Oncogene-driven post-transcriptional regulation in lung cancer

April Lo

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Alice Berger, Chair

David MacPherson

Robert Bradley

Program Authorized to Offer Degree:

Genome Sciences

University of Washington

Abstract

Oncogene-driven post-transcriptional regulation in lung cancer

April Lo

Chair of the Supervisory Committee:
Alice H. Berger
Genome Sciences

Lung cancer is often caused by genetic mutations which alter the activity of proteins in the RAS family, with somatic mutation of the *KRAS* gene occurring in ~30% of lung adenocarcinoma tumors. A gene of the same family, *RIT1*, is mutated or amplified in 10-15% of lung adenocarcinomas. Although therapies targeting some *KRAS*-mutated tumors have begun to enter the clinic, a comprehensive understanding of the Ras family and their regulatory functions continues to elude us, preventing effective treatments and complete cures for all Ras-driven lung cancers. I employed high-throughput multi-omic methods to interrogate the regulation of gene expression and protein abundance in oncogenic *KRAS* and *RIT1* cells. Quantitative proteomic and transcriptomic profiling revealed that both mutant KRAS and mutant RIT1 promoted canonical Ras signaling, and overexpression of wild-type RIT1 resulted in phenotypes similar to those of oncogenic RIT1 and KRAS, including induction of epithelial-to-mesenchymal transition. Thus,

RIT1 amplifications which result in increased gene expression may be tumorigenic. Then, with quantitative phosphoproteomics, I found that KRAS variant cells differentially phosphorylated SR protein phosphosites. To determine how KRAS may regulate alternative splicing, I analyzed unique RNA sequencing profiles of lung adenocarcinoma cells ectopically expressing 75 different wild-type or variant alleles across 28 genes implicated in lung cancer. Mutant KRAS induced the greatest number of differential alternative splicing events when compared to the wild-type gene, second only to the differential splicing between RNA binding protein RBM45 and its mutant $RBM45^{M126I}$. These data suggest that in addition to widespread transcriptional changes, Ras signaling pathways in cancer promote post-transcriptional splicing changes that may contribute to oncogenic processes. In sum, I provided a baseline for how oncogenic KRAS and RIT1 affect the proteome and transcriptome of lung cancer cells, and exposed potential oncogenic mechanisms and therapeutic vulnerabilities in RIT1 or KRAS altered cells.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I could not have completed my Ph.D. at the University of Washington and the Fred Hutchinson Cancer Center without the support of an amazing network of people throughout graduate school and my life.

To my thesis committee, Rob Bradley, David MacPherson, Cole Trapnell, and Daniela Witten, thank you for your invaluable time and advice through the years. From helping direct my projects to providing general career guidance, you taught me to be a more creative and confident scientist. I also thank Rob Bradley and David MacPherson for serving on my reading committee, and the Bradley and MacPherson labs for innumerable insightful conversations and group meetings.

My love of science was nurtured by many great mentors throughout my life. To my primary school teachers Miss King and Mr. Thums, thank you teaching us the joy of experiments and guiding me to earn a school-wide science award. To Dr. Lowery who advised my high school science fair project and fostered a great chemistry classroom, thank you. To Mrs. Carlson, thank you for not only teaching us AP Biology and inciting in me a driving curiosity in molecular genetics, but also for believing in me and advocating for my first research experience. Shawna Guillemette and Sharon Carter, that summer at UMass Medical under your mentorship gave me my first taste of cancer biology and lab research and I've never looked back, thank you. Rachel Karchin, Noushin Niknafs, and Peter Kharchenko thank you for mentoring my undergraduate research experiences and teaching me cancer bioinformatics. And Jean Fan and Lindsey Fernandez, thank you for that summer showing me what a fulfilling life can be like in research.

To my advisor, Alice Berger, it has been an absolute honor learning from you and being your mentee for the past 5 years. Thank you for teaching me everything about lung cancer biology, but more importantly for teaching me how to be a good scientist who enjoys the process and never forgets the real-world and clinical implications for what we do.

To Berger lab members past and present, thank you for being awesome people to learn and work with. James Adams, the lab simply would not function without your always steady hand and assistance with all things administrative. Maria McSharry, Naomi Nkinsi, and Kristin Holmes, the early days were tough but the foundation you set for the lab is indispensable. Athea Vichas and Sitapriya Moorthi, thank you for your mentoring and encouragement from the other side of graduate school. Shriya Kamlapurkar, thank you for your hard work across the board but especially for the RIT1 profiling project. Phoebe Parrish and Amanda Riley, in the high and low times of graduate school, between science conversations and life vents, you both have been the greatest fellow students in the lab. Amanda Bradley and Callie Rominger, the last year with you in the lab has been so fun and fulfilling and I am glad to know the lab will be in good hands. And Akshay Vakharia, I am so proud of how much you learned as a high school intern in less than a year, thank you for reinvigorating our excitement in studying RBM45.

To my Genome Sciences cohort, I feel incredibly lucky to have gone on this journey with you all. From research reports practice to celebrating passing general exams to bar trivia, it has been a delightful time knowing you and watching all of you become amazing scientists. I want to especially thank Cindy Yeh and Andria Ellis for not only being great friends but also spearheading the Genome Hackers camp and letting me lead sessions year after year. I also thank Women in Genome Sciences and Fred Hutch Girls Who Code, groups with which I have learned to create more truly inclusive environments in STEM.

Before graduate school was a formative time in college where I met some of my best friends. Sasha Stevenson, thanks for giving me knitting tips and always having my back. Emily Hendrickson, thank you for our regular video chats and never-ceasing friendship.

My family has always been my greatest source of support. Thank you baba (Harry) and mama (Linda) for providing me with every opportunity I could have asked for, and literally moving across the world and the country to support my education and dreams. To my brother, Hans, you have always been my role model, teaching me to have fun and laugh, and later to code and figure out what I want. I can't wait to get dim sum with everyone again. And to my extended family in Hong Kong and across the world, I miss you all a lot and I hope to see you soon.

Mitchell Vollger, you are my rock and my partner in everything. I cannot imagine going through graduate school and life without you. Thank you, for all of it.

DEDICATION

This work is dedicated to my family:

To my parents who taught me to ask questions of the world around me, who moved countries and states to help me achieve my dreams, and who made sure I always put myself first.

And to my brother who taught me to laugh, to code, and so much more.

Chapter 1. INTRODUCTION

In 2004, cancer biologists brimmed with excitement at the prospect of treating cancers successfully and painlessly with the help of genetic biomarkers and targeted therapies. The first few studies had just been published showing the remarkable, almost complete efficacy of Tyrosine Kinase Inhibitors (TKIs) specifically in lung cancers with mutations in the EGFR gene (Paez et al. 2004; Pao et al. 2004). It seemed possible now to treat all cancers in a targeted manner by matching up drugs to cancers with genetic alterations conferring specific drug sensitivities. This prospect seemed especially within grasp given that, just a year earlier in 2003, the Human Genome Project had announced the completion of the first human reference genome, providing the full map of possible genetic alterations (Lander et al. 2001).

Twenty years later, we now know that successful targeted cancer therapy is more complex. Despite the initial success, EGFR-mutant cancers treated with TKIs invariably develop drug resistance within a year (Passaro et al. 2021). EGFR-mutant tumors also account for just 10% of lung adenocarcinomas. Another 30% are KRAS-mutant, and the remaining are a mix of lower frequency alterations and cases of unknown genetic cause (Figure 1.1A) (Cancer Genome Atlas Research Network 2014). Recently, the first KRAS-mutant inhibitors were introduced into the clinic (Canon et al. 2019; Skoulidis et al. 2021). However, as with TKIs, these inhibitors only delay tumor progression and acquired resistance is a universal occurrence (Moore and Malek 2021). To design truly effective targeted therapies, we need to map all genetic alterations in cancer and, crucially, understand how these alterations contribute to cancer formation and growth.

Figure 1.1. Ras pathway genetic alterations in lung adenocarcinomas (TCGA).

1.1 UNDERSTANDING THE MOLECULAR LANDSCAPE OF LUNG CANCER

Of all types of cancers, lung cancer is the second most prevalent in both men and women. In addition, lung cancer causes the most deaths out of any cancer type (American Cancer Society 2021). While the past two decades have seen these numbers decrease with the help of public health smoking cessation efforts, successful treatment of lung cancer remains out of our grasp.

In this dissertation, I focus on lung adenocarcinomas (LUADs), a histological subtype of Non-Small Cell Lung Cancers (NSCLCs), which make up 40% of all lung cancers. However, while histological subtypes are the predominant classifications used in the clinic, it is just as important to understand the diversity of cancers with a lens on the molecular and genetic differences. The Cancer Genome Atlas (TCGA) is a multi-institutional network that catalogued the cancer and matched normal genomes of over 11,000 patients across 33 cancers. This massive tumor sequencing effort identified many known genetic changes which contribute to cancer. For example, the lung adenocarcinoma study of 230 patient tumors found that 50% had mutations in the tumor suppressor gene *TP53*.

Quantifying the variety of genetic changes in cancers has already brought ground-breaking advances to the clinic. Approximately 50% of lung adenocarcinomas harbor a genetic alteration in a member of the RTK-Ras signaling cascade. Other than the 27% with *KRAS* mutations, this includes alterations in the receptor tyrosine kinases *EGFR*, *MET, ALK,* and *RET*, and mutations in downstream genes *BRAF*, *NF1*, and *P13KCA* (Imielinski et al. 2012). Since the characterization of these mutations, many targeted therapies have been developed and introduced in the clinic to varied success (Mok et al. 2017; Rotow and Bivona 2017).

1.2 Regulating Ras signaling network in cancer

A crucial signaling network in development, the RTK-Ras pathway is frequently co-opted by cancer. This pathway consists of a receptor tyrosine kinase (RTK) protein such as EGFR which, when bound by a growth factor, activates a small guanosine triphosphatases (GTPase) protein of the Ras family. When Ras is in its activated GTP-bound state, it activates the Raf-MEK-ERK kinase cascade and other parallel pathways such as PI3K-AKT signaling (Figure 1.2). These signaling cascades regulate a variety of cellular processes including apoptosis, membrane trafficking, and transcriptional changes promoting cell survival and proliferation (Mukhopadhyay, Vander Heiden, and McCormick 2021).

Cancer mutations in the Ras pathway predominantly promote tumor growth by causing abnormal constitutive activation of one of the pathway members. *KRAS* mutations occur not only in lung cancers but also in 90% of pancreatic cancers and 40% of colorectal cancers, and many of these mutations are at codon 12 including $KRAS^{G12C}$ and $KRAS^{G12D}$ (Kim, Xue, and Lito 2020). Substitutions in codon 12 and 13 allow the mutant protein to persist in a GTP-bound activating state (Hobbs, Der, and Rossman 2016). Although KRAS was considered “undruggable” for decades, $KRAS^{G12C}$ inhibitors have recently been developed and introduced in the clinic to much

excitement. Still, much work must still be done to develop more potent inhibitors, to develop inhibitors for other KRAS variants, and to address the inevitable resistance to the inhibitors (Moore and Malek 2021).

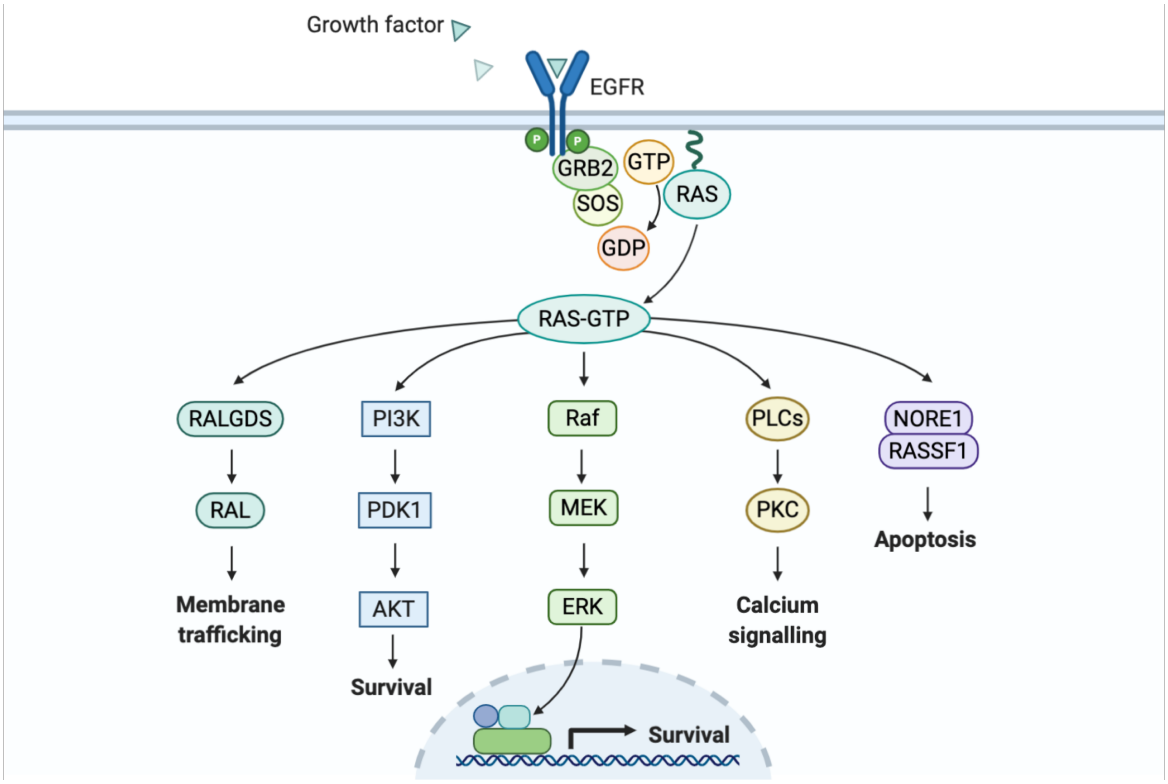


Figure 1.2. The RTK-Ras signaling cascade (BioRender).

Interestingly, cancer genome sequencing also identified mutations in a different small GTPase called Ras-like in all tissues or *RIT1* (A. H. Berger et al. 2014). Mutations in *RIT1* occur in 2% of lung adenocarcinomas and present in a mutually exclusive manner to mutations in other RTK-Ras genes. Unlike *KRAS* however, *RIT1* is also amplified in 10% of lung adenocarcinomas and these copy number alterations include chromosome arm amplifications as well as many focal amplifications around the *RIT1* locus (Cancer Genome Atlas Research Network 2014). Understanding RIT1's role in cancer when it is mutated or amplified will not only allow for the

design of targeted therapies for RIT1-driven tumors, it will also help us better understand the complete Ras signaling network.

In **Chapter 2**, I ask how KRAS and RIT1 differ in their signaling activity and post-transcriptional regulation of lung epithelial cells. To answer this question, I characterize and analyze proteomic, phosphoproteomic, and transcriptomic profiles of cells ectopically expressing wild-type or variant alleles of *KRAS* and *RIT1*.

1.3 ALTERNATIVE RNA SPLICING IN CANCER

One significant outcome of the TCGA effort and subsequent cancer genome studies was the identification of somatic mutations in genes and pathways previously not considered part of the cancer landscape. Among these findings were mutations in genes involved in RNA splicing and processing.

Eukaryotic cells require RNA splicing to produce functional proteins. Even more so, by splicing in or out different introns and exons, cells produce different mRNA and protein isoforms from single genes. This phenomenon, termed alternative splicing, contributes to the diversity and functional complexity of the human proteome. RNA splicing also regulates gene expression because alternatively spliced isoforms can result in antagonistic protein products or mRNA that is subjected to nonsense-mediated decay (Singh and Cooper 2012). The highly regulated process allows different cells to have specialized proteomes to perform functions specific for their tissue type and biological context. However, cancer cells often take advantage of alternative RNA splicing to promote oncogenesis and tumor growth (Escobar-Hoyos et al. 2020).

For example, the splicing factor *U2AF1* is mutated in 4% of LUADs, while *RBM10* is mutated in another 9% (Figure 1.4) (Cancer Genome Atlas Research Network 2014). The S34F/Y mutation in U2AF1 dysregulates the protein's sequence specificity and results in widespread mis-

splicing (Brooks et al. 2014). Similarly, the splicing factor RBM10 functions as a tumor suppressor by regulating the splicing of genes such as the Notch inhibitor *NUMB* (Bechara et al. 2013). Furthermore, pan-cancer studies revealed the prevalence of both splicing factor mutations and splice-site mutations in cancer (Seiler et al. 2018; Jayasinghe et al. 2018).

Many studies have used TCGA data to identify splicing-associated mutations in cancers and these studies reveal key characteristics of cancer-associated aberrant splicing (Dvinge and Bradley 2015). These studies investigated mis-splicing by largely focusing on mutations that directly affect splicing such as splicing factor mutations which are examples of *trans*-acting splicing dysregulation. *Cis*-acting mechanisms such as splice site mutations also cause aberrant splicing of cancer-associated genes. In 4% of LUADs, splice site mutations in the *MET* gene cause the deletion of exon 14 of the gene which leads to hyperactivation of the MET protein and downstream signaling pathways (Cancer Genome Atlas Research Network 2014).

However, we have only just begun to determine what events lead to and what effects stem from aberrant RNA splicing (Climente-González et al. 2017). While a subset of aberrant splicing events result from the aforementioned alterations in splicing factors and splice sites (Dvinge et al. 2016), we do not know whether the remainder of mis-splicing events are tightly regulated events which contribute to tumor growth or are simply a result of the high transcriptional and proliferation rates in cancer cells. Motivated by previous studies which identified individual cases where alternative splicing of a protein occurs downstream of signaling pathways (Muñoz et al. 2012; Zhou et al. 2012; Yea et al. 2008), my work takes an oncogene screening approach to probe whether such regulation is widespread across proteins and regulatory networks, and whether splicing regulation occurs through Ras signaling in particular. By elucidating the drivers and effects of splicing dysregulation, we can better understand the role of splicing in cancer.

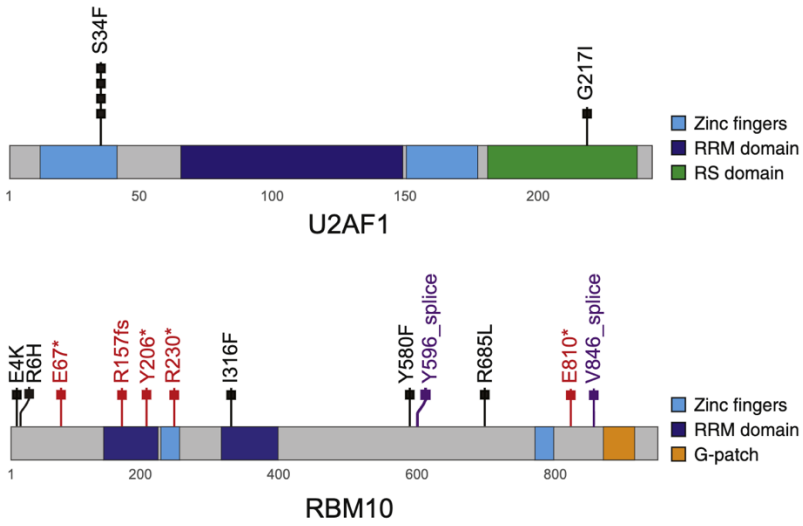


Figure 1.3. Splicing factor mutations in lung adenocarcinomas (Imielinski et al., 2012).

Targeting aberrant alternative RNA splicing can reduce disease burden. Recent *in vivo* work has shown that spliceosome inhibitors can reduce tumor burden in leukemia as assessed by survival benefit in mice with a splicing factor mutation (Lee et al. 2016). Therapies can also be developed to address specific splicing events. In spinal muscular atrophy patients, the functional isoform of a survival motor neuron *SMN2* is deficient. To address this deficiency, a splice-switching antisense oligonucleotide (ASO) therapy was developed to ensure the functional isoform is completely transcribed and this therapy has encouraging clinical efficacy (Finkel et al. 2016). However, ASOs can be unstable and efficient delivery of oligonucleotide-based therapies to target organs remains a significant challenge.

Splice-switching ASOs may not be the only therapeutic option for aberrant alternative splicing. In lung cancer patients whose tumors harbor the *MET* exon 14 exon skipping event, TKIs provide clinical benefit (Lu et al. 2017; Frampton et al. 2015). Understanding the relationship between signaling pathways and RNA splicing will pave the way for development of similar therapeutics to target either signaling components or splicing events.

In **Chapter 3**, I use a large-scale high-throughput RNA-seq screen combined with phosphoproteomic analyses to ask how KRAS regulates RNA splicing and processing activity in cells and how KRAS differs in these activities from other genes associated with lung cancer. By using a controlled experimental system, I functionally associate pathway changes with alternative splicing in lung adenocarcinoma cells.

1.4 GOALS OF THIS DISSERTATION

In the following chapters, I present and discuss my work interrogating the regulatory networks in lung cancer cells using transcriptomic, genomic, and proteomic profiling and analyses. Although the questions are anchored in observed oncogenic genetic mutations, it is the resulting post-transcriptional modulation of protein abundance and function that I seek to understand.

In characterizing RIT1 signaling in **Chapter 2**, I look to how RIT1 and KRAS levels and mutations affect protein abundance, and how these changes in turn drive tumorigenic processes such as epithelial-mesenchymal transition (EMT).

In profiling aberrant RNA splicing due to KRAS activity in **Chapter 3**, I ask how Ras signaling can moderate RNA isoform levels and thus protein isoforms and protein abundance.

Finally, in **Chapter 4**, I discuss the implications of identifying mechanisms of post-transcriptional regulation in lung cancers and where I see the future of targeted cancer therapies.

