Title
Design, Synthesis and Delivery of a Self-Cleaving Replicon System Yielding Biologically Active MicroRNAs in vivo

Author
Brandt, Devin

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Design, Synthesis and Delivery of a Self-Cleaving Replicon System Yielding Biologically Active MicroRNAs *in vivo*

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Chemistry

by

Devin Brandt

2018

ABSTRACT OF THE DISSERTATION

Design, Synthesis and Delivery of a Self-Cleaving Replicon System Yielding Biologically Active MicroRNAs *in vivo*

by

Devin Brandt

Doctor of Philosophy in Chemistry

University of California, Los Angeles

Professor William Gelbart, Chair

Non-coding RNAs play a pivotal role in the regulation of a myriad of cellular processes. MicroRNAs (miRNAs) – short, single-stranded, RNAs about 20 nucleotides long – have been shown to be key regulators of various oncogenes, and dysregulation of their expression is strongly correlated with the cancerous state. In particular, the exogenous reintroduction of these miRNAs, has been shown by many groups to reduce the ability of cancer cells to spread in vitro. However, the delivery of short-hairpin RNA is inherently transient, as the RNA is exogenously produced and transfected into the cell; and shRNA-based gene silencing relies on the activity of the DICER complex, which often has reduced efficiency in cancerous cells. Here, we utilize the modified, self-replicating, genome of the single-stranded RNA virus Nodamura to deliver mature, self-amplifying, miRNAs cells in vitro. By inserting into this replicon a cassette consisting of a miRNA of interest (here miR-34A) flanked by two ribozymes, we have created a means to up-regulate the expression of particular miRNAs in a persistent manner. We have quantified the levels of both the replicon and miRNA, and demonstrate correlation of their expressions. Finally, we show that these replicon-derived microRNAs are biologically active.

This dissertation of Devin Brandt is approved.

Yung Ya Lin

Emil Reisler

Harold G. Monbouquette

William M. Gelbart, Committee Chair

University of California, Los Angeles

2018

Table of Contents

List of Figures

List of Tables:

Acknowledgements

I thank Bill Gelbart and Charles Knobler for immeasurable assistance throughout this endeavor. We thank UC Cancer Research Coordinating Committee (CRCC) and The Hirshberg Foundation for Pancreatic Cancer Treatment for the funding provided.

Vitea

Devin Brandt attended the University of Colorado, Boulder where he double-majored in Chemistry and Biochemistry. He graduated *magna cum laude* in Biochemistry, and his honors thesis involved the study of calcium-sensing, membrane-targeting proteins involved in insulin secretion. After graduating, Devin worked as a Laboratory Manager at the University of Colorado, Denver, in the Lab of Jefferson Knight, where he published two papers on the equilibrium and kinetic properties of synaptotagmin proteins.

Devin then joined the lab of Charles Knobler and William Gelbart. Here Devin studied the use of agrobacterium for genetic transformations, before extending his research to the use of self-replicating RNA to deliver functional RNAs to cancer cells. This research was awarded grants by both the Hirshberg Foundation as well as the UC SPORE program. While at UCLA, Devin developed a love of teaching. As a testament to this, he was awarded the Hanson-Dow teaching award in 2012. Devin is also dedicated to laboratory safety; he has been the safety officer since time immemorial.

Chapter 1: Background and Introduction

Obligate parasites known to infect all kingdoms of life, viruses straddle the line between living organism and cellular automaton. All viruses discovered so far are composed of a replicating genome surrounded and protected by a capsid composed of protein subunits and, in the case of enveloped viruses, a lipid coat. While they lack an intrinsic metabolism, due partially to their inherently limited genome size, viruses are capable of self-replication, a key trait of a living organism. They achieve this feat by commandeering the cellular machinery of the organism they infect. Owing in part to their short life-cycle and simple biology, viruses are the most successful form of replicating genomes on the planet[1]. This means that viruses constitute a large trophic pump for all ecosystems, and this fact is most obvious when one considers viral activity within the world's oceans. In the ocean, there are an estimated 10^23 viral infections initiated per second, ultimately infecting and killing 20-40% of the total marine bacteria each day[2]. Given that the number of offspring from a given single prokaryotic infection can vary between 100 and 1000, it is no surprise that an estimated 10^31 viruses are thought to exist in the oceans, outnumbering any other genome by orders of magnitude[1]. A brief overview of these numbers is striking and leads to the following conclusion: viral genomic replication is one of the most frequent biological events and a significant portion of the total metabolic output of the planet is dedicated to their continued existence and reproduction[3].

The gene-products of all viruses are separated into two classes: structural and non-structural proteins. Structural proteins include all the relevant proteins associated with the mature capsid, such as scaffold proteins found in the formation of the procapsid in T4, and also include all relevant, portal proteins, tail spike proteins etc. as found in lambda phage, and membrane proteins in the case of enveloped viruses. The number of structural proteins varies greatly from virus to virus. In Lambda virus, there are 14 distinct structural proteins that form the mature bacteriophage. Conversely, the number of

structural proteins can be a few as a single gene product[4,5]. In both Cowpea Chlorotic Mottle Virus (CCMV) and Brome Mosaic Virus(BMV), there exists a single structural protein, known as the capsid protein which forms the mature protein shell[6]. Non-structural proteins are all proteins that are not involved directly with either the formation of the capsid or the capsid itself. In double-stranded DNA (dsDNA) viruses these includes the motor proteins, as well as the polymerases, ligases, nucleases etc. Following the above trend, the number of non-structural proteins in single-stranded RNA (ssRNA) viruses is smaller, sometimes by an order of magnitude; Nodamura – a ssRNA virus – encodes only 3 non-structural proteins and one structural protein[7].

The high replication rate, limited ability to proofread, and short size of viral genomes – features especially characteristic of the RNA genome viruses – make individual viral "species" prone to genetic drift and variation, and as such rigorous phylogenetic analysis is functionally impossible[8]. This effect is worsened by the amount of horizontal gene transfer seen in viruses, particularly in phages [9]. However, viral genomes can be initially segregated based on the composition of their genome (DNA or RNA), and this schema can be quite useful for predicting the replication strategies and overall structure of a virus. The Baltimore classification system, first proposed by David Baltimore, further classifies viruses by the strandedness of the genome (either single-stranded (ss) or double-stranded (ds)), as well as if the genome is directly translated (positive or negative). Additional groups are included for viruses that employ reverse-transcription during their life-cycle, denoted with an "RT." The resulting combinations that make up the full suite of viral "families" is thus: positive-sense single-stranded RNA (+ssRNA), negative-sense single-stranded RNA (-ssRNA), double-stranded RNA (dsRNA), single-stranded DNA (ssDNA), double-stranded DNA (dsDNA), single-stranded RNA reverse transcription (ssRNA-RT) and double-stranded DNA reverse-transcription (dsDNA-RT). Of these classifications, dsDNA and +ssRNA viruses make up the bulk of the sequenced genomes[10]. Whether the current distribution is wholly representative of the true distribution is unknown; every year hundreds to thousands of new viruses are

detected via sequencing, and many of them follow more non-traditional replication strategies [11].

All viruses contain genomes that are protected in some way by a proteinaceous structure known as a capsid. These capsids vary greatly from virus to virus. The simplest viral capsid is comprised of a single gene-product that self-assembles into a mature virion. More complicated capsids, often found in dsDNA bacteriophages, can include many different types of protein subunits, including scaffolding proteins etc., that help ensure the final structure of the capsid is strong enough to withstand the pressure of the rigid dsDNA genome found within. Many simple viruses are often icosahedrally symmetric, and as such are further categorized based on the triangulation number of their capsids. This is represented as a "T-number" and can be used a valuable metric for viral size and structure[12]. For many viruses, numerous protein-protein and protein-nucleic acid interactions are present in the final virion, which greatly increases the stability of the mature virion [13].

Regardless of the exact shape, the mere presence of the nucleocapsid establishes a firm selective pressure; the genome must "fit" into the capsid. Neither the genome or its gene products can change in ways that prevent the formation of a mature, infectious, virion. Mutations that directly affect the aforementioned protein-nucleic acid interaction, such as the deletion of a positively charged amino acid that interacted with the negatively charged genome, will affect the formation of the mature virion [14,15]. But the effects of mutations can be subtler. Even synonymous, single-nucleotide, mutations can affect the structure and shape of the genomic cargo, by changing for example which regions of the ssRNA are base-paired, preventing the formation of a mature virion, even if the gene-products remain unchanged [16]. Sequence, shape and size all play important roles in the selection of viral genomes and this has profound implication when one considers the innate difference between single- and double-stranded nucleic acids.

Double-stranded DNA is perhaps the most well characterized biological polymer studied. Due to the base pairing and pi-stacking interactions found between complementary nucleotides, coupled with

the self-repelling nature of the negatively-charged phosphate backbone, dsDNA is rigid, with a persistence length of about 50nm, weakly dependent on sequence and ionic strength[17]. This poses interesting problems for viruses that encode their genes in dsDNA form. The average gene length in these viruses is around 10,000 bp, with the largest recorded dsDNA genome belonging to *Klebsiella* Phage vB_KleM-RaK2, which has a genome that is 346,000 bp long [18].The average capsid diameters for these viruses, however, is measured in tens of nm. This means that the density of DNA inside these DNA viruses in high – almost close-packed, and thus the DNA must be strongly crowded on itself and forced into the capsid at curvatures smaller than its natural persistence length. This requires a profound amount of work to be done on the genome, and the resulting pressures within the capsid, which arise from a combination of bending energies as well as osmotic pressure effects, can reach up to 50 atmospheres[19]. The capsids of these viruses therefore must be very rigid and resist deformation, and are often formed independently prior to the packaging of their genomic cargo. These "pro-capsids" often undergo structure rearrangements during genomic packaging[20]. Nano-indentation studies performed on these capsids indicate that they can withstand indentation forces of 100s of pN[21]. The other consequence of the need for high pressures is that these viruses must package their genomes with incredibly strong motor proteins. Double-stranded DNA viruses have the strongest motor proteins known, and are capable of exerting more than 50pN of force[22] when packaging DNA. The benefit of the high pressures established by these viruses is that their gene-delivery is assisted by the high internal pressures; the genome effectively "explodes" out from the virus[19] due to the large pressure differential.

Unlike dsDNA viruses, ssRNA viruses must contend with a genome that is flexible, globular and dynamic[23]. Where double-stranded DNA is rigid, single-stranded nucleic acids are relatively flexible; the persistence length of dsDNA is tens of nanometers, while the persistence length of ssRNA is much smaller, on the order of nanometers [24]. In addition, there are many secondary and tertiary structure configurations of RNA that are within *k*T of each other, which means there is not a single structure found

at room temperature. Therefore ssRNA viruses do not require work to be done for the their genomes to be packaged; rather, formation of the capsid around the genome occurs as a spontaneous co-self-assembly of RNA and capsid protein Single-stranded RNA viruses usually do not do work on their genomes, instead they rely on self-assembly to package their RNA genomes with their protective capsids[15].

Self-assembly is a process in which discrete disordered components order themselves with the help of close-range associative interactions into supramolecular organizations. Single-stranded RNA viruses use self-assembly to effectively package the RNA within the capsid without any work being done on their genetic cargo. The assembly of these capsids is a delicate and poorly understood topic, with large physical constraints placed on many of the interacting subunits[14]. The viral RNA must be of the appropriate size; an RNA that is too large will not allow the capsid to form. RNA that is too short may not interact sufficiently with its capsid protein to provide enough driving force to induce encapsidation[15]. Capsids formed around nucleic acid that is too short may not be the same T number as those formed around wild-type RNA, and are less stable overall [14].The interaction strength between the capsid protein and RNA, as well as between capsid subunits, must be appropriate. Interactions that are too strong will result in kinetically formed aggregates that do not give rise to mature virions, and weak interactions will not lead to RNA-protein associations strong enough to generate mature particles.

A closer look at +ssRNA viruses

The length of ssRNA virus genomes is generally much shorter than those of dsDNA viruses, and it remains unclear what the exact evolutionary reasons are for this[25]. While geometric considerations, such as size of the genome in solution and the volume of the final capsid, must play a role in the limiting of these genomes, other factors also impose limitations on these viruses. The lack of strong motor

proteins means that there is a limit to how much "work" can be done on packing the genomes into their respective capsids. The replication of ssRNA is much more error prone than the replication of DNA, owing to the lack of intrinsic RNA proofreading and repair enzymes[26]. Because of these factors, a longer RNA genome would incorporate more mutations, any of which could prove fatal, than a shorter, more efficient genome. The above factors result in a genome that codes for a surprisingly low number of proteins with high genomic densities[3]. Both Nodamura virus and Flock House Virus(FHV) have only 3 ORFs that code for a scant 4 proteins [7].

To circumvent these stringent size requirements, ssRNA viruses employ a variety of techniques to ensure a high gene density per unit length. These include genome segmentation, subgenomic replication, and the use of alternate reading frames[27]. Genome segmentation is perhaps the simplest method counteracting these size requirements imposed by the capsid protein. In segmented viruses, the whole genome is broken up into two or more individual RNA molecules, that are then packaged into either one capsid, as seen in FHV/NoV, or multiple capsids, such as CCMV, which packages its tripartite genome into three separate capsids[28,29]. In these viruses, infection is not initiated unless all three virions (corresponding to a complete genome) enters into a cell.

Subgenomic promotion is a strategy employed by a large number of ssRNA viruses, and involves the transcription of a shorter RNA than the full length RNA. Subgenomic RNAs are often produced in high copy number, owing partially to their smaller size, and allow for the production of two separate gene products from the same RNA sequence[30]. In NoV, protein B1 is expressed subgenomically, and is simply the N-terminal portion of the replicase protein, Protein A. Protein B1, however, has a distinct biological function from protein A, and thus allows for the virus to code two useful proteins from a single open reading frame. Subgenomic promotion also makes the virus more responsive to single point mutations, as both the full length and subgenomic protein products will be altered upon mutation of a single nucleotide[28].

Alternate reading frames provide yet another method of ensuring a high gene density in ssRNA viruses. Found primarily in subgenomic RNA transcripts, alternate reading frames involve shifts by 1 nucleotide in the reading frame of the transcript. By shifting the reading frame 1 nucleotide, yet another protein product is generated from the same sequence space. In NoV, the Subgenomic transcript B2 is a 1 nt frame-shifted transcript that codes for the protein B2. Unlike the aforementioned B1 protein, the frame-shifted B2 involves a completely different set of amino acids; B2 is known to shut down host RNAi is eukaryotes and is associated with higher levels of RNA accumulation in cells[31] .

+ssRNA Viruses: Lifecycle of an Infectious mRNA

The genome of a +ssRNA virus serves as messenger RNA for the translation of all of its necessary gene products. The life-cycle of a +ssRNA virus begins upon entry into the cytoplasm of the host cell, where the genome (mRNA) is uncoated or released from the capsid protecting it. Little is known about the exact nature of genomic uncoating in +ssRNA viruses. Unlike cellular RNAs, viral genomic RNAs (vgRNAs) lack a polyadenylated tail, and sometimes a 5' methylated cap. Many viruses, such as CCMV/BMV and NoV, employ structured 5' and 3' ends, known as either untranslated regions (UTRs) or replication elements (REs), which help bind both translation machinery and cofactors, as well as viral polymerase proteins. These structured RNA elements help the virus evade the large number of exoribonucleases found in the cytoplasm of eukaryotic cells [32]. The 5' UTR associates with many cellular cofactors involved in translation initiation, such as eif4, and helps recruit the vgRNA to the ribosome, where it is translated [33]. Active translation further helps protect the RNA by binding the molecule to ribosomes, preventing access to nucleases.

Among the first gene products translated in ssRNA viruses are the non-structural proteins involved in the replication of viral genome. These gene products are known collectively as the replicase

proteins, and serve as the primary organizers for the later stages of the viral lifecycle. Replicase proteins are expressed in high numbers, and have many functional elements that aid in the replication of the virus. Replication in this sense refers to the synthesis of the reverse complement of the vgRNA, from which new viral genomes are subsequently transcribed, a process depicted below in Figure 1. There is a direct competition between the binding of the RNA to ribosomes and the binding of the RNA to the viral replicase[34]. This interplay surely plays a role in the regulation of the viral lifecycle. As the number of viral replicases increases, the equilibrium will shift from translation to the binding of the RNA by its replicase. The replicase begins replication by binding to the 3' UTR of the vgRNA, and uses it as a template to form the minus-strand RNA, which is the reverse complement of the initial vgRNA . These newly synthesized RNA molecules form a transient double-stranded RNA intermediate that serves as the template for subsequent generation of new vgRNA (seen in Figure 1) [35].

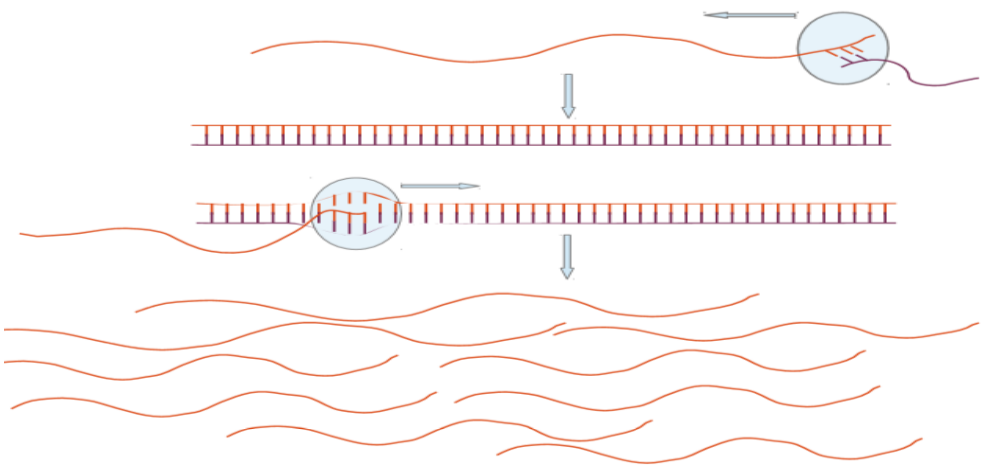


Figure 1: RNA replication: The genomes of +ssRNA viruses are replicated using a transient double-stranded intermediate formed after the replication of the reverse compliment of the viral genome. The positive strand of the viral genome is first replicated by viral RNA-dependent RNA Polymerases (RdRps) proteins – see light-blue bubbles, which bind to untranslated RNA sequences on the 5’ and 3’ end of the RNA. These regions, known as

untranslated regions (UTR) contain important secondary-structures that recruit the replicase proteins. Genome replication results in in the formation of a double-stranded intermediate, from which many copies of the initial RNA are transcribed. Genomes replicated this way commonly amplify themselves 100,000s of times in a single cell.

However, double-stranded RNA is readily detected by the host cells natural defense mechanisms. The eukaryotic interferon response is activated by double-stranded RNA, so it is critical that these intermediates, which are an important step in the viral lifecycle, be hidden from the host cell's immune responses. In order to prevent the detection of the vital double-stranded intermediates, all known +ssRNA viral replication does not occur directly in the cytoplasm of the cell [36]. Instead, the replicase proteins contain domains that are known to actively remodel internal membranes, transforming membranous organelles into viral replication factories that protect the double-stranded RNA intermediate [36]. Nodamura virus, in fact, requires the presence of phosphatidylethanolamine (PE) (found in these internal membranes) to replicate viral RNA [37], and this trait is shared with CCMV. Each ssRNA virus replicates in different replication complexes. CCMV is known to remodel the endoplasmic reticulum in both plants and yeast, while Nodamura virus is known to actively remodel the membranes of the mitochondrion [36]. These replication complexes vary in size and shape, from spherules to tubules, but are always connected to the cytoplasm through necks around 10 nm in diameter [38]. This hole is smaller than the 3D size of the viral genome, which is roughly 25nm[39], and thus the genomes cannot simply depart the replication factory via diffusion. The newly synthesized progeny RNA that are transcribed from the double-stranded intermediates are exported out of these pores into the cytoplasm for packaging and translation. These complexes also serve to concentrate the relevant materials needed for transcription and translation (NTPs, tRNAs etc.), which helps the virus ensure robust replication.

After synthesis of the double-stranded intermediates, often there is a switch within the replicase, allowing it to transcribe new progeny RNA that are identical to the initial vgRNA RNA that entered into the cell. The replicase can bind to several sites on the dsRNA intermediate. If it binds to the 5' end of the RNA, it will transcribe a full length vgRNA that is ready to be packaged by the viruses' capsid proteins. The replicase can also bind to internal sites on the dsRNA, resulting in the transcription of subgenomic RNAs, which are shorter than the full length molecules and often code for separate proteins[27]. There is often a physical interaction between replicase and capsid protein that facilitates packaging of replicating RNA by the capsid protein; because the self-assembly of ssRNA viruses is less specific than that of other viruses, this physical interaction assures that only replicating viral RNA is packaged in the capsids.

Replicons: self-amplifying RNA

Segmented +ssRNA viral genomes can provide intriguing and useful molecules for gene delivery when all the requisite molecules for genomic replication are encoded in a single vgRNA molecule. In this case, no other RNAs or gene products are needed to establish conditions that allow for the replication of the initial RNA molecule within a suitable host cell. Because these molecules do not code for the structural proteins that comprise a capsid, they do not represent a true virus or even a viral genome. Instead, they are a class of self-amplifying RNA molecules known as RNA replicons. An RNA replicon is any RNA that can replicate itself solely from its own gene products. In their simplest forms, these RNA will contain 5' and 3' UTRs needed for recruitment of the replicase and a gene that encodes for the replicase molecule. The replication conditions of these molecules can vary dramatically from "natural" infections, as these replicon molecules do not express the structural proteins, which can affect the replication cycle of the virus by binding newly synthesized genomes. If the genome is bipartite,

interactions between separate RNA molecules can also affect the replication of either molecule. Because the replication of a replicon molecule is fundamentally different than an active viral infection, we call this state a “pseudoinfection”, as it exhibits some, but not all, of the traits associated with true viral infection.

Many replicons have been developed for the treatment and study of pathogenic diseases, as the replicon itself is not directly infectious, and therefore poses less of a threat to researchers involved in its study. RNA replicons derived from wild-type yellow fever virus, as well as Dengue virus, have been used to discover novel small molecule treatments for these viruses[40]. These replicons have also helped researchers elucidate the life-cycle of these pathogens. Other replicons have been generated specifically for the delivery of transgenes; RNA 1 from the bipartite alphanodavirus Flockhouse is an excellent example of this type of molecule [41].

The lifecycle of +ssRNA viruses makes them ideal for the delivery of transgenes. Because they do not replicate in the nucleus, and do not go through a DNA intermediate, RNA replicons do not pose a risk of integrating into the host genomes. Many viral vectors that use DNA genomes, such as adeno-associated viruses, tend to integrate their genome into host genomes effectively at random, which can often cause host mutations and cancer[42] . Several RNA replicons have been engineered from +ssRNA viruses specifically to deliver transgenes. A replicon derived from the alphavirus Sindbis has been designed for the stable expression of heterologous genes that persists for many cell divisions [43]. Replicons such as this have been further used as platforms for cancer immunotherapy and HIV vaccine delivery[44]. These replicons are ideal for vaccines, as they inherently include adjuvants (ss and dsRNA) and replicate to high copy number[45]. However, Sindbis and other replicons from the alphavirus family are limited by their cell specificity and cell apoptosis rate [46]. Being derived from mammalian viruses, these replicons replicate very well in mammalian cells, and in circumstances where this is not needed, such as for early cancer detection and treatment, these replicons may be paradoxically too effective.

One can also use the genomes of ssRNA viruses to deliver functional, noncoding RNAs to the cytoplasm of cells. Unlike gene products, functional RNAs do not require direct translation to be effective. This broad group of RNA molecules includes miRNAs, long non-coding RNAs (lncRNAs) as well as RNAs involved in transcriptional/epigenetic regulation, such as piRNAs[47] and SnoRNAs[48] Many functional RNAs, such as microRNAs, have a large half-life and are biologically effective at even low concentrations, which make them ideal candidates to be delivered in an amplifying RNA construct [49]. Previous studies have used the genome of both teschovirus and flaviviruses to deliver pre-miRNAs to cells *in vivo*. They do this by engineering long double-stranded duplex RNA into the genome of the virus. This duplex RNA does not affect the replication/translation of the viral protein, and is actively recruited to the DICER complex, where it is processed into mature miRNA, which can then regulate protein expression transcriptionally/post-transcriptionally. Because this system relies on the genome of the virus being physically processed by DICER, a cytoplasmic protein involved in RNA-induce gene silencing, it is only effective when DICER is present, and can access the genome to process it. As such, delivery of pre-miRNAs to cells using the minus strand of the virus is ineffective, as the minus-strand RNA never leaves the replication factories to be processed by DICER. In addition, DICER is often down-regulated in many cancer cells, and the above technique is dependent on the appropriate activity of DICER to function [50]. Finally, the requirement for DICER to be present prevents the use of other functional RNAs, as they are not directly processed by DICER[51].

Nodamura and the nodamuraT2A replicon: lifecycle.

Nodamura virus (NoV) belongs to the super-family of alpha-nodaviruses, a family which includes other well-characterized viruses such as black-beetle virus (BBV) and Flockhouse virus. The super family is further divided into alphanodavirus and betanodavirues. Alphanodavirus are well known for their

abilities to infect three different orders of animals (insects, mammals and fish) spanning both phyla, whereas betanoavirsuses infect fish. Alphanodavirus contain a bipartite genome (shown in detail in Figure 2) consisting of two ssRNA molecules, often referred to as RNA 1 and RNA 2[30]. RNA 1 codes for protein A, which is a multi-component enzyme containing a mitochondrial recruitment sequence, membrane targeting sequence, RNA-dependent RNA polymerase sequence and a self-association sequence. This protein is the single non-structural protein needed for the replication of the RNA *in vivo* and *in vitro*[37]. Within the cellular milieu, protein A will replicate both RNA 1 and RNA 2, resulting in the formation a double-stranded intermediate (Figure 1). However, in the absence of glycolipid, subsequent transcription of progeny positive-sense RNA does not occur[52]. In the presence of these lipids, the replicase is capable of subsequent transcription *in vitro*. This mimics the natural life cycle of the alphanodavirus, which replicate within compartments formed in the outer membranes of mitochondria, which are rich in glycolipid. RNA 1 in Nodamura also codes for two additional proteins: B1 and B2. B1 protein is a 111-amino acid protein whose function is unknown in alphanodavirus. In betanoavirsuses, this protein is thought to prevent mitochondrial membrane depolarization, which prevents early cell death during the initial stages of replication. Early stages of Nodamura infections are generally characterized by the modification of the mitochondria, creating replication factories for the protection of the double-stranded RNA intermediates, and this remodeling can depolarize the mitochondria, which can kill the cell[36]. Protein B2 is a strong inhibitor of cellular immune responses, and it is thought that this protein assists in the establishment of infection, particularly in insects, which rely heavily on RNAi as a kind of innate immune system[53]. RNA 2 codes for the capsid protein. Little is known about the process in which capsid protein self-assembles around genomic RNA. It has been shown, however, that upon self-assembly the procapsid Protein α, splits into two protease products β and γ, and that this proteolysis results in the formation of the mature virion[54].

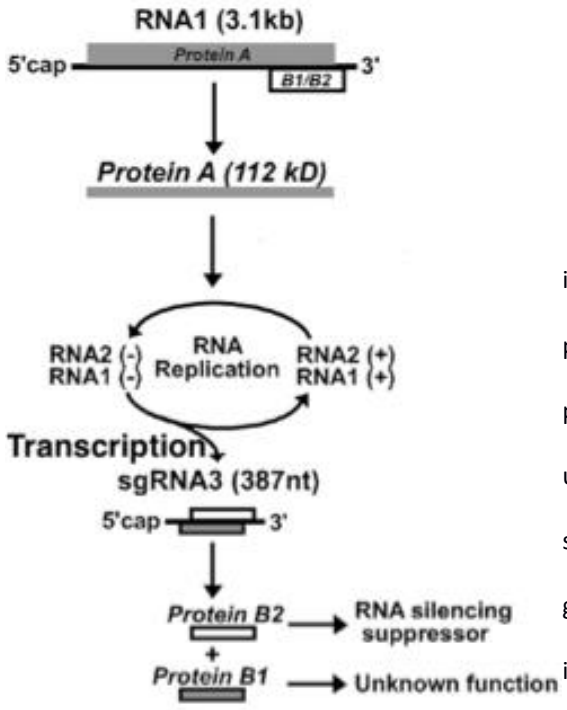


Figure 2: Schematic of *Nodamura* Replication cycle.

RNA 1, which codes for the replicase protein is initially translated, yielding multiple copies of the polymerase. RNA 2 codes exclusively for the capsid protein of *Nodamura*. This polymerase then binds to the 3' untranslated region of either viral RNA, and synthesizes a strand that is the reverse compliment of the initial viral genome (known as the negative-sense strand.) This results in a double-stranded intermediate from which new plus-strands are transcribed. Transcription can either be genomic, which results in a copy of the initial genome, or subgenomic, where only a small portion of the genome is transcribed. In *Nodamura,* subgenomic promotion transcription results in two distinct subgenomic transcripts: sgRNA 1 codes for protein B1, which is simply the N-termini of protein A, and whose function is not known. sgRNA2 is a single-nucleotide frameshift of sgRNA 1, and codes for protein B2, which is a potent inhibitor of host-cell RNAi.

RNA 1 from both FHV and NoV have been modified to generate replicons expressing heterologous genes. These replicons have several interesting properties that make them compelling molecules for protein expression. They work in a large number of cell types, and their range of natural hosts ranges from insects to mammals and fish. Flockhouse RNA 1 has been engineered to deliver a

variety of transgenes both to cells in culture as well as mice[55]. Unlike most mammalian viruses, NoV replicons work more efficiently at 30˚C, owing to their primary hosts being insects, fish and other endotherms[34]. This means that the expression of their transgenes, as well as replicase proteins, is attenuated at 37˚C[56]. Because much of the cellular apoptosis is thought to be from protein-mediated mitochondrial restructuring, it was thought that these vectors would provide a platform for a more limited expression of heterologous genes in general across the organism, limiting off-target replication. In addition, this temperature dependence would allow for tissue specific targeting of tumors that replicate in intrinsically cooler environments, such as melanomas and testicular cancers.

