziorski C**. 2013. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinforma Oxf Engl **29**:1035–1043.

106. **Perelson AS**, **Neumann AU**, **Markowitz M**, **Leonard JM**, **Ho DD**. 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science **271**:1582–1586.

107. **Jiao Y-M**, **Weng W-J**, **Gao Q-S**, **Zhu W-J**, **Cai W-P**, **Li L-H**, **Li H-J**, **Gao Y-Q**, **Wu H**. 2015. Hepatitis C therapy with interferon-α and ribavirin reduces the CD4 cell count and the total, 2LTR circular and integrated HIV-1 DNA in HIV/HCV co-infected patients. Antiviral Res **118**:118–122.

108. **Pan X**, **Baldauf H-M**, **Keppler OT**, **Fackler OT**. 2013. Restrictions to HIV-1 replication in resting CD4+ T lymphocytes. Cell Res **23**:876–885.

109. **Zhou Y**, **Zhang H**, **Siliciano JD**, **Siliciano RF**. 2005. Kinetics of human immunodeficiency virus type 1 decay following entry into resting CD4+ T cells. J Virol **79**:2199–2210.

110. **Zack JA**, **Arrigo SJ**, **Weitsman SR**, **Go AS**, **Haislip A**, **Chen IS**. 1990. HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. Cell **61**:213–222.

111. **Brass AL**, **Dykxhoorn DM**, **Benita Y**, **Yan N**, **Engelman A**, **Xavier RJ**, **Lieberman J**, **Elledge SJ**. 2008. Identification of host proteins required for HIV infection through a functional genomic screen. Science **319**:921–926.

112. **Zhou H**, **Xu M**, **Huang Q**, **Gates AT**, **Zhang XD**, **Castle JC**, **Stec E**, **Ferrer M**, **Strulovici B**, **Hazuda DJ**, **Espeseth AS**. 2008. Genome-scale RNAi screen for host factors required for HIV replication. Cell Host Microbe **4**:495–504.

113. **Schoggins JW**, **Wilson SJ**, **Panis M**, **Murphy MY**, **Jones CT**, **Bieniasz P**, **Rice CM**. 2011. A diverse range of gene products are effectors of the type I interferon antiviral response. Nature **472**:481–485.

114. **Weil AF**, **Ghosh D**, **Zhou Y**, **Seiple L**, **McMahon MA**, **Spivak AM**, **Siliciano RF**, **Stivers JT**. 2013. Uracil DNA glycosylase initiates degradation of HIV-1 cDNA containing misincorporated dUTP and prevents viral integration. Proc Natl Acad Sci U S A **110**:E448–457.

115. **Guo B**, **Godzik A**, **Reed JC**. 2001. Bcl-G, a novel pro-apoptotic member of the Bcl-2 family. J Biol Chem **276**:2780–2785.

116. **Giam M**, **Okamoto T**, **Mintern JD**, **Strasser A**, **Bouillet P**. 2012. Bcl-2 family member Bcl-G is not a proapoptotic protein. Cell Death Dis **3**:e404.

Abbreviations

Alanine aminotransferase, ALT

Cycle threshold, Ct

Hepatitis C virus, HCV

Hepatocellular Carcinoma, HCC

Human immunodeficiency virus, HIV

Infectious clone of HCV in cell-culture, HCVcc

Injection drug users, IDU

Institutional Review Board, IRB

Interferon-stimulated gene, ISG

IU, international units; Cycle Threshold

Interquartile range, IQR

Lower limit of quantification, LLOQ

MicroRNA, miRNA

Reverse transcription-quantitative polymerase chain reaction, RT-qPCR

RNA sequencing, RNAseq

U, units

Curriculum Vitae

RAMY MONIR EL-DIWANY

2221 Fleet St. Baltimore, MD 21231
Born April 3rd, 1986 in Sacramento, CA to Drs. Seham (mother) and Monir (father) El-Diwany

EDUCATION

MD-PhD Candidate, expected graduation 2017
Medical Scientist Training Program
Johns Hopkins University School of Medicine

Bachelor of Arts with High Honors, 2008
Major in Molecular and Cell Biology (emphasis in infectious diseases), Minor in Philosophy
University of California, Berkeley

RESEARCH EXPERIENCE

PhD student, Biochemistry, Cellular, and Molecular Biology Graduate Training Program (2011-Present)
Johns Hopkins School of Medicine Viral Hepatitis Research Center, Baltimore, MD

- Identification of novel interferon-induced restriction factors for HIV-1 using computational genomics to explore naturally occurring variation in innate immunity.
  - Mentors: Ashwin Balagopal, Robert Siliciano, and David Thomas
- Characterization of naturally occurring resistance polymorphisms to broadly neutralizing antibodies for HCV.
  - Mentors: Justin Bailey and Stuart Ray
- Identification of HIV RNA and inflammatory gene signatures in activated CD4 T-cells from patients on HAART.
  - Mentors: Ashwin Balagopal, Robert Siliciano, and David Thomas
- Characterization of circulating miRNAs signatures of acute HCV infection.
  - Mentors: Ashwin Balagopal, Kenneth Witwer, Andrea Cox, Stuart Ray, David Thomas
- Single-cell analysis of hepatocytes from HCV infected patients to evaluate patterns and determinants of hepatitis C virus replication within the liver.
  - Mentors: Ashwin Balagopal and David Thomas
- METHODS USED: RNA sequencing, RNAi, virus purification, tissue culture, immunohistochemistry, Laser-

Capture micro-dissection, quantitative PCR array, flow-cytometery, Fluorescence-activated cell sorting (FACS), HCV sequence analysis, statistical analysis and bioinformatics in R.