Chapter 2. CHARACTERIZATION OF RIT1 SIGNALING BY PROTEOME, PHOSPHOPROTEOME, AND TRANSCRIPTOME PROFILING

A version of this chapter has been published:

A. Lo, K. Holmes, S. Kamlapurkar, F. Mundt, S. Moorthi, I. Fung, S. Fereshetian, J. Watson, S. A. Carr, P. Mertins, A. H. Berger, Multiomic characterization of oncogenic signaling mediated by wild-type and mutant RIT1. *Sci. Signal.* **14**, eabc4520 (2021).
http://dx.doi.org/10.1126/scisignal.abc4520

I led this work and my contributions to this paper included performing experiments, analyzing transcriptome data, and writing the manuscript.

Aberrant activation of the RAS family of guanosine triphosphatases (GTPases) is prevalent in lung adenocarcinoma, with somatic mutation of *KRAS* occurring in ~30% of tumors. We previously identified somatic mutations of the gene encoding RAS family GTPase RIT1 in lung adenocarcinomas. To explore the biological pathways regulated by RIT1 and how these relate to the oncogenic KRAS network, we performed quantitative proteomic, phosphoproteomic, and transcriptomic profiling of isogenic lung epithelial cells in which we ectopically expressed wild-type or cancer-associated variants of RIT1 and KRAS. We found that both mutant KRAS and mutant RIT1 promoted canonical RAS signaling, and that overexpression of wild-type RIT1 partially phenocopied that of oncogenic RIT1 and KRAS, including induction of epithelial-to-mesenchymal transition. Our findings suggest that RIT1 protein abundance is key to its pathogenic function. Therefore, chromosomal amplification of wild-type *RIT1* in lung and other cancers may be tumorigenic.

2.1 INTRODUCTION

The RAS-family GTPase-encoding gene *RIT1* was found to harbor somatic mutations in lung cancer (A. H. Berger et al. 2014) and myeloid leukemias (Gómez-Seguí et al. 2013). Disrupted regulation of RAS family genes and RAS effector pathways is a predominant feature of many human cancers. In particular, somatic mutation of the *KRAS* proto-oncogene is prevalent in lung adenocarcinoma, where *KRAS* is mutated in up to 30% of tumors. Cancer-associated KRAS variants include G12V and Q61H, which alter the normal regulation of KRAS GTPase activity by disrupting GTP hydrolysis or reducing the RAS protein's ability to physically associate with GTPase-activating proteins (GAPs) (Trahey and McCormick 1987; Emil F. Pai, Wolfgang Kabsch, Ute Krengel, Kenneth C. Holmes, Jacob John & Alfred Wittinghofer 1989). This uncontrolled increase in GTP-bound KRAS results in heightened downstream cellular signaling through the canonical RAS effector pathways RAF–MEK and PI3K–AKT, as well as others. Following the discovery of cancer-associated RAS mutations in the 1980s (Nakano et al. 1984; Rodenhuis et al. 1987), thousands of studies have delineated the critical pathways involved in RAS-mediated cellular transformation, metastasis, and metabolism.

Notably, in addition to somatic mutations in cancer, germline *RIT1* mutations are found in families with Noonan syndrome, a developmental "RAS"-opathy involving altered craniofacial morphology and cardiac abnormalities (Aoki et al. 2013), that is caused by germline mutations in *KRAS* itself or other RAS-pathway genes such as *SOS1, SOS2, LZTR1,* and *SHOC2* (https://omim.org/). In cancer and Noonan syndrome, *RIT1* mutations are found in tumors or individuals that lack canonical *KRAS* mutations, suggesting that activated RIT1 may impart the same phenotypes conferred by activation of RAS.

Although prior studies have characterized the role of RIT1 in neural development (Cai et al. 2012) and we and others have described the role of mutant RIT1 in cellular transformation (A. H. Berger et al. 2014; Rusyn et al. 2000; Sakabe et al. 2002), knowledge of the function of cancer- and Noonan-associated RIT1 variants is relatively limited. Unlike in KRAS, mutations in RIT1 are rarely observed near the critical glycine residues involved in GTP hydrolysis, which are $Gly^{12/13}$ in KRAS and $Gly^{30/31}$ in RIT1. Instead, RIT1 mutations occur most frequently near the switch II domain, the region also altered by Gln^{61} mutations in KRAS (**Figure 2.1A**). Gln^{61} and $Gly^{12/13}$ mutations in RAS reduce association with regulatory GAPs or alter intrinsic hydrolysis rate (Smith, Neel, and Ikura 2013). It is possible these differences in mutational spectrum reflect functional differences in regulation of RIT1 and KRAS GTPase activity, and no GAPs that regulate RIT1 have been identified to date.

Whereas the regulation of enzymatic activity is thought to be the predominant determinant of RAS oncogenic function, both protein abundance and GTP binding may be key to RIT1 function. Biochemical assays have established that oncogenic RIT1 variants do enhance GTP-bound levels of RIT1 to varying degrees (Castel et al. 2019; Fang et al. 2016). However, the molecular consequences of RIT1 switch II domain mutations may additionally be linked to the loss of RIT1's physical interaction with LZTR1, a ubiquitin-conjugating enzyme responsible for stimulating degradation of RIT1 (Castel et al. 2019). Cancer- and Noonan-associated RIT1 variants lose the ability to interact with LZTR1 and consequently are highly overexpressed, resulting in increased signaling activity through the RAF–MEK pathway (Castel et al. 2019). Because germline loss-of-function mutations in *LZTR1* result in increased abundance of wild-type RIT1 protein and also confer Noonan syndrome, overexpression of RIT1 in the absence of mutation may be sufficient to drive altered RIT1 signaling.

The pathogenic role of increased *RIT1* expression in *LZTR1*-mutant Noonan syndrome raises the possibility that wild-type *RIT1* overexpression in cancer may also play a pathogenic role. Indeed, 7-14% of lung adenocarcinomas (TCGA 2014; Campbell et al. 2016) harbor amplification of the *RIT1* gene on 1p21.3 which is typically associated with overexpression of *RIT1* mRNA. Amplification of *RIT1* was found in an independent lung adenocarcinoma genomic study of samples from the Fudan University Shanghai Cancer Center (Chen et al. 2019). Focal *RIT1* amplifications have also been noted in uterine carcinosarcoma (Cherniack et al. 2017), and *RIT1* amplification and/or overexpression may play a role in hepatocellular carcinoma and endometrial cancer (Song et al. 2019; Xu et al. 2015).

Based on RIT1's homology to RAS proteins, candidate effectors of RIT1 function have been identified (Shao and Andres 2000; Shi and Andres 2005; Castel et al. 2019, 2020) but no unbiased mapping of downstream RIT1-regulated pathways in human cells has been performed to date. Here, we sought to broadly describe the changes in the cellular proteome, phosphoproteome, and transcriptome induced by overexpressed wild-type RIT1 and by expression of $RIT1^{M90I}$, a cancer- and Noonan-associated variant, and to compare these changes to those induced by expression of wild-type and oncogenic KRAS. Because RIT1 mutations occur frequently in lung cancer, we focused this study on analysis of RIT1 and KRAS signaling in a human lung epithelial cell line.

2.2 RESULTS

2.2.1 *Multi-omic profiling of RIT1- and RAS-transformed human lung epithelial cells*

Although the cellular signaling regulated by KRAS has been extensively characterized, little is known about how RIT1 or its oncogenic variants perturb cellular signaling. To address this

question in the context of lung cancer, we ectopically expressed either wild-type or variant forms of KRAS and RIT1 in AALE cells, which are immortalized, non-transformed human lung epithelial cells (A. S. Lundberg et al. 2002). To confirm the functional activity of each expressed variant, we quantified the anchorage-independent growth in soft agar of each isogenic cell line (**Figure 2.1B**). Expression of wild-type or oncogenic variants of KRAS and RIT1 promoted growth in soft agar, confirming the biological activity of each expressed variant (**Figure 2.1, B and C**).

To comprehensively identify the cellular signaling events associated with this phenotypic transformation, we profiled each isogenic cell line using both RNA-seq and deep proteome and phosphoproteome profiling by liquid chromatography tandem mass spectrometry (LC-MS/MS) **Figure 2.1, D and E**). For LC-MS/MS, creation of two overlapping 10-plex tandem mass tag (TMT) pools enabled the relative quantification of the proteome and the phosphoproteome across replicates of the five isogenic cell lines compared to a vector control line ($RIT1^{WT}$, $RIT1^{M90I}$, $KRAS^{WT}$, $KRAS^{G12V}$, $KRAS^{Q61H}$; **Figure 2.1D and Figure 2.5, A to C**). In total, 10,131 proteins were identified, 9002 of which were detected and quantified in every sample; additionally, 29,140 phosphopeptides were detected, 12,325 of which were identified in every sample, enabling comparative analysis across cell lines (**tables S1 and S2 and Data Files S1 and S2**).

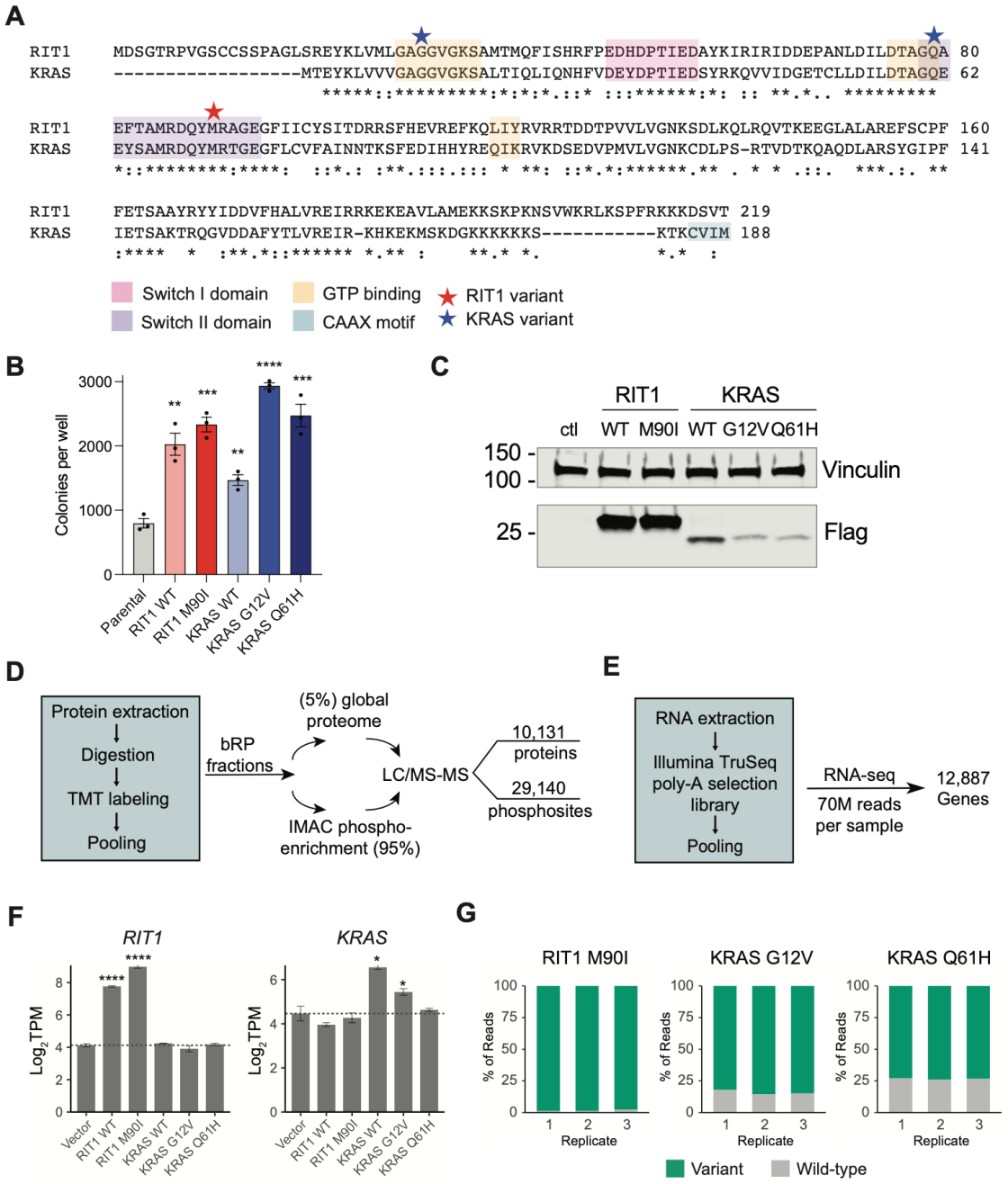


Figure 2.1. Multi-omic profiling of KRAS- and RIT1-mutant human lung epithelial cells.

Protein alignment of KRAS-4B (Uniprot #P01116-2) and RIT1 Isoform 1 (Uniprot #Q92963-1) generated by ClustalW2 (Larkin et al. 2007). Stars indicate the position of the $RIT1^{M90I}$, $KRAS^{G12V}$, and $KRAS^{Q61}$ variants used in this study. Asterisks indicate fully conserved residues; colons indicate strongly conserved residues; periods indicate weakly conserved residues. **(B)** Soft agar colony formation assay of isogenic AALE human lung epithelial cells. N = 3 wells, data shown is representative of 3 independent experiments: ** $p < 0.01$, *** $p < 0.001$, and **** $p<0.0001$ by two-tailed t-test against the parental cell data. **(C)** Western blot confirmation of

ectopic expression of FLAG-tagged RIT1 and KRAS in isogenic AALE cells using an anti-FLAG antibody. Vinculin was used as a loading control. The data shown is representative of N > 3 independent experiments. **(D)** Workflow for the LC-MS/MS–based generation of proteome and phosphoproteome profiles. bRP, basic reverse phase chromatography. IMAC, immobilized metal affinity chromatography. **(E)** Workflow for the RNA-seq analysis. **(F)** mRNA quantification in transcripts per million (TPM) showing mean ± SD of *RIT1* (left) or *KRAS* (right) in isogenic AALE cells, N = 3 RNA-seq libraries per cell line. * $p < 0.05$ and **** $p < 0.0001$ by two-tailed Student's t-test compared to vector controls. **(G)** RNA-seq quantification of variant allele expression. Data shown is the percentage of reads at the M90I, G12V, or Q61H mutation site for the variant allele or wild-type allele. N = 3 RNA-seq libraries for each isogenic AALE cell line.

For transcriptome analysis, RNA sequencing was performed in triplicate for each isogenic cell line to a median read-depth per replicate of 70.1 million reads (**Figure 2.1E**, **tables S3 and S4 and Figure 2.5D**). 12,887 gene transcripts were quantified, including 7,749 whose corresponding protein was detected in the LC-MS/MS profiling. We note that no compensatory feedback regulation of *RIT1* to *KRAS* expression or vice versa was observed (**Figure 2.1F**). Despite relatively low protein expression of KRAS variants in the AALE lines (**Figure 2.1C**), the majority of *KRAS* transcripts in each isogenic cell line corresponded to G12V or Q61H variants, respectively, with 84.1% of reads harboring the G12V mutation in $KRAS^{G12V}$ cells, and 73.3% of reads harboring the Q61H mutation in $KRAS^{Q61H}$ cells (**Figure 2.1G**).

Next, we interrogated the phosphoproteome regulated by KRAS and RIT1. Protein phosphorylation is a reversible and dynamic mechanism of intracellular signaling that enables rapid intracellular transduction of signals controlling cell proliferation, survival, and metabolism. Although RIT1 and KRAS act as GTPase switches, they both additionally stimulate activation of cellular protein kinases, such as BRAF, resulting in substantial reorganization of the phosphoproteome. Because the stoichiometric ratio of phosphorylated to non-phosphorylated proteins is often believed to play an important role in cell signaling, we normalized phosphosite abundance to the total protein abundance for each phosphoprotein (**table S2**). Recognizing that in

some cases absolute abundance of a phospho-peptide may be relevant, non-normalized phosphosite data is also provided (**table S5**).

2.2.2 *Global similarity between oncogenic RIT1 and KRAS signaling networks*

With our LC-MS/MS proteome profiles, we identified differentially abundant proteins in each condition compared to the vector control cells using a two-tailed moderated t-test (**Figure 2.2A**). One of the top proteins with increased abundance in $RIT1^{WT}$, $RIT1^{M90I}$, $KRAS^{G12V}$, and $KRAS^{Q61H}$ cells was the transcription factor FOSL1 (**Figure 2.2B**), also known as FRA1, a basic leucine zipper transcription factor in the FOS family (M. R. Young and Colburn 2006). Activation of RAS is known to promote transcriptional upregulation and protein stabilization of FOSL1 (Mechta et al. 1997; Casalino, De Cesare, and Verde 2003). Western blot of independently-derived AALE isogenic lines confirmed greater abundance of FOSL1 in $RIT1^{WT}$, $RIT1^{M90I}$, and $KRAS^{G12V}$ cells compared to vector expressing cells (**Figure 2.2B**). Consistent with FOSL1 upregulation downstream of the MAPK pathway, analysis of our previous L1000 mRNA profiling dataset revealed common regulation of *FOSL1* by *RIT1*, *KRAS*, and other RAS pathway oncogenes *ARAF*, *BRAF*, *EGFR*, and *NRAS*, but not other oncogenes such as *MYC*, *MDM2*, or *NFE2L2* (**Figure 2.6A**).

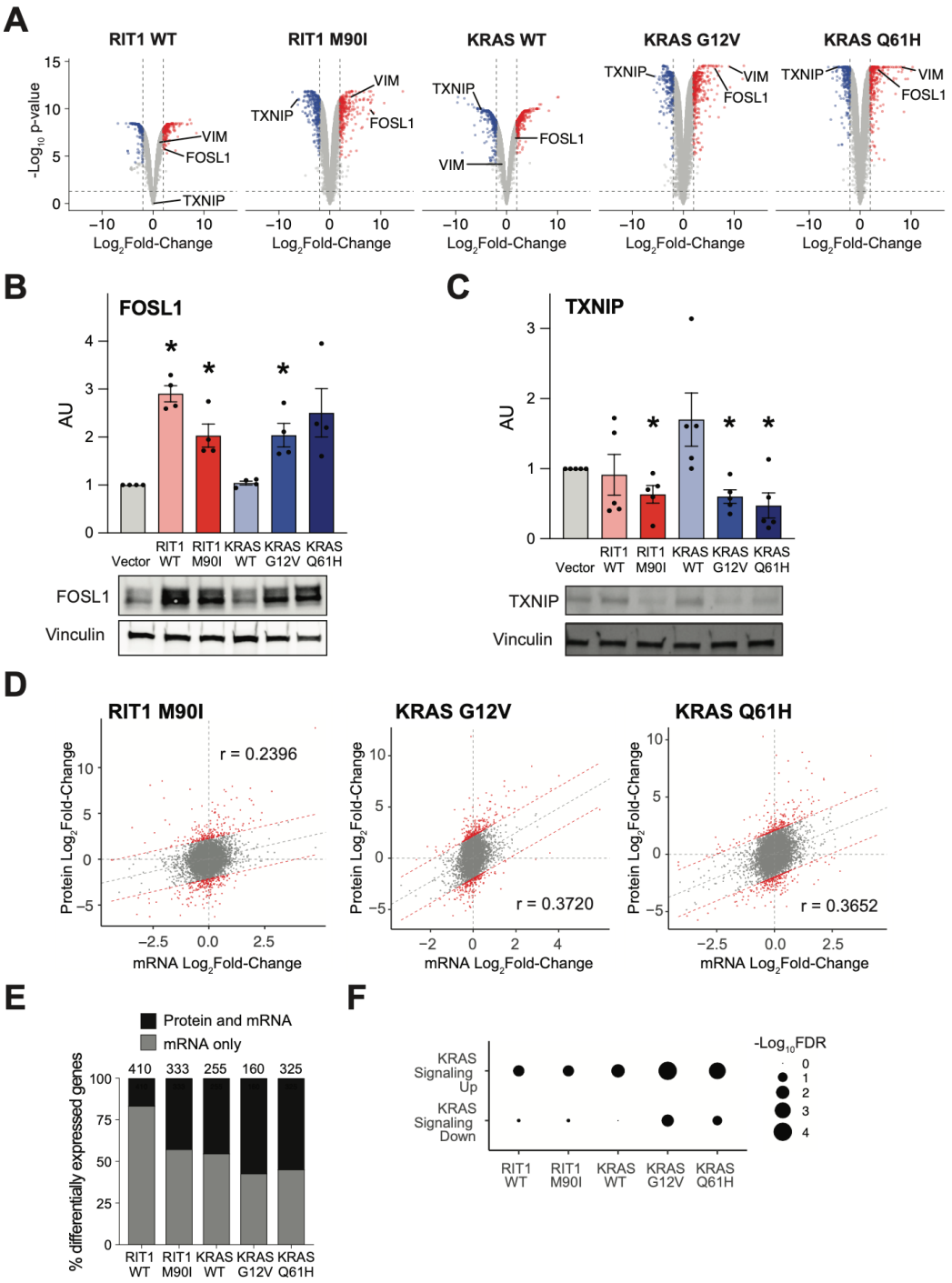


Figure 2.2. Quantitative proteomic and transcriptomic profiling identifies similarity of oncogenic RIT1 and KRAS signaling.

