UNIVERSITY OF CALIFORNIA
SANTA CRUZ

NON-CODING RNAs REGULATE INNATE IMMUNE SIGNALING

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MOLECULAR, CELLULAR, AND DEVELOPMENTAL BIOLOGY

by

Haley Lynne Halasz

September 2023

The Dissertation of Haley Lynne Halasz
is approved:

_____
Professor Susan Carpenter, Chair

_____
Professor Victoria Auerbuch Stone

_____
Professor Jacqueline Kimmey

_____
Peter F. Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

ABSTRACT

Non-Coding RNAs Regulate Innate Immune Signaling
Haley Lynne Halasz

As our understanding of the human genome has progressed, so has our interest in a class of molecules that challenge the central dogma of biology: DNA is transcribed into mRNA, mRNA gets translated into protein. These molecules have become known as "Non-coding RNAs" because they carry out cellular functions as RNAs without coding for proteins. One of the many compelling things about non-coding RNAs, is that their expression is highly context and tissue specific. The work presented here, focusses on a specific class of non-coding RNAs, long non-coding RNAs (lncRNAs), and how they function in the context of innate immunity and inflammation. Chapter 1 reviews the importance and clinical implications of lncRNAs in inflammatory diseases. Chapter 2 describes a high-throughput CRISPRi screening approach to identifying lncRNAs that regulate a prominent inflammatory signaling pathway, the NFkB pathway, in human monocytes. Chapter 2 also focusses on uncovering the mechanism of one such lncRNA, LOUP (lncRNA originating from upstream regulatory element of SPI1 [also known as PU.1]).

Our current understanding of lncRNAs that regulate inflammatory signaling in human monocytes is quite limited. Monocytes are precursors to macrophages, and both are critical effector cells of the innate immune system. Monocytes and macrophages are some of the first cells to respond to pathogens and are characterized by their ability to phagocytose. As presented in Chapter 2, we have used a human monocytic cell line (THP1 cells) as a model system to conduct a reporter based CRISPRi screen to identify lncRNAs that regulate NFkB signaling. NFkB is a transcription factor that activates transcription of hundreds of inflammatory genes. Dysregulation of this pathway underlies many diseased states. Our screen successfully identified numerous lncRNAs that regulate NFkB positively or negatively. One of the topmost significant candidates was a previously described lncRNA, LOUP that neighbors the myeloid lineage determining factor SPI1. In addition to driving myeloid differentiation, SPI1 is a transcription factor known to also control activation of inflammatory genes. Interestingly, we found that when we knockdown LOUP with CRISPRi, the TLR4/NFkB-driven inflammatory response is broadly upregulated, designating LOUP a negative regulator of NFkB. Previously, it's been found that the lncRNA LOUP transcript directly mediates interactions between an upstream response element (URE) and SPI1's promoter, hence regulating transcription of SPI1. Consistent with this previous work, we also found that expression of LOUP enhances SPI1 expression, but that this does not account for LOUP's inflammatory regulation. Remarkably, knowledge of the complexity of the human genome continues to develop, and it's

now appreciated that some designated “non-coding” RNAs produce very short functional peptides. Upon further investigation of LOUP’s coding potential, we discovered that it produces a small peptide responsible for LOUP’s ability to negatively regulate NFkB.

The studies described in Chapter 2, reveal new insights for lncRNAs and short ORF-encoded peptides (SEPs) in the context of inflammation. Relatively little is known about the mechanisms underlying how lncRNAs function in this context, let alone SEPs. While some lncRNAs have been identified as regulators of inflammation, the proportion of these genes that have been ascribed functions is still quite small, and the ability of some of these genes to produce short peptides has only been recognized somewhat recently. To our knowledge, this is the first successful CRIPSRi lncRNA screen performed in monocytes, making this not only a great technological advancement, but an invaluable resource. Using reliable high-throughput methods to screen for functional lncRNAs genome-wide is a highly efficient way to identify and further study the mechanisms of this under-examined class of genes.

ACKNOWLEDGEMENTS

First, I must acknowledge that this dissertation was not possible on my own. I've had colossal support, not only from academic mentors and collaborators, but from a community of family, friends, and healthcare providers that have made this great task at all feasible.

I first owe utmost thanks to my primary mentor Dr. Susan Carpenter. As those reading this will know, three years into this experience I wanted to accept a Master's degree and leave the PhD program. Susan encouraged and persuaded me to stay in the program and offered me a position in her lab. She had advocated for me throughout my experience leading up to my attempt to exit the program, demonstrating her genuine dedication in seeing students succeed. In my time in her lab, she has consistently convinced me that what seems unachievable is in fact achievable. Her expertise and highly skilled practice of the scientific method has been vital to my success. Clearly, I would not have completed this dissertation without her.

I also owe many thanks to all the members of the Carpenter Lab. Those that have been especially instrumental in helping me advance this research are Dr. Sergio Covarrubias, Eric Malekos, Dr. Lisa Sudek, Michelle Ramos, Steven Liang,

and Leila Namvar. Big thanks to Dr. Elektra Robinson for her friendship, guidance, and support throughout my time in the Carpenter Lab and beyond. Overall, all the members of the Carpenter Lab have offered a great deal of support during this inherently challenging experience.

I would also like to thank all the other UCSC faculty and staff that have supported me throughout my time as a graduate student. I want to thank Drs. Bill Saxton and Susan Strome who were very supportive during a time I was experiencing some difficulty in the program prior to beginning my work in Susan's lab. I also want to thank those that agreed to sit on my qualifying exam committee under unconventional circumstances – Drs. Bill Saxton, Rohinton Kamakaka, Vicki Stone, and Jeremy Sanford- who were all extremely instrumental in getting me on track towards a PhD after beginning a new thesis project in Susan's lab. And finally, thanks to my thesis committee members who have offered me guidance the last few years- Drs. Vicki Stone and Jacqueline Kimmey. I also received a lot of support from Dr. Bari Nazario, who not only offered moral support, but her great skills and expertise have made significant contributions to my research. I greatly appreciate all the active roles these members of the UCSC community have had in my success.

I next owe enormous thanks to those that mentored me prior to attending graduate school. Soon after I competed my B.S. in Biology at UCSC, Dr. Michael German welcomed me into his lab at UCSF. I had very little research experience at

the time and an extremely kind, patient and talented post-doc in his lab, Dr. Stephanie Nalle, agreed to mentor me. In the 1.5yrs that I spent there, I gained an immense amount of training and a much greater understanding of molecular and developmental biology. After my time at UCSF, Dr. Rohinton Kamakaka welcomed me as a Junior Specialist into his lab at UCSC. In his lab Dr. Namrita Dillon, a truly brilliant scientist, mentored me. Here I learned a great deal more about gene regulation and how to approach important mechanistic questions using molecular biology. The knowledge that Stephanie and Namrita imparted on me made graduate school a possibility, and my time in both labs really inspired my decision to pursue a PhD.

I want to deeply thank all of those that have supported me outside of the academic community. My thanks to both my husband and my parents cannot be overstated. They have truly supported me unconditionally throughout, and not only emotionally but financially as well. I hope that they can see how instrumental their love and support have been in this accomplishment. I can't begin to thank my in-laws enough for their undying support and interest in my success. Finally, I have so much gratitude to all the close friends that have cheered for me non-stop the last several years. To have the undying support of family and friends is something I cannot imagine having done this without.

CHAPTER 1- Challenges and Future Directions of LncRNAs and Inflammation

The text of Chapter 1 of this thesis includes a reprints of the following previously published material:

Halasz, H., Carpenter, S. (2022). Challenges and Future Directions for LncRNAs and Inflammation. In: Carpenter, S. (eds) Long Noncoding RNA. Advances in Experimental Medicine and Biology, vol 1363. Springer, Cham. https://doi.org/10.1007/978-3-030-92034-0_10

The co-author listed in this publication, Susan Carpenter, directed and supervised the research which forms the basis for the thesis.

1.1 Abstract

Until somewhat recently, the complexity of the human genome has not been well understood. With advancements in sequencing technology, we now know that nearly the whole genome is transcribed but a very small portion of those transcripts code for proteins. As the research of non-coding genes and transcripts has evolved rapidly in the last decade, it has become clear that many of them serve important biological functions in many previously well-studied cell processes. As the previous chapters in this book have reviewed, the field of noncoding RNA research has provided new insights into specific disease states, especially those driven by inflammation. Understanding the basic mechanisms of non-coding RNAs in the context of inflammation has led to prospective therapeutics that may overcome many of the challenges faced in diagnosing and treating inflammatory diseases. In this final chapter we discuss the current state of the field of non-coding RNA therapeutics and how it may evolve to overcome the short cummings we currently face with diagnosing and treating inflammatory diseases.

1.2 Non-coding RNAs and Inflammation

A well-adapted immune response is defined by acute activation of immune cells that results in transient release of inflammatory mediators. This rapid response aids in ridding the body of pathogens and repairing tissue damage. This process must be tightly regulated, and the acute response must be adequately resolved once the perceived threat is no longer present. Prolonged activation of inflammatory

signaling cascades can lead to chronic expression and secretion of inflammatory mediators. Even at low levels, chronically circulating inflammatory mediators can instigate mal-adaptions in a variety of tissues throughout the body (Barnig et al., 2019). It is now well understood and accepted that unresolved inflammation drives cardiovascular, lung, neurodegenerative, and metabolic diseases, including cancer and stroke (Barnig et al., 2019; Franceschi et al., 2018; Lin et al., 2018; Zuo et al., 2019). All of which the CDC has defined as leading causes of death (https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm).

The invasion of a pathogen leads to infiltration and activation of innate immune cells such as macrophages and lymphocytes. Pattern Recognition Receptors (PRRs) present in the membranes of these cells specifically recognize invading pathogens. Engagement of these receptors leads to transcriptional activation and translation of many pro-inflammatory chemokines and cytokines. In a well-functioning immune response these inflammatory signals get resolved once the pathogen has been eliminated. But other factors at the tissue-environmental interface can influence immune function and cause constitutive activation of inflammatory signals in the absence of a pathogen (Renz et al., 2011). Constantly elevated levels of cytokines can disrupt the function of many crucial cellular processes, such as mitochondrial function, protein folding, DNA repair, cell differentiation and cell regeneration (Barnig et al., 2019; Franceschi et al., 2018; Renz et al., 2011).

While many mechanisms by which chronic inflammation drives diseases have been defined, there is still a great need to better control such aberrant processes. While aging, genetics, and environmental pressures seem to be the main factors that determine immune function, developing a more complete understanding of the complex signaling networks that lead to chronic inflammation is central to preventing and treating inflammatory diseases.

As many of the chapters in this book have demonstrated, we can say with growing certainty that non-coding genes play an essential role in inflammatory disease pathology. While the non-coding RNA field is still somewhat in its infancy, advancements in screening technologies and transcriptomics have helped us identify hundreds of non-coding genes involved in various cell types and cell processes (Covarrubias et al., 2017; S. J. Liu et al., 2017; Napoli et al., 2020; Xue et al., 2017). Recent studies have shown that over 80% of the genome is transcribed but less than 3% of those transcripts are coding, revealing newfound complexity of the human genome (Salama, 2022). This means that most of the transcriptome is non-coding, and despite increasing evidence that many of these transcripts have specialized functions, especially in diseased states, only roughly 2% of non-coding transcripts have been characterized. Long non-coding RNAs (lncRNAs) account for the largest class of non-coding RNA transcripts and are defined as being greater than 200 nucleotides in length. These genes have now been found to participate in

diverse regulatory functions ranging from signaling molecules to transcriptional modifiers.

The previous chapters in this book have described in detail the mechanisms of many well-studied lncRNAs. These lncRNAs have been found to play crucial roles in mediating inflammatory processes underlying autoimmune diseases, cardiovascular diseases, and cancer (Brodnicki, 2022; Hennessy, 2022; Reggiardo et al., 2022; Wijesinghe et al., 2022). As we continue to better understand the basic mechanisms of action by which lncRNAs fine-tune many complex inflammatory signaling pathways and develop better biochemical assays for studying lncRNA structure and function (Sanbonmatsu, 2022), we anticipate that lncRNAs will provide novel and desirable therapeutic targets (Pierce et al., 2022; Reggiardo et al., 2022). Collectively the field of non-coding RNA research may offer promising new resolutions to many of the challenges we face in diagnosing and treating chronic inflammatory conditions.

1.1 Current State of Inflammatory Disease Diagnostics and Treatment

One particular challenge in diagnosing inflammatory diseases has been the detection of biomarkers that can truly distinguish a diseased state in an individual (Huang et al., 2017). Standard diagnostic blood tests can somewhat reliably assess systemic inflammation through circulating levels of interleukins IL-6, IL-8, IL-10, tumor necrosis factor-α (TNFα), C-reactive protein, and fibrinogen, etc.… (Pearson

et al., 2003) and while these markers are useful for flagging general inflammation or risk for a particular disease, they lack the specificity needed to make a clear diagnosis. Further diagnostic methods that may accompany blood tests involve rather invasive practices such as imaging, endoscopy, and tissue biopsies, but are currently necessary to confirm specific disease states. Many inflammatory pathologies require multiple diagnostic approaches and the time to diagnosis, let alone treatment, can be quite long. This can pose a challenge for patients already experiencing debilitating or severe symptoms and may make certain treatment options less effective once a diagnosis is determined.

Drug development and treatment of diseases driven by inflammation has been somewhat underwhelming. Current pharmacological treatment strategies for inflammatory diseases still rely mainly on non-steroidal anti-inflammatory drugs (NSAIDs), disease-modifying anti-rheumatic drugs (DMARDs), glucocorticoids and biologic agents (Pearson et al., 2003). While each of these classes of drugs work through slightly different mechanisms, they all work by inhibiting immune signaling globally and ubiquitously. Fortunately, these drugs have been effective at relieving symptoms in patients, but inhibition and suppression of these important signaling cascades in a non-specific manner can be very detrimental to immune homeostasis in the long term, and these drugs fail to address the underlying biology driving disease (Her and Kavanaugh, 2016; Sanbonmatsu, 2022). This again highlights the need to better understand the diverse biological roles of many of these

targets and the various aspects of physiology they control, to develop more effective treatments with much better specificity and precision. With the mounting evidence that lncRNAs can fine-tune and very tightly control immune related pathways in a very tissue specific manner, these molecules may act as much more promising therapeutic targets (Pierce et al., 2022).

With the advancement of the complete human genome sequence also came the ability to better study gene variants associated with disease. Genome-wide association studies (GWAS) have suggested the relationship between specific single nucleotide polymorphisms (SNPs) and specific diseases. These studies have served as a valuable guide to better understand the biology of many diseases, with the hope of identifying more specific drug targets (Dugger et al., 2018). Although, GWAS studies have also revealed the fact that most SNPs associated with disease are present in non-coding regions, therefore the biology of many of these SNPs are still unknown (Castellanos-Rubio and Ghosh, 2022). Hence, better understanding the actions of non-coding variants could reveal drug targets better tailored to specific diseases.

1.4 Clinical Potential for Non-coding RNAs

Developments in genomics and transcriptomics, have led to a much clearer understanding of the individual variation or stratification that exists within a certain

diseased population. A "one-size-fits-all" approach to treating disease is simply not optimal. In 2015 the U.S. government launched the Precision Medicine Initiative to fund more research and implement a clinical revolution that emphasizes individual, personalized healthcare. This includes a push to start employing more advanced sequencing and biomarker detection technologies that yield in-depth information about healthy and diseased individuals (Dugger et al., 2018). A more personalized approach to medicine means considering a person's genetics, microbiome, metabolism, lifestyle, and health history to determine the biological cause of the disease and tailor treatments to address that cause. The focus on precision medicine has dawned a new approach to drug development and diagnostics, and a push to detect subclinical disease before the onset of overt or debilitating symptoms.

RNA-based diagnostics and therapeutics have begun to show great promise in precision medicine. As mentioned earlier, such a small portion of the genome is translated into proteins and a majority of proteins have been deemed undruggable (Damase et al., 2021). An eruption of evidence for functional non-coding transcripts has really expanded the potential for tailored healthcare. The highly tissue and context specific expression of non-coding transcripts and their deregulation in diseased states makes them very attractive drug targets. Growing evidence suggests that many non-coding RNAs are secreted from cells and are protected in extracellular vesicles (Everaert et al., 2019). Therefore the tissue-

specific expression profiles of ncRNAs are often reflected in bodily fluids, making them very attractive as diagnostic markers as well (Bolha et al., 2017; Reggiardo et al., 2022; Yuan et al., 2020). Having this type of diagnostic power would not only eliminate the need for invasive tissue biopsies but could also help detect diseases in their infancy, leading to much better management and prevention.

To date, targeting ncRNAs therapeutically has relied on various RNA interference (RNAi) methods. These therapeutic strategies include the use of anti-sense oligos (ASOs), aptamers, small interfering RNAs (siRNAs), and micro-RNAs (miRNAs). These strategies work to modulate transcription or post-transcriptional RNA processing of their targets (Pierce et al., 2022). The biggest challenges with these methods have been related to RNA instability, difficult delivery across cell membranes, and the immunogenicity or cell toxicity of exogenous nucleic acids. Advancements in nanotechnology have facilitated stabilization, more efficient delivery, and better safety profile of therapeutic nucleotides, mainly overcoming such challenges (Damase et al., 2021).

While RNAi therapeutics have had success modulating their target genes, the only RNAi drugs that have been FDA approved so far, target transcripts of coding genes. For instance, Alnylam Pharmaceuticals has developed FDA approved siRNA therapies for treating rare genetic diseases that include hereditary amyloidosis (patisiran), acute hepatic porphyria (givosiran), and primary hyperoxaluria type 1 (lumasiran). Ionis Pharmaceuticals and Serepta Therapeutics have successfully

developed ASOs to treat spinal muscular atrphy (nusinersen), hereditary amyloidosis (inotersen), and Duchenne muscular distrophy (eteplirsen) respectively. Currently Haya Therapeutics is dedicated to targeting lncRNAs that drive fibrotic disease and have developed an ASO that successfully targets the lncRNA Wisper in the treatment of cardiac fibrosis, which is now undergoing clinical trial (Micheletti et al., 2017).

Perhaps the most promising aspect of the translational potential of lncRNAs is the fact that they can be detected in circulation, and their expression profile often corresponds with incipient disease. With advancements in RNA sequencing technology, the ability to detect circulating RNA in blood and other bodily fluids may offer a truly non-invasive way to not only diagnose presence of disease with high tissue specificity, but also prognose the potential for developing disease (Pierce et al., 2022; Reggiardo et al., 2022). Companies such as Freenome, Grail, and Guardant Health have now developed liquid biopsy tests ready for the clinic that can detect the earliest stages of cancer through circulating tumor DNA (Roy and Tiirikainen, 2020). This technology relies on DNA that gets released into the blood from cancer cell death (tumor shedding), but there is now evidence that live cancer cells secrete RNA through extracellular vesicles (Larson et al., 2021). In this sense, circulating RNA may offer a more reliable and even more specific biomarkers of cancer and other inflammatory conditions.

1.5 Future Insights for Non-coding RNA Therapeutics and Inflammation

While lncRNA research is advancing rapidly, our understanding of this class of molecules is still somewhat limited. Given that such a large portion of the genome produces non-coding transcripts, we must continue to rely on high throughput screening methods that can efficiently identify functional lncRNAs in specific cell types and under certain disease conditions. As we continue to identify functional lncRNAs in the context of inflammation, we will continue to gain a more complete understanding of the basic molecular mechanisms by which lncRNAs modulate inflammatory processes. Defining these mechanisms has begun to highlight the translational potential for lncRNAs yet getting lncRNA therapeutics into clinical trials has been somewhat challenging. This is primarily because non-coding RNAs tend to be poorly conserved, making the jump from preclinical studies in animals into human clinical trials challenging. But with the implement of high throughput methods across species, we have the power to quickly identify molecules that may lack sequence conservation but maintain functional conservation. Hopefully this will help focus our basic research efforts on lncRNAs with the utmost clinical potential. Additionally, rapidly evolving transcriptomic technologies have facilitated the possibility to detect circulating lncRNAs. The hope is that these lncRNAs will provide highly specific diagnostic markers for a range of conditions. As mentioned above, being able to detect subtle, chronic levels of inflammation or autoimmunity at their onset in an individual, has the power to

revolutionize how we manage and prevent inflammatory conditions. So while the lncRNA field has gained momentum, we have really only begun to shine a light on this so-called RNA “dark matter”, but what we have uncovered so far offers great promise and we look forward to seeing more progress made in both basic and translational lncRNA research.

CHAPTER 2- A Reporter-based CRISPRi Screen Reveals lncRNA LOUP that Regulates NFkB by Producing a Small Functional Peptide

2.1 Abstract

Long non-coding RNAs (lncRNAs) account for the largest portion of RNA from the transcriptome and yet most of their functions remain unknown. Here we performed a high throughput CRISPRi reporter-based screen to identify lncRNAs that regulate TLR4-NFkB signaling in human monocytes. We successfully identified numerous non-coding and protein-coding genes that can positively or negatively regulate this pathway. To understand the functional roles of lncRNAs in TLR4/NFkB signaling, we chose to further study one top candidate from our screen, LOUP (lncRNA originating from upstream regulatory element of SPI1 [also known as PU.1])(Trinh et al., 2021). It's been previously shown by Trinh et al. that the LOUP transcript directly mediates interactions between a nearby enhancer element of SPI1 and the SPI1 promoter. SPI1 is a transcription factor that drives myeloid cell fate, is highly expressed in monocytes, and regulates transcription of inflammatory genes. Here we've demonstrated that while LOUP expression can enhance SPI1 expression in monocytes, knockdown of LOUP also leads to a broad upregulation of NFkB targeted genes, both at baseline and upon TLR4-NFkB activation, and that this mechanism is independent of LOUP's regulatory effects on SPI1. We found that LOUP harbors three small open reading frames (sORFs) capable of being translated, and that one sORF in particular, is responsible for LOUP's ability to negatively regulate TLR4/NFkB signaling. This work

emphasizes the value of high-throughput screening to rapidly identify functional lncRNAs in the innate immune system.

2.2 Introduction

According to the latest Gencode release (version 43), the human genome encodes 19,928 long noncoding RNAs (lncRNAs) making it the largest group of genes produced from the genome. Due to their cell type specificity this number continues to increase as more sequencing is performed (Cabili et al., 2011) LncRNAs are transcripts over 200 nucleotides that are often spliced and polyadenylated with low protein coding potential. Over the last decade a number of lncRNAs have been functionally characterized and shown to play diverse roles in a variety of biological processes from cell differentiation and cancer, to immunity (Bhan et al., 2017; Guttman et al., 2011; Robinson et al., 2020). Yet the functions of the vast majority of these transcripts remains unknown. Historically one of the largest challenges in studying lncRNAs has been the lack of reliable and specific approaches to target these transcripts, especially in a high-throughput manner. Because lncRNAs lack open reading frames, these genes are not susceptible to frameshift mutations induced by classic CRISPR/Cas9. Recently the adoption of the modified CRISPR/Cas9 technology- CRISPR inhibition (CRISPRi), has become a powerful tool for interfering with transcription of lncRNAs by inducing repressive chromatin marks at the transcription start site, rather than altering the DNA with an indel or point mutation making it an attractive approach to discover

functional lncRNAs. Advanced computational developments in sgRNA library design coupled with targeted transcriptional repression induced by the components of CRISPRi, have made it possible to rapidly identify many functional lncRNA loci in a single pooled screening experiment (S. J. Liu et al., 2017).

A small number of high throughput screens have been performed to identify functional lncRNAs (Haswell et al., 2021; S. J. Liu et al., 2017; Liu et al., 2018; Zhu et al., 2016) but very few have been performed in immune cells (Arnan et al., 2022; Covarrubias et al., 2017). To explore new insights into innate immunity, more specifically monocyte and macrophage biology, we've conducted a screen in human monocytes to identify lncRNAs that regulate inflammation through NFkB signaling. Upon recognition of pathogens by TLR receptors, the cytoplasmic NFkB heterodimer (p50/p65) is allowed to translocate to the nucleus and activate transcription of hundreds of inflammatory response genes (Gilmore, 1999; Kaikkonen et al., 2023). Acute activation of NFkB signaling in monocytes is imperative to proper resolution of pathogenic invasions, therefore it's important that we gain a more complete molecular understanding of NFkB regulation in this context, as dysregulation of this pathway can lead to inflammatory diseases (Zinatizadeh, et al., 2021). To screen for regulators of NFkB signaling in monocytes, we first generated an NFkB (RelA/p65) reporter in the monocytic cell line THP1. We then developed an sgRNA library targeting over 2,000 THP1-expressed lncRNAs and conducted a pooled CRIPSRi screen in the NFkB-reporter

THP1s. With this, we identified numerous non-coding genes that positively or negatively regulate NFkB activation.

Some lncRNAs exhibit bimodal functions whereby they can function as an RNA in cis or in trans, act as enhancers or function through small encoded proteins reviewed in (Malekos and Carpenter, 2022). By definition lncRNAs lack coding potential, despite the fact that the majority of these genes contain small open reading frames (sORFs) less than 100 amino acids (Ruiz-Orera et al., 2020). With advances in sequencing and proteomics, it has come to light that many sORFs present throughout the genome are actively translated, and some cytoplasmic lncRNAs encode small functional peptides (Harrison, 2002; Ji et al., 2015; Wright et al., 2022; Zheng et al., 2023). Interestingly, we've identified a lncRNA that regulates inflammation in a bimodal fashion. LOUP was previously published to function in cis to regulate its neighboring protein SPI1. We confirmed this phenotype and identified LOUP as a top hit from our screen as a negative regulator of NFkB. Interestingly LOUPs function as a negative regulator of inflammation is dependent on its ability to produce a sORF-encoded peptide (SEP). Together this work highlights the power of CRISPRi screening to identify functional lncRNAs.

2.3 Results

2.3.1 CRISPRi screen identifies lncRNAs that regulate NFkB

To begin to define lncRNAs that regulate NFkB signaling in human macrophages, THP1 cells containing five NFkB (p65) binding sites upstream a minimal CMV-driven EGFP, as well as deactivated Cas9-KRAB (dCas9-KRAB), were transduced with pooled lentivirus (MOI = 0.3) generated from our custom sgRNA library containing ~25,000 individual sgRNAs. The sgRNA library was designed using the hCRISPRi-v2.1 algorithm (S. J. Liu et al., 2017), with 10 sgRNAs targeting the transcription start sites of 2,342 lncRNAs annotated in the human genome assembly GRCh37 (hg19). LncRNA targets were determined based on expression in RNA-seq data from THP1s generated by our lab and previously published p65 ChIP-seq data (Kaikkonen et al., 2023). The same design and cloning strategy were used as previously described for the CRiNCL library (S. J. Liu et al., 2017). Cells were stimulated with LPS for 24 hours and then sorted by FACS. The top and bottom 20% of GFP+ gated cells were collected, and genomic DNA was isolated from each population (GFPhi and GFPlo). This gating strategy was determined based on earlier reporter-based sorting screens to sufficiently capture non-targeting controls (Covarrubias et al., 2017; Kampmann et al., 2013; She et al., 2023). Since lentiviral sgRNAs integrate into the genome, genomic DNA was isolated from each GFPhi and GFPlo and unsorted population. In each resulting population sgRNAs were PCR amplified and then sequenced (Fig 2.1A-B). MAUDE (de Boer et al., 2020) analysis was performed comparing each GFPhi and GFPlo population to the unsorted (Fig 2.1C-D). Genes with combined Z-scores of less than -3 were defined as significant positive regulators of NFkB while genes

with combined Z-scores above 3 were defined as significant positive regulators. Given that our sgRNA library included lncRNA genes regardless of genomic location, many of the targets are very near to or overlapping with neighboring genes. Heterochromatin induced by the dCas9-KRAB has been shown to reach as far as 1kb so we considered this caveat when choosing top candidate lncRNAs to further functionally validate (Gilbert et al., 2014; S. J. Liu et al., 2017). 10 of the 35 top significant hits are intergenic; defined as having their own promoters at least 1kb away from promoters of neighboring genes (Table 1). 3 top hits, LRF1, UHRF1, and KMT2B, that were previously annotated as non-coding transcripts on hg19 have now been established as coding genes with unknown functions (Table 2.1). The other top hits are annotated either from a bidirectional promoter or antisense and overlapping to coding genes. In these cases, it is likely that the dCas9-KRAB has disrupted transcription of both bidirectional and overlapping transcripts. Interestingly, none of the coding genes targeted have been previously identified as regulators of NFkB and represent novel coding regulators of the pathway and therefore are still of interest (Table 2.1).

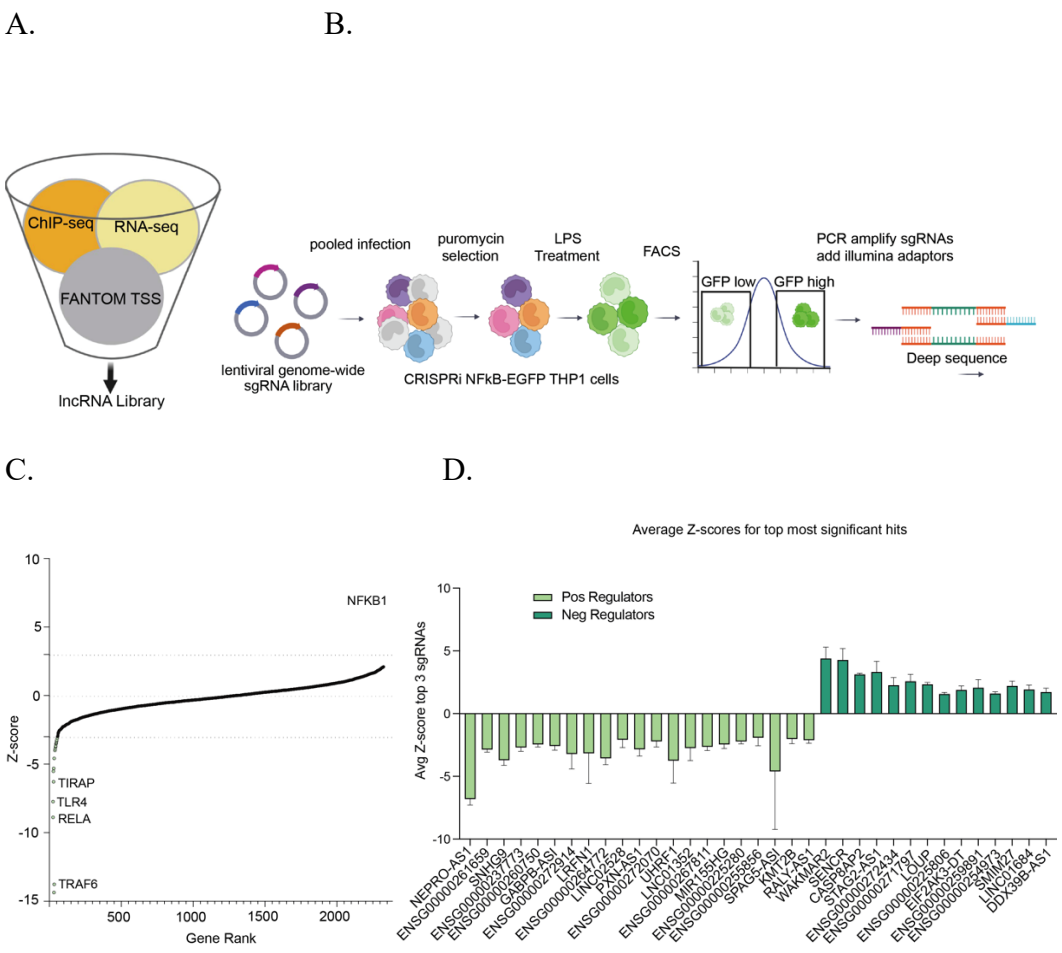


Figure 2. 1

Figure 2.1- CRISPRi screen identifies positive and negative regulators of NFkB.

A. Overview of sgRNA library design. sgRNAs were designed to target the transcription start sites of over 2000 Gencode hg19 annotated lncRNAs. Transcription start sites were predicted using consortiums of data from FANTOM and ENCODE. THP1 lncRNA expression was estimated from THP1 RNAseq data. **B. Overview of the screen.** NFkB-EGFP-CRISPRi-THP1 cells were infected with pooled the sgRNA libraries, selected, stimulated, and then sorted based on the top and bottom 20% of EGFP fluorescence. sgRNAs from the resulting populations were PCR amplified and sequenced. **C. MAUDE screen analysis.** MAUDE was performed on each of 3 screen replicates comparing sqRNA enrichment in the GFP low population or the GFP high population to the unsorted population. Z-scores across replicates were combined by Stouffer's method and ranked in order of Z-score. Z-score cutoffs of -3 and 3 are considered significant as highlighted in light green (positive regulators) and dark green (negative regulators). Genes labeled are

significant known protein coding regulators of NFkB that acted as positive controls in the screen. **D. Significant hits.** The average of the top 3 best scoring sgRNAs for all significant hits across three replicates with standard deviation.

