Identifying Long Non-Coding RNAs Associated with a Protective Type I Interferon Signature in Melanoma

Identifying Long Non-Coding RNAs Associated with a Protective Type I Interferon Signature in Melanoma

By Stephanie Yuan

Advisor: Dr. Carl Novia

A Thesis Submitted to the Faculty of

the Harvard Medical School

in Partial Fulfilment of the Requirements

for the Degree of Master of Medical Sciences in Immunology

Harvard Medical School

Boston, Massachusetts

May 2022

Abstract

Long non-coding RNAs (lncRNAs) have emerged as a new class of regulators of type I interferon (IFN) signaling and can potentially serve as therapeutic targets.[1] In this study, we identified lncRNAs that are associated with both a strong type I IFN response, characterized by increased interferon-stimulated gene (ISG) expression, and high survival rate in cancer patients. In particular, we are interested in the effect of lncRNA-induced differential ISG expression in the context of melanoma. To study the effects of lncRNA on ISG expression, we performed gain-of-function experiments in two systems – A375 melanoma cell line and melanoma short-term culture (MSTC). We used Lipofectamine to transfect plasmids containing lncRNA of interest into these systems, and then performed RT-qPCR to assess ISG expression. We used the term "ISG signature" to refer to the unique gene expression differences observed upon lncRNA overexpression in our experimental set up. Here, we identified one lncRNA, *HCP5-204*, to be significantly associated with a protective ISG signature in melanoma, consisting of ISG15, STAT1, IRF9, JAK1, and TLR3. Future experiments will explore the mechanisms of action through which *HCP5-204* modulates type I IFN response in melanoma.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Firstly, I would like to thank my mentor, Dr. Carl Novina, for accepting me into the lab and for trusting me to work on this project. I would also like to thank Drs. Eleni (Lena) Anastasakou, Kushani Shah, and Aya Kobayashi for providing me guidance and support. I also enjoyed the working with the rest of the members of the Novina lab – Dr. Sabine Studer, Dr. Mohamad Sarmini, Chris Rauchet, Jonas Bohnlein, Shuming Liu, Simon Goller, Leo Motta-Zacks, Alina Shen, and Jovan Tormes – they made my time at lab a fun experience. I also thank the valuable work done by Dr. Leon Wert-Lamas and Ksenia Morozova – they laid the groundwork of my project and have also been very helpful at answering my questions.

Additionally, I would like to thank everyone in the Master of Sciences program in Immunology at Harvard Medical School. To Drs. Shiv Pillai, Michael Carroll, and Bobby Cherayil - thank you for having me in the program, sharing your experience and knowledge, and giving me invaluable guidance. To Dr. Gavin Porter – thank you for your amazing lectures last year. To Faith, Selina, and Naima – thank you so much for providing administrative support. The Class of 2022, especially my best friends Yidan Gao and Hang Yin, deserves a special mention here – thank you for your unconditional support during the past year.

Lastly, I would like to extend my gratitude to my parents – Dr. Jiandong Yuan and Wenhui (Ellen) Liang. My master's research journey would not have been possible without any one of the people mentioned.

Chapter 1: Background

1.1 Introduction to lncRNAs

Since the advent of advanced whole genome sequencing technologies, it has been revealed that majority of the human genome comprises non-protein coding sequences.[2,3] Increasing numbers of studies on the non-protein coding genome have uncovered the essential roles of long non-coding RNAs (lncRNAs) in regulating cellular and organismal activities. In 1991, Willard et al., discovered human X-inactive specific transcript (*XIST*) as one of the first lncRNAs, leading to the elucidation of the mechanism of X chromosome inactivation.[4] In brief, *XIST* is a 17,000-nucleotide-long transcript expressed from the X chromosome during embryonic development.[5] In female embryos which have 2 copies of the X chromosome, *XIST* coats one of the X chromosomes at random, marking it for long-term silencing by epigenetic modifications.[5,6] Since then, many more lncRNAs have been identified and subsequently structurally and functionally profiled, and their importance is getting increasingly recognized.

For several reasons, including but not limited to their relatively recent discovery, low abundance in cells, and limitations in current technologies, only a handful of lncRNAs have been well characterized and much more is to be investigated. LncRNAs have been broadly defined as a strand of non-protein coding transcript more than 200 nucleotides in length.[7] Only about 3% of all known lncRNAs have been functionally annotated.[8] Due to the plethora of RNA molecules that fit the general description, lncRNAs can be further categorized based on their genomic origins, hence some subcategories of lncRNAs include long intergenic non-coding RNAs (lincRNAs) – which are lncRNAs encoded in distinct genomic regions without overlapping protein-coding loci, and natural antisense transcripts, which are found antisense to protein-coding genes and

pseudogenes.[9,10] The biological significance of lncRNAs has been the topic of heated debate, as many argue that the majority of these molecules are simply "transcriptional noise," belittling lncRNAs as transcriptional byproducts.[11] However, with the advancement of sequencing tools, functional assays, and novel computational methods, lncRNAs have become much better annotated and functionally validated.[12]

1.1.1 Biogenesis, Processing, and Degradation

LncRNAs bear striking similarities to protein-coding messenger RNAs (mRNAs) but at the same time are fundamentally different. For instance, the biogenesis of most lncRNAs is almost identical to that of mRNAs – they are transcribed by RNA polymerase II (Pol II), co-transcriptionally spliced, 5' capped, and 3' polyadenylated, although there are exceptions.[13] LncRNAs can receive the same types of post-transcriptional modifications as mRNAs, such as N6-Adenosine methylation, subjecting them to metabolic regulation.[14]

While mRNAs are produced in the nucleus and then exported to the cytoplasm for participation in protein synthesis, lncRNAs may localize to the nucleus, cytoplasm, and even organelles.[15] The specific localization of each lncRNA varies and is closely linked to its function. Another distinction is that lncRNAs are generally present in much lower quantities, some as low as 0.0006 copies per cell.[16] Tissue- and temporal-specificity are both common aspects of mRNAs.[3] For example, in the developing *Drosophila* embryo a concentration gradient of mRNAs coding for transcription factors regulating genes with competing functions drive the regulation the posterior-anterior axis formation.[17] In the case of lncRNAs, cell type and temporal regulation are even more prominent than mRNAs, suggesting that lncRNAs have highly specialized functions relevant to a

specific cell type or developmental stage, respectively. Furthermore, the non-coding genome is under much less evolutionary pressure compared to protein-coding genes, suggesting that lncRNAs are also species-specific.[18]

Lastly, lncRNAs can be degraded via several different mechanisms. In the nucleus, lncRNAs can be degraded by exosomes.[19] In the cytoplasm, they might be subject to de-capping, and degraded by the exonuclease Xrn1.[20] Finally, another mechanism of lncRNA degradation is nonsense-mediated decay (NMD), a process that usually destroys mRNA containing premature translation-termination codons.[21] Due to structural similarities to mRNAs, lncRNAs can often be degraded through this mechanism, when the NMD machinery accidentally recognizes the lncRNA sequence as a termination codon.[7] For example, the cytoplasmic lncRNA Growth Arrest Specific 5 (*GAS5*) is found to be bound by NMD complex component, regulator of nonsense transcript 1 (UPF1) and then degraded.[22]

1.1.2 Functions and Modes of Actions of LncRNA

Following or during transcription, lncRNAs can act either in *cis*, in proximity to the site of transcription, or in *trans*, far away from the site it originated from. Some of the functions of lncRNAs characterized to date are regulating gene transcription, modifying pre-mRNA splicing, and facilitating protein-protein interactions (Figure 1).[13] Notably, lncRNAs usually exert their functions with help from associated proteins.[23]

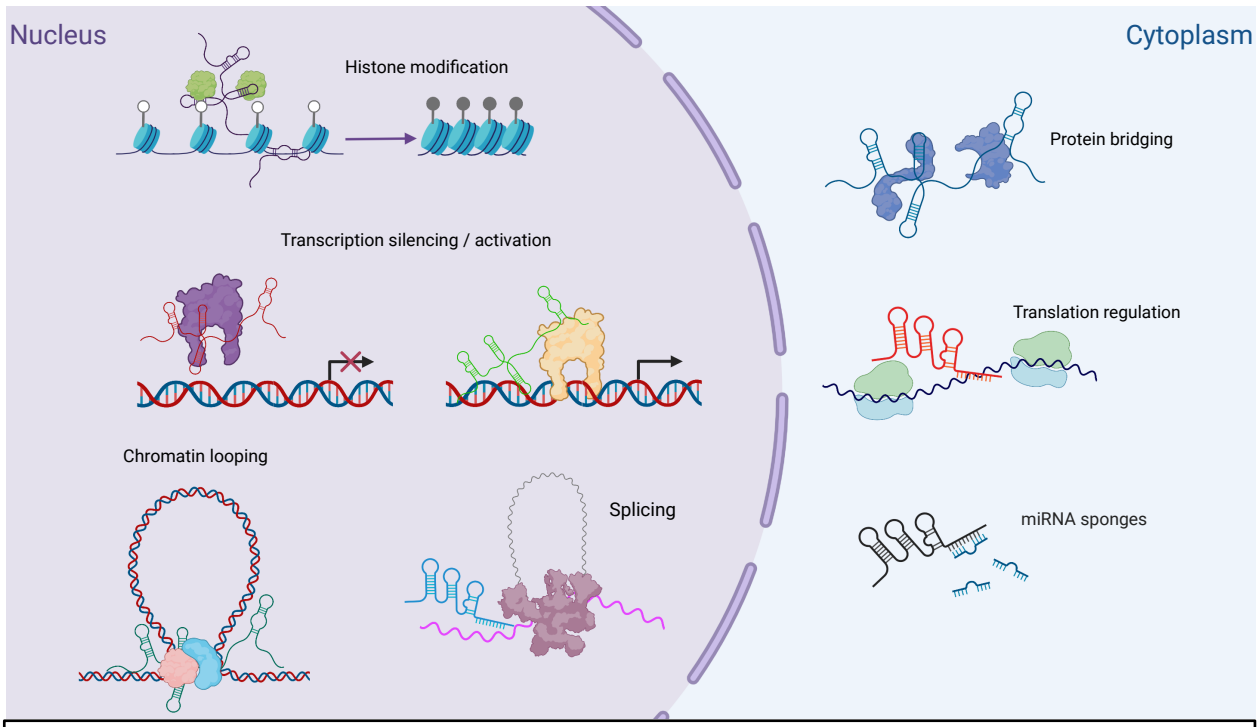


Figure 1: Illustration of known lncRNA modes of actions in the nucleus and cytoplasm. The light purple structure with porous membrane represents the nucleus, and light blue background represents the cytosol. Many lncRNAs are retained in the nucleus and exert several gene regulatory functions: recruiting histone modifying complexes to the chromatin; repress transcription by acting as decoys, thus blocking transcription factor binding to DNA; activate transcription by recruiting transcription factors to promoter; activate gene transcription by bringing enhancer elements and promoter together; finetune pre-mRNA splicing by interfering with spliceosomes. In the cytosol, lncRNAs can mediate protein-protein interactions by acting as scaffolds, or bridges; regulate translation by obstructing the translational machinery; disable miRNA function by acting as "sponges". Figure created with Biorender.

In the nucleus, lncRNAs can regulate gene transcription in *cis* or in *trans* through various modes of actions. Acting as decoys and guides for transcription factors and/or epigenetically modifying the chromatin, lncRNAs can activate or repress gene expression. An example of a chromatin modifying lncRNA is the HOXA distal transcript antisense RNA (*HOTTIP*), which recruits chromatin modifiers and forms a chromatin loop, bringing enhancers to the promoter region to activate HOXA gene transcription.[24] Notably, our study aims to characterize many antisense

lncRNAs, which can act in *cis* from their corresponding protein-coding sense genes,[25] modulating them at the transcriptional level , and sometimes even co-transcriptionally.

In the cytoplasm, lncRNAs can interact with proteins, mRNAs, and microRNAs (miRNAs), among others (Figure 1).[7] For instance, the lncRNA *GUARDIN*, also known as long non-coding transcriptional activator of miR34a, acts as a "sponge" for miR-23a, preventing the miRNA from binding to its target mRNA.[26] In addition to playing important roles in many cellular processes in the nucleus and cytoplasm, lncRNAs can also localize to and function in specific organelles.[27] An example is the survival associated mitochondrial melanoma specific oncogenic non-coding RNA (*SAMMSON*), which is localized to the mitochondria and participates in the regulation of mitochondrial metabolism.[28]

1.1.3 LncRNAs in Cancer

Because of their ability to regulate gene transcription, emerging studies are revealing the indispensable role of many lncRNAs in cancer.[29] Aberrant lncRNA expression is a reliable marker for cancer prognosis and diagnosis.[15] LncRNAs are highly tissue-specific and are generally expressed at different levels in tumors compared to healthy cells, making them ideal biomarkers.[13,29] For instance, overexpression of HOX antisense intergenic RNA (*HOTAIR*) in early stages of breast cancer is predictive of metastasis.[30,31] In addition to cancer detection, several studies recognized that lncRNAs can partake in pivotal cellular processes during oncogenesis. For instance, Leucci et al. found that aberrantly high expression of the lncRNA *SAMMSON* in melanoma correlated with uncontrolled cell growth.[32] Through RNA antisense purification-mass spectrometry (RAP-MS) and chromatin isolation by RNA purification mass spectroscopy (ChIRP-MS), it was demonstrated that *SAMMSON* facilitates the production of the mitochondrial protein

p32, a direct transcriptional target of the oncogene Myc.[28] To further investigate the therapeutic potential of this lncRNA, the researchers used locked nucleic acid (LNA) – modified antisense oligonucleotides called GapmeRs to inhibit the effects of *SAMMSON*.[28] In mouse models, the administration of GapmeRs against *SAMMSON* significantly curbed the growth of tumors and decreased the rate of cell proliferation, suggesting *SAMMSON* is an effective therapeutic target for melanoma treatment.[28] The example of *SAMMSON* highlights the potential therapeutic value of lncRNAs in the treatment of melanoma and serves as an example for the targeting of other lncRNAs for the specific treatment of different types of cancer.