Gitlin et al. generated a single molecule RNA replicon that is an ideal platform for this type of transgene expression[57]. By engineering a stop codon at the start of the B2 reading frame, the levels of RNA replication and translation are attenuated compared to that of the wild-type replicon. In addition, a self-cleaving peptide was introduced into the replicon downstream of the ORF from protein A, and the wild-type stop codon at the end of the ORF was removed. This self-cleaving peptide (T2A), derived from porcine Teschovirus, allows for the co-expression of two gene products from one ORF. By using the fact that RNA 1 is replicated and translated to high levels throughout the replicon infection, the inclusion of the peptide allows for continuous expression of a transgene. The expression of this transgene far exceeds that achieved by delivery of a similar amount of mRNA, demonstrating the high amplification of the replicon RNA within cells. The general schema for this replicon is reproduced below (Figure 3)

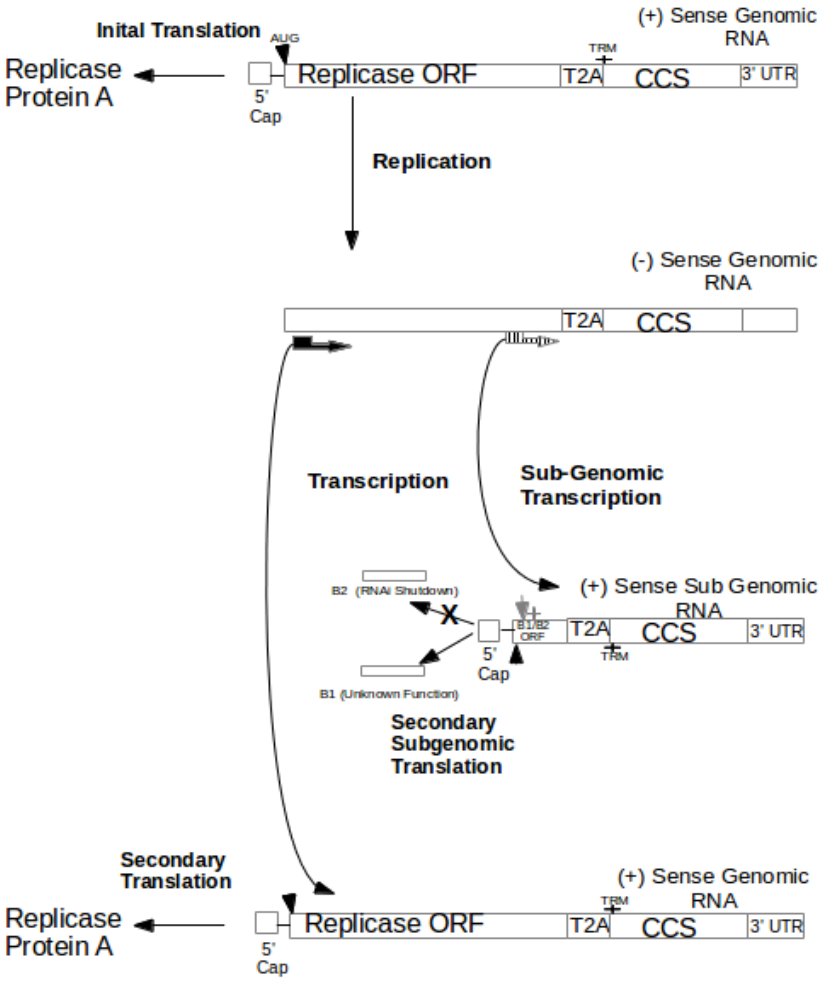


Figure 3: The T2A replicon. The T2A replicon contains the open reading frame for Protein A, a multi-functional protein that is responsible for all facets of viral replication. The stop codon has been removed, and additional sequence encoding a self-cleaving T2A peptide, as well as a poly-linker region have been added to the 3' end of the RNA molecule. This allows for the insertion of a gene of interest into the replicon RNA. This gene (or functional RNA as presented in Figure 7) will be translated concomitantly with protein A, with the T2A peptide cleaving during translation, resulting in functional Protein A as well as the reporter gene. The exact nature of this CCS insert is described in Figure . Upon entry into a mammalian cells, the replicon RNA is initially translated by host-cell machinery, yielding copies of protein A and the reporter gene of interest. Protein A then replicates the RNA, yielding an RNA that is the reverse compliment of the replicon mRNA, known as the minus strand. The minus-strand RNA forms a transient double-stranded intermediate from which new copies of the replicon mRNA are

synthesized via transcription. One mRNA replicon molecule can thereby be replicated to produce upwards of one million molecules.

MicroRNA: overview, origin, biological activity and role of miR-34a in cancer

MicroRNAs are short (20-25 nt) single-stranded RNA molecules that are one of the keystone regulatory molecules in the eukaryotic transcriptional pathway. There are thousands of microRNAs that have been either predicted theoretically[58] or detected experimentally[59]. While they were originally thought only to be involved in sequence-specific, post-transcriptional gene-regulation, it is now known that they are also involved in transcriptional silencing[60], and interact strongly with a variety of proteins, including p53[60]. These non-canonical interactions increase the capacity for each microRNA to regulate a myriad of protein targets, increasing the versatility of these regulatory molecules.

Discovery of these small regulatory molecules immediately led to the realization that one could deliver small, double-stranded, RNA duplexes to cells, which would be recruited into DICER, and converted into active microRNAs within the RISC. Researchers began to use such short-hairpin RNA molecules to knock down protein expression using exogenous RNA sequences. RNAs produced in this method are known as short interfering RNAs (siRNAs) and can be used to temporally silence protein translation in cells. This knockdown, however, is never complete, owing partially to the exogenous nature of the hairpin RNA, and provides compelling evidence that siRNAs and miRNAs are fundamentally different[61].

MicroRNAs are transcribed within the nucleus of the cell, and begin as long branched structures known as primary microRNA molecules (pri-miRNA.) These transcripts are generated stochastically by action of RNA polymerases within the nucleus. This pri-miRNAs are then processed by a restriction endonuclease known as Drosha, which digests the majority of the pri-miRNA, leaving a single 50-80 nt

double-stranded RNA duplex remaining, known as the pre-miRNA. Drosha is capable of differential processing, and can generate many different pre-miRNAs from a single pri-miRNA, further increasing the number of microRNAs that can be generated form a single transcript in the nucleus[62].

The pre-miRNA is then exported into the cytoplasm of the cell, where it interacts with an endonuclease known as DICER, which is responsible for pre-miRNA processing (Figure 4). DICER interacts with the helical RNA molecule and cleaves off the single-stranded loop, leaving a double-stranded duplex with a 2 nt overhang on the 5' and 3' ends of the duplex[63]. From here, the nascent RNA-induced silencing complex (RISC) is formed via the recruitment of two proteins: Ago1 and TRBP. This complex is the structure ultimately involved in gene silencing. The RISC complex must then choose which one of the two strands to incorporate into the mature RISC complex[64]. This step is of fundamental importance because the RISC complex uses the microRNA to knock-down genes in a sequence specific manner; as the two strands of the RNA within the RISC complex have different sequences, they will knockdown different targets[65]. This is known as off-target silencing and is a significant problem in designing siRNAs, as it is difficult to predict which strand will be incorporated into the final complex[64]. Thus, one pre-miRNA can regulate at least two separate sequences, and this number increases due to the fact that absolute sequence alignment is not needed[66], due to the fact that the RISC complex can also prevent translation of templates that only hybridize partially with the miRNA in question[67].

After selecting a strand, the RISC complex is now considered a "mature" RISC complex that is capable of directing gene silencing. This post-transcriptional regulation can be either due to digestion of the transcript, or translational blocking. When the sequence of the miRNA hybridizes completely with the target mRNA, Ago1 within the RISC complex actively digests the RNA in question. When the hybridization is less stringent, the complex prevents translation by preventing the binding of ribosomes to the mRNA of interest. This digestion occurs in regions known as p-bodies [68]

Not surprisingly, the levels of microRNAs within a cell are directly correlated with its cancerous

state[69]. Being one of the main regulatory molecules means that microRNAs are involved in the suppression of oncogenes, as well as the transcriptional regulation of a variety of housekeeping genes needed for maintenance of cellular homeostasis. There is no singular definition of the cancerous state, particularly at the single-cell level, and thus the cancerous state is not directly induced by the dysregulation of a single microRNA. Instead, cancerous states are often associated with changes in expression of many miRNAs within an organism, and thus represent an accumulation of many changes in microRNA expression[69]. Moreover, functional mutations within DICER itself also induce a cancerous state. Mutations that reduce the efficiency of DICER reduce its ability to process pre-miRNAs, and thus reduces its overall ability to silence potential oncogenes. Complete suppression of DICER results in cell death[70].

Of the many microRNAs that are thought to be critical regulators of the cancerous state, the miR-34 class of miRNAs has gained the most attention[71]. This microRNA was originally predicated theoretically[72], then detected in tissue[73]. Of the three members of this family (miR-34a,b,c) miR-34a is well known to be well correlated with the cancerous state[74]. MiR-34a represses the protein SIRT 1, which is a key part of the p53 regulatory network[75]. It is also inhibits prostate cancer cell spread by inhibiting CD44 [76] and inhibits metastasis by inhibiting c-met translation in carcinoma cells[77]. It also interacts with the E2F pathway in colon cancer cells [78]. Beyond just knocking down protein synthesis, miR-34a also interacts directly with p53[79] to regulate many different pathways [80]. This direct interaction allows for this single microRNA to regulate a broad array of genes in the cell [81]

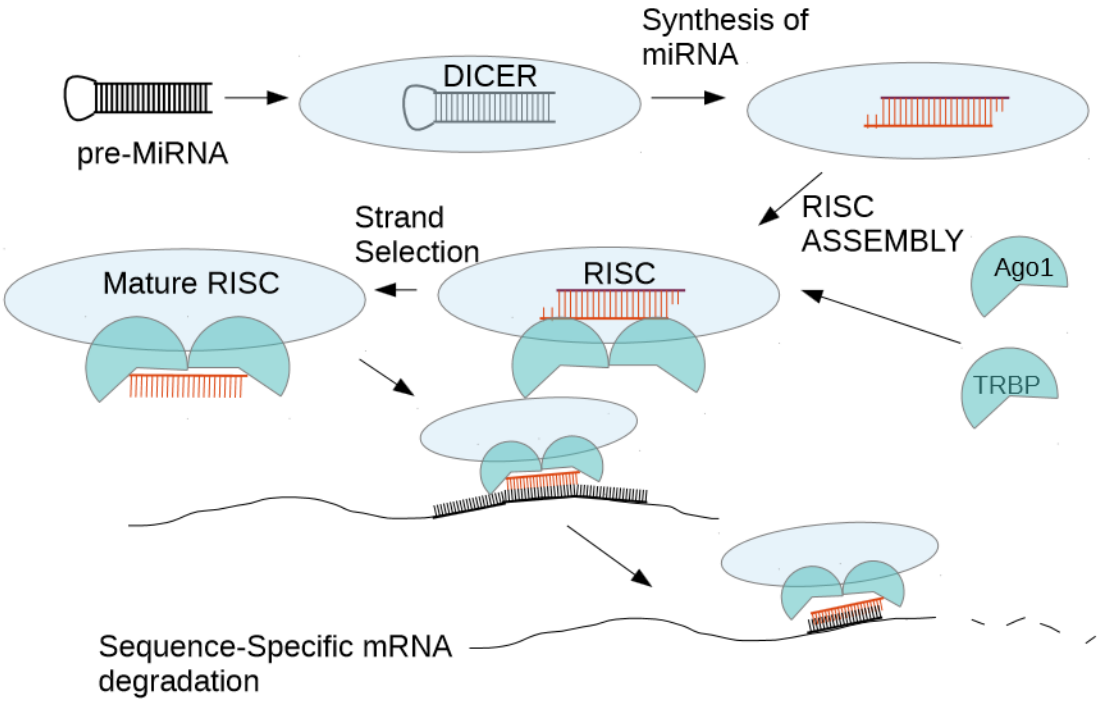


Figure 4: THE RISC Complex: After being exported out of the nucleus, pre-MiRNAs are initially processed by DICER, a cytoplasmic protein responsible for MiRNA processing and digestion. DICER cleaves the short-hairpin of the RNA into two ssRNA strands, which are staggered by 2 nucleotides. DICER then assembles with other cellular cofactors, including Ago1 and TRPB, to form the mature RISC complex. Of the two strands generated by DICER, only one is used for gene silencing. This RNA strand is subsequently incorporated into the RISC complex, where it prevents translation via either sequence-specific RNA digestion or translational blocking.

NodaMiRNA: using replicons to upregulate MiRNAs in cancerous cells

Given their ability to replicate to high copy number in cells, replicons suggest themselves as ideal vehicles for introducing non-coding RNA to tumors and cancerous stem cells. Unlike gene-delivery, these functional RNAs do not need to be translated, so there is no need to be concerned with the

correct translation and folding of exogenous proteins. In addition, because these functional molecules are ssRNA, they can be included in either strand of the replicon RNA. In other words, one can place the functional RNA sequence in the positive-sense genome, or Instead introduce the reverse complement of the sequence, such that when the minus strand is synthesized it yields the correct sequence. We needed a way to separate these functional RNAs from the rest of the viral RNA the same way the T2A self-cleaving peptide cleaves gene-products from the protein A molecules.

Previous studies that have used ssRNA viruses to deliver functional RNAs used the replicons to deliver pre-siRNA hairpins that were embedded within the viral genomes. These long, double-stranded, RNA regions did not affect the replication of the viral RNA, but were effective in knocking down the expression of luciferase *in vivo,* but only when the RNA was incorporated into the plus-sense genome of the replicon molecules. Inclusion into the minus strand did not result in knock down, and this is thought to be because the negative-sense RNA molecules do not enter the cytoplasm and interact with DICER. We sought to create a system that allowed for the delivery of the functional RNA in *either* strand, and one that was DICER independent.

To allow for this dual-strand delivery system, and to eliminate the need for DICER, we used a system of ribozymes that splice themselves out of the replicon genome during replication, resulting in a 22nt long, mature, microRNA in a DICER-independent fashion. This system allows us generate a single microRNA copy per replicon molecule, which means it could generate millions of microRNAs per cell. By introducing this self-amplifying microRNA to cancer cells, we can thereby restore homeostatic levels of microRNA via RNA replication and induce cancer cell apoptosis and prevent metastasis, as restoration of these microRNAs is tied to cancer cell apoptosis and a reduction in metastatic capability. Figure 5 shows a cartoon summarizing the NodaMiRNA strategy. It is also possible for the mature miRNAs to enter the RNA-induced transcriptional silencing complex (Figure 6.)

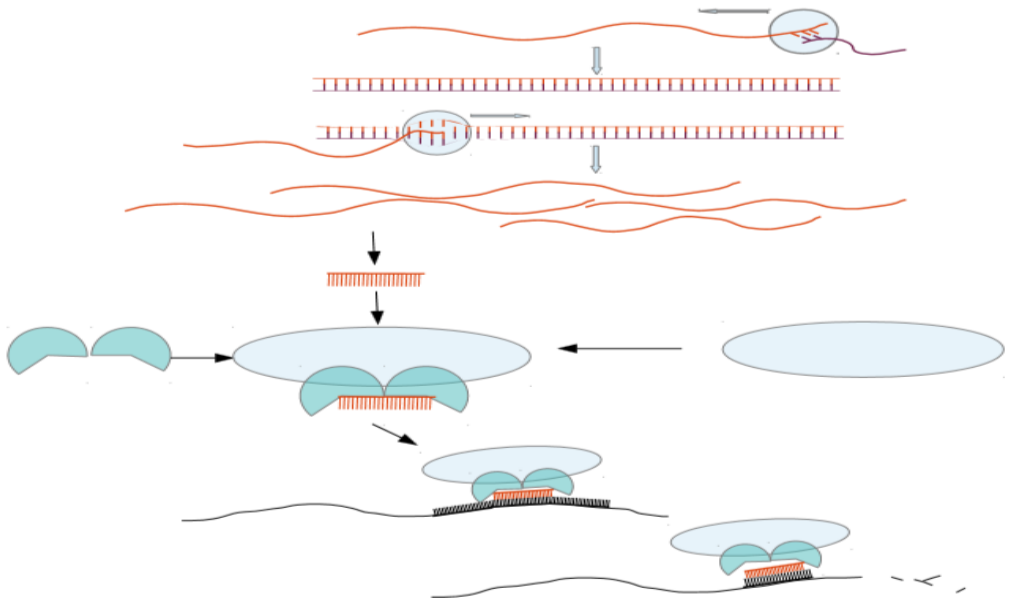


Figure 5: Exploiting Replicon RNA to generate Functional MicroRNA: The ability of replicon RNA to amplify itself makes it a compelling vehicle for delivering functional RNA to cells. One can modify the genome of the replicon, via ribozymes, to produce ssRNA guide strands in a DICER independent manner, or to upregulate biologically relevant miRNA in cells. Guide strands generated in this way are DICER-independent, which is relevant for many cancer cell types, as they are often associated with DICER mutations.

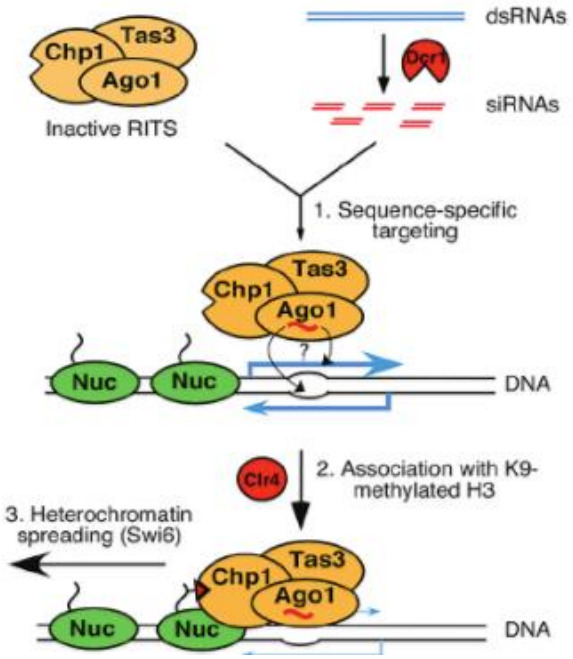


Figure 6: The RNA induced transcriptional silencing complex(RITS). MicroRNAs are also involved in transcriptional silencing. Here, the RITS complex binds to methylated DNA. It digests the nascent RNA (blue) and uses the degraded RNA to further degrade RNA and methylate histones. This leads to a spread of transcriptionally silent heterochromatin[79].

Chapter 2: *In vitro* studies of the Core Catalytic Sequence and NodaMiRNA Sequence

Choice of microRNA and overview of the core catalytic sequence

In order to engineer a self-replicating microRNA, we first needed to choose an appropriate microRNA target. There are 1000s of microRNAs that have been identified via computational studies. Of these, hundreds of candidates have been detected using either deep-sequencing techniques or rational design and testing. With such a broad set of RNA to choose from, we decided to use a class of miRNAs that are thought to regulate a wide set of functions related to tumor proliferation and apoptosis. Here it is again important to distinguish between the seemingly similar notions of siRNA and microRNAs. siRNAs are synthetic RNA sequences that knockdown protein expression post-transcriptionally through incorporation into the RISC complex. siRNAs are not involved in other methods of protein silencing, such as the RITS complex etc., and are not considered to be true regulatory molecules. MicroRNAs, on the other hand, are endogenous RNA molecules that are involved in a myriad of regulatory processes and play key roles in many different regulatory pathways. As such, restoration of homeostatic levels of microRNA does more than simply knocking down the translation of a particular protein; it restores a vital part of an entire regulatory pathway that involves disparate cellular processes such as apoptosis and cell division.

The Mi34 super family of microRNAs was first predicted computationally[31], and later detected experimentally in neuronal cells[59]. This superfamily of miRNAs is further divided into three members, a, b, c. All three miRNAs share some degree of sequence homology, and are key regulators of numerous cellular processes. They are involved in targeting proteins involved in cancer proliferation, cell-cycle control, differentiation and apoptosis[74]. In addition, these molecules affect cancer cell stemness, and

metastatic abilities as well as chemioresistance and response to phytonutrients. Of this super family, Mi34a has been implicated in many cell cancer types. It is an important member of the p53 cell signaling pathway[79] and is involved in the regulation of apoptosis via SIRT 1 interactions[75]. Furthermore, this microRNA has been shown to effectively slow PC3 cancer cell proliferation, while making the cells more sensitive to chemotherapy agents[76].

Choice of ribozymes

In order to generate mature miRNA from the replicon molecule, we employed a class of catalytic RNAs known as ribozymes. Ribozymes are sequences of RNA that possess catalytic activity. These molecules have a diverse set of functions, including splicing as well as small-molecule detection and response. They are also found in viruses, which use them in the synthesis of their genomes[82] (see below). A large number of ribozymes are involved in the maturation of mRNAs transcribed within cells, in particular by cutting out introns. Introns are stretches of transcribed RNA that do not code for amino acids, which must be removed from an mRNA transcript prior to it being translated. In the nuclei of higher eukaryotes, this RNA is spliced out of the RNA by a complex of proteins known as the spliceosome. However, both group 1 and group 2 introns are capable of auto-catalytically splicing themselves out of a mRNA, yielding a mature mRNA that is ready for translation. Group 1 introns, originally discovered in the protest *tetrahymena,* utilize an exogenous guanosine plus magnesium co-factors to catalyze a two-step mechanism that removes the intron and splices together the remaining RNA, resulting in a mature mRNA[83]. The second step requires an additional guanosine at the end of the intron sequence, termed the omega guanosine; removing this internal guanosine allows the intron to cleave itself from the RNA, but prevents the subsequent ligation of the ends of the RNA. Importantly,

the actions of the intron leave a 3' hydroxyl on the RNA of interest, which is a requirement for biological activity in microRNAs.

Many RNA viruses use ribozymes as a convenient and efficient way to cleave their own genomes. Ribozymes are ideal for RNA viruses because they do not require translation and are short in length, which means they “fit” into the crowded viral genome, where space is limited. This class of ribozymes, which includes the hairpin, hammerhead and twister ribozymes, are very short (44-75 nt) and often serve to generate single-genome segments from concatamer molecules comprised of many copies of the genome[84]. This is the case for the hammerhead ribozyme, which is found in satellite viruses and viroids. The genomes for these organisms are generated via rolling circle replication in which a circular genome is replicated by a highly processed polymerase, generating a concatamer of individual genomes separated by active ribozymes. These ribozymes then serve to auto-catalytically splice themselves out of the viral genome, resulting in many copies of the individual genome, which is subsequently ligated into its circular form. These ribozymes are on average much smaller (45 nt) than intrinsic ribozymes (500nt), and generally proceed via a single catalytic step that does not require large structural rearrangements or exogenous small molecule cofactors, such as guanosine[85].

The action of these ribozymes involves the rotation of two phosphodiester bonds on the backbone of the RNA molecule, and results in a cyclic 3-5 phosphate bond on the 3' end of the ribozyme, and a free hydroxyl on the target strand. We chose to place a hammerhead ribozyme upstream of our microRNA, so that after splicing, the microRNA would contain the aforementioned hydroxyl group; we were concerned that the presence of a 5' 3-5 cyclic phosphate group on the 5' end of the miRNA might prevent it from being catalytically active. To avoid this, we used a modified group 1 intron as the 3' ribozyme, as it yields an RNA with appropriate ends that can be phosphorylated by endogenous kinases[86].

Design of CCS+

The wild-type sequences for both the group 1 intron (derived from the tetrahymena intron) and hammerhead ribozymes were obtained and engineered to auto-catalytically splice themselves out of a transcript, yielding a 22-nt, mature microRNA (miRNA-34a). The last guanosine present in the Group 1 intron was removed in order to prevent the ligation of the 3' end sequence to the microRNA. This plasmid, was synthesized with appropriate restriction sites included for later cloning and is depicted in Figures. 7 and 9. This sequence was termed the core-catalytic-sequence (CCS+) and was be used in all further experiments to generate mature miRNA coupled to replication of the replicon. Figure 7 depicts this process below.

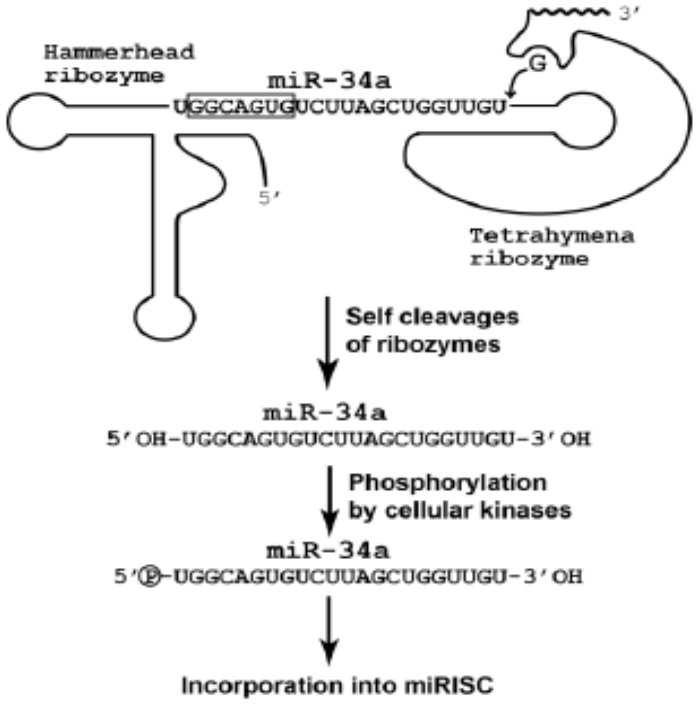


Figure 7: Schematic of the Core Catalytic Sequence, including microRNA of interest (miR-34a). The general schema of splicing, phosphorylation, and incorporation is shown, leading to upregulation of the chosen microRNA, miRNA34-a. This miRNA is a vital part of the p53 feedback loop, and is partially responsible for inducing apoptosis in mammalian cells. Ribozymes were chosen to ensure the released microRNA is biologically active and contains the correct 3' and 5' ends.

Choice of strand: Nod(+) and Nod(-)

The nature of ssRNA viral genomes makes them intriguing candidates for the delivery of functional RNAs to cancer cells. Not only do these molecules replicate to high copy number, effectively amplifying the initial RNA input by thousand fold – but because they generate a negative-sense strand during their lifecycle – one can deliver functional RNA in either strand of the genome. This allows for one to "hide" their functional RNA, ensuring that it is only produced via the replication of the negative strands. Previous studies found that if a pre-miRNA hairpin was incorporated into the negative-sense RNA strand of a flavivirus, the construct does not knock-down protein expression, and the RNA is not processed by DICER. However, delivery of this hairpin via the positive-sense RNA genome was successful, and the miRNA delivered was capable of gene silencing[51]. This is thought to occur because the negative-sense RNA is bound-up in double-stranded intermediates that do not leave the viral replication factories, while the positive-sense RNA is actively exported out of the replication complexes. Our system is more flexible, as it does not depend on the export of whole RNA genomes, instead relying

on the diffusion of small RNAs through the pores of the replication complexes. Because of this, we thought it likely that we would be able to engineer a construct that is both catalytically active and biologically effective using the minus strand of the virus.

For our purposes, delivering our CCS construct in the minus strand was compelling due to the activity of the ribozymes selected to generate the miRNA. Both ribozymes splice co-transcriptionally, which means that many of the plus-sense RNA molecules produced during in-vitro transcriptions were already spliced and would not replicate in cells. We chose therefore to pursue a CCS construct that was only active in the minus-sense strand that was generated during replication, ensuring in this way that all the molecules delivered to a cancerous cell were replication competent. This construct would be comprised of the exact reverse complement of the CCS, such that upon replication and synthesis of the negative strand the CCS would become active and splice itself, yielding a mature microRNA from the nascent negative strand. The sequence for this core catalytic sequence that is active in the negative strand (CCS-) is shown in Figure 13. We termed this family of microRNA-containing NodaT2A replicons NodaMiRNA. Replicons that contained the CCS in the positive-sense RNA were termed Nod(+); similarly, replicons containing CCS(-) were termed Nod(-). Note that they all contain the replicase gene in its positive-sense (as it must be in order for the replicase protein to be synthesized): it is only the CCS insert that is alternately + or -.

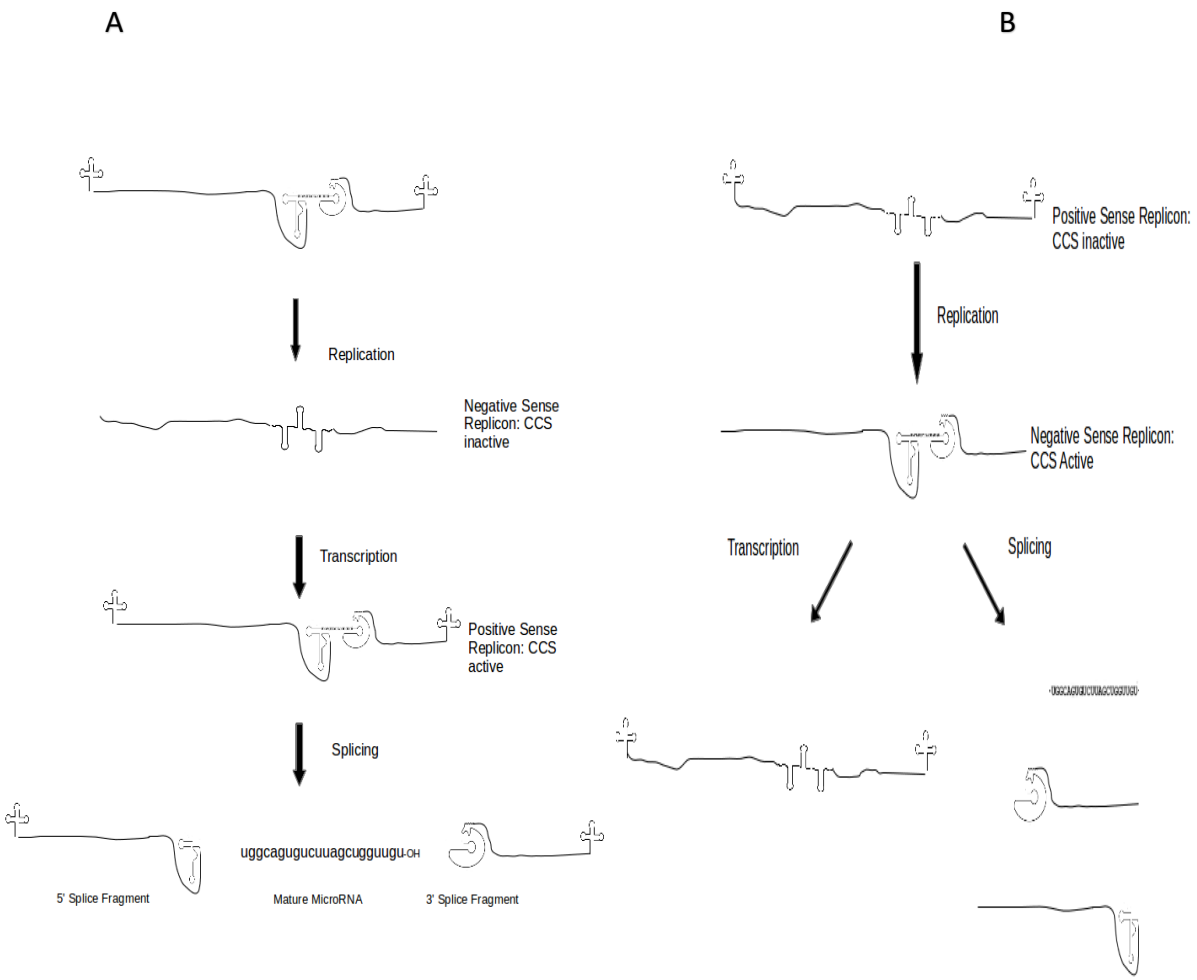


Figure 8: Two strands, both alike in dignity: The nature of Nodamura replication allows for the use of either strand of the genome for functional RNA delivery. Choice of the positive-sense strand results in the schema seen in **A,** on the left, where unspliced replicon molecules are first replicated by replicase proteins, resulting in a catalytically inactive negative-sense RNA strands. Subsequent transcription of those strands results in catalytically active RNA molecules, which then splice, resulting in the release of mature miRNA34-a. Choice of the negative strand results in the system shown in **B**, on the right. Initial replicon RNA is inactive, and unspliced. This RNA is replicated by protein A, resulting in catalytically active

Methods

Design and synthesis of CCS+ and CCS-.

Both DNA sequences for CCS (+) and CCS(-) were purchased from Genscript. The plasmids containing the sequences were transformed into DH5a competent bacterial cells(NEB), and were subsequently purified from bacterial stock using a miniprep column. These plasmids were used to clone the CCS sequences into the NodaT2A replicon, as well as to generate a linear PCR template for *in vitro* transcription reactions of the CCS(+) construct. Both CCS sequences were designed to include a Nde1 and Age1 restriction site on their 5' and 3' ends respectively, and these restriction sites were used for all sub cloning performed. The sequence of the CCS(+) construct is shown in Figure 9.

CCCTCTACGGTGACTACTCAGGCACTCCTGCTTTGCCATA(HAMMERHEADRIBOZYME)TGGCAGACCG
TCACAGAATCGA(MicroRNA)CCAACATGTTGGTCGATTCTTTATCGTTATAAATGGAAACCTCCCTTTTC
AATAGTCCGTACGTGGACCATCGATCAGAAATTTGGTTATCTAACGTAGCCAAATTTTCCGTTCTGGCAG
TTTAACGCCCTTTCCCCAGTTGTCGGCAAGTCATGGTTCAGAGTCCCCTTTGAAACTCTACCGGAACGTTT
CCCATACCATTATTCGACTGCCTGTACCAGGATTGGTGCGTCGGTTCAGGATTCAGTTGTCTAGAAGACA
ACTATACCTACGTCAAGTGTCTGATTTACAGCCAGCCCCTTCTACATAAGAAGAGTATTCTATATCAGCCT
GGAGAGGAATTACCCTCGATCGCCTACTTCACTACGTTGTGACCTCGGCGACCCTTGATTAAACATACGC
TTTCATATAACTAATCAAAACCTCATGA(GROUP 1 INTRON)GGAGCTC(3' RESTRICTION
SITE)accggtgggcgtacg(RNA
1,RNA3)CCAAGTAAGCGGCAGGCGGAAGGCCGCAGCCGCAGCTCCCGTGGCCCAGCTGGGTCGCGTG
GGCGTGGTAAATGAGTGATTCATCGTCCCATCTGACGAAACCCGAACTAGGCTTATGCCAGTGGT(RNA3
)

Figure 9: Sequence of CCS(+) Construct. All relevant sequences are color coded and labeled.

