2023

Investigating miRNA Regulation of the Human APOBEC3 Enzymes

William Dietrich
Minnesota State University, Mankato

Investigating miRNA regulation of the human APOBEC3 enzymes

By
William Dietrich

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of
Master of Science
In
Biological Sciences

Minnesota State University, Mankato,
Mankato, Minnesota
April 2023

April 2023

Investigating miRNA regulation of the human APOBEC3 enzymes

William Dietrich

This thesis has been examined and approved by the following members of the student's committee.

_____

Advisor

_____

Committee Member

_____

Committee Member

Acknowledgments:

A very big thanks to my advisor and mentor, Dr. Allison M. Land for all the support, knowledge, and kindness given to me. She helped foster my love of science and was crucial in the making and completing of this project. Thanks to my committee members Dr. David Sharlin, and Dr. Keenan Hartert for their knowledge and support. I would like to thank the Department of Biology and the College of Graduate Studies and Research for their financial contribution and opportunity to support me in my degree. I would also like to thank Joshua Vetter for his help in the lab, Hunter Herber, Joelle Anderson, and Madeleine Kennedy for their work on discovering and creating miRNAs, and everyone in Lands lab and RISEbio. Thank you, to all these people along with all the faculty in the Department of Biology for their kindness, teaching and support given to me throughout my tenure at Minnesota State University Mankato.

Table of Contents

Acknowledgements ii

Page

Abstract 1

Chapters

1. Introduction 2
    a. APOBEC3 3
    b. Cancer and APOBEC3s 9
    c. MiRNAs 11
    d. MiRNAs and APOBEC3s 14
    e. Experimental summary 14
2. Methods 15
    a. Identifying candidate miRNAs and their sequences 15
    b. Sequencing miRNAs in pMiniT 2.0 vectors 16
    c. Cloning miRNAs into pcDNA-3.1 expression vectors 17
    d. Cell culture conditions 17
    e. Optimization of transfection conditions 18
    f. Testing miRNA constructs 19
    g. Data analysis 20
3. Results 21
    a. Identifying candidate miRNAs 21
    b. Optimization of dual luciferase assay 21

c. Testing A3 3’ UTR constructs expression 24

d. Testing miRNA constructs 25

4. Discussion 35

5. References 40

Abstract

The human apolipoprotein B mRNA editing enzyme catalytic polypeptide 3 (APOBEC3, A3) are a family of proteins consisting of seven enzymes, A3A, B, C, D, F, G and H, which function as cytosine deaminases. The enzymes' purpose in the cell is to mutate viral DNA during infection hindering or stopping replication of viruses such as human papillomavirus, herpes simplex virus, and HIV-1. Several of the A3 enzymes have also been implicated in contributing to cancers such as head and neck and breast cancers by mutating cellular genomic DNA, making the ability to control A3 expression an attractive target for cancer therapy. The A3 enzymes are regulated by small fragments of RNA called miRNA. In this thesis, twenty-four miRNAs were identified for their predicted silencing of A3 mRNA, cloned into expression vectors, and tested in a dual luciferase assay to determine their silencing ability. Three miRNAS, *hsa-miR-1207-5p*, *hsa-miR-1227-3p,* and *hsa-miR-548a-5p,* indicated a possible silencing interaction and I was able to show a significant silencing of A3C by *hsa-miR-1207-5p.* This thesis contributes to the study of a family of human immune proteins that are not only important in the innate immunity against viruses that infect millions of people each year, but are also molecular drivers of cancers. This research advances our understanding on miRNA control of the human APOBEC3 family and the potential for miRNA to be used as therapeutic agents against disease.

Introduction

Human cells have evolved numerous ways to defend themselves against viral invaders, one example is the human APOBEC3 (A3) class of enzymes. The A3 enzymes are part of the cell's innate immune response to viruses and are constantly expressed in many tissues at differing levels. When a susceptible virus invades the cell, the A3 enzymes are incorporated into viral particles and cause DNA damage by converting the cytosine in DNA to uracil. These mutations inhibit viral replication and stop the virus from infecting more cells. The A3 enzyme's ability to mutate DNA can also be problematic for our own cells. In cancers such as head and neck and breast cancer A3 enzymes are expressed at abnormally high levels. It is hypothesized that the high levels of A3 expression leads to a higher mutation rate of the cells genome and a faster progression of the cancer. This leads to a more severe outcome for the cancer patient.

The expression of A3 enzymes in the cell are controlled by microRNAs (miRNA). These small RNA fragments bind to the messenger RNA (mRNA) transcripts that encode for the protein and inhibit translation of the mRNA into protein. For a cell, miRNAs are essential for normal function, it is important for a cell to maintain control over the proteins it expresses and the ability to quickly inhibit translation of proteins allows the cell to respond to environmental changes. Since miRNAs are such an important part of protein expression, they need to be highly regulated to avoid protein mis-regulation. When disease strikes such as cancer, the expression of miRNAs can change which leads to abnormal expression of the proteins the miRNAs regulate such as A3 enzymes. However, using miRNAs as a way to control cellular protein expression could be a way to treat diseases with over expression of proteins, such as oncogenic proteins in cancer.

The goal of my research was to develop a library of miRNAs predicted to interact with A3 enzymes and test the miRNAs using a dual luciferase assay to determine the miRNA's ability

to interact with the A3 enzymes. This research would increase our understanding of how miRNA regulation of A3 enzymes works. This research would also serve as a potential avenue for future research in developing a therapeutic technique miRNA drug for those whose suffer from cancers with A3 involvement.

APOBEC3:

The human A3 enzyme family consists of seven enzymes: A3A, A3B, A3C, A3D, A3F, A3G and A3H, which are all encoded on chromosome 22 (1). These genes arose from gene duplication events in primate ancestors (2). The A3 enzyme genes belong to a large gene family which also include activation induced cytosine deaminase (AID), APOBEC1, APOBEC2, and APOBEC4. This large family of genes is highly conserved in vertebrates, with APOBEC1 and A3 enzymes being unique to mammals. AID, A3, and APOBEC1 all perform similar functions by

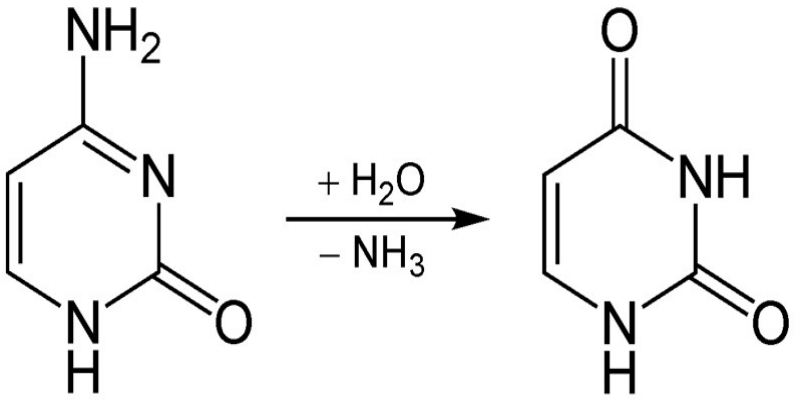

Figure 1. The process of cytosine (left) being deaminated into uracil (right). APOBEC3 enzymes are able to catalyze this reaction on mRNA strands (**71**)

acting as cytosine deaminases (2). Cytosine is one of the four bases found in DNA, the others being thymine, guanine, and adenine. Structurally, cytosine is very similar to another nucleotide found in RNA called uracil. In double stranded DNA or RNA, adenine binds to thymine or uracil in RNA and cytosine binds to guanine (3). Cytosine deaminases such as the APOBEC/AID enzymes bind

cytosine in DNA remove the amine group, replacing it with an oxygen molecule and converting it into uracil **Figure 1** (3). When the DNA is replicated, uracil templates for an adenine instead of the guanine that cytosine templates. This mutation can alter protein coding or regulatory function of the affected DNA. In genomic DNA, DNA repair machinery recognizes the uracil and reverts it back to cytosine. In heavily mutated genomic DNA however, the repair mechanisms can be overwhelmed, and permanent DNA damage can occur.

An enzyme that is related to the A3 enzymes and shares a similar cytosine deaminase activity is Activation-induced cytidine deaminase (AID). AID is an enzyme that is active in B cells during secondary antibody diversification during an immune response. AID causes random mutations in the genes that encode for antibodies and allows for a wide range of antibodies to be created (4). This is important for the adaptive immune response as it allows for antibodies to be selected that specifically bind the antigen the body is facing. This mutation does come with a risk of damaging the B cell's genome and can lead to B cell lymphomas. Since AID and APOBEC3 enzymes can cause mutations in cells due to their cytosine deaminase activity, it is important to study both as a breakthrough about one can lead to a breakthrough in the study of the other.

While all A3 enzymes share the same function of being cytosine deaminases, the enzymes differ in the preferred sequence context of the cytosine, and the number of deaminase protein domains (1,5). A3A, B, C, D, F, and H prefer to interact with cytosines that follow thymine while A3G prefers cytosines that occur after another cytosine. (3). A3A A3C, and A3H have a single cytosine deaminase domain, while A3 B, D, F, and G have two functional domains (3,6). These functional domains are called zinc coordinating domains (ZCD). In A3 ZCDs, a zinc ion is bound by histidine and two cysteines, and is able to interact with a cytosine, oxidizing it into an uracil (3).

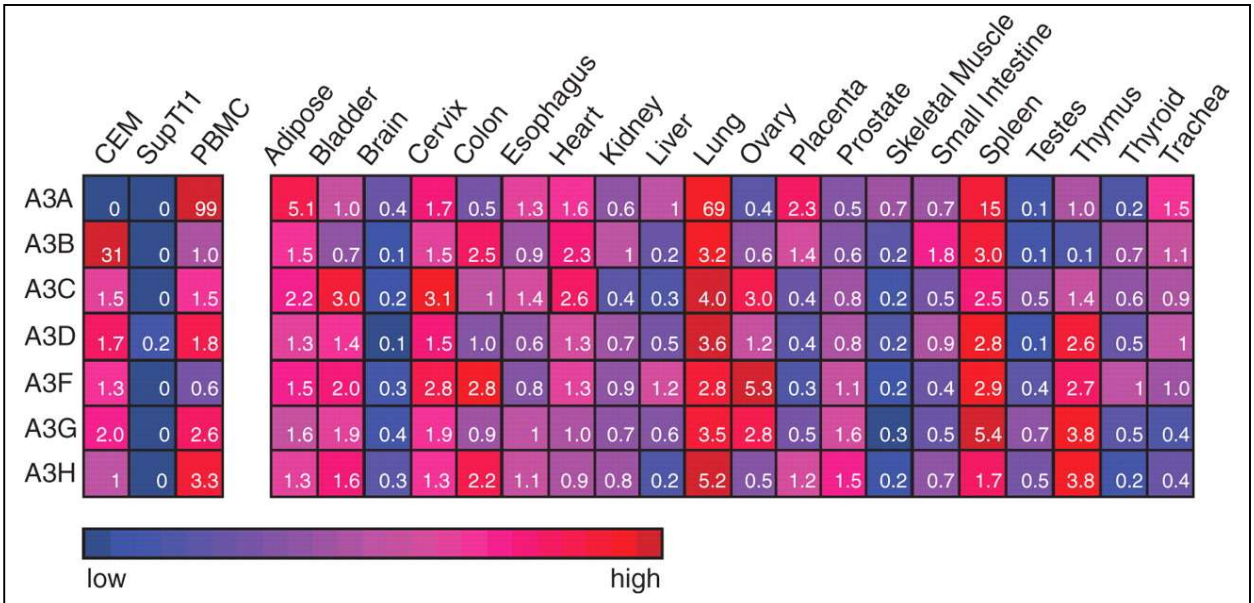

Figure 2. Expression of A3 enzymes in various tissues (7). Tissues with important immune functions and activity show high levels of A3 expression.

The A3 enzymes are widely expressed in many tissues at different levels of expression **Figure 2**, with immune tissues such as the spleen and peripheral blood mononuclear cells showing high levels of A3 expression. A3C is expressed in the most tissues, while A3A and A3G are highly expressed in lung tissue (4,7). The A3 enzymes are also found in differing locations inside the cell. A3B is found in the nuclear compartment of the cell, while A3A, H and C are found cell wide. A3D, F, and G are found in the cytoplasm of the cell (6,4,8,7).

The role of the A3 enzymes in our innate immune response is to restrict the replication of viruses. It is hypothesized that the evolution of the A3 gene family is closely tied to viral evolution, especially to the retroviruses that our primate ancestors faced (9). These viral evolution events have led to the existence of the seven members of the human A3 family. Each member while similar in function and structure exhibits differences as described below.

A3A is expressed in the nucleus and cytoplasm of cells. The A3A enzymes have been described to restrict the replication of DNA viruses such as human papillomavirus (HPV), herpes simplex virus type 1 (HSV-1), and hepatitis B virus (HBV) by causing high levels of mutations in the viral genomes (10,62). The role of A3A in restricting HPV is especially important as HPV infections are widespread in the United States, infecting nearly 78 million people and can lead to cancers affecting both men and women (11). In contrast to several other A3 members, A3A is not known to have a significant role in restricting HIV-1 replication. In addition to these beneficial, viral restriction roles, A3A has also been implicated in driving mutation in some cancers and has also been shown to damage DNA in mitochondria (62).