Gene	lncRNA Function	Type	NFkB Regulation	GenomeCRISPR THP1	Nearest Neighbor	Promoter Distance	Neighbor Function	GenomeCRISPR THP1
NEPRO-AS1	Unknown	Antisense	Positive	No data	NEPRO	10bp	Notch signaling	No data
ENSG00000261659	Unknown	Bidirectional	Positive	Viability hit	WDR24	One promoter	Autophagy, TOR signaling	Viability hit
SNHG9	miRNA sponge, phase separation, Wnt signaling, cancer	Intergenic	Positive	No data	RPS2	80bp	Ribosomal protein	Viability hit
ENSG00000237773	Unknown	Antisense	Positive	No data	AHR	400bp	Cytochorme p450 regulation, aromatic hydrocarbon sensing	Not a hit
ENSG00000260750	Unknown	Intergenic	Positive	No data	JPH3	6kb	Junctional ER and plasma membrane comlex	Not a hit
GABPB1-AS1	miRNA sponge, cancer	Antisense	Positive	No data	GABPB1	50bp	Transcription factor, organelle biosynthesis	Not a hit
ENSG00000272914	Unknown	Intergenic	Positive	No data	ZNF438	10kb	Zinc finger protein	Not a hit
LRFN1	N/A	Coding	Positive	N/A	N/A	N/A	Unknown	Not a hit
ENSG00000264772	Unknown	Sense	Positive	No data	EIF4A1	One promoter	Translation initiation factor	Viability hit
LNC02528	Unknown	Intergenic	Positive	No data	TNFaIP3	60kb	Ubiquintation enzyme, NFkB inhibition	Not a hit
PXN-AS1	Regulation of PXN, cancer	Antisense	Positive	No data	PXN	45kb	Actin membrane attachment, cell adhesion	Not a hit
ENSG00000272070	Unknown	Intergenic	Positive	No data	PCDHGA1	4kb	Subunit of protocadherin gene cluster gamma	Not a hit
UHRF1	N/A	Coding	Positive	N/A	N/A	N/A		No data
LINC01352	Cancer	Intergenic	Positive	No data	MTARC1	40kb	Molybdopterin ion cofactor binding, cellular detox	Not a hit
ENSG00000267811	Unknown	Antisense	Positive	No data	TAF6L	one promoter	Transcription inititation by RNApoII	Not a hit
MIR155HG	Cancer	Intergenic	Positive	No data	MRPL39	30kb	Mitochondrial ribosomal protein	Not a hit
ENSG00000225280	Unknown	Antisense	Positive	No data	NKX2-4	25bp	Transcription regulation and cell differentiation	Not a hit
ENSG00000255856	Unknown	Antisense	Positive	No data	BCL7A	2kb	Gene translocation	Not a hit
SPAG5-AS1	SPAG5 regulation, mTOR signaling, Autophagy	Antisense	Positive	No data	SPAG5	100bp	Mitotic spindle apparatus protein	Not a hit
KMT2B	N/A	Coding	Positive	N/A	N/A	N/A	CXXC and PHD Zinc finger and SET domain protein	Not a hit
RALY-AS1	Unknown	Antisense	Positive	No data	RALY	400bp	hnRNP, pre-mRNA splicing	Not a hit
WAKMAR2	Enhancer of immune genes, Inflammation, wound healing, cancer	Antisense	Negative	No data	TNFaIP3	600bp	Ubiquintation enzyme, inhibits NFkB	Not a hit
SENCR	miRNA sponge, interaction with CKAP4, cell proliferation and migration	Antisense	Negative	No data	FLI1	2kb	Proto-oncogene transcription factor	Not a hit
CASP8AP2	N/A	N/A	Negative	N/A	N/A	N/A		No data
STAG2-ASI	Unknown	Antisense	Negative	No data	STAG2	50bp	De novo guanine nucleotide biosynthesis.	No data
ENSG00000272434	Unknown	Bidirectional	Negative	No data	IMPDH2, QRICH1	One promoter	De novo guanine nucleotide biosynthesis, unfolded protein response, apoptosis, and transcription	Not a hit, Not a hit
ENSG00000271797	Unknown	Antisense	Negative	Not a hit	PGGT1B	80bp	Geranyl geranyl transferase	Not a hit
LOUP	Regulation of SPI1	Intergenic	Negative	No data	SPI1	15kb	Transcription factor myeloid and B-cell development	Viability hit
ENSG00000225806	Paternally expressed, may regulate imprinting at locus	Bidirectional	Negative	No data	GNAS	One promoter	Complex imprinted gene expression	Not a hit
EIF2AK3-DT	Unknown	Antisense	Negative	No data	EIF2AK3	150bp	Tanslation initiation factor	Not a hit
ENSG00000259891	Unknown	Intergenic	Negative		CHRAC1	3kb	Histone fold protein	Not a hit
ENSG00000254973	Unknown	Intronic, possibly Intergenic	Negative	No data	IQANK	30kb	Unknown, predicted in actin filament capping	No data

Table 2. 1

Table 2.1- Description of top hits from NFkB screen.
Here each significant hit gene and its closest protein-coding neighbor are described. Data from the Genome CRISPR database was used to note whether the targeted lncRNAs or their neighbors have been previously indicated as hits in viability screens, if so they have been labelled "viability hit". Four of the hit genes once annotated as lncRNAs have been updated to protein coding genes and therefore their neighboring genes were not assessed (LRFN1, UHRF1, KMT2B, and CASP8AB2). Distances between targets and their neighbors were measured between promoters predicted by the ENCODE Registry of Candidate Cis-regulatory Elements on the UCSC genome browser.

2.3.2 LOUP regulates its neighbor SPI1

The top hit from our CRISPRi screen is the lncRNA LOUP (lncRNA originating from the upstream regulatory element of SPI1). Given the recent evidence that LOUP may act as an enhancer for its neighboring gene SPI1 (Trinh et al., 2021), and that this is true for other lncRNAs neighboring critical protein regulators such as the lncRNA PVT1 with its neighbor MYC, (Cho et al., 2018) and lncRNA p21 and its protein neighbor p21 (Groff et al., 2016), we measured levels of SPI1 protein in THP 1 upon CRISPRi knockdown of LOUP and over a time course of LPS stimulation. (Fig 2.2A-B). In THP1 cells SPI1is present at low levels at baseline and is induced following LPS stimulation. While the same non-targeting control cells were used for each experiment (sgRNA_NT) some inconsistencies in the timing of this induction are observed with SPI1 being induced highly at 2 or 6 hours following LPS stimulation (Fig 2.2B). Perhaps most interesting is the comparison of SPI1 levels between each of the control replicates and three different LOUP knockdowns at baseline where SPI1 is downregulated when LOUP is knocked down. In the LOUP knockdown lines SPI1's induction peaks after 24 hours of LPS to the levels observed in the control lines after 2 to 6 hours, indicating that LOUP acts as an enhancer for SPI1 but LOUP alone is not essential for SPI1 expression. ChIP-seq (chromatin immunoprecipitation with sequencing), ATAC-seq (assay for transposase-accessible chromatin with sequencing) and Hi-C (chromosome conformation capture) data from THP1s also provides evidence of LOUP's enhancer qualities. Regions of open chromatin

dictated by ATAC-seq and H3K27Ac ChIP-seq reads correspond with both LOUP and SPI1 promoters. LOUP and SPI1 occupy the same topologically associated domain (TAD) determined by HI-C interactions and CCCTC binding factor (CTCF) ChIP-seq (Fig 2.3A). Together this data confirms that LOUP can act in cis to enhance expression of its neighboring protein SPI1.

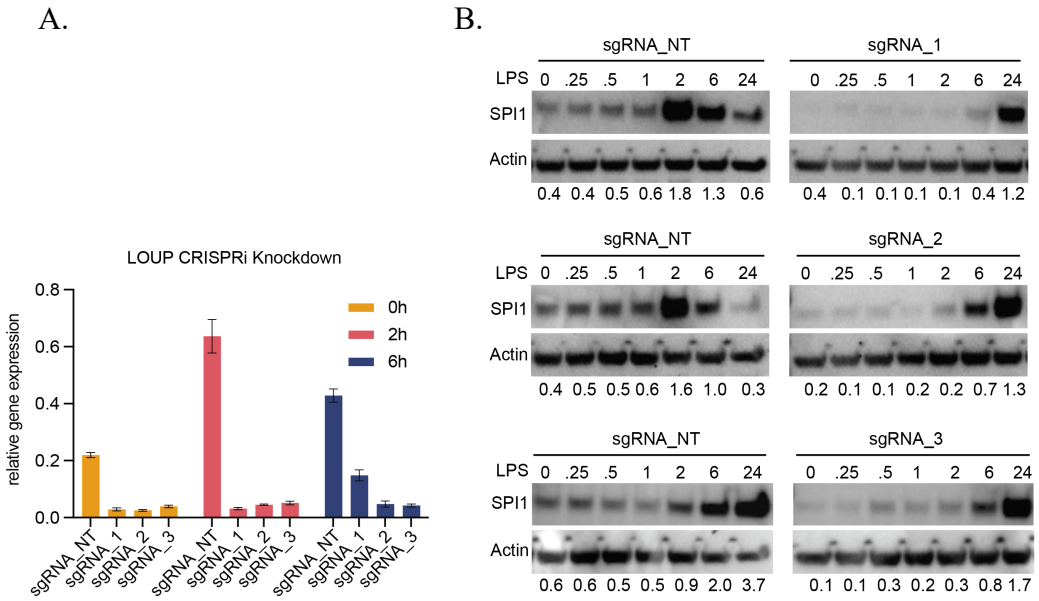


Figure 2. 2

Figure 2.2- LOUP regulates SPI1 in monocytes.
A. CRISPRi Knockdown of LOUP in THP1s. Three additional sgRNAs were designed to target top lncRNA candidate LOUP. qPCR measurement of LOUP across three replicate experiments shows knockdown of LOUP by all three sgRNAs (p-values<0.05) vs a non-targeting control sgRNA (NT) before LPS stimulation (0 hours) and after 2 and 6 hours of LPS stimulation. Values are normalized to HPRT and error bars represent standard deviation. **B. Western blot analysis of SPI1.** Non-targeting control (NT) vs. 3 LOUP knockdowns (sgRNAs 1, 2, 3) over a time course (hours) of LPS treatment. Samples were collected at baseline (0), 15min (.25), 30min (.5), 1 hour (1), 2 hours (2), 6 hours (6) and 24 hours (24). Quantitative values indicate densitometry ratios of SPI1:Actin.

A.

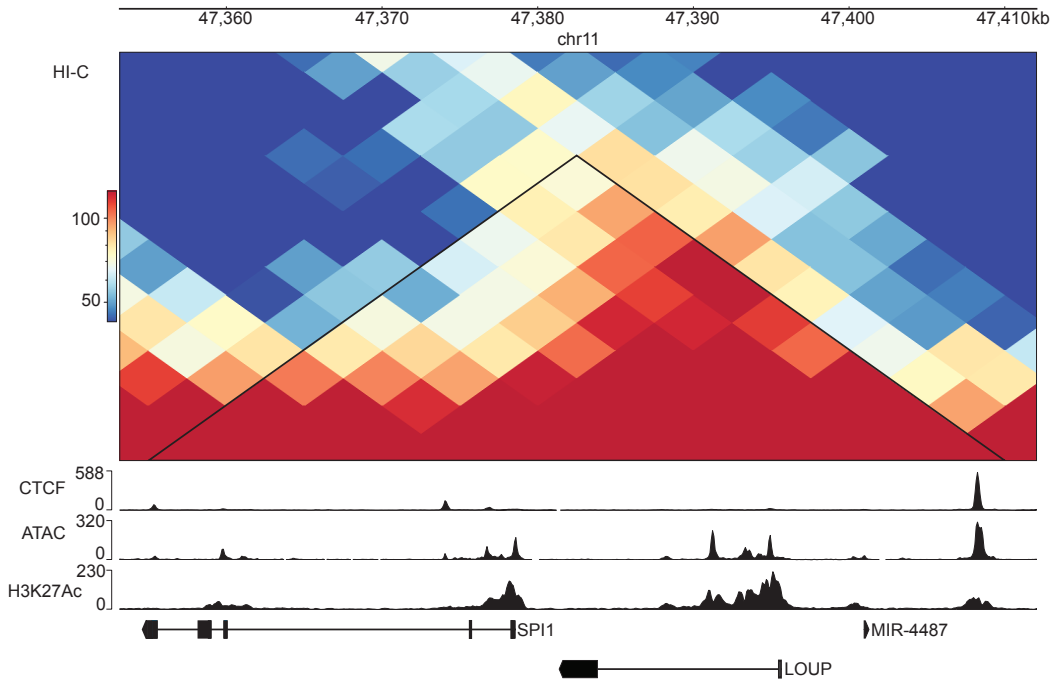


Figure 2. 3

Figure 2.3- LOUP acts as an enhancer of SPI1.
A. HI-C data from THP1s displaying predicted TAD for LOUP and SPI1. Browser tracks display regions of CTCF binding, ATAC peaks, and regions with of high H3K27 acetylation. Numbers on the Y-axis represent normalized read counts.

2.3.3 LOUP acts to negatively regulate NFkB target genes at the RNA and protein level

Based on the results of the screen, LOUP can negatively regulate NFkB activation (Fig 2.1C). To begin to validate LOUP as a negative regulator of NFkB, we looked broadly at inflammatory gene expression upon LOUP knockdown (Fig 2.4A). RNA was collected from two of the LOUP CRISPRi knockdown THP1 cell lines (sgRNA_1 and sgRNA_2) both before and after 6h of LPS stimulation. Using Nanostring technology to directly quantify RNA transcripts of over 500 immune genes, it was evident that knockdown of LOUP broadly upregulates transcription of inflammatory genes both before and after LPS stimulation, including transcripts that comprise the TLR4/NFkB signaling pathway (Fig 2.4B-C). To determine if knockdown of LOUP affects protein levels of inflammatory genes, we collected supernatant 24h post LPS treatment from all three LOUP knockdown THP1 cell lines and performed cytokine arrays testing 45 proteins. 16 of the 45 inflammatory cytokines were significantly increased compared to controls (Fig 2.4D).

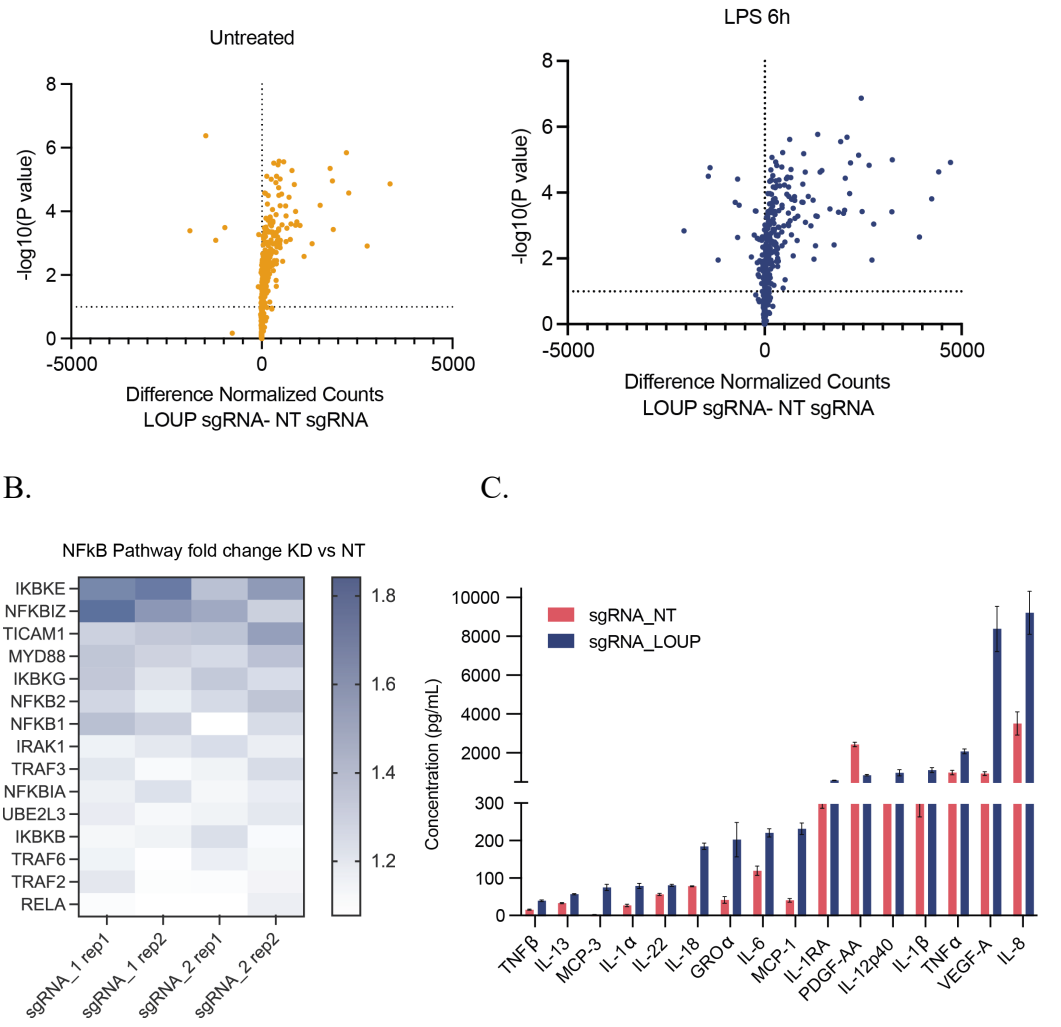


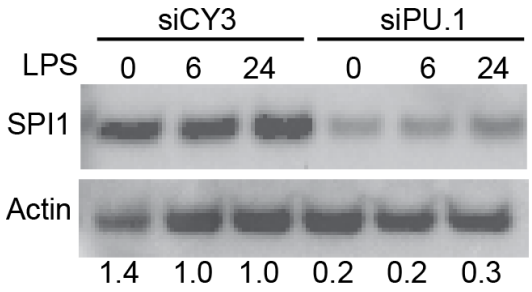
Figure 2. 4

Figure 2.4- Knockdown of LOUP upregulates NFkB targeted genes.
A-B. Multiplexed analysis of immune related transcripts upon LOUP knockdown. A. Transcripts of 580 immune genes were measured in RNA from a non-targeting control and 2 of the sgRNA LOUP Knockdowns at baseline and after 6 hours of LPS stimulation. Each sample was measured in duplicate. All transcript counts were normalized to 6 housekeeping genes then both knockdowns and duplicate measurements were averaged. All genes were plotted regardless of p-value. B. Heat map representing fold change of knockdowns vs non-targeting control at baseline for NFkB pathway genes significantly up-regulated in transcript analysis (A) (p-values<0.05). **C. Multiplexed ELISA analysis of cytokines upon LOUP knockdown.** ELISA was performed on the supernatant from three non-

targeting controls and all three LOUP knockdowns after 24 hour LPS treatment. All bars represent an average of all three non-targeting (NT) or an average of all three knockdowns with standard deviation. All differences observed between NT and knockdowns for each cytokine are significant (p-values<0.05).

To determine if a reduction in SPI1 could be responsible for the negative regulation of NFkB or inflammatory signaling, THP1s were transfected with siRNA targeting SPI1 or a non-targeting siRNA control (siCY3). IL8 was the most measurably upregulated secreted cytokine in the LOUP knockdowns (Fig 2.4C). IL8 was measured in the SPI1 knockdown vs control by ELISA. Not surprisingly, knockdown of SPI1 resulted in significantly decreased levels of IL8 (Fig 2.5A-B). This is consistent with the role of SPI1 as a positive regulator of inflammation (Ghisletti et al., 2010; Karpurapu et al., 2011; Turkistany and DeKoter, 2011), and suggests that loss of SPI1 upon LOUP knockdown is not responsible for the increase in inflammatory gene expression.

A.



B.

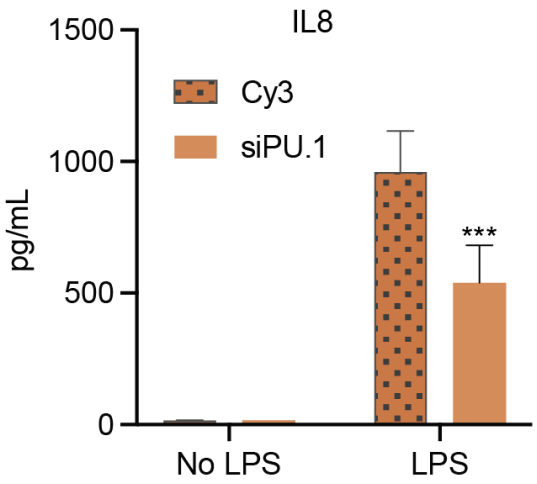


Figure 2. 5

Figure 2.5- The effect of loss of SPI1 on inflammation.
A. Knockdown of SPI1. Western blot analysis of SPI11 in WT THP1 cells transfected with a SPI1 targeting siRNA (siSPI1) compared to a control non-targeting siRNA (siCy3). Samples were collected at baseline (0) and after 6 and 24 hours of LPS treatment. Quantitative values indicate densitometry ratios of SPI1:Actin. Blot is representative of 3 separate experiments. **B. Effect of SPI1 knockdown on IL8.** IL8 ELISA was performed on supernatant from control and LOUP siRNA THP1s treated with LPS for 24 hours. Bars represent an average of two separate experiments each measured in triplicate with standard deviation (p-value<0.0005).

2.3.4 A LOUP sORF encoded peptide (SEP) functions as a negative regulator of NFkB

LOUP RNA was detected in both the nucleus and the cytoplasm (Fig 2.6A), which is consistent with its roles in regulating transcription of SPI1 and perhaps the effects seen more broadly on gene targets of NFkB. It's been found that some cytoplasmic lncRNAs harbor short open reading frames (sORFs) capable of producing peptides less than 100 amino acids in length, and that these peptides can carry out important functions (Ruiz-Orera et al., 2020). To investigate the coding potential of LOUP we first examined existing Ribo-seq datasets from THP1s and primary macrophages (Fig 2.6B). Based on ribosome footprints we established that LOUP harbors three potential sORFs. To test the translational potential of these three sORFs, each of them was cloned in frame with GFP to create a translational fusion and introduced into THP1-NFkB-CRISPRi-LOUP knockdown cells. Measurable GFP expression was driven by all three sORFs (Fig 2.6C). sgRNAs were designed to target the sORF regions as outlined in (Fig 2.6D). To determine if disruption of any sORFs resulted in an increase of inflammatory cytokines at baseline as it does when transcription of the entire locus is suppressed, we measured transcript levels of IKBKE and NFKBIZ. These two transcripts were most significantly upregulated at baseline in the LOUP CRISPRi knockouts (Fig 2.3C), and interestingly both of these genes were significantly upregulated by the sgRNA targeting sORF1 and sORF2, but not the sgRNA uniquely targeting sORF2 (Fig

2.6E). It is possible that this sgRNA did not work or that only ORF1 is functional. Together this data confirms that LOUP is a bimodal locus capable of regulating its neighboring gene through an enhancer mechanism as well as encoding an SEP that functions to negatively regulate NFKB genes in monocytes.

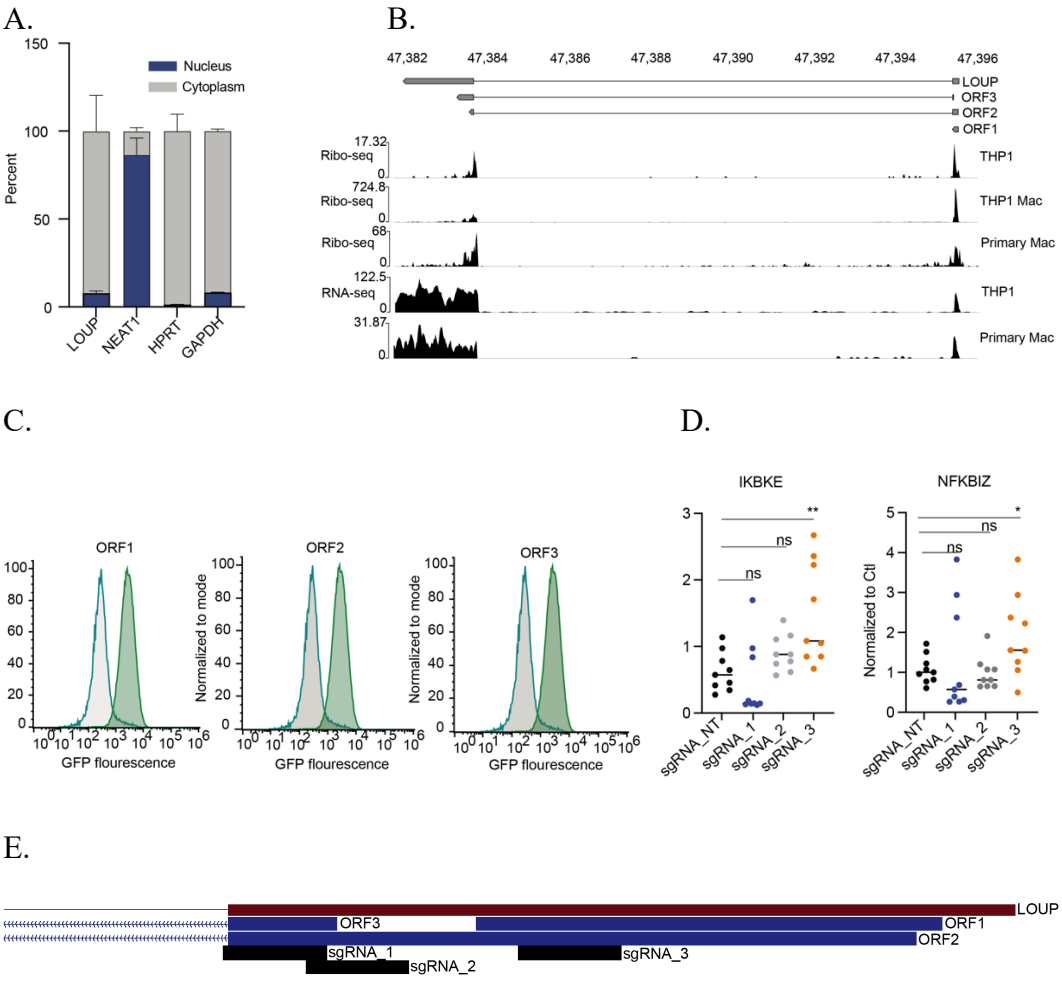


Figure 2. 6

Figure 2.6- LOUP encodes a SEP that can regulate NFkB target genes.
A. Cellular localization of LOUP in THP1s. Nuclear/cytoplasmic fractionation of WT THP1 cells after 6 hours of LPS stimulation. Data is from a single experiment with propagated error calculated for triplicate measurements. **B. LOUP coding potential.** Annotations of two LOUP isoforms and of three small open reading frames (sORFs) predicted for LOUP gleaned from Ribo-seq tracks generated from THP1s, differentiated THP1s and primary macs. Also included are tracks for short read RNA-seq from THP1s and long read RNA-seq (R2C2) from primary macs. **C. Expression of LOUP ORF-GFP.** Flow cytometry measuring GFP expression in LOUP knockdown THP1s transfected with sORF-GFP fusion constructs. **D. sORF regulation of inflammatory genes.** Two of the top most upregulated genes in the LOUP CRISPRi knockdowns were measured in LOUP Cas9 sORF targeted cells. Expression of IKBKE and NFKBIZ were measured at baseline by qPCR in each of 3 Cas9 cell lines (sgRNA_1, sgRNA_2, sgRNA_3) along with a non-targeting control (sgRNA_NT) (ns = not significant, * = p-

value<0.05, ** = p-value<0.001). **E. Cas9 targeting of sORFs.** Depiction of sgRNAs targeting sORFs with CRISPR-Cas9. sgRNAs 1-3 correspond to panel D.

2.4 Discussion

Here we've described a genome-wide pooled CRISPRi screen to identify novel lncRNA genes that regulate NFkB signaling in human monocytes. NFkB signaling pathways have been previously well studied and protein coding genes that regulate the signaling cascades have thought to have been mostly resolved (Gilmore, 1999; Neumann and Naumann, 2007; Verma, 2004; Zinatizadeh et al., 2021). Not only have we successfully identified novel lncRNA regulators of the TLR4/NFkB pathway, but we have also unveiled novel protein regulators as well. lncRNAs that share a promoter or overlap with protein coding regions pose a great challenge for functional characterization. By targeting these genes with CRISPRi we may have simultaneously disrupted the function of the overlapping coding gene, regardless of how specifically targeted the dCas9-KRAB was to the lncRNA TSS. Despite this being a caveat for identifying functional lncRNAs, we have identified novel coding genes that function in this context. Further studies using classic CRISPR Cas9 to disrupt coding potential of these genes and measure effects on NFkB signaling are warranted to confirm that it is indeed the protein mediating the phenotype (Table 1). Although phenotypes associated with lncRNAs can often be more subtle than those of proteins, and reporter-based screens are inherently noisy, we've used MAUDE analysis with a stringent z-score threshold in an effort to uphold confidence in our list of functional lncRNA candidates. One candidate in particular, LOUP, that has been previously identified as an enhancer lncRNA in

myeloid cells (Trinh et al., 2021), served as an excellent target for further functional validation. Not only were we able to confirm its ability to function as an enhancer in a monocyte cell line, but we've also begun to unravel its novel mechanistic role as a negative regulator of TLR4/NFkB signaling.

It is known that activation of TLR4 signaling can prime defined enhancer regions in both human and mouse macrophages (Hah et al., 2015; Kaikkonen et al., 2013). In addition to the specific histone modifications that define enhancer regions, the actual transcription of enhancer RNAs from these regions (eRNAs) seems to be an important mechanism for enhancer function during TLR4/NFkB-driven gene expression (Kaikkonen et al., 2013). Although the question still remains whether the production of eRNAs is only a consequence of transcriptional regulation by these loci or if these eRNA transcripts themselves are functional. It has been established that many lncRNA genes can also exert transcriptional regulation on nearby genes without having the definitive histone marks that delineate an enhancer (Chen et al., 2017; Gil and Ulitsky, 2020). Cis upstream regulatory elements of SPI1 have also been previously identified (Bonadies et al., 2010; Leddin et al., 2011; Li et al., 2001; Okuno et al., 2005), but more recently it has been shown that the functional and myeloid-specific lncRNA LOUP, is transcribed from an upstream element of SPI1 (Trinh et al., 2021). Here the authors demonstrated that the lncRNA transcript itself mediates the interaction between the SPI1 promoter and the transcription factor RUNX1, but that the LOUP gene does

exhibit histone modifications consistent with an enhancer (high H3K4me1, low H3K4me3). While we first identified LOUP for its ability to regulate the activity of TLR4/NFkB signaling in our screen, we were able to confirm the previous finding that it acts as an enhancer of SPI1. SPI1 is a myeloid lineage-determining transcription factor that is also active in the inflammatory response of monocytes and macrophages (Ghisletti et al., 2010; Karpurapu et al., 2011; Turkistany and DeKoter, 2011). In addition to the fact that LOUP and SPI1 occupy the same TAD, we've shown that SPI1 is induced in monocytes (Fig 2A) upon treatment with LPS and that absence of LOUP decreases SPI1 expression, especially at baseline. Interestingly, expression of SPI1 in these monocytes seems to be able to overcome the absence of LOUP by 24 hours, perhaps indicating the essential role of the protein in these cells, and that there are multiple routes that can lead to its regulation (Bonadies et al., 2010; Bulger and Groudine, 2010; Spitz and Furlong, 2012).

Interestingly, we found that hundreds of inflammatory genes are upregulated upon suppression of the LOUP locus both at baseline and upon LPS stimulation. Of the secreted cytokines measured in the LOUP knockdowns (Fig 2.4C), IL-8 increased most measurably, but was very significantly decreased in response to knocking down SPI1, indicating that LOUP's effect on IL8 is independent of its role as an enhancer of SPI1. This was not surprising given SPI1's known role in positively regulating expression of inflammatory genes. Given that LOUP was a top hit in our NFkB-reporter based screen, along with the fact that

LOUP is present in the cytoplasm and harbors three small ORFs, led us to investigate the possibility that LOUP encodes a functional peptide capable of negatively regulating the NFkB-driven inflammatory response in these cells. Based on the association of LOUP's sORFs with ribosomes together with the fact that targeting one sORF in particular resulted in upregulation of inflammatory genes, we have reason to believe that LOUP produces a functional peptide, although more work needs to be done to characterize the peptide. The minimum cutoff employed for annotated protein coding genes is typically 100 amino acids. It's now appreciated that many annotated lncRNAs harbor sORFs that would result in peptides well below this cut-off (~44 amino acids) (Ruiz-Orera et al., 2020). CRISPR-Cas9 disruption of sORFs 1 and 2, which are predicted to be 30 and 59 amino acids respectively, resulted in an increase in inflammatory gene expression, while disruption of sORF3, predicted to be 123 amino acids, did not have an effect (Fig 2.6D). While questions remain about the functionality of these peptides under 100 amino acids, there is mounting evidence suggesting that they have evolved to carry out important immunological functions (Couso and Patraquim, 2017; Jackson et al., 2018; Ji et al., 2015; Malekos and Carpenter, 2022). While limited, there is some evidence supporting the potential for both the lncRNA transcript itself along with SEPs harbored within the lncRNA gene to have distinct functional roles. For instance, the lncRNA *Aw112010,* has been shown to downregulate the transcription of IL-10 through interactions with the histone demethylase KMD5, thereby controlling T-cell differentiation (Yang et al., 2020). *Aw112010* also produces a

LPS-inducible SEP (<100 amino acids) that regulates the mucosal inflammatory response in mice without effects on IL-10 (Jackson et al., 2001). Our work here illustrates that a similar bi-modal mechanism might be the case for the LOUP gene, in which the lncRNA regulates transcription of its neighboring gene in *cis* while the SEP modulates the NFkB-driven inflammatory response in *trans*.

Here we have made clear the value of pooled CRISPRi screening as an efficient method to identify functional lncRNAs in the context of innate immunity. Meticulous control of TLR4/NFkB signaling is crucial for a proper immune response and the screen performed here has demonstrated that lncRNAs play an important role in maintaining the pathway. Understanding lncRNA function in this context has led to novel insights into inflammatory gene regulation and even the bi-modal functional capabilities of lncRNAs to regulate gene expression in *cis* and *trans*. Here we described a novel role for a myeloid-specific lncRNA as a potent regulator of inflammatory gene expression. This work will serve as a valuable resource of both lncRNAs and coding genes previously undiscovered in this pathway, as well as an important foundation for further mechanistic understanding of functional SEPs.

2.5 Methods

2.5.1 Cell lines

Wildtype (WT) THP1 cells were obtained from ATCC. All THP1 cell lines were cultured in RPMI 1640 supplemented with 10% low-endotoxin fetal bovine serum (ThermoFisher), 1X penicillin/streptomycin, and incubated at 37°C in 5% CO2.

Lentivirus production

All constructs were cotransfected into HEK293T cells with lentiviral packaging vectors psPAX (Addgene cat#12260) and pMD2.g (Addgene cat#12259) using Lipofectamine 3000 (ThermoFisher cat# L3000001) according to the manufacturer's protocol. Viral supernatant was harvested 72 hours post transfection.

THP1-NFkB-EGFP-dCasKRAB

We constructed a GFP-based NF-κB reporter system by adding 5x NF-κB-binding motifs (GGGAATTTCC) upstream of the minimal CMV promoter-driven EGFP. THP1s were lentivirally infected and clonally selected for optimal reporter activity. Reporter cells were then lentivirally infected with the dCas9 construct that was constructed from a pSico lentiviral backbone with an EF1a promoter expressing T2A flanked genes: blastocidin-resistant (blast), blue fluorescent protein, and humanized dCasKRAB. Cells were clonally selected for knockdown efficiency greater than 90%.

THP1-NFkB-EGFP-dCasKRAB-sgRNA (LOUP knockdown)

NFkB-EGFP-CRISPRi-THP1 cells were lentivirally infected with sgRNAs. sgRNA constructs were made from a pSico lentiviral backbone driven by an EF1a promoter expressing T2A flanked genes: puromycin resistance and mCherry. sgRNAs were expressed from a mouse U6 promoter. 20-Nucleotide forward/reverse gRNA oligonucleotides were annealed and cloned via the AarI site.

THP1-NFkB-EGFP-dCasKRAB- LOUP-/- sORF+

sORF-GFP fragments were synthesized by Twist Biosciences and cloned into a pSico lentiviral backbone. Constructs were then packaged into lentiviral particles as described above. Unstimulated GFP+ cells were sorted by FACS on a BD FACSAria II two times to achieve a 100% GFP positive population assuming that GFP expression in unstimulated cells was not activation of the reporter. Cells were consistently cultured under blasticidin and puromycin to maintain active dCas9 and sgRNA expression.

THP1-NFkB-EGFP-Cas9

The NFkB reporter was introduced as described above. The Cas9 construct was constructed from a pSico lentiviral backbone with an EF1a promoter expressing T2A flanked genes: blastocidin-resistant (blast), blue fluorescent

protein, and humanized *Streptococcus pyogenes* Cas9.Cells were clonally selected for knockdown efficiency greater than 90%.

2.5.2 Screen

sgRNA library design and cloning

10 sgRNAs were designed for each TSS of hg19 annotated lncRNAs expressed in THP1s at baseline and upon stimulation. The sgRNA library also included 700 non-targeting control sgRNAs, and sgRNAs targeting 50 protein coding genes as positive controls. The sgRNA library was designed and cloned as previously described in (S. J. Liu et al., 2017).

CRISPRi FACS Screen

Library infected and selected THP1-NFkB-EGFP-CRISPRi-sgRNA cells were expanded to 2000X coverage. Cells were stimulated with LPS (200 ng/mL) for 24 h to induce expression of NFkB-EGFP. Flow cytometry and PCR amplification of genomic sgRNA sequences were conducted as previously described in detail in (Covarrubias et al., 2020).

Screen Analysis

SgRNA guide adapters were removed with cutadapt (Martin, 2011) and counts were obtained with the mageck count function from MAGeCK (Li et al., 2014). Z-scores for each gene were found by analyzing each replicate

independently with MAUDE (de Boer et al., 2020). For each gene in the experiment, aggregate z-scores were generated using Stouffer's method and a combined false-discovery rate was calculated.