In vitro transcription of CCS(+) sequence

The DNA sequences for the DNA template used in the initial *in vitro* transcription reactions were generated using PCR to add the full T7 promoter to the 5' end of the CCS. The resulting 505nt blunt-end product was purified with a Qiagen Qiaquick column, and used for subsequent *in vitro* transcription reactions. This sequence contains both active ribozymes, as well as the miR34-a sequence. Upon transcription, these ribozymes should splice, resulting in the release of the microRNA.

Initial *in vitro* transcription reactions were performed using 0.5µg of purified DNA template in transcription buffer (10mM MgCl2, 20mM Tris, 4mM DTT, 2mM Spermidine) supplemented with RNAse inhibitor RNasein, from ThermoFisher. During transcription, 5uL aliquots of transcription mix were taken at indicated time points (see Figure 10) and immediately added to 5uL gel-loading dye (95% formamide, 10mM EDTA, 10mM Tris) and heat denatured at95˚C for 5 minutes prior to loading onto a 15%, 8M urea PAGE gel. The gel was pre-run for 30 minutes at 20 W, the samples loaded, and the gel was subsequently run at 10W for 30 minutes. It was then rinsed extensively with deionized water, and stained for 5 mines with GelRed. The resulting gel was imaged using a UV transilluminator.

Cloning into single-molecule Nodamura constructs

After verifying the activity of the CCS, the construct was sub-cloned out of its initial vector via digestion with Nde1 and Age1. This yielded a single, clean fragment that was then ligated into a custom vector to be used in transfection and transcription experiments. The presence of the target sequence in the Nodamura-replicon vector was determined by both restriction digest (using Age 1 and SacII) as well

as sequencing, and this vector was used to generate the full length Nodamura replicon – Nod(+) – containing the CCS, for transcription studies and transfection experiments. This replicon schema, seen in Figure 2, will generate thousands of copies of itself *in vivo*, effectively restoring non-cancerous levels of mir34a.

Nodamura splicing

The above plasmids were linearized, and replicon molecules were generated via *in vitro* transcription with an Invitrogen mMESSAGE mMACHINE T7 transcription kit. Nod(+) template was linearized and transcribed using an mMESSAGE mMACHINE T7 transcription kit. Aliquots were collected at 0, 15, 30, 45, 60, and 120 minutes and were immediately denatured in gel loading buffer (7% formaldehyde, 50% formamide, 10mM EDTA) at 65˚C for 10 minutes. They were then frozen at -80 until loading onto a 1.2% denaturing formaldehyde agarose gel and running at 80v for 3 hrs. Gels were extensively de-stained in deionized water for 2 hours, then stained for 30 minutes with GelRed and imaged using a UV trans illuminator.

Measuring co-transcriptional splicing of Nod(+) and Nod(-) RNA.

The resulting gel image was analyzed using the image analysis program imageJ to determine the relative intensity of the spliced and unspliced Nod(+) construct. This resulted in peaks which could be integrated to determine their respective overall intensity. By measuring the ratio of spliced to unspliced RNA as a function of time, one can calculate the percent RNA spliced as a function of time. This method measures the rate of one of the ribozymes only, and is unable to determine if both ribozymes spliced

completely. Most importantly, this technique can give a good estimate for the amount of unspliced RNA present after 1 hour of transcription, which is important when calculating the amount of RNA to deliver during transfections.

Inhibition of Nod(+) splicing by antisense oligos (ASOs)

The high amount of cotranscriptional splicing made us seek a way to inhibit RNA splicing. We decided to use antisense oligos to inhibit the ribozymes as they had the least off-target effects. These anti-sense oligonucleotides are designed to hybridize to the nascent RNA transcript during transcription, prevent the ribozyme from folding into the correct catalytic structure. The anti-hammerhead oligonucleotide (AHHASO) was designed to hybridize to the majority of the hammerhead ribozyme (31 of the total 44 nts) while the anti-group-1-intron oligonucleotides (AG1ASO) targeted specific interactions between the microRNA and the P1.1 loop of the intron. Unlike other inhibitors, such as arginine or neomycin, ASOs do not interfere with the production of full length replicon (data not shown). The sequences for the oligonucleotides used are presented in Table 1 [87] .

Table 1: Sequence of oligos used to inhibit either the 5' hammerhead ribozyme or 3' group 1 intron.

Intron

GCCTGATAACTTTGATTGCAAATATATTTAC

Hammerhead

CCGTTTCGTCCTCACGGACTCATCAGTGGCA

Results

Initial Splicing of CCS(+) in vitro

Figure 10 shows the gradual accumulation of 22nt RNA during the *in vitro* transcription of CCS(+) template in a 15%, 8M Urea PAGE gel. At the specified times, aliquots from the *in vitro* transcription reaction were denatured in RNA loading dye and loaded onto the gel. Starting at around 20 minutes, a strong band at 22nt appears in each lane studied. This band is not present at earlier time points, and gradually accumulates. After about 1 hour, the bands intensity appears to level off, indicating that either the *in vitro* transcription reaction has stopped, or all of the RNA has spliced. Also visible, but not shown, are the 5' and 3; splice products, which are roughly 44 nts and 460 nt long, respectively. Presence of both these bands, coupled with the high intensity of the band running at 22nt, provides strong evidence that both ribozymes are splicing effectively. The reaction conditions of *in vitro* transcription are ideal for splicing of both the ribozymes; the mix contains high levels of magnesium ions, as well as the guanosine cofactor needed for splicing.

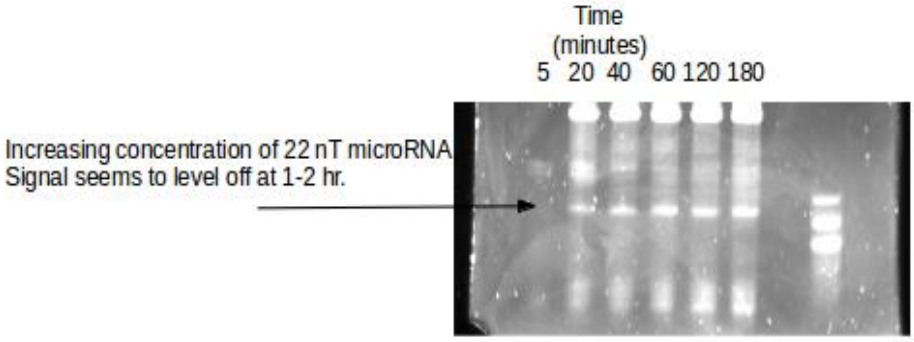


Figure 10: 15% denaturing Urea-PAGE image showing the gradual accumulation, co-transcriptionally, of miR-34a. The far right lane corresponds to the microRNA ladder with lengths (from top to bottom) of 25, 21, and 17 nt. Template for *in vitro* transcription was generated using PCR primers containing the full T7 promoter. During transcription, aliquots of transcription mix were taken at certain time points and immediately added to gel loading dye (60% formamide, 10mM EDTA) and heat denatured at95˚C for 5 minutes prior to loading. Shown above is 15% denaturing Urea-PAGE image, demonstrating the gradual accumulation, co-transcriptionally, of miR-34a. This species runs as a clean, sharp band at 22nt, and there appears to be no alternative splicing (no other bands) occurring. The signal intensity for this band levels off at around 60 minutes, indicating near-complete splicing at 60 minutes. A 10% Urea-PAGE gel (not shown) was also run, this time looking for the presence of the 44 nt hammer-head fragment and the approximately 400 nt intron. Both these bands were present in this gel, and their relative intensities match that of the microRNA levels, providing good evidence that they are splicing correctly

In vitro transcription of Nod(+) RNA shows CCS(+) is unaffected by replicon secondary structure

After verifying that the CCS(+) was able to splice co-transcriptionally, we subsequently cloned the insert into the NodaT2A replicon. This construct would allow for the release of mature microRNA

coupled to NodaT2A replication. The ability of this construct to splice co-transcriptionally was also assessed, as it was thought that the RNA secondary structure of the replicon could affect the splicing of the CCS(+) The plasmid was linearized, and an *in vitro* transcription reaction was performed. Aliquots were taken at six time points (t= 0, 15, 30, 45, 60, 120 minutes), shown left to right in Figure 11 A. A ladder is shown in the far-right lane, with the 3000 nt band having increased intensity. The samples were immediately denatured in gel loading buffer (7% formaldehyde, 50% formamide, 10mM EDTA) at 65˚C for 10 minutes. Samples were frozen at -80 until loading onto a 1.2% denaturing formaldehyde agarose gel and run at 75V for 3 hrs.

Figure 11 A initially shows 2 strong bands: one running at approximately 3.9 knt (corresponding to the unspliced Nodamura replicon RNA,) and one at 3.2 knt (corresponding to the 5' splice product of the replicon.) There is also a faint band at approximately 700nt, which corresponds to the 3' splice fragment.

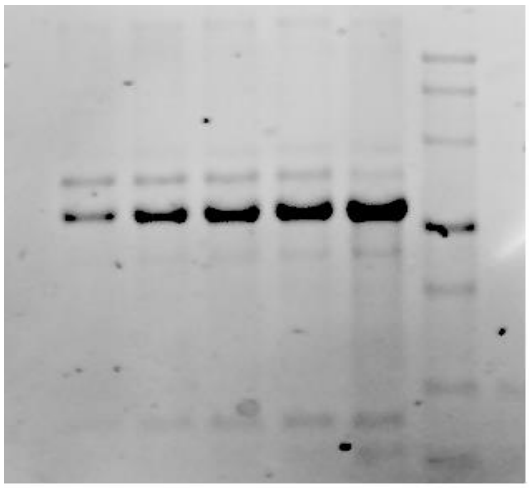
The intensity of full-length molecule is always markedly less than that of the spliced band, and further quantification using image analysis software provides insight into the ratios of spliced and unspliced replicons. Plot-profiles of lanes 1-6 were generated with ImageJ, and representative traces for the t=15 min and t= 120 min time points are shown. While the intensity of the spliced band increases steadily, the amount of unspliced remains surprisingly invariant, indicating a subpopulation of RNA too misfolded to be catalytically active. Indeed, a plot of the relative ratio as a function of time (Figure 11 C), shows a leveling off at around t=60 minutes, with about 20% of the total RNA remaining unspliced even after 2 hours.

The ratio of spliced to unspliced RNA is very low at later time points, which is inconvenient, as only the full-length RNA molecule is able to be fully-replicated *in vivo*. In addition, the 5' splice product is able to be replicated initially, generating a truncated negative-sense strand RNA molecule that lacks the critical 3' UTR. Without this UTR, the minus-sense RNA is unable to be transcribed back into positive-

sense RNA. Moreover, even if this truncated RNA were able to be transcribed (that is converted from minus strand to plus strand), it would inherently lack the CCS sequence and would be non-functional. As such, the 5' spliced RNA fragments may act as sort of defective interfering RNA, limiting the usefulness of this construct. Therefore, it is critical to deliver as much full-length RNA as possible to the cell, which means suppressing as much as possible splicing during transcription.

Figure 11:

A

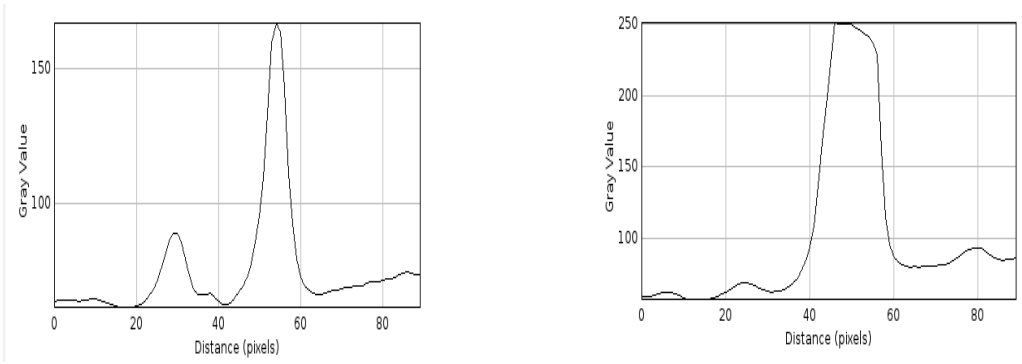


B

t=15mins

t=120mins

C



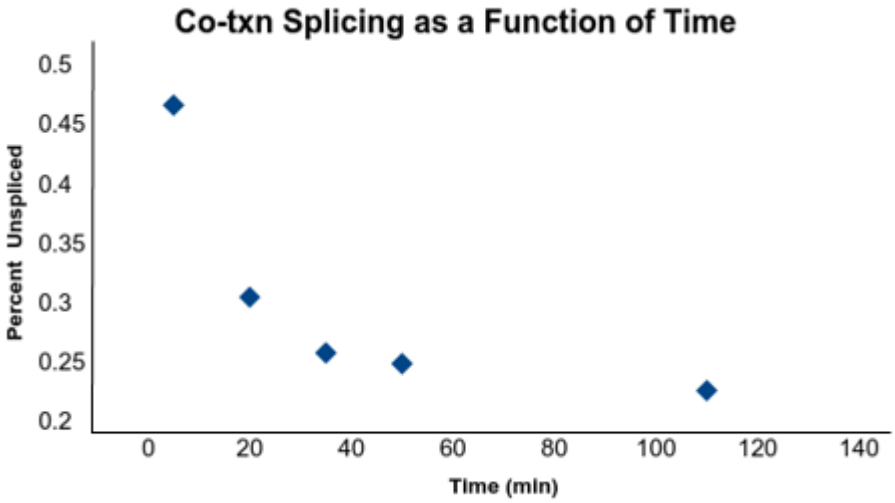


Figure 11: In-Vitro Splicing of Nod(+). **A**: Replicon molecules were generated via *in vitro* transcription with an Invitrogen mMESSAGE mMACHINE T7 transcription kit. Aliquots were taken at six time points (t= 0, 15, 30, 45, 60, 120 minutes), shown left to right above**.** A ladder is shown in the far right lane, with the 3000nt band having increased intensity. The samples were immediately denatured in gel loading buffer (7% formaldehyde, 50% formamide, 10mM EDTA) at 65˚C for 10 minutes. Samples were frozen at -80 ˚C until being run on a 1.2% denaturing formaldehyde agarose gel for 3 hrs. The faint bands at 3.9 knt represent unspliced Nod(+) molecules (*), while the intense bands at 3.2 knt represent the 5' splice product(**). The 3' splice product is visible at approximately 700 knt at 1 hour, at which point approximately 75 % of the molecules produced have spliced. **B** Pixel-intensity plots of two representative time points (t=15mins and t=120mins) from A, generated with ImageJ software suite. Note the overall larger values for the spliced species at approximately 3.2knt (50 pixels) for both time points. The integrated areas can be used to determine the ratio of spliced to unspliced replicon. **C** Measure of splicing of the full molecule replicon as a function of time. By t=60 minutes, the majority of the replicon has spliced at least once, removing its ability to replicate *in vivo*. Note a subpopulation (about 20%) that does not splice even at 2hrs.

Splicing inhibition by anti-sense oligonucleotides(ASOs)

Figure 12 shows the results of experiments using a combination of anti-sense oligonucleotides to inhibit the splicing of both the hammerhead ribozyme, and the group 1 intron at two different total oligonucleotide concentrations (25 (even lanes) and 50 µM(odd lanes).) Co-transcriptional splicing is significantly reduced in the presence of oligonucleotides. This is particularly obvious at the two-hour time point: in transcription mix without oligonucleotides, only 15% of the replicon remains unspliced at t=120 minutes. Compare this to the transcription done with 25 µM oligonucleotides, which has about 65% of the RNA unspliced at long times scales. Being stable for this amount of time is critical, as it will maximize the total yield of full length RNA, which are the only molecules that are capable of establishing a pseudoinfection in cells.

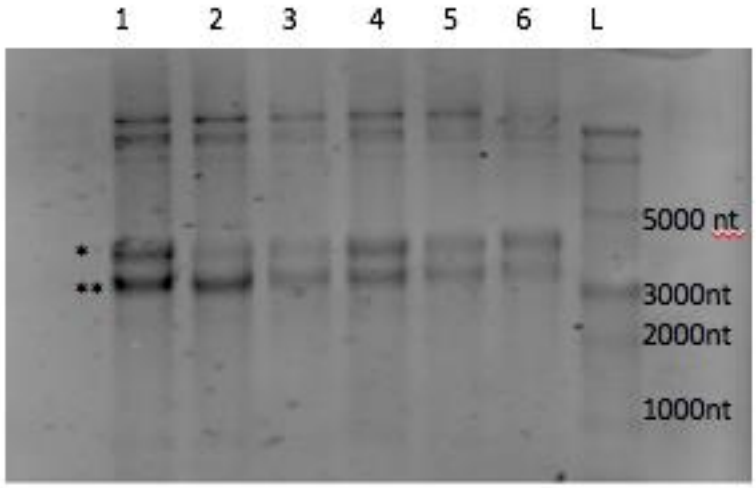


Figure 12: Total ribozyme inhibition: *In vitro* transcription reactions were run in the presence of 25 µM (odd) or 50 µM(even) mixed ASO targeting both ribozymes. Aliquots were taken at t=30 minutes (lanes 1&2) t=60 minutes (lanes 3&4) and t=120 minutes (lanes 5&6.) The samples were immediately denatured in gel loading buffer (7%

formaldehyde, 50% formamide, 10mM EDTA) at 65˚C for 10 minutes. Samples were frozen at -80 until loading onto a 1.2% denaturing formaldehyde agarose gel for 3 hrs. Much more unsliced replicon RNA (*) is visible at all time points compared to transcriptions performed without theses ASOs.

Nod(+) in vitro transcription reactions contain primarily spliced RNA, even after ASO inhibition.

After settling on an appropriate concentration of ASOs, all *in vitro* transcription reactions were carried out at 50uM blocking oligonucleotides for future experiments and for the generation of RNA to be used for transfections. Without the oligo, purified Nod(+) RNA contains little to no active full length RNA molecules. Inclusion of these oligos left approximately 30% of the RNA unspliced after purification.

It is thought that the purification of the RNA also induced the RNA to splice, as the purification process includes both magnesium ions as well as guanosine analogues.

All RNAs used for transfections were generated as follows. Linearized template was *in vitro* transcribed at 37˚C for one hour, and then immediately treated with DNase for 15 minutes and purified using a RNeasy RNA purification column. The yields for these transcriptions were consistently very low, as the large amounts of oligonucleotide were thought to inhibit transcription by chelating magnesium ions needed for transcription. Gel analysis of the purified RNA indicated that only about 25 % of the RNA remained unspliced after purification. However, given the self-replicating nature of these RNA molecules, even this low amount was considered to be suitable for transfection experiments. Of concern was the fact that transfection would deliver large numbers of replication-deficient RNAs (the 5' splice products) which would compete for translation with the full-length RNA molecules, i.e., that these molecules may act as a sort of interfering RNA molecule, attenuating or eliminating active infections. But their lack of suitable 3’ UTRs would make them replication incompetent, and their lack of the CCS would make them biologically inactive. Moreover, the truncated RNA may form dsRNA intermediates lacking appropriate 3’ ends, which may affect the switch from minus strand synthesis to + strand synthesis.

Cloning and analysis of CCS-: catalytic sequence active in the minus strand.

Analysis of the splice products from *in vitro* transcription of Nod(+) constructs indicates that about 75% of the total RNA produced is either singly or doubly spliced co-transcriptionally. The 5’ splice product that results from this is around 3.2 knt long, while the full-length molecule is 3.9 knt. *In vitro* packaging experiments using CCMV indicate that the capsid protein of CCMV prefers RNA lengths closer to 3.2 knt, which would mean that packaging by CCMV could potentially exclude the full length molecules of Nod(+), resulting in particles that would not replicate in cells, as they lacked the critical 3’

replication elements. While we anticipated being able to package both the truncated molecules and full-length molecules by increasing the overall amount of capsid protein used during packaging, we wanted to avoid the delivery of particles that could serve to interfere with the activity of the replicon in cells. Delivery of these particles to cells would not result in an increase in mature microRNAs, as the whole of the catalytic sequence would by definition not be included. Because previous co-transcriptional inhibition experiments demonstrated that we could only prevent about 50% of the splicing from occurring, we sought another avenue to generate a more monodisperse RNA that is more suitable for packaging with CCMV capsid proteins[14].

The nature of ssRNA viruses uniquely allows for the delivery of functional RNA in either sense of the RNA genome. Gene are of course traditionally only delivered in the positive-sense, as the negative-stranded RNA lacks the requisite sequence required for recruitment by the cell's translation machinery. In addition, because the negative-sense RNA is only generated in replication factories, they are effectively prevented from interacting with the ribosome present in the cytoplasm. The CCS, however, does not require translation to be functional, and the microRNA produced does not need to interact with DICER. In addition, the product is small enough to diffuse from the replication factories and enter the cytoplasm, and was not thought to be bound in double-stranded form, due to the strong secondary structures found in the ribozymes.

The CCS- sequence was generated from the CCS+ sequence by taking the reverse complement of the CCS sequence using APE. This sequence is shown below in Figure 13. In doing so, this would ensure that the active CCS sequence would be generated during the replication of the positive strands into minus strands. This sequence was then purchased from Genscript, with the addition of the relevant restriction sites on the 5' and 3' ends to facilitate cloning into the Nodamura-replicon containing plasmid. Both the plasmid containing the CCS- sequence, as well as the original Nodamura-replicon containing plasmid were double-digested with AgeI and NdeI. The two digests were run on a 1% agarose

gel, and the Nodamura-replicon backbone as well as the CCS- insert were extracted from the gel, and purified via a Qiaquick gel extraction system. After purification, the backbone plasmid was SAP-treated, to prevent self-ligation, and subsequently heated to 65˚C to inactivate the phosphatase. The insert and backbone were mixed at a mole ratio of 1:3 vector-to-insert, and ligated overnight at 16c in the presence of T4 ligase. The resulting plasmids were used to transform dh5α ultracompetant cells, and then plated on LB amp plates. Five colonies were taken for screening, and sequencing. The presence of the insert was first confirmed by a double digests and analysis on a 1% agarose gel, and further verified by sanger sequencing performed by Laragen. Plasmid and cell stocks were then grown and stored at -20c and -80c respectively.

ACCACTGGCATAAGCCTAGTTCGGGTTTCGTCAGATGGGACGATGAATCACTCATTTACCACGCCCACGCGACCCA
GCTGGGCCACGGGAGCTGCGGCTGCGGCCTTCCGCCTGCCGCTTACTTGGcgtacgcccaccggtAGCTCCTCATGAGG
TTTTGATTAGTTATATGAAAGCGTATGTTTAATCAAGGGTCGCCGAGGTCACAACGTAGTGAAGTAGGCGATCGA
GGGTAATTCCTCTCCAGGCTGATATAGAATACTCTTCTTATGTAGAAGGGGCTGGCTGTAAATCAGACACTTGACG
TAGGTATAGTTGTCTTCTAGACAACTGAATCCTGAACCGACGCACCAATCCTGGTACAGGCAGTCGAATAATGGTA
TGGGAAACGTTCCGGTAGAGTTTCAAAGGGGACTCTGAACCATGACTTGCCGACAACTGGGGAAAGGGCGTTAA
ACTGCCAGAACGGAAAATTTGGCTACGTTAGATAACCAAATTTCTGATCGATGGTCCACGTACGGACTATTGAAAA
GGGAGGTTTCCATTTATAACGATAAAGAATCGACCAACATGTTGGTCGATTCTGTGACGGTCTGCCATATGGCAAA
GCAGGAGTGCCTGAGTAGTCACCGTAGAGGG

Figure 13: DNA sequence of the CCS(-) construct. This construct exhibits no catalytic activity until it is replicated. Upon replication, it will have the exact sequence of CCS(+) shown in Figure 9 and will be catalytically active.

Prior to *in vitro*-transcription, the Nodamura replicon with the active ribozymes in the negative-sense (NOD-) was purified and linearized with XbaI, whose restriction site is directly downstream of the 3' UTR of the Nodamura replicon. The linearization was performed overnight, and the linearized plasmid was purified with a Qiaquik DNA purification column. This template was then used in the *in vitro* transcription reaction.

In vitro transcription of the NOD(-) sequence was done using an mMESSAGE mMACHINE T7 *in vitro* transcription kit. Oneμg of template was used, and the reaction was carried out at 37˚C for 3 hours, followed by a DNase digestion for 15 minutes at 37˚C. The resulting RNA was purified via a RNeasy RNA purification column, and 1μg of the purified RNA was run on a 1% agarose gel using 50% formamide as a denaturant. Unlike the NOD(+) sequence, which splices readily in-vitro, there was no splicing visible during the transcription of Nod(-). Instead, the RNA ran as a single band at approximately 3.9 knt, which is the length of the full genome.

Discussion

It is clear from comparison of the Nod(+) and CCS(+) splicing gels that the local RNA secondary structure around the ribozymes affects their activity. In the initial transcription of CCS(+) RNA, almost all of the RNA splices co-transcriptionally, and the amount of miRNA produced levels off after about 1 hour. This leveling off coincides with the depletion of NTPs in solution, and the cessation of transcription. Because miRNA does not continue to accumulate after transcription has ceased, we can conclude that the RNA is fully spliced contemporaneously in the CCS(+) transcription. The case is different for the Nod(+) transcription, where there is a constant amount of RNA that remains unspliced even after 2 hours post-transcriptional initiation. We suspect that these RNAs represent misfolded replicons whose secondary structures are in some way inhibiting the activity of the ribozymes, possibly by base-pairing to

either the miRNA sequence or one or both of the ribosomes in question.

In vitro transcription reactions indicate that both the CCS(+) and Nod(+) constructs splice efficiently outside the cell, and do so in a co-transcriptional manner. For the CCS(+), both the ribozymes can be seen to splice almost immediately upon transcription. The activity of these ribozymes can be attenuated by the inclusion of anti-sense oligos that bind to critical regions of both the ribozymes. Because ribozymes depend fully on their ability to base-pair with their target sequence, any disruption of this interaction can cause the ribozyme to become inactive. There is a myriad of molecules that inhibit ribozyme splicing[88]. These include EDTA, neomycin as well as arginine[85] and other guanosine analogues. Unfortunately, many of these molecules are also potent inhibitors of *in vitro* transcription reactions. Neomycin chelates magnesium needed for ribozyme activity, yet this also effects the T7 polymerase. Many molecules were tried, and all small-molecule inhibitors proved to be too deleterious to *in vitro* transcription, yielding a smear of RNA or inhibiting the action of the polymerase completely.

We chose ASOs to inhibit the action of both ribozymes because their method of inhibition does not affect the activity of T7 polymerase. These oligos work by hybridizing to important regions of the ribozymes, which prevents them from splicing as intended. Initial inhibition experiments were done using just the CCS construct, and allowed us to develop a pair of oligos that worked for the short sequence. These oligos also worked for the full-length Nod(+) sequence, and allowed us to purify unspliced RNA. After purification, the RNA is much more stable, as it is stored in deionized water that contains neither magnesium ions or guanosine. The RNA is also stable in lipoplexes, as the large amount of positively charged lipid presumably help keep the RNA in non-active secondary structures.

Even with these oligos, there was still much more spliced RNA than unspliced, and there was a concern that the nature of RNA replication would result in poor RNA amplification *in vivo*. RNA replication is dependent on both the 3’ and 5’ UTRs, and any spliced RNA molecule (either the 5’ or 3’ splice product) would only contain one of the two needed RNA regions. In transfections of Nod(+), we

would be delivering RNA that would be translated, but not replicated in the form of spliced replicons. At the time we were unsure if transfection with a small amount of replicon RNA would be enough to induce pseudoinfection. Clearly in the wild-type viral infection only one copy of the genome is sufficient to enter the cell and set up an active infection, and we wanted to find a way to determine if small amounts of replicon RNA could do the same. To do so, we designed a series of experiments to selectively deliver controlled amounts of active NodaT2A replicon molecules to cells using transfection. These experiments help us further understand what was occurring in Nod(+) transfection and are described in the next chapter.

These results also underline why the minus-strand is such a compelling choice for the delivery of the CCS. We initially designed the construct because we wanted a more homogenous RNA for assembly. Because the splicing only occurs in the negative-sense RNA strand, there is no co-transcriptional splicing, and 100% of the RNA produced is replication competent. As such, no ASOs are needed, and as a result the transcriptional yields of Nod(-) are much higher than those of Nod(+). In addition, these transcription reactions can be run for longer periods of time, as there is no concern about splicing. This results in more RNA produced per reaction. However, we did not have direct evidence that the Nod(-) strand would splice *in vivo* or *in vitro*. Even though we knew Nod(+) would splice *in vitro*, the sequence of the minus strand replicon RNA is different, even though the CCS(-) sequence in the minus strand of Nod(-) is identical to the CCS(+) sequence in Nod(+) and has a different secondary structure, which could have inhibited the splicing of the ribosomes after replication.

Chapter 3: Inverse Dose-dependence and Temperature Dependence of Protein Expression by Nodamura RNA in BHK-21

Background

Lipofectamine-based transfections

There is a myriad of techniques used to introduce genomic material into cultured mammalian cells The use of viral agents, such adeno-associated virus (AAV) or lentivirus, is well-characterized and has high transfection efficiency overall. However, these viral vectors often integrate their genomes into the host-cell, which ensures constitutive expression at the cost of being potentially oncogenic[42].

A more common and safer technique is to use amphiphilic lipid molecules to generate complexes that are capable of delivering nucleic acids. These complexes often involve the use of lipids as the complexing agent, which helps the nucleic acid enter the cell by interacting with the plasma membrane. Often these complexes are incorrectly termed liposomes, which are double-layered, spherical lipid vesicles that are tens of nanometers to microns in diameter. The large number of genes and persistence length of dsDNA would prevent all but the smallest DNA molecules from being packaged into a proper liposome. Instead, complexes of lipid and nucleic acid form complex lamellar/ hexagonally close-packed structures with the DNA acting as rigid rods surrounded by lipid bilayers[89]. The lipid molecules routinely used in transfections molecules are cationic in character, as this helps

them bind to the outer leaflet of cellular membranes, which often has a net negative charge. The positive charge on the lipids also aids in the association of the negatively charged nucleic acid with the lipid[90]. The adsorbed lipoplex enters the cell via cholesterol mediated clatherin-coated pits that result in the formation of an endosome[91]. All nucleic acids delivered via lipoplexes end up in endosomes, and escape from these endosomes is a significant rate-limiting step in the transfection process[92]. If the nucleic acid to be delivered is DNA, it must then enter the nucleus of the cell, which is yet another barrier to successful transfection. This means the DNA must pass another lipid bilayer: the nuclear envelope. Many plasmids designed for transfection have sequences known as nuclear localization signals, which help the plasmid enter the nucleus[93]. Often these sequences first bind proteins, and the resulting protein-nucleic acid complex is then transported into the nucleus via nuclear pore proteins. For DNA-based transfection, the state of the cell is important; actively mitotic cells are more readily transfected than cells not dividing. This is thought to be due to the fact that mitotic cells have partially permeable nuclei, which are easier to enter than intact nuclei[94].