A, Volcano plots of global proteome data from isogenic AALE cells showing the Log_2Fold-Change in protein abundance in each cell line compared to vector control cells. The y-axis displays the negative Log_{10}(*p*-value) calculated from a one sample moderated t-test with multiple hypothesis correction by the Benjamini-Hochberg method. Data shown are generated from 2-4

technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line. **B,** Western blot validation of FOSL1 increased protein abundance in RIT1- and KRAS-variant AALE cells. Quantification of N=4 independent experiments is shown with representative blot. *, $p < 0.05$, change in protein abundance from vector cells, by two-tailed t-test. **C**, Western blot validation of TXNIP protein abundance in RIT1- and KRAS-variant AALE cells. Quantification of N=5 independent experiments is shown with representative blot. *, $p < 0.05$, change in protein abundance from vector cells, by two-tailed t-test **D,** Global proteome-transcriptome correlation analysis for RIT1- and KRAS-variant AALE cell lines. Gray dashed line = linear regression between Log_2Fold-Change of mRNA to respective Log_2Fold-Change of protein. Red = genes outside the 95% prediction interval. *r* = Pearson correlation coefficient. Data shown are generated from 2-4 technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line. N = 3 RNA-seq replicates for each isogenic cell line. **E,** In RIT1- or KRAS-expressing AALE cell lines compared to vector control AALE cells, proportion of genes that are differentially expressed at the mRNA level (Log_2Fold-Change > 1, FDR < 0.05) which are also differentially abundant at the protein level (Log_2Fold-Change > 1, adjusted p-value < 0.05). Number of total differentially expressed mRNA transcripts is shown above the bar for each cell line. Data shown are generated from 2-4 technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line. N = 3 RNA-seq replicates for each isogenic cell line. **F,** Gene set enrichment analysis of differentially expressed gene transcripts between KRAS or RIT1 AALE cell lines and vector controls using GOseq (M. D. Young et al. 2010). mSigDB hallmark gene sets specific to KRAS signaling are shown. Circle size = $-Log_{10}$FDR of enrichment significance determined by GOseq. Analysis performed with N = 3 RNA-seq replicates per isogenic cell line.

Among the top proteins with decreased abundance in $RIT1^{M90I}$, $KRAS^{G12V}$, and $KRAS^{Q61H}$ cells was TXNIP, an inhibitor of thioredoxin involved in both redox regulation and glucose metabolism (Yoshihara et al. 2014; Parikh et al. 2007). Prior studies indicate that $HRAS^{G12V}$ suppresses TXNIP transcription and protein translation (Elgort et al. 2010; Z. Ye and Ayer 2018). TXNIP abundance was also decreased in Western blot analysis of $RIT1^{M90I}$, $KRAS^{G12V}$, and $KRAS^{Q61H}$ cells compared to vector control (**Figure 2.2C**). These validation data demonstrate the utility of LC-MS/MS to describe protein expression changes and additionally suggest the mechanism of RAS-mediated modulation of FOSL1 and TXNIP is shared with $RIT1^{M90I}$.

Previous studies of steady-state cell lines show that approximately 40% of relative protein abundance is explained by mRNA expression levels (E. Lundberg et al. 2010; Liu, Beyer, and

Aebersold 2016). Similarly, in the stable AALE cell lines overexpressing KRAS variants or $RIT1^{M90I}$, we see a moderate correlation between protein abundance and gene transcript levels (Pearson r = 0.2 to 0.4, **Figure 2.2D and Figure 2.6B**). Correspondingly, many genes which are differentially regulated at the mRNA level are also dysregulated at the protein level (**Figure 2E**).

As expected, both KRAS variants and, to a lesser degree, wild-type and variant RIT1 induced KRAS transcriptional signatures (**Figure 2.2F**). To determine whether oncogenic RIT1 and KRAS also induced similar phosphoproteomic changes, we performed Kinase-Substrate Enrichment Analysis (KSEA) (Wiredja, Koyutürk, and Chance 2017), which uses kinase-substrate pairings from PhosphoSitePlus (Hornbeck et al. 2011) and NetworKIN (Linding et al. 2007) to identify differential phosphorylation of kinase-substrate families (**table S6**). This analysis identified ribosomal S6 kinase (RPS6KA1), MAPKAPK2, AKT1, and protein kinase C (PRKCA) as top kinases driving phosphorylation changes in $RIT1^{M90I}$ and $KRAS^{G12V}$/$KRAS^{Q61H}$ cells (**Figure 2.7A-G,** and **table S5**). Substrates of Aurora kinase B (**Figure 2.7A-C, H**) and CDK1 and PAK1 were suppressed in RIT1- and KRAS-variant cells (**Figure 2.7B, C**). Although the total phosphorylation of each substrate reflects the balance between kinases and phosphatases in the cell, these data suggest that $RIT1^{M90I}$, like oncogenic KRAS, can activate canonical RAS effector pathways.

2.2.3 *Wild-type RIT1 overexpression phenocopies RIT1 mutational activation*

Calculation of a correlation matrix from the proteome and phosphoproteome data showed high correlations of all profiles with the exception of the wild-type KRAS cells. $KRAS^{G12V}$ and $KRAS^{Q61H}$ proteomes and phosphoproteomes were highly correlated with each other (proteome r = 0.85 and phosphoproteome r = 0.79; **Figure 2.3A**), and with $RIT1^{WT}$ and $RIT1^{M90I}$ cells (r = 0.70-0.80 and 0.72-0.75 for proteome and phosphoproteome, respectively; **Figure 2.3A**). In

contrast, wild-type KRAS replicates were the most divergent of all profiles, showing limited correlation to either the KRAS variant or RIT1 profiles.

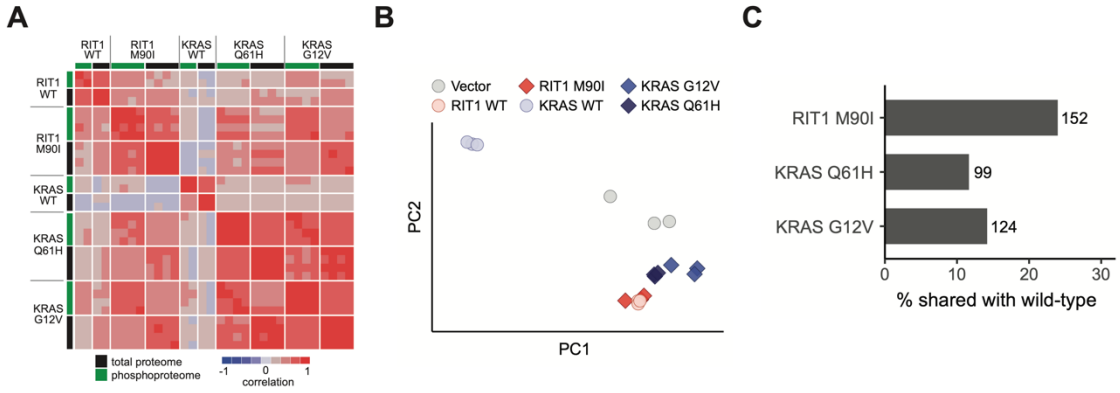


Figure 2.3. Wild-type RIT1 overexpression mimics oncogenic RIT1 activation.

A, Correlation heatmap showing pairwise correlations of each proteome and phosphoproteome replicate to every other replicate. To enable correlation of proteome with phosphoproteome, phosphosites were collapsed to the protein level by taking the median of all phosphosites for each protein. Rows are Spearman correlation; columns are Pearson correlation. Proteome and phosphoproteome data were generated from 2-4 technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line. **B,** Principal component analysis of RNA-seq data from AALE cells. Circles denote individual control vector, $RIT1^{WT}$, or $KRAS^{WT}$ transcriptomes, as labeled; diamonds denote RIT1-variant or KRAS-variant transcriptomes, as labeled. PC1 explains 96.2% of variance, PC2 explains 2.16% of variance. N = 3 RNA-seq libraries per cell line. **C,** Bar plot showing percentage of differentially abundant proteins ($|Log_2Fold\text{-}Change| > 2$) in wild-type AALE cells that are also differentially abundant in variant AALE cells. Numbers next to bars indicate number of proteins represented. Data shown are generated from 2-4 technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line.

A relatively recent study found that RIT1 mutations, including M90I, may function by relieving negative regulation of RIT1 by a LZTR1-dependent proteasomal degradation mechanism (Castel et al. 2019). Accordingly, overexpression of wild-type RIT1 should largely phenocopy expression of $RIT1^{M90I}$. Consistent with this idea, $RIT1^{WT}$ cells more closely resembled both $RIT1^{M90I}$ (proteome $r = 0.49$) and KRAS variant cells (proteome $r = 0.36\text{-}0.48$) than $KRAS^{WT}$ cells (proteome $r = 0.19$; **Figure 3A**). These data highlight a critical divergence between KRAS and RIT1: expression of wild-type KRAS is not capable of activating downstream oncogenic RAS

pathways, whereas expression of wild-type RIT1 in part resembles activation of RIT1 or KRAS by mutation. We confirmed this observation in a principal component analysis of the transcriptome data, which further revealed a high degree of similarity between $RIT1^{WT}$ and $RIT1^{M90I}$-regulated gene expression (**Figure 2.3B and table S4**). Further, of the differentially abundant proteins in $RIT1^{WT}$ cells, 25% were differentialy abundant in $RIT1^{M90I}$ cells. In contrast, 7-10% of proteins with altered abundance levels in $KRAS^{WT}$ cells had differentially altered abundance levels in $KRAS^{mut}$ cells **(Figure 2.3C)**. In the phosphoproteome, unsupervised hierarchical clustering of the phospho-signatures identified the $RIT1^{M90I}$ phosphoproteome as most similar to $KRAS^{G12V}$ and $KRAS^{Q61H}$ phospho-signatures (**Figure 2.7I**), suggesting that $RIT1^{M90I}$ induces similar cellular phosphorylation changes as KRAS variants. Additionally, the phosphoproteome signatures of $RIT1^{WT}$ cells clustered more closely than wild-type KRAS to the $RIT1^{M90I}$, $KRAS^{G12V}$, and $KRAS^{Q61H}$ phosphoproteome signatures (**Figure 2.7I**).

2.2.4 *RIT1 promotes an epithelial-to-mesenchymal transition*

To further investigate the phenotypic changes conferred by these proteome and transcriptome changes, we performed gene set overlap analysis using MSigDB Hallmark Pathway gene sets (Liberzon et al. 2015) (**Figure 2.4A**). The epithelial-to-mesenchymal transition (EMT) pathway was the most significant gene set enriched among up-regulated proteins across all KRAS- and RIT1-expressing cells compared to vector control (**Figure 2.4A**). EMT is a cellular transdifferentiation process promoted by cell-extrinsic signaling proteins and orchestrated by activation of transcription factors such as Twist, Snail, and Zeb family transcription factors (X. Ye and Weinberg 2015). It has long been observed that oncogenic RAS proteins—including KRAS variants—promote EMT, but little is known about the involvement of RIT1 in this process.

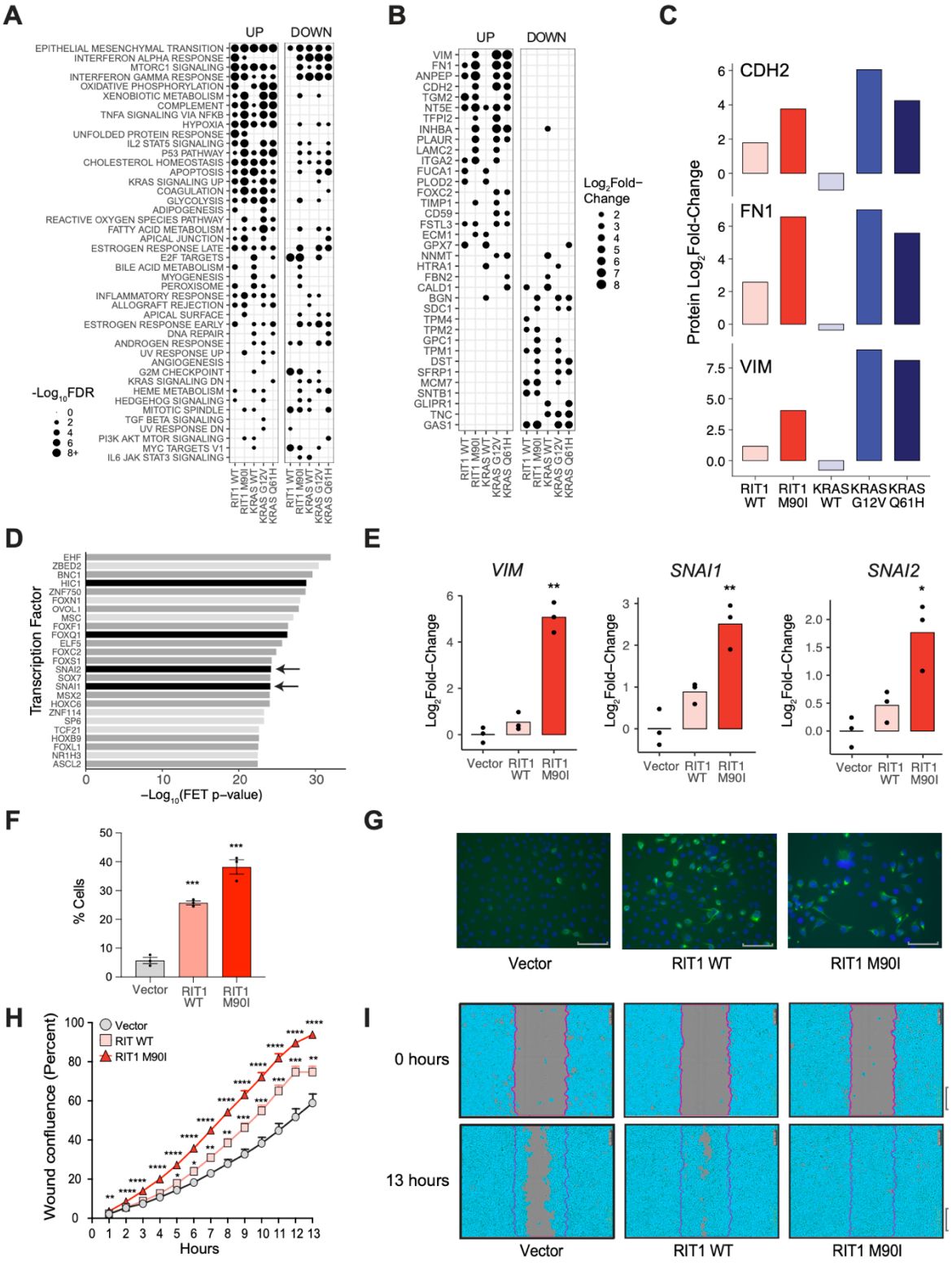


Figure 2.4. Wild-type RIT1 and RIT1^{M90I} promote an epithelial-to-mesenchymal (EMT) transition phenotype.

A, Gene set overlap analysis of up-regulated ("Up"; Log_2Fold-Change>2) and down-regulated ("down"; Log_2Fold-Change<-2) proteins in each condition in AALE cells, using MSigDB Hallmark Pathways (Liberzon et al. 2015). Gene sets sorted by descending average -Log_{10}FDR across all conditions. Circle size = -Log_{10}FDR of gene set overlap analysis determined by MSigDB. The analysis was based on data generated from 2-4 technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line. **B**, Log_2Fold-Change of protein abundance of hallmark EMT genes in RIT1- or KRAS-expressing AALE cells relative to vector control AALE cells. Proteins sorted by average Log_2Fold-Change across all conditions. The analysis was based on data generated from 2-4 technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line. **C,** Log_2Fold-Change in RIT1- or KRAS-expressing AALE cells of abundance of EMT-marker proteins N-cadherin (CDH2), fibronectin (FN1), and vimentin (VIM). The analysis was based on data generated from 2-4 technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line. **D,** Transcription factor target enrichment analysis of differentially expressed genes in $RIT1^{M90I}$ AALE cells using Enrichr libraries through ChEA3. FET, Fisher's exact test. Analysis performed on N = 3 RNA-seq libraries compared to 3 vector controls. Arrows indicate proteins of the Snail family. Black denote confirmed EMT genes in dbEMT (M. Zhao et al. 2019); dark gray denote EMT-associated genes in the literature. **E,** Real-time PCR of relative Log_2Fold-Change of mRNA expression in perturbed AALE cells of vimentin (*VIM*), Snail (*SNAI1*), and Slug (*SNAI2*). N=3 biological replicates; * p<0.05, ** p<0.01 by two-tailed t-test. **F,** Immunofluorescence of Vimentin stained cells. N=3 biological replicates; *** p<0.001 by two-tailed t-test. **G,** Representative immunofluorescence images of Vimentin stained cells. Scale bar corresponds to 89.1μm. N = 3 replicates per cell line. **H,** Wound healing scratch assay. Percent wound confluence at hourly time points are shown with SEM error bars. N=16 replicates per cell line, data shown is representative of N = 3 independent experiments; * p<0.05, ** p<0.01, *** p<0.001, and **** p<0.0001 by two-tailed t-test each time point, corrected for multiple testing. **I,** Representative images of wound healing at time of wound making (0 hours) and at end time point (13 hours). Scale bar, 300μm

Notably, $RIT1^{M90I}$ and, to a lesser degree, wild-type RIT1 were capable of promoting expression changes of key EMT markers, including up-regulation of vimentin, N-cadherin, and fibronectin, and downregulation of keratin-19 (**Figure 2.4, B and C, and Figure 2.8A**). Although the canonical EMT transcription factors Snail (SNAI1) and Slug (SNAI2) were not detected by proteomic analysis, transcriptomic analysis of $RIT1^{M90I}$ cells showed increased activity of these EMT transcription factors by transcription factor target enrichment analysis (**Figure 2.4D and Figure 2.8, B and C**). Real-time qPCR confirmed induction of vimentin, Snail, and Slug by

RIT1^{M90I} (**Figure 2.4E**). Quantitative immunofluorescence analysis revealed both RIT1^{WT} and RIT1^{M90I} induced vimentin protein expression in AALE cells (**Figure 2.4, F and G**).

To functionally determine if EMT characteristics are present in RIT1-transformed AALE cells, we performed a scratch assay to quantify cellular migration. Consistent with a functional EMT, wild-type RIT1 and RIT1^{M90I} cells showed enhanced migration, closing the scratch wound up to 1.5x and 1.9x faster than vector control cells respectively (**Figure 2.4, H and I**). Therefore, we find that both wild-type RIT1 overexpression and RIT1^{M90I} promote an EMT phenotype in lung epithelial cells involving increased expression of vimentin and increased cellular migration.

Together these data strengthen the notion that oncogenic RIT1 hijacks canonical RAS effector pathways to promote tumorigenesis. In addition, these data demonstrate that wild-type RIT1 overexpression can induce many of these same processes to a weaker degree, supporting a role for protein abundance in regulation of RIT1 activity and *RIT1* amplification/overexpression in contributing to malignant phenotypes.

2.3 DISCUSSION

Herein, we described quantitative proteomic, phosphoproteomic, and transcriptomic datasets that provide a systems-level view of the RIT1-regulated signaling network in lung epithelial cells. These datasets were generated from isogenic human lung epithelial cells to provide a physiological view of the consequences of RIT1 activation in the same cellular compartment that is involved in lung adenocarcinoma, a tumor type with prevalent mutations in *KRAS* and *RIT1*. Broadly, we find that proteome and transcriptome signatures from RIT1M90I-expressing cells partially phenocopy those from cells with overexpression of wild-type RIT1. This finding lends further support to the notion that oncogenic RIT1 variants function at least in part through increasing RIT1 abundance (Castel et al. 2019). This is in contrast to KRAS, where overexpression of wild-type KRAS induces

signatures less related to that of oncogenic variants KRASG12V and KRASQ61H. The opposing functions of wild-type and variant KRAS is consistent with recent evidence suggesting that KRAS functions as a dimer and that wild-type KRAS directly inhibits the function of oncogenic KRAS variants via physical dimerization (Ambrogio et al. 2018).

This divergence in the function of wild-type RIT1 and KRAS hints at fundamental differences in molecular regulation of each wild-type GTPase. The ability of RIT1 to promote downstream RAF/MEK/ERK signaling when aberrantly expressed suggests that RIT1 may not be subject to the same tight regulation by GTPase-activating proteins (GAPs) that normally keep RAS in an inactive state. Furthermore, these data point to the importance of studying wild-type *RIT1* amplification and overexpression in cancers. *RIT1*, on chromosome 1q, is frequently amplified in uterine carcinosarcoma, liver hepatocellular cancer, cholangiocarcinoma, breast cancer, lung adenocarcinoma, and ovarian cancer (**Figure 2.9A**) (Cancer Genome Atlas Research Network et al. 2013). *RIT1* mRNA expression is increased in amplified cases, regardless of tissue type (**Figure 2.9B**), raising the possibility that RIT1 overexpression could play a role in tumorigenesis in these cancers.

Although this resource should enable identification of differences between RIT1 and KRAS networks, we observe marked similarity between the transcriptional, proteomic and phosphoproteomic changes induced by $RIT1^{M90I}$, $KRAS^{G12V}$, and $KRAS^{Q61H}$. Oncogenic RIT1 and KRAS share the ability to activate canonical RAS effector pathways such as those involved in regulation of FOSL1 and TXNIP abundance. These cellular signaling changes are accompanied by phenotypic changes to the lung epithelial cells including acquisition of anchorage-independent growth capability as well as induction of an EMT phenotype with enhanced cellular migration and increased mRNA or protein abundance of EMT markers Vimentin, Snail, and Slug.

Together, these results demonstrate the power of quantitative proteomics and transcriptomics to provide global views of cancer oncogene signaling and provide a resource for understanding the altered cell signaling in Noonan syndrome and cancer.