2.5.3 Sequencing Data

Data for ATACSeq, ChIPSeq and HiC analyses in THP-1s were originally reported in (Phanstiel et al., 2017) and are available at GEO: GSE96800 and SRA: PRJNA385337. RiboSeq data is from (Ansari et al., 2022; Fritsch et al., 2012; Su et al., 2015), with data available at GSE208041, GSE66810, GSE39561 respectively.

ATAC and ChIP Seq

Adapters were trimmed with NGmerge and mapped to Gencode GRCh38 primary assembly with `Bowtie2 (--very-sensitive --maxins 1000)` (Langmead and Salzberg, 2012). Replicates were merged and alignments were converted to BigWig tracks with the `bamCoverage (--binsize 1)` module from deepTools (Ramírez et al., 2016).

HiC

Paired-end reads from untreated THP-1s were deduplicated with BBMap `clumpify (dedupe=t ziplevel=3 reorder=t compresstemp=f deleteinput=t)` (https://sourceforge.net/projects/bbmap/). Fastqs were converted to Pairs format with Chromap (`-e 4 -q 1 --split-alignment -`

`-pairs`) (Zhang et al., 2021) and replicates were merged with Pairtools (Open2C et al., 2023). Cool format files were generated with Cooler (`cooler cload pairs`) at 5kb resolution, then normalized with `hicNormalize (-n smallest -sz 1.0)` and Knight-Ruiz corrected with `hicCorrectMatrix (--correctionMethod KR)` (Abdennur and Mirny, 2020). TADs were called with `hicFindTADs ( --minDepth 15000 --maxDepth 150000 --step 15000 --thresholdComparisons 0.05 --delta 0.01 --correctForMultipleTesting fdr)` from the HiCExplorer suite (Ramírez et al., 2016).

RiboSeq

Adapters were removed with Cutadapt (Martin, 2011). A tRNA/rRNA index was built with Bowtie2 (Langmead and Salzberg, 2012) and mapped reads were discarded. The remaining reads were mapped with STAR to the Hg38 genome, guided by a custom annotation set composed of Gencode v41 merged with the published isoform atlas from primary macrophages (Vollmers et al., 2021). Multimapping reads were discarded.

2.5.4 Western Blots

Cell lysates were quantified by the Pierce™ BCA Protein Assay Kit. Equal amounts of protein (15 ug) of each sample were denatured at 70°C for 10 min prior to loading on 12% SDS-PAGE. Samples were transferred to polyvinylidene difluoride (PVDF) membranes using the Trans-Blot Turbo Transfer System (Bio-

Rad), blocked with TBST (1x Tris buffered saline with 0.1% Tween 20) supplemented with 5% (wt/vol) BSA for 1 hr and blotted with PU.1 (9G7) (1:1000, Cell Signaling #2258) at 4°C overnight. Horseradish peroxidase (HRP)-conjugated goat anti-rabbit (1:2000, Bio-Rad, #1706515) secondary antibody was used. Western Blots were developed using SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Scientific cat# 34577). After imaging HRP was inactivated for 2 hours in 0.2% Sodium Azide in TBST supplemented with 5% (wt/vol) BSA. B-Actin (C4) monoclonal antibody (1:500, Santa Cruz Biotechnology cat #47778) was subsequently used as a loading control followed by HRP-conjugated goat anti-rabbit secondary antibody (1:2000).

2.5.5 siRNA knockdown of SPI1

WT THP1 cells were transfected with 60 pmol of SPI1 targeting (ThermoFisher cat# HSS186060) or Cy3-conjugated non-targeting siRNA (ThermoFisher cat# AM4621) for 72 hours using Lipofectamine 3000 (ThermoFisher cat# L3000001) according to the manufacturer's protocol.

2.5.6 Nanostring multiplexed transcript analysis

RNA was isolated using the Direct-zol RNA Miniprep Plus Kit (Zymo cat# R2052) from one control and two LOUP knockdown THP1 cell lines at baseline and after 6 hours of LPS treatment. RNA was quantified on a Nanodrop. 100ng of RNA was used for each sample hybridization (Nanostring Master Kit cat# 100052)

and detection with the Nanostring Human Immunology V2 nCounter GX Codeset (cat# 115000062) on a MAX/Flex nCounter according to the manufacturer's protocol.

2.5.7 ELISA and Multiplexed ELISA

WT THP1 cells transfected with SPI1 targeting siRNAs were seeded at equal densities and stimulated with LPS for 24 hours. Samples were diluted 1:10 and IL8 was measured using the DuoSet IL8 ELISA kit (R&D Systems cat# DY208) following the manufacturer's protocol. THP1-NFkB-EGFP-CRISPRi-sgRNA LOUP knockdown and control cells were seeded at equal densities and treated with LPS for 24 hours. 75ul of supernatants were collected and analyzed undiluted by EVE Technologies using their Human Cytokine Panel A 48-Plex Discovery Assay (cat# HD48A).

2.5.8 Nuc/Cyt fractionation and RT-qPCR

WT THP1 cells were fractionated according to the NE-PER kit protocol (ThermoFisher cat# 78833) with RNAse inhibitor (Superase-IN, ThermoFisher cat# AM2696) added to the cytosolic and nuclear lysis buffers. 3 volumes of Trizol (TRI Reagent, Sigma T9424) was added to the fractions and RNA was isolated using the DIrect-zol RNA Miniprep Plus Kit (Zymo cat# R2052). 16uL of RNA isolated from fractions was reverse transcribed (iScript cDNA synthesis kit, Bio-Rad cat# 1708840) followed by qPCR (iTaq SYBRgreen Supermix, Bio-Rad cat#

1725121) using the cycling conditions as follows: 50C for 2 min, 95C for 2 min followed by 40 cycles of 95C for 15s, 60C for 30s and 72C for 45s.

2.6 Acknowledgments

I would like to thank Dr. Sergio Covarrubius at the UCSC CRISPR Core for creating the reporter-CRISPRi cell lines and conducting the NFkB lncRNA screen. I also sincerely thank Eric Malekos for doing the computational analysis of the screen as well as computational analysis of Hi-C, ATAC-seq, ChIP-seq, and Ribo-seq data sets. I thank Dr. Lisa Sudek for assisting with the western blots. I thank Aswad Kahdilkar for advising and assisting me with the nanostring expereiments, as well as Bari Nazario at the UCSC Institute for the Biology of Stem Cells Flow Cytometry Facility for assisting with FACS during the creation of the sORF-expressing cells. Lastly, I thank Leila Namvar and Susan Carpenter for assisting with the final RNA isolations and qPCR measurements of inflammatory genes in sORF knockdown cells. All these people have made significant contributions to the completion of this work.

CHAPTER 3- Conclusion and future directions

3.1 Conclusions

The work presented here exhibits the molecular importance of a class of molecules that are still relatively under studied. A large majority of the genome is non-coding (Cabili et al., 2011; Uszczynska-Ratajczak et al., 2018), and while the field of study of non-coding RNAs is progressing rapidly, there is still a need for high throughput studies that can efficiently identify functional non-coding genes, particularly in the context of inflammation and immunity. Here we have implemented CRIPSRi pooled screening to gain a better understanding of the functions on non-coding RNAs in a key inflammatory signaling pathway, and more specifically in monocytes. We successfully identified 32 genes that are expressed in monocytes and that have not been previously characterized as regulators of NFkB. While only 10 of the top 32 hits are annotated intergenic lncRNAs, and the others are antisense or bidirectional to coding genes (Table 2.1), this list of genes provides a great resource for the further study of both lncRNAs and proteins that may have been simultaneously targeted by CRIPSRi. Given that the NFkB pathway has been thoroughly studied (Gilmore, 1999; Neumann and Naumann, 2007; Verma, 2004; Zinatizadeh et al., 2021), the finding that there are still so many undiscovered cell-specific regulators of the pathway is an important breakthrough that has progressed our understanding of the great complexity of inflammatory signaling.

Beyond identifying a number of functional lncRNA genes in a high throughput manner, we were able to begin to uncover the intricate mechanism by which one such lncRNA, LOUP functions to regulate inflammation. To our surprise, we first identified LOUP as a negative regulator of NFkB, despite its ability to act as an enhancer for its neighboring gene, SPI1, a known positive regulator of inflammation. Perhaps even more surprising is that LOUP's broad negative regulation of NFkB target genes happens despite its enhancement of SPI1 expression, leading us to hypothesize that while LOUP *enhances* SPI1 expression, especially at baseline and the shorter exposures to LPS, it is not the only regulator of SPI1 expression. In trying to understand how LOUP negatively regulates inflammation, we discovered that LOUP is primarily present in the cytoplasm, immediately raising the question of whether LOUP has coding potential. In further investigation of LOUP's coding potential, we uncovered evidence that LOUP does harbor three short ORFs and does in fact produce a functional peptide capable of negatively regulating inflammatory genes. Not only does this work advance our understanding of lncRNAs that control gene regulation and inflammation, but it also advances our knowledge of the functional roles of small peptides, a field that is still quite immature.

Understanding how inflammatory signaling is regulated has broad implications for translational science. Tight control over inflammatory signaling is crucial for a proper immune response. Chronic expression and secretion of inflammatory cytokines, even at low levels can drive maladies, including

cardiovascular, neurodegenerative, and metabolic diseases. The ultimate hope in doing this research, and developing comprehensive knowledge of inflammatory signaling pathways, is that this expertise may lead to better much prevention and treatments of one of the lead causes of disease.

3.2 Future directions

While this work contributes novel discoveries that span the fields of lncRNAs, small peptides, inflammatory signaling, and innate immunity, it brings up many new questions that warrant investigation. For one, the results from the screen have left us with many genes to mechanistically characterize and understand in the context of NFkB signaling. Likewise, despite the great deal of work done here, LOUP's molecular mechanism is still a bit of a mystery. While we could see that LOUP's sORFs cloned into an artificial system get translated, and targeting LOUP's sORF regions with classic CRISPR-Cas9 results in an inflammatory phenotype, we still don't know which sORF produces a functional peptide. Isolation of short peptides from THP1s followed by mass spectrometry could help identify the functional SEP. Introducing tagged sORFs to THP1 cells followed by pull-down assays to determine any molecular interactions with translated SEPs could also shed light on SEP function.

In addition to defining which of LOUP's SEPs are truly functional, the question remains whether LOUP functions the same in primary monocytes as it

does in THP1s. The experiments done here in THP1s may be too technically challenging in primary cells, but the biological relevance of cell lines is always in question. Another way to address the true biological relevance of LOUP's mechanism in THP1s will be to understand its function in a whole organism, particularly a mouse. Hence determining if LOUP is conserved in mice is of great value, and not terribly straight forward for non-coding regions given that they are subject to different selective pressures than coding regions (Ulitsky et al., 2011), and many have arisen in the human genome relatively recently (Ulitsky and Bartel, 2013). So far, there does appear to be syntenic conservation of the LOUP locus in mice, but currently no evidence of sORFs. The next step will be to confirm expression of LOUP in mouse monocytes.

Overall, this work has established an exciting basis for future studies and will serve as the foundation for research that will continue to expand our understanding of genes that control inflammatory signaling and innate immunity.

3.3 Outstanding questions

- Which of LOUP's SEP is functional?
- How does LOUP's SEP function to regulate NFkB?
- How does LOUP or LOUP's SEP function in primary monocytes?
- Is LOUP or LOUP's SEP functionally conserved in mice*?*

APPENDIX 1- Supplemental information to Chapter 2

Supplemental tables for the NFkB CRISPRi lncRNA screen (presented in Chapter 2 -Figure 2.1, Table 2.1) will be available for download once puplished- they will include the sgRNA library and the detailed results of the MAUDE analysis. Since we screened over 2000 genes, with 10 sgRNAs per gene, the tables are too large to include in this printed document.

APPENDIX 2- Mutant KRAS Regulates Transposable Element RNA and Innate Immunity via KRAB Zinc-Finger Genes

Roman E. Reggiardo, Sreelakshmi Velandi Maroli, Haley Halasz, Mehmet Ozen, Eva Hrabeta-Robinson, Amit Behera, Vikas Peddu, David Carrillo, Erin LaMontagne, Lila Whitehead, Eejung Kim, Shivani Malik, Jason Fernandes, Georgi Marinov, Eric Collisson, Angela Brooks, Utkan Demirci, Daniel H. Kim.

My contributions to the following manuscript:

The following work focuses on understanding the earliest stages of cellular transformation driven by a common and highly pathogenic mutation in the KRAS gene. I worked on the very initial stages of this project. I optimized the transformation protocol for the AALE cells by transient transfection of a construct harboring the KRAS G12D mutation. I subsequently collected RNA from the cells. I made RNA-seq libraries from the RNA of these cells as described in the methods section. One of the lab's technicians, Lila Whitehead, assisted me in the library preparation protocol. The RNA-sequencing data that was generated contributes to Figure 1 and Supplemental Figures 1, 3, and 5 of this manuscript

and formed the basis for which subsequent studies were built. Roman Reggiardo and Georgi Marinov did the computational analysis of this RNA-seq data.

Acknowledgments

We thank members of the Kim lab, the Demirci lab, the Brooks lab, the Collisson lab, the Haussler lab, and the Carpenter lab for helpful discussions. This work was supported by funds from the Baskin School of Engineering at University of California, Santa Cruz (to D.H.K.), the Ken and Gloria Levy Fund for RNA Biology (to D.H.K.), the Department of Defense Congressionally Directed Medical Research Program (W81XWH-20-1-0746 to U.D. and D.H.K.), and the National Cancer Institute (R01CA227807, R01CA239604, R01CA230263 to E.C.). R.E.R. is supported by the National Institutes of Health (1F99DK131504-01), S.V.M. is supported by the California Institute for Regenerative Medicine (EDUC4-12759), and V.P. (T32DT4904) and D.C. (T30DT0997) are supported by the Tobacco-Related Disease Research Program.

Author contributions

Conceptualization, D.H.K.; methodology, D.H.K. and R.E.R.; investigation, S.V.M., H.H., M.O., E.H.-R., A.B., D.C., E.L., L.W., E.K., and S.M.; formal analysis, R.E.R., V.P., and G.M.; data curation, R.E.R.; writing – original draft, D.H.K. and R.E.R.; writing – review & editing, D.H.K., R.E.R., and U.D.;

funding acquisition, D.H.K., U.D., A.B., and E.C.; resources, D.H.K., U.D., A.B., E.C., and J.F.; supervision, D.H.K., U.D., A.B., and E.C.

2.1 Abstract

RAS genes are the most frequently mutated oncogenes in cancer, yet the effects of oncogenic RAS signaling on the noncoding transcriptome remain unclear. We analyzed the transcriptomes of human airway and bronchial epithelial cells transformed with mutant KRAS to define the landscape of KRAS-regulated noncoding RNAs. We find that oncogenic KRAS signaling upregulates noncoding transcripts throughout the genome, many of which arise from transposable elements (TEs). These TE RNAs exhibit differential expression, are preferentially released in extracellular vesicles, and are regulated by KRAB zincfinger (KZNF) genes, which are broadly downregulated in mutant KRAS cells and lung adenocarcinomas in vivo. Moreover, mutant KRAS induces an intrinsic IFN-stimulated gene (ISG) signature that is often seen across many different cancers. Our results indicate that mutant KRAS remodels the repetitive non-coding transcriptome, demonstrating the broad scope of intracellular and extracellular RNAs regulated by this oncogenic signaling pathway.

2.2 Introduction

Most of the human genome is noncoding and transcribed into RNA (Djebali et al., 2012). Moreover, about half of the human genome is comprised of transposable elements (TEs) (International Human Genome Sequencing Consortium et al., 2001), and TEs contribute substantially to the noncoding transcriptome (Kelley and Rinn, 2012; Rinn and Chang, 2020). TE RNAs (Burns, 2017) and other classes of noncoding RNAs are often altered during cancer (Slack and Chinnaiyan, 2019) and epigenetic reprogramming (Kim et al., 2015), where activation of RAS signaling leads to the repression of microRNAs (Kent et al., 2010) and the upregulation of long noncoding RNAs (lncRNAs) (Kim et al., 2015), respectively, via changes in chromatin accessibility. In lung cancers, RAS mutations are present in one-third of lung adenocarcinomas (The Cancer Genome Atlas Research Network, 2014) and serve as driver mutations that initiate tumorigenesis (Jackson et al., 2001). Although RAS genes are among the most frequently mutated oncogenes in cancer (Simanshu et al., 2017), how oncogenic RAS signaling regulates the noncoding transcriptome remains unknown. To investigate the role of mutant KRAS in reprogramming the

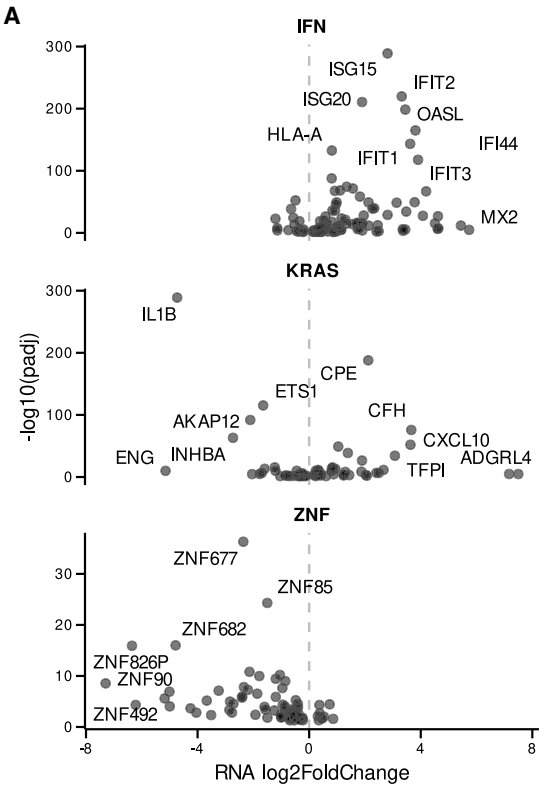
transcriptome during early stages of cellular transformation, we characterized the composition of both intracellular and extracellular RNA, including protein-coding RNA, lncRNA, and TE RNA, using human airway epithelial cells (Lundberg et al., 2002) and human bronchial epithelial cells (Ramirez et al., 2004) with constitutively active mutant KRAS. We show that oncogenic KRAS induces TE RNA and cell-intrinsic interferon (IFN)-stimulated gene (ISG) signatures and that

KRAB zinc finger (KZNF) genes are globally downregulated both in vitro and in mutant KRAS lung adenocarcinomas in vivo. Moreover, our findings indicate that significant upregulation and extracellular secretion of TE RNAs and ISGs are transcriptomic signatures of mutant KRAS signaling.

2.3 Results

2.3.1 Transcriptomic Reprogramming by Mutant KRAS

To determine the transcriptomic landscape of protein-coding and noncoding RNAs regulated by oncogenic RAS signaling, we performed RNA sequencing (RNA-seq) on human airway epithelial cells (AALE) that undergo malignant transformation upon the introduction of mutant KRAS (Lundberg et al., 2002). We compared the transcriptomes of AALE cells transduced with control lentiviral vector to AALEs that were transduced by mutant KRAS-containing lentiviral vector and performed differential expression analysis. We identified thousands of significantly differentially expressed protein-coding RNAs (n = 1,028 upregulated, n = 1,194 downregulated), including ISGs, KRAS signaling genes, and zinc-finger genes (Apendix 2 Figures 1A and S1A), as well as hundreds of significantly differentially expressed lncRNAs (n = 116 upregulated, n = 163 downregulated) (FigureS1A), demonstrating the broad extent to which mutant KRAS reprograms the transcriptome (Figures S1A and S1B).



Appendix 2 Figure 1

Apendix 2 Figure 1- Mutant KRAS Signaling activates an Intrinsic ISG Signature

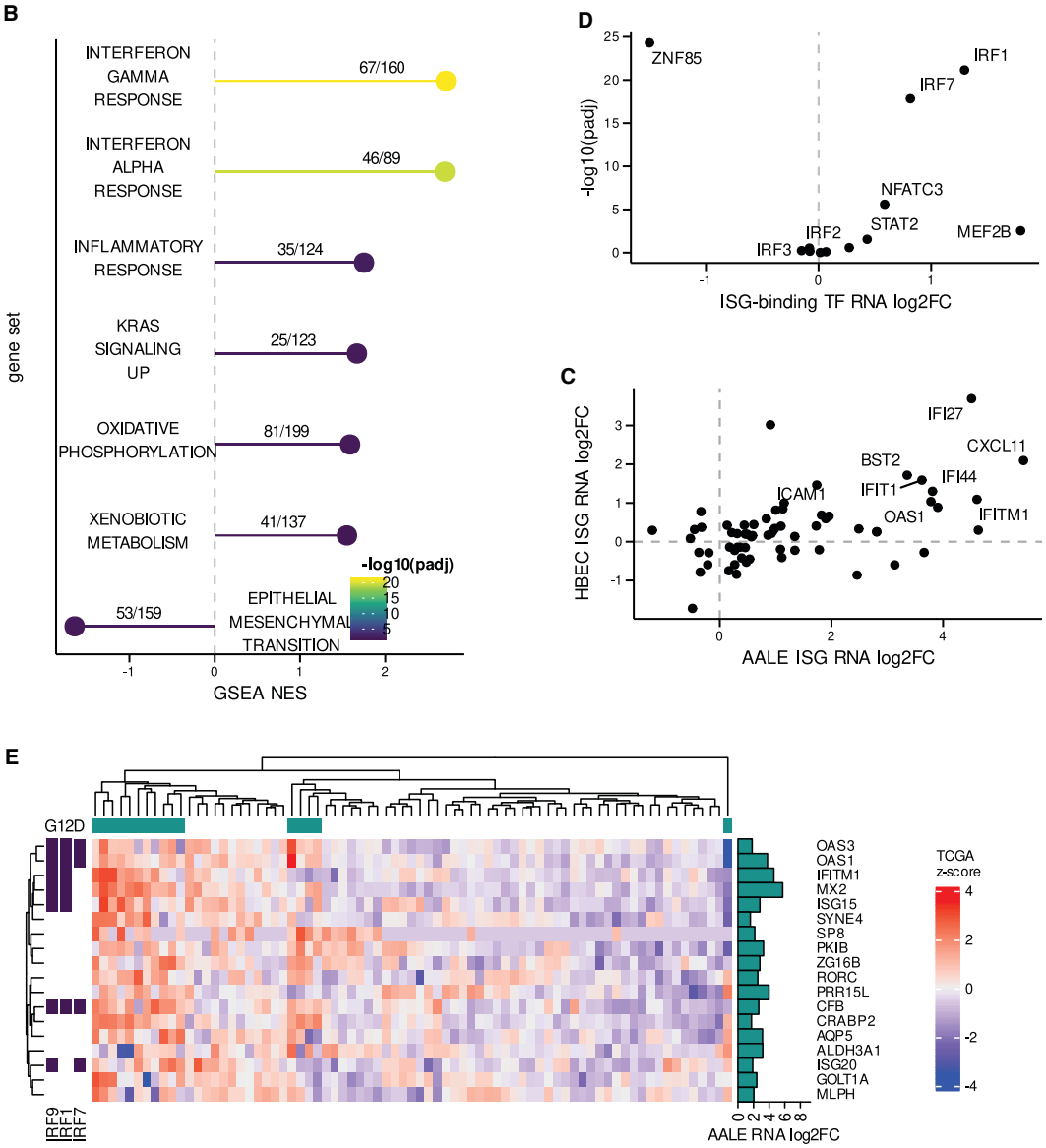
A. Volcano plots depicting significant differential expression observed in key gene sets (interferon [IFN] response alpha/gamma: IFN, KRAS signaling up: KRAS, zinc-finger genes: ZNF).

2.3.2 Mutant KRAS induces intrinsic ISG expression

To explore the biological pathways that are perturbed by onco- genic RAS signalling, we performed gene set enrichment analysis (GSEA) (Powers et al., 2018) using genes that were differentially expressed in our mutant KRAS AALE cells. GSEA revealed that the three most significantly enriched pathways were the IFN-a and -g responses, as well as the hallmark inflammatory response (Appendix 2 Figure 2B), along with increased KRAS signalling from mutant KRAS(G12D), increased metabolic gene expression, and decreased expression of epithelial-to-mesenchymal transition (EMT) genes (Appendix 2 Figure 2B).

To further validate the connection observed between mutant KRAS and ISG expression, we compared mutant KRAS- induced ISGs in AALE cells to those that were induced in human bronchial epithelial cells (HBECs) in response to mutant KRAS(G12V) (Appendix 2 Figure S1C). We observed a strong concordance between mutant KRAS-induced ISGs in AALE and HBEC cells (Appendix 2 Figure 2C), confirming our previous results. We then examined the promoter regions (±500 bp) of upregulated ISGs and identified motifs enriched in comparison to non-differentially ex- pressed (DE) ISGs (Appendix 2 Figure 2D), including the key IFN response regulators IRF1, IRF7, and STAT2 (Jefferies, 2019). To deter- mine the *in vivo* relevance of our findings in both mutant KRAS AALE and HBEC cells, we examined ISG expression in mutant KRAS(G12D) lung adenocarcinomas

(LUAD) from The Cancer Genome Atlas (TCGA), which revealed a subset of ISGs that were upregulated in KRAS(G12D) tumors when compared to lung cancer samples with wild-type (WT) KRAS (Figure 1E). These results indicate that mutant KRAS signalling activates an intrinsic ISG response in lung cells both *in vitro* (AALE, HBEC) and *in vivo* (TCGA LUAD).



Appendix 2 Figure 2

Appendix 2 Figure 2- Mutant KRAS Signaling Activates an Intrinsic ISG Signature.

B. Significant gene set enrichment analysis (GSEA) results observed in mutant KRAS AALE differentially expressed genes ranked by adjusted p value (padj), normalized enrichment score (NES), and annotated with the number of genes observed out of the total genes in each gene set. C. Differential expression of ISGs in mutant KRAS AALEs compared to mutant KRAS HBECs. D. Differentially expressed transcription factors (TFs) with binding motifs enriched in differentially expressed ISG promoter regions. E. Hierarchical clustering of expression *Z* score

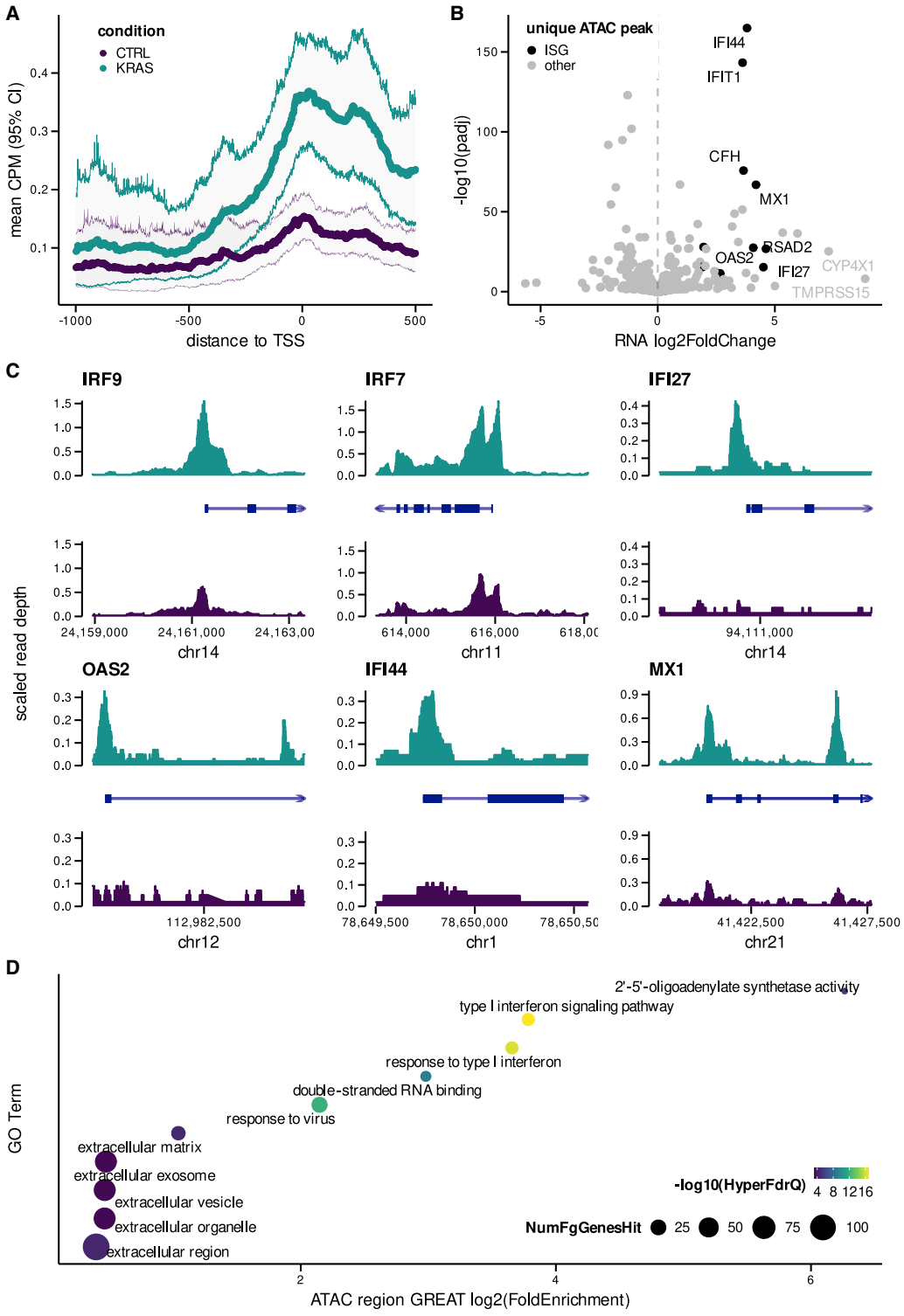
in TCGA LUAD RNA-seq data for ISGs upregulated in mutant AALE and exhibiting strong segregation in TCGA LUAD samples based on KRAS G12D mutation status; presence of IRF9/1/7 binding motifs in promoter regions of labeled ISGs.

2.3.3 Epigenetic reprogramming of ISGs by mutant KRAS.

To elucidate potential mechanisms involved in inducing ISG signatures in mutant KRAS AALE cells, we performed the assay for transposase-accessible chromatin using sequencing (ATAC-seq) (Buenrostro et al., 2015). In mutant KRAS AALEs, open chromatin was significantly enriched at gene promoters for upregulated ISGs (Appendix 2 Figure 3A). Open chromatin peaks were uniquely present in mutant KRAS AALEs when compared to control AALEs at 183 transcriptional start sites (TSSs), including 11 ISGs that were specifically and significantly upregulated by mutant KRAS signaling (Appendix 2 Figure 3B). In addition, we observed strong enrichment of ATAC signal at the TSS of the significantly upregulated IRF9 gene, which forms the ISGF3 transcription factor (TF) com- plex with STAT1 and STAT2 (MacMicking, 2012), and also strong enrichment at the TSS of IRF7, IFI27, OAS2, IFI44, and MX1 (Appendix 2 Figure 3C). In conjunction with the motif enrichment analysis (Appendix 2 Figure 2D), these results show that oncogenic KRAS signaling induces the epigenetic activation of ISG TFs and their downstream ISG targets.

The genome-wide effects of mutant KRAS-mediated epige- nomic reprogramming were further assessed with the Genome Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010). GREAT analysis orthogonally confirmed the enrichment of accessible chromatin regions near ISGs and showed

the enrichment of related molecular functions, including double- stranded RNA binding. Notably, the cellular components most enriched were extracellular in nature, including extracellular vesicle and extracellular exosome (Appendix 2 Figure 3D).



Appendix 2 Figure 3

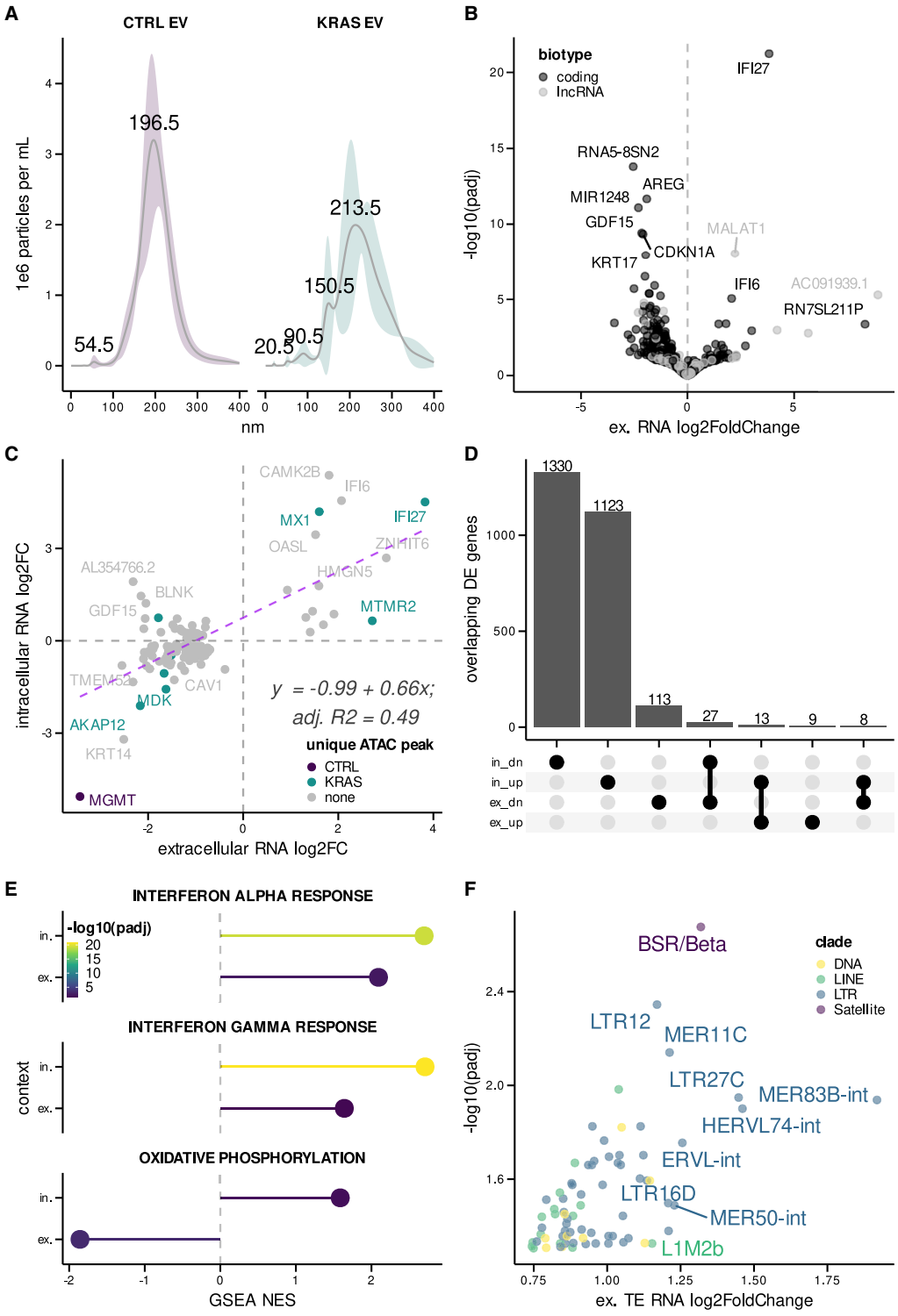
Appendix 2 Figure 3- Mutant KRAS signaling mediates epigenomic reprogramming of ISGs.
A. Mean ATAC-seq counts per million (CPM) (95% confidence interval [CI]) in promoter regions of upregulated ISGs (log2 fold change >1.5) in both mutant KRAS and control (CTRL) AALEs. B. Differential expression of ISGs with unique peaks near TSS (only present in mutant KRAS or control AALEs). C. ATAC-seq coverage in both mutant KRAS and CTRL AALEs for subset of ISGs with unique peaks detected near TSS. D. Significant Gene Ontology (GO) term enrichment over unique peaks detected in mutant KRAS AALEs as determined by genomic regions enrichment of annotations tool (GREAT) analysis.