Research from the Novina lab interrogated the tumor-promoting role of another melanoma-associated lncRNA, *SLNCR*.[33] Through ChIP-seq and RNA immunoprecipitation (IP), our lab discovered that *SLNCR* binds to both androgen receptor (AR) and transcription factor early growth response protein 1 (EGR1). Through this interaction, *SLNCR* functions as an anchor by recruiting AR to EGR1-occupied genomic loci. Finally, our lab found that co-localization of the AR-*SLNCR*-EGR1 regulatory triad at EGR1 binding sites negates EGR1-mediated p53-independent upregulation of the p21 tumor suppressor gene, unleashing melanoma proliferation.[33] In summary, findings from our lab connected lncRNA biology to androgen-independent oncogenic activity of AR in melanoma.[33] Given that AR is a hormonal receptor specialized in male-hormone signaling and is therefore more abundant in male tissues, this discovery helped to further explain why males are more prone to developing melanoma.[34,35]

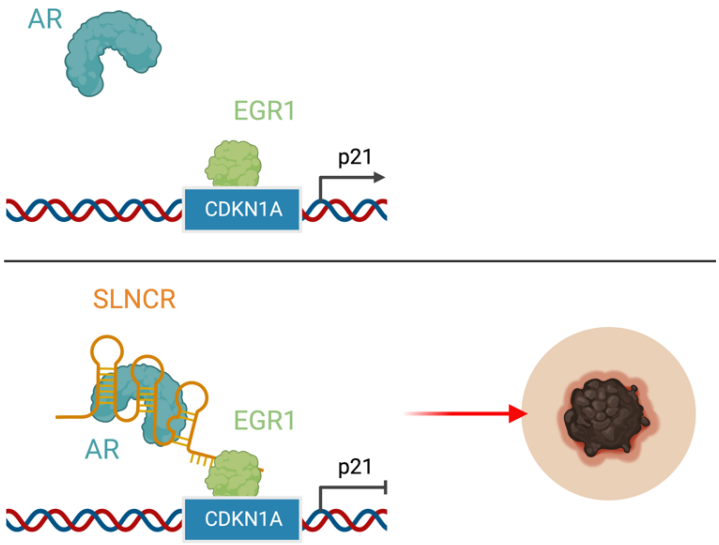


Figure 2: Example of a lncRNA that plays a regulatory role in melanoma – *SLNCR* – which represses transcription of p21, leading to uncontrolled melanoma cell proliferation. In normal cells, where there is low abundance of *SLNCR*, the transcription factor early growth response protein 1 (EGR1) binds to the cyclin dependent kinase inhibitor 1A (CDKN1A) promoter, turning on p21 transcription. p21 is a tumor suppressor that inhibits the activity of CDK proteins, curbing cell growth. In melanoma cells, there is a high abundance of *SLNCR*, which recruits androgen receptors (AR) to the EGR1 binding site. This regulatory triad inhibits the transcription of p21, leading to uncontrolled cell growth, and therefore melanomagenesis. Figure created with BioRender.

1.2 Type I Interferon Response

Type I interferons (IFNs) constitute a family of cytokines that are primarily involved in inducing the antiviral state.[36] These cytokines are also known to play significant roles in cancer, although their effect could be beneficial or detrimental, depending on context.[37] The type I IFN family consists of the 13 partially homologous IFNα subtypes and a single IFNβ subtype in humans, among other less well characterized subtypes.[38] The type I IFN response is an essential immune pathway – almost all nucleated cells express the IFNα receptor (IFNAR) and can produce these cytokines.[39] The production of type I IFNs can be induced by an array of different viral pathogen-

associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs).[40] When pattern recognition receptors (PRRs), which are highly specialized pathogen sensors, detect these foreign molecules, a series of downstream events that ultimately lead to anti-pathogen effects is triggered (Figure 3).

Type I IFN binding to IFNARs activates the Janus Kinase – Signal Transducer and Activator of Transcription (JAK-STAT) signal transduction pathway, which in turn can activate transcription of IFN-stimulated genes (ISGs)[41] (Figure 3). The JAK-STAT pathway is the canonical pathway through which type I IFN response is signaled, though there are other pathways that can induce transcription of ISGs.[42] There are hundreds of known ISGs which include signal transducers like STATs, PRRs like toll-like receptors (TLRs) and cyclic GMP-AMP synthase (cGAS), and IFN response factors (IRFs), among others.[43] These genes code for proteins that not only have anti-viral functions but are also pro-apoptotic and pro-inflammatory. Transcription activation of ISGs following detection of PAMPs and DAMPs, therefore, also enhance and prolong the type I IFN response. Interestingly, one mechanism inhibiting the JAK-STAT signal transduction pathway is also mediated through type I IFN signaling. In a type I IFN-dependent manner, the ISG ubiquitin specific peptidase 18 (USP18) negatively regulates JAK-STAT signaling by outcompeting JAK1 binding to IFNAR subunit 2 (IFNAR2).[44,45] Through this negative feedback loop, type I IFNs induce the transcription of an ISG that downregulates its response.

Elevated ISG expression is an important indicator of the induction of type I IFN response.[43] In this study, we focused on a subset of ISGs provided by Takeda Pharmaceuticals, which had been previously linked with a desirable signature in cancer patient samples and increased patient

survival (data not shown). Hence, in this study we treat ISG upregulation as an indicator of both induction of type I IFN response and its associated anti-cancer effects.

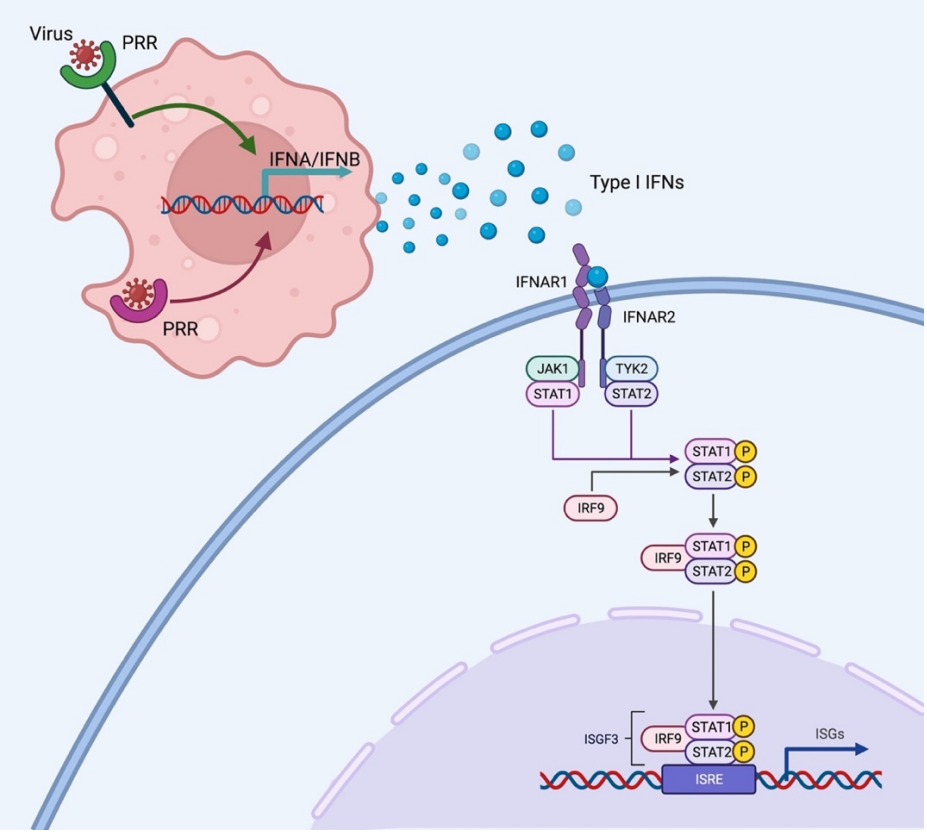


Figure 3: The Janus kinase (JAK) - signal transducer and activator of transcription (STAT) signal transduction pathway. Type I interferons (IFNs) are produced in response to recognition of pathogen associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs) by pattern recognition receptors (PRRs). Several different pathways downstream of PRR signaling initiates transcription of IFN genes and the subsequent production of type I IFNs. These cytokines then bind to IFNα receptors (IFNARs), after which the receptor subunits (IFNAR1, IFNAR2) dimerize, activating the JAK1 and tyrosine kinase 2 (TYK2) kinases. The activation of these receptor-bound kinases then leads to the phosphorylation of STAT1 and STAT2, causing them to heterodimerize and dissociate into the cytoplasm. The heterodimer is then bound by IRF9, forming a triplex, now called the IFN-stimulated gene factor 3 (ISGF3). ISGF3 translocates to the nucleus and binds to the IFN-stimulated response element (ISRE) in the IFN-stimulated gene (ISG) promoters, initiating transcription of ISGs. ISGs code for proteins that establish the antiviral state. Figure created with BioRender.

1.3 LncRNAs and Type I IFN Signaling

To date, only a handful of lncRNAs have been characterized in the context of the type I IFN response during viral infections. For example, Kambara et al. performed high-throughput RNA sequencing (RNA-seq) analysis on IFNα-induced primary human hepatocytes to identify lncRNAs that are involved in type I IFN signaling in the context of hepatitis C virus (HCV) infection.[46] From hundreds of lncRNAs that were induced by type I IFN signaling in hepatocytes, the researchers found cytidine/uridine monophosphate kinase 2 (*CMPK2)* to be most highly responsive. Subsequently, they found that knocking down the lncRNA-CMPK2 gene using short hairpin RNA (shRNA) results in the upregulation in a number of antiviral ISGs.[46] This finding suggests that (1) *lncRNA-CMPK2* is a downstream effector of type I IFN and (2) it inhibits the IFN response, engaging in a negative feedback loop.

From another study conducted by Kambara et al., they discovered the lncRNA BST2 IFN-stimulated positive regulator (*BISPR*) to be a positive regulator of BST2 (an ISG) transcription.[47] First, *BISPR* expression was highly induced upon IFNα stimulation. However, shRNA knockdown of STAT2 and treatment with ruxolitinib, a JAK inhibitor, independently suppressed *BISPR* expression in the presence of IFNα treatment.[47] Therefore, *BISPR* induction is mediated through the JAK-STAT signal transduction pathway. Furthermore, knocking down and overexpressing *BISPR* in various cells types resulted in downregulation and upregulation of BST2, respectively.[47] Because the genes for *BISPR* and BST are located in proximity and even share a bidirectional promoter, Kambara et al. hypothesize that this lncRNA mediates the expression of the neighboring protein-coding gene in *cis*.[47] Similarly in our study, we aim to study the effect of lncRNA on ISG

expression, although we are interested in characterizing the lncRNAs with a set of differentially upregulated ISGs.

Despite efforts to study the effect of lncRNAs in immune responses, many questions in this field of research remain to be answered. This study explores the interplay between lncRNAs, type I IFN response, and cancer. Former members of the Novina lab have identified abundantly expressed lncRNAs that are associated with strong IFN response based on cancer patient data.[48]

1.4 Prior Work from the Novina lab

Previously in the Novina Lab, postdoctoral fellow Leon Wert-Lamas and his student Ksenia Morozova used bioinformatic analysis to identify lncRNAs relevant in melanoma. Specifically, Leon and Ksenia scouted The Cancer Genome Atlas (TCGA)[3,49] database for melanoma-related lncRNAs associated with (1) a "desirable" IFN signature, defined by upregulation of a subset of ISGs that had previously been linked to higher tumor immunogenicity and (2) with higher patient survival rates (Figure 4A). Individual lncRNAs were assigned an IFN score (IFNS) to express the correlation of those two factors in a numerical way.[48] A total of 31 lncRNAs with highest IFNS, i.e. they were most differentially expressed in melanoma patients with high survival rates and correlated with high expression of desirable interferon response genes, were chosen for downstream validation (Figure 4B).

For the experimental validation of the selected 31 melanoma-associated lncRNAs, Ksenia Morozova preformed a small interfering RNA (siRNA)-mediated knockdown experiment in HEK Lucia Null cells that have a *Renilla* luciferase reporter gene coupled to an IFN-responsive

ISG54/ISRE promoter.[48] Luciferases are naturally existing proteins which catalyze the breakdown of their substrate luciferin and emit light as a reaction by-product.[50] Hence, in this experimental setup, where the production of luciferase protein is IFN-responsive, increased IFN signaling can be represented by bioluminescence emitted by *Renilla* luciferase upon addition of luciferin. The goal was to investigate whether the lncRNAs function upstream of type I IFNs, manipulating the signaling of these cytokines.

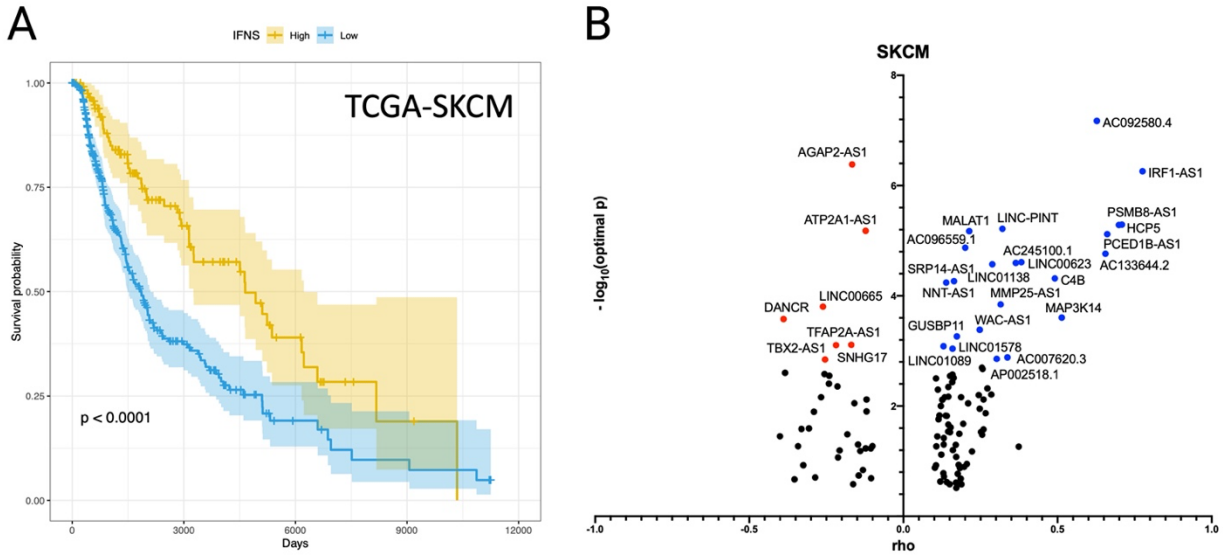


Figure 4: Strong type I interferon signaling is significantly associated with better disease outcomes in melanoma patients. (A) Survival plot of melanoma patients with high and low interferon scores. High interferon score (IFNS) represents a stronger type I IFN response, defined by the upregulation of the list of 59 IFN-stimulated genes (ISGs). Survival probability of melanoma patients are plotted on the y-axis and number of days are plotted on x-axis. Average survival probability of high IFNS patients is shown in yellow, whereas that of low IFNS patients is in blue. Difference between high and low IFNS patients is statistically significant, with p-value <0.0001 (log-rank test). (B) Scatter plot of lncRNAs significantly associated with melanoma. Spearman's correlation with the IFNS is plotted on the x-axis and statistical significance on the y-axis. Positively correlated lncRNAs are highlighted in blue, while negatively correlated ones are in red. Data were obtained from the Cancer Genome Atlas (TCGA, project ID: TCGA-SKCM). Data analyzed and graph generated by Leon Wert-Lamas and Ksenia Morozova.