A3B is found only in the nucleus of cells. When a cell goes through mitosis A3B is rapidly imported into the new cells' nuclei, this location of A3B means it can defend against viral DNA within the nucleus (63). There is evidence that A3B inhibits the replication of HPV, HSV-1, and HIV-2. (12,64). Since A3B is found in the nucleus, A3B can access the cells genomic DNA and is the A3 protein most strongly linked to mutation in human cancers, most notably, breast cancer (21,65).

A3C is found throughout the cell and A3C along with A3F, A3G, and A3H can form structures that localize into specific subcellular structures. (2) These A3 enzymes can be found in two forms in the cell, an active low molecular mass form (LMM) and an inactive high molecular mass form (HMM). A3C binds with proteins and ribonucleotides in these structures and is inactivated until cytokine signals from viral infections activate it (2). There is evidence that A3C can restrict HPV, HIV-1, and human herpes viruses such as Epstein-Barr virus, however it is the weakest A3 enzyme against these viruses (67,68). A very small percentage of the human

population, however, contains a variant that is more effective as an antiviral protein, this mutation causes A3C to form a dimer with itself (67).

A3D has seven splice variants described, more than any other human A3 protein. These variants are found throughout the cell, with four containing functional cytosine deaminases (69). The A3Dv1 variant has two functional deaminases, Cytosine Deaminase 1 and 2 (CD1, CD2) while A3Dv2 only has CD1 and A3Dv6, and A3Dv7 contain CD2. All variants show antiviral activity against HIV-1 and foreign DNA restriction, however the antiviral activity of A3D is weaker compared to some other A3 enzymes (69). Hepatitis B and HIV-1 have been shown to be restricted due to A3D (2). Two of the variants A3Dv2 and A3Dv6 are found in the nucleus and can have the potential to mutate cellular DNA.

A3F is found in the cytoplasm of cells and like A3C can be found in LMM and HMM forms (2). A3F can be co-expressed with A3G in CD-4 positive immune cells, and when these cells are infected with HIV-1, both A3F and A3G can be found in the same viral particles (70). Inside the viral particle, A3F stimulates A3G activity by forming an oligomer with A3G, this new structure is a more active mutagen than either enzyme alone (70). While A3F is weaker in antiviral activity compared to A3G, A3F is more resistant to viral proteins such as Vif (70).

A3G is found in the cytoplasm of cells and like A3C and A3F can be found in LMM and HMM (2). A3G was the first A3 protein to be characterized as a viral restriction factor, for its ability to restrict the replication of HIV-1. HIV-1 is a virus that has infected 37 million people worldwide, 1.2 million people in the United States, and has killed 22 million people since the discovery of the virus (13,14). The role of A3G restriction of HIV-1 replication shown in **Figure 3**, is well characterized, and can be applied to the other A3 enzymes that restrict HIV-1. A3 enzymes target HIV-1 genetic material by being packaged in the viral particles as they begin to

assemble inside the cell. When the viral particle attaches and enters a new cell, the viral genome and A3 enzymes are released. The A3 enzymes can deaminate viral genome cytosine into uracil, and if enough mutations occur the virus may become defective and unable to replicate (15). HIV-1 can stop A3 from targeting its genome by producing viral infectivity factor (Vif). Vif is a protein that will attach itself to the A3 enzymes and mark them for destruction (12). Studies have shown that HIV with nonfunctional Vif have much lower infection rates and replication because of the A3 mediated inhibition. (12).

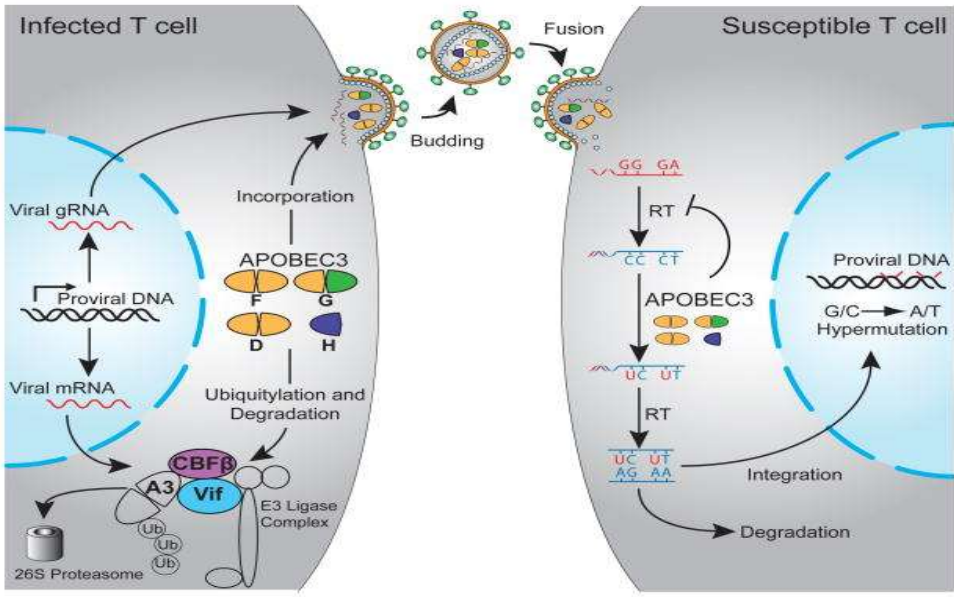

Figure 3 APOBEC3 enzymes limit HIV-1 replication. APOBEC3 enzymes of HIV infected cells are packaged into viral particles and travel with the virus to a new cell. When the virus enters the newly infected cell, A3 enzymes mutate cytosines to uracils on the viral cDNA causing damaging levels of genomic mutation. (15).

A3H is found in both the cytoplasm and nucleus of the cell, it can also be found in LMM and HMM forms in the cytoplasm (2,16). A3H has seven major haplotypes I-VII, as well as four different splice variants. (17,18,19**)** It is unclear why so many variations exist, but it is thought that geographic distribution of different populations and the viruses they faced led to the different

types. (17,18). Of the seven haplotypes, haplotype II, V, and VII are the best able to restrict HIV-1 that lacks Vif. Haplotypes I and VI could also have antiviral activity but they are poorly expressed in healthy cells (19). The other haplotypes are structurally unstable and do not exist in a biologically stable form in the cell at any appreciable levels. A3H haplotype I can appear in the cell nucleus while the other haplotypes appear in the cytoplasm of the cell (19).

Cancer and APOBEC3:

Cancer is characterized by the presence of cells that grow and divide rapidly without control. This uncontrolled growth can lead the formation of tumors and colonization of other tissues by the cancer. According to the American Cancer Society, cancer is the second most common cause of death in the United States after heart disease (20). Cancer is most common in tissues that have rapidly dividing cells such as skin or immune tissue; it is also common in tissues that interacts with harmful chemicals such as the liver and lung (20). Internal factors such as hormones can cause cancer as well as normal cell DNA replication events in where the cells machinery makes a mistake, turning off cancer stopping genes or turning on pro cancer genes. Environmental factors called carcinogens can also cause cancer. Examples of common carcinogens are tobacco smoke, radiation, ultraviolet light, lead, and processed meat. These factors can lead to the mutation of cellular DNA or interfere with cell growth and death signals, causing the cell to begin to divide uncontrollably.

Cells growth is not the only thing affected in cancer, often cellular processes such as protein regulation are also affected. In certain cancers such as neck and head and breast cancer, A3A and A3B enzymes have been shown to be present in abnormally high levels (21,22,23). It is hypothesized that these enzymes could be the cause of certain cancers or help the progression and mutagenesis of these cancers (21,22,23).

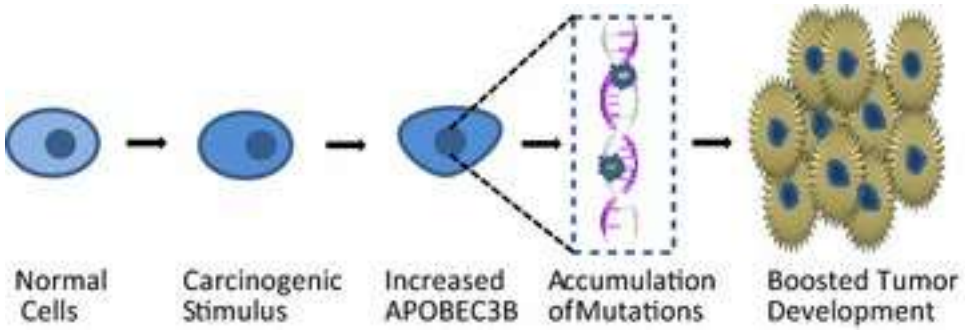

Figure 4. A hypothesized process on how APOBEC enzymes can lead to increase cellular DNA mutations and cancer development (22). During a disease such as cancer, the cell loses control of its normal protein expression leading to mis regulation of proteins such as A3. When over expressed the A3 proteins can mutate cellular genomic DNA, causing the cancer to progress more quickly.

The A3 enzyme's ability to mutate DNA can be problematic for our cells. Because A3B and A3H-I are localized in the nucleus there is a chance they can mutate the cellular genome. Additionally, during mitosis, A3A, C and H can access cellular genomic DNA and potentially mutate it. (8). These mutations in cellular DNA can cause the cell to lose control of its growth and division. APOBEC3 mediated genomic mutation occurs when the cell loses control over the regulation of these enzymes, likely as a result of an earlier oncogenic event (**Figure 4)**. Increased expression of A3 enzymes that can access genomic DNA in the nucleus such as A3A or A3B can lead to increased cellular DNA mutations which can overwhelm the cellular DNA repair processes and lead to permanent mutations (22,23).

When treating cancer, a cancer with a slower mutation rate has a more favorable prognosis than one that rapidly mutates (24). A rapidly mutating cancer can become resistant to chemotherapy making treatment more difficult for the patient. The phenotype of the cancer can also change making it more likely to fragment and colonize other tissue (23,24). Because A3 enzymes can cause mutation in cancers, by learning how APOBEC3 enzymes levels are mis

regulated in cancerous tissues, we could better understand cancer progression and aid in the development of anticancer therapeutics (21,22,23,6).

MicroRNAs:

To be able to grow and perform its function a cell must control the expression of the proteins it produces. There can be global regulation from outside sources that activate transcription factors which control the expression and transcription of genes. This mechanism can be used to control large amounts of genes and their products. A more fine-tuned regulation of a specific

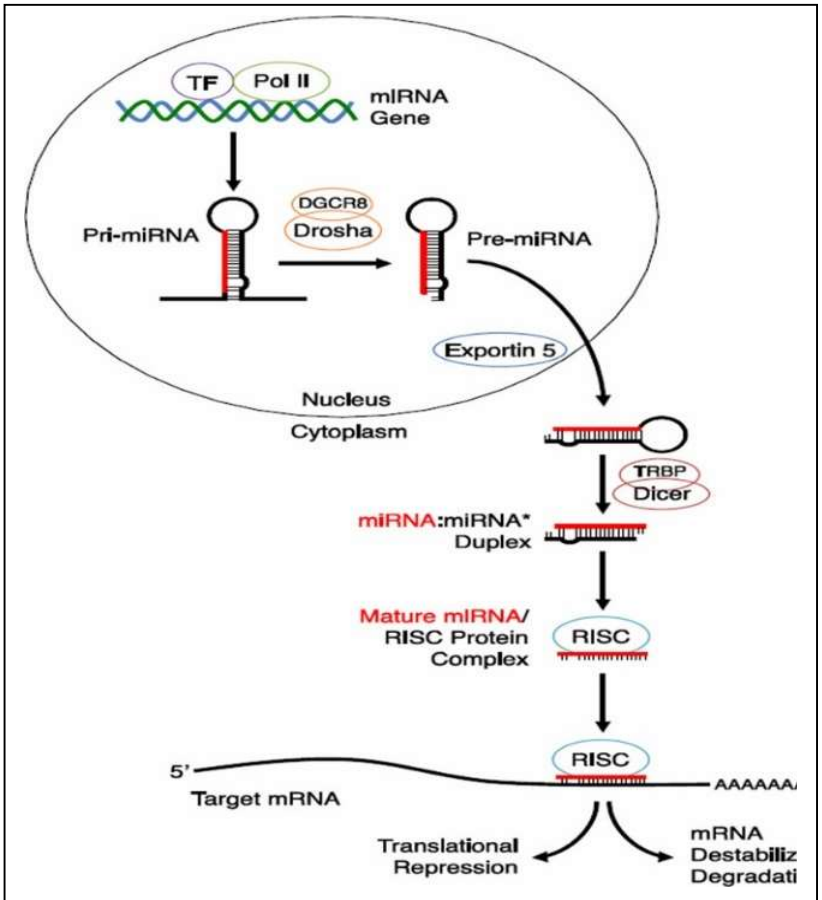

Fig 5. The creation of miRNAs and the RNA induced silencing complex (RISC) protein binding and regulation of mRNA. RISC binds the 3'UTR region of the mRNA, and leading to translation repression and signaling the mRNA for degradation. Adapted from (11).

protein can occur with the use of miRNAs. MiRNAs are small fragments of RNA that regulate protein expression by binding to the messenger RNA (mRNA) on the 3' untranslated region (3' UTR) and signaling the mRNA for degradation. **Figure 5** shows the process of miRNA maturation and silencing of the target mRNA strand. In most cases, miRNA binding causes the destabilization or degradation of the mRNA (25). This destabilization stops protein from being produced from that mRNA, making it a way a cell can control how much protein is produced by each strand of mRNA. MicroRNAs are transcribed from noncoding regions of genomic DNA by RNA polymerase II (25). Once the miRNA is transcribed, it folds into a hairpin structure, allowing the nuclear proteins Drosha and DGCR8 to bind (25). This protein complex processes the miRNA and cuts it into a smaller hairpin around seventy nucleotides long (25). The miRNA then is exported from the nucleus to become activated in the cytoplasm.