2.3.4 Mutant KRAS reprograms the extracellular transcriptome.

To test whether extracellular RNAs secreted from mutant KRAS cells also exhibit differential expression relevant to intracellular reprogramming events, we isolated extracellular vesicles from the culture media of control and mutant KRAS AALEs (Enderle et al., 2015; Liu et al., 2017) (Enderle et al., 2015; F. Liu et al., 2017). Extracellular vesicles isolated from mutant KRAS AALEs comprised different sized vesicles that were 90, 150, and 213 nm in diameter, while vesicles from control AALE media were predominantly 196 nm in size (Appendix 2 Figure 4A).

RNA isolated and sequenced from these vesicles exhibited mutant KRAS-dependent differential expression of both protein-coding genes (n = 17 upregulated, n = 140 downregulated) and lncRNA (n = 5 upregulated, n = 8 downregulated) (Appendix 2 Figures 4B and S2A). We also observed significant correlation between differentially expressed ISGs in our intracellular and extracellular RNA-seq datasets that largely agreed with intracellular epigenetic changes (IFI6, MX1, IFI27, and OASL) (Appendix 2 Figures 4C and 4D). Furthermore, GSEA showed that IFN-a and -g signatures were enriched in both intracellular and extracellular RNA (Appendix 2 Figure 4E), indicating that extracellular RNAs reflect intracellular ISG changes due to mutant KRAS signaling. To determine the effects of oncogenic KRAS on noncoding RNA secretion, we also characterized the TE

RNAs that were preferentially packaged and released in extracellular vesicles. We found significant upregulation of predominantly long terminal repeat (LTR) RNAs such as LTR12, MER11C, and LTR27C, along with LINE, DNA, and Satellite repeat RNAs in mutant KRAS AALE extracellular vesicles (Appendix 2 Figure 4F). Moreover, TE RNAs represented approximately 50% of the extracellular RNA released from mutant KRAS AALE cells, suggesting their preferential secretion in extracellular vesicles (Appendix 2 Figures S2C and S2D).



Appendix 2 Figure 4

Appendix 2 Figure 4- Mutant KRAS signaling induces secretion of TE RNAs and ISGs in EVs.

A. Size distribution of extracellular vesicles (EV) isolated from control (CTRL) and mutant KRAS AALEs. B.Volcano plot of differentially secreted GENCODE protein-coding RNAs and lncRNAs between mutant KRAS and CTRL AALE EVs. C. Scatterplot comparing differentially expressed genes between intracellular and extracellular mutant KRAS AALE RNA-seq libraries; linear regression fit with formula and goodness of fit displayed. D. Upset plot summarizing overlap of differentially expressed upregulated (up) and downregulated (dn) genes across in and ex contexts. E. Significantly enriched gene sets detected in both in and ex contexts. F. Differential secretion of TE RNAs in EVs from mutant KRAS AALEs when compared to control AALE EVs.
Ex, extracellular; in, intracellular.

2.3.5 Regulation of TE RNAs by mutant KRAS.

Given the prevalence of secreted TE RNAs, we investigated intracellular TE RNA dynamics in response to mutant KRAS signaling in AALE cells. Analogous to extracellular RNAs, LTR RNAs were among the most significantly upregulated TE RNAs in response to oncogenic KRAS signaling, including LTR12C RNAs (Appendix 2 Figure S3A). In addition, LINE RNAs such as L1MEc and DNA element RNAs such as Tigger5 were also significantly enriched in mutant KRAS AALEs (Appendix 2 Figure S3A). Furthermore, we examined TE RNAs in mutant KRAS HBEC cells, which similarly exhibited significant upregulation of TE RNAs in response to mutant KRAS when compared to control HBECs (Appendix 2 Figure S3A). LTR12C RNAs were again the most significantly upregulated TE RNAs in mutant KRAS HBEC cells (Appendix 2 Figure S3A), further vali- dating our intracellular and extracellular RNA analyses in mutant KRAS AALE cells.

Based on the functions of KZNF genes in silencing TE RNAs in other contexts (Imbeault et al., 2017), we examined whether KZNFs could be involved in TE RNA regulation in both mutant KRAS AALE and HBEC cells. Given the broad downregulation of KZNFs in mutant KRAS AALEs (Appendix 2 Figure 1A), we also analyzed KZNF expression in mutant KRAS HBECs, which similarly exhibited significant downregulation of KZNFs, many of which overlap with

KZNFs downregulated in mutant KRAS AALEs (Appendix 2 Figure S3B). To determine the potential relationship between our upregulated TE RNAs and our downregulated KZNFs, we looked for significantly enriched motifs in TE RNAs using a previously described KZNF-specific motif set (Barazandeh et al., 2018), which confirmed the presence of binding motifs for significantly downregulated KZNFs in the significantly upregulated TE RNAs (Figure S3C). We also used the KNZF binding scores generated from previous chromatin immunoprecipitation sequencing (ChIP-seq) experiments (Imbeault et al., 2017) to rank TE RNAs targeted by KZNFs, finding that many of the upregulated TEs were among the top 10–20 targets of downregulated KZNFs in mutant KRAS AALEs (Appendix 2 Figure S3D), their extracellular vesicles (Appendix 2 Figure S4A), and in mutant KRAS HBECs (Appendix 2 Figure S4B). We then computed the average log2 fold change of downregulated ZNFs with putative binding sites within upregulated TE RNAs, which confirmed a negative association across all three contexts of mutant KRAS transcriptional profiling (Appendix 2 Figure S3E). These analyses point to a coordinated, TE-KZNF axis that is dysregulated by mutant KRAS.

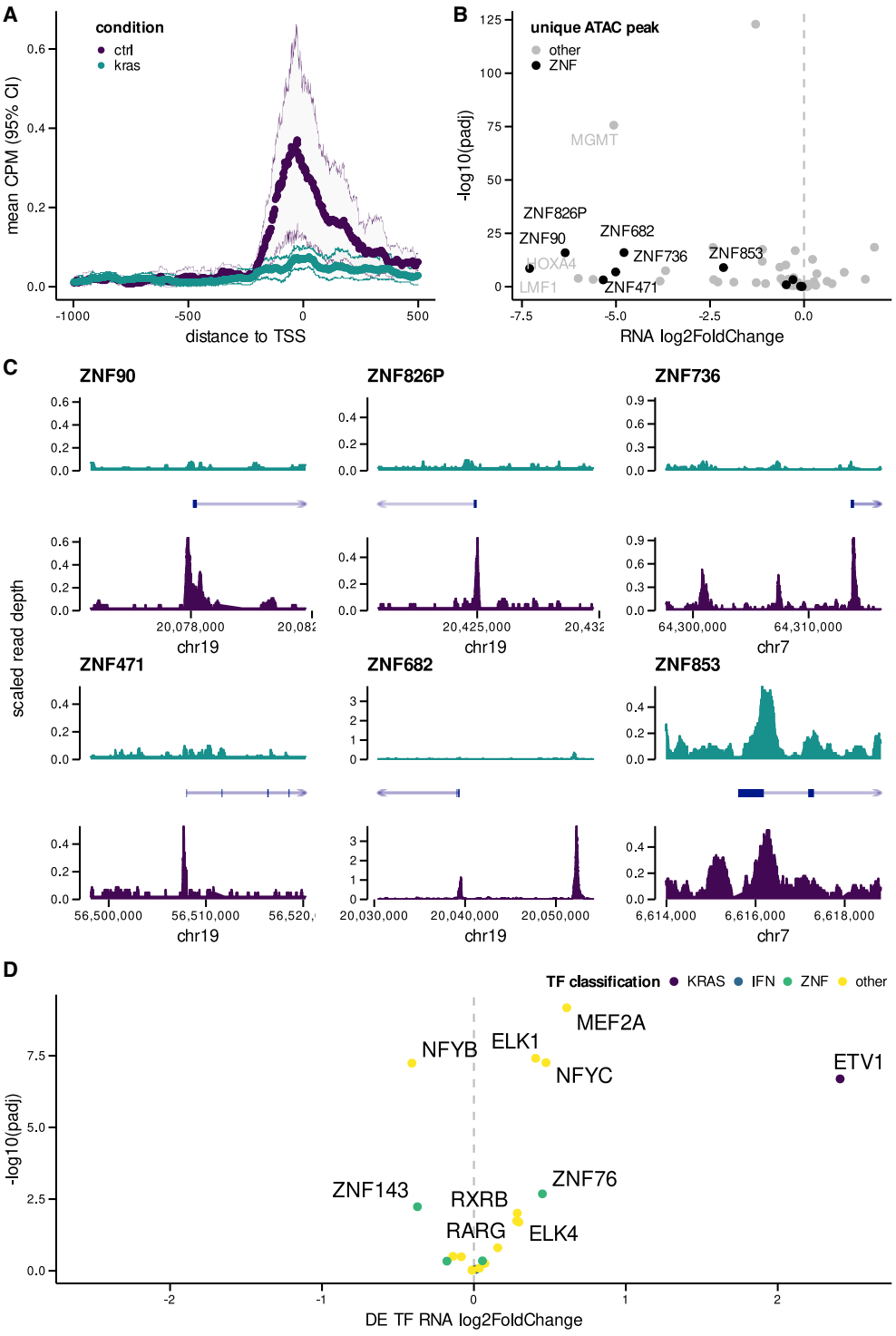
2.3.6 KZNFs repress TE RNAs and ISGs activated by mutant KRAS.

To explore the mechanistic relationship between KZNFs and TE RNA expression, we examined mutant KRAS A549 lung cancer cells that overexpress ZNF257 or ZNF682 (Ito et al., 2020), both of which we found to be significantly downregulated by mutant KRAS signaling in AALE cells and putative regulators

of dysregulated TE families (Appendix 2 Figures 1A and S3). Differential expression analysis of RNA-seq data indicated significant down- regulation of ISGs OAS1 and IRF9 in mutant KRAS A549 cells overexpressing either ZNF257 or ZNF682 (Figure S5A), as well as significant downregulation of TE RNAs that were upregulated by mutant KRAS in AALE cells (Appendix 2 Figure S5B). These findings directly connect mutant KRAS-regulated KZNFs with control of TE RNA and ISG expression.

2.3.7 Epigenetic silencing of KZNFs regulated by mutant KRAS signaling.

To determine the extent to which mutant KRAS signaling epigenetically silences KZNF expression, we examined ATAC-seq data for all significantly downregulated KZNF loci. We found that mutant KRAS signaling substantially reduces chromatin accessibility at TSS regions (Appendix 2 Figure 5A). When we examined genes with ‘‘unique’’ ATAC peaks that were only present in control AALEs but disappeared in mutant KRAS AALEs, we found that many of these genes were KZNFs that were significantly downregulated (Appendix 2 Figure 5B). Six of these downregulated KZNFs, ZNF90, ZNF826P, ZNF736, ZNF471, ZNF682, and ZNF853, had peaks unique to control AALEs (Figure 4C). Downregulated KZNF TSS regions were enriched in motifs for ETS (ETV1) and ELK (ELK1) TFs (Appendix 2 Figure 5D), known downstream effectors of the RAS signaling pathway (Simanshu et al., 2017).



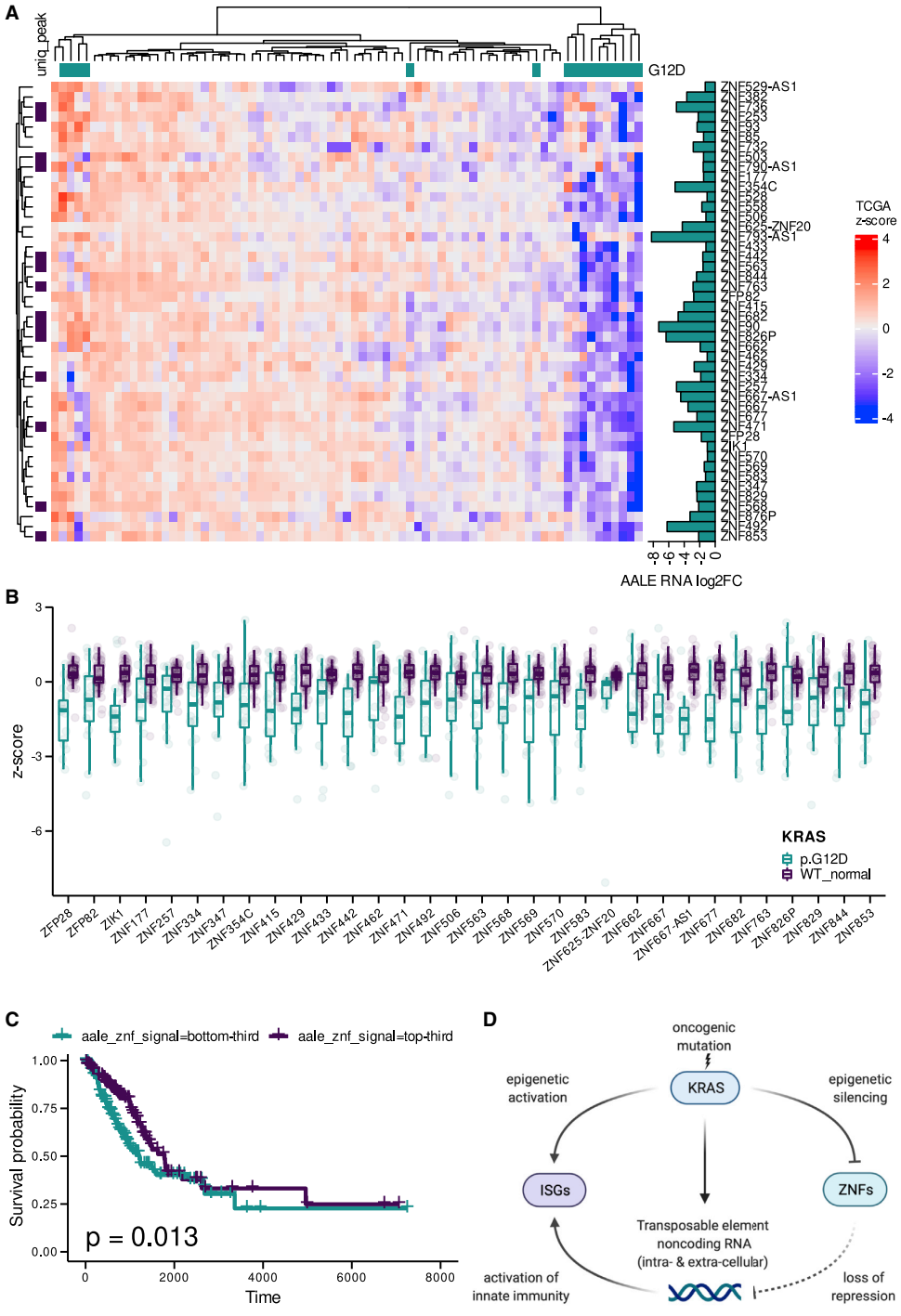
Appendix 2 Figure 5

Appendix 2 Figure 5- Mutant KRAS signaling epigenetically silences KZNFs *in vitro*.

A. Mean ATAC-seq CPM (95% CI) in promoter regions of downregulated KZNFs (<4.5 log2 fold change) in both mutant KRAS and control (CTRL) AALEs. B. Differential expression of KZNFs with unique peaks near TSS (only present in mutant KRAS or control AALEs). C. ATAC-seq coverage in both KRAS and CTRL AALEs for subset of KZNFs with unique peaks detected near TSS. D. Volcano plots of differentially expressed TFs in mutant KRAS AALEs with significant TF motif enrichment in downregulated KZNF gene promoters. chr, chromosome.

2.3.8 Downregulated KZNFs *in vivo* are associated with poor outcomes in lung cancer.

Finally, we explored the clinical significance of the mutant KRAS- induced KZNF silencing we identified in AALE and HBEC cells. Evaluation of KZNF expression in TCGA LUAD RNA-seq data revealed their significant downregulation in mutant KRAS(G12D) samples when compared to WT KRAS lung cancer or matched normal samples, respectively (Appendix 2 Figures 6A and 6B). Furthermore, LUAD samples in the lowest third of KZNF expression demonstrated a significant decrease in overall survival probability (Appendix 2 Figure 6C), highlighting the clinical impact of the mutant KRAS- mediated KZNF downregulation we found in AALE and HBEC cells.



Appendix 2 Figure 6

Appendix 2 Figure 6- Broad downregulation of KZNFs in mutant KRAS LUAD *in vivo*.

A. Hierarchical clustering of expression *Z* scores in TCGA LUAD RNA-seq data for KZNF genes downregulated in mutant KRAS AALEs; KZNFs with unique peaks in their promoter regions in control AALEs are labeled. B. Distribution of *Z* scores for significantly downregulated KZNF genes (Wilcox) in TCGA LUAD RNA-seq data. C. Kaplan-Meier survival curve for patients in the TCGA LUAD dataset stratified into thirds by expression levels of KZNFs downregulated in mutant KRAS AALEs. D. Model of mutant KRAS-mediated regulation of TE RNAs and ISGs by KZNFs. Created with BioRender.com.

2.4 Discussion

Collectively, our findings demonstrate the transcriptomic and epigenomic impact of oncogenic KRAS signaling on TE RNAs and ISGs. Our study suggests that KZNF repression by mutant KRAS signaling leads to de-repression of TE RNAs, triggering an intrinsic ISG response (Figure 5D). This model is supported by broad and significant downregulation of these same KNZFs in mutant KRAS-driven lung adenocarcinomas *in vivo*. Our conclusions are based on deeply sequencing and analyzing the intracellular and extracellular transcriptomes and epigenomes of mutant KRAS-transformed lung cells, building on previous work in which we discovered the coordinate regulation of non- coding RNAs and RAS signaling in the context of epigenomic re- programming (Kim et al., 2015).

The molecular basis for the intrinsic ISG signature we observe in mutant KRAS AALE cells differs from TE RNA-induced IFN responses in cancer cells treated with DNA methyltransferase inhibitors (Enderle et al., 2015; Roulois et al., 2015), as we instead find a prominent role for broad KZNF suppression during the early stages of mutant KRAS-driven cellular transformation. Our studies also suggest that oncogenic KRAS signaling is sufficient to induce at least a subset of the intrinsic ISG signatures that are observed across many cancers and cancer cells lines with ADAR dependencies (Gannon et al., 2018; Liu et al., 2019).

We also present further evidence for the utility of extracellular RNAs in detecting intracellular RNA changes in cancer cells (Reggiardo et al., 2022). Notably, we show the secretion of specific TE RNA and ISG signatures that are aberrantly upregulated in mutant KRAS lung cells. The enrichment of TE-derived noncoding RNAs in extracellular vesicles (Wang et al., 2019) released from mutant KRAS cells highlights their potential utility as RNA biomarkers for diagnosing RAS-driven cancers.

2.5 Methods

2.5.1 Experimental model and subject details.

Immortalized lung epithelial cells (AALE cells; XX), derived at Dana-Farber and immortalized by SV40 large-T antigen (Lundberg et al., 2002) were obtained as a gift from the laboratory of Eric Collison (University of California, San Francisco). The AALE stable cell lines pBABE-mCherry Puro (control) (Lu et al., 2017) and pBABE-FLAG-KRAS(G12D) Zeo (mutant KRAS) were generated using retroviral transduction, followed by selection in puromycin or zeocin. Cells were cultured at 37C and 5% CO_2 in SABM Basal Medium (Lonza SABM basal medium, CC-3119) with supplements and growth factors (Lonza SAGM SingleQuots Kit Suppl. & Growth Factors, CC- 4124).

HBEC3kt cell lines (HBEC cells; XX) were obtained as a gift from the laboratory of Harold Varmus (National Human Genome Research Institute and Weill Cornell Medicine). The HBEC stable cell lines pLenti6/V5-GW/lacZ (control) (Vikis et al.,

2007) and pLenti-KRASV12 (mutant KRAS) were generated using lentiviral transduction, followed by selection in blasticidin. Lentiviral plasmids were obtained as a gift from the laboratory of John Minna (The University of Texas Southwestern Medical Center) (Sato et al., 2013; Vikis et al., 2007). Cells were cultured at 37C and 5% CO_2 in Keratinocyte Serum-Free Media (KSFM) with supplements (Invitrogen, #17005042).

2.5.2 RNA-seq.

For AALE cell lines, bulk RNA was isolated from cells using Quick-RNA MiniPrep kit (Zymogen) and RNA was quantified via NanoDrop-8000 Spectrophotometer. 1ug of total RNA was used as input for the TruSeq Stranded mRNA Sample Prep Kit (Illumina) according to manufacturer protocol. Library quality was determined through the High Sensitivity DNA Kit on a Bioanalyzer 2100 (Agilent Technologies). 6 multiplexed libraries, 3 biological replicates of each condition, were sequenced as HiSeq400 100PE runs.

For HBEC cell lines, cells grown in 10 cm plates (n = 3 per cell line) were washed twice in cold DPBS then collected in Tri-reagent for storage at 80C until the bulk RNA was extracted using Direct-Zol RNA Miniprep Kit (Zymo Research). Concentrations of purified RNA in nuclease-free water were determined by Nanodrop-2000 Spectrophotometer and by Qubit RNA BR Assay (ThermoFisher Scientific). Quality RIN numbers ranging from 9.4-10 were determined by TapeStation 4150 RNA ScreenTape Analysis (Agilent Tech- nologies) before

sending RNA to UC Davis DNA Technologies and Expression Analysis Core Laboratory for poly-A strand specific library preparation to obtain 60 million paired end reads by NovaSeq S4 (PE150) sequencing.

2.5.3 ATAC-seq.

100,000 AALE cells were collected and centrifuged at 500xg for 5 min at 4C. Pellets were washed with ice-cold PBS and centrifuged. Pellets were resuspended in ice-cold lysis buffer. Tagmentation reaction and purification were conducted according to manufac- turer's protocol (Active Motif). 2 Libraries, one from each condition, were sequenced on a NextSeq500 as 2 3 75 paired end reads.

2.5.4 Extracellular RNA-seq.

The exoRNeasy serum/plasma maxi kit (Qiagen) was used to isolate extracellular vesicles, which were quantified using Nanoparticle Tracking Analysis (Malvern, UK). 30 mL of AALE cell culture supernatant was filtered to remove particles larger than 0.8 um. The filtrate was precipitated with kit buffer and filtered through a column to collect extracellular vesicles. These vesicles were then lysed with QIAzol lysis reagent. Total RNA was isolated using the indicated phase separation method and used to make 6 libraries, 3 bio- logical replicates for each condition, for RNA-seq using the Smart Seq HT low input mRNA library prep kit (Takara). Libraries were sequenced on an Illumina NextSeq500.

2.5.5 RNA-seq Analysis

All *fastq* files were trimmed with *Trimmomatic 2 (0.38)* (Bolger et al., 2014) using the Illumina NextSeq PE adapters. The resulting trimmed files were assessed with *FastQC* (Brown et al., 2017)(Brown et al., 2017) and then processed with the following analytical pipeline:

Salmon (1.3.0)

Pseudoalignment of RNA-seq reads performed with *Salmon* (Patro et al., 2017) using the following arguments: –validateMappings –gcBias –seqBias –recoverOrphans –rangeFactorizationBins 4 using an index created from the *GENCODE* version 35 (Frankish et al., 2021) transcriptome fasta file using decoy sequences to enable selective alignment. An additional, TE- aware index was created in a similar fashion but supplemented with sequences generated from the UCSC Repeat Masker track.

DESeq2 (1.32.0)

Salmon output was imported into a DESeq object using *tximeta* (Love et al., 2020; Soneson et al., 2015)(Love et al., 2014; Soneson et al., 2016) and differential expression analysis was performed with standard arguments (Love et al., 2014; Zhu et al., 2019). All results were filtered to have padj <0.05. In the case where R could only generate 0.00 for the padj values, they were reset to the lowest non-zero padj value in the dataset. Where count data was used, it was normalized across samples using DESeq.

Principal component analysis

PCA was performed in *R* using the function prcomp provided by the package *stats* (*4.1.1)*. Input gene abundance data was first variance stabilized using DESeq2 and then filtered for genes with 0 standard deviation across the samples.

Motif discovery and enrichment analysis

All motif-based analysis was performed in *R* using packages *memes (1.1.4)*, *universalmotif (1.10.2)*, *BSgenome.Hsapien- s.UCSC.hg38 (1.4.3)*, *MotifDBGenomicRanges (1.44.0)* and *MotifDB* (1.34.0) (Lawrence et al., 2013; Nystrom and McKay, 2021). Enriched motifs were identified by using the runAME() function provided by *memes* with a control set to 'shuffle' the input sequences unless otherwise noted in the text. Individual motif occurrences were identified with the runFimo() function pro- vided by *memes*.

Zinc finger gene analysis

ChIP-exo data and supplementary information were extracted from supplementary data provided by Imbeault et al. (Imbeault et al., 2017). ZNF genes were cross referenced with *DESeq2* and bed file of Repeat Masked TE inserts from the UCSC Genome Browser to extract relevant differential expression data of ZNF proteins and Transposable Element transcripts using *R*. Promoter and motif analyses performed as described above.

Motif discovery was intersected with repeat-masked insertions and cross referenced with ChIP-exo target data to identify potential regulatory targets of

differentially expressed KZNFs. KZNF targets were ranked by the score provided. Additional ZNF binding motifs were acquired from Barazandeh et al.'s website (Key resources table) and converted to a database compatible with MEME suite (Bailey et al., 2015; Barazandeh et al., 2018).

Gene set enrichment analysis

Differentially expressed genes were ranked by the shrunken log2FoldChange values generated by *DESeq2*. Gene sets were acquired using the *R* package *msigdbr (7.4.1)* (Dolgalev, 2020) and filtered to only contain gene sets with 'Hallmark' status.

The *R* package *fgsea (1.18.0)* (Korotkevich et al., 2016) was used to generate Gene Set Enrichment (Liberzon et al., 2011; Subramanian et al., 2005) estimates which were filtered to results with adjusted pvalues <0.05.

GREAT gene ontology analysis

The *R* package *rGREAT (1.24.0)* (Bioconductor) was used to process ATAC-seq identified peaks with GREAT and identify enriched GO terms. ATAC-seq peaks unique to either CTRL or KRAS contexts were used as input with the background set to the entire peak library comprised from both contexts.

TCGA ZNF analysis

TCGA-LUAD phenotype and normalized count data were downloaded from the UCSC Xena browser TOIL data repository (Key resources table)

(Goldman et al., 2020). The files were combined, and patients were grouped by their KRAS mutation status and identity. Heatmaps and associated hierarchical clustering were performed in *R* using the package *ComplexHeatmap* (2.8.0) (Gu et al., 2016). Survival analysis was performed using the *survival* R package (3.3) (Therneau, 2022)(Therneau and others, 2015).

ATAC-seq analysis

The nf-core ATAC-seq pipeline was used to process ATAC-seq reads to alignments with BWA, narrow peak calls with MACS2, and ultimately annotated peaks. Read count analysis was performed with the *R* package *bamsignals* (1.24.0) (Bioconductor) using the sorted bam files produced by the nf-core pipeline.

2.5.6 Quantification and statistical analysis.

All quantitative data for functional assays have been reported as means ± standard deviation. Statistical significance for these were calculated using a Wilcox-test (*R – wilcox.test()*) unless otherwise noted and p values <0.05 were considered significant. All statistical analyses were performed with R (version 4.1.1) running from the Rocker 'Tidyverse' Docker container (rocker/tidyverse:4.1.1). Linear regression was carried out with the lm() function.

Supplemental information for Appendix 2

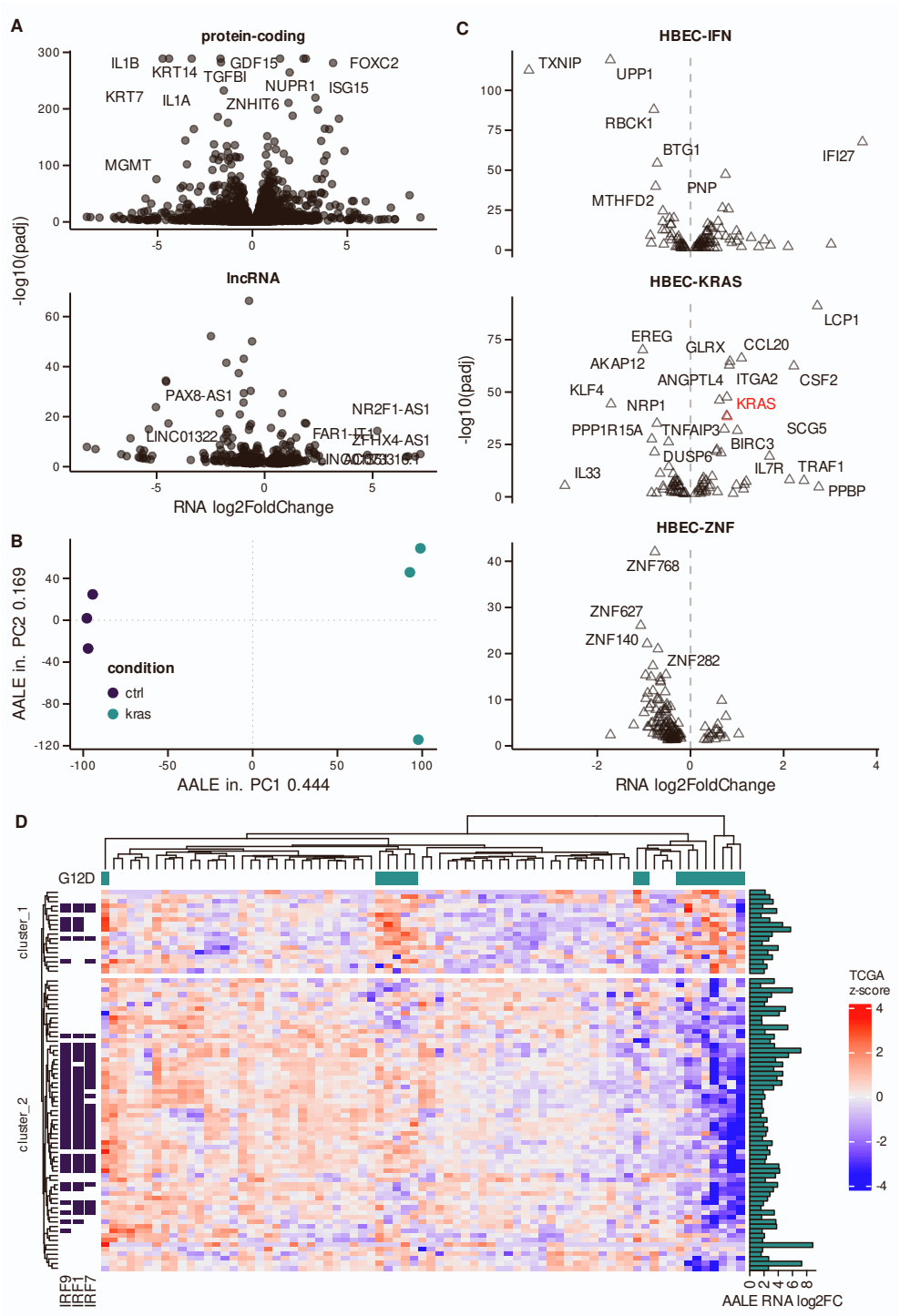


Figure S1. Mutant KRAS signaling induces global transcriptional changes. Related to Figure 1.

A. Volcano plots of differentially expressed protein coding genes and lncRNAs in mutantKRAS AALEs. B. Principal component analysis (PCA) of control (ctrl) and

mutant KRAS (kras) AALE RNA-seq libraries. Related to variance explained by each PC displayed with axis label. C. Volcano plots of differentially expressed protein coding genes and lncRNAs in mutant KRAS HBECs. D. Hierarchical clustering of expression z-scores in TCGA LUAD RNA-seq data for genes upregulated in mutant KRAS AALEs. Related to genes with IRFbinding motifs in their promoter regions are labeled.

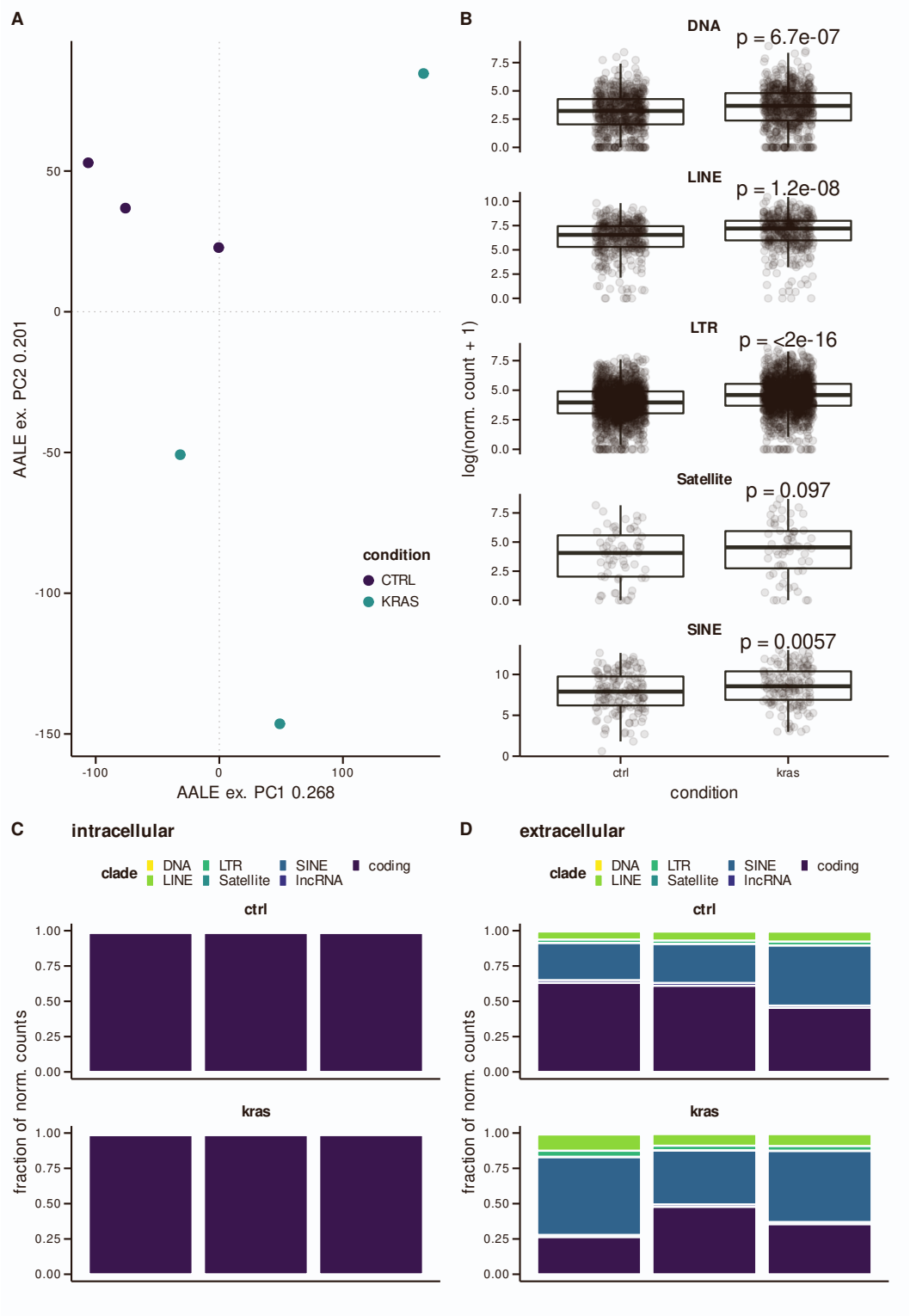

A
AALE ex. PC2 0.201
AALE ex. PC1 0.268
condition
CTRL
KRAS
B
DNA
p = 6.7e-07
LINE
p = 1.2e-08
LTR
p = <2e-16
Satellite
p = 0.097
SINE
p = 0.0057
log(norm. count + 1)
ctrl
kras
condition
C
intracellular
D
extracellular
clade
DNA
LINE
LTR
Satellite
SINE
lncRNA
coding
ctrl
kras
fraction of norm. counts

Figure S2. Mutant KRAS signaling significantly alters the RNA composition of
secreted extracellular vesicles. Related to Figure 3.