2.4 MATERIALS AND METHODS

2.4.1 *Isogenic Cell Line Generation*

Plasmid constructs were cloned using Gateway Technology (Invitrogen/ThermoFisher) using pLX301 destination vector (Broad Institute) and pDONR223-RIT1 donor vectors previously described(A. H. Berger et al. 2014). Lentivirus was generated by transfection of HEK293T cells with packaging and envelope vectors using standard protocols. AALE cells were a kind gift of Jesse Boehm (Broad Institute). Isogenic cells were generated by transduction of lentivirus generated from pLX317-Renilla luciferase or pLX301-RIT1^{WT}, pLX301-RIT1^{M90I}, pLX301-KRAS^{WT}, pLX301-KRAS^{G12V}, or pLX301-KRAS^{Q61H} and selection with puromycin. Stable pools of cells were maintained in small airway growth medium (Lonza).

2.4.2 *Soft Agar Assay*

$1x10^5$ cells were suspended in 1 ml of 0.33% select agar in small airway growth medium (Lonza) and plated on a bottom layer of 0.5% select agar in the same media in six-well dishes. Each cell line was analyzed in triplicate. Colonies were photographed after 3-5 weeks and quantified using Fiji software.

2.4.3 *Transcriptome profiling*

Three replicates per cell line were harvested at ~90% confluence (n = 18 total culture dishes). Cells were lysed and total RNA was extracted using Direct-zol RNA Miniprep plus (Zymo Research).

Libraries were constructed using the non-strand-specific poly-A selection Illumina TruSeq kit for 50bp paired-end reads. Libraries were pooled and sequenced on the Illumina NovaSeq platform (Fred Hutch Genomics Core). Reads were aligned to the human reference genome build hg19/GRCh37 using STAR v.2.5.3a (Dobin et al. 2013). Alignments were annotated for duplicates and read groups then reordered and indexed using Picard Tools v.1.114 ("Picard Tools - By Broad Institute" n.d.). Read statistics for each RNA-seq sample were calculated using RSeQC (Wang, Wang, and Li 2012). Quantification of gene transcripts was performed by the featureCounts program within the Subread package (Liao, Smyth, and Shi 2014), using hg19 gene annotations from UCSC. Gene level counts per million reads (CPM) and reads per kilobase of transcript mapped reads (RPKM) values were calculated with edgeR v.3.22.3 (Robinson, McCarthy, and Smyth 2009), and converted into transcripts per million reads (TPM) values with an in-house script. In total, 12,887 genes were identified with average logCPM at least 0.1 across all samples. Differential expression analyses comparing KRAS or RIT1 perturbed cell lines against vector control lines were performed using the R package edgeR (Robinson, McCarthy, and Smyth 2009).

2.4.4 *High performance liquid chromatography tandem mass spectrometry (LC-MS/MS)*

Cells were washed in ice-cold PBS, scraped into PBS, pelleted, and snap frozen in liquid nitrogen. The experimental workflow for sample processing, TMT-labeling, peptide enrichment, and LC-MS/MS were largely as previously described (Mertins et al. 2018). Briefly, pellets were lysed in 200 µl of chilled urea lysis buffer (8 M urea, 75 mM NaCl, 50 mM Tris (pH 8.0), 1 mM EDTA, 2 µg/ml aprotinin, 10 µg/ml leupeptin, 1 mM PMSF, 1:100 (vol/vol) Phosphatase Inhibitor Cocktail 2, 1:100 Phosphatase Inhibitor Cocktail 3, 10 mM NaF, and 20 µM PUGNAc) for each ~50 mg portion of wet-weight tissue. Lysates were reduced with 5mM DTT, alkylated with 10 mM IAM, and digestion performed in solution with 1 mAU LysC per 50 µg of total protein and trypsin at an

enzyme/substrate ratio of 1:49. Reactions were quenched with FA and brought to pH = 3 with FA. Peptides were desalted on 200 mg tC18 SepPak cartridges and dried by vacuum centrifugation. 340 µg of peptides were labeled with 10-plex Tandem Mass Tag reagents (TMT10, Fisher Scientific), according to manufacturer's instructions. To enable quantification of peptides across all 12 samples, the samples were labeled in sets of 10 across two different TMT10 pools in a crossover design with 8 of 12 samples analyzed in both TMT10 pools. A 50/50 mix of both AALE vector control lysates was used as an internal reference in both TMT10 runs (**Figure 2.5B**).

Each TMT10-plex was desalted in a 200 mg tC18 SepPak cartridge and fractionated using offline HPLC. 5% of each fraction was collected into an HPLC vial for proteome analysis by LC-MS/MS. The remaining 95% was processed for phospho-peptide enrichment via immobilized metal affinity chromatography (IMAC). IMAC enrichment was performed using Ni-NTA Superflow Agarose beads incubated with peptides solubilized in a final concentration of 80% MeCN/0.1% TFA. Phospho-enriched peptides were desalted and collected into an HPLC vial for analysis by LC-MS/MS.

Online fractionation was performed using a nanoflow Proxeon EASY-nLC 1200 UHPLC system (Thermo Fisher Scientific) and separated peptides were analyzed on a benchtop Orbitrap Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) equipped with a nanoflow ionization source (James A. Hill Instrument Services, Arlington, MA). In-house packed columns (20 cm x 75 µm diameter C18 silica picofrit capillary column; 1.9 µm ReprosIl-Pur C18-AQ beads, Dr. Maisch GmbH, r119.aq; Picofrit 10 µm tip opening, New Objective, PF360-75-10-N-5). Mobile phase flow rate was 200 nL/min, comprised of 3 % acetonitrile/0.1 % formic acid (Solvent A) and 90 % acetonitrile /0.1 % formic acid (Solvent B). The 110 min LC-MS/MS method consisted of a 10 min column-equilibration procedure; a 20 min sample-loading procedure; and

the following gradient profile: (min: % B) 0:2; 2:6; 85:30; 94:60; 95:90; 100:90; 101:50; 110:50 (the last two steps at 500 nL/min flow rate). Data-dependent acquisition was performed using Xcalibur QExactive v2.4 software in positive ion mode at a spray voltage of 2.00 kV. MS1 Spectra were measured with a resolution of 70,000, an AGC target of 3e6 and a mass range from 300 to 1800 m/z. Up to 12 MS/MS spectra per duty cycle were triggered at a resolution of 35,000, an AGC target of 5e4, an isolation window of 0.7 m/z, a maximum ion time of 120 msec, and normalized collision energy of 30. Peptides that triggered MS/MS scans were dynamically excluded from further MS/MS scans for 20 sec. Charge state screening was enabled to reject precursor charge states that were unassigned, 1, or >6. Peptide match was set to preferred for monoisotopic precursor mass assignment.

2.4.5 *Protein-peptide identification, phosphosite localization, and quantification*

MS data was interpreted using the Spectrum Mill software package v6.0 pre-release (Agilent Technologies, Santa Clara, CA. MS/MS spectra were merged if they were acquired within +/- 45 sec of each other with the same precursor m/z. Also, MS/MS spectra that did not have a sequence tag length > 0 (i.e., minimum of two masses separated by the in chain mass of an amino acid) or did not have a precursor MH+ in the range of 750-6000 were excluded from searching. MS/MS spectra searches were performed against a concatenated UniProt human reference proteome sequence database containing 58,929 human proteins including isoforms (obtained 10/17/2014) and 150 additional common laboratory contaminants. ESI-QEXACTIVE-HCD-3 scoring parameters were used for both whole proteome and phosphoproteome datasets. Spectra were allowed +/- 20 ppm mass tolerance for precursor as well as product ions, 30% minimum matched peak intensity, and “trypsin allow P” was set as enzyme specificity with up to 4 missed cleavages allowed. Carbamidomethylation at cysteine was set as fixed modification together with TMT10

isobaric labels at lysine residues (N-termini would be considered regardless if it was TMT labelled). Acetylation of protein N-termini andoxidized methionine were set as variable modifications with a precursor MH+ shift range of -18 to 64 Da for the proteome searches. For the phosphoproteome searches the precursor MH+ shift range was set to 0 to 272 Da and variable modifications of phosphorylation of serine, threonine, and tyrosine. Identities interpreted for individual spectra were automatically designated as confidently assigned using the Spectrum Mill autovalidation module to use target-decoy based false discovery rate (FDR) estimates to apply score threshold criteria. For the whole proteome datasets, thresholding was done at the spectral (< 1.2%) and protein levels (< 0.1%) . For the phosphoproteome datasets, thresholding was done at the spectral (< 1.2%) and phosphosite levels (< 1.0%).

Replicates across TMT-plexes were highly correlated (**Figure 2.5C)** with median Pearson r = 0.87 and 0.69 for proteome and phosphoproteome, respectively**.** Technical replicates and biological replicates were merged to generate final total proteome and phospho-proteome profiles for each isogenic cell line (**tables S1 and S2**). Replicate-level profiles are also supplied as JavaScript Object Notation (.json) files that can be visualized and analyzed using the Morpheus Matrix Visualization and Analysis Software at https://software.broadinstitute.org/morpheus (**Data Files S1 and S2**). Differential protein and phospho-site signatures were generated by computing the mean $\log_2$(fold change) of the abundance of each site in each sample compared to the vector control cells. Statistical significance of differentially abundant proteins and phosphosites was determined by performing a one sample moderated t-test with multiple hypothesis correction (**tables S1 and S2**).

2.4.6 *Integrative Proteome and Transcriptome Analysis*

Correlation of changes in protein expression and changes in RNA expression was modeled with a linear model. 95% prediction intervals were calculated to highlight genes with weak concordance between protein and RNA expression. Differentially expressed transcripts were determined by filtering edgeR results (described above) to transcripts with $|\log_2\text{Fold-Change}| > 1$ and FDR < 0.05. Differentially abundant proteins were determined by filtering to proteins with $|\log_2\text{Fold-Change}| > 1$ and adjusted p-value < 0.05.

2.4.7 *Gene Set Enrichment Analysis*

Analysis of enrichment of KRAS signaling in differential RNA expression profiles was performed in R with GOseq (M. D. Young et al. 2010). KRAS signaling gene sets were taken from MSigDB hallmark gene sets (Subramanian et al. 2005; Liberzon et al. 2015).

2.4.8 *Transcription Factor Target Enrichment Analysis*

Analysis of over-representation of transcription factor targets was performed with ChIP-X Enrichment Analysis 3 (ChEA3) by submitting lists of differentially expressed genes ($|\log_2\text{Fold-Change}| > 1$ and FDR < 0.05). ChEA3 performs Fisher's Exact Tests to compare the input gene set to transcription factor target gene sets in six different libraries (Keenan et al. 2019)**.** Analysis of the Enrichr Queries library was selected as the focus of the present study. Transcription factors resulting from this analysis were manually annotated as one of three groups of EMT association. These groups were: 1) confirmed EMT genes defined by dbEMT(M. Zhao et al. 2019), 2) genes shown to be associated with EMT in at least one study in literature, and 3) genes unrelated to EMT.

2.4.9 *Antibodies and immunoblotting*

Antibodies against FOSL1 (D80B4), TXNIP (D5F3E), and Vimentin (D21H3) were purchased from Cell Signaling Technology. Vinculin (V9264) was purchased from Sigma Aldrich. Secondary antibodies StarBright Blue 700 Goat anti-Rabbit IgG, StarBright Blue 520 Goat anti-Rabbit IgG and StarBright Blue 520 Goat anti-Mouse IgG (12005867) were purchased from Bio-Rad. Antibody against RIT1 (#53720) was purchased from Abcam. Cell lysates were prepared in RTK lysis buffer with protease (11836153001, Roche) and phosphatase (04906837001, Roche) inhibitors added and quantified by the BCA assay (Thermo Scientific Waltham, MA). Samples were then boiled in Laemmli buffer (1610747, Bio-Rad, Hercules, CA) and 50 ug of protein was loaded onto 4-15% Mini-Protean TGX (4561084, Bio-Rad, Hercules, CA) gels. Protein gels were run and transferred to PVDF membranes (1704274, Bio-Rad, Hercules, CA) according to manufacturer's instructions. Proteins were detected by specific primary antibody and secondary antibody then visualized using the ChemiDoc MP Imaging System (Bio-Rad, Hercules, CA) or Odyssey Imager (Li-Cor).

2.4.10 *KSEA analysis*

Kinase-substrate enrichment analysis (KSEA) (Casado et al. 2013) was performed using the KSEA App (Wiredja, Koyutürk, and Chance 2017) (https://casecpb.shinyapps.io/ksea/) using kinase-substrate mappings from PhosphoSitePlus (Hornbeck et al. 2011) and a p value threshold of < 0.05. A minimum of five detected phospho-site substrates were needed for kinases to be included in the analysis. The full list of kinase scores and number of substrates are shown in Supplementary Table 6. 36 kinases had sufficient substrate sites detected to be included in the analysis. Kinase-substrate mappings are shown in Supplementary Table 6.

2.4.11 *Scratch Assay*

Cells were seeded in 96-well ImageLock plates (Sartorius) at a concentration of 46,000 cells/well. The cells were then incubated overnight in SAGM media without EGF (Lonza). After incubation, with the cells forming a monolayer, a scratch wound is made using the 96-well WoundMaker tool (Sartorius). Plate was imaged every hour for 13 hours and analyzed using the IncuCyte ZOOM v2016A (Sartorius). Further analysis was performed in GraphPad Prism v9 using multiple comparison two-tailed t-tests and reporting q values.

2.4.12 *qRT-PCR*

Extracted RNA for each biological replicate of the 6 cell lines were used to generate cDNA by reverse transcription with the SuperScript IV First-Strand Synthesis kit (Invitrogen). Quantitative RT-PCR reactions were set up in technical triplicates using the Taqman Gene Expression Master Mix (Thermo Fisher) and a TaqMan Gene Expression Assay for a reference gene, 18S (Hs99999901_s1), as well as for each gene of interest: VIM (Hs00958111_m1), SNAI1 (Hs00195591_m1), and SNAI2 (Hs00161904_m1). PCR reactions were performed and gene expression quantified using the CFX384 Real-Time System (Bio-Rad). Relative gene expression was calculated by the Livak method using an in-house script.

2.4.13 *Immunofluorescence*

For immunofluorescence detection of Vimentin, 2.5 x 105 AALE stable cells (passage 5) were plated in triplicate on 4-well Nunc Lab-Tek II CC2 chamber slides and allowed to adhere overnight at 37°C and 5% CO2. The next day, media was aspirated and cells were washed once with PBS, then fixed with 4% paraformaldehyde in PBS for 10 minutes at room temperature. Cells were then washed twice with ice cold PBS and permeabilized with 0.25% Triton-X-100 for 10 minutes

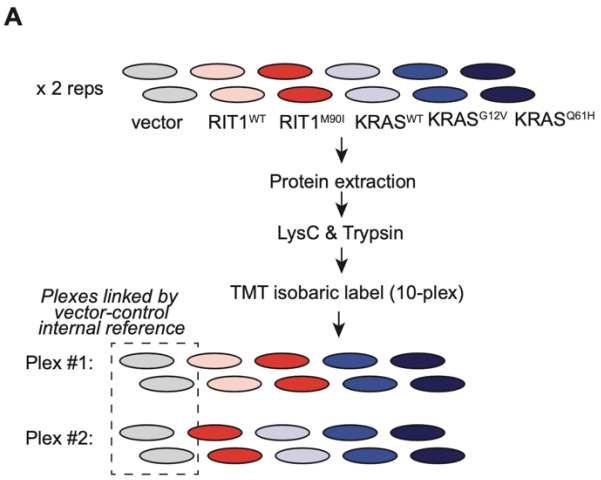
followed by three washes with PBS. Cells were blocked with 10% normal goat serum (Thermo Fisher Scientific) in PBS with 0.1% Tween 20 (PBST) followed by overnight incubation at 4 degree C with Vimentin rabbit antibody (D21H3 XP® monoclonal antibody, Cell Signaling #5741, 1:1000 dilution) in PBST. The next day, cells were washed with PBST three times for 5 minutes then incubated with Alexa Fluor® 488 preadsorbed Goat anti-rabbit IgG antibody (Abcam #ab150117 at 1:1000 dilution) for 1 hour at room temperature in the dark. Cells were then washed three times with PBST and slides mounted using Vectashield antifade mounting medium with 1.5ug/mL DAPI (Vector Laboratories). The slides were sealed with clear nail polish and imaged on a Nikon Eclipse E800 microscope with 20X/0.75 NA PlanApo objective using a DAPI filter set (excitation wavelength: 330-380 nm; emission wavelength: 435-485) and fixed 300 ms exposure using the FITC filter set (excitation wavelength: 460-500 nm; emission wavelength: 510-560 nm) using a AxioCam HRm camera and AxioVision 40 software v4.8.20. To determine the percent of positive cells, images were analyzed using TissueQuest software v7.0.1.139 (TissueGnostics, Vienna, Austria). Images were segmented using nuclear DAPI fluorescence and then a ring mask around nuclei was applied to the Vimentin image. The mean fluorescence intensity within each masked area was quantified and the percentage of cells above a threshold intensity were summarized by the software. Data were analyzed using a two-tailed t-test in GraphPad Prism v9 with a p value less than 0.05 considered significant.

2.5 ACKNOWLEDGEMENTS

We thank Drs. Athea Vichas and Jon Cooper (Fred Hutchinson Cancer Research Center) for advice, discussion, and critical reading of the manuscript. We thank Dr. D.R. Mani (Broad Institute) for guidance on statistical analysis. We thank the Fred Hutch Cellular Imaging Shared Resource for assistance with microscopy and image analysis. We thank the Fred Hutch Genomics

Shared Resource for performing RNA sequencing and assisting with analysis. **Funding:** This research was funded in part through the National Cancer Institute (NCI) K99/R00 CA197762 and R37 CA25205 to AHB, NIH/NCI Cancer Center Support Grant P30 CA015704, NCI Clinical Proteomic Tumor Analysis Consortium grants NIH/NCI U24-CA210986 and NIH/NCI U01 CA214125 to SAC. AL was supported in part by NSF IGERT DGE-1258485. KH was supported in part by PHS NRSA T32GM007270 from NIGMS. **Author contributions:** A.H.B. conceived of and directed the study. S.C. and P.M. supervised the LC-MS/MS experiments. F.M. performed the proteomics data analysis, with contributions from A.H.B. and K.H. A.L. performed the transcriptome analysis. A.L. and A.H.B. wrote the manuscript. A.L., K.H., S.K., S.M., I.F., S.F., J.W., and A.H.B. performed experiments. All authors discussed results and provided input on the manuscript. **Competing interests:** S.A.C. is a member of the scientific advisory boards of Kymera, PTM BioLabs, and Seer. **Data and materials availability:** pLX301 lentiviral expression vectors for expression of RIT1 and KRAS are available from A.H.B. under a material agreement with Fred Hutchinson Cancer Research Center. The RNA-seq data have been deposited in the NCBI Gene Expression Omnibus database with accession number GSE146479. The original mass spectra and the protein sequence database used for searches have been deposited in the public proteomics repository MassIVE (https://massive.ucsd.edu) and are accessible at ftp://MSV000085225@massive.ucsd.edu with username: MSV000085225 and password: oncogenic. All other data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials.

2.6 SUPPLEMENTAL FIGURES



B

Experiment 1	*TMT Tag*
KRAS G12V Rep 1	126
RIT1 WT Rep 2	131
RIT1 M90I Rep 2	127C
KRAS Q61H Rep 2	127N
Vector Rep 1	128C
KRAS G12V Rep 2	128N
RIT1 M90I Rep 1	129C
RIT1 WT Rep 1	129N
KRAS Q61H Rep 1	130C
Vector Rep 2	130N

Experiment 2	*TMT Tag*
KRAS G12V Rep 1	126
RIT1 M90I Rep 1	131
KRAS Q61H Rep 2	127C
KRAS Q61H Rep 1	127N
KRAS WT Rep 2	128C
KRAS WT Rep 1	128N
Vector Rep 1	129C
RIT1 M90I Rep 2	129N
KRAS G12V Rep 2	130C
Vector Rep 2	130N

C

Replicate set	*Proteome*	*Phospho-proteome*
RIT1 WT	0.84	0.60
RIT1 M90I	0.83	0.69
KRAS WT	0.90	0.77
KRAS G12V	0.87	0.69
KRAS Q61H	0.91	0.72

D

Group	Sample	Total Reads	Mapped Reads	rRNA Reads
KRAS-G12V	KRAS-G12V_1	265969454	102506440	1413908
KRAS-G12V	KRAS-G12V_2	236719962	63365494	1239580
KRAS-G12V	KRAS-G12V_3	246633574	77975208	1684446
KRAS-Q61H	KRAS-Q61H_1	230047062	75117724	2306602
KRAS-Q61H	KRAS-Q61H_2	212706418	67632558	1454868
KRAS-Q61H	KRAS-Q61H_3	212464086	80122798	719300
KRAS-WT	KRAS-WT_1	215832806	75204614	1927496
KRAS-WT	KRAS-WT_2	114503272	54598020	915628
KRAS-WT	KRAS-WT_3	275145008	86902056	1710894
Vector	Vector_1	208952294	77102730	1373232
Vector	Vector_2	250094672	79498920	2074506
Vector	Vector_3	188129714	55042564	1131944
RIT1-M90I	RIT1-M90I_1	219121344	53220806	2628260
RIT1-M90I	RIT1-M90I_2	228132502	67300612	2385516
RIT1-M90I	RIT1-M90I_3	180861518	34480460	1295320
RIT1-WT	RIT1-WT_1	223601164	59893560	1913980
RIT1-WT	RIT1-WT_2	254163922	72581670	1783556
RIT1-WT	RIT1-WT_3	211397706	63770336	1824880

Figure 2.5. Workflow and quality control of proteomic and transcriptomic profiling

A, Replicate-level workflow for tandem mass tag (TMT) labeling and LC-MS/MS. Lysates from duplicate sets of six isogenic AALE cell lines were used to generate two TMT-plex sets, with control samples used to link the two sets. **B,** TMT 10-plex layout showing mass tags associated with each replicate set of isogenic AALE cells. **C,** Average pairwise replicate correlations (Pearson *r*) of all replicates from each sample group indicated. **D,** RNA-seq run and mapping statistics show total reads, mapped reads, and reads mapped to rRNA, for each AALE cell line.