1.5 Project objectives

Despite the enormous number of transcripts that fall into the definition of lncRNAs, few have been well characterized to date. Furthermore, the role of lncRNAs in one of the most important immune responses, the type I IFN response, remains largely unstudied. This project aims to fill the gap in our current understanding of the role of lncRNAs in the type I IFN response, in the context of melanoma.

The main objective of my part of this project is to identify lncRNAs that are implicated in type I IFN response and to characterize them with a set of ISG, or what we call the "ISG signature." Previous work has found that a strong type I IFN response, defined by significant upregulation of ISGs resulting in a high IFNS, is significantly correlated with increased survival in melanoma patients (Figure 4).[48] Thus, we aim to identify and characterize lncRNAs that modulate this response by profiling ISG expression in melanoma cells overexpressing the candidate lncRNAs.

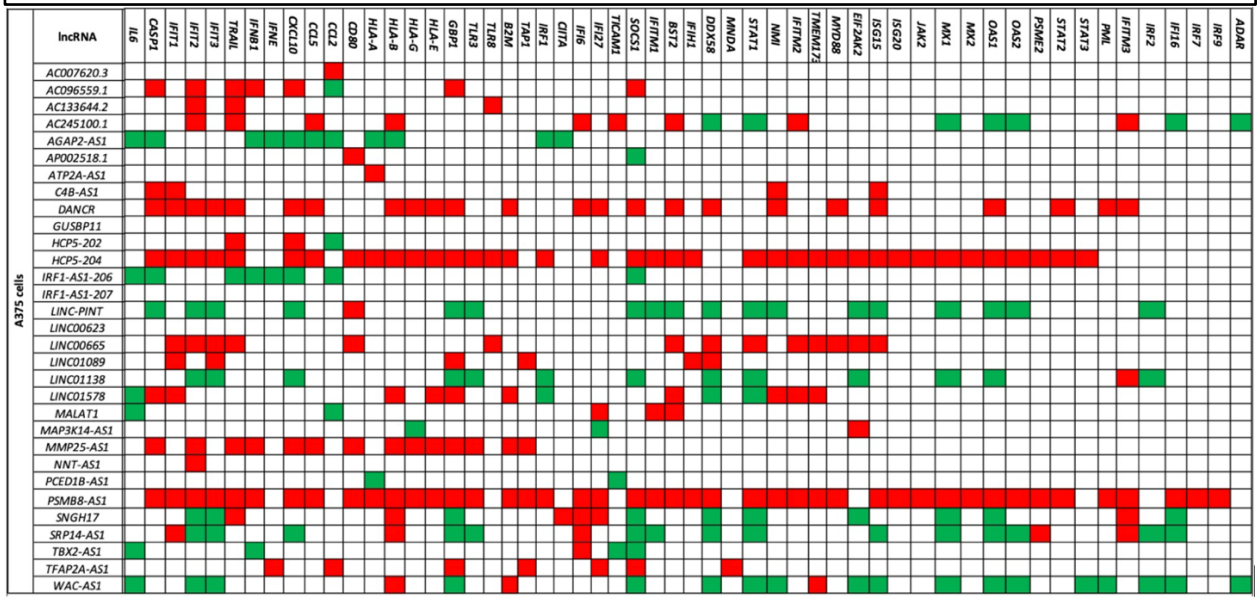
Chapter 2: Results

2.1 Chapter Introduction

The goal of this study is to identify lncRNAs connected to a strong type I IFN signature in the context of cancer. To do this, Dr. Leon Wert-Lamas, a past postdoctoral fellow in our lab combined information mined from TCGA[51,52] database to identify abundantly expressed lncRNAs associated with higher survival rates in patients with different cancer types. The combination of this information generated the "lncRNA IFN score (IFNS)." Through this pipeline, Leon showed that high IFNS in patients with melanoma significantly correlated with high survival rate. Subsequently through bioinformatic analysis, Leon selected the top 31 lncRNAs from TCGA database that are highly correlated with a high IFNS for *in vitro* characterization.

As a next step, Leon and his student Ksenia Morozova characterized the changes in expression of a list of 59 ISGs upon lncRNA overexpression in the A375 melanoma cell line using Nanostring.[53] This set of ISGs was provided by Takeda Pharmaceuticals and had been associated with a desirable IFN signature, positively correlating with increasing tumor immunogenicity without increasing cancer adaptability and maliciousness. Nanostring results revealed the ISG expression profiles in A375s overexpressing different selected lncRNAs (Table 1). These gene expression profiles serve as the basis for our selection of candidate lncRNAs to study in the current stage of the project.

Table 1: Summary of the changes in ISG expression following lncRNA overexpression in A375 cells. Nanostring was used to profiles the interferon-stimulated gene (ISG) expression of lncRNA-overexpressing A375s. RNA samples from at least n=2 biological replicates were sent for each condition, and was normalized to those of A375s transfected with an empty pcDNA3.1-EGFP vector. The "lncRNA" column lists the 31 lncRNAs found to be associated with melanoma from bioinformatics analysis. The ISGs are listed on the top row, and upregulation of each of the ISGs is indicated by green while downregulation is indicated by red. Blank grids indicate no significant differential expression of the corresponding ISG. Data presented in this table were work by Ksenia Morozova and Leon Wert-Lamas.[48]



2.2 Materials and Methods

2.2.1 DNA cloning

The 5 candidate lncRNAs were cloned into a pcDNA3.1 backbone that has a CAG promoter and puromycin selection gene. The lncRNA gene sequences were ordered from Integrated DNA Technologies (IDT) and amplified through polymerase chain reaction (PCR) using the Q5® High-Fidelity PCR Kit (New England BioLabs). The vector pcDNA3.1/Puro-CAG-VSFP-CR used to insert the lncRNA genes into was a gift from Michael Lin (Addgene plasmid #40257; http://n2t.net/addgene:40257; RRID:Addgene_40257). Restriction digestion was used to create sticky ends on the lncRNA gene and backbone, and subsequently ligated at 16°C overnight. Plasmids for lncRNAs *AGAP2-AS1* and *PSMB8-AS1* were directly ordered from Genewiz. Plasmids are amplified using XL10-Gold ultracompetent cells (Agilent) and extracted using HiSpeed Plasmid Maxi Kit (Qiagen). Plasmids are eluted in 32uL Plasmid concentration and purity are checked by measuring absorbance at 260 and 280nm using Nanodrop.

2.2.2 Primer Design and Validation

Sequences of lncRNAs and ISGs were obtained from Ensembl Genome Browser and the National Center for Biotechnology Information (NCBI) databases. Primers for RT-qPCR were aligned to the human transcriptome using the NCBI primer BLAST tool. Subsequently, primers were validated by qPCR and gel electrophoresis. Ensembl IDs and NCBI IDs of the lncRNAs and ISGs are listed in Table 3 and Appendix Table A1, respectively.

2.2.3 Cell culture

A375s were cryopreserved and purchased from the American Type Culture Collection (ATCC). Melanoma short term cultures were from the Wistar Institute collection. All cells were incubated at 37°C and maintained at 95% humidity and 5% CO_2. A375s and MSTCs were grown in

Dulbecco's Modified Eagle Medium (DMEM) ([+] 4.5 g/L D-glucose, [+] L-glutamine, [+] 1mM sodium pyruvate) (Corning, Gibco) supplemented with 10% fetal bovine serum (FBS) (Sigma) and 1% penicillin/streptomycin (Gibco). Adherent cells were detached from cell culture-treated plates with trypsin (ThermoFisher) and passaged when they reached ~90% confluency. Bambanker solution (GC Lymphotec) was used to freeze cells.

2.2.4 LncRNA Overexpression

The lncRNAs of interest were overexpressed in A375 and MSTCs respectively by transfection of an overexpression plasmid into the cells for transient expression. Transfection of A375s was done using Lipofectamine 2000 (Invitrogen) at a 1:1 ratio. Transfection in MSTCs was done using Lipofectamine LTX (Invitrogen). Transfections were performed in 6-well format, with biological triplicates for each condition and following the manufacturer's protocol. Briefly, 0.3 million A375 cells / 0.5-0.7 million MSTC WM1575 cells were seeded and transfected with 1 µg of plasmid at 24 hours post seeding. Cells are maintained in transfection media for 48 hours, and then were changed to selection media containing puromycin (Gibco) at concentration 1.5 µg/mL for an additional 48 hours. IFN-treated groups are changed to media with 1.5 µg/mL puromycin and 10.42 ng/mL human interferon-α1 (hIFN-α1) (Cell Signaling).

2.2.5 Total RNA Extraction

Transfected cells are pelleted at 500 x g for 5 minutes and RNA was extracted using the RNeasy Plus RNA extraction kit (Qiagen). RNA concentration and purity are checked by measuring absorbance at 260 and 280 nm using Nanodrop.

2.2.6 Quantitative Reverse Transcription Polymerase Chain Reaction (RT-qPCR)

Equal amounts of cDNA were generated for each sample through RT-PCR using the SuperScript® III First-Strand Synthesis SuperMix (ThermoFisher). cDNA and qPCR primers were added to SsoAdvanced Universal SYBR® Green Supermix (Bio-rad) for a total volume of 10 µL per reaction. Expression levels of lncRNAs and ISGs were checked by RT-qPCR on the CFX384 Touch™ Real-Time PCR Detection System (Bio-rad). Data analysis was performed using the 2-ΔΔCt method[54] with the aid of Microsoft Excel and GraphPad Prism.

Table 2: List of reagents used for cell-based assays. Each column, in order, lists name of reagent, manufacturer name, and catalog number.

Name	Manufacturer	Catalog number
Dulbecco's Modified Eagle Medium (DMEM)	Gibco	11971025
	Corning	10-017-CV
Sodium pyruvate	Gibco	11360070
Fetal bovine serum (FBS)	Sigma	9048-46-8
Penicillin/streptomycin	Gibco	15140122
Trypsin	ThermoFisher	27250018
Puromycin	Gibco	A1113803
BAMBANKER	GC Lymphotec	BB01
Lipofectamine 2000	Invitrogen	11668027
Lipofectamine LTX with Plus Reagent	Invitrogen	15338030
Human interferon-α1	Cell Signaling	8927SC
Plasmid Mini kit	Qiagen	12123
HiSpeed Plasmid Maxi kit	Qiagen	12663
Plasmid Maxi kit	Zymo	D4202
RNeasy Plus RNA Extraction kit	Qiagen	74034
SuperScript® III First-Strand Synthesis SuperMix	Invitrogen	18080400
SsoAdvanced Universal SYBR® Green Supermix	Bio-rad	1725274

2.3 Results

2.3.1 Selection of lncRNAs significantly associated with type I IFN response and survival in melanoma

Previously in our lab, Ksenia Morozova performed Nanostring analysis[53] on the model melanoma cell line A375 transiently expressing our lncRNAs of interest. Using Nanostring, we profiled the differential ISG expressions upon overexpressing the 31 lncRNAs in A375s (Table 1) and utilized these results to select candidates for subsequent validations. Because we aim to identify lncRNAs linked to a gene signature, our rationale for selecting the lncRNAs was that they show a consistent pattern of change in gene expression. By establishing this criterion, we identified 5 lncRNAs – *AGAP2-AS1, PSMB8-AS1, HCP5-204, MMP25-AS1,* and *TFAP2A-AS1* – whose overexpression correlated with an IFN signature, characterized by global, unidirectional changes in ISG mRNA expression. Those lncRNAs were selected for subsequent *in vitro* validation in human primary cells whilst other lncRNAs which are associated with both up- and downregulation of ISGs were excluded from downstream analysis for the present time.

Table 3: Information in the 5 selected lncRNAs. Each column, in order, lists the lncRNA name followed by isoform number, length in basepairs, genomic locus, forward or reverse strand, and Ensembl ID. "AS" in the names of the lncRNA genes designates "anti-sense".

LncRNA name with isoform number	Length (basepairs)	Genomic locus	Forward or reverse strand	Ensembl ID
AGAP2-AS1-201	1500	Chromosome 12 58,120,054-58,122,139	Forward	ENST00000542466.2
PSMB8-AS1-201	1015	Chromosome 6: 32,844,078-32,846,500	Forward	ENST00000412095.1
HCP5-204	1689	Chromosome 6: 31,463,170-31,478,936	Forward	ENST00000541196.3

TFAP2A-AS1-202	910	Chromosome 6: 10,409,340-10,416,446	Forward	ENST00000443546.1
MMP25-AS1-209	410	Chromosome 16: 3,051,096-3,056, 232	Reverse	ENSG00000261971*

Working with primary cells is limited to a low number of cell passages due to do the short lifespan of those cells outside the patient. Hence, we first needed to establish a solid experimental workflow before moving into working with melanoma short-term cultures (MSTCs). To do so, we set to reproduce the Nanostring results in A375s by RT-qPCR.