A. Principal component analysis (PCA) of control (CTRL) and mutant KRAS (KRAS) AALE extracellular (ex) RNA-seq libraries. Related to variance explained by each PC displayed with axis label. B. Distribution of insertion-level abundance for TE clades in extracellular RNA-seq libraries (Wilcoxon). C, D. Distribution of counts assigned to GENCODE coding, lncRNA, and TE clades in intracellular and extracellular RNA-seq libraries.

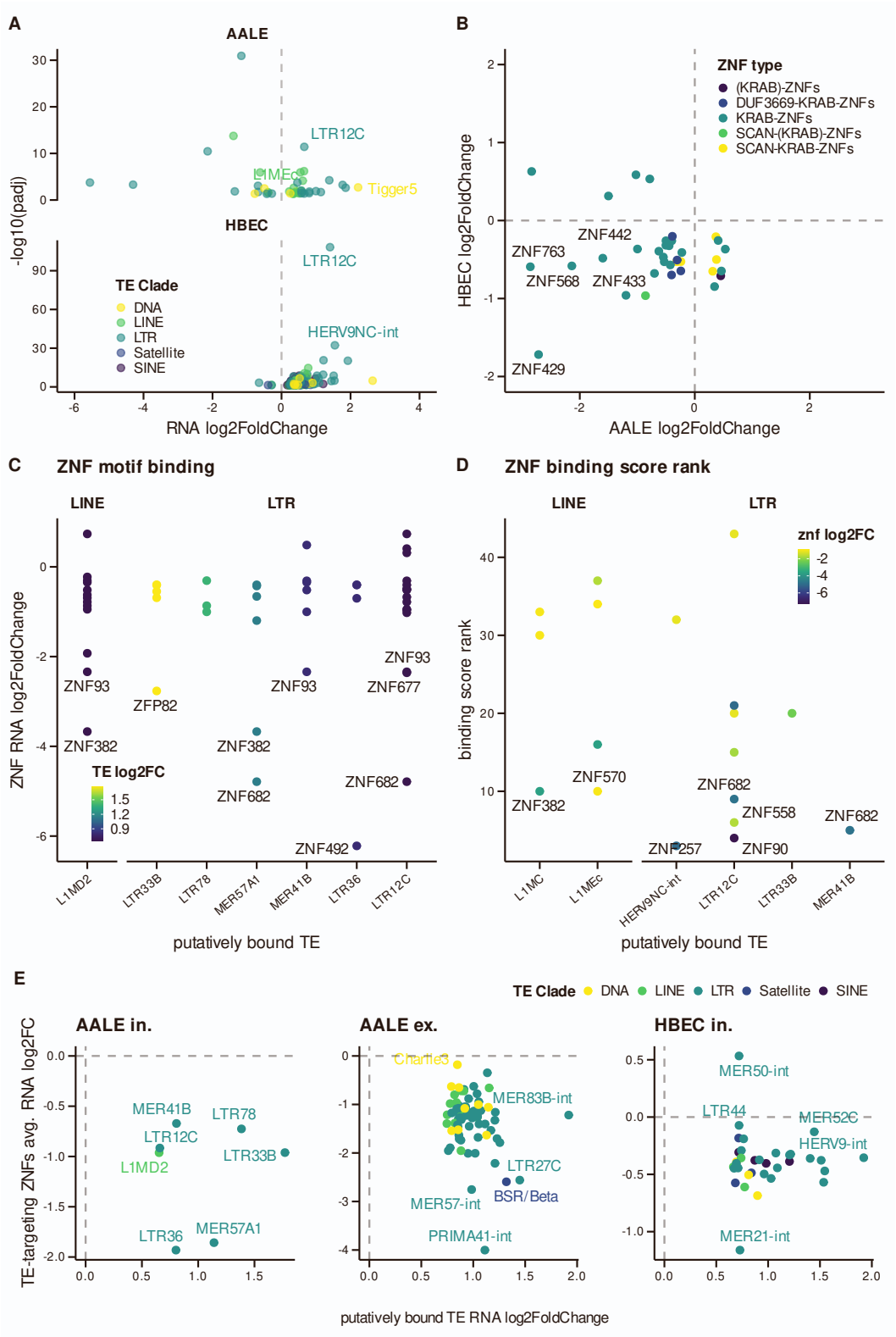

A
AALE
-log10(padj)
LTR12C
L1MEc
Tigger5
HBEC
LTR12C
HERV9NC-int
TE Clade
DNA
LINE
LTR
Satellite
SINE
RNA log2FoldChange
B
ZNF type
(KRAB)-ZNFs
DUF3669-KRAB-ZNFs
KRAB-ZNFs
SCAN-(KRAB)-ZNFs
SCAN-KRAB-ZNFs
HBEC log2FoldChange
ZNF442
ZNF763
ZNF568
ZNF433
ZNF429
AALE log2FoldChange
C
ZNF motif binding
LINE
LTR
ZNF RNA log2FoldChange
ZNF93
ZFP82
ZNF382
ZNF93
ZNF382
ZNF682
ZNF93
ZNF677
ZNF682
ZNF492
TE log2FC
1.5
1.2
0.9
L1MD2
LTR33B
LTR78
MER57A1
MER41B
LTR36
LTR12C
putatively bound TE
D
ZNF binding score rank
LINE
LTR
znf log2FC
-2
-4
-6
binding score rank
ZNF382
ZNF570
ZNF682
ZNF558
ZNF257
ZNF90
ZNF682
L1MC
L1MEc
HERV9NC-int
LTR12C
LTR33B
MER41B
putatively bound TE
E
TE Clade DNA LINE LTR Satellite SINE
AALE in.
AALE ex.
HBEC in.
TE-targeting ZNFs avg. RNA log2FC
MER41B
LTR78
LTR12C
LTR33B
L1MD2
LTR36
MER57A1
Charlie3
MER83B-int
LTR27C
MER57-int
BSR/Beta
PRIMA41-int
MER50-int
LTR44
MER52C
HERV9-int
MER21-int
putatively bound TE RNA log2FoldChange

Figure S3. TE RNAs activated by mutant KRAS are enriched for KZNF motifs. Related to Figures 3 & 4.

A. Volcano plots of differentially expressed TE RNAs in mutant KRAS AALEs and HBECs. B. Comparison of KZNF gene differential expression in mutant KRAS AALEs and HBECs. C. Differential expression of KZNFs with binding motifs in mutant KRAS-activated TE RNAs in AALEs. D. Ranking of KZNFs by binding score for mutant KRAS-activated TE RNAs in AALEs. E. Comparison of TE differential expression (x axis) to the average expression of ZNFs with putative binding sites within the TE based on motif library generated by Hughes et al.

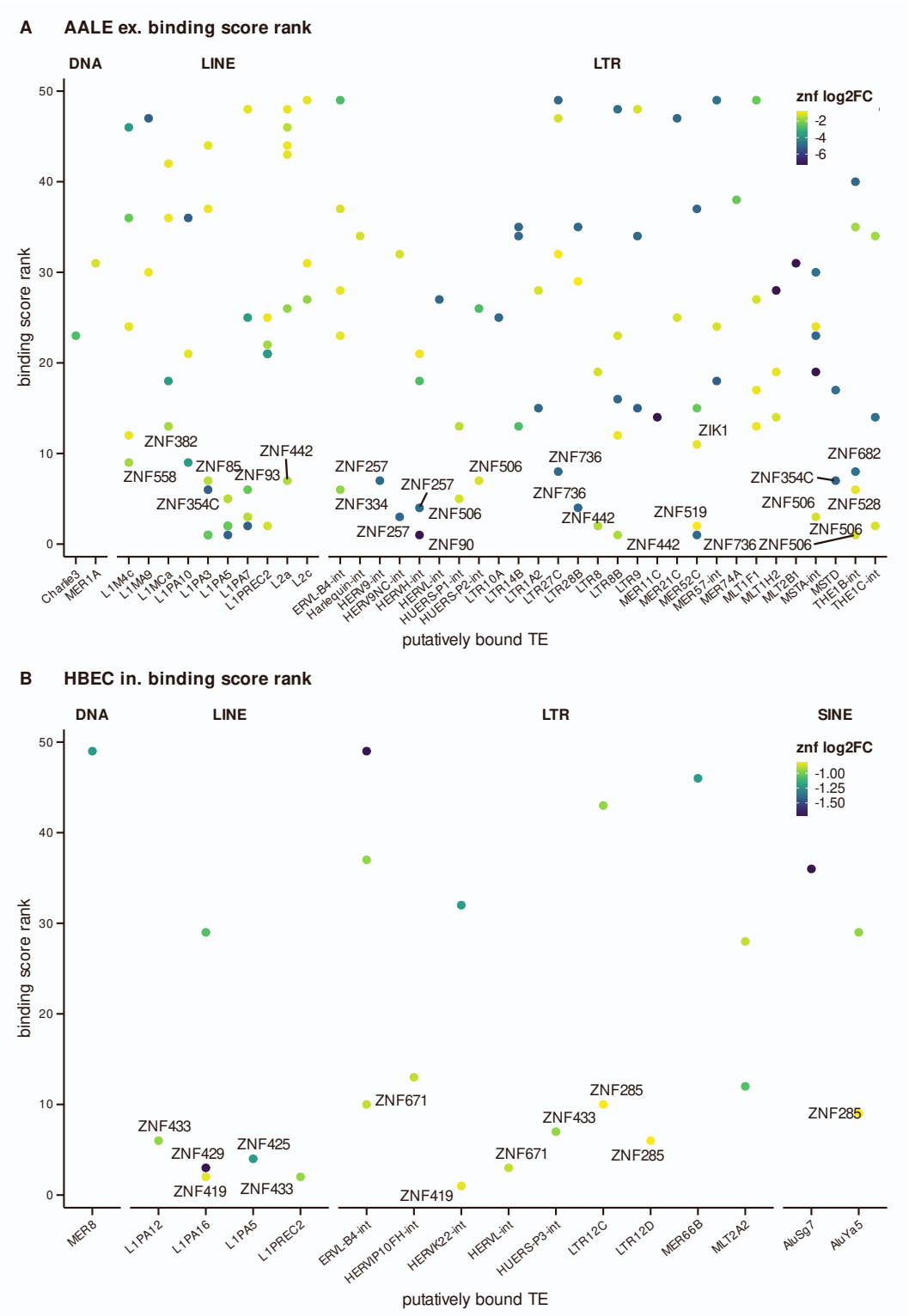

A
AALE ex. binding score rank
DNA
LINE
LTR
znf log2FC
-2
-4
-6
binding score rank
ZNF382
ZNF558
ZNF85
ZNF93
ZNF442
ZNF354C
ZNF257
ZNF334
ZNF257
ZNF506
ZNF257
ZNF90
ZNF506
ZNF736
ZNF736
ZNF442
ZIK1
ZNF519
ZNF442
ZNF736
ZNF506
ZNF682
ZNF354C
ZNF506
ZNF528
ZNF506
Charlie3
MER1A
L1M4c
L1MA9
L1MCa
L1PA10
L1PA3
L1PA5
L1PA7
L1PREC2
L2a
L2c
ERVL-B4-int
Harlequin-int
HERV9-int
HERV9NC-int
HERVH-int
HERVL-int
HUERS-P1-int
HUERS-P2-int
LTR10A
LTR14B
LTR1A2
LTR27C
LTR28B
LTR8
LTR8B
LTR9
MER11C
MER21C
MER52C
MER57-int
MER74A
MLT1F1
MLT1H2
MLT2B1
MSTA-int
MSTD
THE1B-int
THE1C-int
putatively bound TE
B
HBEC in. binding score rank
DNA
LINE
LTR
SINE
znf log2FC
-1.00
-1.25
-1.50
binding score rank
ZNF433
ZNF429
ZNF425
ZNF419
ZNF433
ZNF671
ZNF419
ZNF285
ZNF433
ZNF671
ZNF285
ZNF285
MER8
L1PA12
L1PA16
L1PA5
L1PREC2
ERVL-B4-int
HERVIP10FH-int
HERVK22-int
HERVL-int
HUERS-P3-int
LTR12C
LTR12D
MER66B
MLT2A2
AluSg7
AluYa5
putatively bound TE

Figure S4. KZNF binding scores highlight potential regulation of differentially expressed TEs in mutant KRAS AALE EVs and HBEC cells. Related to Figures 3 & 4.

A,B. Ranking of KZNFs by binding score for each upregulated TE RNA in mutant KRAS extracellular (ex) AALE and intracellular (in) HBEC RNA-seq data.

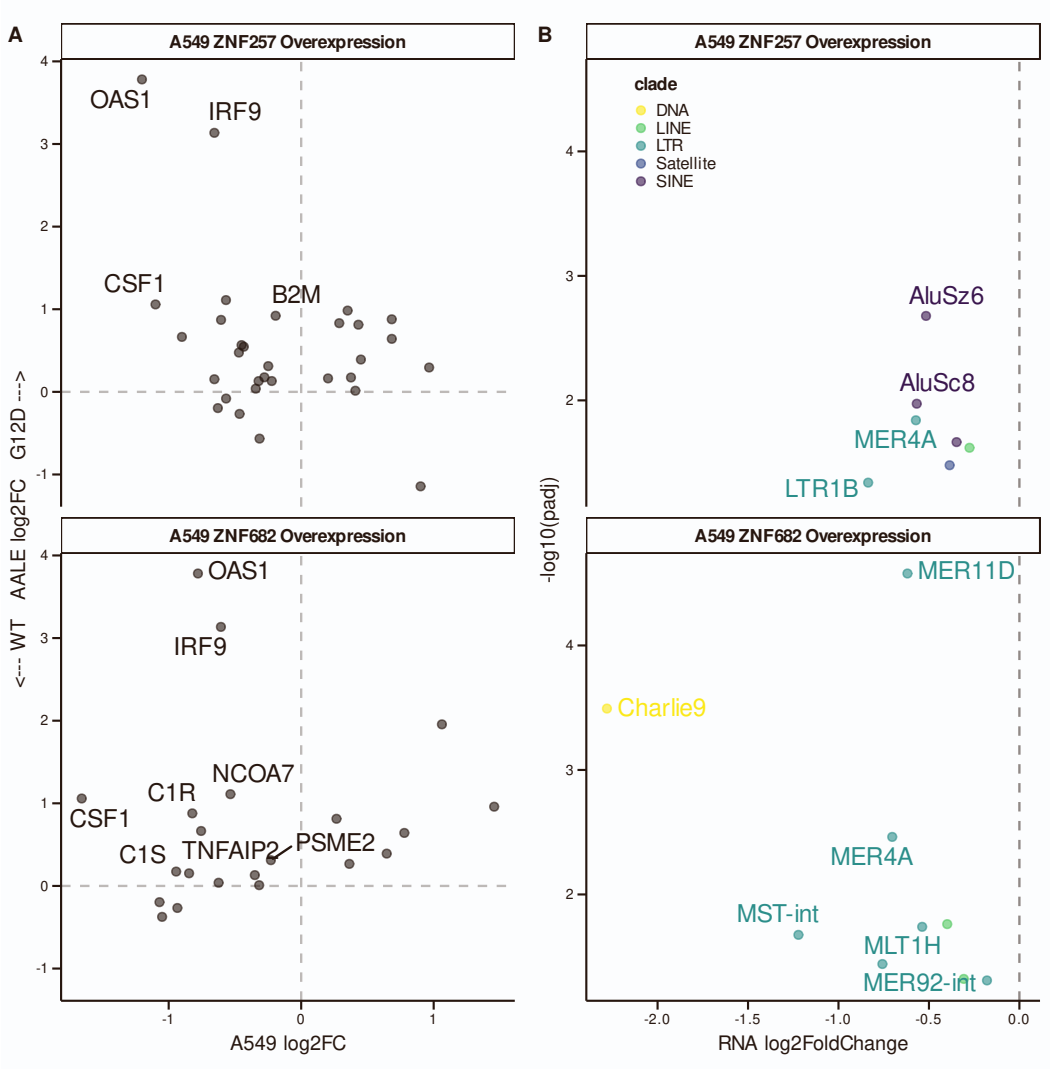

A
A549 ZNF257 Overexpression
OAS1
IRF9
CSF1
B2M
A549 ZNF682 Overexpression
OAS1
IRF9
NCOA7
C1R
CSF1
C1S
TNFAIP2
PSME2
<--- WT AALE log2FC G12D --->
A549 log2FC
B
A549 ZNF257 Overexpression
clade
DNA
LINE
LTR
Satellite
SINE
AluSz6
AluSc8
MER4A
LTR1B
A549 ZNF682 Overexpression
MER11D
Charlie9
MER4A
MST-int
MLT1H
MER92-int
-log10(padj)
RNA log2FoldChange

Figure S5. Mutant KRAS-regulated KZNFs repress ISGs and TE RNAs. Related to Figures 1, 2, 3 & 4.

A. Scatter plots of differentially expressed genes between mutant KRAS A549 lung cancer cells overexpressing ZNF257 or ZNF682 and mutant KRAS AALEs. B. Volcano plots of differentially expressed TE RNAs in mutant KRAS A549 lung cancer cells overexpressing ZNF257 or ZNF682.

APPENDIX 3- High-Throughput CIRSPR Screening Identifies Genes Involved in Macrophage Viability and Inflammatory Pathways

Sergio Covarrubias, Apple Cortez Vollmers, Allyson Capili, Michael Boettcher, Aaron Shulkin, Michele Ramos Correa, Haley Halasz, Elektra K. Robinson, Laura O'Briain, Christopher Vollmers, James Blau, Sol Katzman, Michael T. McManus, and Susan Carpenter

My Contributions to the following manuscript:

This work focusses on identifying genes involved in macrophage cell viability. Here we performed a genome-wide CRIPSPR-Cas9 screen targeting protein coding genes in bone-marrow-derived macrophages. Following the screen, I worked to validate the top significant hits as regulators of macrophage viability. I cloned the top sgRNAs from the screen targeting the top significant hits, created a stable knockdown cell line for each target and performed a mixed cell assay for each of these top candidates. From the work that I did, I generated Figure 3.1F of this appendix.

Acknowledgments

Thanks to all members of the Carpenter Lab and to Dr. Kate Fitzgerald and Dr. Maninjay Atianand for feedback on manuscript. This work was funded by NIH grants R03 AI131019-01 and R21 AI142165 to S. Carpenter and U01 CA217882

to M.T.M. A.C.V. was supported by an NIH Predoctoral Training grant (T32 GM008646), a Ford Predoctoral Fellowship, and a Howard Hughes Medical Institute Gilliam fellowship. E.K.R. was supported by an NIH predoctoral training grant (T32 GM008646).

Author Contributions

S. Covarrubias and S. Carpenter conceptualized and designed the research study. S. Covarrubias and A.C.V. performed the screen. M.B., J.B., and M.T.M. provided the custom whole-genome sgRNA library. A.C., H.H., E.K.R., A.C.V., M.R.C., A.S., and L.O. performed candidate validation experiments. C.V. developed data analysis tools. S.K. performed MAGeCK analysis. S. Covarrubias, A.C.V., and S. Carpenter analyzed and interpreted data and wrote the paper.

3.1 Abstract

Macrophages are critical effector cells of the immune system, and understanding genes involved in their viability and function is essential for gaining insights into immune system dysregulation during disease. We use a high-throughput, pooled-based CRISPR-Cas screening approach to identify essential genes required for macrophage viability. In addition, we target 3' UTRs to gain insights into previously unidentified cis-regulatory regions that control these essential genes. Next, using our recently generated nuclear factor kB (NF-kB) reporter line, we perform a fluorescence-activated cell sorting (FACS)-based high-

throughput genetic screen and discover a number of previously unidentified positive and negative regulators of the NF-kB pathway. We unravel complexities of the TNF signaling cascade, showing that it can function in an autocrine manner in mac- rophages to negatively regulate the pathway. Utilizing a single complex library design, we are capable of interrogating various aspects of macrophage biology, thus generating a resource for future studies.

3.2 Introduction

Macrophages are critical cells of the innate immune system providing one of the first lines of defense against invading microbes. Macrophages arise from precursor monocyte cells that constitute ~10%–20% of the immune cells found in the blood (Verhoeckx et al., 2015). Upon encountering a danger signal, monocytes differentiate into macrophages and rapidly move to the site of infection. Important aspects of macrophage function include their ability to proliferate and migrate, as well as their ability to induce the inflammatory program to aid in clearing infections and initiate tissue repair to maintain homeostasis (Wynn et al., 2013). While much work has been performed to understand the contribution of individual proteins to the processes that control macrophage biology, there has been no systematic approach adopted to study genes involved in macrophage viability and function simultaneously in a high-throughput manner.

Clustered regularly interspaced short palindromic repeat (CRISPR) technology has revolutionized the field of functional genomics, providing an easy-to-use method for disrupting specific genes (Knott and Doudna, 2018). The coupling of CRISPR technology with pooled single-guide RNA (sgRNA) screening al- lows simultaneous knockout of thousands of individual genes in a large population of cells (Shalem et al., 2014; Wang et al., 2014), enabling unbiased reconstruction of biological pathways. Numerous CRISPR screens have probed pathways ranging from cell viability (Tzelepis et al., 2016; Wang et al., 2015) to virus infection (Han et al., 2018; Park et al., 2017), supporting the use of this system for exploring a wide-range of biology. More recently, in macrophages, CRISPR screens have been performed to identify novel regulators of infection and inflammation. (Schmid-Burgk et al., 2016) carried out a genome-wide CRISPR screen to uncover previously unidentified regulators of the NLRP3 inflammasome. They found that knockout of NEK7 rescued macrophages from lethality and was associated with activation of the NLRP3 inflammasome. (Yeung et al., 2019) performed a screen in macrophages to identify regulators of *Salmonella* infection. They identified NHLRC2, showing that it can play a role both in *Salmonella* infection as well as macrophage differentiation. Interestingly, many of the hits identified in their screen are within pathways with known chemical inhibitors providing avenues for future therapeutic targeting. Finally, a recent screen was conducted to identify regulators of *Shigella* infection in macrophages (Lai et al., 2020). Lai et al. identified host factors modulated by *Shigella flexniri* infection.

They could show that inhibiting acetyl-coenzyme A (CoA) production caused by the infection is beneficial to the function of macrophages and limits the infection. In addition, knockout of the TLR1/2 pathway reduced inflammation, enhanced macrophage survival, and limited infection (Lai et al., 2020).

Whereas these screens that have been performed, to date, in macrophages are focused on a single readout, we have utilized pooled screening to address three questions simultaneously: (1) What genes are required for macrophage survival and proliferation? (2) How are those genes regulated? (3) What genes contribute to the downstream inflammatory signaling processes?

Our screens provide a resource that identifies genes essential to macrophage viability and provide insights into potential *cis*-elements within the 3^0 UTRs of these genes that may reveal important means of regulation. Lastly, we identify previously unidentified positive and negative regulators of nuclear factor kB (NF-kB) inflammatory signaling. Unexpectedly, our screen uncovers a role for tumor necrosis factor (TNF) as a negative regulator of NF-kB and shows that this is functioning in an autocrine manner in macrophages. In a single screen, we bring together decades of literature on the complex regulation of TNF. Here, we demonstrate the power of CRISPR pooled screening to identify a plethora of genes with varied and critical roles in macrophage biology.

3.3 Results

3.3.1 Pooled CRISPR Screen Identifies Macrophage-Specific Genes Involved in Viability

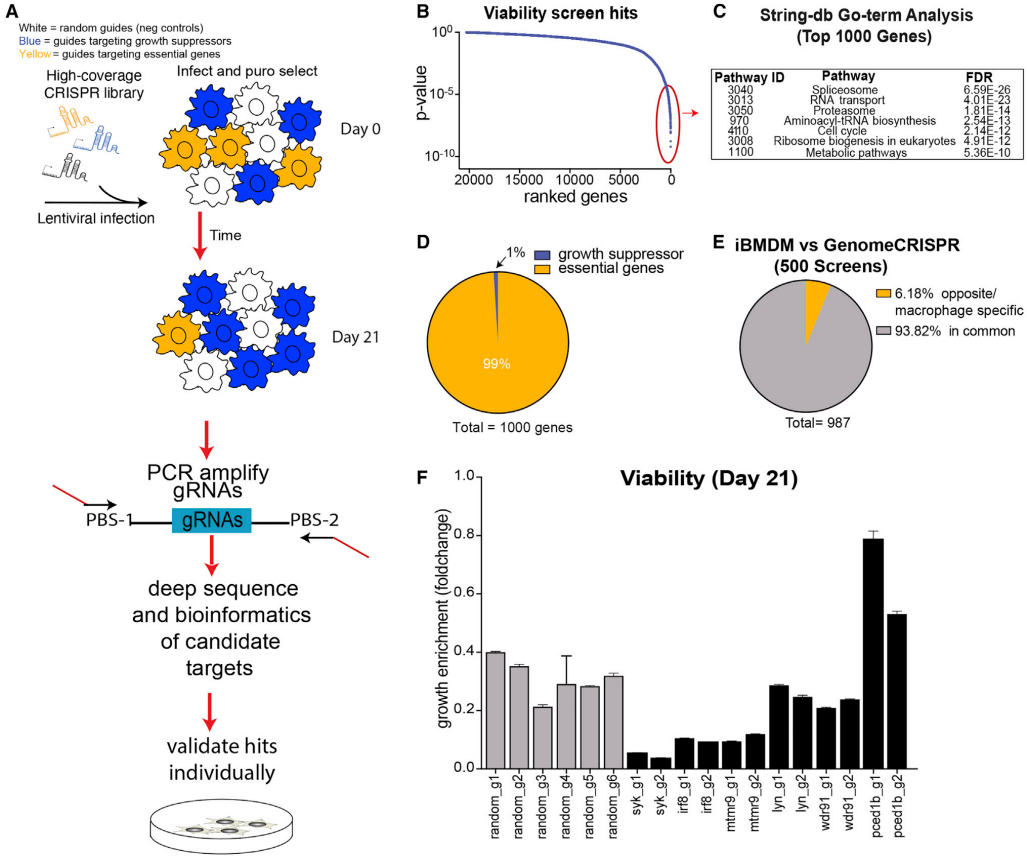
To define all genes essential for macrophage survival, immortalized bone-marrow-derived macrophage (iBMDM)-Cas9 cells were transduced (MOI = 0.3) with pooled lentivirus generated from our custom whole-genome sgRNA library containing ~270K individual sgRNAs targeting all RefSeq annotated coding genes and ~500 microRNAs (miRNAs; 12 guides per gene), along with ~5K non-targeting controls (Appendix 3 Figures 1A and S1A; Tables S1 and S2). Cells were maintained at >1,000 cells per sgRNA throughout the screen. Cells were cultured for 21 days collecting genomic DNA from cells at day 0 and day 21 (Appendix 3 Figure 1A). The libraries were prepared as described previously (Boettcher et al., 2019). Using the Mann-Whitney (MW) U test, we compared the sgRNA repertoire from day 21 to that from day 0 and identified significant genes (Appendix 3 Figure 1B; Tables S3 and S4). We identified expected viability-related genes with roles in spliceosome, proteasome, and cell-cycle functions (Appendix 3 Figure 1C). The majority of the top significant hits were genes essential for viability, while only 1% of genes were growth suppressors (Figure 1D), consistent with previous findings (Gilbert et al., 2014). We also analyzed the data using the model-based analysis of genome-wide CRISPR-Cas9 knockout (MAGeCK) analysis pipeline (Tables S5

and S6). We obtained a strong overlap (88% of genes) using both MW and MAGeCK analyses and no significant difference in the identified top hits (Appendix 3 Figure S1B); however, the MW U test identifies a much larger set of hits, possibly due to its sensitivity and, therefore, identifying more true positives.

Our library also contains within it random barcode sequences associated with each sgRNA (Boettcher et al., 2019). Each sgRNA is associated with ~50 different barcodes. By analyzing the data with the barcodes, it enabled us to generate in-sample replicates providing greater statistical power to identify significant hits (Figure S1C; Tables S7 and S8). The barcodes were assigned to four bins, providing four in-sample replicates that were then used as the input into MAGeCK analysis. Figure S1C compares the single-sample analysis that identified 417 hits (false discovery rate [FDR] < 0.05) to the in-sample replicate analysis that identified 609 hits (FDR < 0.05), showing the power of utilizing the binning approach to increase statistical power.

We compared our viability screen hits to the GenomeCRISPR (http://genomecrispr.dkfz.de/) database, which includes a collection of ~500 CRISPR viability screens performed in ~421 cell types. Over 93% of genes from our screen overlapped those from the database (shown in gray, Appendix 3 Figure 1E), suggesting that these are genes critical for a variety of biological processes common to all cell types. Interestingly, ~6% of genes identified showed the opposite MW *Z*-score phenotype in our screen compared to the database,

suggesting that these genes have unique cell-type-specific functions in macrophages (shown in yellow, Appendix 3 Figure 1E). Furthermore, these macrophage-specific essential hits included genes involved in NF-kB signaling (Appendix 3 Figure S2), suggesting that these genes evolved functions involving not only survival but also inflammatory activation, a critical component of macrophage function. We individually cloned six candidate macrophage-specific essential guides and validated the phenotype using a mix-cell proliferation assay (Appendix 3 Figure 1F). The mix-cell assay involved combining cherry-positive cells (containing sgRNAs) with cherry-negative cells at a 1:1 ratio and monitoring cell growth over time as assessed by changes in the ratio of cherry-positive to cherry-negative cells. As shown in Figure 1F, we confirmed four of the six selected candidate genes. We showed that tyrosine-protein kinase (Syk), Interferon regulatory factor 8 (Irf8), and myotubularin-related protein 9 (Mtmr9) are macrophage-specific essential genes, as targeting the coding sequence of these genes resulted in decreased fitness. While PC-esterase domain containing 1B (Pced1b) is a growth suppressor in macrophages, knocking it out resulted in increased fitness of the cells.



Appendix 3 Figure 1

Appendix 3 Figure 1- Screen Identifies Macrophage-Specific Genes Involved in Viability A. Cas9-expressing iBMDM cells were infected with a whole-genome library targeting all RefSeq annotated coding genes (Table S1). Two days post-infection, cells were harvested for an initial day-0 time point and then again after 21 days in culture. B. The MW U test was performed comparing 12 sgRNAs targeting each gene to the nontargeting controls for samples collected at both day 21 and day 0. Genes are displayed ranked by significance. C. GO-term analysis was performed on the top 1,000 significant hits using STRINGdb. D. We determined the number of essential genes (negative MW *Z* score) and total number of growth suppressor genes (positive MW *Z* score) and plotted them as a fraction of the total (total = 1,000 genes). E. Viability screen hits from our screen were compared to those from GenomeCRISPR, a collection of ~500 CRISPR screens (http://genomecrispr.dkfz.de/). Genes ‘‘in common’’ as well as genes showing ‘‘opposite/macrophage-specific’’ phenotypes are displayed. F. Cas9-expressing iBMDM cells were infected with sgRNAs targeting selected macrophage-specific

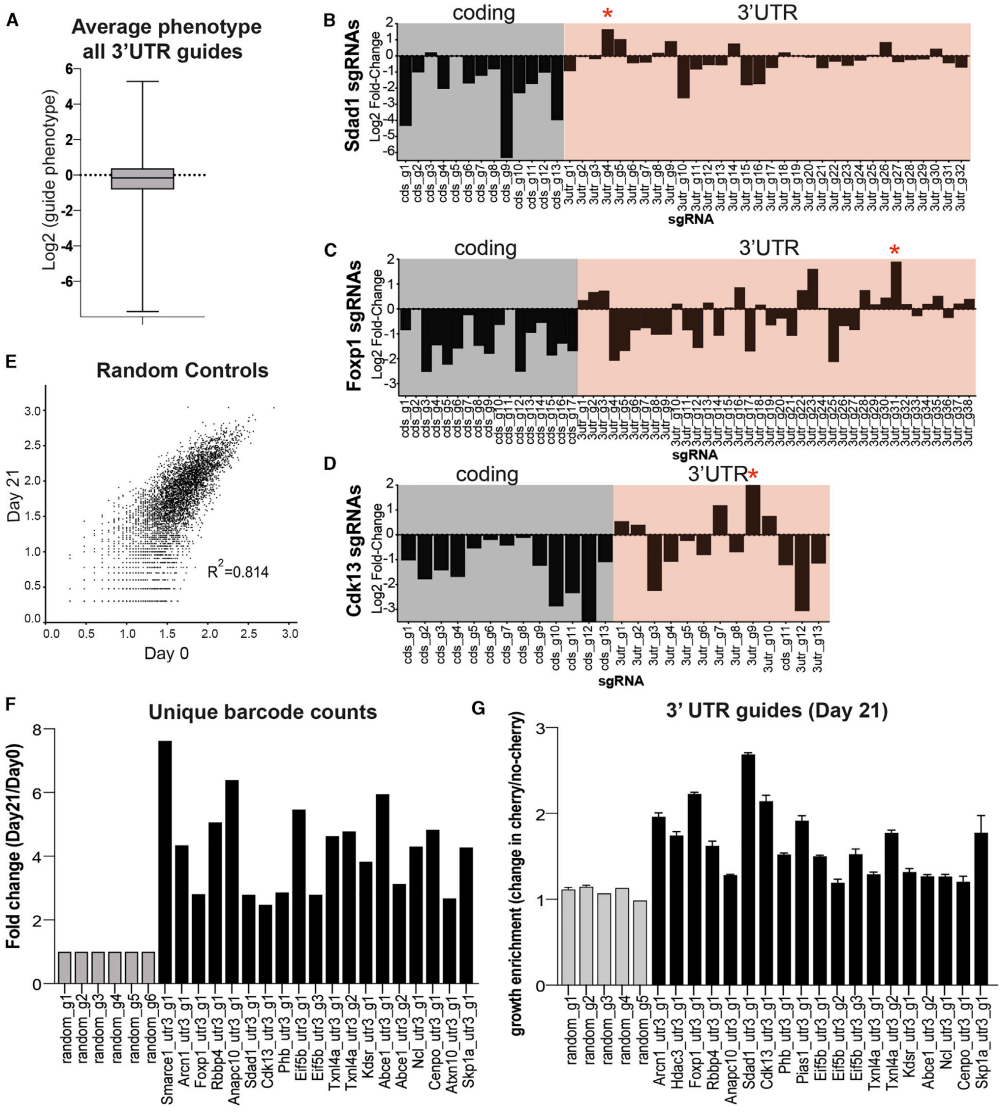
viability genes. We combined cherry-positive cells (containing sgRNAs) with unedited cherry-negative cells at a 1:1 ratio and monitored growth of sgRNA-infected cells (cherry) relative to uninfected reference cells in a mix-cell growth assay for 21 days. Experiment was repeated 2 times, and a representative experiment is displayed.