Once the miRNA is in the cytoplasm it is bound by Dicer and TRBP (25). These proteins remove the hairpin loop on the miRNA, converting the miRNA molecule into single strands. Once the hairpin is removed the miRNA forms a complex with the RNA Induced Silencing Complex (RISC) (60,61). RISC is a protein complex that is made up of RNA binding proteins, which help stabilize the miRNA and allow it to bind to its 3'UTR mRNA target, leading to its inhibition of transcription (26, 60,61).

The miRNA uses a seed site, located from nucleotides 2-7 of the miRNA to bind to the 3' UTR of the mRNA. (25,27). The miRNA can bind fully or partially to the targeted mRNA using the seed site region, and how tightly it binds can lead to different outcomes. Close matches can lead to mRNA degradation while weak matches can lead to inhibition of translation initiation, but not mRNA destruction (25,27). The miRNAs dependance on these seed sites and complementary regions in the mRNA 3' UTR allows for target prediction.

Thousands of miRNAs have been described across many different species, because of this, miRNAs follow a naming nomenclature in which related miRNAs use letters in the suffix to differentiate them, while miRNAs that are identical but are transcribed from different areas of the genome, an additional number is used (28,29). For example, for a human miRNA the prefix *hsa* is used followed by *mir* and the number the miRNA in sequential sequence based on discovery such as *hsa-mir-1, hsa-mir-2* and so on. (28,29). This convention is not always followed and miRNA sequences that are nearly identical can have a letter or number suffix added such as, *hsa-mir-1a or hsa-mir-1-1*. A miRNA identified in an organism and then later discovered in a different organism can be named the number but with the species prefix. Because each miRNAs transcript can give rise to two mature miRNAs, one from each side of the pre miRNA hairpin loop, (5' or 3'), the suffix 5p or 3p can be added to the end of the name (28,29). This naming nomenclature allows many unique miRNAs of to be described and easily identified.

Losing the ability to control gene expression is a hallmark of many diseases such as cancer. In head and neck, lung, ovarian, and breast cancers, studies have found altered levels of miRNAs in the cancer tissue (30,31,32). This can lead to abnormal protein levels that can worsen the outcome of the disease. Because miRNAs are so important to the cells ability to regulate protein levels, the miRNAs can be targets for therapeutic approaches to diseases where abnormal protein expression is present.

Using miRNAs as therapeutic agents presents a few challenges as miRNAs are not stable in the bloodstream, they are negatively charged so they are not easily absorbed by cells, and they need to target specific cells to avoid systemic toxicity (32,33). For tumors that are easily accessible a local injection can be used but for harder to reach tumors the miRNAs would have to be injected intravenously (32,33). To help stabilize the miRNA in the bloodstream and facilitate uptake viral

vectors could be used to target tumors cells. These have drawbacks as they will eventually initiate an immune response to the viral vector. Non-viral vectors such as nanoparticles, exosomes, and even metals such as gold could be used as well but face challenges such as systemic toxicity and need more research to be developed (24). In animal trials however, the addition of miRNAs that regulate oncogenes reduced the growth of tumors (32). MicroRNAs show great potential in treating cancer and with more efficient delivery systems for the miRNAs being developed they will become important tools in cancer therapy.

APOBEC3 and miRNAs:

Many proteins present in cells, including A3 enzymes are regulated by microRNAs (26). In cancers that have abnormal levels of A3 enzymes, altered levels of miRNA control could be a contributing factor. Since human A3 enzymes have important roles in limiting viral infections and enhancing certain cancers, it is important to identify the miRNAs that regulate them. Once a library of miRNAs that interacts with the A3 enzymes mRNA is created, it can be used to better understand A3 regulation, and potentially be used as therapeutic targets to treat A3 mediated cancers.

I hypothesize that the A3 family is regulated by miRNAs, each member by a different repertoire of miRNAs.

Experimental summary:

To test my hypothesis three to four miRNAs identified to interact with each A3 enzyme were predicted using three miRNA databases. The miRNAs were then amplified from HeLa cell genomic DNA and cloned into pcDNA 3.1 (Invitrogen) expression vectors. All seven A3 3' UTRs were cloned into a psi-Check-2 (Promega) vector that encodes two luciferase enzymes. To test the

regulatory effects of the miRNAs, HeLa cells were transfected with the psi-Check-2 expression vector with the A3 3' UTR as well as pcDNA 3.1 vector encoding the miRNA. A dual luciferase assay (Promega) was performed to quantify the miRNA silencing for each A3.

Methods:

Identifying candidate miRNAs and their sequences:

To find candidate miRNAs listed in **Table 1**, A3 3'UTRs sequences were entered into three miRNA databases, microrna.org, mirdb.org, and mirwalk (34,35,36,37,38). The database at microrna.org identifies predicted matches by using the miRanda algorithm (39). The miRanda algorithm matches optimal sequence complementarity between miRNAs and a specific mRNA sequence, giving a score based on the match of base pairs. It also estimates the free energy needed to form a miRNA:mRNA complex. Mirdb is a miRNA database that mines literature for miRNAs as well as allowing submissions of new miRNAs (35,37). This data base integrated an algorithm that uses RNA sequences and cross-linking immunoprecipitation data to determine miRNA and mRNA targeting efficiencies (35,32). The last database mirwalk accesses miRNA information from other data bases such as mirdb, and NCBI, it then uses the TarPmiR algorithm to identify potential interactions between miRNAs and mRNAs (40). The TarPmiR algorithm predicts binding through deep machine learning from data derived from crosslinking, ligation, and sequencing experiments (41). These databases each use different algorithms to predict binding efficiency of miRNAs to the 3' UTR of mRNA and the miRNAs with the highest predicted efficiency for interacting, across all three databases were chosen. Full sequences for the identified miRNAs were then found on GenBank by NIH and then identified in the genome on the University

of California Santa Cruz Genome Browser (42). The identified sequence, with extra base pairs on both sides for the creation of a primer, was then copied from UCSC Genome Browser and entered into Primer3Plus. Forward and reverse primers were then identified from the miRNA transcript by the Primer3Plus identification tool (43).

Sequencing miRNAs in pMiniT 2.0 vectors:

In order to amplify and sequence the candidate miRNAs, the identified primers were ordered from Integrated DNA Technologies. The miRNAs were then amplified from HeLa genomic DNA by PCR following the recommended protocol for Q5 polymerase (NEB) (44). The resulting PCR products were tested for successful amplification by running on a 1% agarose gel and imaged **Figure 6**. To determine if the miRNAs were successfully amplified, the PCR products were purified and transformed into pMiniT 2.0 vectors following the protocol from NEB (45). The plasmids were then purified and sequenced by GeneWiz and the resulting sequences were compared to their original primer sequence.

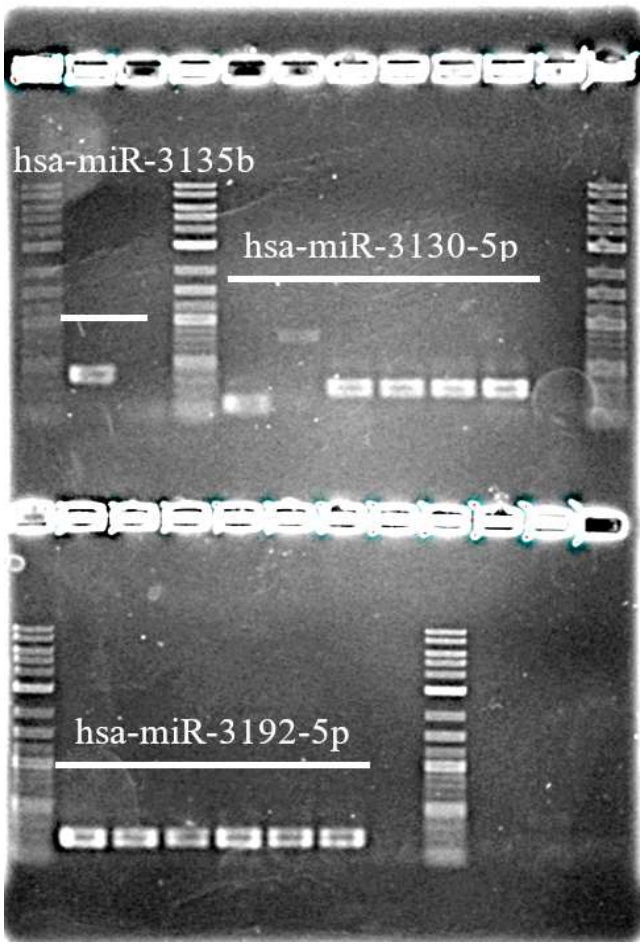

Figure 6. PCR confirmation of miRNA sequences in pcDNA 3.1 vectors. miRNAs show in this figure are (top left band), *hsa-miR-3135b,* the six bands on the top right *hsa-miR-3130-5p, and the bottom hsa-miR-3192-5p.* The samples were then sent to Genewiz to for sequencing.

Cloning miRNAs into pcDNA-3.1 expression vectors:

To prepare the miRNAs for use in the dual luciferase assay, successfully sequenced miRNAs in pMiniT 2.0 vectors were excised by EcoRI and BamHI (NEB) and ligated into pcDNA 3.1 expression vectors with Instant Sticky-End ligase (NEB) **Figure 7**. The plasmids were then purified and sequenced by GeneWiz to confirm correct sequence and orientation.

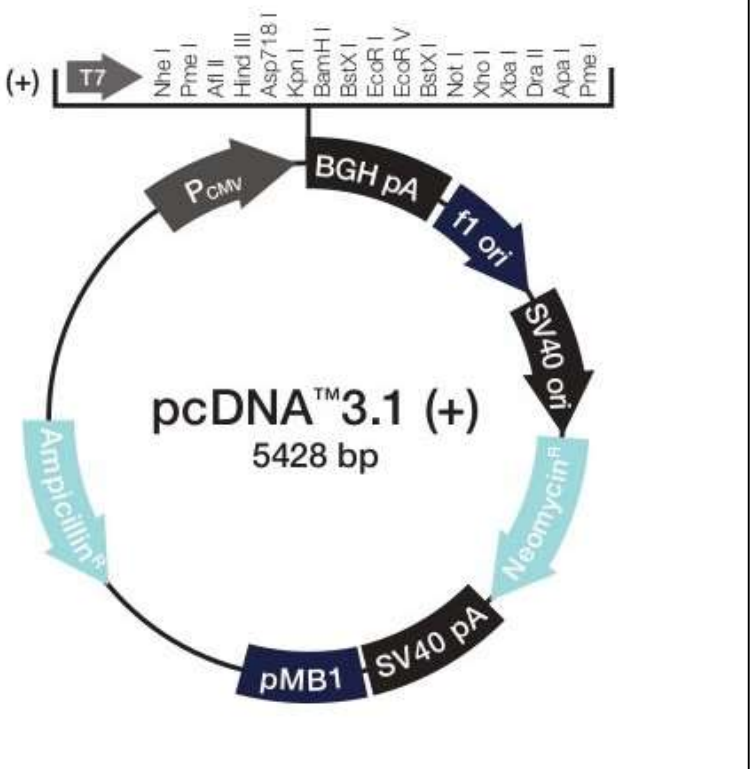

Figure 7. The pcDNA 3.1 vector from Invitrogen, the miRNA constructs will be cloned into (**72**).

Cell culture conditions:

HeLa cells used in the experiment are grown in a 37˚ C incubator at a 5% CO2 atmosphere. The cells are plated on a tissue cell culture plate in Dulbecco's modified eagle media (DMEM), with 10% fetal bovine serum.

Optimization of transfection conditions:

In order to determine the optimal cell number for the assay, HeLa cells were counted on a Luminocyte hemocytometer and differing densities of cells were then seeded into flat bottom 96 well plates. After a 24-hour period, the Hela cells were transfected with 100ng of pEGFP-N3 plasmid, which encodes GFP (Green Fluorescent Protein), by following a polyethylenimine (PEI) transfection procedure (46). After a 48-hour period, The Hela cells were removed from the plate and resuspended in DMEM medium 10 µl of cells were mixed with 10 µl of trypan blue and immediately loaded onto a Luminocyte hemocytometer. Total number of alive cells, dead cells, and cells fluorescing green under UV light were counted and percent dead and percent fluorescing were calculated. The data was analyzed to determine the optimal cell number for maximal transfection efficiency and minimal cell death and the amount was used for the rest of the experiments.

To determine the optimal amount of plasmid DNA to be transfected; HeLa cells were plated in a flat bottom 96 well plate and transfected with 80, 90, 100, 110, and 120ng of pEGFP-N3 plasmid DNA using the same methods above. After a 48-hour period the cells were counted as above, and the optimal amount of plasmid DNA was determined and used for the rest of the experiments.

To determine the optimal time for plasmid encoded protein expression, using the conditions for cell concentration and DNA amount from above, HeLa cells were plated and transfected into a flat bottom 96 well plate. Cells were then counted at 24, 48, 72, and 96 hours after transfection to find the optimal time to allow the cells to express the most fluorescent proteins.