Depending on its sequence, a plasmid that enters the nucleus can either become integrated into the genome of the transfected cell, or it can remain a distinct entity outside the genome. Plasmids that do not incorporate into the host cell genome are known as “transient” plasmids, as they do not replicate with the host genome. Over time, the protein expression of cells transiently transfected will decrease, as any daughter cells that originate from the transfected cell will not contain the plasmid. Plasmids that incorporate into the host-cell genome result in a stable transfection that persists across generations. This type of stable transfection is convenient, as there is no time constraint on protein expression. However, the incorporation into the genome is somewhat random, and the insertion of the gene can be oncogenic[95].

mRNA transfections

Unlike dsDNA transfections, transfection complexes that involve mRNA do not need to enter into the nucleus. Rather, the mRNA that leaves the endosome is exposed to cytoplasmic ribosomes and is directly translated. The exact structure of RNA lipoplexes is less well known than that of DNA lipoplexes due to the fact that RNA does not have a fixed structure in solution. The lamellar and inverted-hexagonal phases described for DNA lipoplexes form in part due to the rigidity of the DNA molecules used in transfection. Lipid-RNA complexes have not been shown to have similar structures to lipid-DNA complexes, and this could be due to the inherent flexibility and effective branching of the mRNA molecules

Low replicon-number transfections:

Traditional transfections are done with very large concentrations of RNA. For every million cells, it is suggested that one transfect 2µg of mRNA, which corresponds to roughly 10^12 molecules of RNA or 1.0^ RNA molecules per cell[96]. The reason for these high numbers is that transfection is actually a low efficiency process. Even though millions of RNA are supposedly transfected, the average number that are actually translated within each cell is thought to be (see discussion below) in the hundreds to thousands. Because the significant majority of RNA molecules delivered with the CCS+ construct would be replication incompetent, the ability of low numbers of replicon molecules to establish a full, persistent infection needed to be assessed. Assaying the accumulation of actual CCS-containing replicons was considered to be too time consuming, so an assay was developed to simulate the conditions of CCS+ transfections. We chose to use NodaEYFP replicon, as it was easy to follow the course of the pseudoinfection via cellular fluorescence, to simulate these low-copy number transfection conditions.

Reports indicate that the average number of RNA molecules in a single lipoplex, for the usual transfection stoichiometry, is around 800 molecules[97]. In a Nod(+) transfection this would mean that only 250 active molecules would be present per lipoplex, with the majority of the RNA being composed of 2 species. While these numbers are greater than the starting concentration in a natural infection, i.e. 1, the attenuated replication of the modified T2A replicon owing to the lack of the B2 protein may prevent active pseudoinfection from occurring in cells when the replicon is delivered in such low copy number.

Further complicating issues was the biological activity of the co-delivered splice products. The 5' splice product, which contains the open reading frame of the replicase protein, would be translated, similar to the full length molecule, resulting in synthesis of the replicase. However, this RNA molecule would lack the critical 3'UTR that is required for recruiting the replicase and subsequent replication, and would thus not contribute to the replication (and splicing), of the full-length molecule. In addition, owing to its shorter length, it would be translated more readily than the-full length molecule. The 3' splicing fragment, however, would contain the 3' UTR needed for replication, but would not code for the replicase. As this fragment would be smaller, it might act as a defective interfering RNA, reducing the overall yield of catalytically active full length molecules. By design, neither of these molecules would contain both of the ribozymes needed for generation of the mature microRNA after replication, and thus neither would yield functional microRNA. When co-delivered with full length molecules, there was concern that the spliced fragment may pose a significant barrier to establishing an active infection.

To assess the ability of these replicons to replicate at lower than optimal concentration, we diluted NodaT2A replicon with a carrier RNA of a similar length. Note that one cannot simply reduce the amount of RNA used in order to generate lipoplexes with fewer RNAs, as this would simply make a smaller total number of lipoplexes with hundreds of RNAs within each. Instead, we dilute our replicon RNA into a pool of non-capped carrier RNA (Figure 14). Because lipoplex formation is thought to be a

function of charge neutralization, we chose uncapped B1 RNA (derived from the genome of BMV) as the carrier molecule, as it has a similar mass to that of our replicon (3.2 knt vs 3.7 knt respectively.). By varying the ratio of the carrier RNA to the NodaT2A replicon, we could tune the amount of active Nodamura RNA delivered to cells, thereby effectively diluting our RNA of interest. This allows us to mimic the conditions seen in the ccs+ transfections. By further diluting our RNA, we can mimic conditions seen in natural infections, which is a single active replicon per complex. We tested a variety of concentration, ranging from 1 molecule of Nodamura replicon per lipoplex to many hundreds (just short of full transfection).

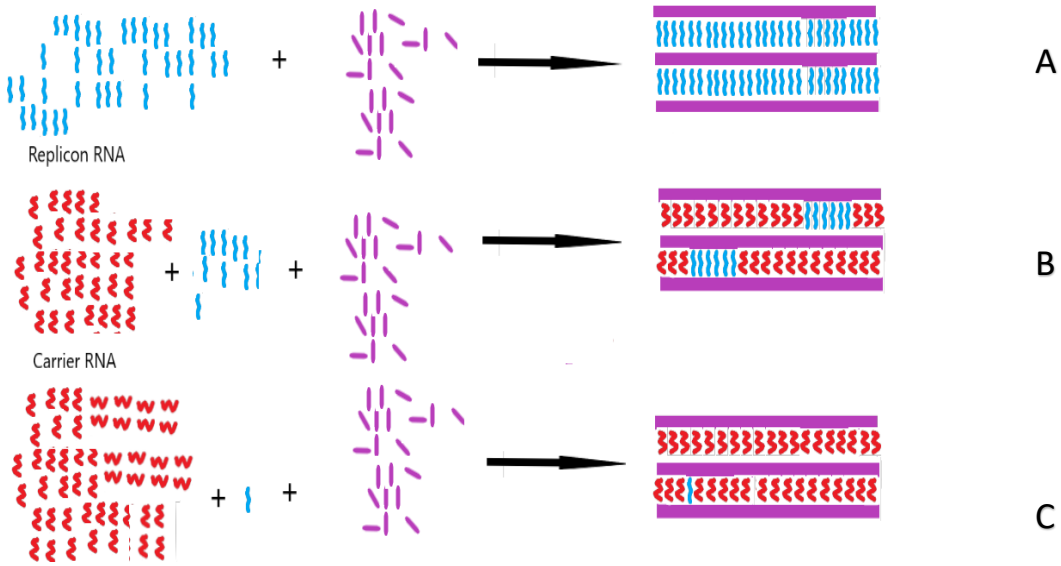


Figure 14: Controlling replicon RNA levels delivered to cells by exploiting stoichiometry of lipoplexes. One can control the ratio of "active" (here, replicon) RNA delivered to cells by adjusting the proportion of "inactive" (here, uncapped B1 RNA) RNA molecules prior to generation of lipoplexes. Lipoplexes formation is dictated by charge neutralization of the cationic lipid and as a result is based on total RNA mass. As such liposomes are formed with set numbers of RNA molecules per lipoplex, and one can take advantage of this to make lipoplexes of varying RNA composition. One can generate lipoplexes with: all one molecule, i.e., only replicon RNA, seen in A, a mixture of

molecules, seen in B, or even make lipoplexes that contain on average only 1 active molecule (C). We “dilute” our replicon molecules with uncapped RNA of a similar length that is not translated in cells.

Methods

BHK-21 cell culturing and transfections

BHK-21 cells were cultured in medium composed of high-glucose DMEM medium with 10% FBS and penicillin/streptomycin. The incubation conditions were 37˚C and 5% CO_2. Standard passaging was performed with the involvement of PBS for washing, and 0.25% trypsin-EDTA for cell detachment. The cells were passaged to 60-90% confluence in 24-well cell-culture-treated plates prior to transfection by electroporation. Prior to transfection, the cells were washed with PBS. For each well, 5 µl of lipofectamine-2000 (ThermoFisher) was diluted with 45 µl Opti-Mem medium (Thermo Fisher) before mixing with 0.5 µg of RNA that is also diluted to an equal volume in Opti-Mem. After a 5-minute incubation at room temperature, the lipofectamine/mRNA mixture is brought up to a volume equaling 200 µl per well. Masses of 12.5 ng of luciferase replicon(3900 nt) mRNA and 3 ng of luciferase mRNA (1200 nt)were diluted with up to 0.5 µg of uncapped carrier RNA (pTRI-XEF [1400 nt] from mMESSAGE kit and pB1 [3200 nt] from Gelbart lab) of similar length, in order to achieve a constant number of liposomes per transfection.

Luciferase assay

A luciferase assay was performed using the Renilla Luciferase Assay System from Promega, and activity was measured using a Moonlight 2010 luminometer. Cells were washed with PBS before directly

lysing using passive lysis buffer and a cell scraper. 100 µl of Renilla luciferase assay reagent was added to a polycarbonate tube and luminescence was measured with 10 seconds of integration and a 2-second delay directly after adding 20 µl of cell lysate and mixing. Luciferase assays were performed in biological and assay duplicates for each time point.

In vitro transcription of RNAs

Plasmids for both the EYFP replicon and B1 sequences were linearized using XbaI and purified using QIAprep Spin Miniprep Kit. The linearized EYFP replicon and Luciferase plasmid was used as a template for *in vitro* transcription using an mMESSAGE mMACHINE T7 Transcription Kit, while the linearized B1 plasmid was *in vitro* transcribed using a MEGAscript™ T7 Transcription Kit. The luciferase mRNA was transcribed with a mmessage mmachine Sp6 plasmid. The RNA transcripts were then purified using a Qiagen RNeasy Mini Kit.

Transfections

Baby hamster kidney cells (BHK-21) were seeded to 70-90% confluence in two 24-well plates at 37˚C. Various amounts of EYFP RNA (0.21, 0.63, 3.13, 6.25, 15.63, 31.25, 46.88, 62.5, 125, 187.5, 250, and 500 ng) were mixed with B1 RNA, such that the total amount of RNA delivered to the cells was 500 ng. This RNA mixture was diluted with Opti-MEM® Medium and mixed in a 1:1 ratio with diluted Lipofectamine® 2000 reagent and incubated for 5 minutes at room temperature to allow the formation of RNA-lipid complexes. The RNA-lipid complexes were then added to the two plates of BHK-21 cells, after which one plate was returned to a 37˚C incubator while the other was placed in a 30˚C incubator. After eight hours, the lipofectamine solution was aspirated and cells were put back into DMEM media.

Fluorescence microscopy

Both bright-field and EYFP fluorescence was visualized using an Amscope Epi-Fluorescent Inverted Microscope with an exposure time of 500 ms; 500 images were taken at various time points post-transfection: 4, 8, 12, 24, 36, 48, and 72 hours.

Results

Dilution Name	EYFP RNA/well (ng)	Carrier RNA/well (ng)
1 replicon RNA/cell	0.21	499.79
1 replicon RNA/Lipoplex	0.63	499.38
5 replicons/Lipoplex	3.13	496.88
10 replicons /Lipoplex	6.25	493.75
25 replicons /Lipoplex	15.63	484.38
50 replicons /Lipoplex	31.25	468.75
75 replicons /Lipoplex	46.88	453.13
100 replicons /Lipoplex	62.50	437.50
200 replicons /Lipoplex	125.00	375.00
300 replicons /Lipoplex	187.50	312.50
400 replicons /Lipoplex	250.00	250.00

800 replicons /lipoplex	500.00	0.00

Table 2: Table of EYFP Dilutions and their corresponding copy number/liposome. In order to deliver small numbers of replicon molecules to mammalian cells, without changing the total number of lipoplexes delivered, replicon RNA was "diluted" with a carrier RNA that was uncapped and therefore unable to be translated. The value assumed for 1 replicon/cell assume that on average 3 lipoplexes enter a cell. These values were chosen to ensure a certain number of replicon molecules were present in each lipoplex. The transfection corresponding to 500ng denotes the manufacturer's specifications for RNA transfections.

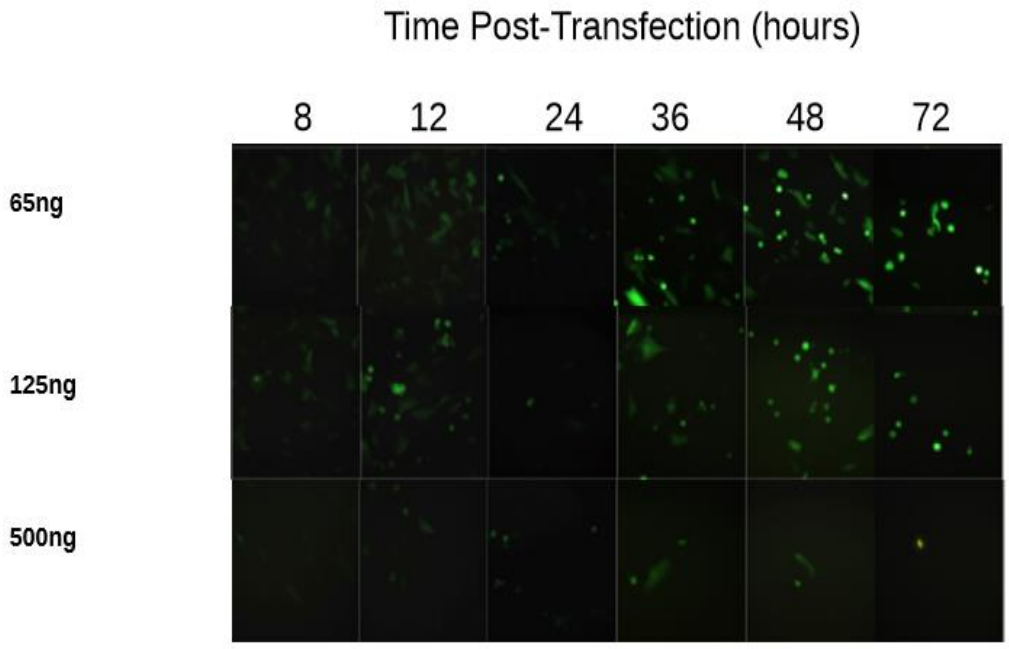


Figure 15: Representative EYFP fluorescent images for several dilutions of Nodamura-EYFP at 37˚C Nodamura-EYFP replicon and B1 RNA were co-delivered to BHK-21 cells in several different ratios using lipofectamine 2000. The cells were incubated with lipoplexes for 18 hours, after which the transfection media was replaced with DMEM supplemented with 10% FBS. At successive time points post-transfection, the cells were imaged using an Amscope AM-220 Epi-fluorescent microscope in the green-fluorescence setting. All images were taken at 500 ms exposure time. Initial translation occurs at short times (4-8 hours post-transfection) and translation coupled to

replication occurs at later time points (24-48 hours post transfection.) Late-stage replication is marked by higher overall protein expression, seen here as brighter cells. Late-stage transfected cells also have changes in their morphologies, with many infected cells become round and semi-adherent.

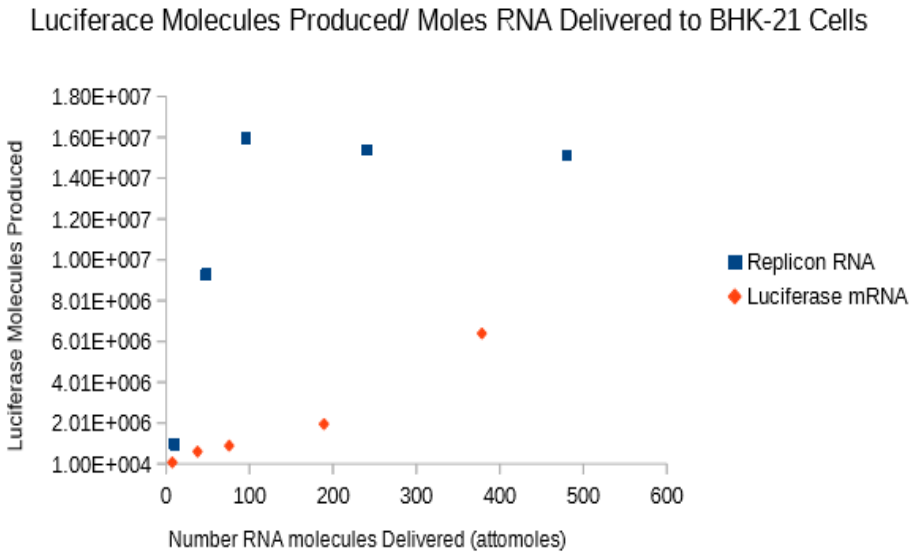


Figure 16: Luciferase replicon/mRNA dilution comparison. Luciferase replicon RNA and luciferase mRNA were transfected at varying dilutions in BHK-21 cells. After transfection, the amount of luciferase produced per 1,000,000 cells was determined via a luminometer. Total amount of RNA delivered to cells was kept constant, with varying amounts of "active" RNA. The luciferase replicon RNA was diluted with 3200 nt B1 RNA, while the mRNA was diluted with a 1.4 knt RNA coding for the xenopus elongation factor gene; both RNAs were uncapped to prevent translation. The luciferase mRNA shows the expected monotonic increase in protein expression as a function of RNA concentration. The replicon RNA shows a sharp response to input concentration, peaking at 100 attomoles then slowly declining.

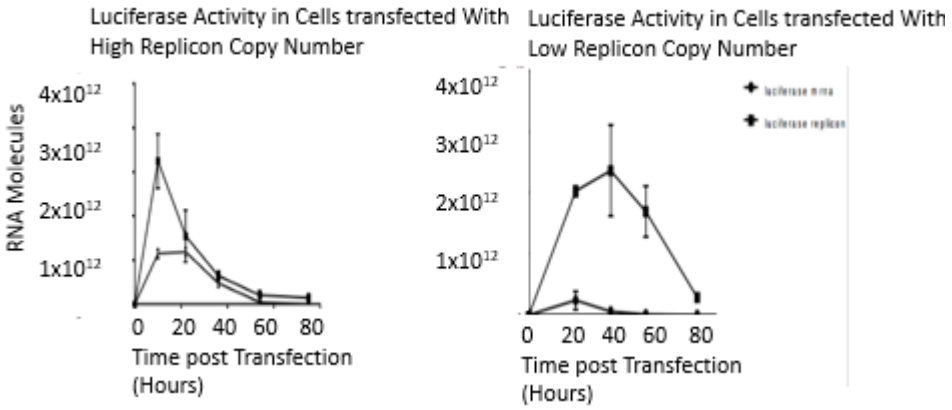


Figure 17: Time-course Comparison between high-concentration luciferase and low-concentration luciferase replicon RNA and luciferase mRNA. Higher concentrations of Nodamura RNA alter the nature of replicon replication. Delivery of large numbers of RNA replicons (125ng) (A) attenuates protein production and shortens the duration of replication. At 48 hours, there is little difference between the curve generated with mRNA and that of the replicon. Higher transfection concentrations also shift the maximum expression to earlier time-points. At 125ng RNA, the peak of expression is around 18 hours' post-transfection, while Lower concentrations (B) show maximum expression at ~36 hours,

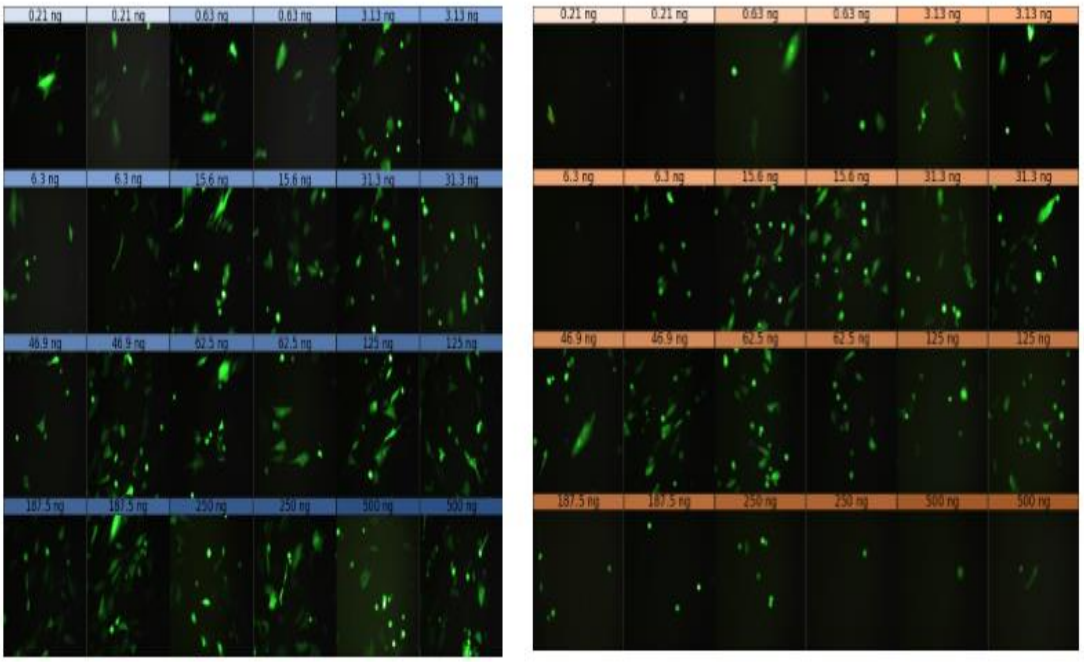


Figure 18: Temperature Effects On Nodamura-EYFP Replication. Various dilutions of Nodamura-EYFP replicon were made using the above protocol and delivered to cells at 37˚C and 30˚C (blue and orange respectively.) At various time-points, these cells were imaged using an Amscope AM-220 Epi-fluorescent microscope in the green-fluorescence setting. All images were taken at 500 ms exposure times. The above image shows these cells after18 hours post-transfection. A full series of images can be seen in the supplemental figures of this paper. The cells incubated at 30˚C show a larger number of transfected cells, and the transfected cells overall are much brighter than those left at 37˚C. In addition, there is less overall change in the morphology than at 37˚C. The inverse-dose dependence seen in BHK-21 cells at 37˚C is not seen in cells at 30˚C; increasing concentrations of replicon does not attenuate protein expression.

EYFP Transfections at 37˚C

Transfections with varying concentration of EYFP replicon show attenuated fluorescence at high replicon concentrations (Figure 15.). Cell transfected with as little as 1 RNA per lipoplex show fluorescence peaking at around 36 hours post transfection, which is indicative of active RNA replication. At early time points (8 and 12 hr.), dilutions that correspond to RNA numbers greater than 25 RNA/lipoplex show some low levels of fluorescence, lower than the fluorescence intensity of the cells during the height of replication, which corresponds to the initial translation of the T2A replicon. The number and intensity of fluorescent cells is relatively uniform for most of the dilutions imaged, indicating that the levels of initial translation are similar across dilutions. However, cells transfected with the higher concentration of replicon RNA show less overall fluorescence at early time points, and this trend continues throughout the duration of infection. Maximum overall EYFP production appears to occur around 50-75 ng replicon RNA delivered, which corresponds to 50-100 replicon RNAs per lipoplex.

The overall *duration* of the infection is also affected by dilution of the replicon RNA; cells transfected with lower amounts of replicon RNA show an infection that persists past 72 hours, while cells challenged with higher RNA levels show far fewer fluorescent cells at this time point. At 96 hours, cells exposed to a higher concentration (125 ng, 500ng) of replicon RNA have almost no detectable EYFP, while cells challenged with less RNA are still actively replicating RNA.

Cell morphology of transfected BHKs was also affected by the amount of RNA initially delivered to cells. As more replicon RNA was delivered to cells, the number of round, semi-adherent cells increased at early time-points. These cells do not stain with trypan-blue, so they are not dead cells, and yet they are noticeably different in shape and size compared to naïve BHK-21 cells. While these round cells appear to be more fluorescent in general, this is a consequence of the fact that they are round and not flat, and as such have a smaller cross-section when imaged, which artificially enhances their fluorescence intensity.

Comparisons between luciferase replicon and luciferase mRNA

Dose-response curves for Luciferase replicon and luciferase mRNA have different overall shapes as seen in Figure 16. Dilutions of luciferase mRNA show the expected response. Decreasing concentrations of active mRNA in the lipoplexes resulted in linearly decreasing luciferase expression levels across several orders of magnitude after 24 hours. The replicon RNA, however, showed different behavior. As the amount of replicon RNA delivered to the cells increased, the response was decidedly non-linear. Low amounts of replicon RNA (10-50 attomoles) show a much stronger expression profile compared to luciferase mRNA, indicating that the replicon is functioning to increase the overall amount of protein expressed. This trend continues until about 100 attomoles of delivered RNA, at which point the cells no longer produce more luciferase as a function of input RNA concentration. Indeed, after 100

attomoles, there is an overall decrease in luciferase protein production.

Effect of initial transfection concentration on luciferase expression during pseudoinfection.

Varying the initial transfection concentration alters the overall protein expression profile of the Nodamura replicon as seen in Figure 17. Transfections of 125 ng of Luciferase replicon showed a diminished response compared to transfections of 12.5 ng RNA. While the higher concentration of Noda RNA showed a higher luciferase expression at 12 hrs., it rapidly declined to levels similar to that of the mRNA. Transfections with an order of magnitude less RNA (12.5 ng) results in a more persistent infection. Levels of protein expression for this dilution reach values similar to that of the higher concentration, but the levels do not decline overall

Temperature effects on Nodamura-EYFP

Cells incubated with different amounts of NodaEYFP replicon at lower temperatures showed different behavior than those at 37˚C. The inverse-dose dependence seen at 37˚C vanishes when the cells are transfected and held at 31˚C. Figure 18 shows the acquired fluorescence micrographs taken at 37˚C and 31˚C, and shows that cells at 31˚C do not have attenuated EYFP fluorescence upon increase in dose. This is in stark contrast to the cells at 37˚C, which show a clear peak in fluorescence at approximately 65 ng replicon RNA delivered. Overall, cells transfected at lower temperatures had more fluorescent cells, and those fluorescent cells had fewer structural changes than those at 37˚C. There is a marked decrease in the number of spherical cells, and the duration of the pseudoinfection is much greater at 30˚C; many cells remained fluorescent even after 90 hours post-transfection.

Discussion

Lipofectamine-based transfections with mRNA are quite common in cell biology, but little has been reported about the optimum amount of RNA to deliver using these techniques. For standard mRNA there is little reason to believe that increases in the amount of mRNA should do anything but increase the amount of protein produced by the cell. Indeed, this trend is seen in Figure 16, as the mRNA transfections increase monotonically with increasing amounts of mRNA delivered. The fact that the NodaLuciferase replicon exhibits different dynamics indicates that there is a fundamentally different biological process occurring. One source of non-linearity can be found in saturating ribosomes; there are a limited number of available ribosomes, and by providing very high numbers of RNA molecules, one could theoretically saturate the ribosomes. At this point, additional RNA molecules would not increase the amount of EYFP or luciferase present at 24 hours, as there would not be ribosomes available for translation. At least initially, the amount of RNA delivered to cells is orders of magnitude less that the amount needed to saturate the ribosomes. Interestingly, qPCR measurements (see below Chapter 4) indicate that during pseudoinfection the replicon does indeed replicate to high copy number (up to millions of copies) and may actually be saturating ribosomes, which could result in the pattern seen in Figure 16.

However, this simple hypothesis does not take into account the fact that this dilution effect is not seen at lower temperatures. Cells incubated at 31 C showed no sensitivity to higher levels of delivered replicon RNA, and the duration of the pseudoinfection was unchanged by input concentration. Here it is important to remember that by design the replicon yields a single copy of the replicase protein for every reported protein synthesized. High levels of luciferase indicates high levels of replicase protein, which has cytopathic properties. Protein A is capable of remodeling internal membranes, and in particular is associated with active remodeling of mitochondria *in vivo*. As such, delivery of large

amounts of replicon molecules may result in the rapid over-expression of protein A, which may trigger the cells apoptotic/immune pathways. In this manner, delivering high amounts of Noda RNA may trigger the cell to undergo phenotypic changes in response to the high levels of protein A being produced; this may be the source of the larger number of spherical cells seen at 37˚C. In addition, this supports the findings of figure 17, which shows that higher concentrations of NodaT2A replicon can change the overall shape of the pseudoinfection time-course plots. Here we see the pseudoinfection has been reduced in intensity and duration by higher amounts of input RNA, which could be triggering the cells to change their respective phenotypes. Protein expression is attenuated at lower temperatures, and this may be the source of the differences in response to different temperature. Lower temperatures mean less translation at early time points, so that fewer cells overall will produce enough protein A to trigger the innate cellular responses.

The lower temperature may also change either the structure of the replicase protein or the RNA itself to make it more amenable to replication. At lower temperatures, the 3' UTR of the replicon molecule adapts a tRNA-like fold that may help the RNA to be translated, and may help the RNA evade the cells own immune system and RNAi. In addition, it is known that the one can replace the C-termini of the replicase protein with GFP and retain much of the replicase's function, which indicates that the protein is somewhat flexible and may have different conformation at different temperatures.

This trend is not seen in all cells lines, and the replicon in general varies in its ability to replicate in cells (data not shown.) In PC3 cells (human prostate cancer cells), there is no inverse-dose dependence: instead, higher concentrations of delivered EYFP replicon yield more fluorescent cells. PC3 cells are known to be resistant to VSV infections, and it is thought that they have similar resistances to other +ssRNA viruses. At 37˚C, PC3 cells are unaffected by even large dosages of NodaEYFP, and are only actively infected when they are transfected at 31˚C.

Clearly there is a complex regulatory network at play that results in the behavior seen here. High

initial translation of the replicon in BHK cells seemingly triggers the cell's own immune response, attenuating the pseudoinfection. Nodamura has evolved such that it only needs one copy of its genome to enter a cell in order to infect it, and perhaps the interactions between the RNA and its polymerase are modified in some way by having many copies of the RNA present at early times during the infection.

Chapter 4: Cell Culture Studies of the NodaMiRNA Replicon

Background

RT-qPCR for detecting RNAs

Because both Nod (+) and Nod (-) lack a reporter gene and have putative biological activity without being translated, it was important to have an assay that accurately measures the absolute number of both replicon RNA as well as miRNA released via the activity of the ribozymes. While the replication of replicon RNA is sufficient to make the band clearly visible on a gel of total RNA extracted from the transfected cell, we wanted a more quantitative method to measure the amount of RNA present in cells at different time points. Reverse-transcription quantitative polymerase chain reaction (RT-qPCR) is the most accurate technique used to measure the amount of RNA in a sample.

RT-qPCR is a highly sensitive technique that utilizes PCR to selectively amplify selected regions of RNA in a sample. RT-qPCR can be either relative or absolute. In *Relative* RT-qPCR measures the levels of selected RNA are compared to a reference gene of interest. Often these reference genes are proteins that are consistently expressed in cells, such as actin or GADPH. *Absolute* RT-qPCR uses a calibration curve generated using known amounts of input RNA. This curve can then be used to measure absolute RNA levels in biological samples.

RT-qPCR can fundamentally be broken into two steps: reverse transcription and qPCR. Commercial 1-step kits are available that allow one to perform both the reverse transcription and qPCR in in one tube, but these are not particularly accurate. The initial reverse transcription reaction uses a DNA primer that is a reverse complement to the RNA sequence to be probed (See Figure 19). The primer hybridizes to the RNA, and a reverse transcriptase, in our case Maloney Murine Leukemia Virus (MMLV) reverse transcriptase, synthesizes a single-stranded DNA strand that is complementary to the RNA

sequence. The region to be detected, known as the amplicon, is roughly 70-300 nt in length, and should have limited secondary structure because secondary structure of the amplicon region greatly affects the initial binding of the DNA primer to the RNA.

We initially planned to probe the 3' UTR of our replicon with RT-qPCR. We chose this region as it would accurately reflect the concentrations of both full-length as well as subgenomic RNA molecules. Multiple attempts were made to generate qPCR conditions that would allow for highly efficient amplification of this region, yet every time the amplification was low efficiency. We suspect that the highly structured RNA at the 3' end of the RNA prevented binding of the primer and subsequent reverse transcription. These considerations are shown in Figure 20, and based on them we chose a region of RNA that was predicted by mfold to be relatively unstructured. This region lies within the T2A coding region, and proved to have the high RT-qPCR efficiencies needed for absolute quantification. It also was sufficiently downstream of the subgenomic region such that it detected all strands transcribed.