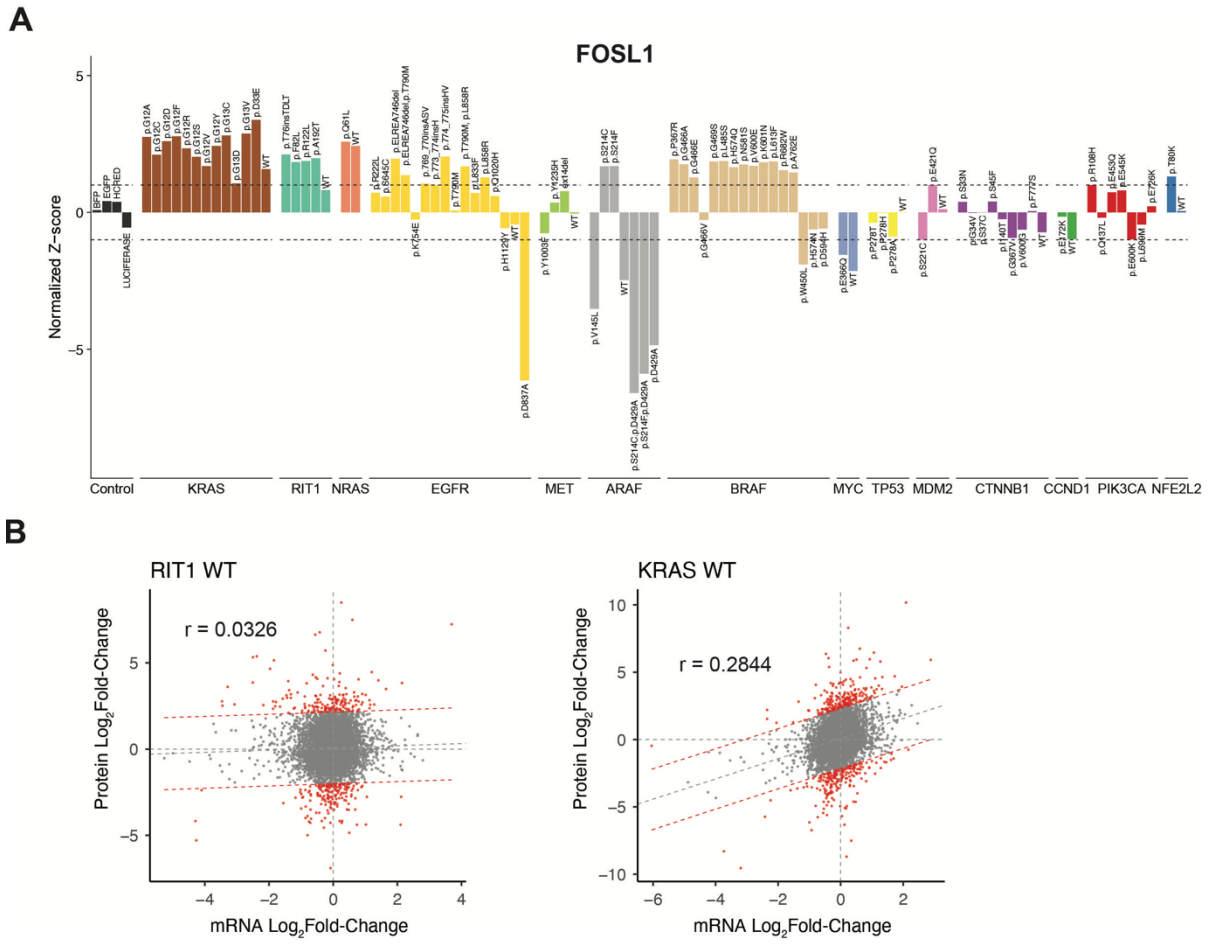


Figure 2.6. Transcriptomic profiling of KRAS and RIT1 cells.

A, Gene expression of *FOSL1* in L1000 assay of isogenic AALE cells. Z-scores are calculated normalized to L1000 plate. Dashed lines indicate Z-scores of 1 and -1. N = 6 to 8 L1000 replicates for each perturbation. Genes are ordered left to right: vector controls, RAS signaling genes, and other oncogenes or tumor suppressors. Within a gene, conditions are shown in the order from left to right: putative oncogenic mutations, wild-types, and lastly any kinase-dead conditions. **B,** Global proteome-transcriptome correlation analysis for wild-type RIT1 and wild-type KRAS overexpressing AALE cell lines. Gray dashed line = linear regression between Log_2Fold-Change of mRNA to respective Log_2Fold-Change of protein. Red = genes outside the 95% prediction interval. *r* = Pearson correlation coefficient. Data shown are generated from 2-4 technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line. N = 3 RNA-seq replicates for each isogenic cell line.

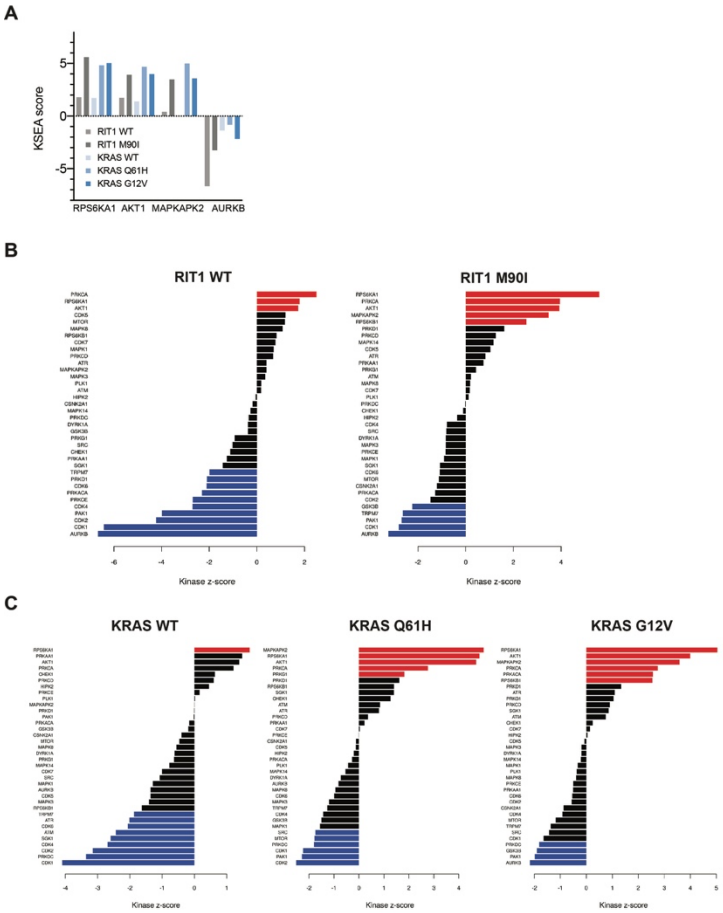


Figure 2.7. Phosphoproteome profiling identifies enhanced phosphorylation of specific kinase substrates in KRAS- and RIT1-mutant cells.

A, KSEA analysis of AALE phosphoproteomes. Top differentially phosphorylated kinase substrates are shown. Marker selection analysis identifies differentially phosphorylated sites in KRAS-mutant cells compared to RIT1-mutant cells. Phosphosites from KRAS-mutant and $RIT1^{M90I}$-mutant replicate-level phosphoproteome profiles were compared by two-tailed t-test. **B,** KSEA of phosphoproteome data for $RIT1^{WT}$ and $RIT1^{M90I}$-expressing AALE cells. The kinase z-score indicates the overall score for each kinase listed, normalized by the total number of substrates. Significant scores ($P<0.05$) are indicated in red and blue. Phosphosites of kinases in red were more highly abundant in the cell line compared to vector control, whereas phosphosites of kinases in blue were more highly abundant in vector control than the indicated cell line. **C,** KSEA of phosphoproteome data for KRAS-expressing AALE cells. Labeling as in B.

Supplementary Figure 3

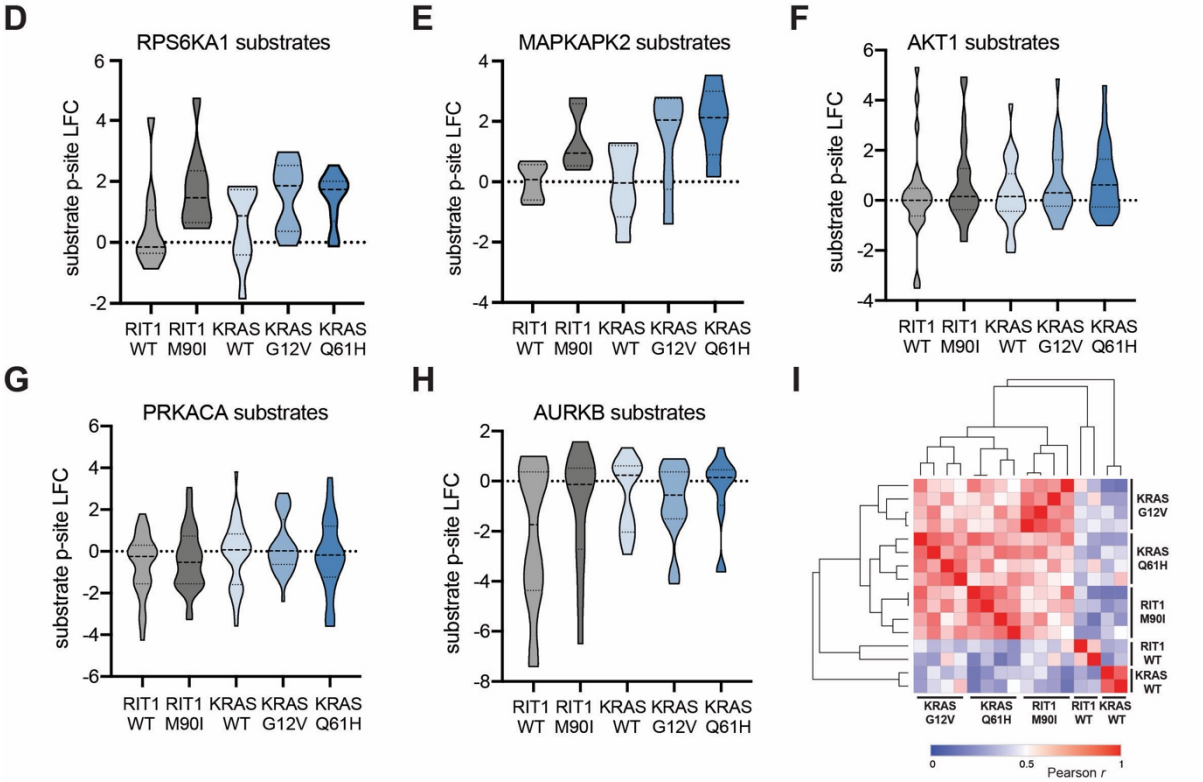


Figure 2.8. Phosphoproteome profiling identifies enhanced phosphorylation of specific kinase substrates in KRAS- and RIT1-mutant cells.

D, Violin plot of phosphosite abundance in RIT1- or KRAS-expressing AALE cells of phosphosites that are RPS6KA1 substrates. **E,** Violin plot of phosphosite abundance in RIT1- or KRAS-expressing AALE cells of phosphosites that are MAPKAPK2 substrates. **F,** Violin plot of phosphosite abundance in RIT1- or KRAS-expressing AALE cells of AKT1 substrate sites. **G,** Violin plot of phosphosite abundance of PRKCA substrate sites. **H,** Violin plot of phosphosite abundance of AURKB substrate sites. **I,** Pairwise replicate correlation (Pearson *r*) heatmap and unsupervised hierarchical clustering of phosphoproteome data from AALE cells. Data shown are generated from 2-4 technical LC-MS/MS replicates per cell line generated from N=2 biological replicates per cell line.

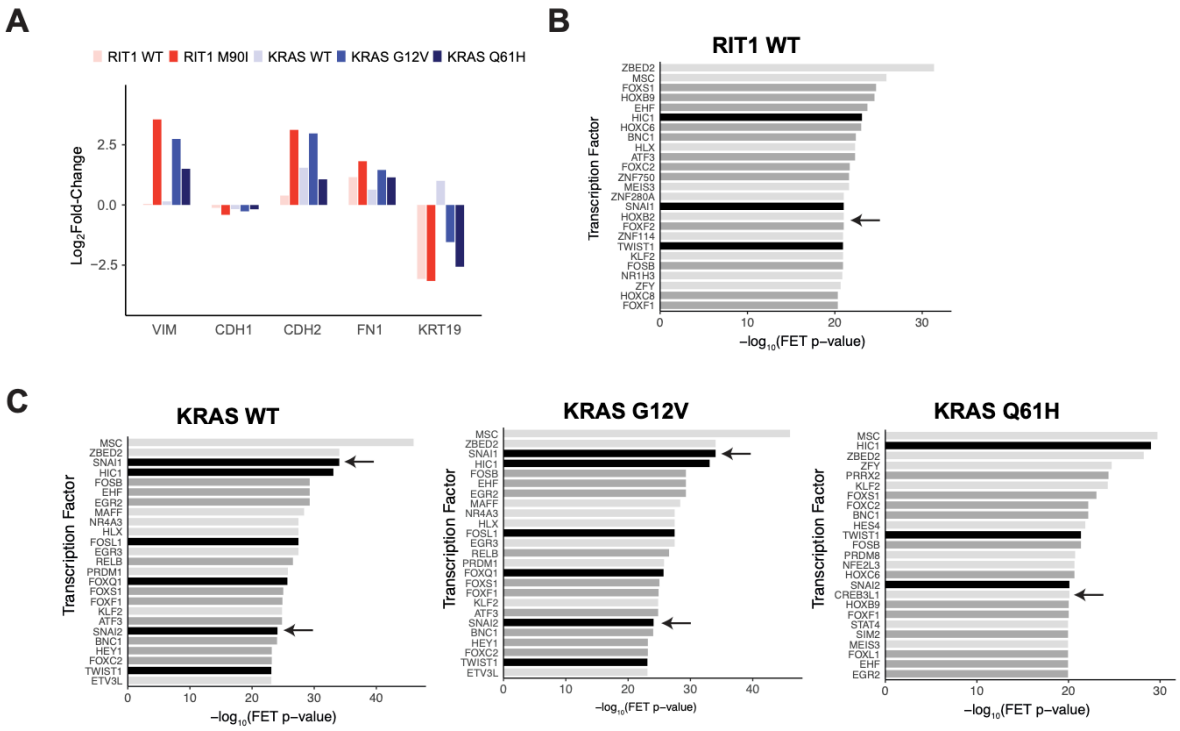


Figure 2.9. RIT1 and KRAS promote epithelial-to-mesenchymal transition.

A, Log_2Fold-Change in mRNA levels of EMT genes *VIM*, *CDH2*, *FN1*, and *KRT19*, in each isogenic cell line compared to vector control. N=3 RNA-seq replicates per cell line. **B,** Transcription factor target enrichment analysis using Enrichr libraries of differentially expressed genes in $RIT1^{WT}$ cells. FET, Fisher's exact test. N = 3 $RIT1^{WT}$ replicates compared to 3 vector controls. Arrows indicate proteins of the Snail family. Black = confirmed EMT genes in dbEMT. Dark gray = EMT-associated genes in literature. **C,** Same as **B** for $KRAS^{WT}$, $KRAS^{G12V}$, and $KRAS^{Q61H}$ cells.

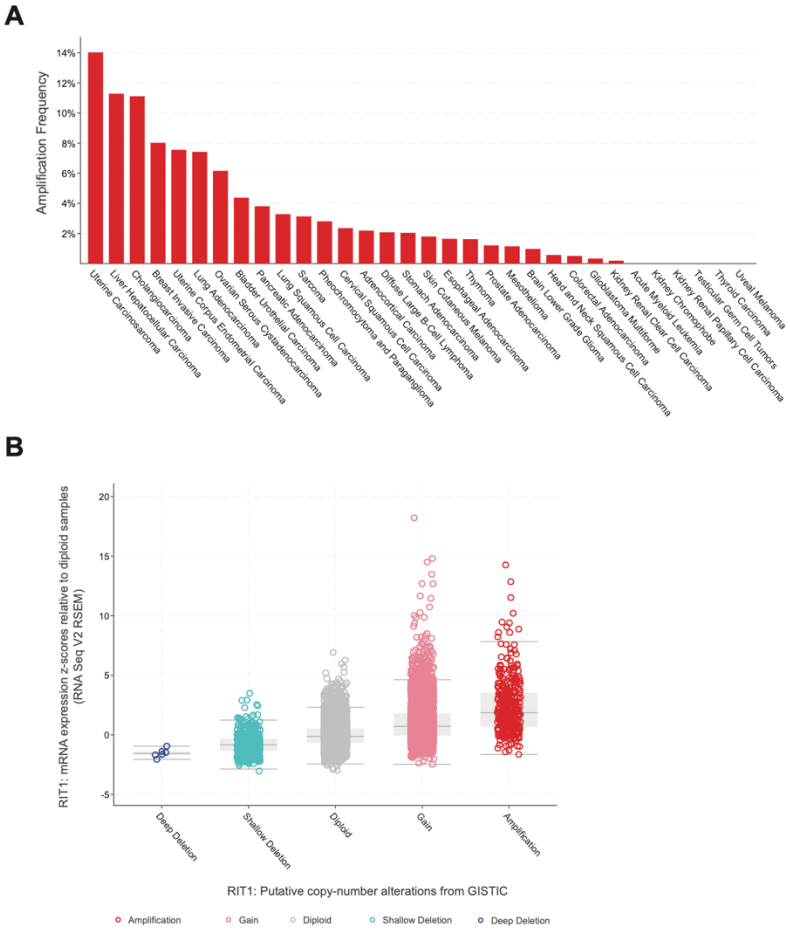


Figure 2.10. Amplification of *RIT1* across cancer types.

A, RIT1 amplification frequency across tumor types in the TCGA Pan-Cancer Atlas. **B,** mRNA expression levels in Pan-Cancer tumors with varying RIT1 copy-number alterations, z-scores normalized to diploid samples. Box plot shows median, 25th percentile, and 75th percentile values. Whisker lines extend to 1.5x the Interquartile range (IQR), or the maximum and minimum data points if these are within the IQR.

Chapter 3. ONCOGENIC KRAS PROMOTES ABERRANT RNA SPLICING IN LUNG CANCER

A version of this chapter is published on bioRxiv:

A. Lo, M. McSharry, A. H. Berger, Oncogenic KRAS alters splicing factor phosphorylation and alternative splicing in lung cancer. *bioRxiv.* 2022. https://doi.org/10.1101/2022.05.20.492866

I led this work and my contributions to this paper included performing experiments, analyzing transcriptome data, and writing the manuscript.

Background: Alternative RNA splicing is widely dysregulated in cancers including lung adenocarcinoma, where aberrant splicing events are frequently caused by somatic splice site mutations or somatic mutations of splicing factor genes. However, the majority of mis-splicing in cancers is unexplained by these known mechanisms. We hypothesize that the aberrant Ras signaling characteristic of lung cancers plays a role in promoting the alternative splicing observed in tumors.

Methods: We recently performed transcriptome and proteome profiling of human lung epithelial cells ectopically expressing oncogenic KRAS and another cancer-associated Ras GTPase, RIT1. Unbiased analysis of phosphoproteome data identified altered splicing factor phosphorylation in KRAS-mutant cells, so we performed differential alternative splicing analysis using rMATS to identify significantly altered isoforms in lung epithelial. To determine whether these isoforms were uniquely regulated by KRAS, we performed a large-scale splicing screen in which we generated over 300 unique RNA sequencing profiles of isogenic A549 lung adenocarcinoma cells ectopically expressing 75 different wild-type or variant alleles across 28 genes implicated in lung cancer.

Results: Mass spectrometry data showed widespread downregulation of splicing factor phosphorylation in lung epithelial cells expressing mutant KRAS compared to cells expressing wild-type KRAS. We observed alternative splicing in the same cells, with 2196 and 2416 skipped exon events in $KRAS^{G12V}$ and $KRAS^{Q61H}$ cells, respectively, 997 of which were shared ($p < 0.001$ by hypergeometric test). In the high-throughput splicing screen, mutant KRAS induced the greatest number of differential alternative splicing events, second only to the RNA binding protein RBM45 and its mutant allele $RBM45^{M126I}$. We identified ten high confidence cassette exon events across multiple KRAS variants and cell lines. These included differential splicing of the Myc Associated Zinc Finger (MAZ). As MAZ regulates expression of KRAS, this splice variant may be a mechanism for the cell to modulate wild-type KRAS levels in the presence of oncogenic KRAS.

Conclusion: Proteomic and transcriptomic profiling of lung epithelial cells uncovered splicing factor phosphorylation and mRNA splicing events regulated by oncogenic KRAS. These data suggest that in addition to widespread transcriptional changes, Ras signaling pathways in cancer promote post-transcriptional splicing changes that may contribute to oncogenic processes.

3.1 Background

Lung cancers are the leading cause of cancer death worldwide, with lung adenocarcinomas accounting for over 40% of lung cancers (American Cancer Society 2021). Lung adenocarcinomas, a form of non-small cell lung cancer, exhibit a high somatic mutational burden, with a median of 8.7 exonic somatic mutations per megabase in the tumor genome (Campbell et al. 2016). While most mutations are not known to have functional consequences, an important subset disrupts the pathways and molecular processes that regulate cell growth and proliferation, contributing to carcinogenesis.