3.3.2 CRISPR Targeting The 3' UTRs of Essential Genes Identifies Cis-Regulatory Elements

Untranslated regions (UTRs) offer a critical source of regulation for messages through specific *cis*-elements, which can bind miRNAs and/or proteins to regulate pathways including RNA decay and translation (Mayr, 2017). To probe for novel *cis*-elements within 3' UTRs, we specifically targeted the 3' UTR within known essential genes. For these essential genes, we expected that guides targeting coding exons would cause a decrease in fitness. In contrast, guides targeting 3' UTR *cis*-elements that result in an increase in fitness will represent regions containing regulatory *cis*-elements with negative roles on gene expression (miRNA binding sites, adenylate-uridylate-rich elements [AU- rich elements] affecting stability, etc.). We assessed the phenotypes of 3' UTR-targeting sgRNAs and found an overall neutral average phenotype for these guides (Appendix 3 Figure 2A), suggesting that the majority of the sites we targeted did not contain any *cis*-regulatory elements. However, a subset of these 3' UTR-targeting guides demonstrated phenotypes of >3-fold change in both directions, suggestive of sites that contain regulatory elements that could improve or decrease fitness of the cells (Appendix 3 Figures 2B–2D; Table S9). We focused on 3' UTR guides that showed positive (>3-fold) enrichment for genes whose coding-targeting guides demonstrated significant negative enrichment (Figures 2B–2D, red stars). We

reasoned that these 3' UTR-targeting guides may disrupt significant *cis*-elements that help explain the opposing phenotype we observed (Appendix 3 Figures 2B–2D, red stars). One possible concern with this analysis is that it is limited to a single sgRNA targeting the sites within the 3' UTR (compared to 10–12 within the screen for each coding gene target), which could lead to false positives. To mitigate this, we made use of the internal barcodes associated with each individual sgRNA. Identifying a sgRNA hit associated with multiple individual barcodes reduces the possibility of the hit being a false positive. In one example, the viability phenotype could be caused by integration-specific effects of an sgRNA on neighboring genes. As a control, we analyzed all barcodes associated with all random control sgRNAs, comparing the day-0 to day-21 viability screens, and can show high correlation, with an R^2 of 0.814. This shows that the distribution of reads for barcodes of random control genes do not change across the course of the experiment, making counting barcodes a valid approach (Appendix 3 Figure 2E). Therefore, we counted all unique barcodes associated with sgRNAs in selected genes or random controls and compared the reads associated with these genes on day 21 to those on day 0. As expected, there is no difference in the read count from day 21 to day 0 for the random controls (shown in gray, Appendix 3 Figure 2F). In contrast, we see positive enrichment for genes (shown in black, Appendix 3 Figure 2F) whose coding-targeting guides demonstrated the opposite effect, which confirmed our previous observations in Appendix 3 Figures 2B–2D.

We individually cloned selected candidate 3' UTR guides and validated the phenotype using a mix-cell proliferation assay as described previously (Appendix 3 Figure 2G). In all the selected hits we could validate, we confirmed that elements targeted within 3' UTRs resulted in an increase in fitness, while those targeting within the coding sequence resulted in a decrease in fitness. The phenotypes for these 3' UTR-targeting guides provide a resource of potentially important *cis*-elements involved in mRNA stability, and further work could involve interrogating whether these sites are miRNA targets or targets of other proteins involved in RNA decay or translation processes.



Appendix 3 Figure 2

Appendix 3 Figure 2- Targeting of the 3' UTRs of Essential Genes to Probe for Previously Unidentified cis-Regulatory Elements. A. Average phenotypes for all 3' UTR-targeting guides was plotted. Error bars represent standard deviation of all sgRNAs. B–D. We summarize the phenotypes for coding-targeting guides (gray) and 3' UTR-targeting guides (pink) for select genes: Sdad1, Foxp1, and Cdk13. E. Scatterplot. The x-coordinates represent the number of unique barcodes associated with each random sgRNA at day 0. The y-coordinates represent the number of unique barcodes associated with each random sgRNA at day 21 ($R^2 = 0.814$). F. Unique barcode counts were obtained for the

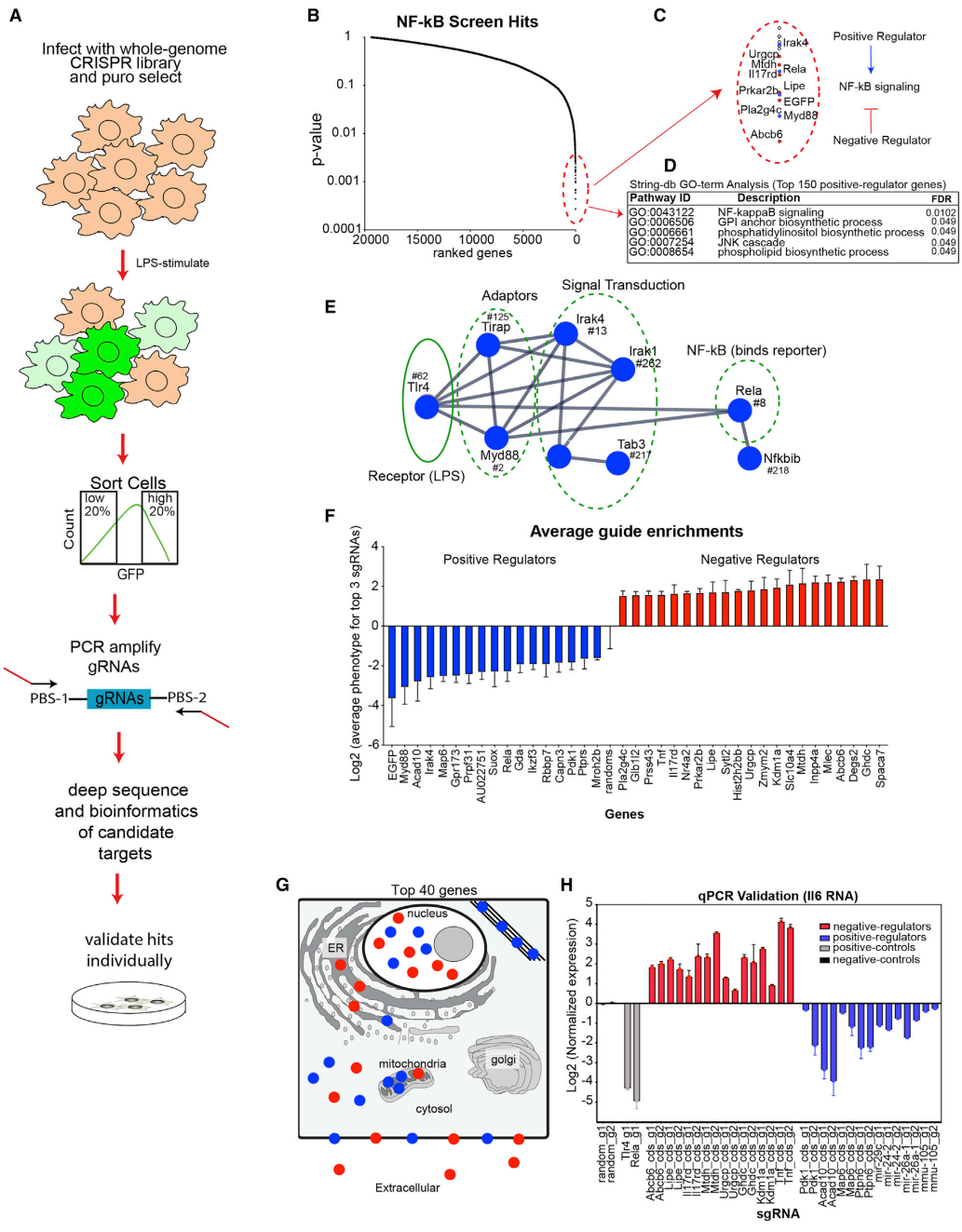
indicated genes, including random controls. The fold change of these counts from day 0 to day 21 was calculated. G. We selected 3' UTR-targeting guides that demonstrated positive enrichment, opposite to their coding-targeting guides, which had negative enrichment. We validated select guides by monitoring growth of sgRNA-infected cells (cherry) relative to uninfected reference cells in a mix-cell growth assay. Error bars represent standard deviation of three technical replicates.

3.3.3 FACS-Based Reporter Screen Identifies Positive and Negative Regulators of NF-kB

Macrophages are critical effectors of the inflammatory response, which involves transcription factors including NF-kB (Liu et al., 2017). We had previously developed NF-kB reporter iBMDMs, adding 5 3 NF-kB-binding motifs (GGGAATTTCC) upstream of the minimal cytomegalovirus (CMV) promoter-driving green fluorescent protein (GFP) and demonstrated lipopolysaccharide (LPS)-dependent activation of GFP fluorescence (Covarrubias et al., 2017). We lentivirally introduced Cas9 into these cells and confirmed its activity (iBMDM-NF-kB-Cas) (Covarrubias et al., 2017). Here, we performed a fluorescence-activated cell sorting (FACS)-based sorting screen using iBMDM-NF-kB-Cas cells and infected with the same library as outlined in Figure 1A. After the library was established in the cells for 7 days, we stimulated with LPS for 24 h (Appendix 3 Figures 3A, S3A, and S3B; Table S10). We sorted the top/bottom 20% of GFP-expressing cells and collected approximately 100 cells per sgRNA (~27 million cells for each top/bottom sort), with the aim of identifying both positive (bottom 20%) and negative (top 20%) regulators of the pathway (Appendix 3 Figures S3A and S3B; Table S11). Due to the inherently noisy nature of pooled screening, we chose to perform a MW U test to identify significant genes, comparing GFP-low to GFP- high sorted samples, and significant genes were ranked by $p < 0.01$ (Figure 3B) (Kampmann et al., 2013). As expected, we found several positive controls,

including EGFP and known regulators myeloid differentiation primary response 88 (Myd88) and Rela (NF-kB/p65) in our top hits (Figure 3C). We performed Gene Ontology (GO)-term enrichment analysis for the top 150 significant positive regulators and found enrichment for pathways that included ‘‘NF-kappaB signaling’’ (Appendix 3 Figure 3D). Within the top 150 genes, we identified numerous genes known to be involved in the Toll-like receptor (TLR)/NF-kB signaling pathway, which include Tlr4 (LPS receptor) and Rela (NF-kB/p65) (Figure 3E; Table S12), confirming that the screen was a success. We plotted the average phenotypes (top 3 guides) for our top 40 candidates, which showed that the average sgRNA enrichments were significant (Appendix 3 Figure 3F). Top candidates were localized throughout the cell (Appendix 3 Figure 3G; Table S13), including in the extracellular compartment. Numerous positive and negative regulators of NF-kB have been identified by their differential expression upon NF-kB activation (Bhatt and Ghosh, 2014). Using previously published data (Zhang et al., 2017), we examined the top 50 negative and positive regulators and found that the majority were not differentially expressed during LPS stimulation and, therefore, could have been missed by previous approaches as regulators of the pathway (Figure S3C). NF-kB has been demonstrated to activate genes that function in positive- or negative- feedback regulation of the pathway (Oeckinghaus and Ghosh, 2009). We assessed whether NF-kB (p65) bound to the promoters of our top candidates using published p65 chromatin immunoprecipitation sequencing (ChIP-seq) data (Lam et al., 2013) (Appendix 3 Figures S3D and S3E). We found

that p65 bound 42% and 54% of the top positive and negative regulators, respectively (Appendix 3 Figures S3D and S3E), further supporting the idea that a significant number of regulators of NF-kB can, themselves, be regulated by NF-kB. We validated our top candidates by re-cloning the top two performing sgRNAs per candidate, generating individual cell lines for each sgRNA, followed by lentiviral infection, selection, and LPS stimulation for 6 h (Appendix 3 Figure 3H; Table S14). The readout for our secondary validation experiments involved measuring Il6 by qPCR. Il6 is a well-known downstream target of NF-kB and a crucial gene in controlling inflammation (Appendix 3 Figure 3H). As expected, ablation of our negative regulators resulted in increased Il6 expression, while targeting of positive regulators led to decreased levels of Il6 relative to non-targeting controls (Appendix 3 Figure 3H). In summary, we performed a genome-wide screen and identified 50 previously unidentified positive and 65 negative regulators of NF-kB signaling. To date, there are 120 known regulators of NF-kB, and our screen has added 115 additional regulators ($p < 0.01$) to the pathway. These will provide a rich source of information going forward to better understand the complex pathways that control inflammation.



Appendix 3 Figure 3

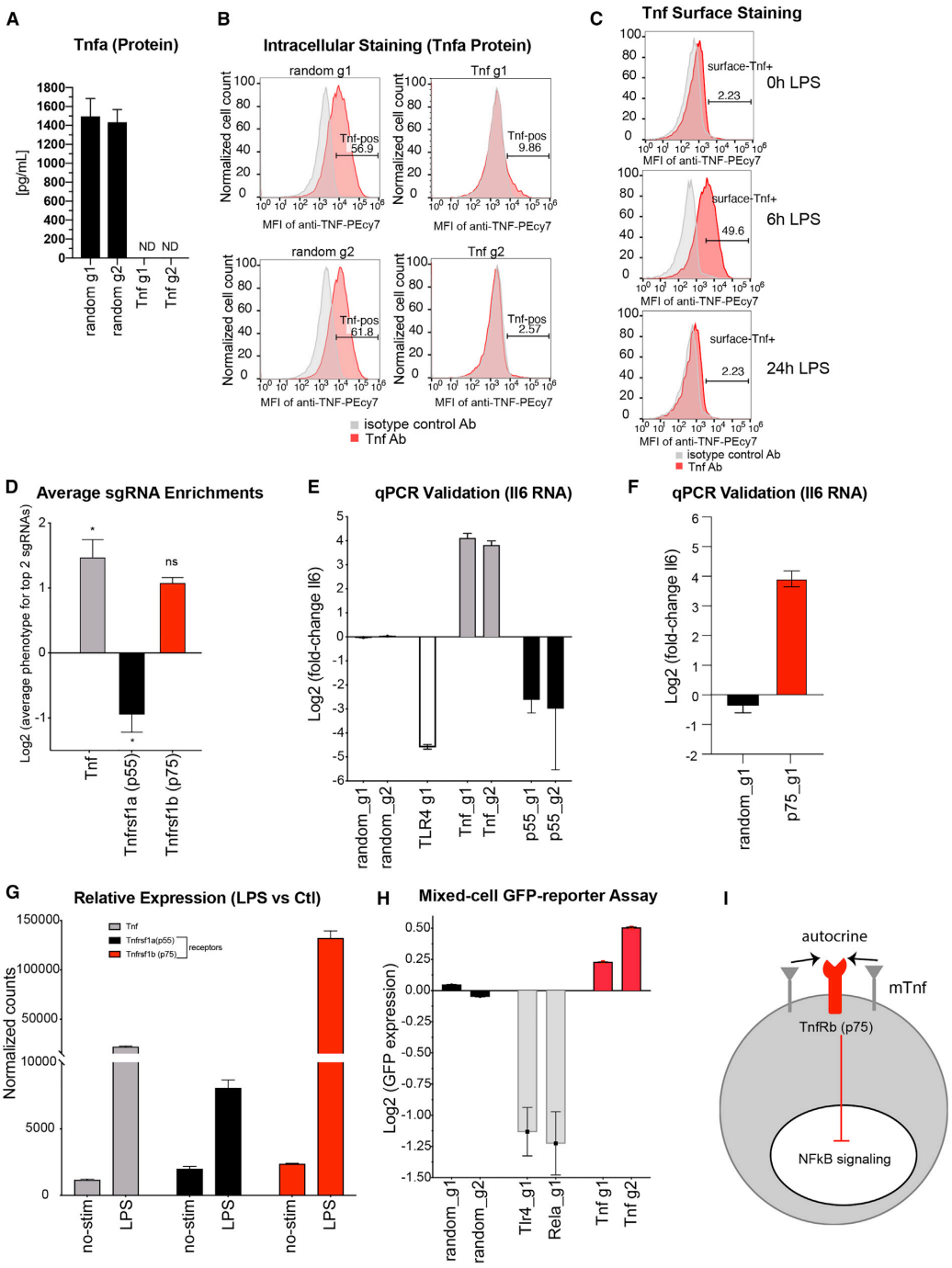
Appendix 3 Figure 3- Screen Identifies Positive and Negative Regulators of NFkB Signaling. A. Overview of the NF-kB screen: sgRNA-library-infected iBMDM-NF-kB-Cas9 cells were stimulated with LPS (200 ng/mL) for 24 h before sorting the top and bottom 20% of GFP-expressing cells. Cells were collected and processed as described in STAR Methods. B. MW U test was performed comparing 12 sgRNAs targeting each gene to the non-targeting controls for GFP-low versus GFP-high sorted samples. Significant genes are displayed, ranked by significance. C. Zoom-in of the top screen hits, displaying positive regulators (blue) and negative regulators (red). Diagram depicts positive- and negative-regulation NF-kB signaling. D. GO-term analysis was assessed for the top 150 positive regulators using STRINGdb. E. Connectivity was determined by STRINGdb for the top 150 positive-regulator candidates. F. Average sgRNA enrichment for the top 3 sgRNAs was calculated for the top 40 screen hits. Error bars represent standard deviation of three biological replicates.
G. Predicted protein localization was determined for the top 40 most significant genes using UniProt's COMPARTMENTS database (https://compartments.jensenlab.org/Downloads). H. Selected candidates were infected with either control (random) or candidate-specific sgRNAs and were stimulated for 6 h with LPS, before RNA harvest. qPCR-based validation was performed by conducting qRT-PCR for Il6 RNA relative to Gapdh. Experiment was repeated 3 times, and a representative experiment is displayed. Error bars represent standard deviation of three technical replicates.

3.3.4 Membrane-Bound TNF Alpha (TNF-a) Acts As A Strong Negative Regulator Of The NFkB Pathway

TNF-a is a well-known pro-inflammatory cytokine with established roles in driving NF-kB-related inflammation (Bradley, 2008). Furthermore, anti-TNF therapy is a proven method for the treatment of various inflammatory-related diseases (Ma and Xu, 2013). Given its established role as a soluble protein functioning as a positive regulator of NF-kB, it was surprising to discover that our pooled-based screen approach identified TNF as a strong negative regulator of inflammation (Appendix 3 Figures 3F and 3H). We confirmed TNF editing by stimulating control or anti-TNF edited cells with LPS for 24 h before collecting supernatant and analyzing TNF protein via ELISA. We found that TNF was undetectable in the supernatant after LPS stimulation (Appendix 3 Figure 4A). We also confirmed near-complete ablation of TNF via intracellular staining (Appendix 3 Figure 4B). Our screen suggested a localized function for TNF, which would be incompatible with its soluble state. Interestingly, TNF can exist as both soluble and membrane bound (Ardestani et al., 2013). We confirmed the presence of membrane-bound TNF on the surface of our iBMDMs and found maximal surface TNF at 6 h post-LPS stimulation (Appendix 3 Figure 4C). TNF mediates its inflammatory effect via binding to its receptor Tnfrsf1a (p55) (Bradley, 2008). Our screen confirmed Tnfrsf1a (p55) as a positive regulator of NF-kB in contrast to what was found for Tnf (Appendix 3 Figures 4D and 4E). However, TNF can bind

to two different receptors, Tnfrsf1a (p55) and Tnfrsf1b (p75), which may have opposing functions (Peschon et al., 1998). While the sgRNAs targeting p75 in our screen show a trend toward it being a negative regulator similar to TNF, it was not significant (Appendix 3 Figure 4D). However, we were able to confirm by qPCR that targeting p75 can result in an increase in Il6 (Appendix 3 Figure 4F) similar to that observed when TNF is knocked down (Appendix 3 Figures 4D and 4E). Interestingly, 6 h post-LPS stimulation, levels of TNF and Tnfrsf1b (p75) increased 11-fold and 66-fold, respectively, while the levels of Tnfrsf1a (p55) increased a moderate 4-fold (Appendix 3 Figure 4G). We evaluated whether the enhanced activation of NF-kB in TNF-edited cells could be rescued by mixing these cells with unedited (TNF-expressing) cells. We combined cherry- positive cells (containing TNF guide RNAs) with unedited cherry- negative cells at a 1:1 ratio. Mixed cells were LPS stimulated for 24 h before FACS analysis to measure GFP mean fluorescence in- tensity (MFI) (NF-kB activation) (Figure 4H). The enhanced NF-kB activation in TNF-edited cells could not be rescued by mixing these cells with unedited cells (even when we increased the cherry-negative cells to >75%). These data suggest that TNF has autocrine properties, acting within the cell from which it is produced, and neither soluble TNF production or membrane-bound TNF from neighboring cells can reverse the increase in inflammatory signaling (Appendix 3 Figure 4I). The expression profiles of TNF, TNFRSF1A, and TNFRSF1B were mirrored in the human THP1 monocytic cell line stimulated with PAM3CSK4 (TLR1/2 agonist), suggesting that this specific regulation of TNF is

conserved (Appendix 3 Figure S4). Here, we confirm that, indeed, it is membrane bound TNF that is functioning as a negative regulator of NF-kB, presumably through interactions with p75. Remarkably, our single screening assay could provide insight into this complex signaling cascade and bring together decades of different approaches to reveal the complexity of TNF signaling in macrophages involving membrane-bound forms of TNF in addition to the roles of the respective receptors.



Appendix 3 Figure 4

Appendix 3 Figure 4- Screening Identifies Negative-Feedbak Regulatory Loop Used by the Tumor Necrosis Factor (TNF) to Regulate Inflammation. A. ELISA was performed in control or TNF-edited iBMDMs stimulated with LPS for 24 h. Supernatant was harvested, and levels of TNF were quantified. Error bars represent standard deviation of three biological replicates. B. Control or TNF-edited iBMDMs were stimulated with LPS followed by brefeldin A treatment and stained to assess intracellular TNF levels by flow cytometry. C. iBMDMs were stimulated for the indicated time points (LPS for 0, 6, and 24 h), and cell surface expression of TNF was measured by flow cytometry. D. Average sgRNA enrichment for the top 2 sgRNAs was calculated for TNF and its receptors: p55 and p75. Error bars represent standard deviation of two biological replicates. E and F. qRT-PCR was performed for Il6 RNA in iBMDMs edited with either control, TNF, p55 (E), or p75 F. sgRNAs and stimulated with LPS for 6 h. Error bars represent standard deviation of three biological replicates. G. Mean fragments per kilobase of transcript per million mapped reads (FPKM) values (4 biological replicates) are plotted for genes Tnf, Tnfrsf1a, and Tnfrsf1b for both un-stimulated (‘‘no-stim’’) or LPS-stimulated (24 h) cells. Error bars represent standard deviation of four biological replicates. H. We combined cherry-positive cells (containing Tnf sgRNAs) with unedited cherry-negative cells at a 1:1 ratio. Mixed cells were stimulated with LPS for 24 h before FACS analysis to measure GFP mean fluorescence intensity (MFI) as proxy for NF-kB activation. Error bars represent standard deviation of three biological replicates. (I) Model of localized TNF negative regulation of NF-kB signaling.

3.4 Discussion

Here, we utilize CRISPR-based pooled genetic screening to reveal genes important for both macrophage viability and inflammation. For the viability screen, we utilized both MW U test analysis as well as the MAGeCK analysis pipelines to identify significant hits. 88% of significant genes identified in the MAGeCK analysis were also identified in the MW U test. A larger proportion of significant hits were identified by MW test, which could repre- sent true positives or that the more stringent MAGeCK pipeline identifies fewer false positives. We identified macrophage-specific viability genes enriched in NF-kB signaling (Figures 1F and S2). Additionally, we identified and validated viability phenotypes for 3' UTR-targeting guides, which may reveal important *cis*-elements involved in mRNA stability (discussed further later). In a separate screen, we utilized our NF-kB -reporter cells to perform a FACS-based screen, resulting in the identification of positive and negative regulators of NF-kB signaling. We go on to describe an unexpected role for TNF as a locally acting negative regulator of NF-kB. Very few miRNAs were identified as hits in either screen (Tables S3, S4, S9, S10, and S11). Although we targeted ~500 miRNAs, it is possible that many of these are not expressed in macrophages or that the sgRNAs targeting the miRNAs did not work.

3.4.1 Macrophage-Specific Genes Involved in Viability

Numerous viability screens have been performed in a wide range of cell types and have demonstrated that the catalog of essential genes can vary significantly among distinct cell types (Wang et al., 2014, 2017). Cell-type-specific differences in viability can be due to genetic differences that could render some pathways inactive, forcing dependency on other pathways. Indeed, Wang et al. (2017) compared essential genes across 14 acute myeloid leukemia (AML) lines and found certain genes to be essential only for a specific subset of the AML lines with certain genotypes. Moreover, a typical cell expresses approximately two thirds of its genes (Hart et al., 2013). Therefore, whether a gene or set of genes is essential can be context specific, varying upon different growth conditions or treatments, etc. (Hart et al., 2015; Viswanatha et al., 2018). In our screen, we identified 61 genes with distinct essentiality in macrophages compared to a database collection of ~421 cell-line CRISPR screens (Figure 1E). We selected 6 of the top candidates that acted as either essential genes or growth suppressors and could validate 4. Not surprisingly, we identified and validated the myeloid-specific transcrip- tion factor Irf8 as a macrophage-specific essential gene (Figure 1F). Interestingly, Irf8 has roles in both myeloid cell viability and inflammatory response (Langlais et al., 2016). Additionally, the macrophage-specific essential genes included ones involved in NF-kB signaling, suggesting that this pathway, similar to Irf8, may be important for both viability and inflammatory activation (Figure S2). Indeed, a significant fraction of them (~10 proteins)—including Syk and Mtmr9, which we

validated (Figure 1F)—are predicted to physically interact supporting a macrophage-specific regulation of protein complexes that control cell viability. In summary, understanding the pathways that control viability in macrophages is essential for development of novel gene targets to control macrophage proliferation in scenarios where the inflammatory response is not properly controlled.

3.4.2 3' UTR-Targeting Guides and Identification of Potential cis-Elements

UTRs of messages are important for regulating stability, translation, and localization (Mayr, 2019). Sequences within the UTRs (*cis*-elements) can function in miRNA binding, structure, and/or protein binding (Mayr, 2017). Furthermore, cell-type-specific expression differences in miRNAs and RNA binding proteins (RBPs) allow for specificity in regulating mRNAs (Erson-Bensan, 2016). Given the rich source of regulation that occurs within the 3' UTR, we built into our library design sgRNAs targeting 3' UTRs within known essential genes with the goal of systematically identifying *cis*-elements that contribute to the regulation of essential genes. For these genes, sgRNAs targeting coding sequences resulted in decreased fitness, as expected. Within these genes, we focused on 3' UTR-sgRNAs that resulted in increased fitness, which could represent the disruption of destabilizing *cis*-elements, such as miRNA binding sites, sites of inter- action with RBPs or AU-rich elements, etc. CRISPR targeting of these destabilizing *cis*-elements may result in stabilization of the mRNA. We determined whether there were any overlapping miRNA target sites using TargetScan, but we did not find any with significant overlap, suggesting that there are other mechanisms of regulation at play besides

miRNA targeting. However, Cdk13-utr3-g1 and Pias1-utr3-g1 (Figure 2G) did target regions near (<50–80 bp) the miRNA binding sites for mir-124 and mir-10, respectively, which could potentially disrupt secondary structure, mir binding, or both (Table S9). Another 3' UTR-targeting guide 1 for Arcn1, is predicted to target near a mushashi element, which is known to negatively regulate RNA stability (Bennett et al., 2016). Given that the majority of our 3' UTR targeting guides resulted in no phenotype, an alternative method to try in future studies would be to use 2 sgRNAs to create larger deletions (Zhao et al., 2017). In summary, we have functionally vali- dated potential 3' UTR *cis*-elements and provided a rich resource for future work aimed at dissecting how these elements contribute to gene expression.

3.4.3 Dissection of the Complex Biology of TNF

We utilized the MW U analysis to identify hits from our FACS- based NF-kB reporter screen. Sorting screens are inherently noisier than viability screens, due to the fact that significant hits are going to be proportional to the sgRNA enrichment, and with our single (LPS) treatment of cells, there is no ability to enrich over time. Here, we activated the NF-kB iBMDM reporter cells for 24 h with LPS and then sorted for the GFP low and high hits (top and bottom 20%) that represent positive and negative regulators of the pathway. Newer tools are beginning to be developed to help with the analysis of these complex pooled sorting-based screens. Recently, de Boer et al. 2020 published a tool, ‘‘MAUDE,’’ that is specifically designed for sorting screens, which could be beneficial for future studies of these

complex assays. Nevertheless, we could show, using MW analysis, that the GO terms for the top 150 hits included ‘‘NF-kappaB signaling,’’ and we confirmed 16 additional NF-kB regulators using individually cloned lines measuring Il6 production as a secondary readout.

One of the surprising findings from our NF-kB screen was the identification of TNF as a negative regulator of NF-kB (anti-inflammatory). We were surprised for two reasons: (1) TNF is an extensively studied pro-inflammatory cytokine with established roles in driving NF-kB-related inflammation (Bradley, 2008). (2) TNF is a secreted protein; therefore, a pooled CRISPR screen would not be expected to capture its biology. Numerous studies spanning decades of research have revealed that TNF biology is much more complex. TNF can exist as both soluble (17 kDa) and membrane-bound (26 kDa) forms, and multiple groups have shown that membrane-bound TNF functions as a negative regulator of inflammation in contrast to its soluble form (Alexopoulou et al., 2006; Ardestani et al., 2013). The dual functions of TNF can also be explained, in part, by its binding to two receptors: Tnfrsf1a (p55) and Tnfrsf1b (p75), which have opposing effects on inflammation (Peschon et al., 1998). Interestingly, the membrane-bound form of TNF has been shown to preferably bind to the inhibitory receptor, p75, allowing for localized regulation of inflammation (Grell et al., 1995). Here, we reveal the complex regulatory biology of TNF, providing evidence that membrane-bound TNF functions as a negative regulator of NF-kB likely through its binding to p75 in macrophages. Our results also support a model in which membrane-bound TNF can

function in an autocrine manner, acting within the cell that produces it. In this study, we bring together decades of research on TNF biology and its role in inflammation using this single-screening approach. Anti-TNF therapy remains one of the most effective methods for the treatment of various autoimmune diseases, including rheumatoid arthritis (RA) and irritable bowel disease (IBD) (Hyrich et al., 2009; Peyrin-Biroulet, 2010), yet as many as 20%– 40% of patients do not respond to treatment (Lopetuso et al., 2017). Our findings show that editing of TNF resulted in elevated Il6 levels, which is another important pro-inflammatory cytokine and might explain the lack of response to anti-TNF therapy. Yimin and Kohanawa (2006) demonstrated that a TNF knockout mouse showed elevated levels of Il6, which is consistent with our findings. More importantly, they presented data showing that production of TNF and IL6 can be negatively regulated by each other (Yimin and Kohanawa, 2006). Therefore, the targeting of TNF could lead to elevated levels of Il6, which may result in elevated inflammation in patients. Our data showed that editing p55 inhibits NF-kB-driven Il6 production (Figures 4D and 4E), while editing p75 caused an increase in Il6 (Figure 4F). Blocking p55 (via antibody or small molecule) could be an alternative that could block the pro-inflammatory effects of p55 while allowing Tnf to, instead, bind to p75 to promote anti-inflammatory signaling (Yang et al., 2018). In summary, with one pooled CRISPR screen, we have revealed interesting complexities of TNF regulation and have presented evidence for alternative therapeutic strategies.

3.4.4 Future Directions

Here, we have demonstrated the power of CRISPR screening in revealing important macrophage biology probing both viability and inflammatory pathways. Macrophages are critical cells of the innate immune system providing one of the first lines of defense against invading microbes. The ability for macrophages to function optimally requires their ability to proliferate and migrate to reach the site of infection and appropriately engage their inflammatory program. Here, we found 60 macrophage-specific viability genes and uncovered 115 additional regulators of NF-kB. The future characterization of these genes will likely yield novel regulatory insights into the complex regulation of viability and inflammatory pathways. From a therapeutic point of view, it would be interesting to explore whether there are drugs that may target some or any of our screen candidates (Wishart et al., 2018). In conclusion, we have revealed important biology insights related to macro- phage function and believe that this work represents a significant resource for the macrophage research community.

3.5 Methods

3.5.1 Experimental Model and Subject Details

Male immortalized bone-marrow-derived macrophages (iBMDMs) with the NF-kB reporter and Cas9 (iBMDM-NFKB-Cas9) cells (previously described in Covarrubias et al., 2017). Cas9 activity was validated via GFP kd (~75% kd) prior

to beginning the screen. Cells were cultured in DMEM, supplemented with 10% low-endotoxin fetal bovine serum (ThermoFisher) and 1X penicillin/strepto- mycin and incubated at 37°C in 5% CO2.

THP-1 cells were cultured in RPMI 1640 supplemented with 10% low-endotoxin fetal bovine serum (ThermoFisher), 1X penicillin/ streptomycin and 2-mercaptoethanol (0.05 mM, Sigma-Aldrich, M6250), and incubated at 37°C in 5% CO2.

3.5.2 sgRNA Library Design and Cloning

We created genome-scale sgRNA library consisting of over 270,000 total sgRNAs (12 sgRNAs per gene) targeting every RefSeq-an- notated (mm9) coding gene, as well as all microRNAs and select 3' UTRs. The library contains > 5,000 non-target control sequences (NTC). The earliest possible ‘‘constitutive’’ exon of each transcript variant was targeted. The criteria for sgRNA selection and the cloning strategy protocol have been previously described (Boettcher et al., 2018, 2019). All sgRNA sequences are shown in Table S1.

3.5.3 Lentiviral Production

HEK293T cells were seeded at 6,000,000 cells per plate in 15 cm dishes in 20 mL media (DMEM, 10% FBS) and incubated overnight at 37°C, 5% CO2. The next morning, 8 mg sgRNA library plasmid, 4 mg psPAX2 (Addgene #12260), 4

mg pMD2.G (Addgene #12259) and 80 mL lipofectamine2000 (Invitrogen) were mixed into 1 mL serum-free OptiMEM (GIBCO), vortexed and incubated for 20 min at RT and added to the cells. At 72 h post-transfection, supernatant was harvested, passed through 0.45 um filters (Millipore, Stericup) and aliquots were stored at 80°C.

3.5.4 CRISPR Screen

iBMDM-NFKB-Cas9 cells were infected with the sgRNA genome-scale library at a at low multiplicity of infection (MOI = 0.3). Three days post infection, cells were puromycin-selected (10 mg/ml) for 5 days to obtain cherry-positive (sgRNA) cells and were maintained at > 1000X coverage at all times.

3.5.5 Growth Screen

Prior to puro-selection, we collected a day 0 time point, consisting of 1000X coverage (270 million cells). We then collected a day 21 time point (also 1000X coverage). Cells from both time points were cryo-preserved in 90% FBS, 10% DMSO for later processing.