After the cDNA is synthesized, it is detected by measuring the amplification of the amplicon region via traditional PCR protocol. The amount of dsDNA (which corresponds to the amount of amplicon DNA produced) is measured by the increase in fluorescence of an intercalating dye, in this case SYBR Green (See Figure 19.) As the amount of DNA present in the tube increases, the increase in fluorescence of this molecule is measured via the optical thermocycler, resulting in a sigmoidal curve seen in Figure 19. Once this fluorescence value reaches a critical threshold level (defined by the cycle number where the signal from the dye is roughly 10X background fluorescence), the thermocycler calculates a "Ct" value, which is representative of the amount of input amplicon and thus the amount of RNA in the initial sample. By using *in vitro* transcribed RNA, or ssDNA oligos, one can generate a calibration curve that converts RNA concentration to Ct values, which allows for rapid and facile determination of absolute RNA concentration for samples purified from cells.

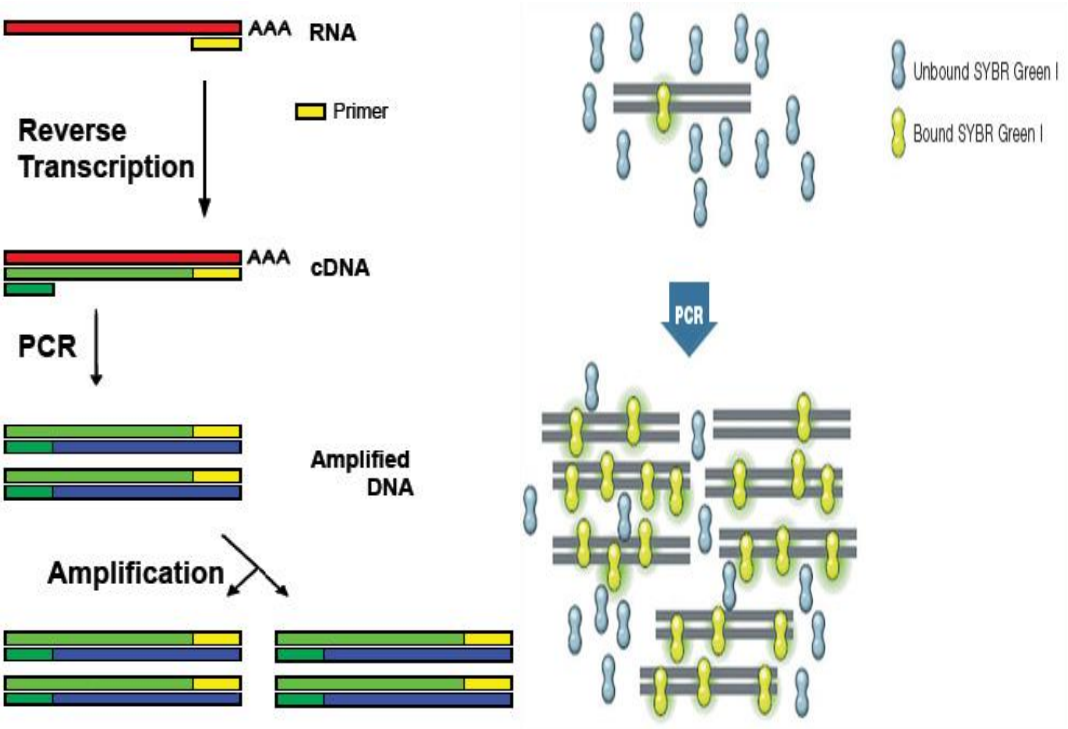


Figure 19: Reverse Transcription Quantitative PCR is used to directly measure RNA levels in a sample. It proceeds via two steps. Initially, a DNA primer is annealed to the RNA of interest. A reverse transcriptase then binds to the DNA/RNA hybrid and generates a single-stranded DNA that is a reverse complement of the RNA. This DNA is used for the qPCR reaction. In the qPCR step, a region of this cDNA is amplified using a different set of primers that anneal to a region known as the amplicon. This region is usually 70-200 nt in length. As the region is amplified, the total amount of dsDNA increases exponentially. The amount of dsDNA is measured by the binding of an intercalating dye – in our case, SYBR Green – which becomes fluorescent only when it is intercalated.

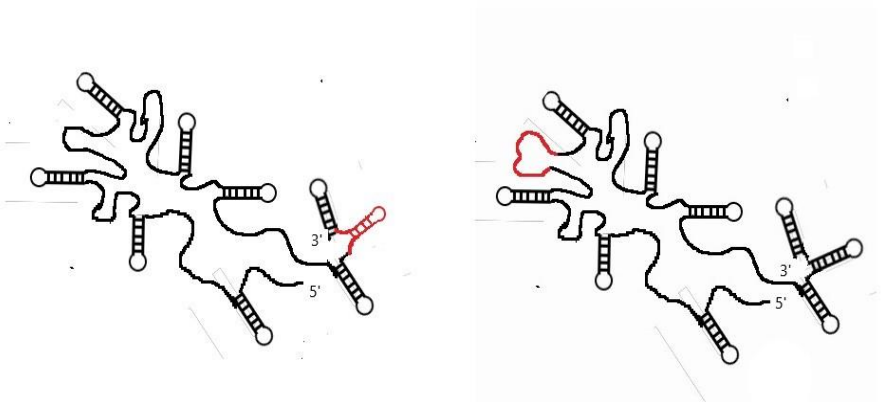


Figure 20. RNA secondary structure greatly affects the efficiency of RT-PCR reactions. Attempting to probe a region (seen in red) is difficult if the RNA molecule has high degrees of secondary structure, as seen in image **A**. This is seen in the NodaT2A replicon, which has a highly ordered 3' UTR. One should choose a region of RNA that is not predicted to be structured by folding algorithms, as seen in **B**. This will ensure the qPCR reaction is high efficiency, and will provide truly quantitative data.

MicroRNA Detection using RT-qPCR

Because mature miRNAs are only 20-25 nt long, they cannot be detected using traditional RT-qPCR techniques. They are roughly the same length as the DNA primer used during the initial reverse transcription reaction, and the MMLV polymerase is not able to generate a cDNA strand due to the short length of the microRNA molecule. The microRNA must be lengthened significantly in order to be detected.

To detect a microRNA, one first polyadenylates it using polyA polymerase, which adds 20-80 adenosine nucleotides to the 3' end of the microRNA. This step requires an available 3' hydroxyl group, and ensures that the 3' end of the microRNA is properly terminated, which indicates the ribozymes are splicing as intended and that the microRNA is biologically active. After polyadenylation, the RNA is

hybridized to a 40-60nt adapter molecule that has a unique 3’ end and has many thymines present on its 5’ end. This ensures that the adapter molecule hybridizes to the polyA tail of the polyadenylated microRNA. After annealing the adapter, a cDNA strand is synthesized using MMLV polymerase, and this results in a 60-80 nt cDNA strand that can have detected via qPCR with primers designed against the microRNA of interest and the adapter molecule used for hybridization[98]. This method is illustrated in Figure 21.

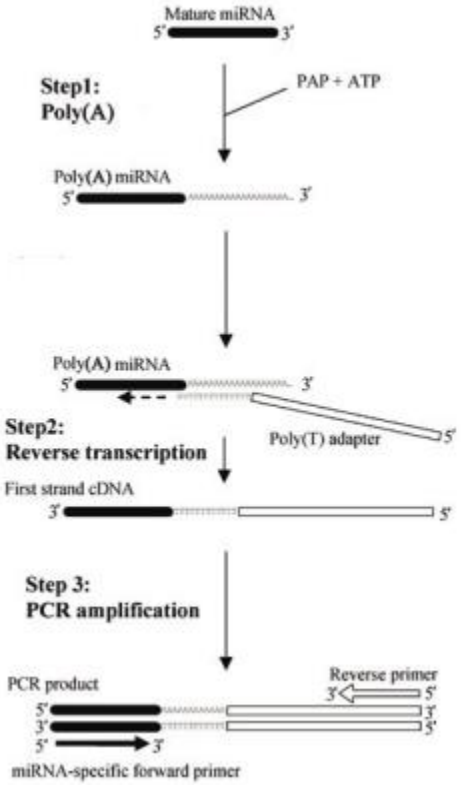


Figure 21: MicroRNA detection via miRNA polyadenylation RT-qPCR. MicroRNAs are too small to detect using traditional RT-qPCR methods, as they are the same length of the primer used in detection. To detect these molecules, we first polyadenylated the RNA using polyA polymerase (PAP); the resulting RNA is then reverse transcribed using an adapter molecule that hybridizes to the newly adenylated molecule. The resulting cDNA is then detected using standard qPCR techniques.

Melting curves

qPCR analysis results in the accumulation of double-stranded nucleic acid that corresponds to the RNA region that was amplified. In an ideal reaction, there is only one amplicon region and the double-stranded DNA present is relatively monodisperse. One convenient and common way to measure the polydispersity of a dsDNA sample is to run a melt-curve analysis on the DNA. In qPCR, the dsDNA is detected by an intercalating nucleic acid stain that is only fluorescent when it is bound to double-stranded DNA and this fact can be used to determine the polydispersity of the DNA present. Double-stranded DNA is not stable at high temperature, and denatures into single-stranded DNA at temperatures ranging from 65-95C. A melting-curve can be generated from dsDNA by incrementally increasing the temperature of a DNA sample and measuring the resulting change in fluorescent that accompanies the DNA phase transition. By plotting the negative derivative of the response curve, one obtains plots which contain peaks that represent unique dsDNA species in solution. A mono-disperse sample will yield a sharp curve at the melting point of the amplicon DNA. The presence of heterogeneity in the sample will appear as a broader peak or even two peaks. This type of analysis is very useful when determining whether the qPCR analysis is efficiently amplifying only one amplicon region. Alternatively, it can be used to analyze alternative splice variants in ribozyme splicing.

Methods

Transfection of BHK-21 cells with CCS+ and CCS – transcripts

BHK-21 cells were split 24 hours prior to transfections, and grown to approximately 70% confluence in 6-well plates. Cells were transfected with 2µg of RNA per well using lipofectamine, according to manufacturer's specification, using 15 2µL of lipofectamine per well. The resulting

transfection mixture was overlaid onto the confluent BHK cells, and the cells were incubated for 8 hours at 37˚C, unless otherwise indicated. After 8 hours of transfection, the transfecting media was removed from the cells, the cells were washed with 2mL warmed PBS, and overlaid with DMEM supplemented with 5% FBS. The medium was left on top of the cells for the remainder of the experiment. Micrographs were taken during transfection to assess the overall health of the cells, using a Amscope Epi-fluorescent microscope with bright-field illumination using the 10 phase contrast setting.

Harvesting and analysis of total cellular RNA.

At successive time points (8, 18, 24, 48, and 72 hrs.), the total cellular RNAs were harvested from transfected cells using a Qiagen RNeasy column, according to manufacturer's specification. RNA was eluted from columns using 50uL of RNase free water, and the concentration of RNA was determined by UV absorption using a nanodrop spectrophotomer. The purified RNAs, containing rRNA, cellular mRNA, and full length and spliced Nodamura replicon RNA were stored (how?) stored at -80C until needed for downstream analysis. Total cellular RNAs were purified from naïve BHK-21 cells as well, using the techniques described above.

Gel electrophoresis.

Either 1 or 5µg of total cellular RNA from each time point was denatured in an equal volume of 2x NEB RNA denaturing buffer containing 96% formamide. The RNA was denatured at 65˚C for 5 minutes, loaded onto a 1.2% agarose gel, and run at 75 V for 2 hours. The gel was then stained with GelRed high sensitivity nucleic acid stain for 30 minutes, prior to imaging.

Cellular mRNAs are difficult to visualize in traditional gel electrophoresis, given their overall low concentration. In addition, it is difficult to identify a particular RNA in gels stained with simple intercalating dyes. But both the spliced and full length Nodamura replicon genomes are of lengths that are not seen in naïve (i.e., untransfected) cells, so electrophoresis is an excellent method for analysis. In addition, it was assumed that they would amplify to sufficient amount to be detected, with both the 5' and 3 ' splice products visible as well.

Generation of positive and negative-strand calibration curves.

In order to accurately quantify the total number of positive, and negative strands generated during replicon replication, a calibration curve for each strand was generated to measure the total number of each RNA from the total purified cellular RNA.

For the positive-strand calibration curve, 1ug of in-vitro transcribed Nod(+) RNA was reverse transcribed using NEB MMLV reverse transcriptase per manufacturer's specification. The RNA region chosen for detection was the 140 nt region spanning from 40 nt upstream of the T2A peptide, through the T2A coding region. This region was chosen due to the low amount of RNA secondary structure, as predicted by mfold. Previous attempts to analyze the 3' UTR of the replicon proved untenable, owing perhaps to the large amount of secondary structure present in the untranslated region. A primer annealing to the 3' end of the amplicon region was used for all reverse transcription reactions. After reverse transcription, the reverse transcriptase was heat denatured at 65˚C for 20 minutes, and a series of dilutions was generated. These dilutions were generated by serially diluting the reverse transcription mixture, at 50 ng/uL, with ddH20, yielding the following concentrations: 1ng, 100pg, 10pg, 1 pg, 100fg, and 10 fg/uL. One µL of each of these dilutions was used in the creation of the calibration plot. The

calibration plot was generated using a Pharos Opticon fluorescent thermocycler. Each calibration point was done in triplicate, and consisted of 1µLof the relevant concentration, 10µLof Bio-Rad SsoAdvanced qPCR Master Mix, 200 nM forward and reverse primer, and water to 20 uL. These samples were loaded onto a 96-well, optically-clear, plate, sealed with transparent film and analyzed. Polymerase chain reaction was performed with an extension temperature of 60C for 30 seconds, during which time the fluorescent emission of the SYBR green dye was measured. This was repeated for 35 cycles, yielding the classic sigmoidal shapes associated with qPCR. Ct values corresponding to each dilution in the series was measured using the Opticon software, and exported for analyses. The Ct values for each of the 6 dilutions were averaged and plotted against the log of the input RNA, yielding the linear curve seen in Figure 24. This curve was used to quantify the number of Nodamura molecules present in the total cellular RNA samples taken at various time points post-transfection.

The calibration curve for the minus strand quantification was generated via qPCR performed on a 100 nt oligonucleotide purchased from IDT, as we did not have a plasmid that allowed for the facile *in vitro* transcription of minus-strand replicon RNA. A region corresponding to the same location in the positive RNA that was used in the above positive-sense quantification scheme (T2A region) was chosen for parity's sake. It should be noted that because we started from the cDNA template, there was no reverse transcription step in the synthesis of the minus-strand calibration curve. This means that the curve does not take into account the efficiency of the reverse transcription reaction, which has been shown to vary from 50-90%. However, given the large amplification factor of this novel RNA, and the tendency for qPCR values to vary up to twofold between biological duplicates, this minor factor was considered to be acceptable. The purchased oligonucleotides were diluted directly to the following concentrations: 1ng, 100pg, 10pg, 1 pg, 100fg, and 10 fg/uL and analyzed as above, yielding the linear curve seen in Figure 25. This curve had an acceptable slope, corresponding to an amplification efficiency of approximately 70%. The curve generated thusly was used to quantify the number of negative-strand

Nodamura molecules present in the total cellular RNA samples taken at various time points post-transfection.

Detection of plus and minus strand RNA from Nod(+) and Nod(-) transfection

Quantification of both plus strands and minus strands resulting from replication of both Nod(+) and Nod(-) from purified total cellular RNA was performed in a protocol described below. Oneμg of total cellular RNA from each time point was reversed transcribed using NEB MMLV reverse transcriptase per manufacturer's specification, using 50uM of the appropriate reverse primer. The reverse transcription reactions were heat denatured at 65 c for 10 minutes, and diluted to the following concentrations: 1ng, 100pg, and 10 pg /uL. Each of these dilutions was analyzed in triplicate using 1μLof the reverse transcription concentration, 10μLof Bio-Rad Ssoadvanced qPCR master mix, 200 nm forward and reverse primer, and water to 20 uL. These samples were loaded onto a 96-well, optically-clear plate, sealed with transparent film and analyzed. Polymerase chain reaction was performed with an extension temperature of 60C for 30 seconds. Ct values corresponding to each dilution in the series were measured using the Opticon software, and exported for analysis. These CT values were converted to concentrations per μg input RNA using the calibration plot

MicroRNA calibration curve and Detection of in vivo spliced miR-34a

Analysis of gels – see Figures 22 and 23 – from both Nod(+) and Nod(-) showed distinct bands running at 3.2 knt, corresponding to the spliced genome. This was a good indication that splicing of one or both ribozymes were occurring during replication. To analyze the efficiency of this splicing, we quantified the amount of miRNA purified from replicon transfected cells.

As explained earlier, owing to the small size of the microRNA, traditional qPCR is impossible as it relies on primers roughly the same size as the microRNA to be detected. We employ the method by Fu et al [98], that uses polyadenylation of the microRNA, followed by reverse transcription using an adapter molecule with a Poly-T annealing section. The reverse transcription product is now an appropriate length for traditional qPCR. A schematic for this technique is shown in Figure 21. The method is ideal for detecting the mature microRNA, as it relies on the microRNA having a 3' hydroxyl, which is present in the microRNA naturally. This hydroxyl is present in the miRNA generated via the splicing of the 3' intron, and would not be present if we used a different ribozyme for the excision.

To generate a calibration plot, 900 ng of Mi34A was first polyadenylated using 20U of NEB polyA-polymerase supplemented with 10mM ATP in MMLV reverse transcriptase buffer (NEB.) This reaction was run at 37 C for 1 hr. To this mixture, adapter molecule was added to a concentration of 50 µM, and the mixture was diluted with water to a final volume of 30uL. The resulting mixture was heated to 65˚C then cooled on ice for 5 minutes. Additional MMLV buffer (2uL) and MMLV reverse transcriptase (2uL) was added to the tube, as well as water, to a final volume of 40 uL. The reaction was run a 42c before being diluted to the following concentrations: 1ng, 100pg, 10pg, 1 pg, 100fg, and 10 fg/uL. OneµLof each of these dilutions was used in the creation of the calibration plot. The calibration plot was generated using a Pharos Opticon fluorescent thermocycler. Each calibration point was done in triplicate, and consisted of 1µLof the relevant concentration, 10µLof Bio-Rad Ssoadvanced qPCR master mix, 200 nM forward and reverse primer, and water to 20 uL. These samples were loaded onto a 96-well, optically-clear, plate, sealed with transparent film and analyzed. Polymerase chain reaction was performed with an extension temperature of 59˚C for 30 seconds, during which time the fluorescent emission of the SYBR green dye was measured. The CT values for each of the 6 dilutions were average and plotted against the log of the input RNA, yielding the linear curve seen in Figure 28. This curve was used to quantify the number of microRNA molecules present in the total cellular RNA samples taken at

various time points post-transfection. The number of the molecules gives a measures the combined splicing efficiencies of both ribozymes.

Collection and analysis of small RNAs purified from BHK-21 cells transfected with Nod(+) and Nod(-) replicons

To detect the presence of microRNAs in the cells transfected with the NodaMiRNA replicons, we first isolated the small RNA fraction of cells. This was done using a MiRVana RNA purification column following manufacturer's specifications. These columns allow for the purification of only small (<100nt) RNA fragments from cells. This helps enrich the sample for microRNA for detection, if need be, and ensures a limited background for downstream qPCR. Because of the size-selection, there is no detection of partially spliced RNA molecules, as they are too long to be collected. By collecting only, the small-RNA fraction, we ensure that only fully spliced 22nt miRNAs are detected by our modified qPCR protocol.

Results

Qualitative studies of NodamuraT2A replicon replication in BHK-21 cells reveals high levels of RNA accumulation.

The accumulation of Nod(+) replicon RNA in transfected BHK-21 cells reaches levels similar to that of ribosomal RNA and far surpasses all other cellular mRNAs. Figure 22 shows the accumulation of replicon RNA in a 1.2 % denaturing agarose gel. The full-length (un-spliced) replicon band (3.9 knt), which runs just a bit faster than the larger rRNA subunit, is clearly visible at 18 hours, and increases in intensity, reaching an apparent peak at around 24 hours. At this time point, the replicon RNA is clearly

visible as an individual band, which is not something traditionally seen in transfections, indicating the large amount of replication present in transfected cells. The smear seen around 500-1500 knt represents all other mRNA present in the cell; as the replicon band is much more intense than this band, it is clear that replicon RNA accumulates to levels much higher than even abundant mRNAs (such as actin or GADPH.) It should also be noted that any minus-strand RNA that accumulates during this time would also be 3.9 knt in length, and is therefore indistinguishable from unspliced Nod(+) RNA when run on a gel. This is because replication proceeds first via the synthesis of the reverse complement of the replicon molecule, which by definition is the same length as the plus-strand genome. However in Nod(+), this strand of the RNA does not contain active ribozymes, and thus does not splice. Further evidence of splicing is clearly visible, as both the 5' and 3' splice products are clearly visible in the gel. In addition to these bands, a band at 1.1 knt is present in both the 18 and 24-hour time points, which represents the subgenomic replication of the RNA which occurs during the replication cycle.

A gel showing the accumulation of Nod(-) RNA is shown in Figure 23. Lanes 1, 3, and 5 shown 2ug total RNA corresponding to 18, 24 and 48 hours' post-transfection, loaded onto a 1.2 % TAE agarose gel. Here one also sees the gradual accumulation of full-length Nod(-) RNA that runs just faster than the 28 S subunit, again with similar intensity to the aforementioned rRNA band. However, unlike the bands seen in the Nod(+) transfection, there is markedly less spliced product present in the sample. As in the Nod(+) gel, one cannot distinguish between unspliced positive and negative strands when they are run on a gel, as both strands of the genome are identical in length. The weaker splicing band seen in the gels obtained from cells transfected with Nod(-) may indicate that the Nod(-) construct splices less efficiently, or that less minus strand RNA is produced overall. The convolution of these two RNAs within a gel provides clear motivation for a more accurate measure of RNA levels in actively transfected cells.

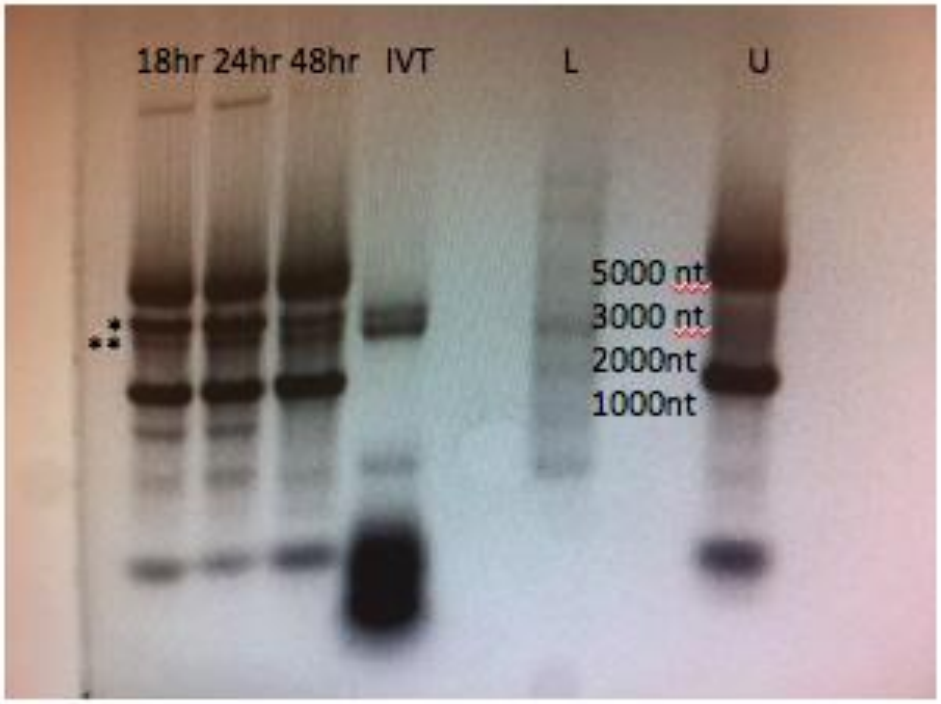


Figure 22: Time Course Gel of Purified Cellular RNAs from BHK-21 Cells Transfected with 2ug of Nod(+) replicon. Total cellular RNAs from approximately 1 million cells were purified using a Qiagen RNeasy column, following manufacturer's instructions at three time-points post-transfection (18, 24 and 48 hours, left-to-right). For each time point, 4µg total RNA was denatured in 50% formamide at 65˚C for 15 minutes, loaded onto a 1.2 % agarose TAE gel, and run for 2 hours. An ssRNA ladder (2ug NEB ssRNA ladder, L), 1ug of in-vitro transcribed Nod(+)(adjacent to 48-hour time point), as well as RNA from untransfected BHK-21 cells (U) were also run as internal standards. Strong amplification of full length Nod(+) RNA is visible in all transfected samples and absent in the naïve RNA (U.) Spliced replicon bands are indicated by **, full-length replicon as *.

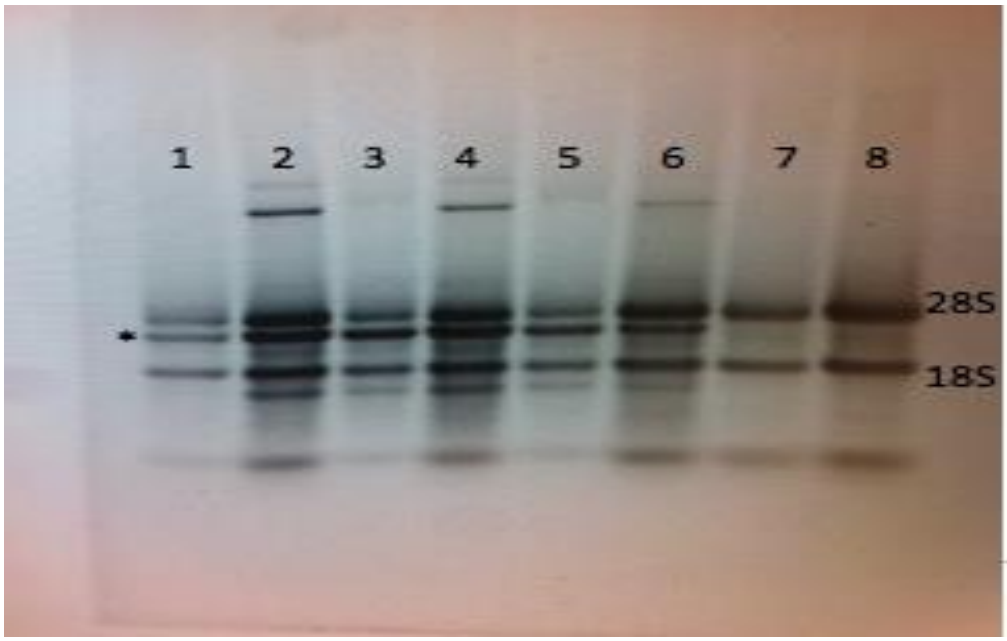


Figure 23: Time Course Gel of Purified Cellular RNAs from BHK-21 Cells Transfected with 2ug of Nod(-) replicon.

Total cellular RNAs from approximately 1 million cells were purified using a Qiagen RNeasy column, following manufacturer's instructions at three time-points post-transfection (18, 24 and 48 hours, left-to-right). Two µg (odd) or four µg (even) total RNA was denatured in 50% formamide at 65˚C for 15 minutes, and loaded onto a 1.2 % agarose TAE gel, and run for 2 hours at 75 V. 28 and 18 S rRNA bands serve as internal references for the Nod(-) RNA, running as sharp bands at roughly 4.2 and 2.8 knt respectively. Strong amplification of full length Nod(-)(*) RNA is visible in all transfected samples and absent in the naïve RNA (rightmost lanes.) Spliced bands present at 3.2 knt are less visible, but still faintly present.

Calibration curves for the detection of NodaMiRNA replicons.

Gels run with total cellular RNA purified from cells transfected with NodaMiRNA replicons showed strong evidence of replication with BHK cells. Because we were interested in restoring homeostatic levels of miRNA to cancerous cells, we were interested in the absolute numbers of both NodaT2A replicon molecules, as well as spliced miRNA levels. These two numbers would help us

understand the degree to which the replicon amplified in cells, and the splicing efficiency of the replicons *in vivo.* Because the replicon molecule is itself a transgene and not present in naïve cells, we could not use qualitative methods to measure the RNA levels. These methods usually compare the RNA levels of two genes: a reference gene (often actin or GADPH) and a gene-of-interest. But this method does not work well for transgenes, and as such we had to be more quantitative in our RNA detection.

To measure absolute levels of RNA present in a sample, one must generate a calibration curve, such as the one seen in Figures 24 and 25, with data obtained by measuring RNA levels in samples where the RNA concentration is already known. This results in a linear plot that can be used to convert from Ct to RNA concentration. These curves should have an ideal slope of -3.3, and should have R^2 values close to unity to ensure accuracy in RNA detection. The slope of -3.3 arises naturally from the fact that DNA amplification is exponential. Both calibration curves generated to detect minus and plus strands have appropriate slopes and regression statistics, and are therefore sufficiently accurate to measure RNA levels from cells transfected with NodaMiRNA replicons.

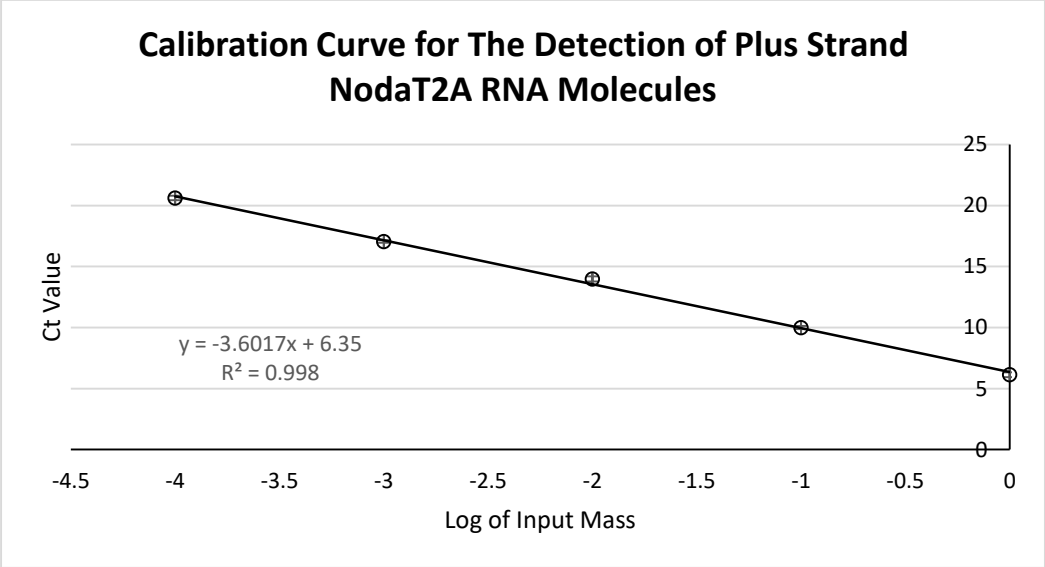


Figure 24: Calibration Curve for Positive-Sense RNA Molecules. One µg of Nod(+) total RNA from an in-vitro transcription reaction was reverse transcribed using NEB MMLV reverse transcriptase following manufacturer's

instructions, using a reverse primer that targeted the T2A region of the RNA molecule. Varying dilutions of this *in vitro* transcription reaction were used to generate a calibration plot for the detection of Nod(+) RNA *in vivo*. Ct values were determined via qPCR using Bio-Rad SsoAdvanced qPCR Master Mix on an Opticon CFXconnect fluorescent Peltier thermocycler. Three replicates of each dilution were used to generate the plot, which ranges over 5 orders of magnitude in terms of RNA copy number. This plot was used to calculate the number of Nod(+) molecules present in transfected cells.