One molecular process that is under a high level of regulation is RNA splicing and processing. RNA splicing is required in eukaryotic cells to produce functional proteins and modulate protein abundance. In a phenomenon known as alternative splicing, cells can also harness RNA splicing to produce different versions of proteins from the same gene, contributing to the diversity and functional complexity of the human proteome. Analyses of cancer genomes and transcriptomes show that alternative splicing is dysregulated in a widespread manner in many cancers including in the lung (Kahles et al. 2018). One mechanism by which splicing can be altered in cancer is *cis*-acting splice site mutations which disrupt the recognition of exon splice sites (Jayasinghe et al. 2018). For instance, a mutation in a splice site of the tyrosine kinase gene *MET* leads to exon 14 skipping and a constitutively active protein (Lu et al. 2017; Onozato et al. 2009). Notably, lung cancers harboring this mutation confer clinical sensitivity to MET inhibitors (Frampton et al. 2015; Drilon et al. 2020). RNA splicing can also be altered in lung tumors by *trans*-acting mutations in splicing factor genes including *U2AF1, RBM10,* and *SF3B1* in lung tumors (Brooks et al. 2014; J. Zhao et al. 2017; Imielinski et al. 2012). Improved knowledge of these splicing aberrations and the resulting mis-spliced genes would improve discovery of new small molecule or gene therapies for cancer.

However, most aberrant splicing observed in cancer cannot be explained by known splice site or splicing factor mutations (Li et al. 2017). Another route to aberrant splicing is through dysregulation of signaling pathways upstream of splicing factors. Splicing factors, the proteins that determine where and when splicing occurs, are themselves regulated by intracellular signaling cascades (Zhou and Fu 2013). SR proteins are a family of splicing factors that plays a large role in splice site selection (Bradley, Cook, and Blanchette 2015). The function of SR proteins as recruiters of the spliceosome and catalyzers of the splicing reaction is heavily dependent on their

phosphorylation state (Mermoud, Cohen, and Lamond 1992). As such, these proteins are regulated by kinases including SRPK1 and Clk/Sty (Shepard and Hertel 2009), and changes in their activity are associated with advanced forms of lung adenocarcinomas (Gout et al. 2012). In particular, previous studies show that alternative splicing can be disrupted through the AKT-SRPK-SR protein axis or the TGFb-CLK/sty-SR axis (Zhou et al. 2012; Nowak et al. 2008; Blaustein et al. 2005).

The AKT-SRPK signaling pathway is one of many arms of the receptor tyrosine kinase (RTK) and Ras signaling cascade, a network that regulates many cellular processes and is often altered in cancers (Mukhopadhyay, Vander Heiden, and McCormick 2021). Mutations in *KRAS* are seen in ~27% of lung adenocarcinomas, and in total, ~76% of tumors harbor mutations in RTK-Ras-Raf pathway genes including the tyrosine kinases *EGFR* and *MET* (Imielinski et al. 2012; Cancer Genome Atlas Research Network 2014; Campbell et al. 2016). Despite a large body of work studying the Ras pathway, therapies targeting KRAS have only recently become viable and are currently effective only for the KRAS G12C variant (Canon et al. 2019). Inhibitors targeting EGFR or other RTKs are effective only for a limited time, with acquired resistance a nearly universal occurrence (Passaro et al. 2021; Drosten and Barbacid 2020). Understanding novel downstream processes of this oncogenic signaling cascade would expand the options for therapeutic intervention. In some instances, splicing factors or alternative isoforms can be amenable to inhibition by small molecule inhibitors, providing opportunities to relieve tumor burden (da Silva et al. 2015). Moreover, oligonucleotide therapies for aberrant splicing have proven effective in treating genetic disorders and could be used in precision oncology as well (Finkel et al. 2016).

In this study, we ask which mRNA splicing processes are downstream of and directed by Ras signaling. We employ a controlled experimental system of lung cell lines in order to investigate the effects of specific genetic perturbations on alternative splicing. In doing so, we avoid confounding influences of other contextual differences on splicing and increase the sensitivity to uncover oncogene-regulated aberrant splicing events.

3.2 RESULTS

3.2.1 *Mutant KRAS suppresses splicing factor phosphorylation in lung epithelial cells*

The process of mRNA splicing is known to involve regulation of phosphorylation states (Cho et al. 2011; Aubol et al. 2016). Previously, we generated human AALE lung epithelial cells stably expressing $KRAS^{WT}$, $KRAS^{G12V}$, $KRAS^{Q61H}$, $RIT1^{WT}$, and $RIT1^{M90I}$, and performed liquid chromatography and tandem mass spectrometry (LC-MS/MS) to profile their proteomes and phosphoproteomes (Lo et al. 2021). As expected, phosphorylation of Ras pathway proteins was significantly altered in cells expressing KRAS or RIT1 variants (Figure 3.1A). Interestingly, in cells expressing $KRAS^{G12V}$ or $KRAS^{Q61H}$ variants, we also observed a marked change in the phosphorylation of splicing factor proteins (Figure 3.1A) with several SR proteins showing decreased phosphorylation state (Figure 3.1B). As each phosphosite was normalized to total protein level, and total protein level of these SR proteins was unchanged (Figure 3.5A), the decreased SR protein phosphorylation was likely due to differences in phosphorylation state itself. The phosphosites that showed the greatest decrease in phosphorylation levels (Figure 3.1B) occurred in proteins known to be precisely regulated by phosphorylation: SRSF7, SRSF1, and SRSF2 (Cho et al. 2011; Aubol et al. 2003; Naro and Sette 2013). Moreover, these phosphosites are primarily located in the RS domains and as such are critical for the shuttling and splicing activity of the proteins (Zheng et al. 2020). Notably, alteration of SRSF7, SRSF1, and SRSF2

protein phosphorylation occurred predominantly in KRAS^{mut} cells compared to KRAS^{WT} cells, and not in the RIT1^{M90I} cells compared to RIT1^{WT} cells (Figure 3.1C, 3.5B). These data suggest that KRAS may uniquely alter SR protein phosphorylation state in a manner distinct from another RAS-family GTPase, RIT1.

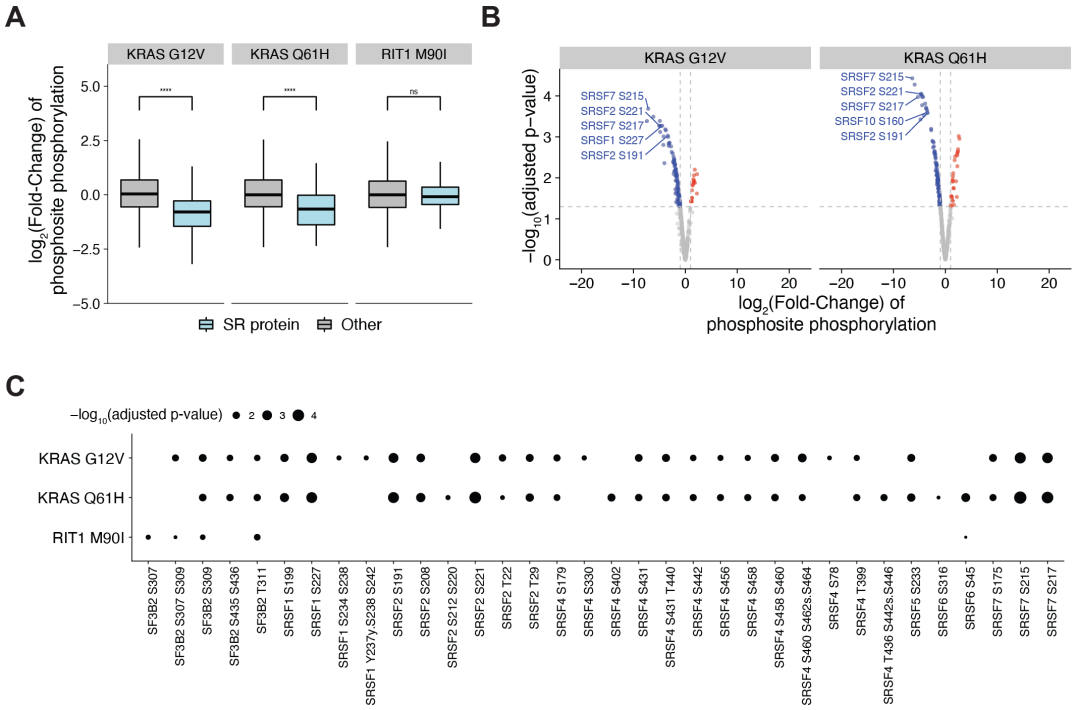


Figure 3.1. Mutant KRAS suppresses splicing factor phosphorylation in lung cells.

A) Phosphorylation changes in SR proteins (blue) or other proteins (gray) in KRAS^{G12V} or KRAS^{Q61H} vs KRAS^{WT} cells, and RIT1^{M90I} vs RIT1^{WT} cells; * $p<0.05$, **** $p<0.0001$ by one-sided Wilcoxon rank-sum test. **B**) Volcano plots of the phosphorylation changes in mRNA splicing proteins. Sites with less phosphorylation in KRAS^{G12V} or KRAS^{Q61H} compared to KRAS^{WT} are shown in blue, sites with more phosphorylation are shown in red. The most significantly altered SR protein phosphosites are labeled. **C**) Phosphorylation changes in phosphosites of RNA splicing proteins of interest SF3B2, SRSF1, SRSF2, SRSF4, and SRSF7.

3.2.2 *Oncogenic KRAS regulates alternative splicing in lung epithelial cells*

Following the observation that splicing factors are differentially phosphorylated in KRAS^{mut} overexpressing cells compared to KRAS^{WT} overexpressing cells, we sought to directly compare RNA splicing regulated by oncogenic KRAS. To this end, we performed RNA sequencing and

differential splicing analysis using rMATS (Shen et al. 2014). Alternative splicing of cassette exons (Figure 3.2A) is tightly regulated by SR proteins (Zhou and Fu 2013; Zheng et al. 2020) and these events are typically the most prevalent form of alternative splicing detected in cancer and development (Scotti and Swanson 2016; Singh and Cooper 2012). The majority of splicing events observed in AALE cells were cassette exon events (70.3%), so we focused on cassette exon events for subsequent analyses. We observed that many cassette exons are more skipped or more promoted in $KRAS^{G12V}$ and $KRAS^{Q61H}$-mutant cells compared to $KRAS^{WT}$ cells (Figure 3.2B). In addition, 997 exons are differentially skipped in both $KRAS^{G12V}$ and $KRAS^{Q61H}$ cells compared to the wild-type cells, indicating that oncogenic KRAS variants cause similar changes in splicing patterns (hypergeometric test for intersect $p < 0.0001$) (Figure 3.2C). The exons most differentially expressed in both $KRAS^{mut}$ cell lines include exons *MOK*, which encodes a member of the MAP kinase superfamily, and *NT5C2*, which encodes a 5′-nucleotidase enzyme implicated in chemotherapy resistance (Qian et al. 2015; Tzoneva et al. 2013) (Figure 3.2D). We observed that despite these changes in splicing activity, mRNA expression of splicing factors is unaltered in these cells, indicating that phospho-signaling is the primary method of regulation (Figure 3.6). These shared cassette exon events point to a change in splicing regulation driven by activated KRAS.

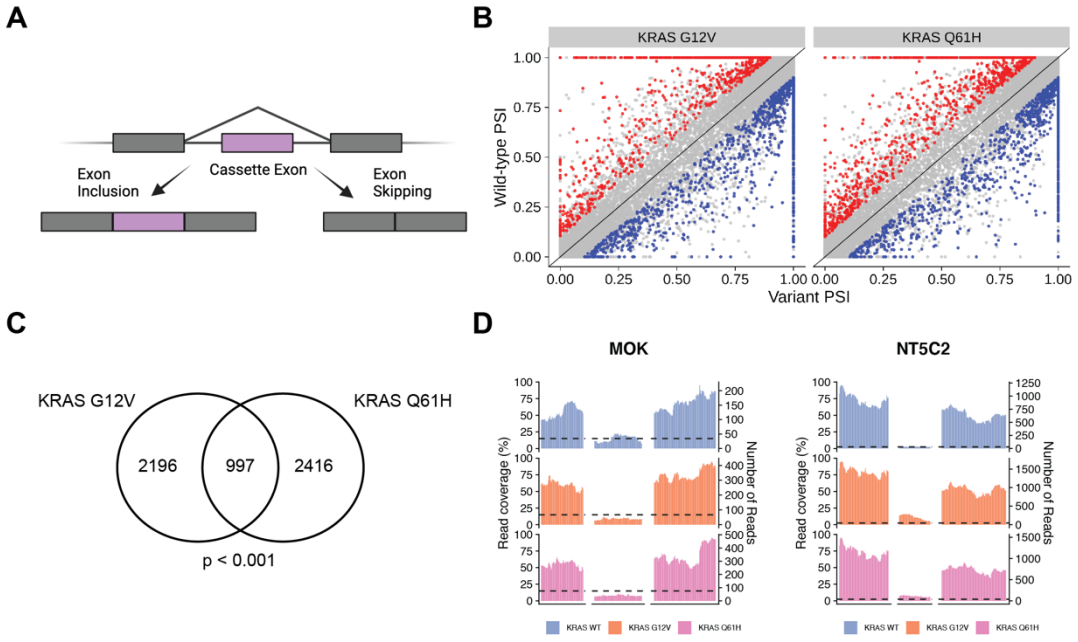


Figure 3.2. Oncogenic KRAS regulates alternative splicing in lung epithelial cell.

A) Schematic of the inclusion or skipping of a cassette exon. **B**) Differentially spliced exons in $KRAS^{G12V}$ and $KRAS^{Q61H}$ cells compared to $KRAS^{WT}$ cells. PSI = Percent Spliced In. Red = exons promoted at least 10% more in $KRAS^{WT}$ than $KRAS^{mut}$. Blue = exons promoted at least 10% more in $KRAS^{mut}$ than $KRAS^{WT}$. FDR < 0.05. **C**) Overlap of differentially spliced exons in $KRAS^{G12V}$ vs $KRAS^{WT}$ and $KRAS^{Q61H}$ vs $KRAS^{WT}$. P-value calculated by modeling skipped exon events as a hypergeometric distribution. **D**) Read coverage plots of skipped exon events in *MOK* and *NT5C2*, normalized by total RNA-seq library size and RNA composition. Left axis = percent coverage relative to region. Right axis = absolute read coverage.

3.2.3 *A large-scale transcriptome screen as a platform for splicing discovery*

To further study how KRAS and related oncogenes regulate alternative splicing, we performed a large-scale perturbation screen in A549 lung adenocarcinoma cells. Isogenic A549 cell lines were generated by lentiviral transduction in 384 well format as described previously (Alice H. Berger et al. 2016) with multiple biological replicates per variant. 4 days after transduction, cells were lysed for transcriptome analysis using the SmartSeq low input method (Figure 3.3A). To identify oncogene-driven effects on transcription and splicing, we chose 28 genes to study based on high mutation frequency in lung adenocarcinomas and relevance in RNA processing and signaling pathways (Alice H. Berger et al. 2016). In total, we generated whole transcriptome profiles for 374

unique replicates of arrayed isogenic cell lines expressing 79 different lentiviral constructs which include vectors for wild-type genes, variant alleles, and controls. For each replicate used in subsequent analyses, we confirmed allele overexpression by direct analysis of the RNA-seq data (Figure 3.7B-F)

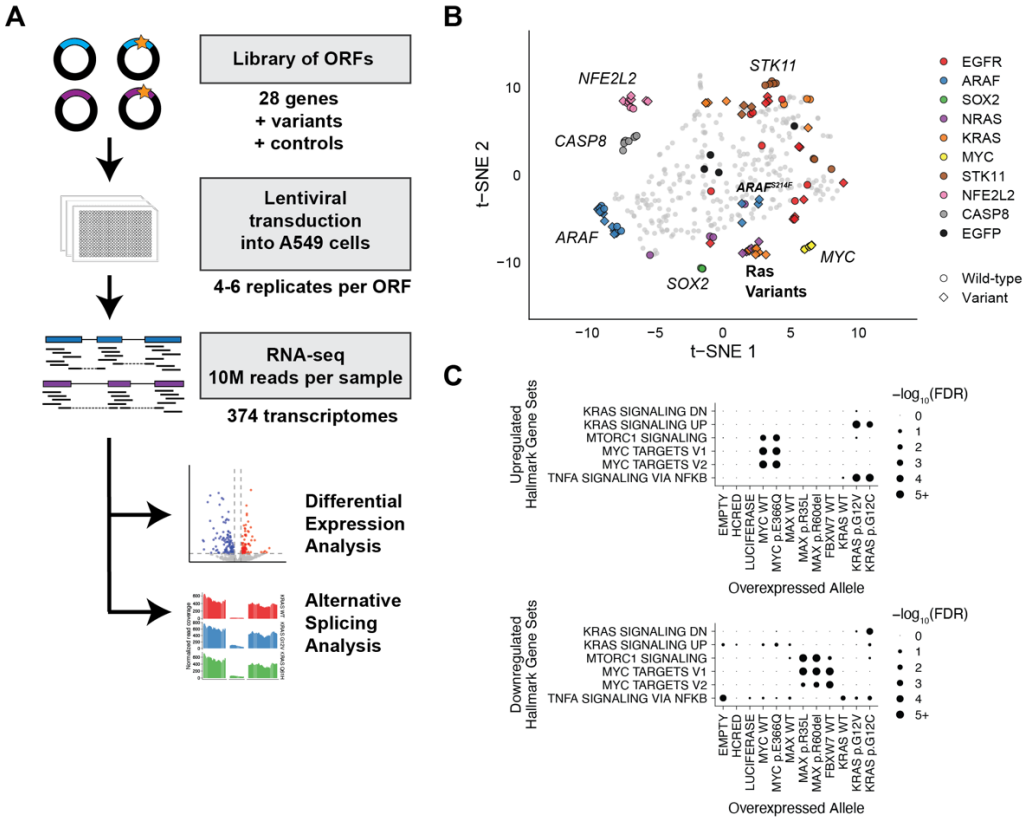


Figure 3.3. A large-scale transcriptome screen as a platform for splicing discovery.

A) Experimental workflow for RNA-seq analysis of lung adenocarcinoma genes and variants. **B**) Transcriptome profiles of cells expressing different gene constructs. Dimensionality reduction performed using t-distributed stochastic neighbor embedding (t-SNE) (Van Der Maaten and Hinton 2008). Each point represents an experimental replicate. Genes and alleles of particular interest are colored and labeled. Controls are circles colored in black. Circle = wild-type allele. Diamond = variant allele. **C**) Gene set analysis with GOseq of up- or down-regulated transcripts in MYC, KRAS, FBXW7, and negative control cells. Hallmark gene sets from mSigDB.

Comparing gene expression signatures across the full experiment highlighted the distinct transcriptomic effects due to the ectopic expression of different open reading frames (ORFs). For

example, ectopic expression of transcription factors such as *MYC, SOX2,* and *NFE2L2* resulted in consistent expression shifts across the transcriptome which are reflected by replicates which group close together (Figure 3.3B). Expression profiles also highlighted the functional differences between wild-type and variant alleles of oncogenes. Consistent with previous findings, cells with activating ARAF^{S214F} showed similarities to *KRAS* and *NRAS* variant cells, while wild-type *ARAF*, neutral variant ARAF^{V145L}, and the kinase-inactivated variants ARAF$^{S214F/D429A}$ and ARAF$^{S214C/D429A}$ did not (Alice H. Berger et al. 2016).

To determine how ectopic expression of each gene affected the transcriptome, we inspected the differentially expressed genes between cell lines expressing an experimental ORF and cell lines expressing a control vector. As expected, overexpression of the transcription factor *MYC* induced high levels of *MYC* expression (Figure 3.7A). Gene set analysis further showed that the transcripts with increased expression in *MYC*-overexpressing cells include known MYC target genes (Figure 3.3C). Conversely, overexpression of known MYC inhibitors such as *FBXW7* decreased the expression of MYC target genes. Expression of two inactivating *MAX* mutations, but not wild-type *MAX*, also suppressed *MYC* target gene expression (Yeh, Bellon, and Nicot 2018; Augert et al. 2020) (Figure 3.3C). These findings demonstrate that this large-scale gene expression profiling screen can recapitulate known biology and support the utility of these data for further discovery.

One potentially confounding factor for this system is the natural endogenous KRAS^{G12S} variant found in A549 cells. To verify that we were observing the activity of activated KRAS from ectopic expression of KRAS variants in this background, we first ensured that KRAS was overexpressed by a significant amount. We defined significant allele overexpression as two standard deviations above endogenous KRAS expression (Figure 3.7B). Then we performed gene set analysis on A549 cells ectopically expressing KRAS^{WT}, KRAS^{G12V}, or KRAS^{G12C}. Genes

involved in upregulation of KRAS signaling were enriched in $KRAS^{G12V}$ and $KRAS^{G12C}$ cells but not $KRAS^{WT}$ cells. Similarly, another gene set we saw enriched in $KRAS^{mut}$ cells was the Nf-kb signaling pathway (Figure 3.3C), which oncogenic KRAS is known to activate (Barbie et al. 2009; Daniluk et al. 2012). These and previous data suggest that despite their endogenous *KRAS* mutation, A549 cells can be used as a model for KRAS signaling upon further KRAS perturbation (Alice H. Berger et al. 2016)

3.2.4 *A screen for splicing alterations in lung cancer identifies alternative splicing events regulated by KRAS*

The use of RNA sequencing rather than probe-based array or bead technologies for transcriptome analysis provides the opportunity to measure alternative splicing in addition to differential gene expression. In both the data from AALE cells and the large-scale screen in A549 cells, we performed differential splicing analysis to first compare replicates with ORF expression against replicates expressing a control vector. We observed a mean of 456 differentially spliced events per overexpressed wild-type or variant allele and fewer than 1000 events in all except wild-type RBM45, which induced 1496 total differential alternative splicing events compared to vector control (Figure 3.9A-B, z-score = 4.51). Overexpression of the cancer-associated $RBM45^{D434Y}$ allele also induced more alternative splicing events than average (z-score = 1.53).