3.5.6 FACS Screen

Library infected and selected iBMDM-NFKB-Cas9 cells were expanded to 2000X coverage. Cells were stimulated with 200 ng/ml of LPS for 24 h to induce expression of GFP (NF-kB responsive). Prior to sorting, cells were collected in

FACS buffer (1XPBS, 1%FBS, 5mM EDTA). Stimulated cells were analyzed by flow cytometry alongside unstimulated cells to ensure the mean fluorescence intensity (MFI) of stimulated cells was > 10-fold compared to unstimulated cells. All flow cytometry experiments and screening were con- ducted on a BD FACSAria II. GFP was excited using a 488-nm laser and detected using a 525/50-nm filter. Sorting was conducted using 4-way purity into 2 tubes and a 100-mm nozzle. Cells were gated by forward (FSC-A) and side scatter (SSC-A) for live cells, then for single cells using FSC-A/FSC-H. Lastly, we evaluated GFP expression (SSC versus GFP), FACS sorted and collected the top/bot- tom 20% into separate tubes. At least 100 cells/sgRNA (100X coverage) for each sorted population were collected and cryo-pre- served in 90% FBS, 10% DMSO for later processing.

sgDNA processing, PCR and sequencing. Genomic DNA was collected from cell pellets (270 millions cells, 1000X coverage) or (27 millions cells, 100X coverage for sorted cells) and was extracted by methods described previously (Boettcher et al., 2018, 2019). A nested PCR strategy was used to 1) allow amplification sgRNA repertoire and 2) to add appropriate Illumina adapters for NGS (detailed protocol is described in Boettcher et al. (2019). For the 100X coverage sorted samples, we scaled the gDNA extraction volumes 1:5 (i.e., 2ml instead of 20ml). Quality and purity of the PCR product were assessed by bioanalyzer (Agilent), and sequencing was performed on an Illumina HiSeq 2500 platform using paired end 50 kits with the custom sequencing primer 5^0-

GAGACTATAAG- TATCCCTTGGAGAACCACCTTGTTGG-3' for reading the sgRNA sequence. Data was submitted to GEO. All tables can be accessed through Mendeley Reserved DOI: 10.17632/vtvxykv2cr.1.

3.5.7 Macrophage specific viability genes and 3' UTR guide validation (Mix-cell growth assay)

sgRNA-infected cells (cherry-pos) were mixed with uninfected cells (cherry-neg) at a 1:1 ratio in triplicate. We used Flow cytometry to monitor the ratio of cherry-pos to cherry-neg cells at 0- and 21-days post plating. All validation cytometry was performed on the Attune NxT Flow Cytometer.

3.5.8 NFkB guide Validation (qRT-PCR)

iBMDM-NFkB-Cas cells infected with indicated guide-expressing lentivirus and were stimulated with LPS (200 ng/ml) for 6 h prior to harvesting for RNA. Total cellular RNA from BMDM cell lines was isolated using the Direct-zol RNA MiniPrep Kit (Zymo Research) according to manufacturer's instructions. RNA was quantified and controlled for purity with a nanodrop spectrometer. (Thermo Fisher). For RT-qPCR, 500-1000 ng were reversely transcribed (iScript Reverse Transcription Supermix, Biorad) followed by RT- PCR (iQ SYBRgreen Supermix, Biorad) using the cycling conditions as follows: 50C for 2 min, 95°C for 2 min followed by 40 cycles of 95°C for 15 s, 60°C for 30 s and 72°C for 45 s. qRT-PCR primer sequences are list below.

3.5.9 ELISA Analysis

For the ELISA, iBMDMs were stimulated with LPS for 24 h and supernatant was collected from triplicate wells. Supernatant was diluted 1:3 and Tnf-alpha levels were measured using the mouse TNF-alpha DuoSet ELISA (R&D Systems) kit following manufacturer's protocol.

3.5.9 Antibody staining for FACS

For intracellular staining, iBMDMs were LPS-stimulated for 0 or 6 h and were treated with Brefeldin A for the last 5 hours of stimulation. Cells were then collected, fixed with 4% PFA and permeabilized with perm buffer (3% BSA, 0.2% Triton-X, 1XPBS), followed by antibody staining with PEcy7 anti-mouse Tnf-alpha (1:150, Thermofisher) or isotype control (1:150, Biolegend). For surface staining, iBMDMs were LPS-stimulated for 0, 6, 24 h. Cells were collected in sorting media (2% Fetal Calf Serum, 5mM EDTA, 1XPBS), treated with Fc receptor block (1:250, BD PharMingen) and were then stained with same Tnf and isotype antibodies used above, all done in sorting media.

3.5.10 RNA isolation and cDNA synthesis and RT-qPCR

Total cellular RNA from THP1 cell lines was isolated using the Direct-zol RNA MiniPrep Kit (Zymo Research) according to manufacturer's instructions. RNA was quantified and controlled for purity with a nanodrop spectrometer. (Thermo Fisher). For RT-qPCR, 500- 1000 ng were reversely transcribed (iScript

Reverse Transcription Supermix, Biorad) followed by RT-PCR (iQ SYBRgreen Supermix, Biorad) using the cycling conditions as follows: 50°C for 2 min, 95°C for 2 min followed by 40 cycles of 95°C for 15 s, 60°C for 30 s and 72°C for 45 s. The melting curve was graphically analyzed to control for nonspecific amplification reactions.

3.5.11 RNA-Sequencing

For generation of RNA-Sequencing libraries the human THP1 cells, RNA was isolated from control or Pam3CSK40-stimulated cells as described above and the RNA integrity was tested with a BioAnalyzer (Agilent Technologies). For RNA-Sequencing target RIN score of input RNA (500-1000ng) usually had a minimum RIN score of 8. RNA-Sequencing libraries were prepared with TruSeq stranded RNA sample preparation kits (Illumina), depletion of ribosomal RNA was performed by positive selection of polyA+ RNA. Sequencing was performed on Illumina HighSeq or NextSeq machines.

3.5.12 Quantification and Statistical Analysis

RNA-Sequencing

RNA-seq 50 bp reads were aligned to the human genome (assembly GRCh37/hg19) using TopHat. The Gencode V32 gtf was used as the input annotation. Differential gene expression specific analyses were conducted with the DESeq R package. Specifically, DESeq was used to normalize gene counts,

calculate fold change in gene expression, estimate p values and adjusted p values for change in gene expression values, and to perform a variance stabilized transformation on read counts to make them amenable to plotting.

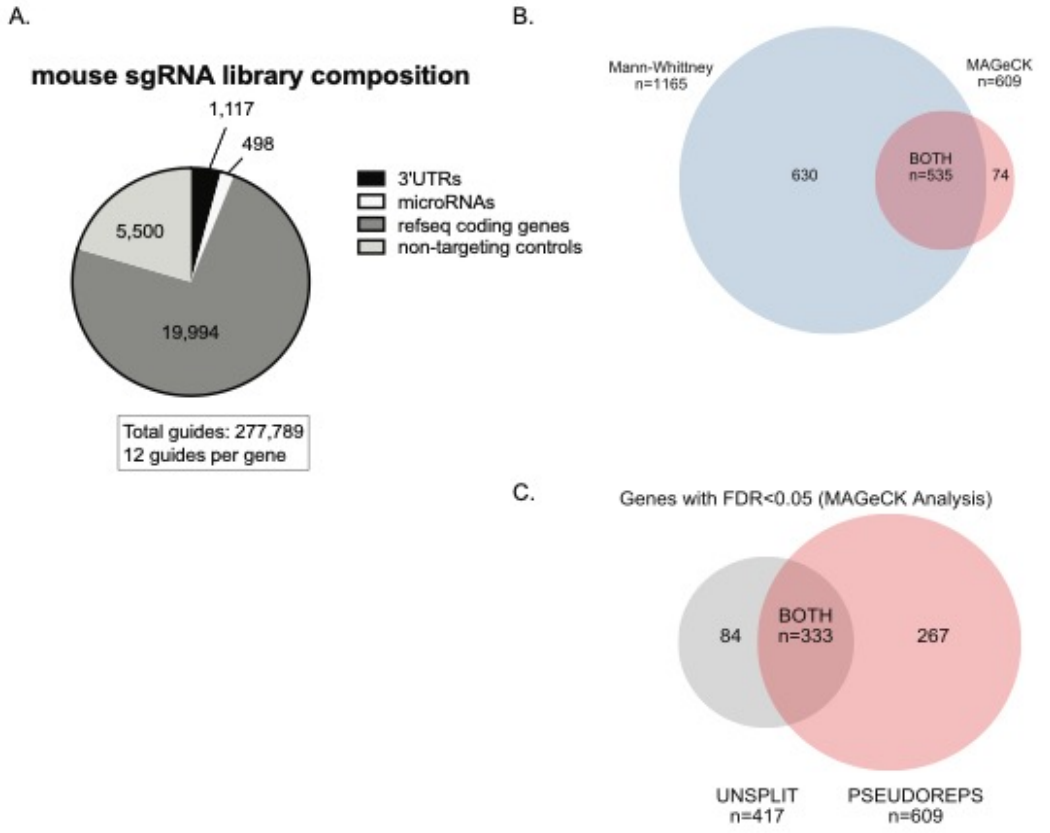
Screen Analysis and generation of hit list

fastq.gz files were analyzed using the gRNA_tool: https://github.com/quasiben/gRNA_Tool. All guide RNA (sgRNA) + barcode reads were collapsed to obtain raw sgRNA counts. Counts were normalized to the median and fold-changes were calculated for each sgRNA. To identify significant genes for the growth screen, the Mann-Whitney U test was performed comparing fold- changes for sgRNAs targeting each gene to non-targeting controls (described in Gilbert et al., 2014) or by following the MAGeCK analysis pipeline (as described in Li et al., 2014). MAGeCK analysis was performed on the full dataset as well as on the data binned into four separate samples (insample replicates) based on the 1st basepair of the random barcode (bins A,T,G, and C). The data was then used as input into the MAGeCK analysis pipeline. For the growth screen, the Day 21 sample was compared to the Day 0 sample. To identify significant genes for the for the FACS screen, GFP low (bottom 20%) sorted cells were compared to GFP high (top 20%) sorted cells.

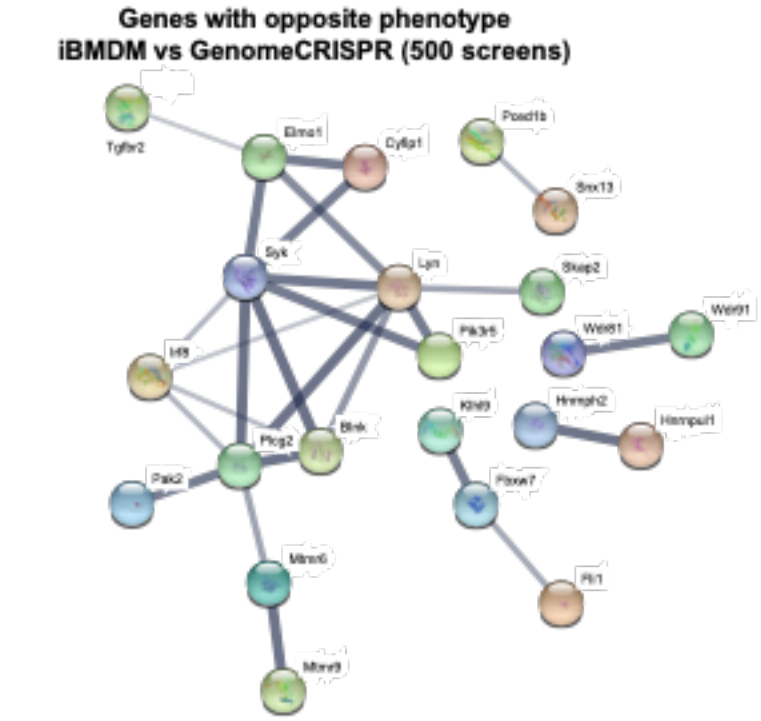
sgRNA selection for screen validation

For the macrophage specific viability genes, we selected 6 targets for validation out of possible 61 hits. We choose 5 essential genes and one growth suppressor. We chose the candidates based on their rankings according to P value. For the 3' UTR-targeting guides, we selected guides with > 3-fold positive enrichment (Day 21 versus Day 0) that targeted the 3' UTRs of essential genes with significant negative enrichment. The criteria for our NF-kB candidate selection was as follows: 1) Most significant: We selected candidates with the lowest Mann-Whitney U test p value (using a p value cut off of < 0.01). 2) Novelty: We focused on genes, which were not previously known to be involved in NF-kB signaling. We used several databases including KEGG, String-DB and Cell Signaling TLR-signaling gene list to determine novelty of gene. 3) Expression: We evaluated expression and confirmed > 10 FPKM for either un-stimulated or LPS-stimulation conditions. 4) Viability: We confirmed that our candidates did not have a significant viability phenotype. We targeted a total of 18 coding genes selecting guides with the strongest enrichment (2 guides/gene). We targeted microRNAs that showed > 3- fold positive enrichment for at least 2 gRNAs. For positive controls we used guides targeting Tlr4.

Supplemental information for Appendix 3

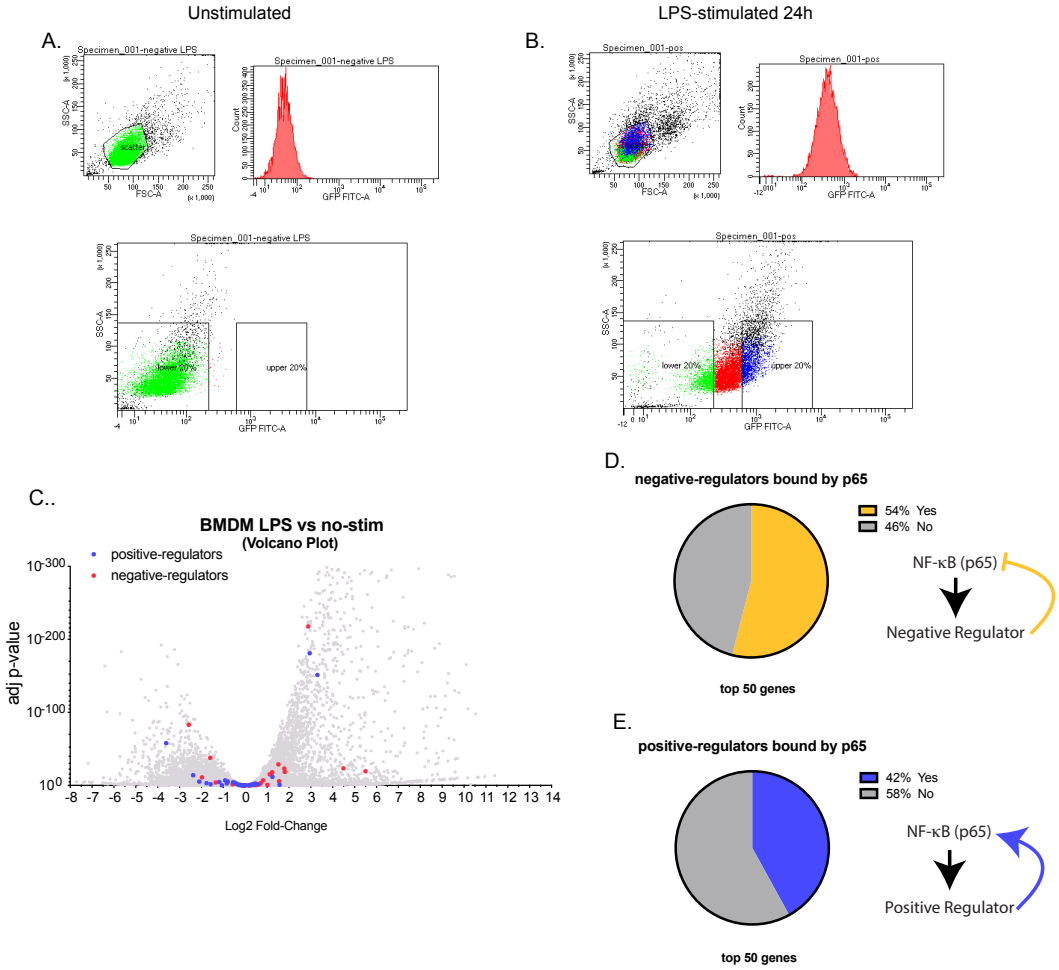


Supplemental Figure 2.1- Comaparing Analysis Screening Tools. Related to Figure 1. A. Breakdown of all genes targeted by our custom mouse sgRNA library is displayed.. B. Venn diagram. Significant genes were determined by Mann-Whittney U- test and MAGeCK analysis (FDR<0.05). 88% of significant genes identified in the MAGeCK analysis were also identified in the Mann-Whittney U-Test. C. Venn diagram. Significant genes were determined on unsplit and insamples replicates by MAGeCK analysis. 80% of the significant genes in the unsplit samples were also identified in the replicate samples.



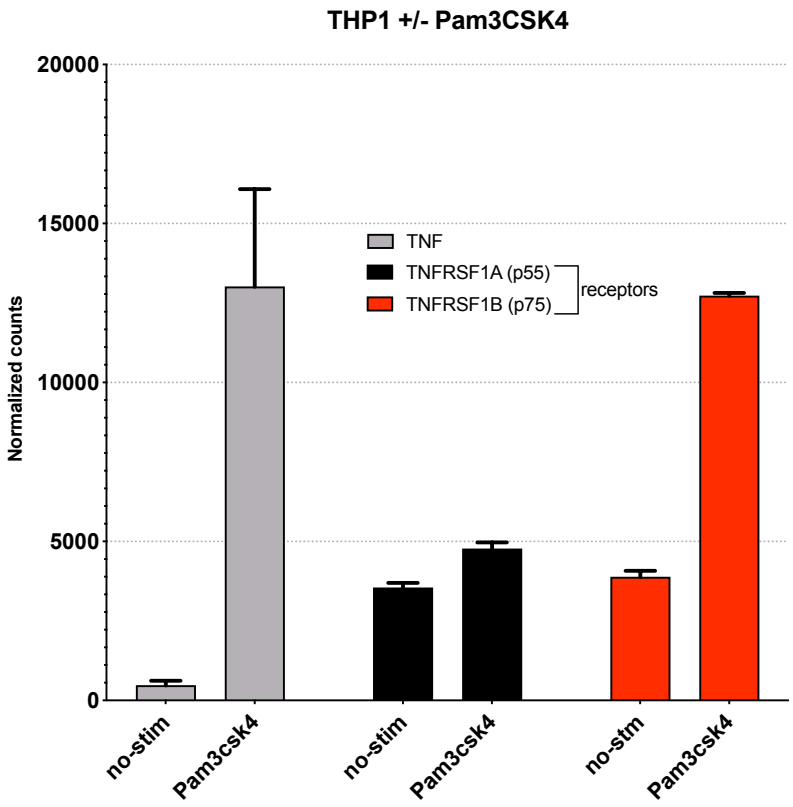
pathway	description	
mmu04662	B-cell receptor signaling pathway	0.00023
mmu04664	Fc epsilon RI signaling pathway	0.0025
mmu04064	NF-kappa B signaling pathway	0.006
mmu04066	HIF-1 signaling pathway	0.0064

Supplemental Figure 2.2- Comparison of Current Screen to CRISPR Screen Database. Related to Figure 1. Genes with opposite phenotypes in our screen compared to the GenomeCRISPR database are displayed visually using String- DB. KEGG pathway Go-term enrichment is also shown.



Supplemental Figure 2.3- Expression of p-65-binding of NFkB Screen Hits. Related to Figure 3. A-B. NF-κB FACS screen gating strategy for unstimulated cells (A) or 24 h LPS stimulated cells (B). C. Differentially expressed genes in 6 h LPS vs unstimulated BMDMs are displayed as log2 fold-change vs. adjusted p-value volcano plot from previously published data (Zhang et al., 2017). Expression of top 50 positive regulators (blue) and top 50 negative regulators (yellow) is shown. D-E. All p65 targets were determined using the ChIP-seq data from (Lam et al., 2013). Positive p65 binding was called if a p65 peak was greater than 10 and was within 1kb of the annotated transcription start site (TSS). P65 promoter binding was then assessed for the top negative (D) and positive (E) regulators.

A.



Supplemental Figure 2.4- Expression of TNF and TNF receptors in human cells. Related to Figure 4. A. Differentially expressed genes were determined for 6h Pam3CSK4-stimulated or unstimulated THP1 (ATCC). Normalized counts +/- SD are displayed for TNF, TNFRSF1A and TNFRSF1B.

Supplemental Tables- supplemental tables are available for download in the published manuscript: https://doi.org/10.1016/j.celrep.2020.108541

Table S1. sgRNA Library Sequences, Related to Figure 1
Table S2. All Targeted Genes, Related to Figure 1
Table S3. Viability Screen Hits, Related to Figure 1
Table S4. Phenotype sgRNAs Viability Hits, Related to Figure 1
Table S5. MAGeCK Analysis for Viability Screen Counts, Related to Figure 1
Table S6. MAGeCK Analysis for Viability Screen Genes, Related to Figure 1
Table S7. Viability Screen in Sample Rep Counts 2020, Related to Figure 1
Table S8. MAGeCK Analysis in Sample Rep Gene Summary, Related to Figure 1
Table S9. Elements within Targeted 3' UTRs, Related to Figure 2
Table S10. Full NFkB Screen Hit List, Related to Figure 3
Table S11. Phenotpe sgRNA NFkB Screen, Related to Figure 3
Table S12. Known Genes in TLR-NFkB Signaling, Related to Figure3
Table S13. NFkB Screen Candidate Localization, Related to Figure 3
Table S14. Validation sgRNA Sequences, Related to Figure 3

REFERENCES

Abdennur, N., Mirny, L.A., 2020. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics 36, 311–316. https://doi.org/10.1093/bioinformatics/btz540

Alexopoulou, L., Kranidioti, K., Xanthoulea, S., Denis, M., Kotanidou, A., Douni, E., Blackshear, P.J., Kontoyiannis, D.L., Kollias, G., 2006. Transmembrane TNF protects mutant mice against intracellular bacterial infections, chronic inflammation and autoimmunity. Eur. J. Immunol. 36, 2768–2780. https://doi.org/10.1002/eji.200635921

Ansari, S.A., Dantoft, W., Ruiz-Orera, J., Syed, A.P., Blachut, S., van Heesch, S., Hübner, N., Uhlenhaut, N.H., 2022. Integrative analysis of macrophage ribo-Seq and RNA-Seq data define glucocorticoid receptor regulated inflammatory response genes into distinct regulatory classes. Comput. Struct. Biotechnol. J. 20, 5622–5638. https://doi.org/10.1016/j.csbj.2022.09.042

Ardestani, S., Li, B., Deskins, D.L., Wu, H., Massion, P.P., Young, P.P., 2013. Membrane versus Soluble Isoforms of TNF-α Exert Opposing Effects on Tumor Growth and Survival of Tumor-Associated Myeloid Cells. Cancer Res. 73, 3938–3950. https://doi.org/10.1158/0008-5472.CAN-13-0002

Arnan, C., Ullrich, S., Pulido-Quetglas, C., Nurtdinov, R., Esteban, A., Blanco-Fernandez, J., Aparicio-Prat, E., Johnson, R., Pérez-Lluch, S., Guigó, R., 2022. Paired guide RNA CRISPR-Cas9 screening for protein-coding

genes and lncRNAs involved in transdifferentiation of human B-cells to macrophages. BMC Genomics 23, 402. https://doi.org/10.1186/s12864-022-08612-7

Bailey, T.L., Johnson, J., Grant, C.E., Noble, W.S., 2015. The MEME Suite. Nucleic Acids Res. 43, W39–W49. https://doi.org/10.1093/nar/gkv416

Barazandeh, M., Lambert, S.A., Albu, M., Hughes, T.R., 2018. Comparison of ChIP-Seq Data and a Reference Motif Set for Human KRAB C2H2 Zinc Finger Proteins. G3 Bethesda Md 8, 219–229. https://doi.org/10.1534/g3.117.300296

Barnig, C., Bezema, T., Calder, P.C., Charloux, A., Frossard, N., Garssen, J., Haworth, O., Dilevskaya, K., Levi-Schaffer, F., Lonsdorfer, E., Wauben, M., Kraneveld, A.D., te Velde, A.A., 2019. Activation of Resolution Pathways to Prevent and Fight Chronic Inflammation: Lessons From Asthma and Inflammatory Bowel Disease. Front. Immunol. 10, 1699. https://doi.org/10.3389/fimmu.2019.01699

Bennett, C.G., Riemondy, K., Chapnick, D.A., Bunker, E., Liu, X., Kuersten, S., Yi, R., 2016. Genome-wide analysis of Musashi-2 targets reveals novel functions in governing epithelial cell migration. Nucleic Acids Res. 44, 3788–3800. https://doi.org/10.1093/nar/gkw207

Bhan, A., Soleimani, M., Mandal, S.S., 2017. Long Noncoding RNA and Cancer: A New Paradigm. Cancer Res. 77, 3965–3981. https://doi.org/10.1158/0008-5472.CAN-16-2634

Bhatt, D., Ghosh, S., 2014. Regulation of the NF-κB-Mediated Transcription of Inflammatory Genes. Front. Immunol. 5. https://doi.org/10.3389/fimmu.2014.00071

Boettcher, M., Covarrubias, S., Biton, A., Blau, J., Wang, H., Zaitlen, N., McManus, M.T., 2019. Tracing cellular heterogeneity in pooled genetic screens via multi-level barcoding. BMC Genomics 20, 107. https://doi.org/10.1186/s12864-019-5480-0

Bolha, L., Ravnik-Glavač, M., Glavač, D., 2017. Long Noncoding RNAs as Biomarkers in Cancer. Dis. Markers 2017, 1–14. https://doi.org/10.1155/2017/7243968

Bonadies, N., Neururer, C., Steege, A., Vallabhapurapu, S., Pabst, T., Mueller, B.U., 2010. PU.1 is regulated by NF-κB through a novel binding site in a 17 kb upstream enhancer element. Oncogene 29, 1062–1072. https://doi.org/10.1038/onc.2009.371

Bradley, J., 2008. TNF-mediated inflammatory disease. J. Pathol. 214, 149–160. https://doi.org/10.1002/path.2287

Brodnicki, T.C., 2022. A Role for lncRNAs in Regulating Inflammatory and Autoimmune Responses Underlying Type 1 Diabetes, in: Carpenter, S. (Ed.), Long Noncoding RNA, Advances in Experimental Medicine and Biology. Springer International Publishing, Cham, pp. 97–118. https://doi.org/10.1007/978-3-030-92034-0_6

Brown, J., Pirrung, M., McCue, L.A., 2017. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics 33, 3137–3139. https://doi.org/10.1093/bioinformatics/btx373

Buenrostro, J.D., Wu, B., Chang, H.Y., Greenleaf, W.J., 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr. Protoc. Mol. Biol. 109. https://doi.org/10.1002/0471142727.mb2129s109

Bulger, M., Groudine, M., 2010. Enhancers: The abundance and function of regulatory sequences beyond promoters. Dev. Biol. 339, 250–257. https://doi.org/10.1016/j.ydbio.2009.11.035

Burns, K.H., 2017. Transposable elements in cancer. Nat. Rev. Cancer 17, 415–424. https://doi.org/10.1038/nrc.2017.35

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., Rinn, J.L., 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 25, 1915–1927. https://doi.org/10.1101/gad.17446611

Castellanos-Rubio, A., Ghosh, S., 2022. Functional Implications of Intergenic GWAS SNPs in Immune-Related LncRNAs, in: Carpenter, S. (Ed.), Long Noncoding RNA, Advances in Experimental Medicine and Biology. Springer International Publishing, Cham, pp. 147–160. https://doi.org/10.1007/978-3-030-92034-0_8

Chen, H., Du, G., Song, X., Li, L., 2017. Non-coding Transcripts from Enhancers: New Insights into Enhancer Activity and Gene Expression Regulation. Genomics Proteomics Bioinformatics 15, 201–207. https://doi.org/10.1016/j.gpb.2017.02.003

Cho, S.W., Xu, J., Sun, R., Mumbach, M.R., Carter, A.C., Chen, Y.G., Yost, K.E., Kim, J., He, J., Nevins, S.A., Chin, S.-F., Caldas, C., Liu, S.J., Horlbeck, M.A., Lim, D.A., Weissman, J.S., Curtis, C., Chang, H.Y., 2018. Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element. Cell 173, 1398-1412.e22. https://doi.org/10.1016/j.cell.2018.03.068

Couso, J.-P., Patraquim, P., 2017. Classification and function of small open reading frames. Nat. Rev. Mol. Cell Biol. 18, 575–589. https://doi.org/10.1038/nrm.2017.58

Covarrubias, S., Robinson, E.K., Shapleigh, B., Vollmers, A., Katzman, S., Hanley, N., Fong, N., McManus, M.T., Carpenter, S., 2017. CRISPR/Cas-based screening of long non-coding RNAs (lncRNAs) in macrophages with an NF-κB reporter. J. Biol. Chem. 292, 20911–20920. https://doi.org/10.1074/jbc.M117.799155

Covarrubias, S., Vollmers, A.C., Capili, A., Boettcher, M., Shulkin, A., Correa, M.R., Halasz, H., Robinson, E.K., O'Briain, L., Vollmers, C., Blau, J., Katzman, S., McManus, M.T., Carpenter, S., 2020. High-Throughput CRISPR Screening Identifies Genes Involved in Macrophage Viability

and Inflammatory Pathways. Cell Rep. 33, 108541. https://doi.org/10.1016/j.celrep.2020.108541

Damase, T.R., Sukhovershin, R., Boada, C., Taraballi, F., Pettigrew, R.I., Cooke, J.P., 2021. The Limitless Future of RNA Therapeutics. Front. Bioeng. Biotechnol. 9, 628137. https://doi.org/10.3389/fbioe.2021.628137

de Boer, C.G., Ray, J.P., Hacohen, N., Regev, A., 2020. MAUDE: inferring expression changes in sorting-based CRISPR screens. Genome Biol. 21, 134. https://doi.org/10.1186/s13059-020-02046-8

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G.K., Khatun, J., Williams, B.A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R.F., Alioto, T., Antoshechkin, I., Baer, M.T., Bar, N.S., Batut, P., Bell, K., Bell, I., Chakrabortty, S., Chen, X., Chrast, J., Curado, J., Derrien, T., Drenkow, J., Dumais, E., Dumais, J., Duttagupta, R., Falconnet, E., Fastuca, M., Fejes-Toth, K., Ferreira, P., Foissac, S., Fullwood, M.J., Gao, H., Gonzalez, D., Gordon, A., Gunawardena, H., Howald, C., Jha, S., Johnson, R., Kapranov, P., King, B., Kingswood, C., Luo, O.J., Park, E., Persaud, K., Preall, J.B., Ribeca, P., Risk, B., Robyr, D., Sammeth, M., Schaffer, L., See, L.-H., Shahab, A., Skancke, J., Suzuki, A.M., Takahashi, H., Tilgner, H., Trout, D., Walters, N., Wang, H., Wrobel, J., Yu, Y., Ruan, X., Hayashizaki, Y., Harrow, J., Gerstein, M., Hubbard, T., Reymond, A., Antonarakis, S.E., Hannon, G.,

Giddings, M.C., Ruan, Y., Wold, B., Carninci, P., Guigó, R., Gingeras, T.R., 2012. Landscape of transcription in human cells. Nature 489, 101–108. https://doi.org/10.1038/nature11233

Dolgalev, I., 2020. msigdbr: MSigDB gene sets for multiple organisms in a tidy data format. R Package Version 7.