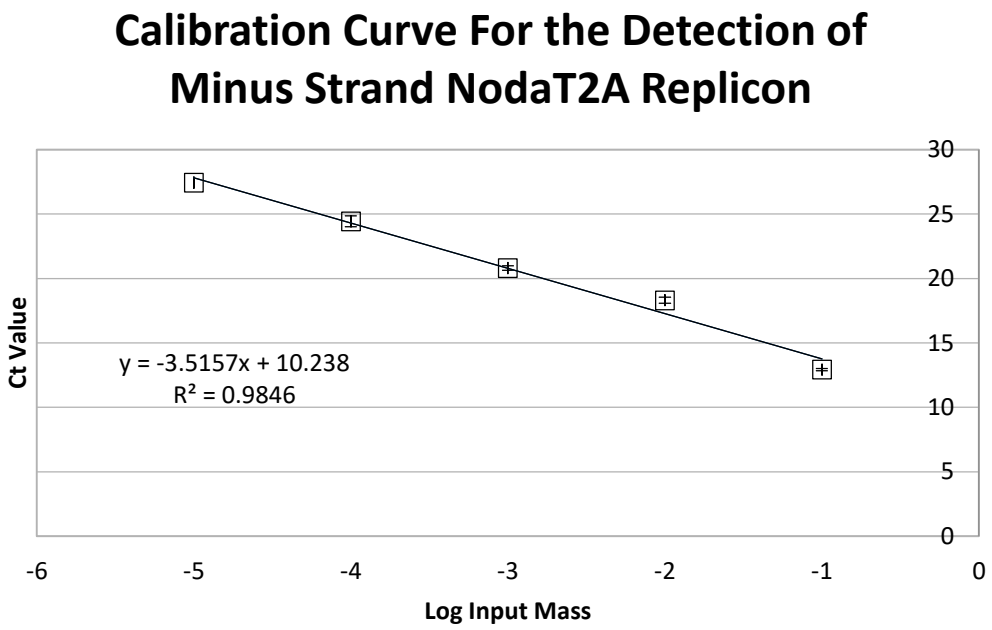


Figure 25: Calibration Curve for the Detection of Negative-sense RNA Molecules: Single-stranded DNA that corresponds to the chosen RNA sequence of the negative-sense replicon was purchased from IDT. This DNA was serially diluted over a range of 5 orders of magnitude, and qPCR was performed on these dilutions. The CT values corresponding to these concentrations was calculated using an Opticon CFXconnect fluorescent thermocycler, and a scatter plot of input concentration and Ct value was generated. This plot allows for the calculation of the number

of negative-sense RNA molecules that accumulate during pseudoinfection.

RNA accumulation in BHK-21 cells

Figures 26 and 27 show absolute RNA levels in BHK-21 cells transfected with either Nod(+) (Figure 26) or Nod(-)(Figure 27.) These values follow a similar trend to that seen in the cells of total RNA purified from cells. At short times (e.g., 8 h) post-transfection, there is little replicon RNA (of either strand) seen in cells transfected with either Nod(+) or Nod(-), with slightly more Nod(-) RNA (plus- and minus-strand) present. for both, the level decreases significantly by 48 hours; by 72 hours, most viral RNA of either sense has been degraded within the cell.

Surprisingly, the amount of minus-strand RNA present in transfected cells is very high, roughly on the same level as the positive strand. This is somewhat strange, as these RNA molecules are not packaged within the virion, and are not actively infectious. In many ssRNA viruses, the amount of minus strands is far less than plus strands because the plus strand is the actual viral genome and the minus strand is simply a way to replicate the initial RNA molecule. This is true for many viral models, such as CCMV, where the levels of minus RNA are much less than the positive strand. However, we are unsure if this replicon, which has evolved to work in the context of the whole virus, follows the same trend, and the lack of an RNA 2 molecule may account for this difference. The time course of RNA accumulation is not altogether different from the time course of a wild-type infection, indicating that incorporation of ribozymes and other moieties into the replicon has not changed the overall duration of its replication.

A B

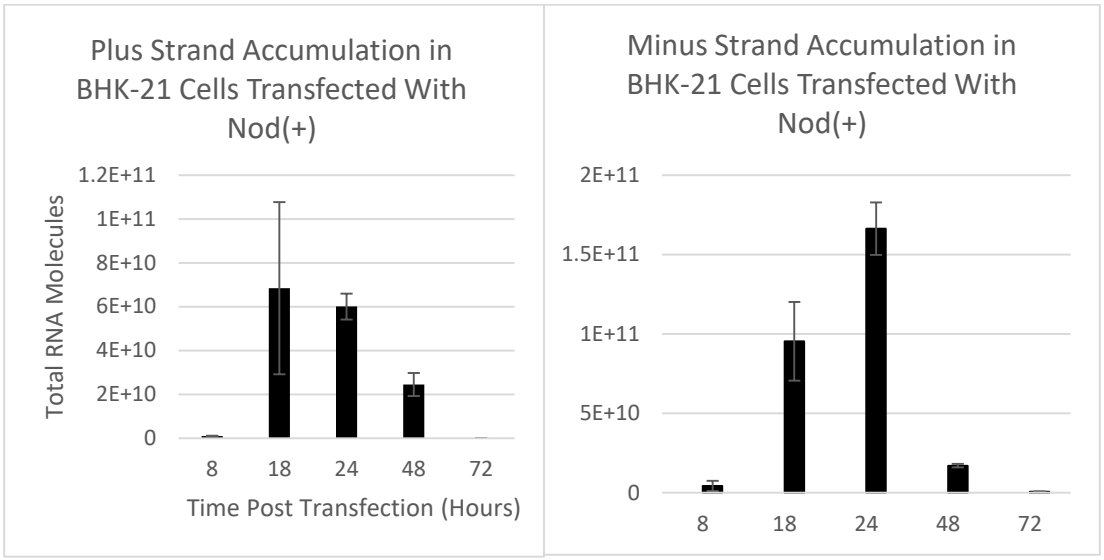


Figure 26: Accumulation of positive- (A) and negative-(B) sense RNA from BHK-21 Cells Transfected with Nod(+): Total RNAs from BH-21 cells were purified with a Qiagen RNeasy column at the indicated time points post-transfection. One µg of this RNA was reverse transcribed using primers that detect either positive or negative-sense replicon RNA, and then subjected to qPCR analysis. The resulting Ct values calculated via qPCR were then converted to absolute numbers via the calibration curve in Figure 24. In A and B, one can see that the accumulation of minus strands is roughly two times the amount of plus-strand RNA that accumulates. Both strands seem to follow similar kinetics, all but disappearing after 72 hours.

A B

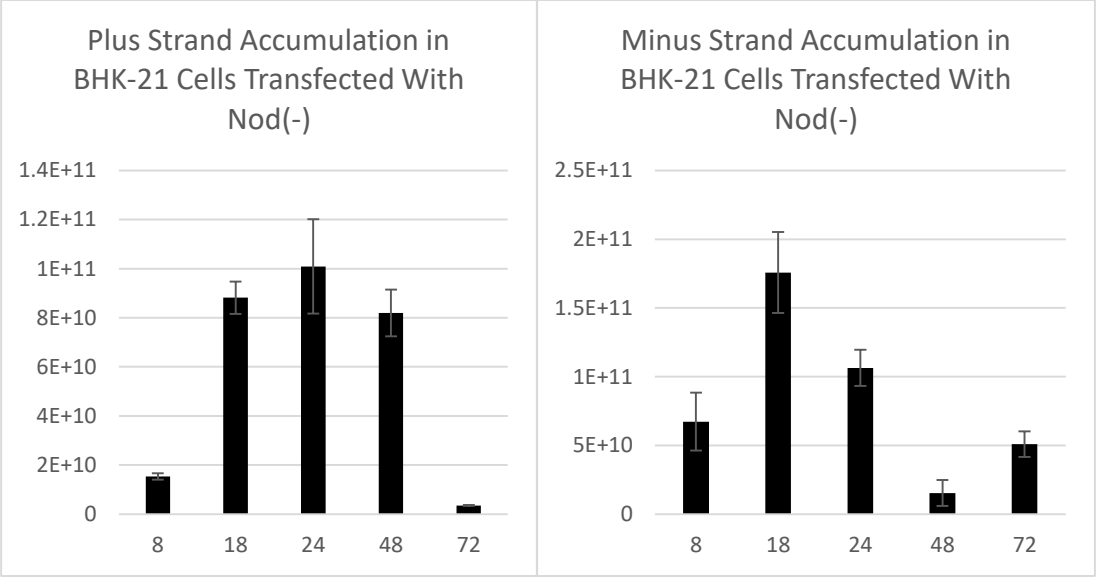


Figure 27: Accumulation of positive-(A) and negative-(B) sense RNA from BHK-21 Cells Transfected with Nod(-): Total RNAs from BH-21 cells were purified with a Qiagen RNeasy column at the indicated time points post-transfection. One µg of this RNA was reverse transcribed using primers that detect either positive or negative-sense replicon RNA, and then subjected to qPCR analysis. The resulting Ct values calculated via qPCR were then converted to absolute numbers with the calibration curve in Figure 25. Nod(-) transfections resulted in levels of plus and minus RNA similar to cells transfected with Nod(+). In A and B, one can see that the accumulation of minus strands is again roughly two times the amount of plus strand RNA that accumulates. Both strands seem to follow similar kinetics, all but disappearing after 72 hours.

Absolute quantification of mi34a levels in transfected BHK-21 Cells: amount of miR-34a produced by splicing depends on the strand used.

Figure 28 shows the calibration curve generated via RT-QPCR analysis of miR-34a purchased from IDT. The curve is linear across 4 orders of magnitude and has very good regression statistics. The

slope is a bit high, and this means that the microRNA detection threshold for this curve is less than those generated for the full length replicons. This curve can be used for the detection of mature microRNAs released during replication of the NodaMiRNA replicons.

Figure 29 shows the absolute number of microRNAs detected using RT-qPCR from cells transfected with NodaMiRNA replicons. The expression profile of miRNAs produced by Nod(+) follows the same general shape as the expression profile for Nod(+) itself. There is little to no expression of the RNA at early time points, expression of the miRNA peaks at about 18-24 hours, and starts to decay after 48 hours. It should be noted that the miRNA levels greatly exceed the replicon RNA levels at later time points. This is thought to be because the miRNAs have very long half-lives in cells (>48 hours) compared to the replicon RNA, which is viral in composition and actively destroyed by the cell.

For the Nod(-) replicon, the expression of the replicon RNA and the appearance of mature miRNA is not so tightly correlated. The amount of miRNA produced by Nod(-) is much lower at all-time points, and its apparent maximum expression – corresponding to about 10^11 molecules – does not occur until 72 hours' post-transfection. It should be noted that the replicon RNA levels in Nod(-) transfections is actually higher than Nod(+), and yet the amount of miRNA produced is orders of magnitude less than the amount produced by the splicing of Nod(+). The means that the Nod(-) replicons appears to be splicing less efficiently than Nod(+), and this splicing must be faster at later time points.

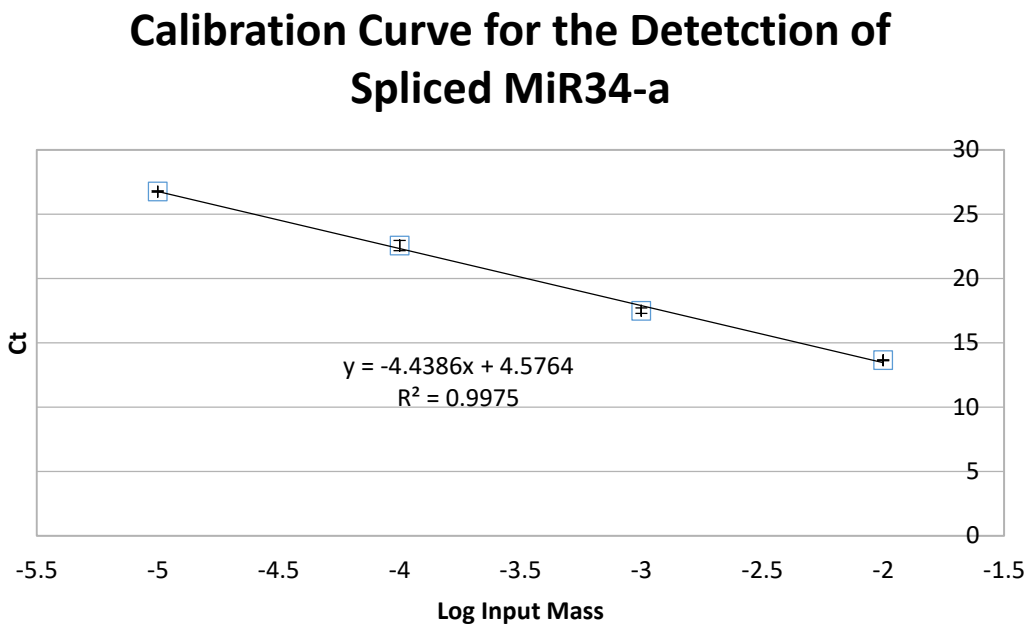


Figure 28: Calibration Curve for miR-34a detection. Full length miR-34a was purchased from IDT, and various dilutions of this RNA were polyadenylated and reverse transcribed using a primer that hybridized to the 3' region of the RNA. The resulting cDNA was them analyzed in triplicate using a standard qPCR protocol, and a calibration curve was generated using this data. This curve was used to calculate the absolute amount of spliced RNA in cells.

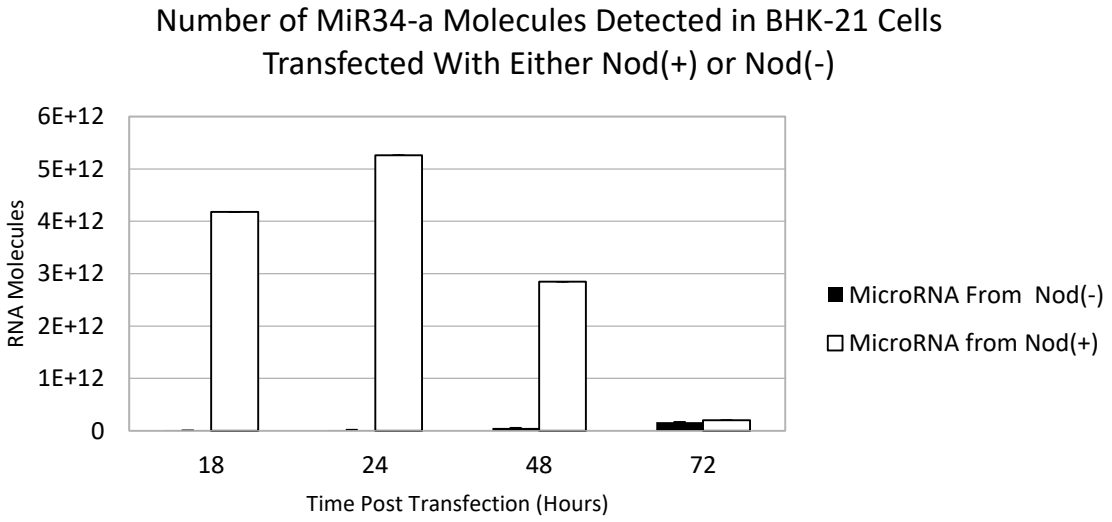


Figure 29: MiRNA34a produced by the splicing of NodaMiRNA replicons *in vivo*. The small-RNA pool of RNAs was harvested from cells transfected with either Nod(+) or Nod(-) RNA, and the presence of miR-34a was detected via a modified RT-qPCR reaction. Note that Nod(+) yields a much larger amount of spliced miR34-a than Nod(-).

Alternative splicing in Nod(-) transfections

Melting curves for the PCR product of the miRNA produced by mi34a, Nod(+) and Nod(-) are shown in Figures 30-32. The melt-curve corresponding to the miR-34a shows a single clean peak at around 65˚C, which is the expected resulted for amplifying a monodisperse RNA sequence of interest. Two time-points for the infected cells are shown: 18 and 72 hours. For cells transfected with Nod(+), there is a single peak in the melting-curves of the PCR products that is slightly broader than the curves seen in the amplification of MiRNA 34aThis increase in width indicates polydispersity in the length of RNA detected. The shifting of the peak to higher temperatures indicates that there is a subpopulation of RNA molecules that is slightly longer than the wild-type miR-34a. The melt-curves for short-RNA purified from Nod(+) transfected cells remains similar across time points, with the RNA produced at later time-points looking similar to the RNA produced initially. However, the spread of the peak decreases somewhat at 72 hours post transfection.

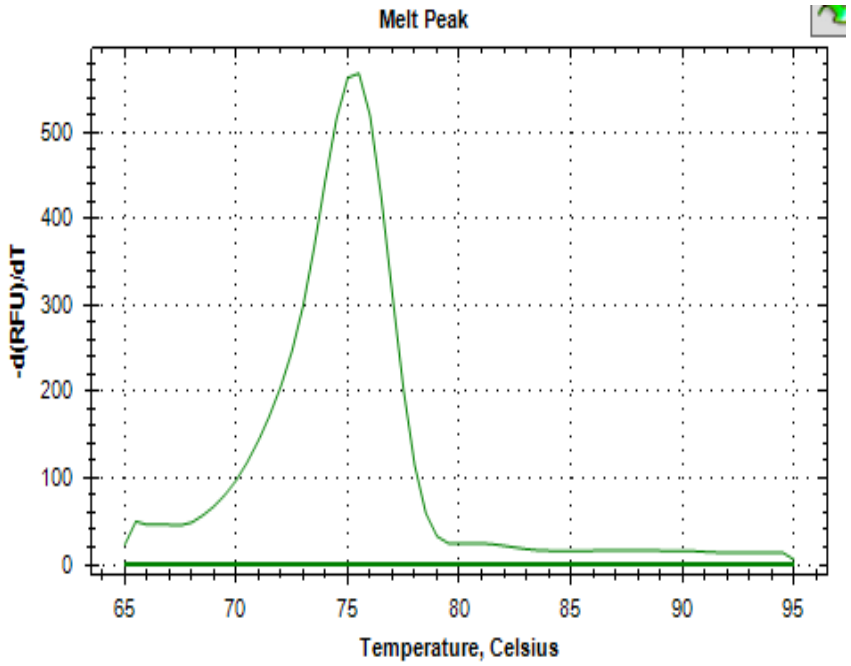


Figure 30: Melting Curve for miR-34a. The melt curve generated after amplification of 100 pg of miR-34a using microRNA RT-qPCR. After amplification, the fluorescent intensity of the PCR mix was measured as a function of temperature, the derivative of which is shown above. The single peak at 75C is indicative of a single species present after amplification.

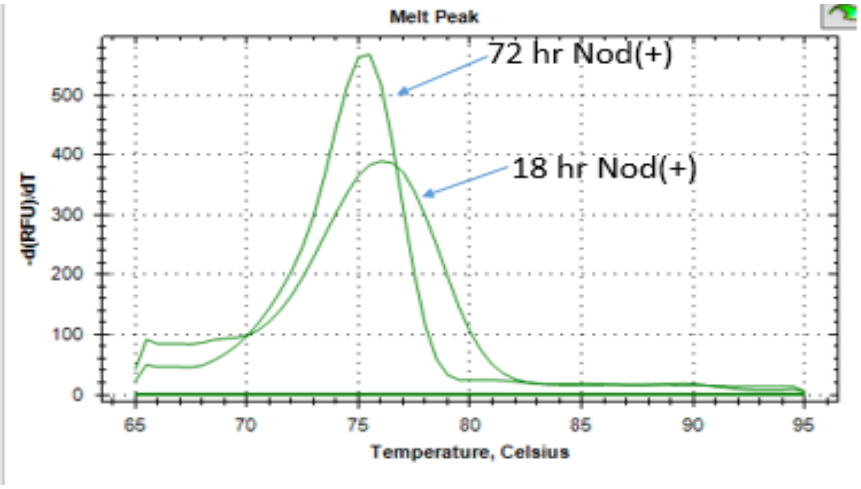


Figure 31: Time-dependent melting-curve Nod(+). The melt curve generated after amplification of 100 pg of small-RNAs from cells transfected with Nod(+) using microRNA RT-qPCR. After amplification, the fluorescent intensity of the PCR mix was measured as a function of temperature, the derivative of which is shown above. The peaks from derived from Nod(+) samples are broader than the peaks generated by the microRNA alone. The spread of the curve decreases over time, with the 72 hour time-point looking similar to the control microRNA.

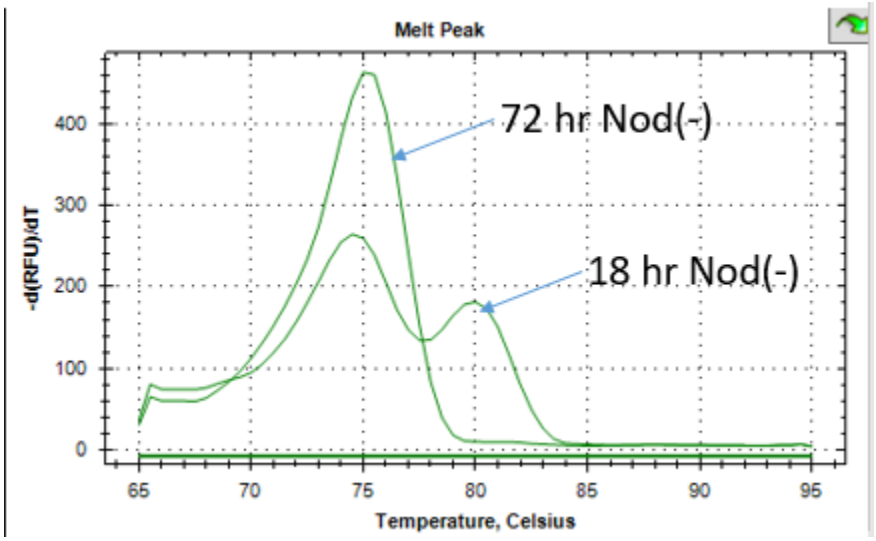


Figure 32: Figure 31: Time-dependent melting-curve for Nod(-). The melting curve generated after amplification of 100 pg of small-RNAs from cells transfected with Nod(+) using microRNA RT-qPCR. After amplification, the fluorescence intensity of the PCR mix was measured as a function of temperature, the derivative of which is shown above. There is a large secondary peak seen in the Nod(-) samples that is not seen in the curves generated from miR-34a RNA. This secondary peak disappears after 72 hours' post-transfection.

The melting curves for the PCR product of the microRNA produced by Nod(-) show a different

behavior (as seen in Figure 32.) At early time points, there are in fact two peaks present: one at 65˚C and one at around 70C, indicating the presence of a longer amplified species. Further complicating matters is the fact that these alternate, higher-molecular-weight species disappear at later time points, indicating they are inherently transient. Melting-curves from RNAs purified from cells transfected with Nod(-) show the expected single peak at later time points. The presence of two peaks in the melt-curve indicates that there are two species present to amplify during the qPCR step of the reaction. The nature of the qPCR probes means that the region must contain both the microRNA region, as well as the adapter molecule. To understand the nature of these other products, the PCR mixes from cells transfected with either Nod(+) or Nod(-) were run on a 1.2% agarose gel for 1 hour and stained with gel red.

Figure 33 shows this gel. The band on the left is taken from the PCR product of cells transfected with Nod(+) at 24 hours post-transfection. Here one sees a strong band at ~65 bp, which is the correct length for amplified miR-34a. In addition, there is a small band at roughly 75 bp, which is fainter than the properly spliced band. This band could account for the slight shoulder seen in the melt peak of amplicons derived from cells transfected with Nod(+). Figure 31 also shows the polydispersity in splicing that resulted in two peaks for the samples taken from cells transfected with Nod(-). Here one sees (in the middle lane) a fainter band at ~65 bp and a secondary band running at approximately 200 bp. This higher molecular weight species is thought to be the origin of the two peaks in the melt-curve.

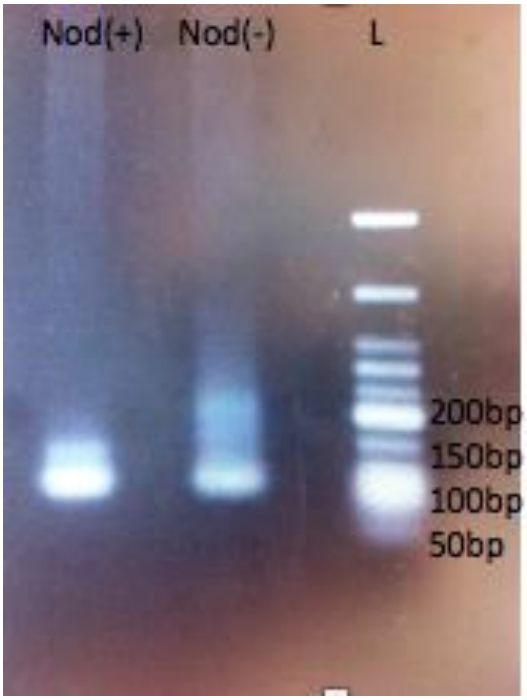


Figure 33: 1.2 % agarose gel of amplicons derived from cells transfected with NodaMiRNA constructs. Amplified DNA regions containing both the miRNA sequence, as well as the adapter sequence are selectively amplified using PCR. Amplicons from cells transfected with Nod(+) – lane 1, left – show a strong peak at 65 bp, with a slight smear running above. Amplicons from cells transfected with Nod(-) – lane 2, center – show a weaker band at 65 bp, and a secondary band appearing at roughly 200 bp. Lane 3, on the right, is a DNA ladder.

Discussion

Replication of Nod(+) and Nod(-)

Both NodaMiRNA constructs were capable of robust replication in BHK-21 cells, even with incorporation of self-splicing elements into the genome. This ability to replicate robustly was a concern, as the incorporation of highly structured RNA into replicons could affect the amplification of the molecule by changing the secondary/tertiary structure of the RNA. While the replication of the NodaT2A

replicon was unaffected by incorporation of either the EYFP or Luc RNA sequence into the replicon, both ribozymes used are much more structured than the RNA for either of these genes, and so their effect on the replication of the RNA molecule was unknown. Also of concern was the ability of the replicase to work through these regions of highly structured RNAs, which could lead to replication pausing and termination.

The ability of both these constructs to replicate also speaks to the ability, or lack thereof, of the ribozymes to splice *in vitro* and *in vivo.* It appears that both constructs generate a subpopulation of RNA that do not splice quickly *in vivo* and last long enough to replicate. There is a delicate interplay between the activity of the ribozymes and the activity of the replicase. Both ribozymes needed to splice sufficiently to produce detectable amounts of mi-R34a, without splicing so fast that the replicase could not operate on full-length molecules and thus continue RNA replication. If either construct spliced completely, and did so at a rate faster than that of replication, the pseudoinfection could not be established or maintained. Interestingly, the conditions needed for replication are different for each construct.

For Nod(+), both ribozymes needed to splice incompletely and slowly enough so that there were enough full-length molecules to enter into cells and begin replicating. Here, it was advantageous for the ribozymes to splice inefficiently *in vitro*. Inefficient here could either refer to either rates or equilibrium distributions, i.e. the ribozymes could either splice slowly, or they could have a subpopulation of RNAs that never splice and are kinetically trapped in a non-catalytic form. For Nod(+), there is compelling evidence from the *in vitro* studies that a small portion or the transcribed RNAs will never splice, because they are misfolded. Figure 11 shows that approximately 25% of the replicon RNA does not splice co-transcriptionally, and is stable for hours after transcription. This indicates that while the rates of splicing were high, the presence of misfolded RNA molecules would ensure sufficient RNA molecules would remain intact in order to establish pseudoinfection.

For Nod(-), it was unknown if the ribozymes would splice too fast *in vivo* for the RNA to replicate. Replication requires a full-length molecule with both 5' and 3' UTRs, and if splicing of nascent minus strands were too fast, the replicon would not amplify. RNA replication in Nodamura is not selective: the replicase will replicate both plus and minus strands effectively stochastically, and as such you need full-length molecules of both strands in order to maintain high replication levels. High amounts of Nod(-) replication also indicates that the incorporation of these highly structured RNA elements does not perturb the lifecycle of the replicon.

We were initially concerned that the small number of full-length Nod(+) molecules delivered to cells would not establish a good pseudoinfection, and we performed many assays to understand what occurs when one delivers small numbers of replicating RNA molecules to cells. Our work with NodaEYFP and NodaLuc shows a clear inverse dose-dependence on protein expression as a function of replicon input levels. For these constructs, higher levels of delivered RNA leads to an attenuated protein expression and weaker expression over time and this trend is also seen in RNA levels. This is not seen in NodaMiRNA constructs, as Nod(+) input levels are about 20% active molecules and still do not replicate to higher copy compared to Nod(-), which is delivered in high copy number. We currently do not understand fully why some constructs show this trend, and others do not. The NodaMiRNA constructs may not show this behavior due to the fact that their replication levels are lower than that of the non-splicing variants. We suspect that the inverse-dose dependence is in someway related to protein expression and/or RNA replication, and both Nod(+) and Nod(-) show lower levels of protein accumulation as well as RNA levels. Perhaps the splicing of the genome is sufficient to overcome any effects on input concentration.

Splicing efficiencies of NodaMiRNA Constructs

Nod(+) produces orders of magnitude more spliced miR-34a than Nod(-), even though their replication levels are similar. Furthermore, this is true even though the Nod(+) replicon produces less active RNA than Nod(-) at every time point. Because the RNA levels are similar for both constructs, the only factor that should affect the amount of miRNA produced should be the splicing efficiencies of the two ribozymes used to produce the miRNA. In general, the efficiencies of ribozymes are greater *in vivo* as the crowded cellular environment helps them fold into catalytically active molecules. The ribozymes here would also be exposed to additional crowding effects from the fact that they are replicated inside replication complexes nanometers in size. The efficiencies of the ribozymes could be affected by a myriad of factors and they all relate to their physical structures.

Looking at Figure 1 one can see that nascent minus-strand RNA molecules are not free to fold into the globular structure seen in many RNAs, as they are bound-up in double-stranded intermediates with the initial genome. The conditions in which the minus strands are synthesized all but guarantee that a minus-strand will form a stable double helix with its counterpart instead of adopting an active conformation. First, the growing RNA molecule is not fully formed, and is not folded into its equilibrium structure. Energetically, the formation of catalytic hairpins needed for ribozyme activity require base-pairing energies to make it spontaneous. The fact that the minus strand RNA is synthesized right next to its reverse complement means that there is a competing structure (the double-stranded intermediate) that is energetically more favorable. This means that instead of forming active strands, newly made minus strands are trapped in these intermediates. The action of the replicase, which unwinds any local secondary structure in the RNA molecule it is replicating, also ensure that the plus-strand is available to base pair with the minus strand.

The minus strand would only fold correctly if the plus strand RNA molecule is not present at high concentrations. At late time points, the levels of minus strand RNA are higher than plus strand RNA, so it is not surprising that this is when we start to see the gradual accumulation of spliced miRNA. The plus

strands may be either degrading or being actively exported outside the replication complexes, leaving the minus strands to begin splicing.

Origin of alternate splicing and differences in splicing rate

The alternate splicing seen in Nod(-) provides further evidence that the ribozymes in the minus strand are behaving differently than those in Nod(+). While Nod(+) showed a slight shoulder on the melt curve, and a corresponding secondary peak on the agarose gel of PCR products, the main product it produced was the expected microRNA sequence. Both qPCR values, as well as the gel, confirm that Nod(+) is generating correctly spliced mi34a. However, the additional peak seen in the melt-curve analysis, coupled with the extra, high-molecular-weight band seen on the gel, indicates that the Nod(-) construct is splicing aberrantly. This would also account for the low amounts of miR-34a that accumulates in Nod(-) transfections as detected by qPCR.

The method in which the mir-34a is detected helps us to begin to understand what is happening with Nod(-) splicing *in vivo*. First, only the small-RNA fraction is purified from cells; this means that any RNAs larger than 100nt are not collected, which prevents the collection of full-length and partially spliced RNA segments. After this, the RNAs are polyadenylated, and reverse transcribed using a long (~40 nt) DNA primer that hybridizes to the poly-A tail and the last 2 nucleotides of the miRNA. The qPCR probes hybridize to the microRNA sequence and the probe sequence on the adapter and this ensures the formation of a relatively monodisperse cDNA for subsequent analysis, which is seen in the Nod(+) lane of Figure 31. Aberrant hybridization of the probe to other parts of the 3' tail would result in a smear of sizes, which is not seen, so it appears that the probe molecule is binding correctly to the microRNA sequence, which means that the 3' end of the RNA is most likely correctly spliced. This means that the alternate spicing is most like the result of the hammerhead ribozyme splicing incorrectly.