Testing miRNA constructs:

To test the ability of the miRNAs to silence A3, a dual luciferase assay was performed. The assay was performed by seeding Hela cells in a 6 well flat bottom plate. After 24 hours the HeLa cells were transfected with a 950 ng of pcDNA3.1 vector containing miRNA and 150 ng of psi-Check-2 A3 3' UTR. After 48 hours the HeLa cells were scraped from the wells and resuspended in 225 µl of serum free DMEM. An equal amount of Dual-Glo® luciferase reagent (Promega) was mixed in and the mixture was allowed to incubate at room temperature for ten minutes (47). 150 µl of the cell and luciferase mixture was added in triplicate to a flat bottom 96 well plate, which was read by a Flouorskan Ascent FL luminometer (Thermo Labsystems), to determine the amount of firefly luciferase present. Next, 75 uL of Dual-Stop and Glo® (Promega) was added to each sample in the 96 well plate to determine the *Renilla* luciferase level (47). After allowing the plates to incubate for ten minutes at room temperature, the *Renilla* luciferase levels were measured in the luminometer. A ratio of *Renilla* luminescence over firefly luminescence was created to determine the ability of the miRNA to interact with the A3 3' UTR.

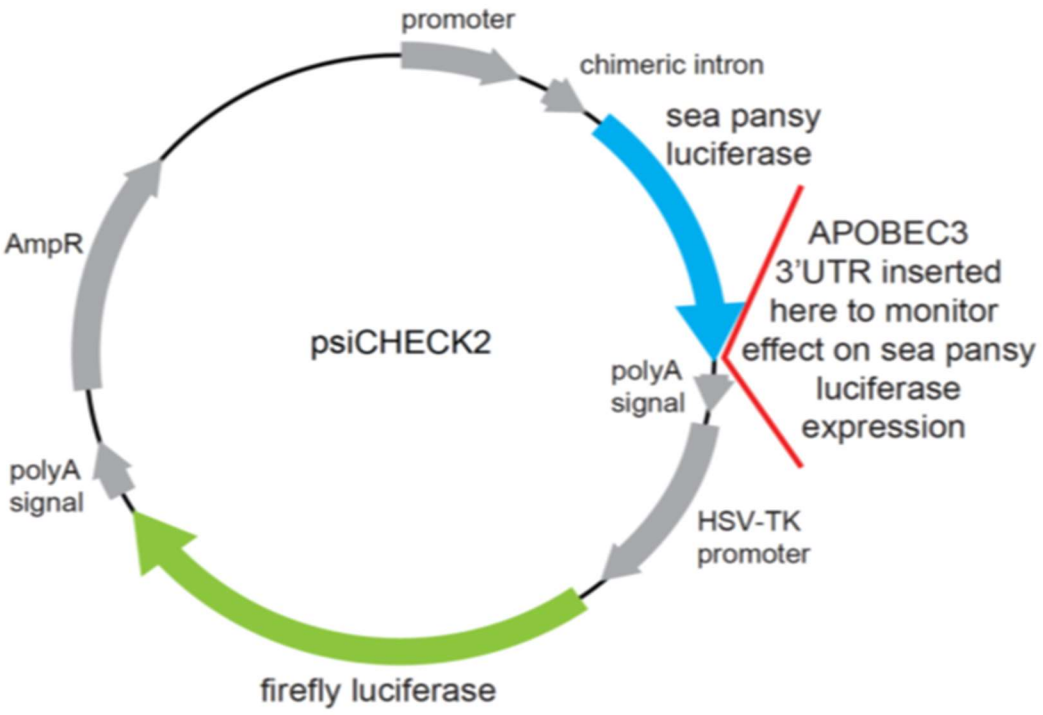

Figure 8. The psi-Check-2 vector that the A3 3' UTR constructs were cloned into. The A3 3' UTR is fused to the sea pansy or *Renilla* luciferase, while the firefly luciferase is independently expressed. The miRNAs are predicted to silence the A3 3' UTRs and decrease the expression of *Renilla* luciferase compared to the firefly.

Data analysis:

All data was added and organized in Microsoft Excel worksheet 365. Graphs and statistical tests were done in GraphPad Prism 8.4 software suite. When comparing two samples an unpaired parametric T test was used with a confidence interval at ninety-five percent. When comparing multiple samples, a one-way ANOVA with multiple comparisons was used with a confidence interval at ninety-five percent.

Results:

Identifying candidate miRNAs:

To create a library of miRNAs predicted to target A3 enzymes, the three algorithms were used to create an expansive list of miRNAs and their predicted efficiency. Of the many miRNAs predicted to target the A3 enzymes, the miRNAs with the highest predicted target efficiency were chosen to be cloned. A total of twenty-four miRNAs were successfully cloned into pcDNA 3.1 vectors and sequenced confirmed, **Table 1.** Many of these miRNAs are predicted to interact with multiple A3 enzyme and were all tested in the dual luciferase assay with their predicted targets.

Optimization of dual luciferase assay:

To determine the optimal conditions for the dual luciferase, multiple cell numbers, different DNA concentrations, and different times allowed for expression were tested. I determined that the optimal concentration of HeLa cells to seed per well was determined to be $1.7*10^4$ cells per mL in 200 μL of medium **Figure 9**. At this concentration the number of cells that expressed GFP fluorescent was the highest, and there was a small percentage of dead cells. For the amount of plasmid DNA to add to transfect with, 110ng was chosen **Figure 9**. This amount of plasmid DNA showed a high level of transfected positive or green cells compared to other amounts and the number of dead cells was acceptable. The optimal time for the HeLa cells to express the luciferase proteins was 48 hours from transfection. **Figure 9.**

Table 1. MicroRNAs predicted to interact with A3s and successfully cloned into pcDNA 3.1 vectors.

	APOBEC3						
miRNA	A3A	A3B	A3C	A3D	A3F	A3G	A3H
hsa-miR-30b-3p	-	-	miRwalk	miRwalk	miRDB miRwalk	-	miRwalk
hsa-miR-373-5p	-	-	-	miRDB miRwalk	miRDB	-	-
hsa-miR-505-5p	miRwalk	miRDB	miRwalk	-	miRwalk	-	-
hsa-miR-518a-5p	-	microRNA.org	miRwalk	-	miRDB	-	-
hsa-miR-527	-	-	miRwalk	-	-	-	-
hsa-miR-548p	microRNA.org	microRNA.org	miRwalk	microRNA.org	microRNA.org	miRwalk microRNA.org	-
hsa-miR-622	miRwalk	miRwalk	-	miRwalk	miRwalk	miRwalk	-
hsa-miR-636	-	-	-	-	miRwalk	-	miRwalk
hsa-miR-890	(73) miRDB	microRNA.org miRDB	-	miRwalk	miRwalk	-	microRNA.org miRDB
hsa-miR-1207-3p	microRNA.org miRDB	microRNA.org miRDB	microRNA.org	microRNA.org	microRNA.org miRDB	miRDB microRNA.org	miRwalk miRDB
hsa-miR-1227-5p	miRwalk	-	miRwalk	miRwalk	miRwalk	-	miRwalk miRDB
hsa-miR-1285-5p	miRwalk	-	-	miRwalk miRDB	miRwalk miRDB	-	-
hsa-miR-3130-5p	microRNA.org	microRNA.org	-	miRwalk miRDB	microRNA.org miRwalk	microRNA.org	-
hsa-miR-3135b	miRDB miRwalk	(73) miRDB miRwalk	miRwalk miRDB	miRwalk	miRwalk	miRDB	-
hsa-miR-3173-5p	-	-	-	miRwalk	miRwalk	-	miRDB microRNA.org
hsa-miR-3192-5p	miRwalk	-	-	-	-	-	-
hsa-miR-4261	-	-	-	-	microRNA.org miRwalk	-	-
hsa-miR-4283	-	-	miRwalk miRDB	miRwalk miRDB	miRDB miRwalk	miRwalk	miRwalk
hsa-miR-4294	miRwalk	-	-	-	-	miRwalk	-
hsa-miR-4297	miRDB	-	-	-	miRwalk	miRwalk miRDB	-
hsa-miR-4314	-	-	-	-	miRwalk	-	-
hsa-miR-4316	-	-	miRwalk	microRNA.org miRwalk	-	-	miRwalk
hsa-miR-4522	miRwalk	(73) miRDB	-	-	miRwalk	miRwalk	-
hsa-miR-6778-3p	-	-	miRDB miRwalk	miRwalk	miRwalk	-	miRDB

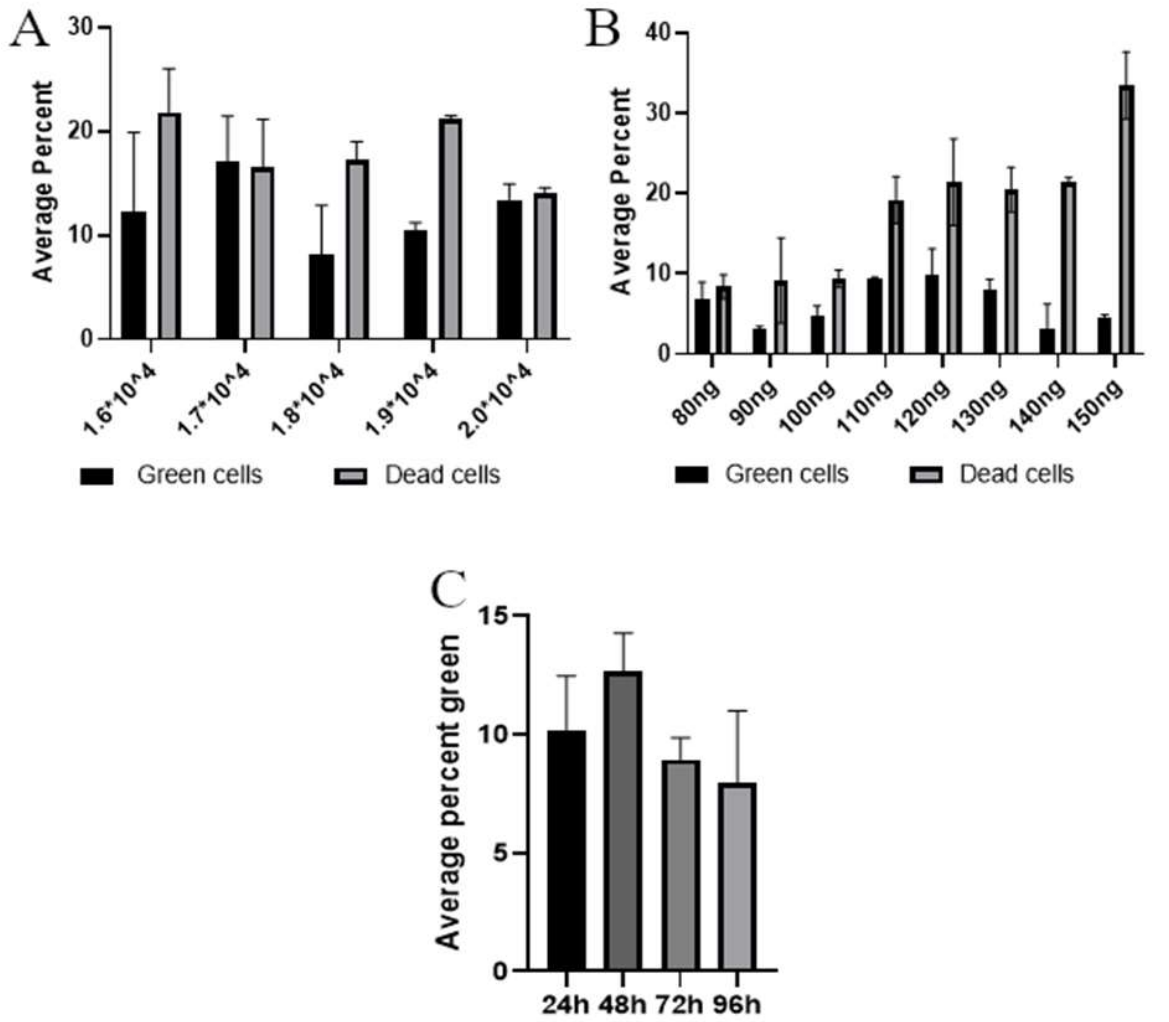

Figure 9. Optimization of transfection conditions. **A**. Optimal cell density results. Hela cells were transfected with GFP and allow to express proteins for forty-eight hours and then counted for expression of GFP. The optimal density of cells to be plated, 1.7* 10^4 cells/ml in 200µL of medium. **B.** Optimal plasmid DNA transfected results. Hela cells were transfected with GFP and allow to express proteins for forty-eight hours and then counted for expression of GFP. The optimal amount of DNA to be added is 110ng based on average percent of green cells and average percent of dead cells. **C.** Optimal time for plasmid expression results. Hela cells were transfected with GFP and allow to express proteins for forty-eight hours and then counted for expression of GFP. The optimal time to measure the cells for expression of proteins is forty-eight hours after transfection.