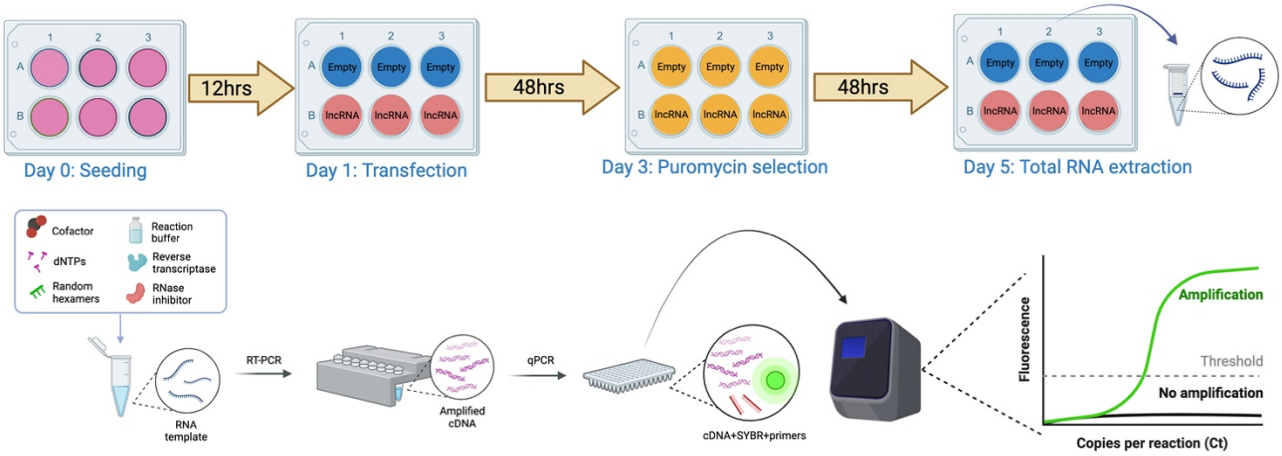


Figure 5: Diagram of experimental design: transfection procedure and subsequent assays for quantification of lncRNA and ISG expression levels. Cells were seeded in 6-well plates on day 0 and transfected with Lipofectamine the next day. Cells were incubated in transfection media for 48 hours before puromycin (1.5ug/mL) selection, which proceeded for another 48 hours. At the 24-hour point of puromycin selection, the IFN-treated group received 10.42ng/mL hIFN-α1 in puromycin media. Total RNA extraction was performed on day 5, during which the concentrations and purity of RNA were determined. The same amount of RNA was added to each RT-PCR reaction, in which cDNA was generated. Expression levels of the genes of interest were then determined using qPCR. Figure created with BioRender.

*The Ensembl ID of the lncRNA gene is listed, instead of transcript ID, because at the time of cloning of this gene into the pcDNA3.1-EGFP vector, information on the *MMP25-AS1-209* isoform in the database shows a 410-basepair lncRNA. This information, however, is changed following a recent update.

First, I established a protocol for transient overexpression of the lncRNAs of interest. I cloned the lncRNA genes into a pcDNA3.1/Puro-CAG-VSFP-CR vector (Appendix Figure 1). The puromycin resistance gene on the vector enabled us to select for the transfected cells by adding puromycin, whereas the CAG promoter allowed for robust overexpression of the lncRNA of interest. After transfection, puromycin selection, and treatment with hIFN-α1, I collected the lncRNA-overexpressing cells, extracted the RNA from them, reverse-transcribed the RNA into cDNA, and used it for subsequent qPCR experiments (Figure 5). Next, I designed primers targeting the lncRNAs and ISGs, respectively, and validated their specificity by performing a melting curve (Appendix Figure 2) followed by visualizing the resulting PCR amplicon by agarose gel electrophoresis (Appendix Figure 3).

2.3.2 Defining a protective ISG signature in lncRNA-overexpressing A375 cells

Before overexpressing our lncRNAs of interest in A375 cells, I first established the endogenous levels of lncRNA and ISG expression. I incubated A375s in media with 10.42 ng/mL of hIFN-α1 for 24 hours. These conditions had been pre-determined to be optimal through a series of dose-response studies performed by Ksenia Morozova[48] (Appendix Figure 4). The control group, on the other hand, was incubated in cell culture media without hIFN-α1. RT-qPCR showed an overall trend of increased lncRNA expression following hIFN-α1 stimulation. Statistical analysis revealed that the expression levels of lncRNAs *PSMB8-AS1*, *HCP5-204*, and *TFAP2A-AS1* were significantly elevated (Figure 6). These results suggested that the transcription of these lncRNAs might be downstream of type I IFN signaling and maybe even be IFN-dependent.

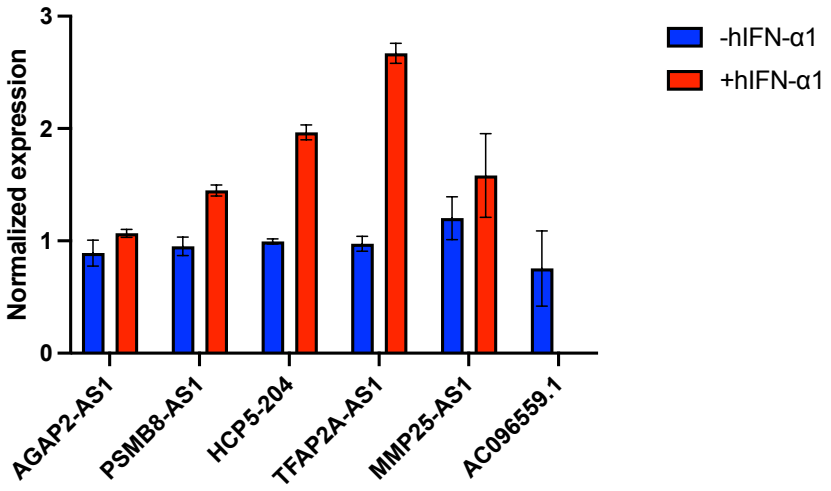


Figure 6: LncRNA expression was elevated following type I IFN treatment in A375s. RT-qPCR with at least n=2 technical replicates was performed to assess endogenous lncRNA expression levels. Results were normalized to actin-β mRNA. Error bars showed standard deviation from n=3 biological replicates. Welch's t-test with α=0.05 was performed to assess the significance of changes in endogenous lncRNA expression upon treatment with hIFN-α1. Expression levels of *PSMB8-AS1, HCP5-204,* and *TFAP2A-AS1* were significantly increased, with p-values of 0.0021, 0.00053, and 0.000024, respectively.

Simultaneously, we profiled the endogenous ISG expression in A375 cells with- and without hIFN-α1 treatment. From the list of 59 ISGs provided by Takeda Pharmaceuticals, a subset of 37 is known to be expressed in melanoma. Through RT-qPCR, we found that there is no consistent up- or downregulation among these ISGs – only two were significantly upregulated while one was significantly downregulated. These results were initially unexpected because hIFN-α1 stimulation should activate the transcription of these responsive genes. However, upon further consideration, it should be noted that many ISGs have redundant functions and not all are transcribed following type I IFN signaling.[42] In other words, only a subset of ISGs is required for initiation of the type I IFN response, and the induction of the set of ISGs is dependent on the specific type of PAMP that triggered the type I IFN response.[55] Moreover, the stimulation from the same class of PAMP might not even activate the same set of ISGs in identical populations of cells.[56] Lastly, we only used one

subtype of type I IFN to stimulate the cells, which further explains the changes in expression of a few, but not all, ISGs.

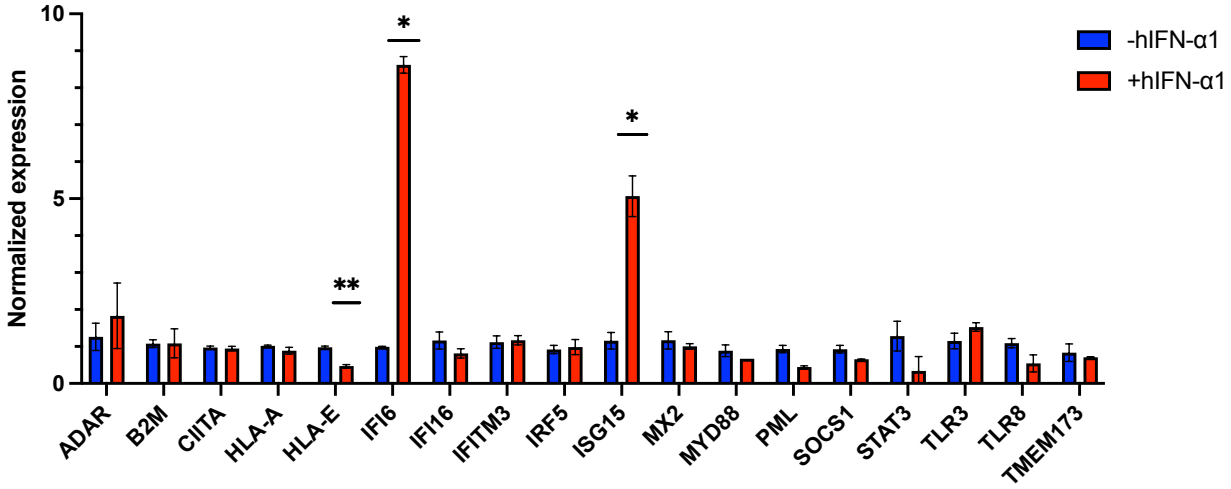


Figure 7: The expression levels of most ISGs remained unchanged following type I IFN treatment in A375s. The endogenous expression levels of interferon-stimulated genes (ISGs) were assessed by RT-qPCR, with at least n=2 technical replicates. Results were normalized to actin-β mRNA. Error bars represent standard deviation from n=3 biological replicates. Asterisks represent significance as determined by Welch's t-test, with α=0.05. Expression levels of IFI6 and ISG15 were significantly elevated, with p-values of 0.013 and 0.036, respectively. Expression level of HLA-E is significantly downregulated, with a p-value of 0.0052.

Statistical analysis revealed that expression level of IFN-α-inducible protein 6 (IFI6), a pro-apoptotic antiviral protein[57], and ISG15, which mainly inhibits viral replication[58], were significantly upregulated (Figure 7). On the other hand, mRNA levels of Major Histocompatibility Complex, Class I, E (HLA-E), responsible for antigen presentation, was significantly decreased (Figure 7). The expression level of another member of the IFI family, IFI16, however, was unaffected following hIFN-α1 stimulation, reinforcing the notion of ISG functional redundancy. Interestingly, the suppressor of cytokine signaling 1 (SOCS1), which is known to inhibit the JAK-

STAT signal transduction pathway, was also downregulated (Figure 7). In murine models, SOCS1 silencing has been shown to slow melanoma cell proliferation and reduce tumor size.[59,60]

After establishing the endogenous lncRNA and ISG levels in A375s, we began the gain-of-function studies in these cells. Performing the overexpression protocol (Figure 5) in A375s, we were able to highly express all the candidate lncRNAs in the melanoma cell line (Appendix Figure 5). We confirmed an average *AGAP2-AS1* and *HCP5-204* expression fold change of 768 and 2462, respectively, after transfection compared to that of empty control (Appendix Figure 5). We simultaneously profiled the ISG expression in these lncRNA-overexpressing A375s and found mRNA levels of components of the JAK-STAT signal transduction pathway – STAT1 and IRF9 – were significantly upregulated, with p-values of 0.0053 and 0.0011, respectively (Figure 8). The activation of STATs and IRF9 is known to sustain signaling through the JAK-STAT pathway in the absence of type I IFN binding.[61] Among other upregulated ISG is TLR3, which is a PRR that detects double-stranded RNA (dsRNA) in the cytosol.[62] We were also interested in the effect of lncRNAs on the two endogenously highly expressed ISGs - IFI6 and ISG15. Surprisingly, IFI6 showed no change in expression after overexpression of either lncRNA, and treatment of hIFN-α1 failed to rescue this phenotype (data not shown). For ISG15, significant increase in mRNA levels was observed following *AGAP2-AS1* and *HCP5-204* overexpression (Figure 8). Lastly, the mRNA level of adenosine deaminase RNA specific (ADAR) was significantly reduced by ectopically expressing the two lncRNAs (Figure 8). Upon overexpression of *AGAP2-AS1* and *HCP5-204*, we showed a full spectrum of unchanged, downregulated, and upregulated genes. From these gain-of-function studies in A375s, we tentatively defined our ISG signature consisting of ISG15, STAT1, IRF9, JAK1, and TLR3 (Table 4).

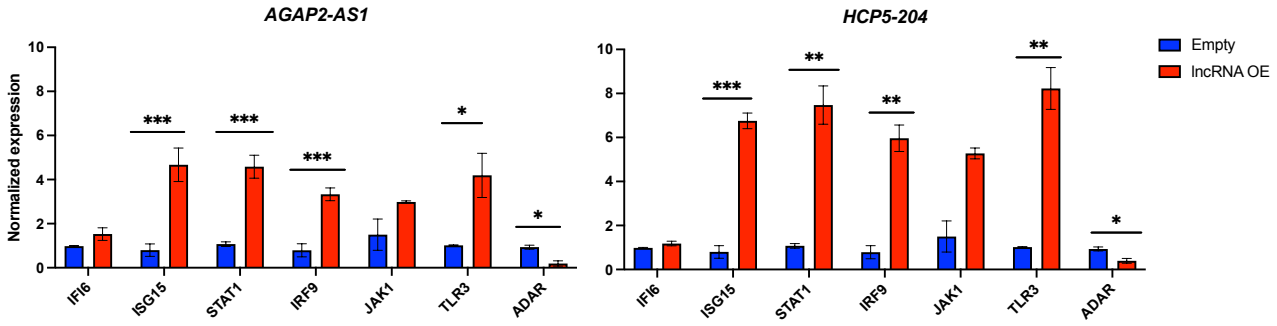


Figure 8: Components of the JAK-STAT signal transduction pathway are upregulated in A375s overexpressing *AGAP2-AS1* and *HCP5-204*. The expression levels of interferon-stimulated genes (ISGs) were assessed by RT-qPCR, with n=2 technical replicates. Results were normalized to actin-β mRNA. Error bars represent standard deviation from at least n=2 biological replicates. Asterisks represent significance as determined by Welch's t-test, with α=0.05. The ISGs ISG15, STAT1, IRF9, TLR3, and ADAR showed statistically significant differences from the empty control.