Unlike proteins, ribozyme catalysis relies solely on the formation of base pairs between the catalytic portion of the RNA and its substrate. The CCS is engineered such that the microRNA target is the substrate for both ribozymes used, and as such the ribozymes must base-pair strongly to this region. Our ability to partially inactivate the ribozymes with oligos provides some insight into the role that base-pairing plays in the activity of the ribozymes. Consider now the conditions seen by the CCS in the minus strand (see Figure 34): here the nascent, catalytically active RNA is being synthesized, and it sees not only its intended splicing substrate (the microRNA sequence) but also its exact reverse complement in the form of the plus-strand RNA genome. The association energy from the formation of the long dsRNA duplex is much greater than that from the formation of a short catalytic helix in the ribozyme, and thus the presence of the plus strand may severely inhibit the ability of the CCS(-) construct to splice. The ribozymes could either splice intramolecularly as intended, yielding the microRNA, or intermoleculary, giving aberrant splice products that would be a hybrid of RNA stands. This splice product would then be detected and amplified with the qPCR primers, resulting in multiple bands, as seen in Figure 31.

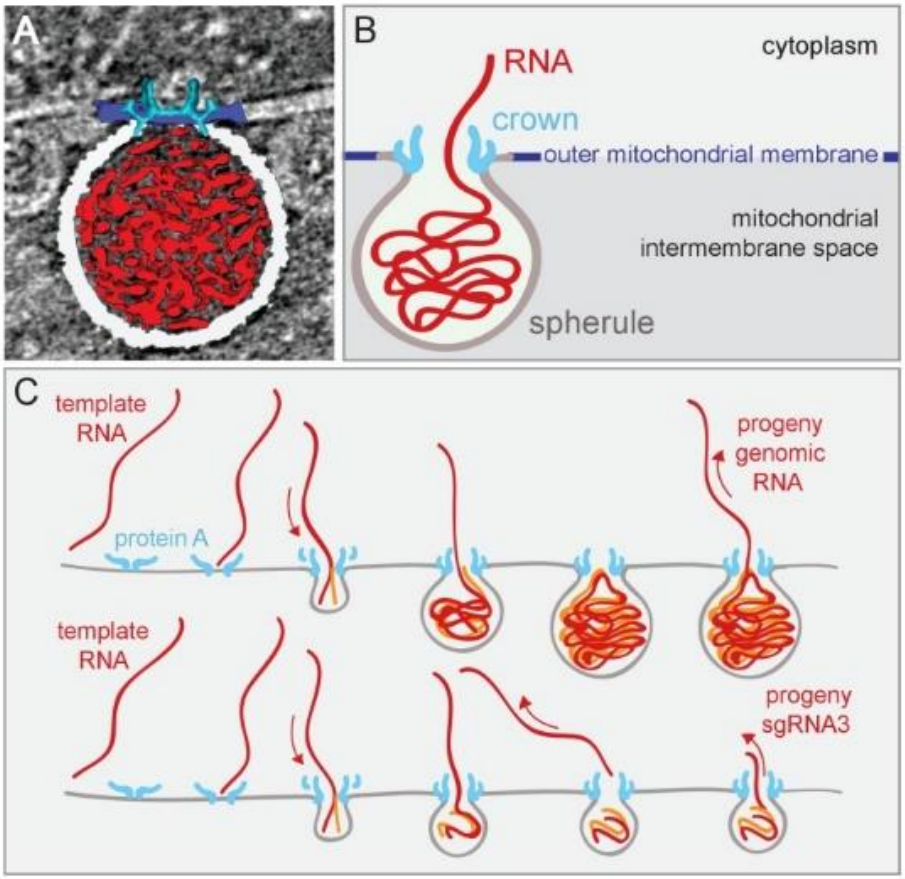


Figure 34: Replication of FHV within replication complexes (taken from reference 98). The replication of FHV (a virus similar to Noda) occurs in remodeled mitochondria, where the RNA is protected from the cellular immune response. These nanometer-sized spherules contain the replicase molecule, as well as both strands of the viral genome. The conditions within the replication complexes are very crowded, and there is extensive base-pairing between the plus-and minus strands. This interstrand hybridization, coupled with macromolecular crowding, is thought to be the origin of alternate splicing seen in both NodaMiRNA constructs[99].

Questions about biological activity

The data presented above indicates that both NodaMiRNA constructs are splicing in cells that

have been transfected with the replicons. However, it is unknown if the microRNAs that are detected using RT-qPCR are biologically active. For one, it is unknown if the microRNAs are able to leave the replication complexes in order to be recruited by the RISC complex. The purification process for the RNA first lyses the cells, so it is not known if the RNA is available and free in solution and not localized in regions where it has no biological activity. It is also necessary for it to be free in solution so that it can be phosphorylated by cellular kinases prior to incorporation into the RISC complex (see Figure 7).

It is known that positive-sense viral RNA that contains a pre-miRNA hairpin can be processed by DICER, while the same construct delivered in the minus sense RNA cannot, owing to the minus strands inability to leave the replication complexes. In the present case it needs to be determined if the microRNA that was produced via the action of the ribozymes would be small enough to leave the complexes via diffusion, unlike the full-length molecule. In addition, there is a direct interaction between DICER and the RISC complex, so it is unknown if the mature microRNA, which never interacts with DICER, would be able to regulate RNA transcription/translation to a similar degree as shRNA, which is strongly associated with the cellular components needed for miRNA activity *in vivo*.

Chapter 5: Analysis of Biological Activity

Background

Replication and cytotoxicity

Many replicons derived from single-stranded viruses exhibit some degree of cytotoxicity[100]. In

particular, alphavirus replicons have been shown to induce apoptosis in cultured cells[101]. While there are many reasons for this, at least part of this cytotoxity derives from the reorganization of interior membrane compartments, which can drastically alter the landscape of the cell. Viruses that replicate within mitochondria risk destroying the mitochondria all together, inducing apoptosis and killing the cell[102]. Viruses that replicate within the ER could change the calcium levels in the cell, which also induces apoptosis. The amount of viral replication certainly affects this delicate balance, with increasing amounts of viral replication further resulting in apoptotic cells. Viral replication also depletes the cell of needed biological cofactors, and surely slows cell growth.

In general, we have found that replicons derived from Nodamura viral genomes are less cytotoxic than other replicons. We suspect7 that this is due to the fact that Nodamura is an insect virus that only replicates somewhat efficiently in mammalian cells. This expression is further educed by the fact that mammalian host cells are warmer than the optimal temperature for replication (33C.) Moreover, we suspect that the protein B1 helps reduce cytotoxity in these replicons. In betanodavirues, B1 serves to limit mitochondrial depolarization and destruction, which is closely associated with apoptosis[103].

Mitochondrial dynamics

The mitochondria of the cell are in a constant state of flux and undergo a heavily regulated series of fission and fusion events. In early parts of the cell cycle, these mitochondria are in an extended networked state that extends though the cytoplasm. As the cell prepares to undergo mitosis, these long extended networks break up into the small kidney-shaped organelles with which we are most familiar. After the cell divides, the mitochondria again regain their tubular structures. The extent of fusion in mitochondria is therefore tied to the cell cycle[104].

There is compelling evidence the viral infections alter mitochondrial dynamics. Cells infected with HIV show mitochondria undergoing significant amounts of fission. This is most likely due to the fact that the viral replication proteins are remodeling the interior membranes, and inducing the mitochondria to undergo fission[105]. It is unsurprising then that there may be some association of viral replication with the cell cycle. Many viruses arrest cells at a certain point of the cell cycle, while other induce changes in the cells own time-keeping machinery. Further, certain viruses, such as SARS require the cells to be in a certain part of the cell cycle to be infectious, a trait which may be shared with the NodaMiRNA replicons as well[101].

Methods:

Cell proliferation assay

PC3 cells were cultured in F12K medium supplemented with 10% FBS and split into 24 well plates at approximately 10% confluence. These cells were then transfected with 1.5µg of Nod(+), Nod(-) or EYFP replicon RNA using lipofectamine 2000 following manufacturer's instruction for 18 hours. As a control, cells were mock transfected using only the transfection reagent with no replicon RNA present. The transfected cells were placed in a 31˚C incubator for the initial 36 hours, after which they were transferred to the 37˚C incubator. The cells were moved from the 31˚C incubator to the 37˚C incubator to allow for the initial replication of the RNA (which is much stronger in PC3 cells at 31˚C than 37˚C) without significantly affecting the cells overall growth profile over the duration of the experiment, because the cells grow much faster at 37˚C than 31˚C. Fluorescent images of transfected PC3 were used to determine when the replication in EYFP-replicon-transfected cells was occurring.

At particular times post-transfection cells were collected and counted using a hemocytometer.

The replicon can induce infected cells to become semi-adherent, and as such the cell-collection protocol was modified to ensure collection of all the cells present. More explicitly, the media on top of the cells, as well as the PBS used to wash the cells, were collected, in addition to the cells that were removed from the plate using trypsin. This allowed for the collection all cells, not just firmly adherent ones. This is particularly important for PC3 cells, as several of the cell's phenotypes are only weakly adherent, a trait made worse by replication. The collected cells were pelleted gently, and resuspended in identical volumes of media supplemented with trypan blue. Ten µL of this resulting mixture was added to a hemocytometer, and the total number of live and dead cells was measured. The effect of the replicon on the cells was determined two ways: by measuring the proliferation of the transfected cells, and by measuring the number of dead cells via trypan-blue exclusion. Previous assays with BHK-21 cells indicated that the replicon was not cytotoxic to those cells, so any increased cytotoxicity must be tied to either the cell type transfected or the biological activity of the spliced microRNA.

Results:

NodaT2A replication affects tumor cell proliferation

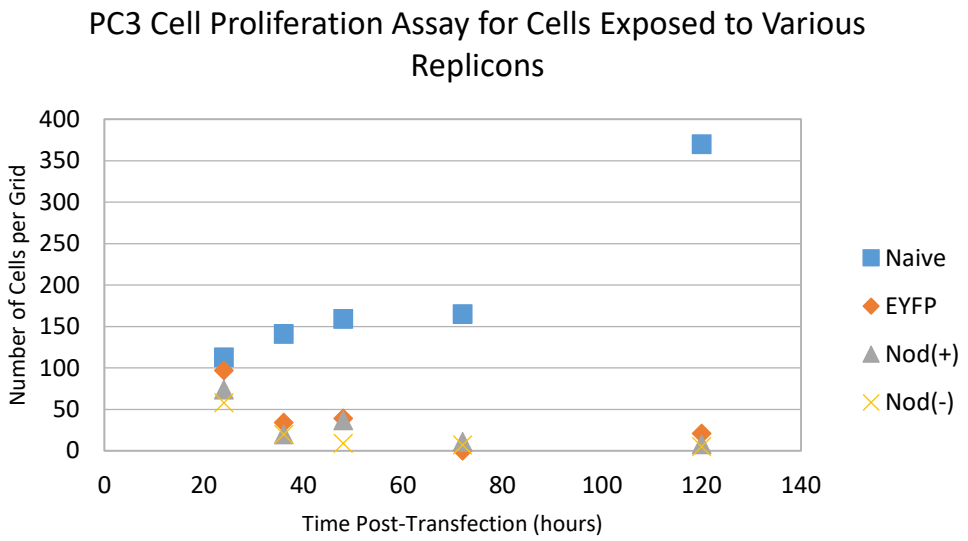


Figure 35: NodaT2A replicons limit PC3 cells' ability to spread in culture. PC3 cells were transfected with 1.5µg of Nod(+), Nod(-) or NodaEYFP RNA or left untransfected (naïve.) These cells were incubated at 31 C for 24 hours, to ensure replication, then transferred to 37˚C. At the indicated times post-transfection cells were collected and counted with a hemocytometer. All replicons limited cell spread, but both MiRNA replicons were more effective in overall suppression of cell growth: cells transfected with NodaEYFP were many times more numerous than cells transfected with either NodaMiRNA construct. The naïve cells spread as expected, doubling every 48 hours or so.

Figure 35 shows the effect of the Nodamura replicon on PC3 cell proliferation is to limit the growth of transfected cells, when compared to the mock-transfected cells. Mock-transfected cells grew as expected, doubling every 48 hours or so, which is consistent with what is known about their growth. All cells treated with the replicon showed significantly reduced growth after 36 hours post-transfection. The EYFP replicon, as well as the two microRNA producing replicons have similar growth curves, with the total number of viable cells decreasing over time. However, the cells transfected with either Nod(+) or Nod(-) have fewer viable cells after 48 hours, and at 120 hours there are 3 times as many viable cells in the EYFP-replicon-treated cells than in cells transfected with Nod(+/-). The number of viable cells decreases fastest for cells treated with Nod(-), while the cells treated with Nod(+) persist for 72 hours

before dwindling to values similar to Nod(-) RNA.

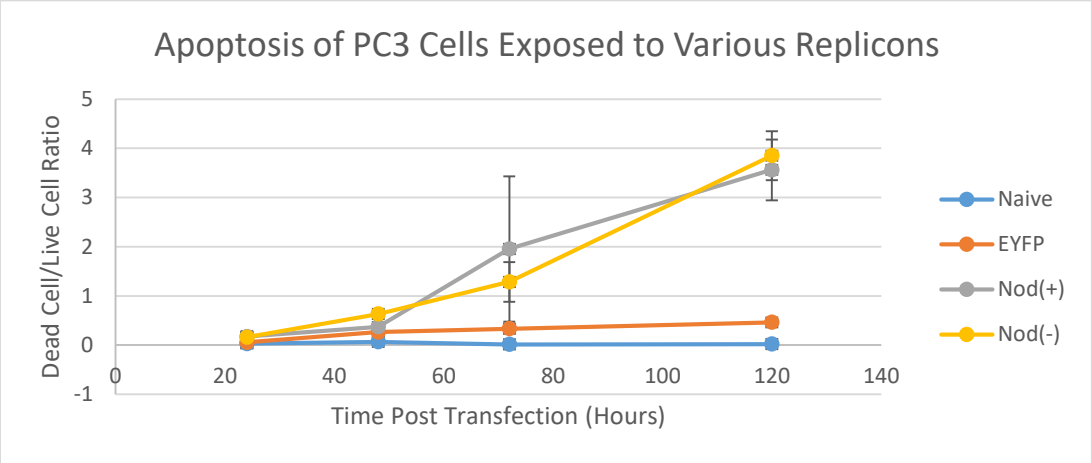


Figure 36: NodaMiRNA constructs induce apoptosis in PC3 cells cultured *in vitro*: PC3 cells were transfected with 1.5µg of Nod(+), Nod(-) or NodaEYFP RNA or left untransfected (naïve.) These cells were incubated at 31 C for 24 hours, to ensure replication, then transferred to 37˚C. At certain time post-transfection, cells were collected, stained with trypan blue, and counted with a hemocytometer. The amount of apoptotic cells for the naïve cells remained constant, as did cells transfected with NodaEYFP. Cells transfected with either NodaMiRNA construct showed dramatically increased levels of apoptosis compared to either control.

Use of trypan-blue allows for the detection of dead cells in culture, and for direct probing of the replicon's cytotoxic effects. These dying cells are stained blue by the trypan dye, and are easily counted during the measurement of viable cells. By calculating the ratio of dead to live cells, we can generate a useful metric for determining the effect of the replicon of the viability of the cell(Figure 36). Mock-transfected cells demonstrated an almost constant ratio of 0.09. Cells transfected by EYFP replicon showed a steadily increasing dead/live ratio, starting with a ratio of 0.04 at 24 hours, and ending at 0.5 at 120 hours. Both curves generated by either Nod(+) or Nod(-) are functionally different than those of either the mock-transfected cells or the EYFP transfected cells. The cells transfected by Nod(+) show

ever-increasing levels of dead cells, with a final dead/live ratio of about 3.5. Similarly, cells transfected with Nod(-) also show monotonically increasing numbers of dead cells, culminating in a ratio of approximately 4.

Discussion

Replicon Replication in PC3 cells

The replication of the NodaT2A family of replicons in PC3 cells varies markedly from replication inside BHK-21 cells. Replication of the RNA molecules only proceeded at 31˚C; no replication of the EYFP replicon was seen at 37˚C. In fact, the strongest expression of EYFP from the NodaT2A replicon that comes from cells that were initially kept at 31˚C for 24 hours, before being transitioned to 37˚C. We also decided that keeping the cells at 31˚C would limit their ability to spread and lead to false-positives. Cells treated in this manner showed a large fraction of fluorescent cells after transferring to the 37˚C incubator. We believe that this effect arises from the fact that while the replicon works better at lower temperatures, the PC3 cells express proteins best at 37˚C. Transitioning from 31 to 37 C allows the RNA to accumulate, and then be translated once the cells are at their preferred temperature. with the replicon.

This behavior is thought to be a result of the PC3 cells own innate immune system, which is quite strong. Unlike other prostate cancer cell lines, PC3 cells have been shown to be resistant to ssRNA viral infections: for example, PC3 cells challenged with Vesicular Stomatitis virus(VSV) show much less RNA accumulation and viral protein expression than LNCaP cells. It should be noted that NoV replicates well in suckling mice, and that BHK-21 cells are derived from hamsters, so there may be a cofactor common to hamsters and mice that is not present at the same levels in human cells, which may explain the differences in NodaT2A replication.

The inverse-dose-dependence seen in transfections of BHK-21 cells is not seen in PC3 cells. Instead, there is a direct correlation between number of transfected PC3 cells and the amount of RNA transfected into cells. Transfections of PC3 cells with the optimal amount of RNA as determined from the dilution experiments in BHK-21 cells produced no visibly fluorescent PC3 cells. The number of fluorescent PC3 cells, however, increased directly with the amount of delivered replicon RNA. It should also be noted that transfection of PC3 cells was decidedly binary: cells were either transfected and replication was robust (the cells are very bright) or no fluorescence was visible. No PC3 cells exhibited low levels of fluorescence, and this trait is also seen in BHK-21 cells. There is little reason to believe that the transfection conditions are such that some cells are not being transfected; instead there is something intrinsic to the cells that determines if the cell is able to be infected. It could be that the cell cycle plays an important role in determining which cells are easy to transfect, or it could be that the mitochondria inside the cells is more amenable to viral replication. A more thorough (and speculative) analysis of this is found in the next chapter.

Biological activity of the nodamura Replicon

The activity of the NodaEYFP replicon is sufficient to slow the growth of PC3 cells *in vivo*. Unlike BHK-21 cells, PC3 cells transfected with NodaEYFP grow much slower than mock-transfected cells. The total number of cells initially decreases in NodaEYFP-replicon-treated cells, before leveling off after 48 hours. While the number of cells present in EYFP transfected cells is still larger than that of either NodaMiRNA construct, this demonstrates that the replicon itself, outside of the delivery of microRNA, attenuates the spread of cancer cells. This decrease in viability could be one or a combination of many factors. The replicon could be simply stressing the cell such that it no longer expresses the proteins etc. needed for replication. The shift to viral protein production alters the production of critical cellular

proteins, such as actin, that are needed for cell division. Another source of this effect may be the active reorganization of cell mitochondria, which could also affect the proliferation of the cells, and this mitochondrial effect may result in less ATP production for the cell. Additionally, the activity of the virus may alter the cells natural division cycle, which is a common trait in many viral infections. Finally, it could be that the intensive RNA replication and protein translation that accompanies replicon pseudoinfection simply depletes the growth media of small molecules needed for cell division (such as NTPs etc.)

It should be noted that even if the NodaEYFP vector limits the spread of PC3 cell *in vitro* it does so with minimal induction of apoptosis. Figure 36 shows this as, even at late time points, the number of dead cells (as measured by trypan blue exclusion) is much smaller than for either NodaMiRNA construct. In addition, the relative number of dead cells in EYFP treated samples remains constant throughout the course of the experiment, indicating that any apoptosis is likely a result of initial replication/transfection effects. While the number of dead cells in these experiments is an order of magnitude greater than the naïve cells, it is still much lower, and has a different profile, than that of either NodaMiRNA construct.

Biological activity of NodaMiRNA constructs

Both Nod (+) and Nod (-) show strong effects both on the proliferation of transfected PC3 cells, as well as their ability to survive *in vitro*. Figure 32 shows a strong response to the activity of both replicons, as both sets of cells transfected with either Nod(+) or Nod(-) do not spread in culture. Instead, the number of cells in both NodaMiRNA-transfected wells continuously decreased. This effect is stronger than the effect of simply the NodaEYFP replicon (the cells treated with EYFP replicon have 6 times as many cells than either NodaMiRNA construct) which implies that there is something unique to the NodaMiRNA construct that is slowing cell proliferation. MiR-34a is known to limit the ability of PC3

cells to proliferate in cell culture. It does this by regulating a myriad of cellular factors, including Jagged 1 and Notch1 that are both associated with tumor cell invasion in cervical cancers. By 120 hours' post-transfection the number of viable cells in both Nod(+) and Nod(-) had dwindled to a scant 15% of the initial number of cells seeded.

This overall decrease in the number of viable cells treated with the NodaMiRNA constructs is coupled with a dramatic increase in the number of apoptotic cells. Figure 32 shows a ratio of dead to live cells in PC3 cells transfected with various NodaT2A replicons. These cells have been stained with trypan blue to detect any cells that are apoptotic. Here there is a distinct difference in the curves generated by either NodaMiRNA construct and the NodaEYFP construct. While the NodaEYFP construct did indeed lower the number of viable cells when transfected, it is not correlated with a strong increase in the number of apoptotic cells. Naïve cells had a D/L ratio of roughly 0.09, while this ratio increased to 0.5 for cells treated with NodaEYFP replicon. However, cells treated with Nod(+) or Nod(-) saw an increase in the total number of apoptotic cells, with D/L ratios of 3.4 and 4.1 respectively. Cells treated with Nod(-) began to die quicker, and reached a steady state faster than cells transfected with Nod(+). We suspect this is because we are delivering more intact RNA than in the Nod(+) transfections, owing to the fact that many of the Nod(+) constructs have spliced prior to entry. Unlike BHK-21 cells, PC3 cells do not exhibit inverse dose-dependence, and fewer transfected RNA means less replication overall.

The strong effect of these replicons on PC3 cancer cells provides compelling evidence that the splicing of the replicons is generating biologically active miRNA *in vivo*. This means that not only do the ribozymes splice correctly to produce the miRNA sequence, but these miRNAs are capable of leaving viral replication complexes and of being actively recruited into the RISC complex. In addition, it means that these miRNAs are being phosphorylated correctly, and are in general interacting with cellular components in an expected manner. Previous studies have shown that miR-34a is effective in slowing PC3 cell proliferation; here we see that transfection of our construct actually decreases the number of

viable cells and dramatically increases the number of apoptotic cells. In many ways it seems that this construct is actually more effective than shRNA for affecting the spread of cancer cells in culture[76].

It should be noted that the activity of both replicon goes beyond simply restoring basal levels of miR-34a; it upregulates them dramatically. Depending on the tissue, the amount of mir-34a within cells varies between 100-1200 copies per cell [49]. In our studies, we detected approximately $1.0x10^{11}$ copies of miR-34a in RNA purified from transfected cells. This amounts to roughly 100,000 copies of mir-34a *per cell,* indicating that we have increased the amount of miRNA in transfected cells by 100 fold. We are unsure what off-target effects this massive influx of miRNA may have on the cell. We do know that single-stranded miRNAs are much less effective than hairpin RNA, owing to the direct interaction between DICER and the RISC complex, so it could be that we are compensating for this by massively increasing the amount of RNA. Interestingly, the less effective construct (i.e. the one that made the least total microRNA) is actually the best at restoring *homeostatic* levels of microRNA to transfected cells. Indeed, if one wanted to truly restore these levels to cells, one would need to attenuate either the replication of the RNA, or the activity of ribozymes, in order to reduce the amount of miRNA present. We do not know if there is a functional difference between the levels of miRNA generated by the two NodaMiRNA replicons, as they both have the same effect on transfected PC3 cells, and further studies probing this may provide insight into what can be considered “optimal” microRNA concentrations.

Chapter 6:

Conclusions and Future Work

Summary of Findings

In vitro studies

We have developed a system wherein we can deliver self-replicating microRNA to cells of interest using either (the positive- or negative- sense) strand of a replicon derived from the insect virus *Nodamura*. To do this, we have introduced a system of coupled, self-cleaving, ribozymes, known as the core catalytic sequence, that is surprisingly effective at splicing both *in vitro* as well as in cell culture. These coupled ribozymes splice quickly both in the CCS itself as well as when embedded into the full length Noda Replicon. Our choice of ribozymes (hammerhead on the 5' end and Group 1 intron on the 3' end) produce a single microRNA species upon splicing, with little off-target, alternative, splicing seen in the *in vitro* splicing experiments presented above. Moreover, these ribozymes are quite efficient at splicing, with the CCS splicing almost completely during the course of transcription. This efficiency is attenuated somewhat when the CCS is incorporated into the full-length replicon, where we find a subpopulation of unspliced replicons that is stable at even at late time points post-transcription.

To further attenuate the co-transcriptional splicing of the CCS when it was incorporated into the replicon, we designed a system of anti-sense oligonucleotides (ASOs) that effectively inhibit the splicing of both the ribozymes in a sequence specific manner. We had tried other small-molecule inhibitors, such as arginine and kanamycin, but these effectors proved also to be inhibitory to the action of the T7

polymerase used in the in vitro transcription reactions. These ASOs reduced the catalytic efficiency of the ribozymes in the Nod(+) replicon by roughly 4-fold, which ensured that sufficient full-length replicon was created during transcription.

We embedded the CCS construct into each of the strands of the Noda replicon RNA, thereby engineering separate constructs that produce the miRNA at different stages during the replicon pseudoinfection. The initial construct tested (CCS+, which splices in the plus-strand of the replicon) was incorporated into the Noda replicon, resulting in the construct known as Nod(+). This construct would produce microRNA from plus-strand replicon molecules produced during replication. An exact reverse complement of the CCS+ sequence, known as CCS-, was incorporated into the same molecule as CCS+, resulting in a construct termed Nod(-). In this construct, the CCS is not active until the initial RNA is replicated, producing the minus-strand RNA molecule that contains the correct RNA sequence to self-splice and yield miR-34a in cell culture. This molecule does not splice during IVT, and as such was a compelling option to deliver the microRNA, as all RNA molecules delivered would be replication competent.

Replicon dilution studies

Even in the presence of inhibitory ASOs, the *In vitro* transcription of Nod(+) RNA resulted in a mixture of both spliced and unspliced Nod(+) RNA. After purification, only about 25% of the total RNA present in these experiments was unspliced and capable of establishing a pseudoinfection. The physics of lipoplex formation dictates that the number of RNA molecules within a single transfecting lipoplex is based on charge-neutralization and inversely dependent on length[97]. There was concern that delivering this small number of replicon molecules to cells would be insufficient to establish a pseudoinfection. Because the detection of these RNA molecules requires RT-qPCR, which is a long process, we needed a

way to quickly asses the effects of this “dilution” on replication in mammalian cells.

To ensure that delivering small numbers of active replicon would not adversely affect the course of the replication, we designed a system that allows us to “dilute” the replicon RNA with another RNA that is not actively translated. We used a carrier RNA molecule that was approximately the same length as the replicon RNA, to ensure that the lipoplexes formed contained on average a representative mixture of the two RNA molecules to be transfected. We tested a variety of ratios, from a single replicon molecule/lipoplex, up to using only the replicon molecule during transfection.

For BHK-21 cells at 37˚C we found an unexpected, inverse dose-dependence for protein translation as a function of replicon RNA input. For transfections with “normal” mRNA, increasing the amount of RNA delivered to cells increases the amount of protein produced. However, for transfections with replicon RNA, larger amounts of input RNA did not increase protein expression, but actually attenuated it somewhat. In addition, the overall duration of pseudoinfection is shorter in cells transfected with less replicon RNA. We suspect that these effects may be due to several factors, owing to either the host cell response, or the lifecycle of the replicon itself.

The first step in +ssRNA viral replication is the translation of the vgRNA into replicase proteins. In many viral infections, only one vgRNA molecule needs to enter the cell to initiate an infection. Transfection with high copy number of replicon molecules delivers 800-2000 RNA molecules into a transfected cell, which is much higher than that seen in a “normal” infection. All these RNA molecules will be initially translated, resulting in a large number of replicase proteins accumulating within the cell in a short period of time. These replicase molecules remodel mitochondrion, and can induce cell immune response if present in large amounts[45]. As such, traditional transfections may actually deliver too much RNA to the cell at once, which produces protein A past a certain threshold, inducing the cellular immune response and reducing the overall amount of protein produced. That this is so is further supported by the fact that this effect is not seen in cells at 31˚C, where translation within cells is

attenuated compared to 37˚C. Less protein production means less protein A and a reduced immune response.

Viral RNA itself is sufficient to induce cellular immune responses in cultured mammalian cells. In addition, long stretches of dsRNA, such as those found in +ssRNA viral replication strategies, are also immunogenic. Most +ssRNA viruses circumvent this effect by replicating themselves within replication complexes formed from reorganized internal membranes of organelles. In doing so, the viral RNAs are sequestered from the host own immune response. Transfections, however, release viral RNA into the cytoplasm directly, where it can be detected by the cells' immune system. As such, introducing large amounts of viral RNA into cells via transfection could trigger the hosts immune system, in a similar manner to the protein A above. Again, we see the temperature studies can provide a compelling argument for this hypothesis. At 37˚C, the 3' UTR of the Noda replicon contains long double-stranded regions, which are good stimulants for the interferon response. However, at 30˚C, this region of RNA adopts a tRNA like fold. This fold, present only at 30˚C, could help the RNA avoid host-cell defenses by mimicking a tRNA within the cytoplasm.

Finally, the native life-cycle of the virus may also be affected by the input amount of RNA. In a "traditional" viral infection, only one genome enters a cells. This genome is translated into Protein A, which is then replicated within replication complexes. At a certain point in a natural infection, the replicase protein shifts from replication to transcription. In other words, it switches from synthesizing minus strands (which generates dsRNA intermediates) to new plus strands. Likewise, RNA replication in general is reduced by large amounts of +ssRNA: at a certain point during infection the virus must transition from making RNA to assembling virions, and this switch occurs often due to accumulation of positive-sense strands. However, again transfection delivers many more RNA molecules than are present in a natural infection and this could have ramifications for replication. A replicase molecule that is translated from a viral infection has only one template molecule (the initial mRNA) and no minus

strands present, and as such it is happy to initiate replication via minus strand synthesis. However, in transfection, one delivers MANY of these molecules to cells, and a newly synthesized replicase molecule would encounter many +ssRNA molecules, and many replicase molecules, which is a condition only seen at the end of replication, when the virus begins to produce virions via self-assembly. Thus, the attenuated protein production may be a due to the fact that transfections mimic the conditions of late-stage infections (large numbers of free, positive-sense RNA), which prevent further replication and protein synthesis.

It is interesting that many of the monolayers transfected with dilute replicon RNA show unexpected fluorescence profiles. When transfecting with even a small amount of replication-competent RNA, one would expect that almost all the cells would be fluorescent during transfections. For traditional mRNA transfection, the transfection efficiencies (i.e., the fractions of transformed cells – ones expressing the mRNA protein) can be up to 90%, while transfections using the Noda replicon are limited to about 60% efficiency. Analysis of the fluorescent micrographs indicates that cells transfected with NodaEYFP exhibit bimodal behavior: cells are either pseudoinfected and *highly* fluorescent, or they are not fluorescent *at all*. Based on the amount of RNA used in all transfection experiments, it is unlikely that these cells are not being exposed to the replicon RNA, yet they do not become fluorescent even at late time points. Clearly there is something intrinsic to the *state* of the cell that makes it more susceptible to replicons. We know that Nodamura replicates within replication complexes within discreet, small mitochondria. Mitochondrial dynamics are closely linked with the cell cycle[106]; mitochondria form more extended tubular structures during the early stages of meiosis, and then undergo fission during cell division[107]. Perhaps only cells that are in the appropriate phase are vulnerable to replicon pseudoinfection. This is not to say that *viral* infection is dependent on the cell cycle, only that this replicon, which lacks many of the evolutionary "tricks" found in the wild-type virus, may be sensitive to this factor.