Notably, several of the ORFs that perturbed splicing the most were transcription factors. These included *NFE2L2, SOX2,* or *MYC* (Figure 3.9A-B). Cells overexpressing wild-type *NFE2L2* harbored 931 alternative splicing events compared to vector controls (z-score = 2.06). Similarly, cells with wild-type *MYC* or the activated MYC^{E366Q} variant exhibited 787 and 726 differential splicing events, respectively (z-scores = 1.44 and 1.17). These increases may reflect a regulatory

role played by mRNA splicing to adjust protein abundance levels when transcription activity and pre-mRNA levels are disrupted (Han et al. 2017).

We next compared cells overexpressing each variant allele to cells overexpressing the corresponding wild-type allele. Again, we observed the highest number of differential splicing events in RBM45 variant cells compared to RBM45 wild-type cells (z-score = 2.94, Figure 3.4A). Among differentially expressed exons, 62% were more included in the variant cells than in RBM45 wild-type cells. This imbalance suggests a role for RBM45 in promoting exon skipping that is disrupted by mutations. Additionally, of the 835 total differentially spliced events in $RBM45^{M126I}$ and 418 in $RBM45^{D434Y}$ cells compared to wild-type RBM45, 154 were shared between the two groups (hypergeometric test of intersect, $p < 0.001$), indicating these variants cause phenotypically similar changes to protein function.

Besides RBM45, the variant with the greatest alternative splicing effect compared to the wild-type gene was $KRAS^{G12C}$ with 729 total differential splicing events including 440 cassette exons. Similarly, $KRAS^{G12V}$ cells differentially splice 409 events compared to $KRAS^{WT}$ cells including 232 cassette exons. (Figure 3.4A, 3.9C). Since G12V, G12C, and Q61H are all KRAS activating variants that promote deregulated downstream MAPK signaling (Cook et al. 2021), we took advantage of our parallel analyses of multiple KRAS variants across two cellular contexts to identify cassette exon splice variants that are similarly differentially regulated in all four $KRAS^{mut}$ AALE and A549 cell lines compared to the respective $KRAS^{WT}$ cells. Using this approach, we identified 10 high-confidence alternative cassette exons that are consistently altered when KRAS is activated across the two different cellular contexts and three KRAS variants (Figure 3.4B). The approximate degree of splicing change (percent spliced in) induced by $KRAS^{G12V}$ was similar in A549 and AALE cells (Figure 3.4B; Pearson's coefficient = 0.71, p-value = 0.020).

As our phosphoproteomic analysis indicated changes in the activity of SR proteins, we asked if these genes are differentially expressed. In most cells, including KRAS and EGFR overexpressing cells, there was no notable change in the gene expression of splicing factors such as *SF3B1, SRSF1, SRSF2,* or *SRSF7* (Figure 3.8A). We thus conclude that, in our experimental system and with our driver genes of interest, splicing factors are not substantially regulated at the mRNA level, and their contribution to the alternative splicing changes we observe are phosphorylation based.

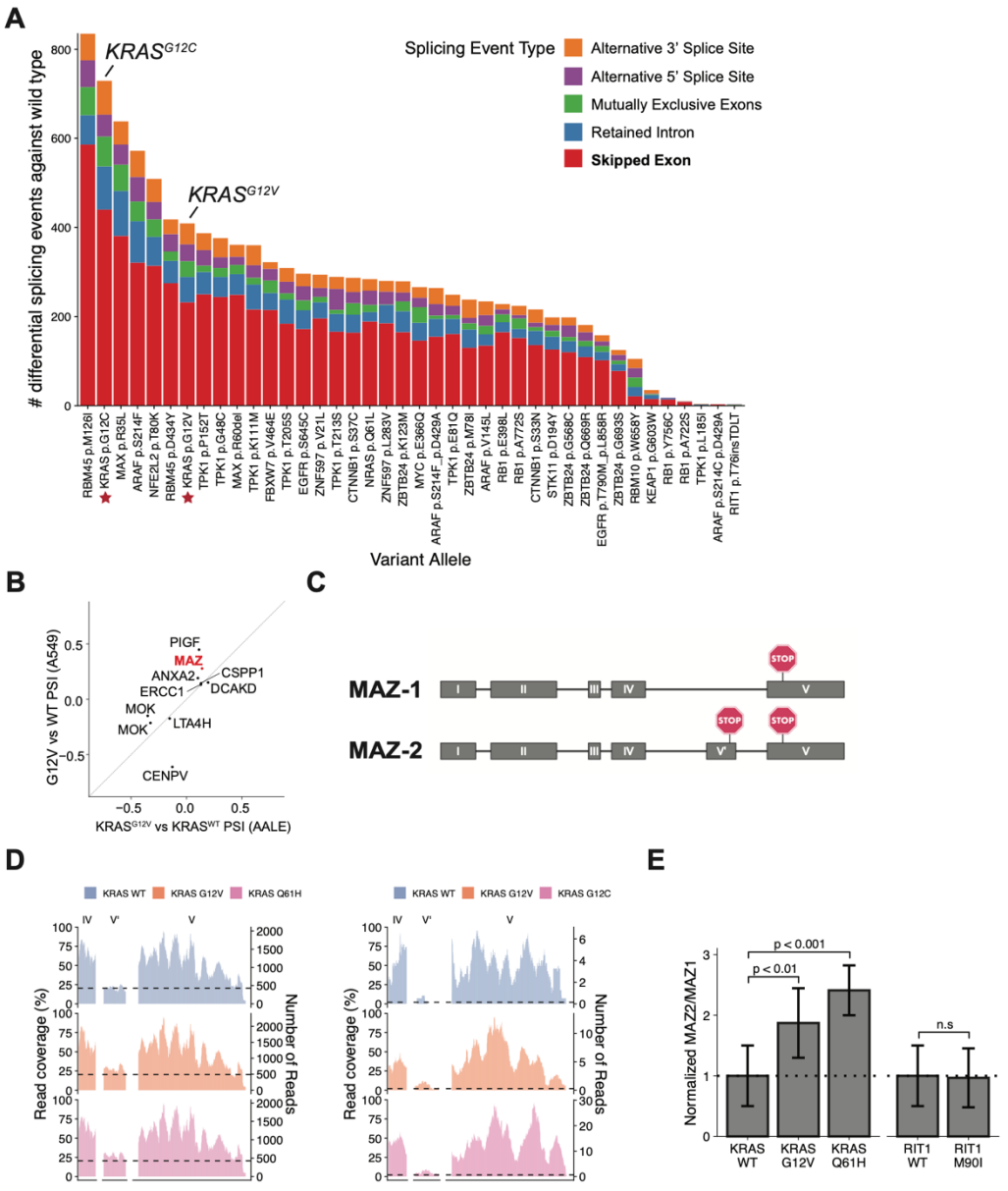


Figure 3.4. A screen for splicing alterations in lung cancer identifies alternative splicing events regulated by KRAS.

A) Number of events differentially spliced between cells with variant alleles and cells with respective wild-type alleles. $KRAS^{G12C}$ and $KRAS^{G12V}$ are starred and labeled. Colors represent the five splicing event categories considered. FDR < 0.05. ΔPSI > 10%. **B**) Change in exon levels in cassette exons differentially expressed in both $KRAS^{G12V}$ compared to $KRAS^{WT}$ and $KRAS^{G12C}$ compared to $KRAS^{WT}$ (Pearson's coefficient = 0.71, p-value = 0.020). ΔPSI = Change in Percent Spliced In. **C**) Schematic of MAZ-1 and MAZ-2 isoforms with labeled stop codons. **D**) Percent read coverage of MAZ exon V' in $KRAS^{WT}$, $KRAS^{G12V}$, and $KRAS^{G12C}$ overexpressing A549 cells (left) and $KRAS^{WT}$, $KRAS^{G12V}$, and $KRAS^{Q61H}$ overexpressing AALE cells (right), relative to the expression of flanking exons. Inset zooms in on cassette exon. **E**) Relative levels of MAZ-2

compared to MAZ-1 isoforms in AALE cells overexpressing KRAS or RIT1 alleles, normalized to $KRAS^{WT}$ or $RIT1^{WT}$, as measured by quantitative RT-PCR. Error bars = standard error. Statistical significance was determined by two-tailed t-test.

Of the ten splice variants identified, we focused on the isoforms of the transcription factor Myc-associated zinc finger protein (MAZ) due to its previously characterized isoforms and its role in increasing the expression of KRAS and HRAS (Yang et al. 2019). The major mRNA isoform MAZ-1 is composed of five exons whereas isoform MAZ-2 has six exons, with the inclusion of exon V' (Figure 3.4C). This additional exon includes an early stop codon, leading to a protein isoform with an alternative C terminus (Ray et al. 2002). In both AALE and A549 cells, expression of $KRAS^{mut}$ resulted in an increase in exon V' inclusion ($\Delta PSI = .14$ to .34, FDR < .01) (Figure 3.4D). These alternatively spliced isoforms were validated through quantitative RT-PCR showing a lower MAZ-2 to MAZ-1 isoform ratio in $KRAS^{WT}$ overexpressing AALE cells (Figure 3.4E). Thus, we found that $KRAS^{WT}$ and $KRAS^{mut}$ overexpressing cells display distinct MAZ isoform levels, suggesting that alternative splicing of the MAZ gene is regulated downstream of KRAS.

3.1 Discussion

Here we asked whether oncogenic signaling by KRAS and other oncogenes in lung cancer can perturb alternative splicing, possibly contributing to the mis-splicing that is so pervasive in human cancer (Climente-González et al. 2017). By expressing KRAS and other oncogenes in lung epithelial and lung adenocarcinoma cells and performing whole transcriptome analysis, we found that KRAS activation leads to differential changes in alternative RNA splicing. One alternative splicing event of note was an exon inclusion event in *MAZ*, which can regulate the expression of *KRAS* itself and *HRAS* (Ray et al. 2002). As a possible explanation for how these alternative splicing changes might occur, we identified decreased phosphorylation of SR proteins SRSF1, SRSF2, and SRSF7 in KRAS-mutant cells by LC-MS/MS. In contrast, we did not observe a change

in protein abundance nor a change in gene expression of known splicing factors by KRAS variants. Thus, KRAS regulation of splicing factors occurs largely through a phospho-signaling network, underlying the importance of studying post-translational modifications. To deepen understanding of how oncogenic KRAS alters splicing factor phosphorylation, further work is needed to determine and validate if the dysregulated sites are phosphorylated by the canonical regulators of SR proteins SRPK1 or Clk/Sty, or if other kinases may be in play (Ngo et al. 2005).

SR proteins are critical in splice site selection in both constitutive and alternative RNA splicing as they are responsible for recognizing exon splicing enhancers (ESEs) and recruiting the spliceosome (Zhou and Fu 2013). Therefore, to modulate isoform abundance, both the abundance and activity of SR proteins in the cell, as well as the presence of ESEs in the alternate splice sites are important. SR protein activity is primarily regulated by the protein kinases Clk1 and SRPK1 (Aubol et al. 2016). In the major isoform of MAZ, exon V is preferentially selected over exon V' likely due to the greater number of ESEs recognized by SR proteins. However, when select SR proteins are downregulated, exon V' may be spliced in more often. In our study of KRAS-associated splicing of *MAZ*, we did not observe a change in SR protein abundances, suggesting that if the shift in splicing of *MAZ* is due to SR protein function, it is caused primarily by changes in SR phosphorylation.

By studying RNA, protein, and phosphorylation levels, we interrogated RNA splicing regulation driven by multiple oncogenes and their tumor-associated variants, identifying splicing factors and alternate exons of interest. Further studies of these alternative splicing mechanisms would open the possibility of designing treatments for mis-splicing in lung adenocarcinomas. For example, oligonucleotides can be designed and engineered to target particular splice sites in tumor cells (Havens and Hastings 2016; Levin 2019). Similarly, existing small molecule inhibitors may

be used to correct splicing factor dysregulation, for example, inhibition of SRPK1 to modulate SR protein activity (Oltean et al. 2012).

Alternative isoforms in cancers also introduce neoepitopes specific to the tumor (Kahles et al. 2018). Thus, understanding oncogene-driven alternative splicing also holds implications for immune-based treatment of tumors with alterations in these oncogenes, including the design of CAR-T cells to target novel tumor neo-antigens resulting from splicing events. Further characterization of the alternate exon events associated with KRAS mutations may explain why some KRAS-mutant tumors respond to immunotherapies more favorably than others (Jeanson et al. 2019). However, our study is limited to effects *in vitro*, isolated from the immune environment, and further investigation of oncogene-driven splicing *in vivo* and in patient tumors is needed. Additionally, although we chose to focus on exon skipping events, other forms of alternative splicing will also be critical to consider as we determine the impact to tumorigenesis and treatment (Smart et al. 2018).

This work shows that diverse oncogenic signaling programs in lung cancer induce both characteristic transcriptional changes as well as alterations in alternative splicing regulation. Together with direct cis-acting somatic splicing genetic variation and trans-acting mutations in splicing factors themselves, these signaling-related splicing patterns can provide a partial explanation for the altered splicing observed in cancer. Continued large-scale studies are needed to map splicing changes to particular signaling programs and mechanistic studies are needed to better understand exactly how these programs connect to splicing. These efforts will in turn create new opportunities for therapeutic intervention in cancer via splicing modulation or immunotherapies, and both modalities may be beneficial in combination with established therapies to finally provide the durable cancer cures that have largely proven elusive.

3.2 Materials and Methods

3.2.1 *LC-MS/MS proteomics of AALE cells*

Proteomic and phosphoproteomic data from AALE cells was generated by LC-MS/MS and quantified as previously described (Lo et al. 2021). Briefly, isogenic AALE cells were generated by transduction with lentiviral vectors encoding wild-type KRAS, KRASG12V, KRASQ61H, or wild-type RIT1 or RIT1M90I. Replicate lysates from each cell line were digested and labeled with TMT mass tags for quantitative mass spectrometry. Two TMT 10-plexes were used with AALE vector control lysates shared across plexes for data demultiplexing. For phosphoproteome analysis, phosphorylated peptides were enriched by IMAC column purification prior to LC-MS/MS analysis. MS data were interpreted using the Spectrum Mill software package v6.0 prerelease. All proteomic data are shared publicly in the public proteomics repository MassIVE (https://massive.ucsd.edu) and are accessible at ftp://MSV000085225@massive.ucsd.edu with username: MSV000085225 and password: oncogenic.

3.2.2 *RNA sequencing*

RNA-seq libraries of KRAS- and RIT1-mutant AALE cells were previously described (Lo et al. 2021). Briefly, three replicates per cell line were used for RNA isolation by Direct-zol RNA mini-prep (Zymo) and libraries generated using the TruSeq RNA Library kit (Illumina). Libraries were sequenced on an Illumina NovaSeq at the Fred Hutch Genomics Shared Resource to an average coverage of 70 million 50 bp paired-end reads per sample. RNA-sequencing reads were mapped to the human genome reference hg19/GRCh37 using STAR 2.5.3a. Data is publicly available at the NCBI Gene Expression Omnibus database with accession number GSE146479.

For the A549 large-scale splicing screen, A549 lung cancer cells were transduced at high multiplicity-of-infection in 384 well plates using pLX317-ORF constructs as previously described with 4-6 biological replicates per ORF (Alice H. Berger et al. 2016). Cells were not selected with puromycin but parallel plates were treated with puromycin or left untreated and then cell viability was determined using CellTiterGlo reagent (Promega) for calculation of infection efficiency. 96 hours post-transduction, cells were lysed using TCL buffer (Qiagen) and lysates were stored at -80 degrees. We adapted library preparation protocols previously developed for single cell RNA sequencing (Trombetta et al. 2014). RNA was isolated from 9 µl lysate and 1 µl of ERCC 1:1000 Spike in control per sample using RNA SPRI beads (Agencourt) in lo-bind microcentrifuge tubes (Eppendorf). Beads were washed several times in 80% ethanol before drying and used for cDNA synthesis. First-strand cDNA synthesis was performed with the Smart-seq v4 Ultra Low Input kit (Takara) using the 3' Smart-seq CDS Primer IIA and Smart-seq V4 oligonucleotide. Next, whole transcriptome amplification and clean-up was performed using the Kapa HiFi HotStart kit with the following thermal cycling protocol: 3min 98°C; 20 cycles of 15s at 98°C, 15s 67°C, 6min at 72°C; final extension of 5min at 72°C. DNA was then cleaned up using DNA SPRI beads (Agencourt). Individual samples were inspected by Qubit quantitation and TapeStation (Agilent) analysis and diluted to 0.1 to 0.2 ng/µl prior to library construction using the Nextera XT sequencing kit (Illumina). Final libraries were pooled and sequenced on an Illumina HiSeq (Northwest Genomics Center) to an average coverage of 10M 75 bp paired-end reads per replicate. RNA-sequencing reads were mapped to the human genome reference hg19/GRCh37 using STAR 2.5.3a.

3.2.3 *Differential gene expression analysis*

Transcripts were quantified with featureCounts from the Subread package v.1.5.3 (Liao, Smyth, and Shi 2014) using RefSeq gene annotations. Differential expression was then determined using edgeR v.3.30.3 (Robinson, McCarthy, and Smyth 2009), normalizing transcript counts by library size and RNA composition scale factors computed by using trimmed mean of M-values (TMM) between sample pairs (Robinson and Oshlack 2010). Whole transcriptome profiles were quality checked for sufficient RNA-seq coverage. Samples were filtered for successful expression of the lentiviral vector, determined by a 2 standard deviation increase in expression of the target gene in the sample compared to negative controls. Gene transcripts were also filtered for sufficient detection across the experiment as determined by a mean logCPM > 0.1 for each gene. After applying quality control filters, the A549 dataset included 374 transcriptome profiles each quantifying 12543 genes. Analysis pipeline is available as a Snakemake workflow at https://github.com/bergerbio/RNA-splicing-screen.

3.2.4 *Differential splicing analysis*

After alignment of RNA-seq reads by STAR, alternative splicing events were identified and quantified by rMATS-turbo 4.1.1 using gene transcript annotations from gencode v.19 (Shen et al. 2014). Both reads spanning exon junctions and reads covering single exons were used for splicing quantification. Analysis pipeline is available as a Snakemake workflow along with the in-house R scripts used to aggregate rMATS results and perform additional analyses and interpretations (https://github.com/bergerbio/RNA-splicing-screen).

3.2.5 *Gene Set Analysis*

Analysis of enrichment of KRAS signaling in differential RNA expression profiles was performed in R with GOseq (M. D. Young et al. 2010). KRAS signaling gene sets were taken from MSigDB hallmark gene sets (Liberzon et al. 2015).

3.2.6 *Quantitative reverse transcription polymerase chain reaction (qRT-PCR)*

Total RNA was extracted using Direct-zol RNA Miniprep plus (Zymo Research) from each biological replicate of AALE cell lines overexpressing KRAS or RIT1 alleles or control vector. From the extracted RNA, 1 µg was reverse transcribed into cDNA using SuperScript IV First-Strand Synthesis (Invitrogen). Quantitative RT-PCR reactions were set up in technical triplicates with BioRad iQ SYBR Green kit and analyzed on a BioRad CFX384 Real-Time System. PCR primers were designed to detect MAZ isoforms without alternate exon V' (Forward: ATG GGA GGC AGC TTT CGC ; Reverse: TCA CCA GTA CCT TTG TTG CA) and with exon V' (Forward: AGCTCTGCAACAAAGGCTTC ; Reverse: GGGCAGGGGTCTTGCA). PCR products were confirmed to be specific by molecular weight analysis via gel electrophoresis. Quantification of isoforms in experimental samples were normalized against vector control samples and relative quantification of MAZ isoforms 1 and 2 were calculated with the Livak method.

3.3 ACKNOWLEDGEMENTS

We thank Rob Bradley, Daniela Witten, Cole Trapnell, David MacPherson, and Berger and MacPherson lab members for critical advice and discussion. We thank Angela Brooks, Xiaoyun Wu, Jesse Boehm, and the Broad Institute Genetic Perturbation Platform for our previous collaboration in generating the large-scale screening platform. We thank Fil Mundt, Philip Merins, and Steve Carr and the Broad Institute Proteomics Platform for our prior collaboration in

generating the proteomic data. We thank the Northwest Genomics Center and Fred Hutch Genomics Shared Resource for generation of RNA sequencing data. This research was funded, in part, through the National Cancer Institute (NCI) R37 CA25205 to A.H.B., NIH/NCI Cancer Center Support grant P30 CA015704, and NSF IGERT DGE-1258485 to A.L. A.H.B. reports consulting fees from Puma Biotechnology.

Proteomic data analyzed in the current study are available in the public proteomics repository MassIVE (https://massive.ucsd.edu) and are accessible at ftp://MSV000085225@massive.ucsd.edu with username: MSV000085225 and password: oncogenic. RNA-seq data generated and analyzed in the current study are available in the Gene Expression Omnibus (GEO) repository and are accessible at [GEO 1], [GEO 2], [GEO 3].

A.H.B. designed and supervised the study and provided study resources. A.L., M.M., and A.H.B. performed experiments. A.L. performed all computational analyses of the proteomics, transcriptomics, and splicing. A.L. and A.H.B. wrote the manuscript. All authors read and approved of the final manuscript.