Testing A3 3' UTR constructs expression

To confirm that the psi-Check-2 plasmid with the A3 enzymes worked they were tested in the dual luciferase assay using the optimized assay conditions **Figure 10**. Two different amounts of A3 3' UTR plasmids were added to test for scaling and to determine if changing the amount of DNA would change the luciferase expression. In general, the relative luminescence for the A3 3'UTR plasmids was similar in the two levels tested, which was expected. Some of the A3 3'UTR

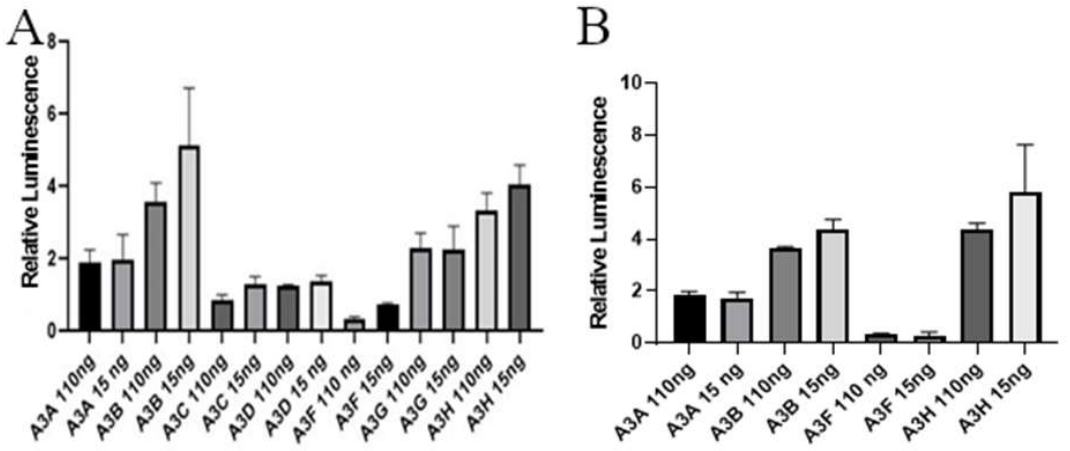

Figure 10. A. Relative luminescence of A3 enzymes in the psi-Check-2 vector at differing concentrations. Hela cells were transfected with A3 3'UTR containing psi-Check2 vectors at 110 and 15ng per well and allowed to express proteins for forty-eight hours, after which the cells were collected, lysed and dual luciferase light levels measured. Relative luminescence was determined by the ratio of the Renilla luciferase over firefly luciferase. **B.** Relative luminescence of A3A, A3B, A3F, and A3H 3' UTRs in the dual luciferase assay. Hela cells were transfected with A3A, B, F and H 3'UTR containing psi-Check2 vectors at 110 and 15ng a well. The cells then were allowed to express proteins for forty-eight hours, after which the cells were collected, lysed and dual luciferase light levels measured. Relative luminescence was determined by the ratio of the *Renilla* luciferase over firefly.

constructs showed higher relative luminesce values when more of the 3'UTR plasmid was delivered to the cell, but the difference was not significant. Of the seven A3 enzymes, A3A, A3B, A3G and A3H showed the highest relative luminescence expression. The A3F plasmids had a very low relative luminescence compared to the other A3 enzymes. This was driven by low Renilla luminescence levels, compared to the other A3 constructs, but comparable firefly luminescence

levels. This suggests that HeLa cells may express endogenous miRNAs that bind the A3F 3'UTR to meditate silencing. As a consequence, I predict that it may be more challenging to identify miRNAs that silence A3F, compared to the others A3 enzymes.

Testing miRNA constructs:

A3A: To determine which miRNAs could silence A3A, the twelve miRNA constructs that were predicted to target A3A from **Table 1** were tested with the A3A 3' UTR plasmid. The relative luminescence of the A3A 3'UTR-plasmid without an added miRNA was compared to the luminescence in combination with each of the identified twelve miRNAs as seen in **Figure 11**. Additionally, the miRNA *hsa-miR-4297* was included to be an example of a miRNA that did not silence A3A. Five miRNAs showed potential targeting, as demonstrated by decreased relative luminescence compared to the empty vector (pcDNA) control. The assay was repeated with these five miRNAs; **Figure 11B** and has *miR-1227-5p* appeared to have a lower relative luminance when compared to the no miRNA control. This miRNA was tested a final time **Figure 11C.** However, no significant difference between the control and *hsa-miR-1227-5p* was observed.

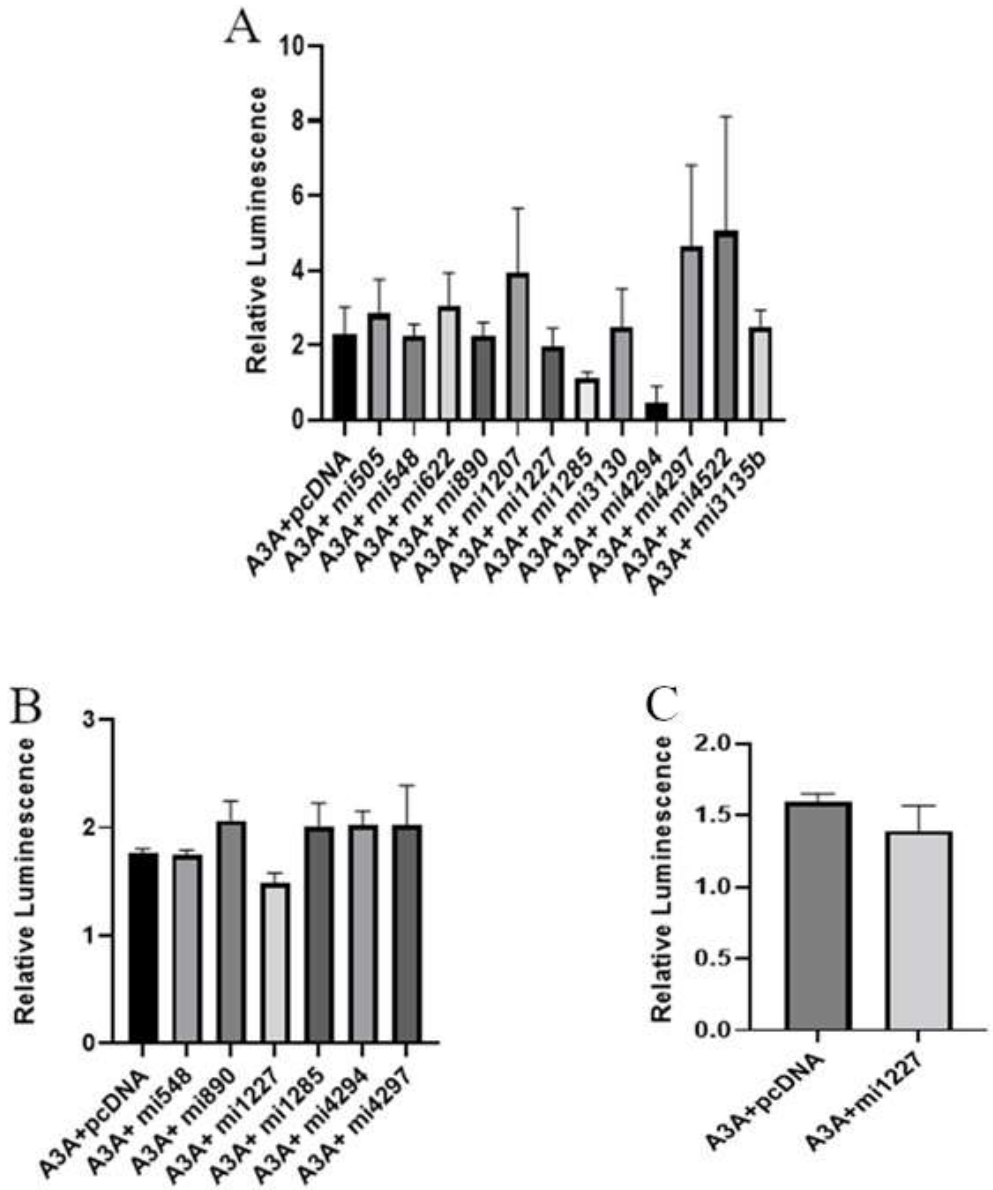

Figure 11 A. Relative luminescence for selected miRNAs and A3A. Hela cells were plated and transfected with A3A 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing miRNAs. After forty-eight hours the cells were collected, and luciferase levels were measured **B.** The p value for the difference between pcDNA and *hsa-miR-1227-5p* is 0. 0550. **C.** Relative luminescence for A3A and *hsa-miR-1227-5p* versus the control. Hela cells were plated and transfected with A3A 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing *hsa-miR-1227-3p*. After forty-eight hours the cells were collected, and luciferase levels were measured.

A3B: To evaluate which miRNAS could silence A3B, the eight of the highest scoring miRNAs for predicted for interaction with A3B 3' UTRs from **Table 1** were tested. The relative luminescence of the A3B 3'UTR-plasmid without an added miRNA was compared to the luminescence in combination with each of the identified eight miRNAs, with *hsa-miR-505* used as a non-silencing control **Figure 12.** Four miRNAs were retested for potential silencing due to their decreased relative luminescence compared to the pcDNA control **Figure 12B**. The replicate testing showed there was no significant difference between the pcDNA control and the four miRNAs.

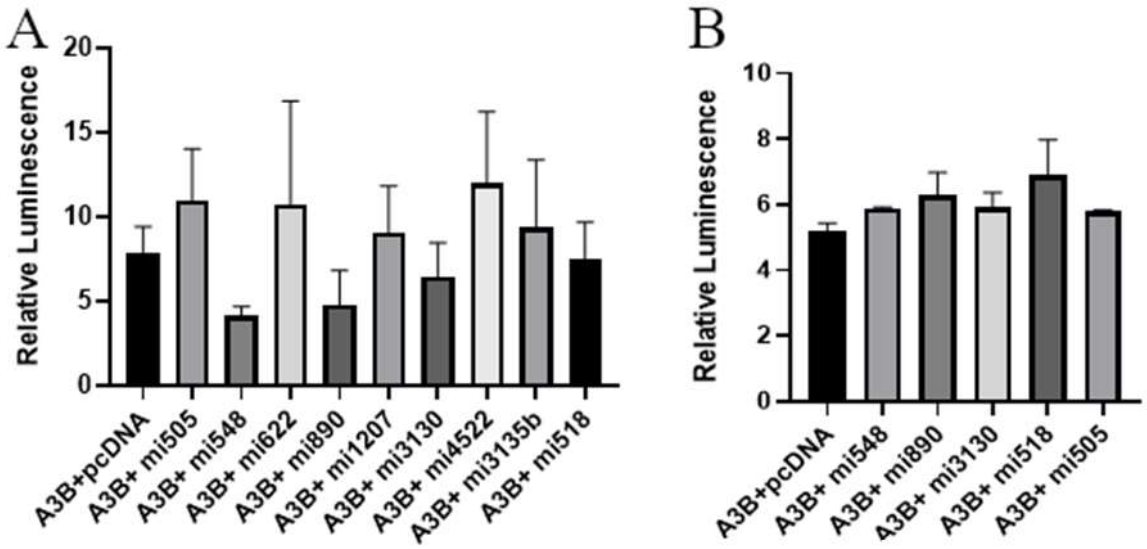

Figure 12. Relative luminescence for all miRNAs predicted to interact with A3B. Hela cells were plated and transfected with A3B 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing miRNAs. After forty-eight hours the cells were collected, and luciferase levels were measured. **B.** Replicate testing of chosen miRNAs.

A3C: The twelve miRNAs that were predicted to silence A3C from **Table** 1, were tested with A3C 3' UTR. The relative luminescence of the A3C 3'UTR-plasmid without an added miRNA was compared to the luminescence in combination with each of the identified twelve miRNAs as seen in **Figure 13.** Five miRNAs showed potential targeting, as demonstrated by decreased relative luminescence compared to the pcDNA control. The assay was repeated with these five miRNAs,

with *hsa-miR-548a-5p* as a non-silencing control **figure 13B**. A single miRNA, *hsa-miR-1207-5p* appeared to have a lower relative luminance when compared to the pcDNA control. To confirm this interaction, *hsa-miR-1207-5p* was retested for its ability to silence APOBEC3C. When tested a final time (**Fig 13C**), I observed a 33% reduction in relative luminesce, compared to the no miRNA control. The difference was statistically significant $p = 0.002$, t test.