Intramural Research Training Award Fellow (7/08-7/09)
Laboratory of Jake Liang, National Institute of Diabetes and Digestive and Kidney Diseases, Liver Disease Branch, Bethesda, MD

- Evaluating the mechanism and efficacy of a novel treatment for chronic hepatitis C.
- Genome-wide evaluation of miRNAs that modulate HCV replication in tissue culture.
- METHODS USED: Tissue culture, ChiP, siRNA transfection, high-through-put assay optimization, quantitative PCR.

Undergraduate Honors Thesis Student (6/06-6/08)
Laboratory of Laurent Coscoy, University of California, Berkeley, Berkeley, CA

- Developed a novel protocol for herpesvirus mutagenesis.
- Examined molecular host-pathogen interactions of Kaposi's sarcoma-associated herpesvirus (KSHV) associated with immune evasion including the MIR proteins and Open Reading Frame 45.
- METHODS USED: Tissue culture, flow cytometery, western blot, Bacterial Artificial Chromosome recombineering, viral genomics, tissue culture, plaque assay, and various PCR and cloning techniques.

Student Researcher (6/03-6/05)
Laboratory of Yueh-hsiu Chien, Stanford University School of Medicine, Stanford, CA

- Elucidated structural components of the γδ T Cell Receptor (TCR) conferring ligand specificity.
- METHODS USED: Mouse Intestinal Intraepithelial Lymphocyte and Splenocyte Isolation, Flow Cytometery, Fluorescence-activated cell sorting (FACS), population TCR sequencing, and various other PCR techniques.

AWARDS AND OTHER FELLOWSHIPS

- **Young Investigator Award** (2015)
  - The International Antiviral Society/CROI Foundation
- **Excellence in Public Health Award** (2014)
  - The United States Public Health Service
- **Young Investigator Award** (2013)

        - The American Association for Study of Liver Diseases
    - **Albert Schweitzer Fellow for Life** (2011)
        - The Albert Schweitzer Fellowship
    - **Department of Molecular and Cell Biology Outstanding Undergraduate Award** (2008)
        - University of California, Berkeley

PUBLICATIONS

Acute Hepatitis C Virus Infection Induces Consistent Changes in Circulating miRNAs that are Associated with Non-Lytic Hepatocyte Release.
El-Diwany R, Wasilewski L, Witwer KW, Bailey JR, Page K, Ray SC, Cox AL, Thomas DL, Balagopal A. 2015. Journal of Virology.

IFNL3 genotype is associated with differential induction of IFNL3 in primary human hepatocytes.
Kurbanov F, Kim Y, Latanich R, Chaudhari P, **El-Diwany R**, Knabel M, Kandathil AJ, Cameron A, Cox A, Jang Y-Y, Thomas DL, Balagopal A. 2015. Antiviral Therapy.

Naturally selected hepatitis C virus polymorphisms confer broad neutralizing antibody resistance.
Bailey JR, Wasilewski LN, Snider AE, **El-Diwany R**, Osburn WO, Keck Z, Foung SKH, Ray SC. 2015. Journal of Clinical Investigation.

S-adenosyl methionine improves early viral responses and interferon-stimulated gene induction in hepatitis C nonresponders.
Feld JJ, Modi AA, **El-Diwany R**, Rotman Y, Thomas E, Ahlenstiel G, Titerence R, Koh C, Cherepanov V, Heller T, Ghany MG, Park Y, Hoofnagle JH, Liang TJ. 2011. Gastroenterology.

Antigen Recognition Determinants of Gamma Delta T Cell Receptors.
Shin S, **El-Diwany R**, Schaffert S, Adams EJ, Garcia KC, Pereira P, Chien Y. 2005. Science.

PRESENTATIONS

Acute Hepatitis C Infection is Associated with an Increase in Circulating microRNA-122
Ramy El-Diwany, Kimberly Page, Stuart C. Ray, Andrea L. Cox, David L. Thomas, Ashwin Balagopal
Oral abstract at the 2013 Annual National Liver Meeting of The American Association for the Study of Liver Diseases

Using Natural Variation in HCV Envelope to Identify Determinants of Broadly Neutralizing Antibody Sensitivity
Ramy El-Diwany, Lisa Wasilewski, Stuart Ray, Justin Bailey
Oral abstract at the 2014 International HIV Dynamics and Evolution Meeting

COMMUNITY SERVICE

Charm City Clinic Cofounder
Executive Board of Directors (2009-present), Executive Director (2011-2013)
Charm City Clinic, Inc. Baltimore, MD www.charmcityclinic.org

- Cofounded an accompaniment-model non-profit free-clinic with students, community leaders, and faculty members.
- Represent CCC in an official capacity to the Johns Hopkins medical student body, School of Medicine administration, hospital, and greater community.

As Executive Director

- Oversaw and developed fund-raising and grant writing efforts.
- Developed capacity to become the main referral center for uninsured patients for several local emergency rooms and clinics including Union Memorial and The Shepherd's Clinic.
- Oversaw strategy to secure funding for two full-time and three part-time employees including the following Executive Director.

500-Mile Runner and Team Captain
Runner (2001-2014), Team Captain (2008-2010)
500-Mile Native American Spiritual Marathon, Northern California

- Run annual 500-mile relay over a five-day period from sunrise to sunset with the message "All Life is Sacred." Train continuously at various competitive and noncompetitive events across the state of California.

Neonatal Intensive Care Unit (NICU) Volunteer and Mentor (1/06-6/08)
Oakland Children's Hospital, Oakland, CA

- Held, soothed, played with, and sung to NICU patients, many of who were in medically and socially fragile condition.
- Trained and mentored new NICU volunteers.

OTHER INTERESTS

Sailing, hiking, running, and skiing.