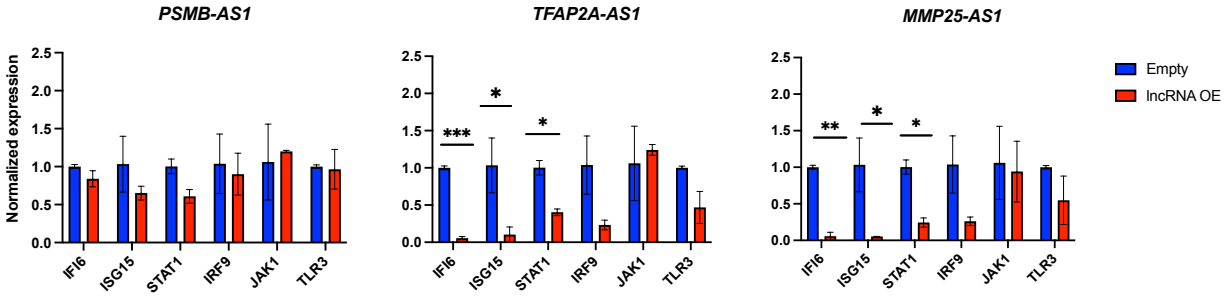


Figure 9: Overexpression of some candidate lncRNAs in A375 cells did not induce upregulation of interferon stimulated genes in the protective signature. The expression levels of interferon-stimulated genes (ISGs) were assessed by RT-qPCR, with n=2 technical replicates. Results were normalized to actin-β mRNA. Error bars represent standard deviation from at least n=2 biological replicates. *TFAP2A-AS1* and *MMP25-AS1* significantly induced decreased expression of IFI6, ISG15, and STAT1. Asterisks represent significance as determined by Welch's t-test, with α=0.05. Error bars represent standard deviation from at least n=2 technical replicates.

Table 4: Summary of interferon-stimulated genes in our interferon signature found in *AGAP2-AS1*- and *HCP5-204*-overexpressing A375 cells.

Gene Symbol	Gene Name	Function	Ensembl ID
ISG15	Ubiquitin-like protein 15	Inhibits viral replication	ENSG00000187608
STAT1	Signal transducer and activator of transcription 1	Transduces signal in JAK-STAT pathway	ENSG00000115415
IRF9	Interferon Regulatory Factor 9	Transduces signal in JAK-STAT pathway	ENSG00000213928
JAK1	Janus kinase 1	Transduces signal in JAK-STAT pathway	ENSG00000162434
TLR3	Toll-like receptor 3	Recognizes pathogens, activate innate immune responses	ENSG00000164342

2.3.3 *HCP5-204*-overexpressing melanoma short-term culture induced upregulation of genes in the interferon signature

We wanted to establish a robust ISG signature across different melanoma models, such as the A375 immortalized malignant melanoma cell line, as well as more representative models like patient derived melanoma short-term culture cells (MSTCs). MSTCs are primary cells directly derived from patient tissues which have been passaged for very few times outside of the patient. Thus, MSTCs represent a more clinically relevant model of melanoma than the immortalized melanoma cell line A375. We transiently overexpressed the lncRNAs of interest in MSTCs. By overexpressing the selected lncRNAs in these MSTCs, we expected to see the same set of differentially expressed ISGs as in A375s, thus validating the clinical significance of the ISG signature identified in A375 cells.

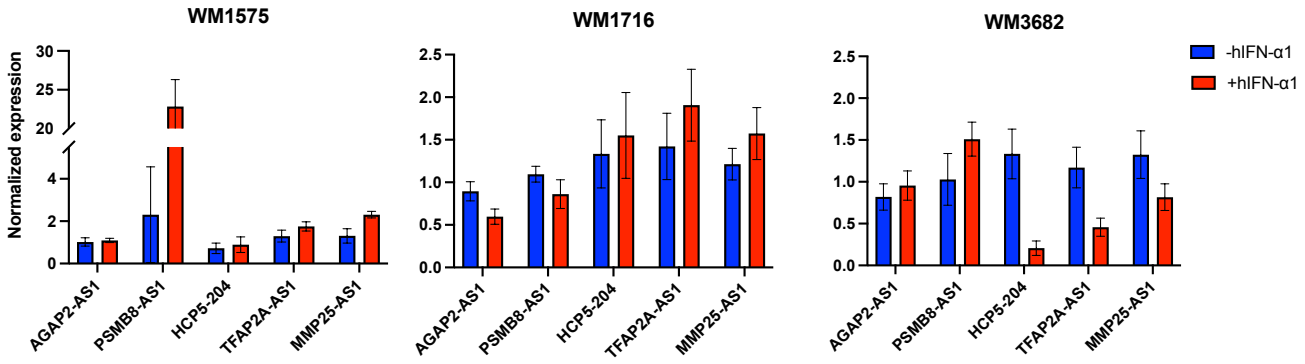


Figure 10: Melanoma short-term culture WM1575 showed elevated lncRNA endogenous levels upon type I interferon treatment. Endogenous lncRNA expression levels in the presence and absence of hIFN-α1 treatment were assessed by RT-qPCR, with n=3 technical replicates. All results were normalized to actin-β mRNA. Error bars show standard deviation from at least n=2 biological replicates.

Similar to the approach we used for the study of A375s, before overexpressing the lncRNAs of interest in MSTCs, we first characterized their responsiveness to hIFN-α1 treatment. Based on availability in the lab, we started with 3 different MSTCs - WM1575, WM1716, and WM3682. Among the MSTCs, there was not an observable trend of endogenous lncRNA expression. For instance, while *AGAP2-AS1* expression remained roughly the same following hIFN-α1 stimulation in WM1575, it was downregulated in WM1716 and upregulated in WM3682 (Figure 10). In parallel, we characterized the endogenous expression levels of selected ISGs – IFI6, ISG15, IRF9, SOCS1, and STAT1 – in MSTCs to assess whether those cells maintain IFN responsiveness. We found that MSTCs WM1716 and WM1575 showed significant upregulation of these ISGs, while WM3682 showed less pronounced IFN responsiveness (Figure 11). For instance, the average IFI6 expression fold changes in WM1575 and WM1715 were 591.32 and 35.36, respectively, while it was 3.87 in WM3682 (Figure 11). Following these characterizations, we chose to perform subsequent validation in WM1575 for their high IFN responsiveness.

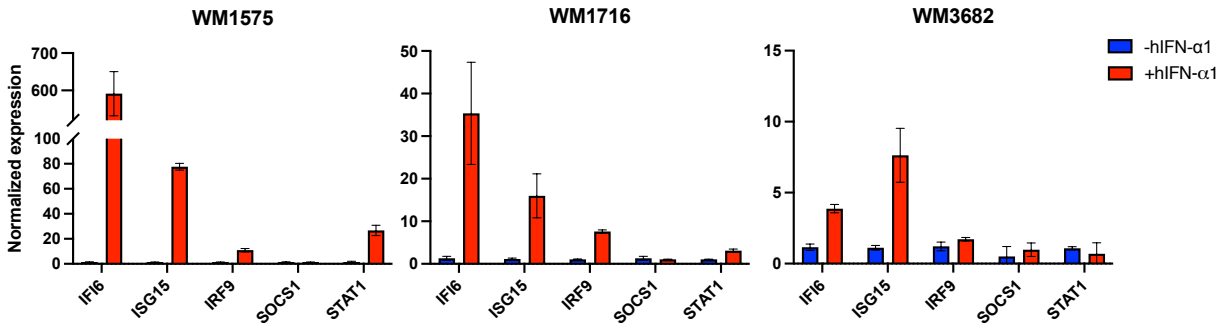


Figure 11: Selected ISGs showed upregulation in melanoma short-term cultures following human interferon-α1 treatment. Endogenous lncRNA expression levels in the presence and absence of hIFN-α1 treatment were assessed by RT-qPCR, with n=2 technical replicates. All results were normalized to actin-β mRNA. Error bars show standard deviation from at least n=2 biological replicates.

To validate the ISG signature found in lncRNA-overexpressing A375s, we ectopically overexpressed our lncRNAs of interest in MSTC WM1575. Using RT-qPCR, we confirmed overexpression of all 5 candidate lncRNAs (Appendix Figure 6). We then proceeded with ISG expression profiling in these lncRNA-overexpressing WM1575s. Results showed upregulation of signature ISGs in this MSTC overexpressing *HCP5-204*, with IRF9 and TLR3 most highly induced, having average expression fold changes of 13.70 and 14.19, respectively (Figure 12). This finding was consistent with our characterization in A375 cells. Therefore, we have identified *HCP5-204* to be a positive regulator of the type I IFN response, indicated by its induction of crucial ISGs in the IFN signature. In *PSMB8-AS1*-overexpresssing WM1575, we did not detect significant differential expression for most ISGs, similar to our observations in A375s, except for IRF9 that displayed a 1.73-fold change, with a p-value of 0.027 (Appendix Figure 7).

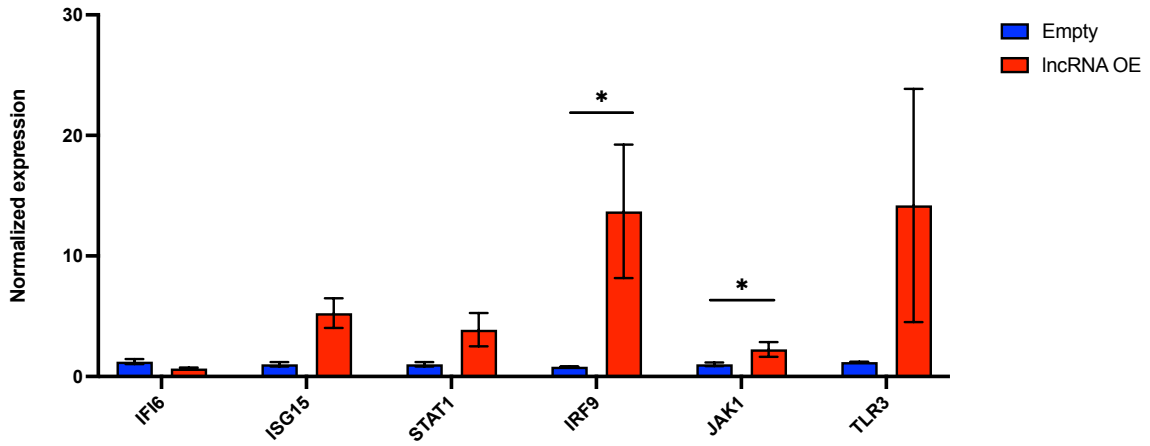


Figure 12: Overexpression of *HCP5-204* induced expression of signature interferon-stimulated genes. Interferon-stimulated gene (ISG) expression levels were assessed by RT-qPCR, with n=2 technical replicates. All results were normalized to actin-β mRNA. Error bars show standard deviation from at least n=2 biological replicates. Welch's t-test with α=0.05 was performed to assess the significance of changes in endogenous ISG expression upon treatment with hIFN-α1.

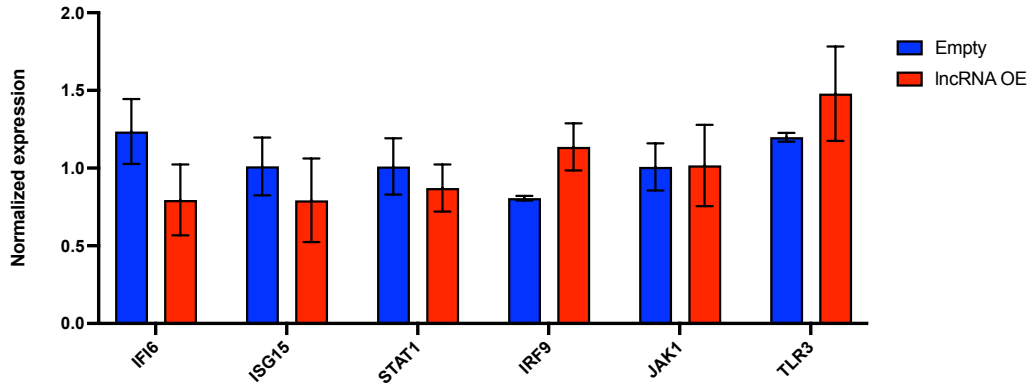


Figure 13*: AGAP2-AS1* overexpression in WM1575 did not upregulate any signature gene, inconsistent with its characterization in A375s. Interferon-stimulated gene (ISG) expression levels were assessed by RT-qPCR, with n=2 technical replicates. All results were normalized to actin-β mRNA. Error bars show standard deviation from at least n=2 biological replicates. Welch's t-test with α=0.05 was performed to assess the significance of changes in endogenous ISG expression upon treatment with hIFN-α1.

For other lncRNAs, however, overexpression in WM1575 showed different ISG expression patterns from that in A375 cells. None of the expression levels of the ISGs in the signature were significantly affected by ectopic expression of *AGAP2-AS1* in WM1575 (Figure 13), even though this lncRNA was shown to be an inducer of signature genes in A375 cells. Meanwhile, all signature genes were significantly upregulated following *TFAP2A-AS1* overexpression in WM1575 (Figure 14), opposite of its characterization in A375 cells (Figure 9). The difference in ISG15 expression between A375s and WM1575, which was 0.10 and 5.91 compared to respective empty control, was particularly remarkable (Figures 9 and 14). The significant upregulation of IFI6 upon *TFAP2A-AS1* overexpression, which we did not observe by ectopically expressing other lncRNAs in WM1575 or in A375s, was also noteworthy (Figure 14). To summarize, findings from *AGAP2-AS1* and *TFAP2A-AS1* overexpression in WM1575 deviated from our characterization in A375s, and further investigations are required to understand the molecular basis for these discrepancies.