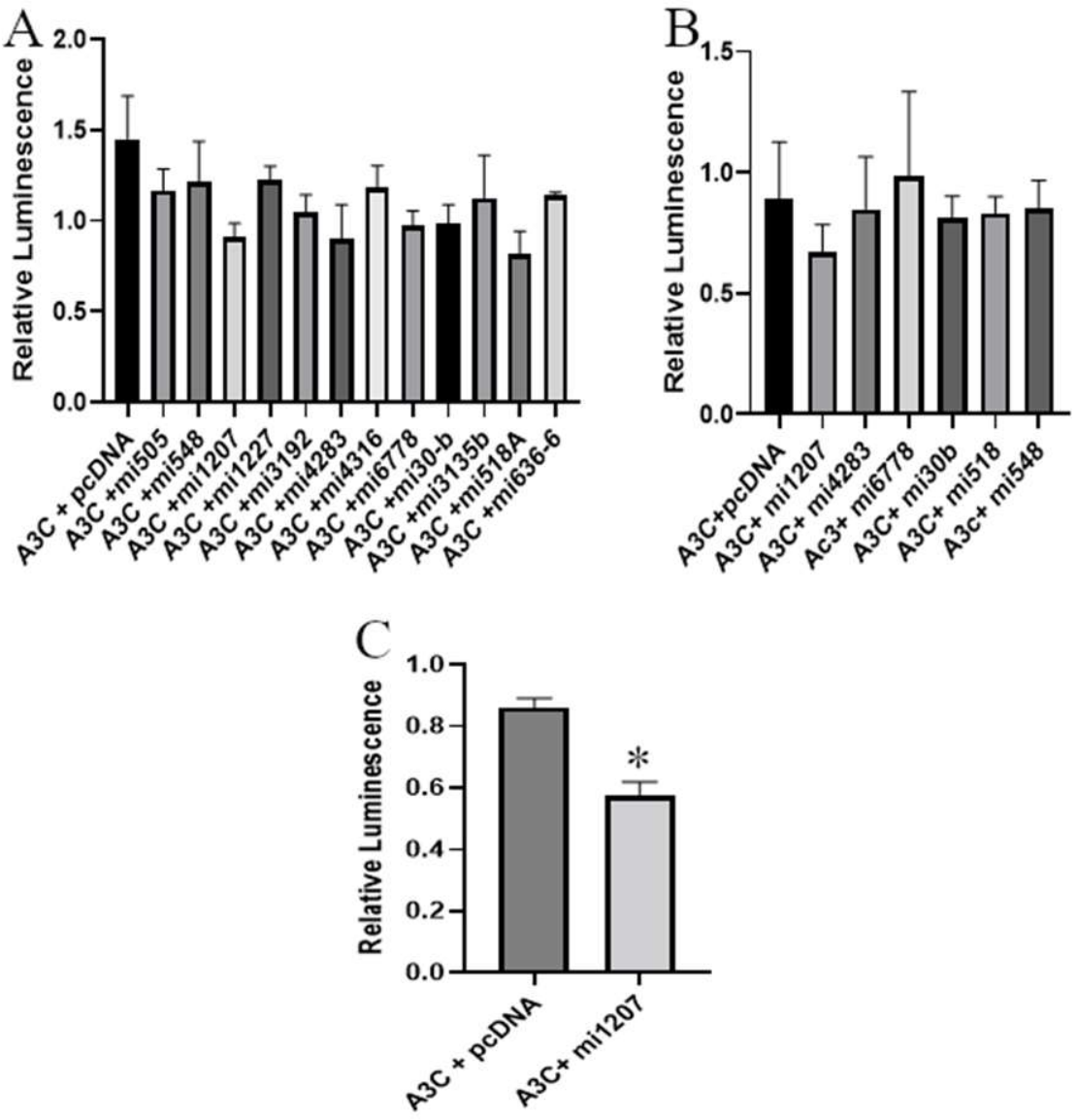

Figure 13. A. Relative luminescence for all miRNAs predicted to interact with A3C. Hela cells were plated and transfected with A3C 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing miRNAs. After forty-eight hours the cells were collected, and luciferase levels were measured. **B.** Replicate testing of chosen miRNAs. **C.** Relative luminescence for A3C and *hsa-miR-1207-5p* versus the control. The experimental value was significantly different $p < 0.005$.

A3D: To determine which miRNAs could silence A3D, fourteen miRNA constructs that were predicted to interact with A3D from **Table 1** were tested in the dual luciferase assay with A3D 3' UTR. The relative luminescence of the fourteen miRNAs with the A3D 3' UTR was compared to the relative luminescence of the A3D 3' UTR plasmid without an added miRNA **Figure 14A**. Four of the miRNAs were tested again because they demonstrated decreased relative luminescence compared to the pcDNA control, with *hsa-miR-373* used as a non-silencing control **Figure 14B.** The miRNA *hsa-miR-548a-5p* had a p value of 0.1083 from the control and was tested a final time. No significant difference in the relative luminescence values between the control and *hsa-miR-548a-5p* **Figure 14C** was observed**.**

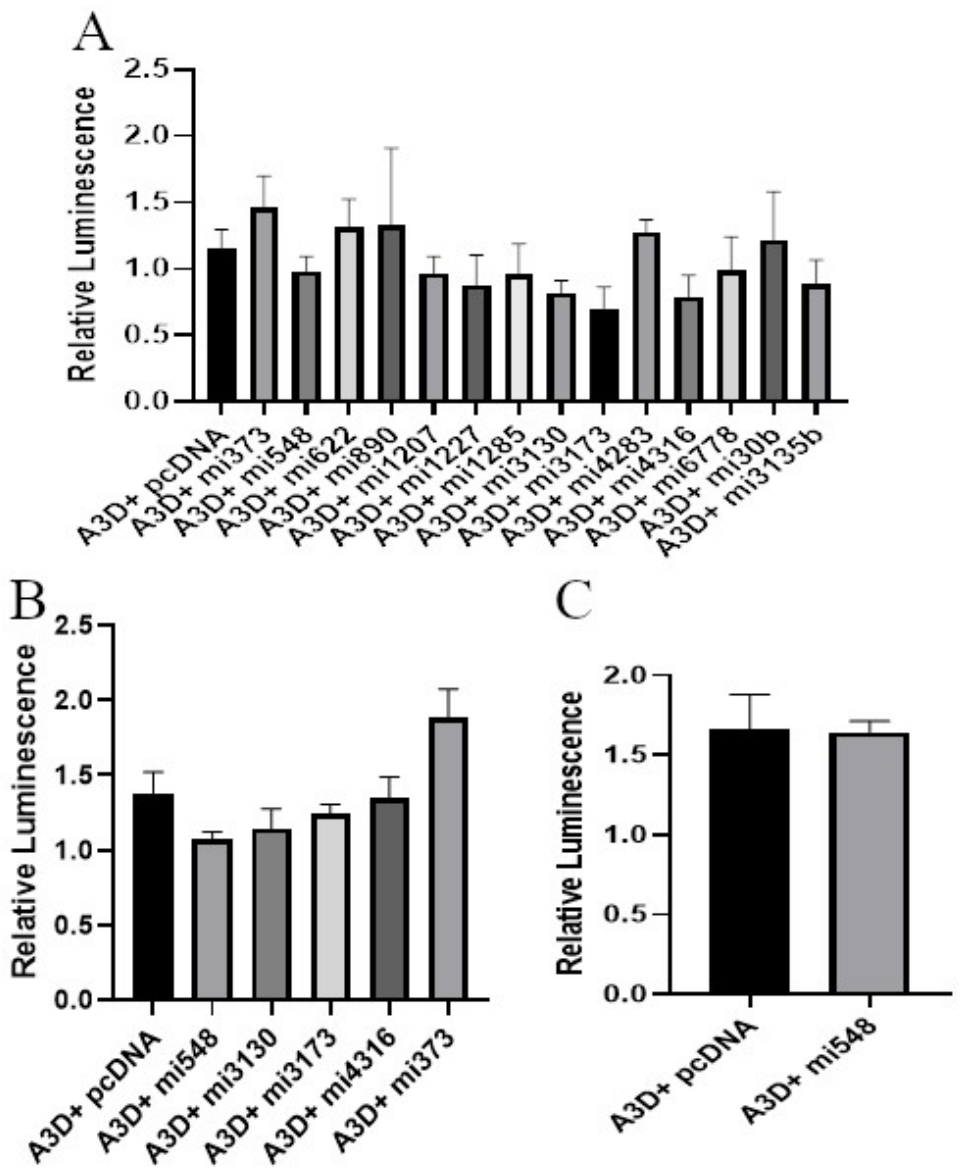

Figure 14. Relative luminescence for all miRNAs predicted to interact with A3D. Hela cells were plated and transfected with A3D 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing miRNAs. After forty-eight hours the cells were collected, and luciferase levels were measured. **B.** Relative luminescence for selected miRNAs and A3D. P value of the difference between pcDNA and *hsa-miR-548a-5p* is 0.1083. Hela cells were plated and transfected with A3D 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing miRNAs. After forty-eight hours the cells were collected, and luciferase levels were measured. **C.** Relative luminescence for *hsa-miR-548a-5p* and control for A3D 3' UTR. Hela cells were plated and transfected with A3D 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing *hsa-miR-548a-5p*. After forty-eight hours the cells were collected, and luciferase levels were measured.

A3F: To evaluate which miRNAs could silence A3F, twenty-two miRNAs from **Table 1** that were predicted to interact with A3F and were tested with the A3F 3' UTR plasmid. The relative luminescence of the A3A 3'UTR-plasmid without an added miRNA was compared to the luminescence in combination with each of the identified twenty-two miRNAs as seen in **figure 15A**. Of those twenty-two miRNAs, three miRNAs demonstrated decreased relative luminescence compared to the pcDNA control and were tested again along with miRNA *hsa-miR-6778* as a non-silencing control **figure 15B**. The miRNA *miR-548a-5p,* had a significantly lower relative luminescence value than the control with p equal to 0.0017 (**Fig 15B**) and was tested a final time. The final test however, failed to observe a significant difference between the pcDNA control and the miRNA **figure 15C.**

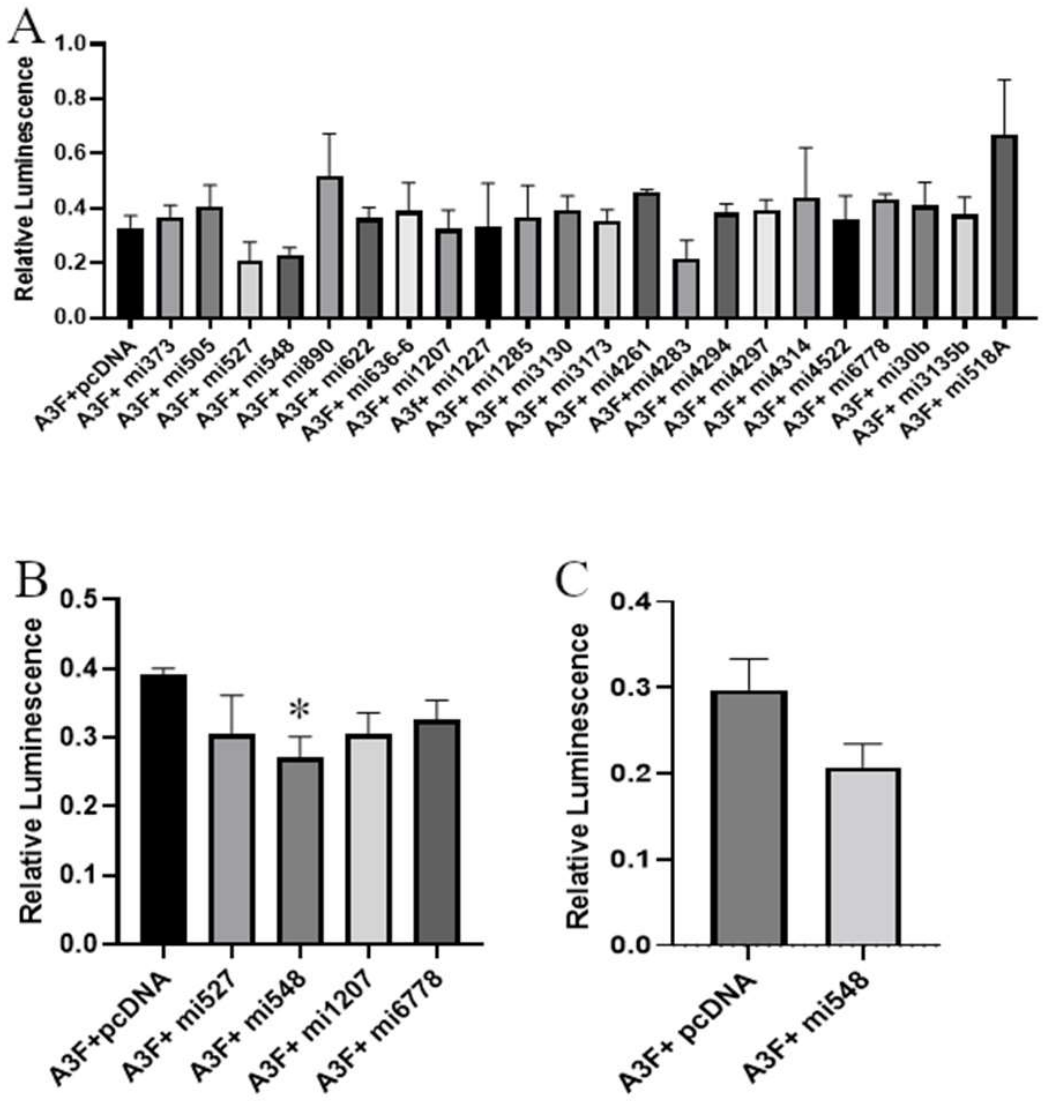

Figure 15. A. Relative luminescence for all miRNAs predicted to interact with A3F. Hela cells were plated and transfected with A3F 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing miRNAs. After forty-eight hours the cells were collected, and luciferase levels were measured. **B.** Relative luminescence for selected miRNAs and A3F. P value of the difference between pcDNA and *hsa-miR-548a-5p* is 0.017. **C.** Relative luminescence for *hsa-miR-548a-5p* and control for A3F 3' UTR.

A3G: The eight miRNAs predicted to silence A3G from **Table 1** were tested with A3G 3' UTR in the dual luciferase assay. The relative luminescence of the eight miRNAs with the A3G 3' UTR was compared to the relative luminescence of the A3G 3' UTR plasmid without an added miRNA, and two miRNAs *hsa-miR-4283 and hsa-miR-3135b*, were found to have relative luminescence

values significantly lower from the control with p values of 0.0021 and 0.0043 respectively **Figure 16A**. These two miRNAs along with two others demonstrating decreased relative luminescence compared to the pcDNA control were tested again **Figure 16B**. However, retesting showed none of the miRNAs demonstrating any significantly lower relative luminescence compared to the control.