Both NodaMiRNA constructs replicated to high copy number in BHK-21 cells. Samples of total cellular RNA purified from transfected cells shows that both NodaMiRNA constructs accumulate to levels similar to that of ribosomal RNA. This RNA persists for about 72 hours, before eventually being degraded by the cell. We developed a method to accurately measure the accumulation of both positive and negative-sense RNA strands via RT-qPCR, and this method allowed us to quantify absolutely the numbers of each strand of the replicon RNA molecules in an infected cell. This is the first time this has been done for cells transfected with a modified Nodamura replicon. Interestingly, we found that unlike other viral infections, roughly twice as many minus strands are synthesized than plus strands for both NodaMiRNA constructs; in many viral infections, the amount of minus strands is an order of magnitude *less* than the amount of plus-strand genomes. In studies on RNA accumulating in viral infections of wild-type Nodamura, this was also found to be true. However, in many ways the Noda replicon is different from the Nodamura genome; most critically, it lacks RNA 2. It is understood that the interaction between RNA 1 and RNA 2 help the replicase switch from synthesizing minus strands to plus-strand synthesis, and these interactions further drive the synthesis of *sub*genomic RNA. The NodaT2A replicon is a single molecule, however, and any regulatory effects of RNA 2 are not present. This could be the source of the increased amount of minus strands found in cells transfected with the replicon; the lack of the second molecule, which contains the ORF for the capsid protein, prevents the replicase from efficiently transitioning from replication to transcription.

We developed a method to detect mature miR-34a molecules produced by the NodaMiRNA replicons using RT-qPCR. This method was both accurate and background-free, as the microRNA produced was exogenous to the BHK-21 cells used for the studies. Both NodaMiRNA constructs were able to release microRNAs via splicing of the replicon molecule. Interestingly, even though the minus

strands accumulated to high values at early time points for both Nod (+) and Nod(-), the amount of microRNAs produced from Nod(-) was orders of magnitude less than that produced by Nod(+). Only at 72 hours does the amount of microRNA produced by Nod(-) exceed that of Nod(+). However, both constructs were capable of releasing their microRNA cargo to the transfected cells, and were capable of doing so in large numbers. Nod(+) released millions of mature microRNAs to transfected BHK-21 cells, far exceeding the normal distribution of miR-34a within tissue[108]. Nod(-) released far fewer mature molecules, but in doing so, still managed to restore the microRNA to "biological" levels.

Both NodaMiRNA constructs showed evidence of time-dependent alternative splicing, with more correctly spliced microRNA being produced at later time points. Nod(+) showed a small fraction of its RNA to be alternatively spliced, resulting in a slightly larger RNA sequence than anticipated. This was confirmed via melting-curve analysis, which showed an increase in the width of the curve and an increase in Tm at early time points of Nod(+) pseudoinfection. This shoulder disappeared after 72 hours, yielding a single peak at the expected Tm. The alternate splicing seen in Nod(-) is even more pronounced, with a significant secondary peak seen in the early time-point melt-curves. These secondary melt-peaks also disappeared by 72 hours. It is suspected that the differences in RNA splicing is a result of the unique environment seen in NodaMiRNA replication. The replication complexes seen in the life-cycle of the replicon surely change the activity of the ribozymes used to generate the microRNA. Either though crowding or alternative base pairing, the presence of both strands of the replicon genome in close proximately invariably alters the activity of these ribozymes. The minus strands are never exported from the replication complexes, and therefore cannot fold into the proper structure. The positive-sense genomes, however are exported from the complexes and can then fold into the correct, catalytically active, structure. This also accounts for the eventual increase in minus strand splicing; at the end of the life-cycle, both plus strands and replication complexes are degraded, which then allows the

minus strands to fold correctly and produce microRNA.

Biological efficacy

The NodaT2A replicon itself slowed the spread of transfected PC3 cells in culture. Exposure to a replicon that did not express a microRNA still prevented the transfected cells from actively dividing *in vitro*. This attenuation is suspected to be due either to the remodeling of internal mitochondria, to the depletion of cellular cofactors, or both. Nodamura is known to remodel mitochondria, and the dynamics of these organelles is closely associated with apoptosis and cell division. Many other cytoplasmic viruses are apoptosis-inducing, often due to their interactions with the mitochondria, and by altering the activity within them this replicon could easily slow down cell-division.

Both NodaMiRNA constructs were effective in slowing the spread of PC3 cells in culture, more so than the replicon alone, indicating that the microRNA released from the self-replicating RNA molecule was biologically active. In addition, both constructs induced much more apoptosis than the replicon alone, indicating again the biological effectiveness of these constructs. MiR-34a is known to be involved in both cell-division as well as apoptosis, and the increase in apoptosis and slowing of cell division is strong evidence that the molecules produced by the activity of the ribozymes are catalytically active. There was some concern that these spliced small RNA molecules would not be able to diffuse from the replication complexes, and would not be biologically active. In addition, we did not know if the construct would work in the minus strand. Previous studies that incorporated the preMiR-34a duplex into the genome of an ssRNA virus found that it did not work in the minus sense, because the minus-strand could not leave the replication complex and interact with DICER. We avoid this problem by generating the microRNA in a DICER independent manner.

Conclusions

In many ways, it is surprising that the NodaMiRNA constructs work *in vivo*, as it depends on many disparate factors working. For the CCS to function, both the hammer head ribozyme as well as the Group 1 intron must splice efficiently, which is made difficult by the fact that they are both dependent on both local and global secondary structures. It was unknown if the extended secondary structures found within the nodaT2A replicon would also interfere with the splicing of the replicons *in vitro* and *in vivo.*

We were more surprised at the overall efficiency of these ribozymes in splicing co-transcriptionally, to the point that we spent a non-trivial amount of time attempting to inhibit the splicing of the Nod(+) replicon during transcription. Many inhibitors have been shown for both group 1 introns and hammerhead ribozymes, including arginine[88] and neomycin[85]. However, many of these inhibitors also inhibit transcription, and were useless to us. While we able to somewhat successfully inhibit the splicing with the use of anti-sense DNA oligos, these too reduced overall transcriptional yields, which we suspect is due to the chelation of catalytically needed magnesium ions. There has been little research into inhibiting ribozymes using small molecules, and this area of research should be further explored so that there exist many ways to inhibit ribozymes, which will certainly aid in structural studies. In addition, more work on co-transcriptional splicing is needed, as almost all studies use artificially derived “trans” ribozymes to study these catalysts.

Our initial concern that the delivery of splice products in Nod(+) transfection would limit the ability of the replicon to replicate *in vivo* led us directly to the discovery that the replicon’s overall protein expression is attenuated by higher input doses of the replicon itself. We feel confident that this is due to the initial overexpression of replicase molecules during initial translation, which may cause the cell to shut down protein synthesis or other cellular pathways needed by the replicon molecule. The

finding is interesting in that it is the first example in the literature of an inverse-dose-dependence for a transfected RNA molecule. This unique aspect of our replicon system is actually a powerful tool for working around the problem of delivery. One of the main problems with many gene-delivery systems is that one needs to target a large number of the molecules to a cell in order to generate a large amount of gene products. However, with the NodaT2A system, we are actually *penalized* for delivering too many RNA molecules to the cell; ideally we would prefer a system of gene delivery that is *highly selective* and overall inefficient, such that we deliver a small number of molecules to a specific class of cells. It should be noted that this fact also makes the replicon very prone to off-target activity, as it only takes a small number of molecules to establish a pseudoinfection in non-cancerous cells.

The amount of replicon accumulation in cells indicates that the NodaT2A replicon is capable of replicating faster than the rates of ribozyme splicing *in vivo.* This is demonstrated by both replicons replicating to high numbers when transfected. If the rate of splicing were faster than the rate of replication, the NodaMiRNA constructs would be incapable of replicating because spliced RNA molecules are replication incompetent. We also learn from Nod(+) that the presence of potentially inhibitory splice products does not prevent the replicon from working, but may attenuate the overall replication of the RNA molecule. Inclusions of these ribozymes also does little to perturb the overall lifecycle of the replicon, which indicates that they did not affect the activity of the replicase in detectable ways.

The surprising result that while Nod(-) makes more RNA, it make orders of magnitude less miR-34a, is easily explained by the nature of viral replication. Because the Nod(-) RNA is both crowded within the replication complex and base-paired extensively to the positive-sense genome, it is not able to effectively fold into the correct catalytic conformation needed for activity. The gradual accumulation of microRNA seen in cells transfected with Nod(-) is certainly evidence that at late time-points the minus strand can fold correctly and splice. We suspect that this may be due to the fact that at late time-points,

the replication complexes are degrading and the minus strands are then free to fold in the cytoplasm of the cell, away from crowding agents and +ssRNA that could interfere with the activity of the ribozymes.

Most importantly, both NodaMiRNA constructs inhibit cancer cell proliferation and induce apoptosis. This indicates that the microRNA is behaving as it should, and regulating protein production. That these catalytically derived miRNAs inhibit both cell spread and induce apoptosis (two fundamentally different cellular processes) indicate the broad importance of miR-34a in regulating the cancerous state. In addition, it means the microRNAs produced within replication complexes are capable of entering the cytoplasm of the cell. Other works using single-stranded RNA viruses to deliver functional RNAs focused on the inclusion of pre-miRNA hairpins[51] into the genome of the virus[50]. These hairpins must first be digested by DICER, prior to insertion into the RISC complex. Minus strands containing the hairpin do not leave replication complexes, and therefore were not incorporated into the RISC complex. The NodaMiRNA replicon system is the first such system that is functional in either strand of the replicon.

Future Work

How much is too much? attenuating replicon effects to increase persistence

Our RT-PCR studies have shown that our replicon is very good at making many copies of itself within transfected cells. In fact, it replicates to levels similar to rRNA. This of course raises the question as to whether or not this level of replication is *necessary.* Compared to endogenous microRNA levels, the Nod(+) construct drastically over-produces our molecule of interest, and while this may be good for cancer treatment a subtler approach may be needed if this system is to be extended to other treatment

avenues (see below.) Often times, one doesn't need 1,000,000 copies of the gene, but – rather – 1,000 will suffice, and this gene should be expressed at a constant slow rate, which is *not* seen in viral infections.

There is a direct correlation between the total amount of viral replication and the duration of infection. In other words, the more protein a virus makes, the faster the infection runs its course. Outside of any influence from its host, a virus has its own lifecycle, and eventually replication ceases. However, certain viruses can change expression profiles and become persistent. Persistent viral infections remain "hidden" producing sufficient RNA and protein to survive, without producing so much viral gene product that the cell dies. There are many avenues to generate persistent replicons, including directed evolution as well as incorporating miRNA targets into the replicon molecule, but these all serve to attenuate the production of RNA so that the replicon is more persistent[109]. This is appropriate for miRNA delivery because the amount of miRNA needed is relatively low, due to its catalytic nature.

Building a better replicon: temperature effects and cell types

The relationship between cell type, temperature and the activity of this replicon must be explored in more detail. The replicon only works in a select set of cell types. While we have seen robust replication in BHK-21 and HEK cells, it is more difficult to establish pseudoinfections in other cell lines. PC3 prostate cancer cells will not allow for robust pseudoinfection at 37˚C, and are only capable of supporting replication at. Human liver cancer cells (HEP-G2 cells) were also not readily capable of being pseudoinfected.

Clearly it is a requirement that this system be capable of working in a variety of cells types, and more importantly at biological temperatures. While the low-temperature activation of this replicon makes it an ideal system for the treatment of certain types of cancers (such as skin and testicular), it

must be able to work at 37˚C to be an effective general treatment. Again it is possible to generate temperature-insensitive mutants via directed evolution, and then to sequence these mutants to determine what residues have changed during infection. These replicons are ideal candidates for directed evolution due to their high mutation rates[8].

Beyond cancer: delivery of assorted self-amplifying functional RNAs (SAfuRNAs)

The success of this NodaMiRNA system in both preventing the spread of cancer, as well as inducing apoptosis, indicates that the microRNA produced interacts as expected with the microRNA processing machinery. This machinery goes beyond the canonical RISC complex, and includes other regulatory systems, such as the RNA-induced *transcriptional* silencing complex, which is responsible for the spreading of heterochromatin within the nucleus. Because the NodaMiRNA system is capable of producing functional miRNAs without recruiting DICER, this means that it can deliver any number of non-coding RNAs that are not processed by DICER. These include potent regulatory molecules such as snoRNAs[110] as well as long-non-coding RNAs[48].

The NodaT2A replicon is inherently transient, and as such it is important to use it to deliver long-lasting biological effectors. Using the replicon to deliver gene products certainly produces a large amount of transgene expression, but it is transient. It is more useful, then, to delivery molecules that change the regulatory pathways of the cells, as that is longer lasting and more effective. In other words, the NodaMiRNA system could also be used for epigenetic regulation. This type of regulation is much more persistent than simple gene-delivery. For example, one could deliver an siRNA to knock-down an oncogene (which would be useful but transient) or one could deliver a lncRNA that changes the *endogenous expression* of the oncogene without mutating the hosts genome, effectively inducing long-term genetic changes with a transient expression system.

Alternative splicing in alternative environments: ribozymes as secondary structure sensor

We believe the major source of differences in splicing efficiency seen between the NodaMiRNA constructs is the location in which the CCS is replicated. For Nod(+) replication, the catalytically active molecules, the plus-sense viral genomes, are exported *out* of the replication complexes, and into the cytoplasm. In the viral infection, this is an important step, as the assembly of the mature virion must occur outside these complexes. These exported plus-strands are not confined within the small replication complexes, they are not crowded by macromolecules, and they are not extensively based-paired with their minus-strand brethren. As such, the Nod(+) molecules are free to fold and unfold extensively, thus increasing the amount of spliced microRNA produced.

From these experiments we can see that these ribozymes are exquisite sensors of both local and global secondary structures, and can even provide some insight into the degree of confinement or crowding experienced by an RNA molecule in solution. By attaching a slow-splicing ribozyme to the end of an RNA molecule, one might be able to understand its local environment by studying the alternative splicing and slicing rate of the ribozyme. Here the ribozyme would serve as a "tag" that would detail how "free" the RNA molecule is to fold and unfold in the cytoplasm. One could then couple this with a riboswitch, which would allow one to gain information about what types of small molecules are within these replication complexes.

References

1. Suttle, C. A. Viruses in the sea. *Nature* **437,** 356–361 (2005).

2. Suttle, C. A. Marine viruses - Major players in the global ecosystem. *Nat. Rev. Microbiol.* **5,** 801–812 (2007).

3. Lane, N. & Martin, W. The energetics of genome complexity. *Nature* **467,** 929–934 (2010).

4. Rajagopala, S. *et al*. The protein interaction map of bacteriophage lambda. *BMC Microbiol.* **11,** 213 (2011).

5. Ogawa, T. & Tomizawa, J. Replication of bacteriophage DNA. I. Replication of DNA of lambda phage defective in early functions. *J. Mol. Biol.* **38,** 217–25 (1968).

6. Speir, J. A., Munshi, S., Wang, G., Baker, T. S. & Johnson, J. E. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure* **3,** 63–78 (1995).

7. Kaesberg, P. *et al.* Structural homology among four nodaviruses as deduced by sequencing and X-ray crystallography. *J. Mol. Biol.* **214,** 423–435 (1990).

8. Sanjuan, R. *et al.* Viral Mutation Rates. *J. Virol.* **84,** 9733–9748 (2010).

9. Lindell, D. *et al.* Transfer of photosynthesis genes to and from Prochlorococcus viruses. *Proc. Natl. Acad. Sci.* **101,** 11013–11018 (2004).

10. Baltimore, D. Expression of animal virus genomes. *Bacteriol. Rev.* **35,** 235–241 (1971).

11. Coutinho, F. H. *et al.* Marine viruses discovered via metagenomics shed light on viral strategies throughout the oceans. *Nat. Commun.* **8,** 1–12 (2017).

12. CASPAR, D. L. & KLUG, A. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* **27,** 1–24 (1962).

13. Manuscript, A. *Viral Molecular Machines*. **726,** (2012).

14. Cadena-Nava, R. D. *et al.* Self-Assembly of Viral Capsid Protein and RNA Molecules of Different Sizes: Requirement for a Specific High Protein/RNA Mass Ratio. *J. Virol.* **86,** 3318–3326 (2012).

15. Garmann, R. F. *et al.* The assembly pathway of an icosahedral single-stranded RNA virus depends

on the strength of inter-subunit attractions. *J. Mol. Biol.* **426,** 1050–1060 (2014).

16. Tubiana, L. *et al*. Synonymous mutations reduce genome compactness in icosahedral ssRNA viruses. *Biophys. J.* **108,** 194–202 (2015).

17. Bednar, J. *et al.* Determination of DNA persistence length by cryo-electron microscopy. Separation of the static and dynamic contributions to the apparent persistence length of DNA. *J. Mol. Biol.* **254,** 579–594 (1995).

18. Šimoliunas, E. *et al.* Klebsiella Phage vB_KleM-RaK2 - A Giant Singleton Virus of the Family Myoviridae. *PLoS One* **8,** (2013).

19. Bauer, D. *et al*. Herpes virus genome, the pressure is on. *J. Am. Chem. Soc.* **135,** 11216–11221 (2013).

20. Chen, C. & Guo, P. Magnesium-induced conformational change of packaging RNA for procapsid recognition and binding during phage phi29 DNA encapsidation. *J. Virol.* **71,** 495–500 (1997).

21. Hernando-Pérez, M. *et al.* Direct measurement of phage phi29 stiffness provides evidence of internal pressure. *Small* **8,** 2366–2370 (2012).

22. Fuller, D. N. *et al.* Measurements of Single DNA Molecule Packaging Dynamics in Bacteriophage λ Reveal High Forces, High Motor Processivity, and Capsid Transformations. *J. Mol. Biol.* **373,** 1113–1122 (2007).

23. Erdemci-Tandogan, G., Orland, H. & Zandi, R. RNA Base Pairing Determines the Conformations of RNA Inside Spherical Viruses. *Phys. Rev. Lett.* **119,** 1–5 (2017).

24. Chen, H. *et al.* Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc. Natl. Acad. Sci.* **109,** 799–804 (2012).

25. Breitbart, M. *et al.* Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci.* **99,** 14250–14255 (2002).

26. Drake, J. W. & Holland, J. J. Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. U. S. A.* **96,**

13910–3 (1999).

27. Lindenbach, B. D., Sgro, J. Y. & Ahlquist, P. Long-distance base pairing in flock house virus RNA1 regulates subgenomic RNA3 synthesis and RNA2 replication. *J Virol* **76,** 3905–19. (2002).

28. Rosskopf, J. J. *et al.* A 3' terminal stem-loop structure in Nodamura virus RNA2 forms an essential cis-acting signal for RNA replication. *Virus Res.* **150,** 12–21 (2010).

29. Allison, R.*et al.* Regeneration of a functional RNA virus genome by recombination between deletion mutants and requirement for cowpea chlorotic mottle virus 3a and coat genes for systemic infection. *Proc. Natl. Acad. Sci. U. S. A.* **87,** 1820–4 (1990).

30. Eckerle, L. D. & Ball, L. A. Replication of the RNA Segments of a Bipartite Viral Genome Is Coordinated by a Transactivating Subgenomic RNA. *Virology* **296,** 165–176 (2002).

31. Johnson, K. *et al*. Nodamura Virus Nonstructural Protein B2 Can Enhance Viral RNA Accumulation in both Mammalian and Insect Cells Nodamura Virus Nonstructural Protein B2 Can Enhance Viral RNA Accumulation in both Mammalian and Insect Cells. **78,** 6698–6704 (2004).

32. Rao, A. L. N. & Cheng Kao, C. The brome mosaic virus 3’ untranslated sequence regulates RNA replication, recombination, and virion assembly. *Virus Res.* **206,** 46–52 (2015).

33. Yi, G. H. *et al.* Selective repression of translation by the brome mosaic virus 1a RNA replication protein. *J. Virol.* **81,** 1601–1609 (2007).

34. Ball, L. *et al*. Replication of nodamura virus after transfection of viral RNA into mammalian cells in culture. *J. Virol.* **66,** 2326–2334 (1992).

35. Li, Y. & Ball, L. a. Nonhomologous RNA recombination during negative-strand synthesis of flock house virus RNA. *J. Virol.* **67,** 3854–3860 (1993).

36. den Boon, J. A. & Ahlquist, P. Organelle-Like Membrane Compartmentalization of Positive-Strand RNA Virus Replication Factories. *Annu. Rev. Microbiol.* **64,** 241–256 (2010).

37. Gant, V. U. *et al.* Two Membrane-Associated Regions within the Nodamura Virus RNA-Dependent

RNA Polymerase Are Critical for both Mitochondrial Localization and RNA Replication. *J. Virol.* **88,** 5912–5926 (2014).

38. Blanchard, E. & Roingeard, P. Virus-induced double-membrane vesicles. *Cell. Microbiol.* **17,** 45–50 (2015).

39. Gopal, A. *et al*. Visualizing large RNA molecules in solution. *Rna* **18,** 284–299 (2012).

40. Kato, F. & Hishiki, T. Dengue virus reporter replicon is a valuable tool for antiviral drug discovery and analysis of virus replication mechanisms. *Viruses* **8,** 1–11 (2016).

41. Maharaj, P. D. *et al.* Nanoparticle encapsidation of Flock House virus by auto assembly of Tobacco Mosaic virus coat protein. *Int. J. Mol. Sci.* **15,** 18540–18556 (2014).

42. Deyle, D. R. & Russell, D. W. Adeno-associated virus vector integration. *Curr. Opin. Mol. Ther.* **11,** 442–7 (2009).

43. Frolov, I. *et al.* Selection of RNA replicons capable of persistent noncytopathic replication in mammalian cells. *J. Virol.* **73,** 3854–65 (1999).

44. Nordström, E. K. L. *et al.* Enhanced immunogenicity using an alphavirus replicon DNA vaccine against human immunodeficiency virus type 1. *J. Gen. Virol.* **86,** 349–354 (2005).

45. Lundstrom, K. Alphavirus-based vaccines. *Methods Mol. Biol.* **1581,** 225–242 (2017).

46. Näslund, T. I. *et al.* Role of innate signalling pathways in the immunogenicity of alphaviral replicon-based vaccines. *Virol. J.* **8,** 1–9 (2011).

47. Zuo, L. *et al*. piRNAs and their functions in the brain. *Int. J. Hum. Genet.* **16,** 53–60 (2016).

48. Kung, J. T. Y., Colognori, D. & Lee, J. T. Long Noncoding RNAs: Past, Present, and Future. *Genetics* **193,** 651–669 (2013).

49. Guo, Y. *et al.* Characterization of the mammalian miRNA turnover landscape. *Nucleic Acids Res.* **43,** 2326–2341 (2015).

50. Rouha, H. *et al*. Functional microRNA generated from a cytoplasmic RNA virus. *Nucleic Acids Res.*

38, 8328–8337 (2010).

51. Varble, A. *et al.* Engineered RNA viral synthesis of microRNAs. *Proc. Natl. Acad. Sci.* **107,** 11519–11524 (2010).

52. Miller, D. J. *et al.* Flock House Virus RNA Replicates on Outer Mitochondrial Membranes in Drosophila Cells. *Proc. Natl. Acad. Sci.* **75,** 11664–76 (2001).

53. Fenner, B. J. *et al.* Betanodavirus B2 Is an RNA Interference Antagonist That Facilitates Intracellular Viral RNA Accumulation. *J. Virol.* **80,** 85–94 (2006).

54. Bothner, B. *et al.* Evidence of Viral Capsid Dynamics Using Limited Proteolysis and Mass Spectrometry. *J. Biol. Chem.* **273,** 673–676 (1998).

55. Kramps, T. & Elbers, K. *Introduction to RNA vaccines*. *Methods in Molecular Biology* **1499,** (2017).

56. Ball, L. A. Requirements for the self-directed replication of flock house virus RNA 1. *J. Virol.* **69,** 720–7 (1995).

57. Gitlin, L. *et al.* Rapid Evolution of Virus Sequences in Intrinsically Disordered Protein Regions. *PLoS Pathog.* **10,** (2014).

58. Nam, J. W. *et al.* Human microRNA prediction through a probabilistic co-learning model of sequence and structure. *Nucleic Acids Res.* **33,** 3570–3581 (2005).

59. Dostie, J. *et al*. Numerous microRNPs in neuronal cells containing novel microRNAs Numerous microRNPs in neuronal cells containing novel microRNAs. 180–186 (2003). doi:10.1261/rna.2141503.which

60. Shimada, Y., Mohn, F. & Bühler, M. The RNA-induced transcriptional silencing complex targets chromatin exclusively via interacting with nascent transcripts. *Genes Dev.* **30,** 2571–2580 (2016).

61. Dana, H. *et al.* Molecular Mechanisms and Biological Functions of siRNA. *Int. J. Biomed. Sci.* **13,** 48–57 (2017).

62. Han, J. *et al.* The Drosha – DGCR8 complex in primary microRNA processing. *Genes Dev.* 3016–

3027 (2004). doi:10.1101/gad.1262504.mic

63. Pratt, A. J. & MacRae, I. J. The RNA-induced silencing complex: A versatile gene-silencing machine. *J. Biol. Chem.* **284,** 17897–17901 (2009).

64. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115,** 209–216 (2003).

65. Schwarz, D. S. *et al.* Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115,** 199–208 (2003).

66. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. **431,** 343–349 (2004).

67. Petersen, C. P. *et al* P. A. Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell* **21,** 533–542 (2006).

68. Sen, G. L. & Blau, H. M. Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nat. Cell Biol.* **7,** 633–636 (2005).

69. Peng, Y. & Croce, C. M. The role of MicroRNAs in human cancer. *Signal Transduct. Target. Ther.* **1,** 15004 (2016).

70. Dedes, K. J. *et al.* Down-regulation of the miRNA master regulators Drosha and Dicer is associated with specific subgroups of breast cancer. *Eur. J. Cancer* **47,** 138–150 (2011).

71. Lodygin, D. *et al.* Inactivation of miR-34a by aberrant CpG methylation in multiple types of cancer. *Cell Cycle* **7,** 2591–2600 (2008).

72. Lim, L. P. *et al*. Vertebrate microRNA genes. *Science (80-.).* **299,** 1540 (2003).

73. Houbaviy, H. B., Murray, M. F. & Sharp, P. A. Embryonic stem cell-specific MicroRNAs. *Dev. Cell* **5,** 351–8 (2003).

74. Misso, G. *et al.* Mir-34: A new weapon against cancer? *Mol. Ther. - Nucleic Acids* **3,** e194 (2014).

75. Yamakuchi, M., Ferlito, M. & Lowenstein, C. J. miR-34a repression of SIRT1 regulates apoptosis.

Proc. Natl. Acad. Sci. **105,** 13421–13426 (2008).

76. Liu, C. *et al.* The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. *Nat. Med.* **17,** 211–216 (2011).

77. Li, N. *et al.* miR-34a inhibits migration and invasion by down-regulation of c-Met expression in human hepatocellular carcinoma cells. *Cancer Lett.* **275,** 44–53 (2009).

78. Tazawa, H.*et al* growth arrest through modulation of the E2F pathway in human colon cancer cells. *Pnas* **104,** 2–7 (2007).

79. He, L. *et al.* A microRNA component of the p53 tumour suppressor network. *Nature* **447,** 1130–1134 (2007).

80. Navarro, F. & Lieberman, J. miR-34 and p53: New insights into a complex functional relationship. *PLoS One* **10,** 1–23 (2015).

81. Chang, T. C. *et al.* Transactivation of miR-34a by p53 Broadly Influences Gene Expression and Promotes Apoptosis. *Mol. Cell* **26,** 745–752 (2007).

82. Hammann, C., Luptak, A., Perreault, J. & de la Pena, M. The ubiquitous hammerhead ribozyme. *Rna* **18,** 871–885 (2012).

83. Strobel, S. A. *et al.* Complementary sets of noncanonical base pairs mediate RNA helix packing in the group I intron active site. *Nat. Struct. Biol.* **5,** 60–66 (1998).

84. Pley, H. W. *et al.* Three-dimensional structure of a hammerhead ribozyme. *Nature* **372,** 68–74 (1994).

85. Stage, T. K. *et al.* Inhibition of the hammerhead ribozyme by neomycin. *RNA (New York, N.Y.)* **1,** 95–101 (1995).

86. Scott, W. G. *et al*. Capturing the structure of a catalytic RNA intermediate: The hammerhead ribozyme. *Science (80-.).* **274,** 2065–2069 (1996).

87. Zhang, L. *et al* Antisense oligonucleotides effectively inhibit the co-transcriptional splicing of a

Candida group I intron in vitro and in vivo: Implications for antifungal therapeutics. *FEBS Lett.* **583,** 734–738 (2009).

88. Liu, Y. & Leibowitz, M. J. Bidirectional effectors of a group I intron ribozyme. *Nucleic Acids Res.* **23,** 1284–1291 (1995).

89. Ma, B. *et al.* Lipoplex morphologies and their influences on transfection efficiency in gene delivery. *J. Control. Release* **123,** 184–194 (2007).

90. Tros de Ilarduya, C. *et al.* Gene delivery by lipoplexes and polyplexes. *Eur. J. Pharm. Sci.* **40,** 159–170 (2010).

91. Zuhorn, I. S. *et al*. Lipoplex-mediated transfection of mammalian cells occurs through the cholesterol-dependent clathrin-mediated pathway of endocytosis. *J. Biol. Chem.* **277,** 18021–18028 (2002).

92. Avital, Y. *et al.* The thermodynamics of endosomal escape and DNA release from lipoplexes. *Phys. Chem. Chem. Phys.* 2591–2596 (2015).

93. Dowty, M. E. *et al*. Plasmid DNA entry into postmitotic nuclei of primary rat myotubes. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 4572–4576 (1995).

94. Brunner, S. *et al.* Cell cycle dependence of gene transfer by lipoplex, polyplex and recombinant adenovirus. *Gene Ther.* **7,** 401–407 (2000).

95. Felgner, P. L. *et al.* Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci.* **84,** 7413–7417 (1987).

96. Support, I. T. Lipofectamine 2000 Transfection Reagent. 1–7 (2005).

97. Leonhardt, C. *et al.* Single-cell mRNA transfection studies: Delivery, kinetics and statistics by numbers. *Nanomedicine Nanotechnology, Biol. Med.* **10,** 679–688 (2014).

98. Shi, R. & Chiang, V. L. Facile means for quantifying microRNA expression by real-time PCR. *Biotechniques* **39,** 519–524 (2005).

99. Ertel, K. J. *et al.* Cryo-electron tomography reveals novel features of a viral rna replication compartment. *Elife* **6,** 1–24 (2017).

100. Nomura-Takigawa, Y. *et al.* Non-structural protein 4A of Hepatitis C virus accumulates on mitochondria and renders the cells prone to undergoing mitochondria-mediated apoptosis. *J. Gen. Virol.* **87,** 1935–1945 (2006).

101. Samson, M. *et al.* HHS Public Access. **155,** 3–12 (2017).

102. Khan, M. *et al*. Mitochondrial dynamics and viral infections: A close nexus. *Biochim. Biophys. Acta - Mol. Cell Res.* **1853,** 2822–2833 (2015).

103. Chen, L. J. *et al.* Betanodavirus non-structural protein B1: A novel anti-necrotic death factor that modulates cell death in early replication cycle in fish cells. *Virology* **385,** 444–454 (2009).

104. O. Shaul, M. V. M. & D. I. *Cell Cycle Control*. *Annals of Botany* **78,** (1996).

105. Korenaga, M. *et al.* Mitochondrial dysfunction in hepatitis C. *J. Clin. Gastroenterol.* **39,** S162-6 (2005).

106. Mishra, P. & Chan, D. C. Mitochondrial dynamics and inheritance during cell division, development and disease. *Nat. Rev. Mol. Cell Biol.* **15,** 634–646 (2014).

107. Wai, T. & Langer, T. Mitochondrial Dynamics and Metabolic Regulation. *Trends Endocrinol. Metab.* **27,** 105–117 (2016).

108. Abdelrahim, M. *et al* RNAi and Cancer: Implications and applications. *J RNAi Gene Silenc.* **2,** 136–145 (2006).

109. Kamrud, K. I. *et al.* In vitro and in vivo characterization of microRNA-targeted alphavirus replicon and helper RNAs. *J. Virol.* **84,** 7713–7725 (2010).

110. Stepanov, G. A. *et al.* Regulatory Role of Small Nucleolar RNAs in Human Diseases. **2015,** (2015).