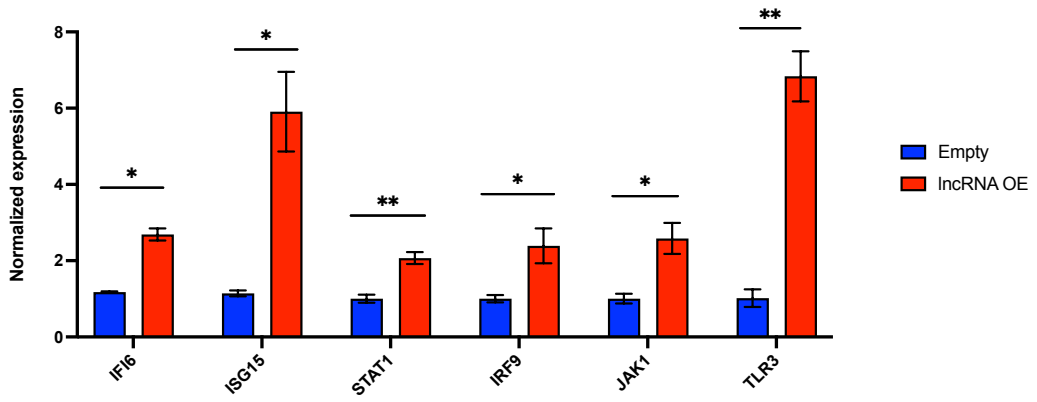


Figure 14: Overexpressing *TFAP2A-AS1* displayed a protective ISG signature, contrary to its characterization in A375s. Interferon-stimulated gene (ISG) expression levels in the were assessed by RT-qPCR, with n=2 technical replicates. All results were normalized to actin-β mRNA. Error bars show standard deviation from at least n=3 biological replicates. Welch's t-test with α=0.05 was performed to assess the significance of changes in endogenous ISG expression upon treatment with hIFN-α1.

2.4 Chapter Discussion

In this study, we validated the function of 5 candidate lncRNAs that were previously found to potentially modulate type I IFN response (unpublished data from the Novina lab). First, we characterized the IFN behavior of the model melanoma cell line A375 and determined a protective ISG signature consisting of ISG15, STAT1, IRF9, JAK1, and TLR3 upregulation. Notably, STAT1, IRF9, JAK1 are crucial ISGs that are involved in the JAK-STAT signal transduction pathway. Subsequently, using a gain-of-function study, we identified *HCP5-204* as an inducer of type I IFN response in melanoma cells, indicated by the significant upregulation of several important ISGs upon ectopic *HCP5-204* overexpression. Finally, we validated this signature in melanoma short-term culture cells, derived from patient samples, supporting the clinical relevance of our findings.

Having identified *HCP5-204* as a positive IFN response regulator, we aim to further validate *HCP5-204* as a potentially protective lncRNA in melanoma. Our next objective is to assess the ISG signature linked to *HCP5-204* overexpression in several other MSTCs, as well as to investigate ISG protein expression levels by immunoblot. Furthermore, using siRNAs, we can knock-down the HCP5 transcript in A375s and MSTCs to study the effect of *HCP5* depletion on the IFN gene signature. In a complementary experiment, we can use the CRISPR-Cas9 system to knockout the gene for HCP5 from those cells. Importantly, these two experiments could yield very different results because some lncRNAs can have distinct functions at the transcriptional and post-transcriptional levels. By applying this dual approach, we hope to understand the functional mechanism of *HCP5* on both levels. Finally, we will also perform RNA-sequencing (RNA-seq) in

melanoma cells overexpressing *HCP5-204* to interrogate the mechanism by which this lncRNA exerts its effect in IFN signaling in the context of cancer.

Overexpression of *AGAP2-AS1*, induced the upregulation of the same set of ISGs found in A375s overexpressing *HCP5-204*. However, we did not detect significant changes in gene expression in WM1575 overexpressing *AGAP2-AS1*. Ectopic expression of other candidate lncRNAs in WM1575 cells did not lead to differences in the expression levels of signature ISGs. Neither A375 cells nor the primary cells exhibited any difference in ISG expression levels upon *PSMB8-AS1* overexpression. *TFAP2A-AS1* and *MMP25-AS1* acted as negative regulators of type I IFN response in A375s. However, we observed the opposite effect in WM1575 overexpressing *TFAP2A-AS1* – all the ISGs in the signature were significantly upregulated.

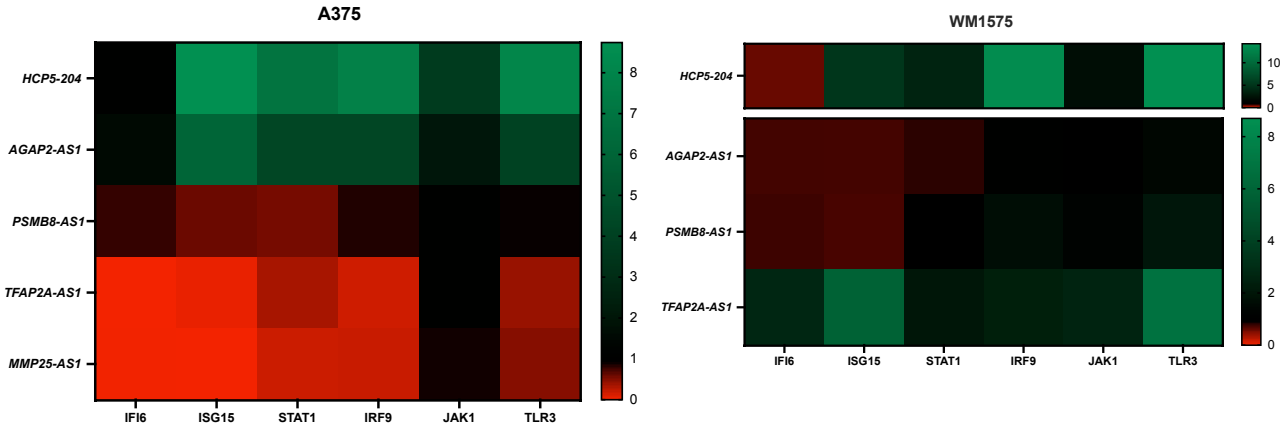


Figure 15: Summary of expression profiles of interferon signature genes in lncRNA-overexpressing A375 and WM1575 melanoma cells. Interferon-stimulated gene (ISG) expression was assessed by RT-qPCR. Each grid represents the average of at least n=2 biological replicates and at least n=2 technical replicates for RT-qPCR. Scale of expression fold change is indicated by color gradient next to each graph: upregulation is indicated in green, no change in expression is indicated in black, and downregulation in indicated in red.

Chapter 3: Discussion and Perspectives

3.1 Summary of results

In this study, we performed lncRNA gain-of-function studies to identify lncRNAs correlating with a robust positive ISG signature in melanoma. First, we profiled ISG expression in melanoma cell line A375 and defined a protective gene signature consisting of upregulated ISG15, STAT1, IRF9, JAK1, and TLR3. We identified the lncRNA *HCP5-204* to be associated with the protective ISG signature in A375 cells, as shown by the remarkable upregulation of the signature ISGs upon *HCP5-204* overexpression.

Previously published work has explored the role of *HCP5* in tumorigenesis.[63] To date, the best characterized mechanism by which *HCP5* is involved in tumor growth is by sponging miRNAs, which are otherwise regulating the expression levels of protein coding mRNAs. For instance, in the context of triple negative breast cancer, *HCP5* is found to promote expression of baculoviral IAP repeat-containing protein 3 (BIRC3), an inhibitor of apoptosis, by sponging miR-219a-5p.[64] *HCP5* also promotes bladder cancer metastasis by sponging miR-29b-3p, preventing its interactions with high mobility group box 1 (HMGB1).[65] Interestingly, HMGB1 can interact with TLR4 and subsequently activate nuclear factor kappa B (NF-κB) signaling, inducing cytokine production.[66] This finding suggests another potential role of *HCP5* in immune pathways. In our study, we have identified a new axis for tumor progression by *HCP5-204* via IFN-independent triggering of IFN-signature genes. Given that *HCP5* has been characterized as a competing endogenous RNA before,[67] it is likely that its involvement in IFN follows a similar mechanism and therefore it would be worth to explore this hypothesis in the future stages of this study.

Furthermore, we assessed the expression of SOCS1, an inhibitor of the JAK-STAT signal transduction pathway and regulator of TLR signaling,[68] in *HCP5-204* overexpressing melanoma cells. We found that exogenous overexpression of *HCP5-204* abolished the expression of SOCS1 in A375 and WM1575 cells (data not shown). Consequently, we observed upregulation of ISGs whose expression is commonly regulated by the JAK-STAT pathway.

Similar to that of *HCP5-204*, overexpression of the lncRNA *AGAP2-AS1* in A375 led to upregulation of ISG signature genes ISG15, STAT1, IRF9, JAK1, and TLR3; however, we failed to observe any differential ISG expression in WM1575. In fact, none of the expression levels of the ISGs in the signature were affected by overexpression of *AGAP2-AS1* in WM1575 cells. Additionally, *TFAP2A-AS1*, which has a transcription inhibitory effect on our signature IFN genes in A375 cells, proved to be an ISG inducer in WM1575. One potential source of the discrepancies we observed for *AGAP2-AS1* and *TFAP2A-AS1* was the different transfection reagents used for studying in the two systems – we transfected the plasmids using Lipofectamine 2000 in A375s and Lipofectamine LTX in MSTCs. Some biological explanations for the discrepancies could be that WM1575 carry a mutation in some upstream effector proteins. For instance, a mutation in the RNA-binding site of a protein X that theoretically binds *AGAP2-AS1* or *TFAP2A-AS1* would prevent it from performing any lncRNA-dependent modulation. To investigate this potential explanation, the effect of *AGAP2-AS1* and *TFAP2A-AS1* in type I IFN gene expression should be assessed in different melanoma cells, such as other MSTCs and other melanoma cell lines like SK-MEL-3. This experiment will indicate whether there is a consensus pattern in IFN gene transcription regulation in response to ectopic lncRNA overexpression. Once this has been established, the next step would be to characterize the interaction partners of *AGAP2-AS1* and

TFAP2A-AS1 first in A375 cells, where we see a clear functional effect upon lncRNA overexpression, and then in WM1575.

Finally, treatment with hIFN-α1 in the lncRNA-overexpressing cells had a consistently repressive effect on our ISG signature. Since prolonged type I IFN signaling is also responsible for suppressing JAK-STAT signaling,[43,44] this could explain differences in the final ISG expression levels we observed because we harvested the cells at different time points owing to experimental design compromises. Hence, it would be interesting to perform a time course experiment to assess early and late effects of lncRNA overexpression on IFN gene regulation in melanoma cells.

3.2 Critical assessment of the study

Endogenously expressed lncRNAs are generally lowly abundant in cells.[23] Hence, overexpression on such a magnified scale, some as high as 13,000 times the endogenous levels (Figure A6), might not have relevant biological implications. We understand that our approach provides a highly artificial environment in which to study the effect of our lncRNA of interest. Nevertheless, from this artificial set-up we wish to gain valuable information about the function and specific targets of our lncRNAs of interest which we aim to further validate in appropriate, biologically relevant models. For instance, once we have identified interesting target lncRNAs based on the overexpression experiments, e.g. *HCP5-204*, we aim to generate CRISPR knock-out and knock-in stable cell lines to study its effects when it is silenced and endogenously expressed, respectively.

Our gain-of-function experiments rely on transfection, a process that introduces foreign nucleic acids into cells.[69,70] Since type I IFN signaling is triggered upon detection of foreign nucleic acids,

such as viral genomes, our experimental approach could be inducing the type I IFN response. In agreement with this scenario, we have observed upregulation of TLR3, a double stranded RNA PRR, upon overexpression of *HCP5-204* and *AGAP2-AS1* in A375s. To control for the effect of the process of transfection of IFN signaling, we have included an empty (no lncRNA) overexpression vector transfection control. This experimental design allows us to discriminate between any transfection-related changes in IFN gene expression levels and those directly related to lncRNA overexpression. Nevertheless, the presence of foreign a nucleic acid vector could still be triggering the differential expression of ISGs due to the PAMP-specific nature of type I IFN signaling.

Finally, our study was limited to studying the *trans*-regulatory effects of our lncRNAs of interest. As we used ectopic overexpression of the lncRNAs, we did not assess the potential *cis*-regulatory effects of these lncRNAs. As many of our lncRNAs of interest are antisense lncRNAs and thus could be potentially *cis*-acting, this approach further limits the scope of our study. To address this issue, we propose to generate CRISPR-Cas9 knock-out stable cells and study the expression patterns of local protein-coding genes when lncRNA transcription is prevented. By studying the lncRNA gene in its genomic context, we can observe a full range of effects, on the transcriptional, post-transcriptional, translational levels and beyond.

3.3 Future directions

We have identified the lncRNA *HCP5-204* as a potential stimulator of a protective type IFN response in melanoma. To further our study, we will perform experiments to solidify our finding. Specifically, we aim to: (1) perform siRNA-mediated knockdown of *HCP5* expression; (2)

generate stable cells using CRISPR-Cas9 knocking-out/down and knock-in lncRNA genes to study the effects of endogenous lncRNA expression; (3) identify differentially expressed genes in both *HCP5-204* over- and under-expressing melanoma cells using RNA-seq. This set of proposed experiments can be used for the characterization of other lncRNAs acting in the type I IFN response and is not limited to *HCP5-204*.

Because our ultimate goal is to find lncRNAs that can act as cancer immunotherapy targets, it is therefore important to investigate the mechanism through which the selected lncRNA modulates the type I IFN response. To elucidate the mechanism of action of our lncRNA target, we propose to perform RNA pull-down mass spectroscopy (RP-MS) and yeast three-hybrid (Y3H) screening[71,72] to identify its interaction partners. We will use mutagenesis analysis to elucidate the active motifs on the lncRNA that are essential for its function. In brief, by introducing mutations or deletions of functional motifs on the lncRNA and then analyze ribonucleoprotein complex formation, we will elucidate the mechanisms of action of the lncRNA in single nucleotide resolution. Our lab has used Y3H to decipher the mechanism of *SLNCR* in melanomagenesis and metastasis.[33,73] Additionally, it would be useful to know where the lncRNAs predominantly localize to. We can explore this by fractionating melanoma cells and then perform RNA *in-situ* hybridization (RNA ISH).

Having established *HCP5-204* or other lncRNAs as potential therapeutic targets, we will enable the manipulation of those interactions to enhance or inhibit the type I IFN response. Precisely, we wish to identify small molecules able to interfere with the native lncRNA-protein complexes. To do this, we will re-purpose the Y3H set up to assess lncRNA-protein interactions in the presence

of small molecule inhibitors. We will then move on to validate the potential of identified small molecules as novel immunotherapy agents first *in cellulo* using cell culture cells and then *in vivo* using melanoma mouse models. Indeed, there are several mouse models available at Harvard Medical School and specifically in the Department of Cancer Immunology and Virology at Dana-Farber Cancer Institute. Finally, our study will set the ground for potential clinical trials testing the therapeutic application of lncRNA-based immunotherapy in the future.