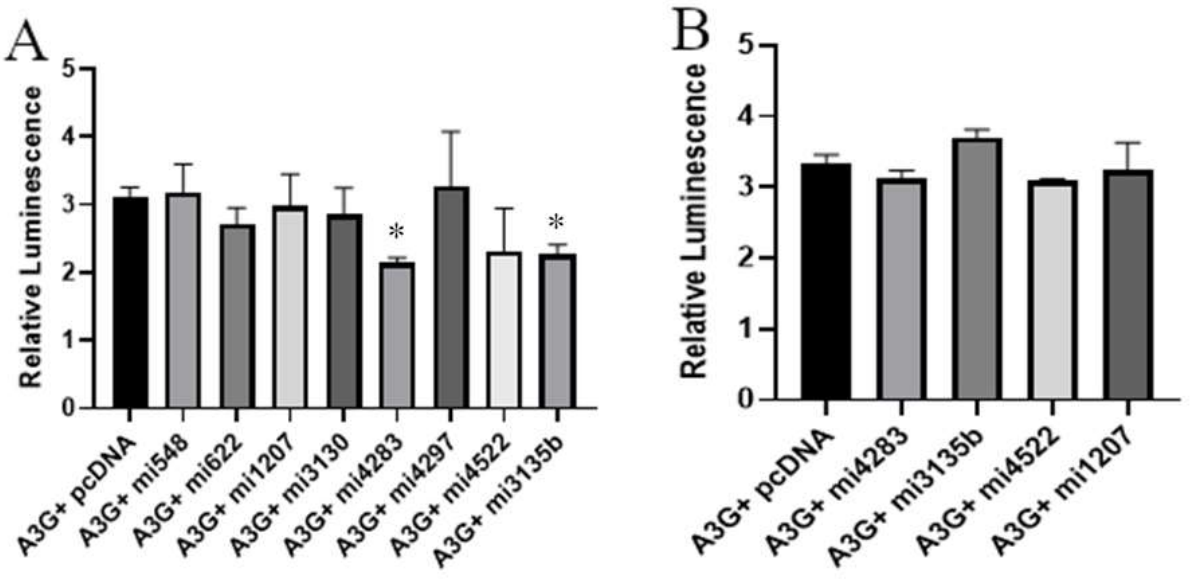

Figure 16. A. Relative luminescence for all miRNAs predicted to interact with A3G. miRNAs *hsa-miR-4283 and hsa-miR-3135b*, have relative luminescence values significantly different from the control with p values of 0.0021 and 0.0043 respectively. Hela cells were plated and transfected with A3G 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing miRNAs. After forty-eight hours the cells were collected, and luciferase levels were measured. **B.** Relative luminescence for selected miRNAs and A3G. Hela cells were plated and transfected with A3G 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing miRNAs. After forty-eight hours the cells were collected, and luciferase levels were measured.

A3H: To determine which miRNAs could silence A3H, nine miRNA constructs that were predicted to interact with A3H from **Table 1** were tested in the dual luciferase assay with A3H 3' UTR. The relative luminescence of the A3H 3'UTR-plasmid without an added miRNA was compared to the luminescence in combination with each of the identified twelve miRNAs as seen

in **figure 17A.** Five of the miRNAs showed potential targeting and were tested again. However, none of the miRNAs were shown to target the A3H 3' UTR. **Figure 17B.**

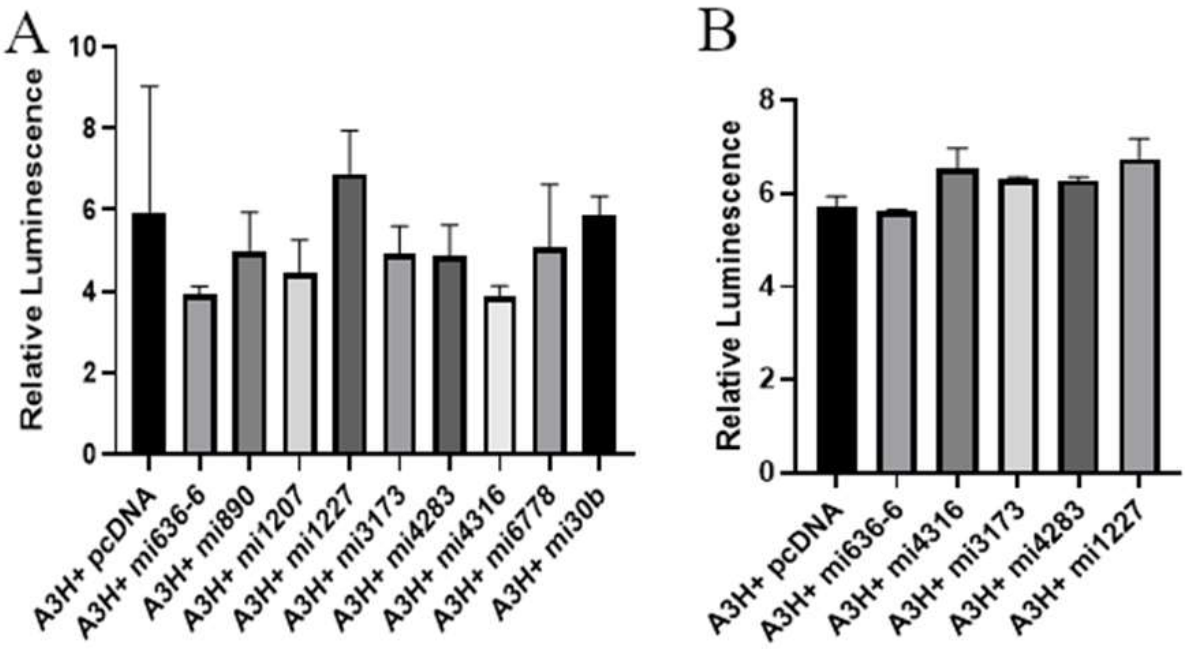

Figure 17. A. Relative luminescence for all miRNAs predicted to interact with A3H. Hela cells were plated and transfected with A3H 3' UTR in psi-Check-2 vectors and pcDNA 3.1 plasmids containing miRNAs. After forty-eight hours the cells were collected, and luciferase levels were measured. **B.** Replicate testing of chosen miRNAs.

Discussion:

The objectives of this project were to create a library of miRNAs that were predicted to regulate translation of the human APOBEC3 cytosine deaminases, clone these miRNAs into mammalian expression vectors, design and optimize a luciferase assay to test the miRNAs and identify miRNAs that silence the APOBEC3 proteins via the 3' UTR of the miRNA. These objectives were important in testing my hypothesis that the A3 family is regulated by miRNAs, each member by a different repertoire of miRNAs.

A total of twenty-four miRNAs were identified as candidates and cloned into plasmid vectors. These miRNAs are the beginning of a library of potential targets for A3 enzyme regulation. Many of candidate microRNAs were predicted to interact with multiple A3 3'UTRs, however, this experiment was unable to show significant interaction from these miRNAs. This does not mean there is no interaction and more optimization of the assay could yield more conclusive results, The objective to create a library of miRNAs however was successful and this library of miRNAs will need to be tested more with potentially different assays before any definitive conclusion can be made.

Even with the challenges presented by the assay, three miRNAs did show results that indicated the possibility of an interaction, *hsa-miR-1207-5p*, *hsa-miR-1227-3p,* and *hsa-miR-548a-5p*. These three microRNAs were predicted to interact with the most A3 3' UTRs and showed promising yet inconclusive results in the experiment with *hsa-miR-1207-5p* potentially silencing A3C, , *hsa-miR-1227-3p* requiring more tests to discern any silencing with A3A, and *hsa-miR-548a-5p* potentially silencing A3D and A3F.

The miRNA *hsa-miR-1207-5p* is a twenty-nucleotide long miRNA, whose gene is located on chromosome 8 (48). This miRNA was predicted to interact with all seven A3 3' UTRs. In this thesis, the most likely interaction was shown with A3C. This interaction was tested three times and significance decrease in relative luminescence was found once. This miRNA has significance for viral infection, as it has been demonstrated to decrease expression of Colony Stimulating Factor 1 (CSF-1) (49). This mis-regulation of CSF-1 can be caused by severe *SARS-CoV-2* infection, which lowers *hsa-miR-1207-5p* allowing CSF-1 to become overexpressed and cause inflammation (49). In the lung cancer cell line A549, *hsa-miR-1207-5p* reduced proliferation and invasion into tissue by the cancer cells by targeting HBEGF mRNA, which codes for a growth factor protein

(3,50,51). In prostate cancer the expression of *hsa-miR-1207-5p* serves as a biomarker for a more aggressive strain of the disease but its function in the disease is unknown (52). As a potential candidate for all seven A3 3' UTRs *hsa-miR-1207-5p* remains an attractive target for more research in its ability to regulate A3 enzymes.

The miRNA *hsa-miR-1227-3p* is a nineteen-nucleotide long miRNA, whose gene is located on chromosome 19 (53). We predicted an interaction between A3A, C, D, F, and H and *hsa-miR-1277-3.* The miRNA only showed potential silencing with A3A, but this was not reproducible. *Hsa-miR-1277-3p* has been described to target the mRNA of proteins important for the progression of human disease, such as the transport protein gene SEC23A, protein kinase gene PRKAB2, and Bone Morphogenetic Protein 4 (BMP4) gene (54,55,56). The over expression of *hsa-miR-1277-3p* can alter glioblastoma cancer phenotypes through the downregulation of SEC23A gene expression causing extracellular vesicle to form and invade surrounding tissue (55). In hepatocellular carcinoma cells, BMP4 induced cell proliferation was promoted by the suppression of *hsa-miR-1277-3p,* and cellular apoptosis was induced when the miRNA was overexpressed or BMP4 was suppressed (54). In the disease Fetal Growth Restriction, *hsa-miR-1277-3p* is highly downregulated and increased expression of its targets such as PRKAB2 are indicated the pathology (56). While we were unable to demonstrate interaction with any A3 3'UTRs, *hsa-miR-1277-3p* will remain a high priority to test for any interactions.

The miRNA *hsa-miR-548a-5p* is twenty-one nucleotides long and its gene is found on chromosome 8 (57). In this experiment *hsa-miR-548a-5p* was predicted to interact with the 3' UTR of all the A3 enzymes except A3H. Of the six A3 3' UTRs, only A3D and A3F had any potential interaction with *hsa-miR-548a-5p*. The miRNA showed silencing when first tested with A3F, but the result was not reproducible. In hepatocellular carcinoma cell lines, *hsa-miR-548a-5p*

has been shown to increase cellular proliferation (58). When the miRNA is present in high levels, the anti-tumor gene TG737 is repressed, and cell apoptosis is prevented (58). If *hsa-miR-548a-5p* is shown to interact with the A3 enzymes 3' UTR, it could possibly be used to control the over expression of ABOBEC3 enzymes in cancer cells.

The dual luciferase used in this experiment has been used by other research groups, for example, to show that *hsa-miR-9* inhibits *NF- kB* (59). In the above experiment, the researchers used the cell line HEK293 (59). HeLa cells were used in this thesis as they had been demonstrated by a previous lab member to have relatively consistent results when testing the APOBEC3 3'UTRs and HeLa cells are able create and process their own miRNAs. However, each cell line would endogenously express miRNAs, which could confer some natural regulation of the A3 3' UTR as seen in from the low relative luminescence levels produced by the psi-Check-2 A3F 3'UTR expression construct. The endogenous miRNAs expressed by the HeLa cells could interact with the A3 3'UTR expression constructs affecting the results. Finding a different cell line could be a potential pathway to optimize the assay. The Promega dual luciferase reporter assay used creates a long-lasting reaction that can be used for up to two hours (47). This "slow" reaction may not produce as much luciferase luminance as a faster flash reaction (47). Testing and using different dual luciferase products are another way this assay could be optimized, and once fully optimized, will serve as a valuable tool for identifying interaction between miRNAs and A3 3' UTRs.

This thesis built a plasmid library of miRNAs predicted to regulate the translation of human APOBEC3 cytosine deaminases, cloned those miRNAs into expression vectors, designed and optimized a luciferase assay to test the miRNAs, and assayed the miRNAs for their ability to silence their A3. Of the twenty-four miRNAs tested, *hsa-miR-1207-5p*, *hsa-miR-1227-3p,* and *hsa-miR-548a-5p* are the most likely to interact with the A3 3' UTRs based on the results presented in

this thesis. These three miRNAs will be high priority targets for future research. Understanding the interactions between miRNAs their mRNAs is important for understanding the role of APOBEC3 proteins in cancer.

References:

1. Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, Scott J, Navaratnam N. 2002. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. Genomics 79:285–296.
2. Smith, H. C., R. P. Bennett, A. Kizilyer, W. M. McDougall, and K. M. Prohaska. 2012. Functions and Regulation of the APOBEC Family of Proteins. *Semin Cell Dev Biol* 23: 258–268.
3. Ebrahimi D, Alinejad-Rokny H, Davenport MP. Insights into the Motif Preference of APOBEC3 Enzymes. *PLoS One*. 2014;9(1). doi:10.1371/journal.pone.0087679
4. Seishima N, Kondo S, Wakae K, Wakisaka N, Kobayashi E, Kano M, Moriyama-Kita M, Nakanishi Y, Endo K, Imoto T, Ishikawa K, Sugimoto H, Hatano M, Ueno T, Koura M, Kitamura K, Muramatsu M, Yoshizaki T. 2018. Expression and subcellular localisation of AID and APOBEC3 in adenoid and palatine tonsils. 1. Scientific Reports 8:918.
5. Chiu Y-L, Greene WC. Multifaceted antiviral actions of APOBEC3 cytidine deaminases. *Trends in Immunology*. 2006;27(6):291-297. doi:10.1016/j.it.2006.04.003
6. Granadillo Rodríguez M, Flath B, Chelico L. The interesting relationship between APOBEC3 deoxycytidine deaminases and cancer: a long road ahead. *Open Biol*. 2020;10(12). doi:10.1098/rsob.200188
7. Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS. 2010. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. Nucleic Acids Res 38:4274–4284.
8. Lackey L, Law EK, Brown WL, Harris RS. 2013. Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. Cell Cycle 12:762–772.
9. Münk C, Willemsen A, Bravo IG. 2012. An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. BMC Evolutionary Biology 12:71.
10. Warren, C. J., J. A. Westrich, K. V. Doorslaer, and D. Pyeon. 2017. Roles of APOBEC3A and APOBEC3B in Human Papillomavirus Infection and Disease Progression. *Viruses* 9: 233.
11. 2019. STD Facts - Human papillomavirus (HPV) Retrieved from: https://www.cdc.gov/std/hpv/stdfact-hpv.htm
12. Harris, R. S., J. F. Hultquist, and D. T. Evans. 2012. The Restriction Factors of Human Immunodeficiency Virus. *J Biol Chem* 287: 40875–40883.
13. Global HIV & AIDS statistics — 2019 fact sheet. Retrieved from: https://www.cdc.gov/std/hpv/stdfact-hpv.htm
14. todaysepidemic-508.pdf. CDC factsheet. Retrived from https://www.cdc.gov/nchhstp/newsroom/docs/factsheets/todaysepidemic-508.pdf
15. Harris, R. S., and J. P. Dudley. 2015. APOBECs and Virus Restriction. *Virology* 10: 131–145.
16. Salamango DJ, Becker JT, McCann JL, Cheng AZ, Demir Ö, Amaro RE, Brown WL, Shaban NM, Harris RS. 2018. APOBEC3H Subcellular Localization Determinants Define Zipcode for Targeting HIV-1 for Restriction. Mol Cell Biol 38:e00356-18.