Dugger, S.A., Platt, A., Goldstein, D.B., 2018. Drug development in the era of precision medicine. Nat. Rev. Drug Discov. 17, 183–196. https://doi.org/10.1038/nrd.2017.226

Enderle, D., Spiel, A., Coticchia, C.M., Berghoff, E., Mueller, R., Schlumpberger, M., Sprenger-Haussels, M., Shaffer, J.M., Lader, E., Skog, J., Noerholm, M., 2015. Characterization of RNA from Exosomes and Other Extracellular Vesicles Isolated by a Novel Spin Column-Based Method. PLOS ONE 10, e0136133. https://doi.org/10.1371/journal.pone.0136133

Erson-Bensan, A.E., 2016. Alternative polyadenylation and RNA-binding proteins. J. Mol. Endocrinol. 57, F29–F34. https://doi.org/10.1530/JME-16-0070

Everaert, C., Helsmoortel, H., Decock, A., Hulstaert, E., Van Paemel, R., Verniers, K., Nuytens, J., Anckaert, J., Nijs, N., Tulkens, J., Dhondt, B., Hendrix, A., Mestdagh, P., Vandesompele, J., 2019. Performance assessment of total RNA sequencing of human biofluids and extracellular vesicles. Sci. Rep. 9, 17574. https://doi.org/10.1038/s41598-019-53892-x

Franceschi, C., Garagnani, P., Parini, P., Giuliani, C., Santoro, A., 2018. Inflammaging: a new immune–metabolic viewpoint for age-related diseases. Nat. Rev. Endocrinol. 14, 576–590. https://doi.org/10.1038/s41574-018-0059-4

Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J.E., Mudge, J.M., Sisu, C., Wright, J.C., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Boix, C., Carbonell Sala, S., Cunningham, F., Di Domenico, T., Donaldson, S., Fiddes, I.T., García Girón, C., Gonzalez, J.M., Grego, T., Hardy, M., Hourlier, T., Howe, K.L., Hunt, T., Izuogu, O.G., Johnson, R., Martin, F.J., Martínez, L., Mohanan, S., Muir, P., Navarro, F.C.P., Parker, A., Pei, B., Pozo, F., Riera, F.C., Ruffier, M., Schmitt, B.M., Stapleton, E., Suner, M.-M., Sycheva, I., Uszczynska-Ratajczak, B., Wolf, M.Y., Xu, J., Yang, Y.T., Yates, A., Zerbino, D., Zhang, Y., Choudhary, J.S., Gerstein, M., Guigó, R., Hubbard, T.J.P., Kellis, M., Paten, B., Tress, M.L., Flicek, P., 2021. GENCODE 2021. Nucleic Acids Res. 49, D916–D923. https://doi.org/10.1093/nar/gkaa1087

Fritsch, C., Herrmann, A., Nothnagel, M., Szafranski, K., Huse, K., Schumann, F., Schreiber, S., Platzer, M., Krawczak, M., Hampe, J., Brosch, M., 2012. Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. Genome Res. 22, 2208–2218. https://doi.org/10.1101/gr.139568.112

Gannon, H.S., Zou, T., Kiessling, M.K., Gao, G.F., Cai, D., Choi, P.S., Ivan, A.P., Buchumenski, I., Berger, A.C., Goldstein, J.T., Cherniack, A.D., Vazquez, F., Tsherniak, A., Levanon, E.Y., Hahn, W.C., Meyerson, M., 2018. Identification of ADAR1 adenosine deaminase dependency in a subset of cancer cells. Nat. Commun. 9, 5450. https://doi.org/10.1038/s41467-018-07824-4

Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.-L., Ragoussis, J., Natoli, G., 2010. Identification and Characterization of Enhancers Controlling the Inflammatory Gene Expression Program in Macrophages. Immunity 32, 317–328. https://doi.org/10.1016/j.immuni.2010.02.008

Gil, N., Ulitsky, I., 2020. Regulation of gene expression by cis-acting long non-coding RNAs. Nat. Rev. Genet. 21, 102–117. https://doi.org/10.1038/s41576-019-0184-5

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., Qi, L.S., Kampmann, M., Weissman, J.S., 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell 159, 647–661. https://doi.org/10.1016/j.cell.2014.09.029

Gilmore, T.D., 1999. The Rel/NF-κB signal transduction pathway: introduction. Oncogene 18, 6842–6844. https://doi.org/10.1038/sj.onc.1203237

Goldman, M.J., Craft, B., Hastie, M., Repečka, K., McDade, F., Kamath, A., Banerjee, A., Luo, Y., Rogers, D., Brooks, A.N., Zhu, J., Haussler, D., 2020. Visualizing and interpreting cancer genomics data via the Xena platform. Nat. Biotechnol. 38, 675–678. https://doi.org/10.1038/s41587-020-0546-8

Grell, M., Douni, E., Wajant, H., Löhden, M., Clauss, M., Maxeiner, B., Georgopoulos, S., Lesslauer, W., Kollias, G., Pfizenmaier, K., Scheurich, P., 1995. The transmembrane form of tumor necrosis factor is the prime activating ligand of the 80 kDa tumor necrosis factor receptor. Cell 83, 793–802. https://doi.org/10.1016/0092-8674(95)90192-2

Groff, A.F., Sanchez-Gomez, D.B., Soruco, M.M.L., Gerhardinger, C., Barutcu, A.R., Li, E., Elcavage, L., Plana, O., Sanchez, L.V., Lee, J.C., Sauvageau, M., Rinn, J.L., 2016. In Vivo Characterization of Linc-p21 Reveals Functional cis-Regulatory DNA Elements. Cell Rep. 16, 2178–2186. https://doi.org/10.1016/j.celrep.2016.07.050

Gu, Z., Eils, R., Schlesner, M., 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847–2849. https://doi.org/10.1093/bioinformatics/btw313

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., Yang, X., Amit, I., Meissner, A., Regev, A., Rinn, J.L., Root, D.E., Lander, E.S., 2011. lincRNAs act in

the circuitry controlling pluripotency and differentiation. Nature 477, 295–300. https://doi.org/10.1038/nature10398

Hah, N., Benner, C., Chong, L.-W., Yu, R.T., Downes, M., Evans, R.M., 2015. Inflammation-sensitive super enhancers form domains of coordinately regulated enhancer RNAs. Proc. Natl. Acad. Sci. 112. https://doi.org/10.1073/pnas.1424028112

Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye, F., Huang, D., Xu, Y., Huang, W., Jiang, M., Jiang, X., Mao, J., Chen, Y., Lu, C., Xie, J., Fang, Q., Wang, Y., Yue, R., Li, T., Huang, H., Orkin, S.H., Yuan, G.-C., Chen, M., Guo, G., 2018. Mapping the Mouse Cell Atlas by Microwell-Seq. Cell 172, 1091-1107.e17. https://doi.org/10.1016/j.cell.2018.02.001

Harrison, P.M., 2002. A question of size: the eukaryotic proteome and the problems in defining it. Nucleic Acids Res. 30, 1083–1090. https://doi.org/10.1093/nar/30.5.1083

Haswell, J.R., Mattioli, K., Gerhardinger, C., Maass, P.G., Foster, D.J., Peinado, P., Wang, X., Medina, P.P., Rinn, J.L., Slack, F.J., 2021. Genome-wide CRISPR interference screen identifies long non-coding RNA loci required for differentiation and pluripotency. PLOS ONE 16, e0252848. https://doi.org/10.1371/journal.pone.0252848

Hennessy, E.J., 2022. LncRNAs and Cardiovascular Disease, in: Carpenter, S. (Ed.), Long Noncoding RNA, Advances in Experimental Medicine and

Biology. Springer International Publishing, Cham, pp. 71–95. https://doi.org/10.1007/978-3-030-92034-0_5

Her, M., Kavanaugh, A., 2016. Alterations in immune function with biologic therapies for autoimmune disease. J. Allergy Clin. Immunol. 137, 19–27. https://doi.org/10.1016/j.jaci.2015.10.023

Huang, Y., Gulshan, K., Nguyen, T., Wu, Y., 2017. Biomarkers of Cardiovascular Disease. Dis. Markers 2017, 1–2. https://doi.org/10.1155/2017/8208609

Hyrich, K.L., Deighton, C., Watson, K.D., BSRBR Control Centre Consortium, Symmons, D.P.M., Lunt, M., on behalf of the British Society for Rheumatology Biologics Register, 2009. Benefit of anti-TNF therapy in rheumatoid arthritis patients with moderate disease activity. Rheumatology 48, 1323–1327. https://doi.org/10.1093/rheumatology/kep242

Imbeault, M., Helleboid, P.-Y., Trono, D., 2017. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. Nature 543, 550–554. https://doi.org/10.1038/nature21683

International Human Genome Sequencing Consortium, Whitehead Institute for Biomedical Research, Center for Genome Research:, Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W.,

Naylor, J., Raymond, Christina, Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., The Sanger Centre:, Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Washington University Genome Sequencing Center, Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., US DOE Joint Genome Institute:, Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.-F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Baylor College of Medicine Human Genome Sequencing Center:, Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., RIKEN Genomic Sciences Center:, Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Genoscope and CNRS

UMR-8030:, Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Department of Genome Analysis, Institute of Molecular Biotechnology:, Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., GTC Sequencing Center:, Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Beijing Genomics Institute/Human Genome Center:, Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Multimegabase Sequencing Center, The Institute for Systems Biology:, Hood, L., Rowen, L., Madan, A., Qin, S., Stanford Genome Technology Center:, Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., University of Oklahoma's Advanced Center for Genome Technology:, Roe, B.A., Chen, F., Pan, H., Max Planck Institute for Molecular Genetics:, Ramser, J., Lehrach, H., Reinhardt, R., Cold Spring Harbor Laboratory, Lita Annenberg Hazen Genome Center:, McCombie, W.R., de la Bastide, M., Dedhia, N., GBF—German Research Centre for Biotechnology:, Blöcker, H., Hornischer, K., Nordsiek, G., *Genome Analysis Group (listed in alphabetical order, also includes individuals listed under other headings):, Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.-C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G.R., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S.,

Jones, T.A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F.A., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.-P., Yeh, R.-F., Scientific management: National Human Genome Research Institute, US National Institutes of Health:, Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Stanford Human Genome Center:, Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., University of Washington Genome Center:, Olson, M.V., Kaul, R., Raymond, Christopher, Department of Molecular Biology, Keio University School of Medicine:, Shimizu, N., Kawasaki, K., Minoshima, S., University of Texas Southwestern Medical Center at Dallas:, Evans, G.A., Athanasiou, M., Schultz, R., Office of Science, US Department of Energy:, Patrinos, A., The Wellcome Trust:, Morgan, M.J., 2001. Initial sequencing and analysis of the human genome. Nature 409, 860–921. https://doi.org/10.1038/35057062

Jackson, E.L., Willis, N., Mercer, K., Bronson, R.T., Crowley, D., Montoya, R., Jacks, T., Tuveson, D.A., 2001. Analysis of lung tumor initiation and progression using conditional expression of oncogenic *K-ras*. Genes Dev. 15, 3243–3248. https://doi.org/10.1101/gad.943001

Jackson, R., Kroehling, L., Khitun, A., Bailis, W., Jarret, A., York, A.G., Khan, O.M., Brewer, J.R., Skadow, M.H., Duizer, C., Harman, C.C.D., Chang, L., Bielecki, P., Solis, A.G., Steach, H.R., Slavoff, S., Flavell, R.A., 2018. The translation of non-canonical open reading frames controls mucosal immunity. Nature 564, 434–438. https://doi.org/10.1038/s41586-018-0794-7

Ji, Z., Song, R., Regev, A., Struhl, K., 2015. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. eLife 4, e08890. https://doi.org/10.7554/eLife.08890

Kaikkonen, M.U., Spann, N.J., Heinz, S., Romanoski, C.E., Allison, K.A., Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C., Glass, C.K., 2013. Remodeling of the Enhancer Landscape during Macrophage Activation Is Coupled to Enhancer Transcription. Mol. Cell 51, 310–325. https://doi.org/10.1016/j.molcel.2013.07.010

Kampmann, M., Bassik, M.C., Weissman, J.S., 2013. Integrated platform for genome-wide screening and construction of high-density genetic interaction maps in mammalian cells. Proc. Natl. Acad. Sci. 110. https://doi.org/10.1073/pnas.1307002110

Karpurapu, M., Wang, X., Deng, J., Park, H., Xiao, L., Sadikot, R.T., Frey, R.S., Maus, U.A., Park, G.Y., Scott, E.W., Christman, J.W., 2011. Functional PU.1 in macrophages has a pivotal role in NF-κB activation and

neutrophilic lung inflammation during endotoxemia. Blood 118, 5255–5266. https://doi.org/10.1182/blood-2011-03-341123

Kelley, D., Rinn, J., 2012. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. Genome Biol. 13, R107. https://doi.org/10.1186/gb-2012-13-11-r107

Kent, O.A., Chivukula, R.R., Mullendore, M., Wentzel, E.A., Feldmann, G., Lee, K.H., Liu, S., Leach, S.D., Maitra, A., Mendell, J.T., 2010. Repression of the miR-143/145 cluster by oncogenic Ras initiates a tumor-promoting feed-forward pathway. Genes Dev. 24, 2754–2759. https://doi.org/10.1101/gad.1950610

Kim, D.H., Marinov, G.K., Pepke, S., Singer, Z.S., He, P., Williams, B., Schroth, G.P., Elowitz, M.B., Wold, B.J., 2015. Single-Cell Transcriptome Analysis Reveals Dynamic Changes in lncRNA Expression during Reprogramming. Cell Stem Cell 16, 88–101. https://doi.org/10.1016/j.stem.2014.11.005

Knott, G.J., Doudna, J.A., 2018. CRISPR-Cas guides the future of genetic engineering. Science 361, 866–869. https://doi.org/10.1126/science.aat5011

Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., Sergushichev, A., 2016. Fast gene set enrichment analysis (preprint). Bioinformatics. https://doi.org/10.1101/060012

Lai, Y., Cui, L., Babunovic, G.H., Fortune, S.M., Doench, J.G., Lu, T.K., 2020. High-throughput CRISPR screens to dissect macrophage- *Shigella* interactions (preprint). Microbiology. https://doi.org/10.1101/2020.04.25.061671

Lam, M.T.Y., Cho, H., Lesch, H.P., Gosselin, D., Heinz, S., Tanaka-Oishi, Y., Benner, C., Kaikkonen, M.U., Kim, A.S., Kosaka, M., Lee, C.Y., Watt, A., Grossman, T.R., Rosenfeld, M.G., Evans, R.M., Glass, C.K., 2013. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature 498, 511–515. https://doi.org/10.1038/nature12209

Langlais, D., Barreiro, L.B., Gros, P., 2016. The macrophage IRF8/IRF1 regulome is required for protection against infections and is associated with chronic inflammation. J. Exp. Med. 213, 585–603. https://doi.org/10.1084/jem.20151764

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923

Larson, M.H., Pan, W., Kim, H.J., Mauntz, R.E., Stuart, S.M., Pimentel, M., Zhou, Y., Knudsgaard, P., Demas, V., Aravanis, A.M., Jamshidi, A., 2021. A comprehensive characterization of the cell-free transcriptome reveals tissue- and subtype-specific biomarkers for cancer detection. Nat. Commun. 12, 2357. https://doi.org/10.1038/s41467-021-22444-1

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., Carey, V.J., 2013. Software for Computing and Annotating Genomic Ranges. PLoS Comput. Biol. 9, e1003118. https://doi.org/10.1371/journal.pcbi.1003118

Leddin, M., Perrod, C., Hoogenkamp, M., Ghani, S., Assi, S., Heinz, S., Wilson, N.K., Follows, G., Schönheit, J., Vockentanz, L., Mosammam, A.M., Chen, W., Tenen, D.G., Westhead, D.R., Göttgens, B., Bonifer, C., Rosenbauer, F., 2011. Two distinct auto-regulatory loops operate at the PU.1 locus in B cells and myeloid cells. Blood 117, 2827–2838. https://doi.org/10.1182/blood-2010-08-302976

Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., Liu, X.S., 2014. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. 15, 554. https://doi.org/10.1186/s13059-014-0554-4

Li, Y., Okuno, Y., Zhang, P., Radomska, H.S., Chen, H., Iwasaki, H., Akashi, K., Klemsz, M.J., McKercher, S.R., Maki, R.A., Tenen, D.G., 2001. Regulation of the PU.1 gene by distal elements. Blood 98, 2958–2965. https://doi.org/10.1182/blood.V98.10.2958

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., Mesirov, J.P., 2011. Molecular signatures database (MSigDB) 3.0. Bioinformatics 27, 1739–1740. https://doi.org/10.1093/bioinformatics/btr260

Lin, T., Liu, G.A., Perez, E., Rainer, R.D., Febo, M., Cruz-Almeida, Y., Ebner, N.C., 2018. Systemic Inflammation Mediates Age-Related Cognitive Deficits. Front. Aging Neurosci. 10, 236. https://doi.org/10.3389/fnagi.2018.00236

Liu, F., Vermesh, O., Mani, V., Ge, T.J., Madsen, S.J., Sabour, A., Hsu, E.-C., Gowrishankar, G., Kanada, M., Jokerst, J.V., Sierra, R.G., Chang, E., Lau, K., Sridhar, K., Bermudez, A., Pitteri, S.J., Stoyanova, T., Sinclair, R., Nair, V.S., Gambhir, S.S., Demirci, U., 2017. The Exosome Total Isolation Chip. ACS Nano 11, 10712–10723. https://doi.org/10.1021/acsnano.7b04878

Liu, H., Golji, J., Brodeur, L.K., Chung, F.S., Chen, J.T., deBeaumont, R.S., Bullock, C.P., Jones, M.D., Kerr, G., Li, L., Rakiec, D.P., Schlabach, M.R., Sovath, S., Growney, J.D., Pagliarini, R.A., Ruddy, D.A., MacIsaac, K.D., Korn, J.M., McDonald, E.R., 2019. Tumor-derived IFN triggers chronic pathway agonism and sensitivity to ADAR loss. Nat. Med. 25, 95–102. https://doi.org/10.1038/s41591-018-0302-5

Liu, S.J., Horlbeck, M.A., Cho, S.W., Birk, H.S., Malatesta, M., He, D., Attenello, F.J., Villalta, J.E., Cho, M.Y., Chen, Y., Mandegar, M.A., Olvera, M.P., Gilbert, L.A., Conklin, B.R., Chang, H.Y., Weissman, J.S., Lim, D.A., 2017. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science 355, eaah7111. https://doi.org/10.1126/science.aah7111

Liu, Y., Cao, Z., Wang, Y., Guo, Y., Xu, P., Yuan, P., Liu, Z., He, Y., Wei, W., 2018. Genome-wide screening for functional long noncoding RNAs in human cells by Cas9 targeting of splice sites. Nat. Biotechnol. 36, 1203–1210. https://doi.org/10.1038/nbt.4283

Lopetuso, L., Gerardi, V., Papa, V., Scaldaferri, F., Rapaccini, G., Gasbarrini, A., Papa, A., 2017. Can We Predict the Efficacy of Anti-TNF-α Agents? Int. J. Mol. Sci. 18, 1973. https://doi.org/10.3390/ijms18091973

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lu, X., Peled, N., Greer, J., Wu, W., Choi, P., Berger, A.H., Wong, S., Jen, K.-Y., Seo, Y., Hann, B., Brooks, A., Meyerson, M., Collisson, E.A., 2017. *MET* Exon 14 Mutation Encodes an Actionable Therapeutic Target in Lung Adenocarcinoma. Cancer Res. 77, 4498–4505. https://doi.org/10.1158/0008-5472.CAN-16-1944

Lundberg, A.S., Randell, S.H., Stewart, S.A., Elenbaas, B., Hartwell, K.A., Brooks, M.W., Fleming, M.D., Olsen, J.C., Miller, S.W., Weinberg, R.A., Hahn, W.C., 2002. Immortalization and transformation of primary human airway epithelial cells by gene transfer. Oncogene 21, 4577–4586. https://doi.org/10.1038/sj.onc.1205550

Ma, X., Xu, S., 2013. TNF inhibitor therapy for rheumatoid arthritis. Biomed. Rep. 1, 177–184. https://doi.org/10.3892/br.2012.42

MacMicking, J.D., 2012. Interferon-inducible effector mechanisms in cell-autonomous immunity. Nat. Rev. Immunol. 12, 367–382. https://doi.org/10.1038/nri3210

Malekos, E., Carpenter, S., 2022. Short open reading frame genes in innate immunity: from discovery to characterization. Trends Immunol. 43, 741–756. https://doi.org/10.1016/j.it.2022.07.005

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10. https://doi.org/10.14806/ej.17.1.200

Mayr, C., 2017. Regulation by 3′-Untranslated Regions. Annu. Rev. Genet. 51, 171–194. https://doi.org/10.1146/annurev-genet-120116-024704

Micheletti, R., Plaisance, I., Abraham, B.J., Sarre, A., Ting, C.-C., Alexanian, M., Maric, D., Maison, D., Nemir, M., Young, R.A., Schroen, B., González, A., Ounzain, S., Pedrazzini, T., 2017. The long noncoding RNA *Wisper* controls cardiac fibrosis and remodeling. Sci. Transl. Med. 9, eaai9118. https://doi.org/10.1126/scitranslmed.aai9118

Napoli, M., Li, X., Ackerman, H.D., Deshpande, A.A., Barannikov, I., Pisegna, M.A., Bedrosian, I., Mitsch, J., Quinlan, P., Thompson, A., Rajapakshe, K., Coarfa, C., Gunaratne, P.H., Marchion, D.C., Magliocco, A.M., Tsai, K.Y., Flores, E.R., 2020. Pan-cancer analysis reveals TAp63-regulated oncogenic lncRNAs that promote cancer progression through AKT

activation. Nat. Commun. 11, 5156. https://doi.org/10.1038/s41467-020-18973-w

Neumann, M., Naumann, M., 2007. Beyond IκBs: alternative regulation of NF-KB activity. FASEB J. 21, 2642–2654. https://doi.org/10.1096/fj.06-7615rev

Nystrom, S.L., McKay, D.J., 2021. Memes: A motif analysis environment in R using tools from the MEME Suite. PLOS Comput. Biol. 17, e1008991. https://doi.org/10.1371/journal.pcbi.1008991

Oeckinghaus, A., Ghosh, S., 2009. The NF- B Family of Transcription Factors and Its Regulation. Cold Spring Harb. Perspect. Biol. 1, a000034–a000034. https://doi.org/10.1101/cshperspect.a000034

Okuno, Y., Huang, G., Rosenbauer, F., Evans, E.K., Radomska, H.S., Iwasaki, H., Akashi, K., Moreau-Gachelin, F., Li, Y., Zhang, P., Göttgens, B., Tenen, D.G., 2005. Potential Autoregulation of Transcription Factor PU.1 by an Upstream Regulatory Element. Mol. Cell. Biol. 25, 2832–2845. https://doi.org/10.1128/MCB.25.7.2832-2845.2005

Open2C, Abdennur, N., Fudenberg, G., Flyamer, I.M., Galitsyna, A.A., Goloborodko, A., Imakaev, M., Venev, S.V., 2023. Pairtools: from sequencing data to chromosome contacts (preprint). Bioinformatics. https://doi.org/10.1101/2023.02.13.528389

Park, R.J., Wang, T., Koundakjian, D., Hultquist, J.F., Lamothe-Molina, P., Monel, B., Schumann, K., Yu, H., Krupzcak, K.M., Garcia-Beltran, W.,

Piechocka-Trocha, A., Krogan, N.J., Marson, A., Sabatini, D.M., Lander, E.S., Hacohen, N., Walker, B.D., 2017. A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. Nat. Genet. 49, 193–203. https://doi.org/10.1038/ng.3741

Pearson, T.A., Mensah, G.A., Alexander, R.W., Anderson, J.L., Cannon, R.O., Criqui, M., Fadl, Y.Y., Fortmann, S.P., Hong, Y., Myers, G.L., Rifai, N., Smith, S.C., Taubert, K., Tracy, R.P., Vinicor, F., 2003. Markers of Inflammation and Cardiovascular Disease: Application to Clinical and Public Health Practice: A Statement for Healthcare Professionals From the Centers for Disease Control and Prevention and the American Heart Association. Circulation 107, 499–511. https://doi.org/10.1161/01.CIR.0000052939.59093.45

Peschon, J.J., Torrance, D.S., Stocking, K.L., Glaccum, M.B., Otten, C., Willis, C.R., Charrier, K., Morrissey, P.J., Ware, C.B., Mohler, K.M., 1998. TNF receptor-deficient mice reveal divergent roles for p55 and p75 in several models of inflammation. J. Immunol. Baltim. Md 1950 160, 943–952.

Peyrin-Biroulet, L., 2010. Anti-TNF therapy in inflammatory bowel diseases: a huge review. Minerva Gastroenterol. Dietol. 56, 233–243.

Phanstiel, D.H., Van Bortle, K., Spacek, D., Hess, G.T., Shamim, M.S., Machol, I., Love, M.I., Aiden, E.L., Bassik, M.C., Snyder, M.P., 2017. Static and Dynamic DNA Loops form AP-1-Bound Activation Hubs during

Macrophage Development. Mol. Cell 67, 1037-1048.e6. https://doi.org/10.1016/j.molcel.2017.08.006

Pierce, J.B., Zhou, H., Simion, V., Feinberg, M.W., 2022. Long Noncoding RNAs as Therapeutic Targets, in: Carpenter, S. (Ed.), Long Noncoding RNA, Advances in Experimental Medicine and Biology. Springer International Publishing, Cham, pp. 161–175. https://doi.org/10.1007/978-3-030-92034-0_9

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., Manke, T., 2016. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44, W160–W165. https://doi.org/10.1093/nar/gkw257

Reggiardo, R.E., Maroli, S.V., Kim, D.H., 2022. LncRNA Biomarkers of Inflammation and Cancer, in: Carpenter, S. (Ed.), Long Noncoding RNA, Advances in Experimental Medicine and Biology. Springer International Publishing, Cham, pp. 121–145. https://doi.org/10.1007/978-3-030-92034-0_7

Renz, H., von Mutius, E., Brandtzaeg, P., Cookson, W.O., Autenrieth, I.B., Haller, D., 2011. Gene-environment interactions in chronic inflammatory disease. Nat. Immunol. 12, 273–277. https://doi.org/10.1038/ni0411-273

Rinn, J.L., Chang, H.Y., 2020. Long Noncoding RNAs: Molecular Modalities to Organismal Functions. Annu. Rev. Biochem. 89, 283–308. https://doi.org/10.1146/annurev-biochem-062917-012708

Robinson, E.K., Covarrubias, S., Carpenter, S., 2020. The how and why of lncRNA function: An innate immune perspective. Biochim. Biophys. Acta BBA - Gene Regul. Mech. 1863, 194419. https://doi.org/10.1016/j.bbagrm.2019.194419

Roulois, D., Loo Yau, H., Singhania, R., Wang, Y., Danesh, A., Shen, S.Y., Han, H., Liang, G., Jones, P.A., Pugh, T.J., O'Brien, C., De Carvalho, D.D., 2015. DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. Cell 162, 961–973. https://doi.org/10.1016/j.cell.2015.07.056

Roy, D., Tiirikainen, M., 2020. Diagnostic Power of DNA Methylation Classifiers for Early Detection of Cancer. Trends Cancer 6, 78–81. https://doi.org/10.1016/j.trecan.2019.12.006

Ruiz-Orera, J., Villanueva-Cañas, J.L., Albà, M.M., 2020. Evolution of new proteins from translated sORFs in long non-coding RNAs. Exp. Cell Res. 391, 111940. https://doi.org/10.1016/j.yexcr.2020.111940

Salama, S.R., 2022. The Complexity of the Mammalian Transcriptome, in: Carpenter, S. (Ed.), Long Noncoding RNA, Advances in Experimental Medicine and Biology. Springer International Publishing, Cham, pp. 11–22. https://doi.org/10.1007/978-3-030-92034-0_2

Sanbonmatsu, K., 2022. Towards Molecular Mechanism in Long Non-coding RNAs: Linking Structure and Function, in: Carpenter, S. (Ed.), Long Noncoding RNA, Advances in Experimental Medicine and Biology.

Springer International Publishing, Cham, pp. 23–32. https://doi.org/10.1007/978-3-030-92034-0_3

Sato, M., Larsen, J.E., Lee, W., Sun, H., Shames, D.S., Dalvi, M.P., Ramirez, R.D., Tang, H., DiMaio, J.M., Gao, B., Xie, Y., Wistuba, I.I., Gazdar, A.F., Shay, J.W., Minna, J.D., 2013. Human Lung Epithelial Cells Progressed to Malignancy through Specific Oncogenic Manipulations. Mol. Cancer Res. 11, 638–650. https://doi.org/10.1158/1541-7786.MCR-12-0634-T

Schmid-Burgk, J.L., Chauhan, D., Schmidt, T., Ebert, T.S., Reinhardt, J., Endl, E., Hornung, V., 2016. A Genome-wide CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) Screen Identifies NEK7 as an Essential Component of NLRP3 Inflammasome Activation. J. Biol. Chem. 291, 103–109. https://doi.org/10.1074/jbc.C115.700492

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., Zhang, F., 2014. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science 343, 84–87. https://doi.org/10.1126/science.1247005

She, R., Luo, J., Weissman, J.S., 2023. Translational fidelity screens in mammalian cells reveal eIF3 and eIF4G2 as regulators of start codon selectivity. Nucleic Acids Res. 51, 6355–6369. https://doi.org/10.1093/nar/gkad329

Simanshu, D.K., Nissley, D.V., McCormick, F., 2017. RAS Proteins and Their Regulators in Human Disease. Cell 170, 17–33. https://doi.org/10.1016/j.cell.2017.06.009

Slack, F.J., Chinnaiyan, A.M., 2019. The Role of Non-coding RNAs in Oncology. Cell 179, 1033–1055. https://doi.org/10.1016/j.cell.2019.10.017

Soneson, C., Love, M.I., Robinson, M.D., 2016. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research 4, 1521. https://doi.org/10.12688/f1000research.7563.2

Spitz, F., Furlong, E.E.M., 2012. Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. 13, 613–626. https://doi.org/10.1038/nrg3207

Su, X., Yu, Y., Zhong, Y., Giannopoulou, E.G., Hu, X., Liu, H., Cross, J.R., Rätsch, G., Rice, C.M., Ivashkiv, L.B., 2015. Interferon-γ regulates cellular metabolism and mRNA translation to potentiate macrophage activation. Nat. Immunol. 16, 838–849. https://doi.org/10.1038/ni.3205

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., Mesirov, J.P., 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. 102, 15545–15550. https://doi.org/10.1073/pnas.0506580102

The Cancer Genome Atlas Research Network, 2014. Comprehensive molecular profiling of lung adenocarcinoma. Nature 511, 543–550. https://doi.org/10.1038/nature13385

Therneau, T., others, 2015. A package for survival analysis in S. R Package Version 2.

Trinh, B.Q., Ummarino, S., Zhang, Y., Ebralidze, A.K., Bassal, M.A., Nguyen, T.M., Heller, G., Coffey, R., Tenen, D.E., van der Kouwe, E., Fabiani, E., Gurnari, C., Wu, C.-S., Angarica, V.E., Yang, H., Chen, S., Zhang, H., Thurm, A.R., Marchi, F., Levantini, E., Staber, P.B., Zhang, P., Voso, M.T., Pandolfi, P.P., Kobayashi, S.S., Chai, L., Di Ruscio, A., Tenen, D.G., 2021. Myeloid lncRNA *LOUP* mediates opposing regulatory effects of RUNX1 and RUNX1-ETO in t(8;21) AML. Blood 138, 1331–1344. https://doi.org/10.1182/blood.2020007920

Turkistany, S.A., DeKoter, R.P., 2011. The Transcription Factor PU.1 is a Critical Regulator of Cellular Communication in the Immune System. Arch. Immunol. Ther. Exp. (Warsz.) 59, 431–440. https://doi.org/10.1007/s00005-011-0147-9

Tzelepis, K., Koike-Yusa, H., De Braekeleer, E., Li, Y., Metzakopian, E., Dovey, O.M., Mupo, A., Grinkevich, V., Li, M., Mazan, M., Gozdecka, M., Ohnishi, S., Cooper, J., Patel, M., McKerrell, T., Chen, B., Domingues, A.F., Gallipoli, P., Teichmann, S., Ponstingl, H., McDermott, U., Saez-Rodriguez, J., Huntly, B.J.P., Iorio, F., Pina, C., Vassiliou, G.S., Yusa, K.,

2016. A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. Cell Rep. 17, 1193–1205. https://doi.org/10.1016/j.celrep.2016.09.079

Ulitsky, I., Bartel, D.P., 2013. lincRNAs: Genomics, Evolution, and Mechanisms. Cell 154, 26–46. https://doi.org/10.1016/j.cell.2013.06.020

Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., Bartel, D.P., 2011. Conserved Function of lincRNAs in Vertebrate Embryonic Development despite Rapid Sequence Evolution. Cell 147, 1537–1550. https://doi.org/10.1016/j.cell.2011.11.055

Uszczynska-Ratajczak, B., Lagarde, J., Frankish, A., Guigó, R., Johnson, R., 2018. Towards a complete map of the human long non-coding RNA transcriptome. Nat. Rev. Genet. 19, 535–548. https://doi.org/10.1038/s41576-018-0017-y

Verhoeckx, K., Cotter, P., López-Expósito, I., Kleiveland, C., Lea, T., Mackie, A., Requena, T., Swiatecka, D., Wichers, H. (Eds.), 2015. The Impact of Food Bioactives on Health. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-16104-4

Verma, I.M., 2004. Nuclear factor (NF)- B proteins: therapeutic targets. Ann. Rheum. Dis. 63, ii57–ii61. https://doi.org/10.1136/ard.2004.028266

Vikis, H., Sato, M., James, M., Wang, D., Wang, Y., Wang, M., Jia, D., Liu, Y., Bailey-Wilson, J.E., Amos, C.I., Pinney, S.M., Petersen, G.M., de Andrade, M., Yang, P., Wiest, J.S., Fain, P.R., Schwartz, A.G., Gazdar,

A., Gaba, C., Rothschild, H., Mandal, D., Kupert, E., Seminara, D., Viswanathan, A., Govindan, R., Minna, J., Anderson, M.W., You, M., 2007. *EGFR-T790M* Is a Rare Lung Cancer Susceptibility Allele with Enhanced Kinase Activity. Cancer Res. 67, 4665–4670. https://doi.org/10.1158/0008-5472.CAN-07-0217

Vollmers, A.C., Mekonen, H.E., Campos, S., Carpenter, S., Vollmers, C., 2021. Generation of an isoform-level transcriptome atlas of macrophage activation. J. Biol. Chem. 296, 100784. https://doi.org/10.1016/j.jbc.2021.100784

Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., Sabatini, D.M., 2015. Identification and characterization of essential genes in the human genome. Science 350, 1096–1101. https://doi.org/10.1126/science.aac7041

Wang, T., Wei, J.J., Sabatini, D.M., Lander, E.S., 2014. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Science 343, 80–84. https://doi.org/10.1126/science.1246981

Wang, W.-T., Han, C., Sun, Y.-M., Chen, T.-Q., Chen, Y.-Q., 2019. Noncoding RNAs in cancer therapy resistance and targeted drug development. J. Hematol. Oncol.J Hematol Oncol 12, 55. https://doi.org/10.1186/s13045-019-0748-z

Wijesinghe, S.N., Lindsay, M.A., Jones, S.W., 2022. Long Non-coding RNAs in Rheumatology, in: Carpenter, S. (Ed.), Long Noncoding RNA, Advances

in Experimental Medicine and Biology. Springer International Publishing, Cham, pp. 35–70. https://doi.org/10.1007/978-3-030-92034-0_4

Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., Knox, C., Wilson, M., 2018. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 46, D1074–D1082. https://doi.org/10.1093/nar/gkx1037

Wright, B.W., Yi, Z., Weissman, J.S., Chen, J., 2022. The dark proteome: translation from noncanonical open reading frames. Trends Cell Biol. 32, 243–258. https://doi.org/10.1016/j.tcb.2021.10.010

Wynn, T.A., Chawla, A., Pollard, J.W., 2013. Macrophage biology in development, homeostasis and disease. Nature 496, 445–455. https://doi.org/10.1038/nature12034

Xue, C., Zhang, X., Zhang, H., Ferguson, J.F., Wang, Y., Hinkle, C.C., Li, M., Reilly, M.P., 2017. De novo RNA sequence assembly during in vivo inflammatory stress reveals hundreds of unannotated lincRNAs in human blood CD14 $^{+}$ monocytes and in adipose tissue. Physiol. Genomics 49, 287–305. https://doi.org/10.1152/physiolgenomics.00001.2017

Yang, S., Wang, J., Brand, D.D., Zheng, S.G., 2018. Role of TNF–TNF Receptor 2 Signal in Regulatory T Cells and Its Therapeutic Implications. Front. Immunol. 9, 784. https://doi.org/10.3389/fimmu.2018.00784

Yang, X., Bam, M., Becker, W., Nagarkatti, P.S., Nagarkatti, M., 2020. Long Noncoding RNA AW112010 Promotes the Differentiation of Inflammatory T Cells by Suppressing IL-10 Expression through Histone Demethylation. J. Immunol. 205, 987–993. https://doi.org/10.4049/jimmunol.2000330

Yeung, A.T.Y., Choi, Y.H., Lee, A.H.Y., Hale, C., Ponstingl, H., Pickard, D., Goulding, D., Thomas, M., Gill, E., Kim, J.K., Bradley, A., Hancock, R.E.W., Dougan, G., 2019. A Genome-Wide Knockout Screen in Human Macrophages Identified Host Factors Modulating *Salmonella* Infection. mBio 10, e02169-19. https://doi.org/10.1128/mBio.02169-19

Yimin, Kohanawa, M., 2006. A Regulatory Effect of the Balance between TNF-α and IL-6 in the Granulomatous and Inflammatory Response to *Rhodococcus aurantiacus* Infection in Mice. J. Immunol. 177, 642–650. https://doi.org/10.4049/jimmunol.177.1.642

Yuan, L., Xu, Z.-Y., Ruan, S.-M., Mo, S., Qin, J.-J., Cheng, X.-D., 2020. Long non-coding RNAs towards precision medicine in gastric cancer: early diagnosis, treatment, and drug resistance. Mol. Cancer 19, 96. https://doi.org/10.1186/s12943-020-01219-0

Zhang, H., Song, L., Wang, X., Cheng, H., Wang, C., Meyer, C.A., Liu, T., Tang, M., Aluru, S., Yue, F., Liu, X.S., Li, H., 2021. Fast alignment and preprocessing of chromatin profiles with Chromap. Nat. Commun. 12, 6566. https://doi.org/10.1038/s41467-021-26865-w

Zhang, X., Chen, X., Liu, Q., Zhang, S., Hu, W., 2017. Translation repression via modulation of the cytoplasmic poly(A)-binding protein in the inflammatory response. eLife 6, e27786. https://doi.org/10.7554/eLife.27786

Zheng, C., Wei, Y., Zhang, P., Xu, L., Zhang, Z., Lin, K., Hou, J., Lv, X., Ding, Y., Chiu, Y., Jain, A., Islam, N., Malovannaya, A., Wu, Y., Ding, F., Xu, H., Sun, M., Chen, X., Chen, Y., 2023. CRISPR/Cas9 screen uncovers functional translation of cryptic lncRNA-encoded open reading frames in human cancer. J. Clin. Invest. 133, e159940. https://doi.org/10.1172/JCI159940

Zhu, A., Ibrahim, J.G., Love, M.I., 2019. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics 35, 2084–2092. https://doi.org/10.1093/bioinformatics/bty895

Zhu, S., Li, W., Liu, J., Chen, C.-H., Liao, Q., Xu, P., Xu, H., Xiao, T., Cao, Z., Peng, J., Yuan, P., Brown, M., Liu, X.S., Wei, W., 2016. Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR–Cas9 library. Nat. Biotechnol. 34, 1279–1286. https://doi.org/10.1038/nbt.3715

Zinatizadeh, M.R., Schock, B., Chalbatani, G.M., Zarandi, P.K., Jalali, S.A., Miri, S.R., 2021. The Nuclear Factor Kappa B (NF-kB) signaling in cancer

development and immune diseases. Genes Dis. 8, 287–297. https://doi.org/10.1016/j.gendis.2020.06.005

Zuo, Prather, Stetskiv, Garrison, Meade, Peace, Zhou, 2019. Inflammaging and Oxidative Stress in Human Diseases: From Molecular Mechanisms to Novel Treatments. Int. J. Mol. Sci. 20, 4472. https://doi.org/10.3390/ijms20184472