To conclude, we have identified a lncRNA, *HCP5-204*, that is linked to a strong, protective ISG signature in melanoma cells. We validated this finding in MSTC WM1575 overexpressing *HCP5-204*. Future experiments will investigate the functions and mechanisms of action of *HCP5-204* modulating the type I IFN response in melanoma. Ultimately, we are using this system to study lncRNAs in type I IFN response in melanoma and utilizing the knowledge to develop targeted immunotherapy.

Bibliography

1. Jiang, M.-C., Ni, J.-J., Cui, W.-Y., Wang, B.-Y. & Zhuo, W. Emerging roles of lncRNA in cancer and therapeutic opportunities. *Am. J. Cancer Res.* **9**, 1354 (2019).
2. ENCODE Project Consortium, T. *et al.* An integrated encyclopedia of DNA elements in the human genome ENCODE Encyclopedia of DNA Elements. *Nature* **488**, (2012).
3. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **488**, (2012).
4. Brown, C. J., Ballabio, A. & F., H. W. *A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome*. (1991).
5. Cerase, A., Pintacuda, G., Tattermusch, A. & Avner, P. Xist localization and function: New insights from multiple levels. *Genome Biol.* **16**, 1–12 (2015).
6. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for Xist in X chromosome inactivation. *Nature* **379**, 131–7 (1996).
7. Quinn, J. J. & Chang, H. Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **17**, 47–62 (2016).
8. Robinson, E. K., Covarrubias, S. & Carpenter, S. The how and why of lncRNA function: An innate immune perspective. *Biochim. Biophys. Acta. Gene Regul. Mech.* **1863**, 194419 (2020).
9. Kung, J. T. Y., Colognori, D. & Lee, J. T. Long Noncoding RNAs: Past, Present, and Future Historical Overview of the lncRNA Field: Where Did It Come From? What Is It? Where Is It Going? (2013) doi:10.1534/genetics.112.146704.
10. Ma, L., Bajic, V. B. & Zhang, Z. On the classification of long non-coding RNAs. *http://dx.doi.org/10.4161/rna.24604* **10**, 924–933 (2013).
11. Blake, W. J., Kærn, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression.

Nat. 2003 4226932 **422**, 633–637 (2003).

12. Chu, C., Spitale, R. C. & Chang, H. Y. Technologies to probe functions and mechanisms of long noncoding RNAs. *Nat. Publ. Gr.* **22**, (2014).

13. Statello, L., Guo, C.-J., Chen, L.-L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. doi:10.1038/s41580-020-00315-9.

14. Ma, S. *et al.* The interplay between m6A RNA methylation and noncoding RNA in cancer. *J. Hematol. Oncol. 2019 121* **12**, 1–15 (2019).

15. Brunner, A. L. *et al. Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers*. http://genomebiology.com/2012/13/8/R75 (2012).

16. Mercer, T. R. *et al.* Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nat. Biotechnol.* **30**, (2011).

17. Nüsslein-volhard, C. & Wieschaus, E. Mutations affecting segment number and polarity in drosophila. *Nature* **287**, 795–801 (1980).

18. Johnsson, P., Lipovich, L., Grandér, D. & Morris, K. V. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim. Biophys. Acta* **1840**, 1063–1071 (2014).

19. Houseley, J. & Tollervey, D. Leading Edge The Many Pathways of RNA Degradation. doi:10.1016/j.cell.2009.01.019.

20. Marquardt, S., Hazelbaker, D. Z. & Buratowski, S. Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. *www.landesbioscience.com Transcr. 145 Transcr.* **2**, 145–154.

21. Svidritskiy, E., Demo, G. & Korostelev, A. A. Mechanism of premature translation

termination on a sense codon. *J. Biol. Chem.* **293**, 12472 (2018).

22. Tani, H., Torimura, M. & Akimitsu, N. The RNA Degradation Pathway Regulates the Function of GAS5 a Non-Coding RNA in Mammalian Cells. *PLoS One* **8**, 55684 (2013).

23. Zhang, H., Liang, Y., Han, S., Peng, C. & Li, Y. Long Noncoding RNA and Protein Interactions: From Experimental Results to Computational Models Based on Network Methods. (2019) doi:10.3390/ijms20061284.

24. Wang, K. C. *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* (2011) doi:10.1038/nature09819.

25. Liu, B. *et al.* The regulatory role of antisense lncRNAs in cancer. *Cancer Cell Int. 2021 211* **21**, 1–15 (2021).

26. Wang, L., Hu, L. & Jin, A. X. GUARDIN is a p53-responsive long non-coding RNA that is essential for genomic stability. *Nat. Cell Biol.* (2018) doi:10.1038/s41556-018-0066-7.

27. Krause, H. M. New and Prospective Roles for lncRNAs in Organelle Formation and Function. *Trends Genet.* **34**, 736–745 (2018).

28. Leucci, E. *et al.* Melanoma addiction to the long non-coding RNA SAMMSON. (2016) doi:10.1038/nature17161.

29. Schmitt, A. M. & Chang, H. Y. Long Noncoding RNAs in Cancer Pathways. *Cancer Cell* **29**, 452–463 (2016).

30. Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. (2010) doi:10.1038/nature08975.

31. Bhan, A. & Mandal, S. S. LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. *Biochim. Biophys. Acta - Rev. Cancer* **1856**, 151–164 (2015).

32. Vendramin, R. *et al.* SAMMSON fosters cancer cell fitness by concertedly enhancing

mitochondrial and cytosolic translation. *Nat. Struct. Mol. Biol.* | **25**,.

33. Schmidt, K. *et al.* The lncRNA SLNCR Recruits the Androgen Receptor to EGR1-Bound Genes in Melanoma and Inhibits Expression of Tumor Suppressor p21. *Cell Rep.* **27**, 2493-2507.e4 (2019).

34. Swetter, S. M. *et al.* Gender Differences in Melanoma Awareness and Detection Practices Between Middle-aged and Older Men With Melanoma and Their Female Spouses. *Arch. Dermatol.* **145**, 488–490 (2009).

35. Fisher, D. E. & Geller, A. C. Disproportionate Burden of Melanoma Mortality in Young US Men: The Possible Role of Biology and Behavior. *JAMA Dermatology* **149**, 903–904 (2013).

36. Murira, A. & Lamarre, A. Type-I interferon responses: From friend to foe in the battle against chronic viral infection. *Front. Immunol.* **7**, 609 (2016).

37. Boukhaled, G. M., Harding, S. & Brooks, D. G. Opposing Roles of Type I Interferons in Cancer Immunity. doi:10.1146/annurev-pathol-031920-093932.

38. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat. Rev. Immunol.* **14**, 36 (2014).

39. Levy, D. E., Marié, I. & Prakash, A. Ringing the interferon alarm: differential regulation of gene expression at the interface between innate and adaptive immunity. *Curr. Opin. Immunol.* **15**, 52–58 (2003).

40. McNab, F., Mayer-Barber, K., Sher, A. & Wack, A. Type I interferons in infectious disease. *Nat. Publ. Gr.* (2015) doi:10.1038/nri3787.

41. Hu, X., Li, J., Fu, M., Zhao, X. & Wang, W. The JAK/STAT signaling pathway: from bench to clinic. doi:10.1038/s41392-021-00791-1.

42. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat. Rev. Immunol. 2014 141* **14**, 36–49 (2013).

43. Schneider, W. M., Chevillotte, D. & Rice, C. M. Interferon-Stimulated Genes: A Complex Web of Host Defenses. (2014) doi:10.1146/annurev-immunol-032713-120231.

44. ronique François-Newton, V. *et al.* USP18-Based Negative Feedback Control Is Induced by Type I and Type III Interferons and Specifically Inactivates Interferon a Response. (2011) doi:10.1371/journal.pone.0022200.

45. Malakhova, O. A. *et al.* UBP43 is a novel regulator of interferon signaling independent of its ISG15 isopeptidase activity. *EMBO J.* **25**, 2358 (2006).

46. Kambara, H. *et al.* Negative regulation of the interferon response by an interferon-induced long non-coding RNA. *Nucleic Acids Res.* **42**, 10668–10681 (2014).

47. Kambara, H. *et al.* Regulation of interferon-stimulated gene BST2 by a lncRNA transcribed from a shared bidirectional promoter. (2015) doi:10.3389/fimmu.2014.00676.

48. Ksenia Morozova. Characterising Long Non-Coding RNAs (lncRNAs) Associated with a Protective Type 1 Interferon (IFN) Gene Signature in Tumour and Immune cells. (Harvard Medical School, 2021).

49. Clark, M. B. *et al.* Quantitative gene profiling of long noncoding RNAs with targeted RNA sequencing. *Nat. Methods* **12**, 339–342 (2015).

50. Carter, M. & Shieh, J. Biochemical Assays and Intracellular Signaling. *Guid. to Res. Tech. Neurosci.* 311–343 (2015) doi:10.1016/B978-0-12-800511-8.00015-0.

51. Genomic Data Commons (GDC) Data Portal. https://portal.gdc.cancer.gov/.

52. Hoadley, K. A. *et al.* Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291-304.e6 (2018).

53. NanoString: The Leader in Spatial Biology. https://nanostring.com/.

54. Rao, X., Huang, X., Zhou, Z. & Lin, X. *An improvement of the 2ˆ(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis*.

55. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. (2014) doi:10.1038/nri3581.

56. Zhao, M., Zhang, J., Phatnani, H., Scheu, S. & Maniatis, T. Stochastic Expression of the Interferon-b Gene. *PLoS Biol* **10**, 1001249 (2012).

57. Sajid, M. *et al.* The Functional and Antiviral Activity of Interferon Alpha-Inducible IFI6 Against Hepatitis B Virus Replication and Gene Expression. *Front. Immunol.* **12**, 1028 (2021).

58. Perng, Y. C. & Lenschow, D. J. ISG15 in antiviral immunity and beyond. *Nat. Rev. Microbiol. 2018 167* **16**, 423–439 (2018).

59. Li, Z. *et al.* Expression of SOCS-1, suppressor of cytokine signalling-1, in human melanoma. *J. Invest. Dermatol.* **123**, 737–745 (2004).

60. Borin Scutti, J. A. *et al.* Role of SOCS-1 gene on melanoma cell growth and tumor development1role of socs-1 gene on melanoma cell growth and tumor development. *Transl. Oncol.* **4**, 101–109 (2011).

61. Cheon, H. *et al.* IFNβ-dependent increases in STAT1, STAT2, and IRF9 mediate resistance to viruses and DNA damage. *EMBO J.* **32**, 2751–2763 (2013).

62. Zheng, X., Li, S. & Yang, H. Roles of Toll-Like Receptor 3 in Human Tumors. *Front. Immunol.* **12**, 1368 (2021).

63. Zou, Y. & Chen, B. Long non-coding RNA HCP5 in cancer. *Clin. Chim. Acta* **512**, 33–39 (2021).

64. Wang, L. *et al.* LncRNA HCP5 promotes triple negative breast cancer progression as a ceRNA to regulate BIRC3 by sponging miR-219a-5p. *Cancer Med.* **8**, 4389–4403 (2019).

65. Zhao, C. *et al.* LncRNA HCP5 Promotes Cell Invasion and Migration by Sponging miR-29b-3p in Human Bladder Cancer. (2020) doi:10.2147/OTT.S249770.

66. Hulda Sigridur Hreggvidsdóttir *et al.* High Mobility Group Box Protein 1 (HMGB1)-Partner MoleculeComplexes Enhance Cytokine Production by SignalingThrough the Partner Molecule Receptor. *Mol. Med.* **18**, 224–230 (2012).

67. Wu, H., Liu, B., Chen, Z., Li, G. & Zhang, Z. MSC-induced lncRNA HCP5 drove fatty acid oxidation through miR-3619-5p/AMPK/PGC1α/ CEBPB axis to promote stemness and chemo-resistance of gastric cancer. doi:10.1038/s41419-020-2426-z.

68. Sharma, J. & Larkin, J. Therapeutic implication of SOCS1 modulation in the treatment of autoimmunity and cancer. *Front. Pharmacol.* **10**, 324 (2019).

69. Felgner, P. L. *et al. Lipofection: A highly efficient, lipid-mediated DNA-transfection procedure (liposomes/cationic lipid vesides/gene transfer)*. *Proc. Nati. Acad. Sci. USA* vol. 84 https://www.pnas.org (1987).

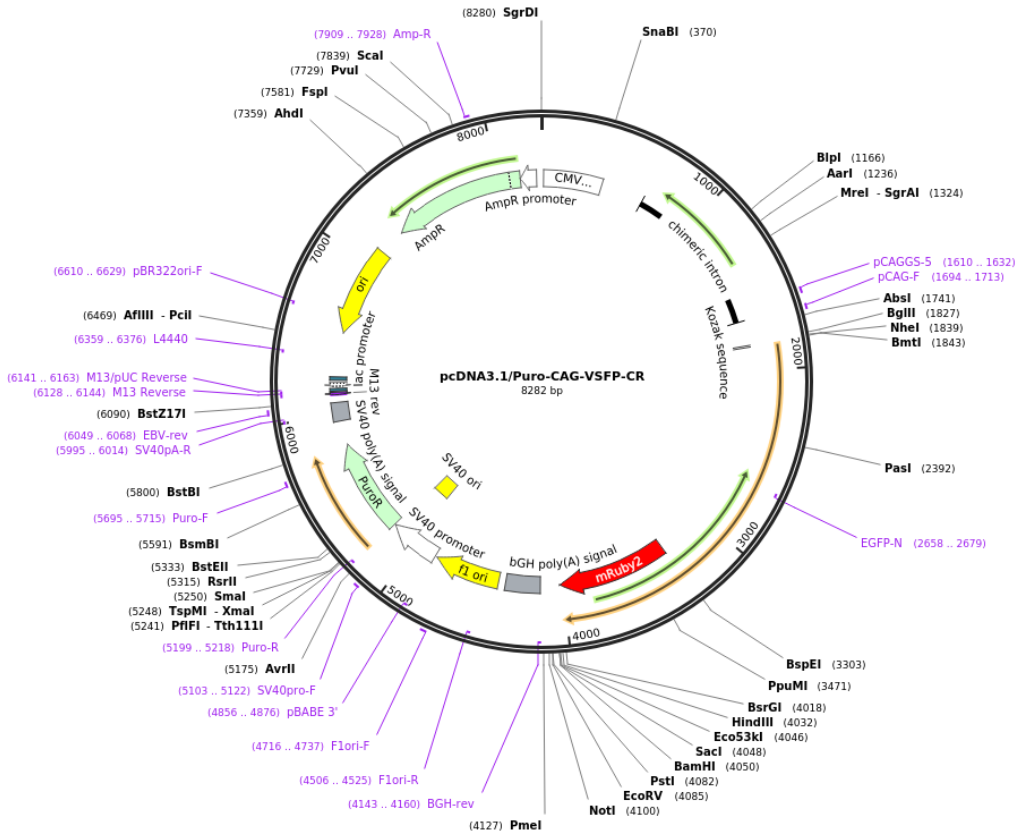
70. Farhood, H., Serbina, N. & Huang, L. The role of dioleoyl phosphatidylethanolamine in cationic liposome mediated gene transfer. *Biochim. Biophys. Acta* **1235**, 289–295 (1995).