17. Ebrahimi D, Richards CM, Carpenter MA, Wang J, Ikeda T, Becker JT, Cheng AZ, McCann JL, Shaban NM, Salamango DJ, Starrett GJ, Lingappa JR, Yong J, Brown WL, Harris RS. 2018. Genetic and mechanistic basis for APOBEC3H alternative splicing, retrovirus restriction, and counteraction by HIV-1 protease. Nat Commun 9:4137.

18. Chesarino NM, Emerman M. 2020. Polymorphisms in Human APOBEC3H Differentially Regulate Ubiquitination and Antiviral Activity. Viruses 12:378.

19. Starrett GJ, Luengas EM, McCann JL, Ebrahimi D, Temiz NA, Love RP, Feng Y, Adolph MB, Chelico L, Law EK, Carpenter MA, Harris RS. 2016. The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. 1. Nat Commun 7:12918.

20. 1930. Cancer Facts & Figures 2022 80.
21. Burns, M. B., L. Lackey, M. A. Carpenter, A. Rathore, A. M. Land, B. Leonard, E. W. Refsland, D. Kotandeniya, N. Tretyakova, J. B. Nikas, D. Yee, N. A. Temiz, D. E. Donohue, R. M. McDougle, W. L. Brown, E. K. Law, and R. S. Harris. 2013. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 494: 366–370.
22. Zou, J., C. Wang, X. Ma, E. Wang, and G. Peng. 2017. APOBEC3B, a molecular driver of mutagenesis in human cancers. *Cell Biosci* 7.
23. Nik-Zainal, S., D. C. Wedge, L. B. Alexandrov, M. Petljak, A. P. Butler, N. Bolli, H. R. Davies, S. Knappskog, S. Martin, E. Papaemmanuil, M. Ramakrishna, A. Shlien, I. Simonic, Y. Xue, C. Tyler-Smith, P. J. Campbell, and M. R. Stratton. 2014. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nat Genet* 46: 487–491.
24. Stratton MR, Campbell PJ, Futreal PA. 2009. The cancer genome. Nature 458:719–724.
25. Gurtan, A. M., and P. A. Sharp. 2013. The role of miRNAs in regulating gene expression networks. *J. Mol. Biol.* 425: 3582–3600.
26. Friedman, R. C., K. K.-H. Farh, C. B. Burge, and D. P. Bartel. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 19: 92–105.
27. Matoulkova, E., E. Michalova, B. Vojtesek, and R. Hrstka. 2012. The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biol* 9: 563–576.
28. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, Dreyfuss G, Eddy SR, Griffiths-Jones S, Marshall M, Matzke M, Ruvkun G, Tuschl T. 2003. A uniform system for microRNA annotation. RNA 9:277–279.
29. Desvignes T, Batzel P, Berezikov E, Eilbeck K, Eppig JT, McAndrews MS, Singer A, Postlethwait JH. 2015. microRNA nomenclature: A view incorporating genetic origins, biosynthetic pathways, and sequence variants. Trends Genet 31:613–626.
30. Rusk, N. 2008. When microRNAs activate translation. *Nat Methods* 5: 122–123.
31. Yang Q, Cao W, Wang Z, Zhang B, Liu J. Regulation of cancer immune escape: The roles of miRNAs in immune checkpoint proteins. *Cancer Letters*. 2018;431:73-84. doi:10.1016/j.canlet.2018.05.015
32. Peng Y, Croce CM. 2016. The role of MicroRNAs in human cancer. 1. Sig Transduct Target Ther 1:1–9.
33. Otmani K, Lewalle P. 2021. Tumor Suppressor miRNA in Cancer Cells and the Tumor Microenvironment: Mechanism of Deregulation and Clinical Implications. Front Oncol 11:708765.

34. Home - miRWalk. http://mirwalk.umm.uni-heidelberg.de/. Retrieved 28 February 2022.
35. Kozomara A, Birgaoanu M, Griffiths-Jones S. 2019. miRBase: from microRNA sequences to function. Nucleic Acids Research 47:D155–D162.
36. Liu W, Wang X. 2019. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. Genome Biol 20:18.
37. Griffiths-Jones S. 2004. The microRNA Registry. Nucleic Acids Research 32:109D – 111.
38. Betel D, Wilson M, Gabow A, Marks DS, Sander C. 2008. The microRNA.org resource: targets and expression. Nucleic Acids Res 36:D149–D153.
39. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. 2004. Human MicroRNA Targets. PLoS Biol 2:e363.
40. Ding J, Li X, Hu H. 2016. TarPmiR: a new approach for microRNA target site prediction. Bioinformatics 32:2768–2775.
41. Helwak A, Kudla G, Dudnakova T, Tollervey D. 2013. Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding. Cell 153:654–665.
42. UCSC Genome Browser Home. https://genome.ucsc.edu/index.html. Retrieved 28 February 2022.
43. Primer3Plus. https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi. Retrieved 28 February 2022.
44. Protocol for Q5® High-Fidelity 2X Master Mix | NEB.
45. NEB® PCR Cloning Kit | NEB.
46. Longo PA, Kavran JM, Kim M-S, Leahy DJ. 2013. Transient mammalian cell transfection with polyethylenimine (PEI). Methods Enzymol 529:227–240.
47. Dual-Glo® Luciferase Assay System. https://www.promega.com/products/luciferase-assays/reporter-assays/dual_glo-luciferase-assay-system/. Retrieved 2 May 2022.
48. 2022. Homo sapiens microRNA 1207 (MIR1207), microRNA.
49. Bertolazzi G, Cipollina C, Benos PV, Tumminello M, Coronnello C. 2020. miR-1207-5p Can Contribute to Dysregulation of Inflammatory Response in COVID-19 via Targeting SARS-CoV-2 RNA. Frontiers in Cellular and Infection Microbiology 10. Oncotarget 7:32421–32432.
50. A549 | ATCC. https://www.atcc.org/products/ccl-185. Retrieved 2 May 2022.
51. Papagregoriou G, Erguler K, Dweep H, Voskarides K, Koupepidou P, Athanasiou Y, Pierides A, Gretz N, Felekkis KN, Deltas C. 2012. A miR-1207-5p Binding Site Polymorphism Abolishes Regulation of HBEGF and Is Associated with Disease Severity in CFHR5 Nephropathy. PLOS ONE 7:e31021.
52. Das DK, Osborne JR, Lin H-Y, Park JY, Ogunwobi OO. 2016. miR-1207-3p Is a Novel Prognostic Biomarker of Prostate Cancer. Transl Oncol 9:236–241.
53. miRNA Entry for MI0006316. https://www.mirbase.org/cgi-bin/mirna_entry.pl?acc=MI0006316. Retrieved 5 May 2022.
54. Cao X, Xu L, Liu Q, Yang L, Li N, Li X. 2019. MicroRNA-1277 Inhibits Proliferation and Migration of Hepatocellular Carcinoma HepG2 Cells by Targeting and Suppressing BMP4 Expression and Reflects the Significant Indicative Role in Hepatocellular Carcinoma Pathology and Diagnosis After Magnetic Resonance Imaging Assessment. Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics 27:301–309.

55. Chin A, Mariscal J, Kim M, Guerra G, Victor B, Qian C, Broseghini E, Posadas E, Freeman MR, Sharma S, Gandellini P, Zaffaroni N, You S, Chan KS, Guarnerio J, Fabbri M, Di Vizio D. 2021. miR-1227 Targets SEC23A to Regulate the Shedding of Large Extracellular Vesicles. 22. Cancers 13:5850.
56. Cui J, Kang X, Shan Y, Zhang M, Gao Y, Wu W, Chen L. 2022. miR-1227-3p participates in the development of fetal growth restriction via regulating trophoblast cell proliferation and apoptosis. 1. Sci Rep 12:6374.
57. miRNA Entry for MI0003612. https://www.mirbase.org/cgi-bin/mirna_entry.pl?acc=hsa-mir-548a-3. Retrieved 5 May 2022.
58. Zhao G, Wang T, Huang Q-K, Pu M, Sun W, Zhang Z-C, Ling R, Tao K-S. 2016. MicroRNA-548a-5p promotes proliferation and inhibits apoptosis in hepatocellular carcinoma cells by targeting Tg737. World J Gastroenterol 22:5364–5373.
59. Bazzoni F, Rossato M, Fabbri M, Gaudiosi D, Mirolo M, Mori L, Tamassia N, Mantovani A, Cassatella MA, Locati M. 2009. Induction and regulatory function of miR-9 in human monocytes and neutrophils exposed to proinflammatory signals. Proc Natl Acad Sci U S A 106:5282–5287.
60. Santhekadur PK, Kumar DP. 2019. RISC assembly and post-transcriptional gene regulation in Hepatocellular Carcinoma. Genes Dis 7:199–204.
61. Redfern AD, Colley SM, Beveridge DJ, Ikeda N, Epis MR, Li X, Foulds CE, Stuart LM, Barker A, Russell VJ, Ramsay K, Kobelke SJ, Li X, Hatchell EC, Payne C, Giles KM, Messineo A, Gatignol A, Lanz RB, O'Malley BW, Leedman PJ. 2013. RNA-induced silencing complex (RISC) Proteins PACT, TRBP, and Dicer are SRA binding nuclear receptor coregulators. Proc Natl Acad Sci U S A 110:6536–6541.
62. Wang Y, Schmitt K, Guo K, Santiago ML, Stephens EB. 2016. Role of the single deaminase domain APOBEC3A in virus restriction, retrotransposition, DNA damage and cancer. J Gen Virol 97:1–17.
63. McCann JL, Klein MM, Leland EM, Law EK, Brown WL, Salamango DJ, Harris RS. 2019. The DNA deaminase APOBEC3B interacts with the cell-cycle protein CDK4 and disrupts CDK4-mediated nuclear import of Cyclin D1. Journal of Biological Chemistry 294:12099–12111.
64. Wu J, Li N, Zhu L, Zhen D, Li M, Chen H, Ye M, Wei Y, Shao G. 2022. Multiple roles of apolipoprotein B mRNA editing enzyme catalytic subunit 3B (APOBEC3B) in human tumors: a pan-cancer analysis. BMC Bioinformatics 23:312.
65. Bandarra S, Miyagi E, Ribeiro AC, Gonçalves J, Strebel K, Barahona I. 2021. APOBEC3B Potently Restricts HIV-2 but Not HIV-1 in a Vif-Dependent Manner. Journal of Virology 95:e01170-21.
66. Constantin D, Dubuis G, Conde-Rubio M del C, Widmann C. 2022. APOBEC3C, a nucleolar protein induced by genotoxins, is excluded from DNA damage sites. FEBS J 289:808–831.
67. Wittkopp CJ, Adolph MB, Wu LI, Chelico L, Emerman M. 2016. A Single Nucleotide Polymorphism in Human APOBEC3C Enhances Restriction of Lentiviruses. PLOS Pathogens 12:e1005865.
68. Duggal NK, Fu W, Akey JM, Emerman M. 2013. Identification and antiviral activity of common polymorphisms in the APOBEC3 locus in human populations. Virology 443:329–337.

69. Takei H, Fukuda H, Pan G, Yamazaki H, Matsumoto T, Kazuma Y, Fujii M, Nakayama S, Kobayashi IS, Shindo K, Yamashita R, Shirakawa K, Takaori-Kondo A, Kobayashi SS. 2020. Alternative splicing of APOBEC3D generates functional diversity and its role as a DNA mutator. Int J Hematol 112:395–408.
70. Ara A, Love RP, Follack TB, Ahmed KA, Adolph MB, Chelico L. 2017. Mechanism of Enhanced HIV Restriction by Virion Coencapsidated Cytidine Deaminases APOBEC3F and APOBEC3G. J Virol 91:e02230-16.
71. Call E. How to get genomic information from museum specimens of Lepidoptera 83.

72. Invitrogen (2010). pcDNA3_1_man.pdf.

73. Revathidevi S, Manikandan M, Rao AKDM, Vinothkumar V, Arunkumar G, Rajkumar KS, Ramani R, Rajaraman R, Ajay C, Munirajan AK. 2016. Analysis of APOBEC3A/3B germline deletion polymorphism in breast, cervical and oral cancers from South India and its impact on miRNA regulation. Tumor Biol 37:11983–11990.