71. Hook, B., Bernstein, D., Zhang, B. & Wickens, M. RNA-protein interactions in the yeast three-hybrid system: Affinity, sensitivity, and enhanced library screening. (2005) doi:10.1261/rna.7202705.

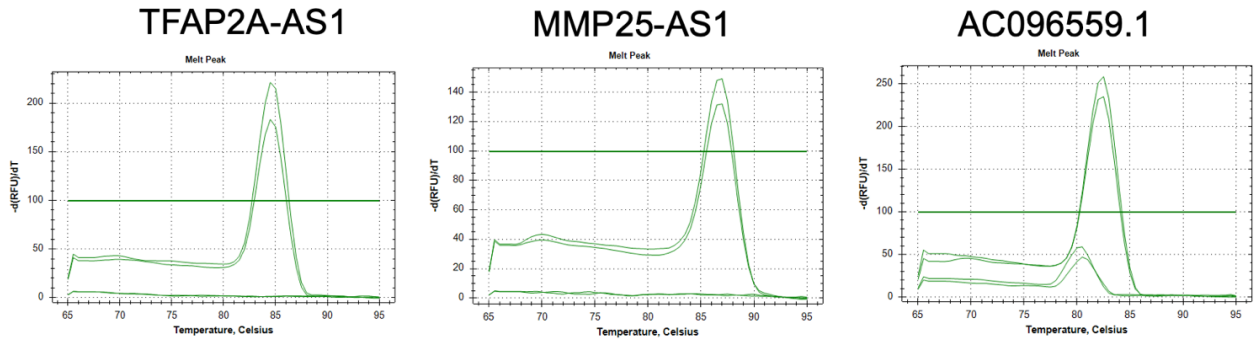
72. Cottier, S. *et al.* The yeast three-hybrid system as an experimental platform to identify proteins interacting with small signaling molecules in plant cells: potential and limitations. (2011) doi:10.3389/fpls.2011.00101.

73. Schmidt, K. *et al.* Targeting the Oncogenic Long Non-coding RNA SLNCR1 by Blocking Its Sequence-Specific Binding to the Androgen Receptor. *Cell Rep.* **30**, 541-554.e5 (2020).

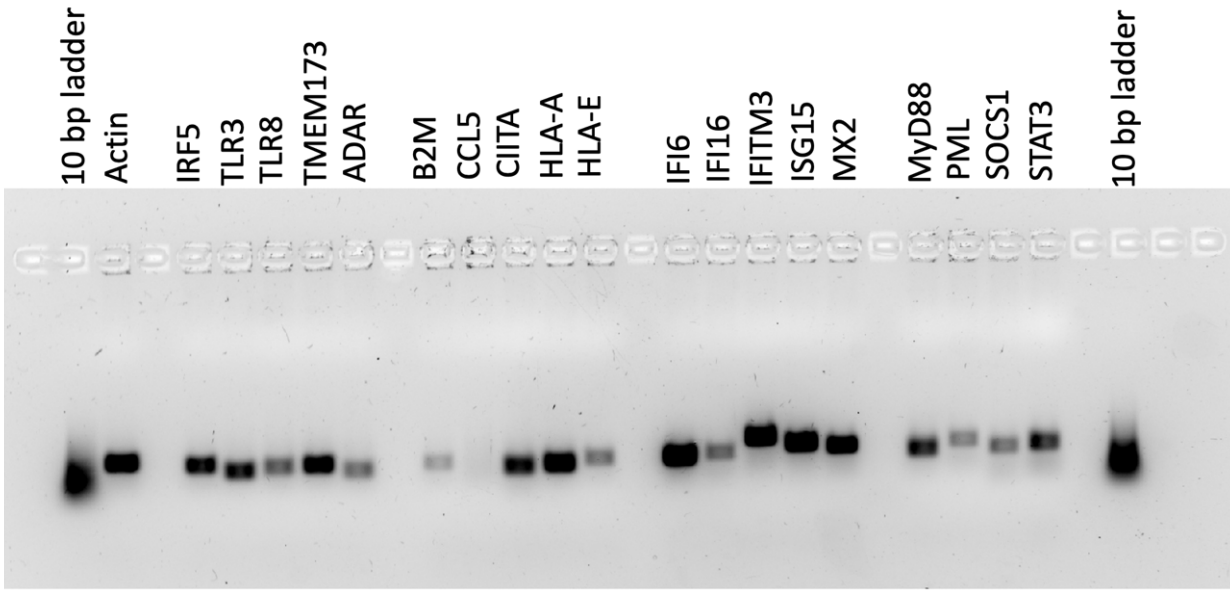
Appendix



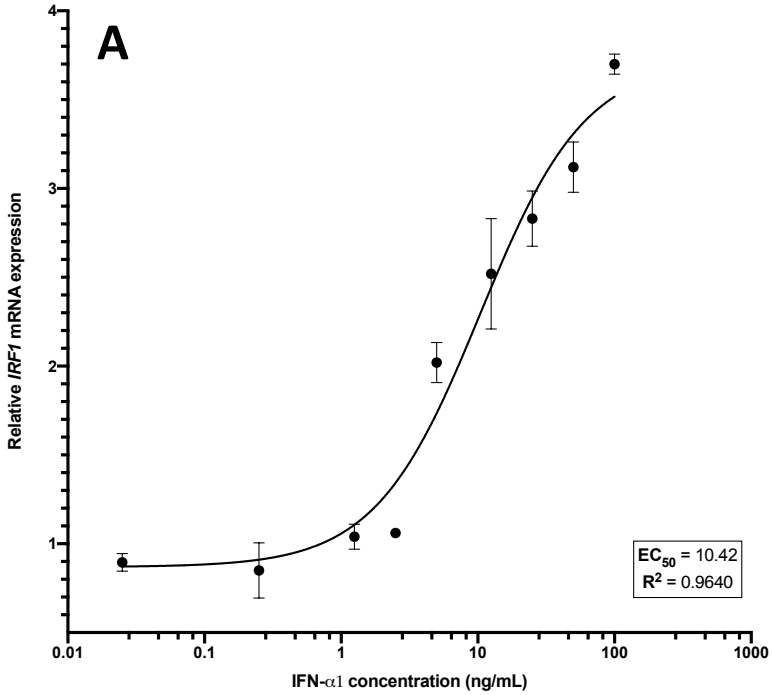
Appendix Figure 1: Plasmid map of pcDNA3.1/Puro-CAG-VSFP-CR vector. This was the backbone in which the 5 selected lncRNAs were inserted.



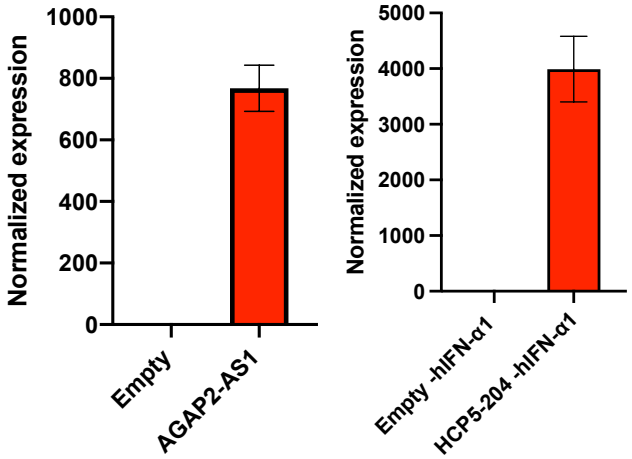
Appendix Figure 2: qPCR melt peak graphs of lncRNA primers. The singular peak in each graph confirms primer binding.



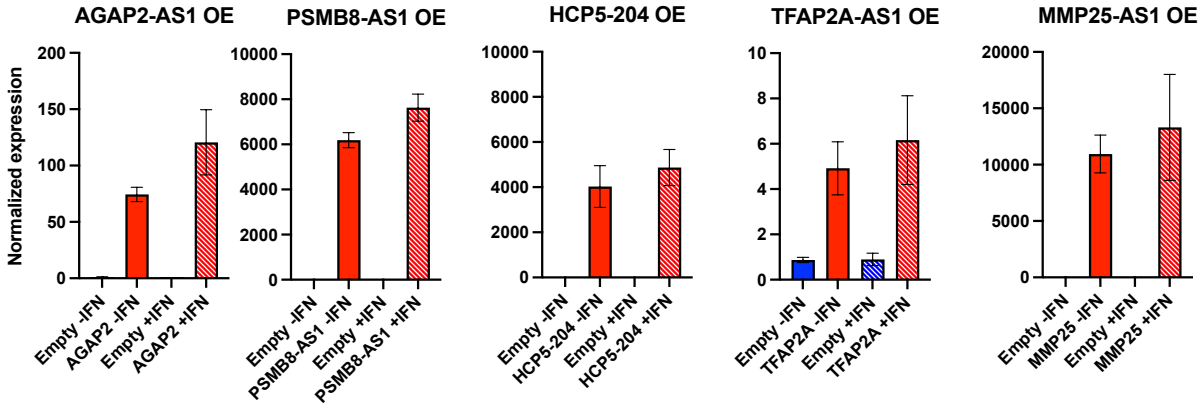
Appendix Figure 3: Using gel electrophoresis to validate ISG qPCR primer specificity. 10-basepair DNA ladder is loaded onto both sides of the gel. qPCR products amplified using ISG qPCR primers are loaded into each as indicated. The bands are all at the expected amplicon sizes.



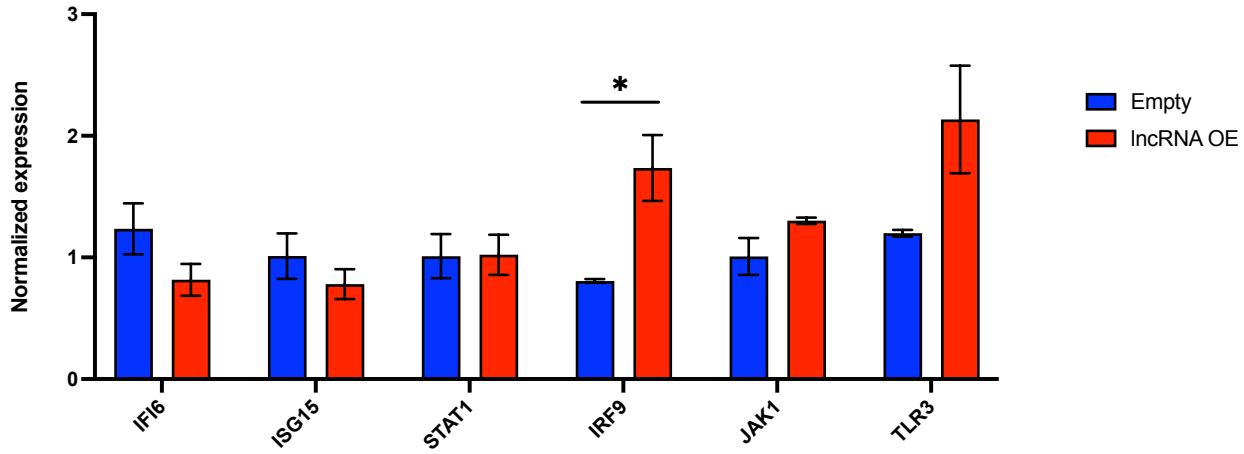
Appendix Figure 4: Dose response curve of hIFN-a1 concentration versus IRF-1 mRNA expression. The y-axis shows IRF1 luciferase activity in IFN-a1-treated HEK Lucia Null cells at the 48-hour timepoint. EC_{50} was extrapolated to be 10.42 ng/mL using the GraphPad Prism nonlinear regression curve fit function: [agonist] vs. response - variable slope (4 parameters). Data by Ksenia Morozova.



Appendix Figure 5: Overexpression of lncRNAs *AGAP2-AS1* and *HCP5-204* was achieved in A375s. The expression levels of interferon-stimulated genes (ISGs) were assessed by RT-qPCR. Error bars represent standard deviation from n=2 technical replicates.



Appendix Figure 6: LncRNA overexpression was achieved in MSTC WM1575, and hIFN-α1 stimulation further increased lncRNA overexpression levels. Transient lncRNA overexpression was achieved through plasmid transfection using Lipofectamine LTX, and 10.42ng/mL of hIFN-α1 was added to IFN-treated cells 24 hours prior to RNA extraction. LncRNA expressions in the presence and absence of hIFN-α1 treatment was assessed by RT-qPCR. Results were normalized to actin-β mRNA. Error bars show standard deviation from n=3 technical replicates. Welch's t-test with α= 0.05 was performed to assess the significance of changes in endogenous lncRNA upon treatment with hIFN-α1.



Appendix Figure 7: *PSMB8-AS1* overexpression in WM1575 did not upregulate any signature gene. Interferon-stimulated gene (ISG) expressions in the presence and absence of hIFN-α1 treatment was assessed by RT-qPCR. All results were normalized to actin-β mRNA. Error bars show standard deviation from n=2 technical replicates. Welch's t-test with α=0.05 was performed to assess the significance of changes in endogenous ISG expression.