3.4 SUPPLEMENTAL FIGURES

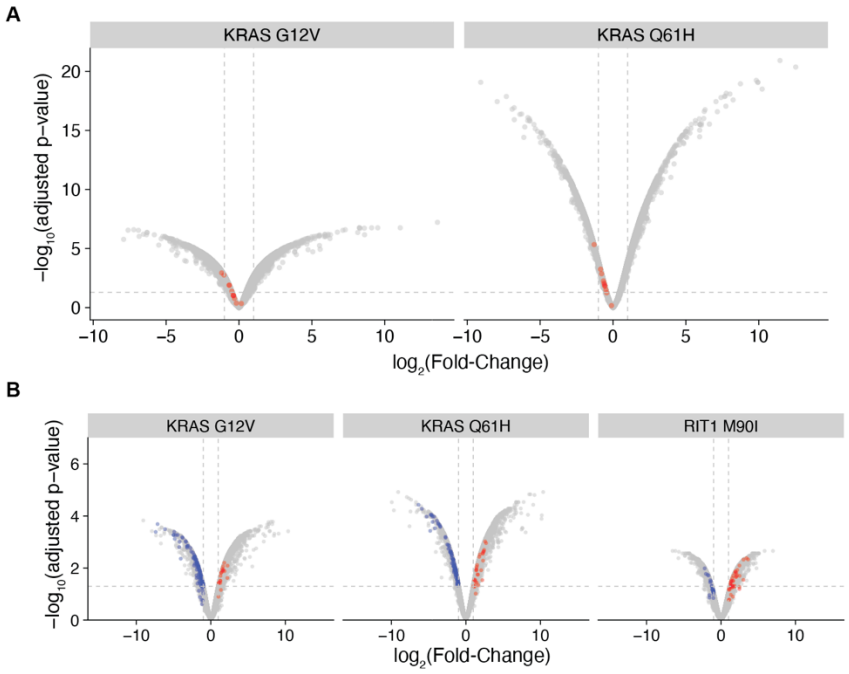


Figure 3.5. Protein abundance and phosphosite changes in splicing factors.

A) Volcano plot of differential protein abundance in $KRAS^{G12V}$ and $KRAS^{Q61H}$ cells compared to $KRAS^{WT}$ cells. Labeled in red are SR proteins. B) Volcano plot of phosphorylation of phosphosites in $KRAS^{G12V}$ and $KRAS^{Q61H}$ cells compared to $KRAS^{WT}$ cells, and $RIT1^{M90I}$ compared to $RIT1^{WT}$. Labeled are phosphosites on proteins in the GO RNA SPLICING gene set which are downregulated (blue) or upregulated (red).

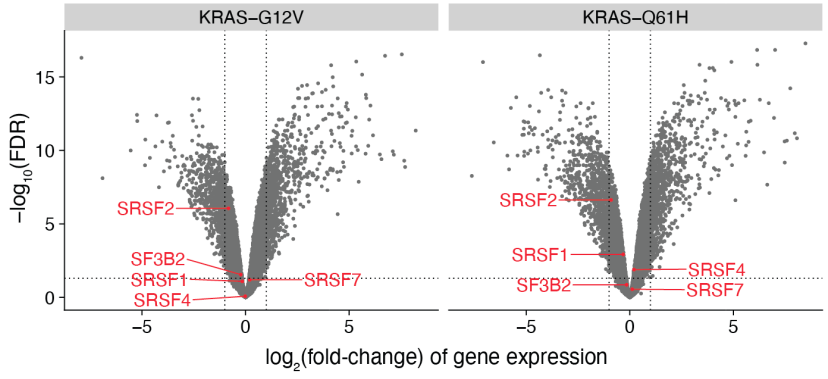


Figure 3.6. Gene expression changes of splicing factors.

A) Volcano plot of mRNA differential expression. Labeled in red are splicing factors of interest SF3B2, SRSF1, SRSF2, SRSF4, and SRSF7. B) Change in overall gene expression compared to change in cassette exon inclusion.

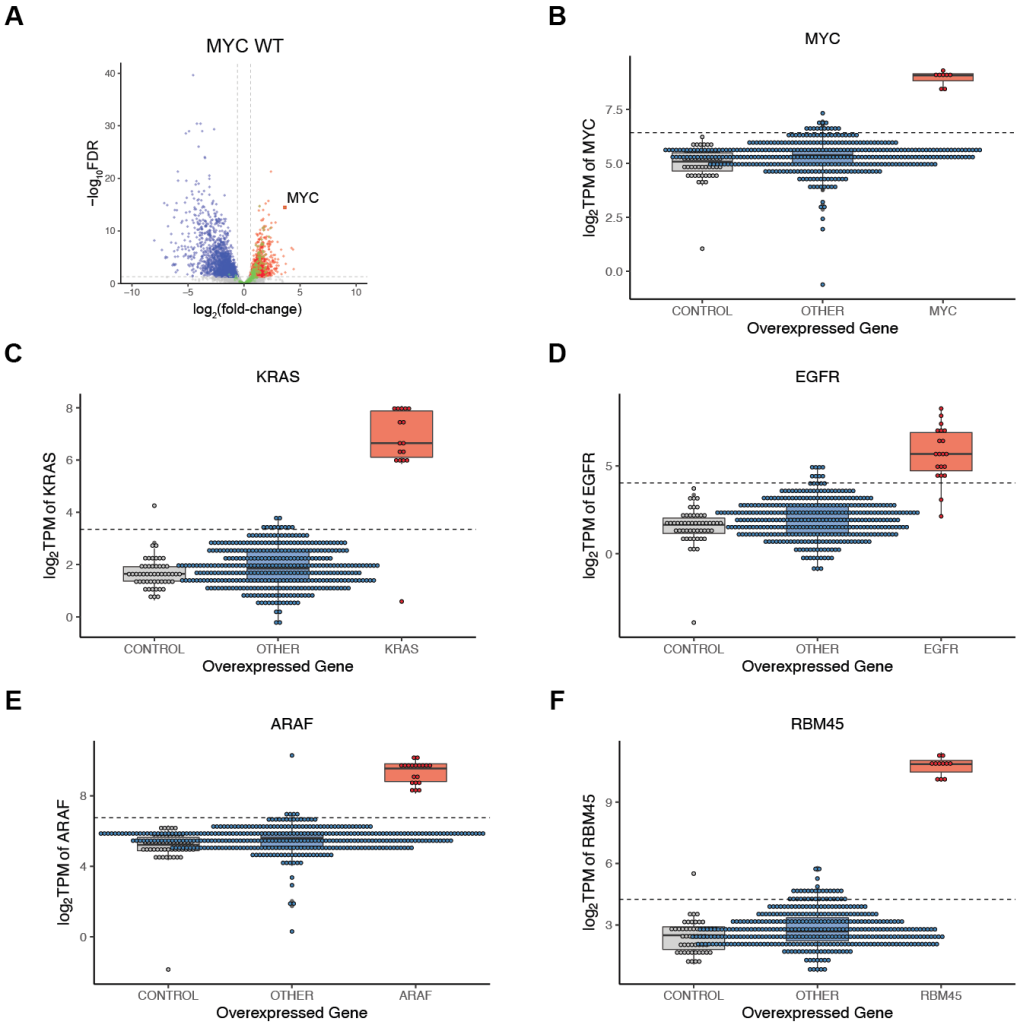


Figure 3.7. Large-scale genetic perturbation screen by overexpression.

A) Volcano plot of differentially expressed genes in wild-type *MYC* overexpressing cells. Blue = downregulated transcripts. Red = upregulated transcripts. Green = transcripts in mSigDB hallmark gene sets MYC targets V1 and V2. B) mRNA expression levels of *MYC* in vector control cells (grey), cells overexpressing non-MYC alleles (blue), and cells overexpressing *MYC* alleles (red). C) Same as B) for *KRAS*. D) Same as B) for *EGFR*. E) Same as B) for *ARAF*. F) Same as B) for *RBM45*.

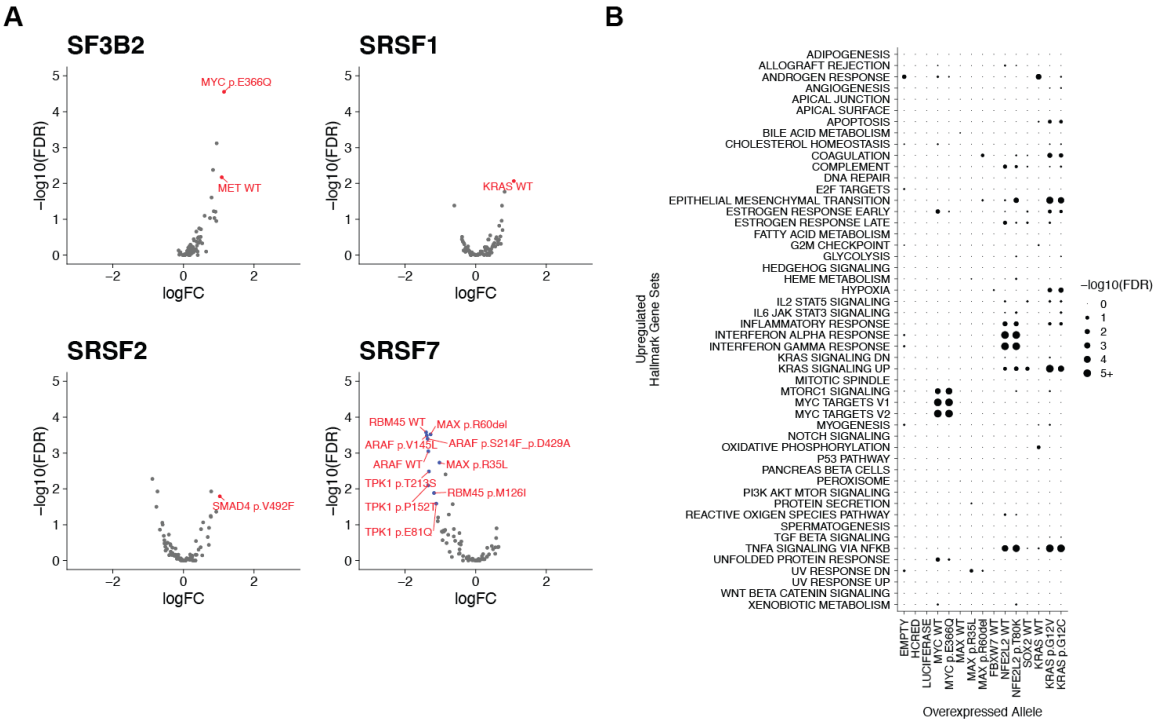


Figure 3.8. Regulation of gene expression of cancer-associated genes.

A) mRNA expression of splicing factors *SF3B1*, *SRSF1*, and *SRSF7*. B) Gene set enrichment analysis of downregulated transcripts in MYC, KRAS, FBXW7, and negative control cells. Hallmark gene sets from mSigDB.

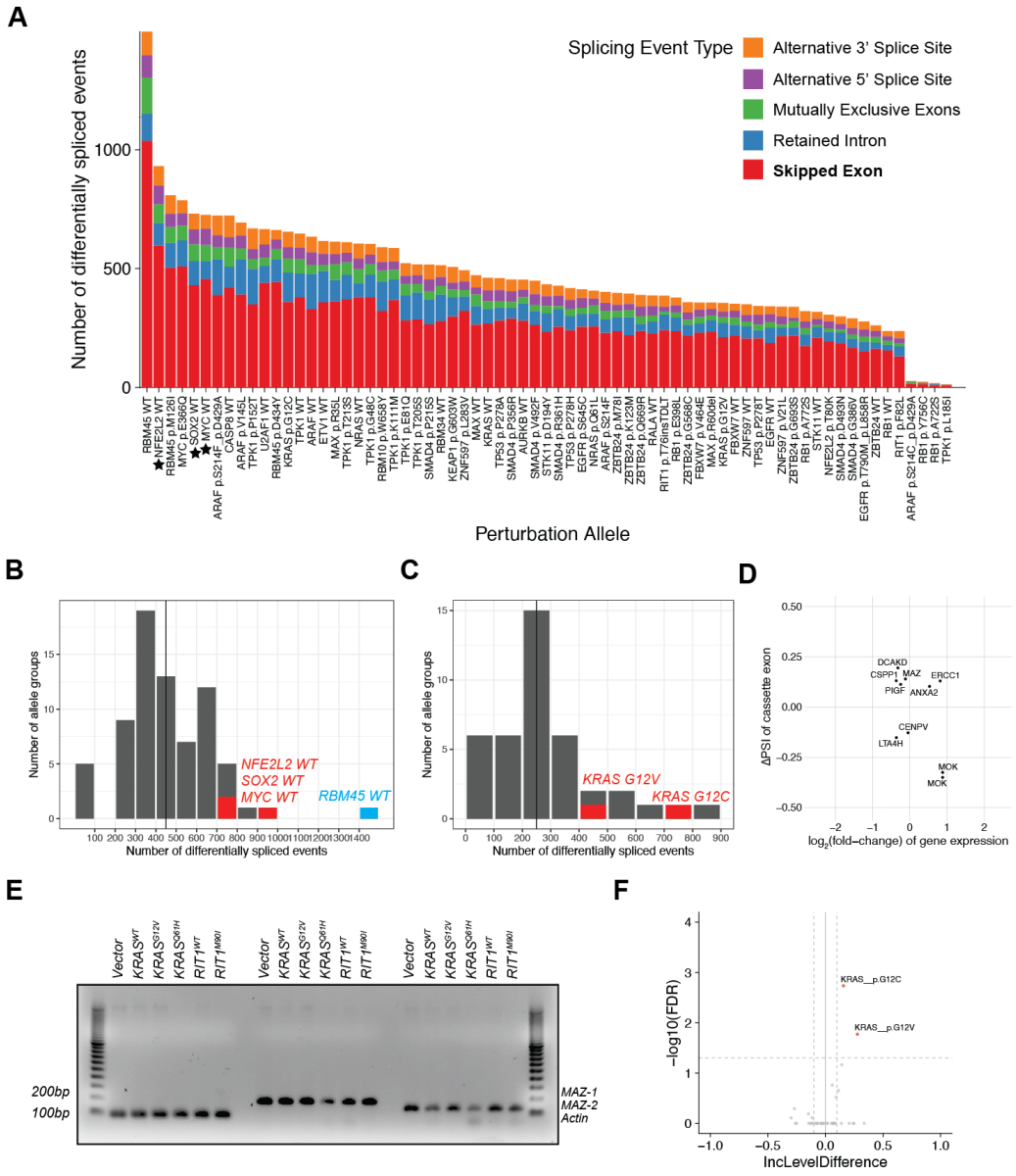


Figure 3.9. Large-scale screen for alternative splicing discovery

A) Number of events differentially spliced between cells with genetic perturbations and cells with vector controls. Colors represent 5 major alternative splicing event categories. Wild-type alleles of transcription factors are starred. B) Number of events differentially spliced between cells with genetic perturbations and cells with vector controls. Vertical line = mean number of events across screen. C) Number of events differentially spliced between cells with variant alleles and cells with respective wild-type alleles. Vertical line = mean number of events across screen. D) Comparing change in Percent Spliced In (ΔPSI) of cassette exons and the differential expression of the corresponding gene transcript. E) PCR gel detecting MAZ-1 and MAZ-2 isoforms in AALE cells overexpressing vector, KRAS, or RIT1 alleles. F) Differential splicing of MAZ exon V' in variant alleles compared to respective wild-type alleles.

Chapter 4. DISCUSSION

Diagnosis and treatment of cancer has transformed in the last two decades. Clinical genotyping is now standard thanks to the developments of Next Generation Sequencing which allow low-cost targeted sequencing panels, whole exome sequencing, and even whole genome sequencing of tumor samples. However, a full picture understanding of how oncogenic mutations lead to disease remains out of our grasp and despite more treatment options than ever, rates of progression-free survival remain low in lung cancer (Soria et al. 2018). In this dissertation, I presented my work towards understanding the cellular processes moderated by the lung cancer oncogenes *KRAS* and *RIT1*. I focused on the regulation of alternative RNA splicing, protein abundance levels, and the post-translational modifications and signaling activity driving their dysregulation.

By examining RIT1- and KRAS-associated proteomes and transcriptomes, we found that increasing cellular RIT1 levels induces similar phenotypes as introducing a RIT1 or KRAS mutation. Though we are in early days of studying RIT1 in the context of cancer, these observations concur with recent literature showing that the hotspot $RIT1^{M90I}$ mutation functions by increasing cellular RIT1 protein levels (Castel et al. 2019). By elucidating the relationship between RIT1 mutants and wild-type RIT1, and between RIT1 and KRAS, we move towards the future of developing therapies effective for the 10-15% of lung adenocarcinomas with *RIT1* mutations or amplifications.

Next, by examining the transcriptomic effects of KRAS on both the differential expression and alternative splicing levels, we expose potential downstream vulnerabilities in Ras-activated cancer cells. As more KRAS inhibitors enter the clinic, knowing more Ras targets and effectors will help us find opportunities for more potent inhibitors and will help us identify possible pathways of resistance.

4.1 FUTURE OF PRECISION ONCOLOGY

KRAS was characterized as an oncogene in the 1980's, yet it was not until 2020 that a targeted therapy for KRAS-mutant cancers was available to patients. This leap has been met with much excitement and enthusiasm in the oncology community, but it is still only the beginning. As with TKIs for EGFR-mutant tumors, the current inhibitors are only appropriate for a subset of KRAS-mutant tumors and in those, acquired resistance develops soon after treatment. Due to these limitations, it is critical to map out all oncogenic processes regulated by Ras and to map the possible paths of escape (Moore and Malek 2021).

Just as genomic sequencing technology led to characterization of somatic variation in cancers, technological developments in the past 5-10 years are now leading to more detailed and context-aware molecular characterizations of tumors. CRISPR screens dissect the exact functional causes of drug resistance and tumor growth, asking not just whether single genes are oncogenes or tumor suppressors, but whether particular isoforms or paralog families promote tumor growth (Parrish et al. 2021; Thomas et al. 2020). Spatial genomics and transcriptomics begin to define the physical constraints of tumor growth in addition to the subclonal landscape of tumors (T. Zhao et al. 2021). Patient-derived organoids supplement patient-derived xenografts as experimental models to understand cancer evolution, drug resistance, and interplay with the immune environment (Karlsson et al. 2022). And long-read sequencing improves characterization of not just somatic variants and known gene fusions, but also novel genome alterations and RNA isoforms (Nattestad et al. 2018). Together, these tools paint a truer and more nuanced picture of cancer.

With technological advances providing new perspectives to cancer development, it is more important than ever to probe how oncogenes dysregulate molecular processes within different

contexts. My hope is that the work I have presented here and the accompanying publicly available data provides a baseline for how KRAS and RIT1 affect the proteome and transcriptome of lung cancer cells and encourages the field to ask more questions about how the Ras signaling pathway modulates cell growth. To extend our findings and move it closer to the clinic, future work is needed to test patterns of post-transcriptional regulation with *in vivo* models and analyses of multi-omic patient data. Though these will be non-trivial efforts, they will define avenues for additional or alternative therapies for Ras-activated tumors.

The field and practice of precision oncology is already benefiting from the above technology advances and better knowledge of splicing regulation and signaling pathways. By taking advantage of tumor-specific splicing factor mutations, new therapies can hone in on cells with certain isoforms (North et al. 2022). One can imagine that similar therapies can be designed for SR proteins and other splicing factors dysregulated by Ras oncogenes. Immunotherapy options including checkpoint inhibitors and CAR-T cells can also be applied and designed in an increasingly individualized manner by taking note of neoepitopes resulting from dysregulated RNA processing (Kahles et al. 2018).

The future of personalized oncology is bright. Despite setbacks, ever better first-, second-, and third-line therapies with fewer comorbidities are now available for cancer patients and more will undoubtedly be introduced in the next few years. Soon, we will be able to map and predict each tumor's subclones and evolutionary path and with this information prescribe highly individualized combination therapies to effectively stop tumor progression in both the present and the future. It will be more complex than oncologists envisioned when prescribing the first TKIs, but it will truly be cures for cancers.

BIBLIOGRAPHY

Ambrogio, Chiara, Jens Köhler, Zhi-Wei Zhou, Haiyun Wang, Raymond Paranal, Jiaqi Li, Marzia Capelletti, et al. 2018. "KRAS Dimerization Impacts MEK Inhibitor Sensitivity and Oncogenic Activity of Mutant KRAS." *Cell* 172 (4): 857-868.e15.

American Cancer Society. 2021. "Cancer Facts & Figures 2021." https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2021/cancer-facts-and-figures-2021.pdf.

Aoki, Yoko, Tetsuya Niihori, Toshihiro Banjo, Nobuhiko Okamoto, Seiji Mizuno, Kenji Kurosawa, Tsutomu Ogata, et al. 2013. "Gain-of-Function Mutations in RIT1 Cause Noonan Syndrome, a RAS/MAPK Pathway Syndrome." *American Journal of Human Genetics* 93 (1): 173–80.

Aubol, Brandon E., Sutapa Chakrabarti, Jacky Ngo, Jennifer Shaffer, Brad Nolen, Xiang-Dong Fu, Gourisankar Ghosh, and Joseph A. Adams. 2003. "Processive Phosphorylation of Alternative Splicing Factor/Splicing Factor 2." *Proceedings of the National Academy of Sciences of the United States of America* 100 (22): 12601–6.

Aubol, Brandon E., Guowei Wu, Malik M. Keshwani, Maliheh Movassat, Laurent Fattet, Klemens J. Hertel, Xiang-Dong Fu, and Joseph A. Adams. 2016. "Release of SR Proteins from CLK1 by SRPK1: A Symbiotic Kinase System for Phosphorylation Control of Pre-MRNA Splicing." *Molecular Cell* 63 (2): 218–28.

Augert, Arnaud, Haritha Mathsyaraja, Ali H. Ibrahim, Brian Freie, Michael J. Geuenich, Pei-Feng Cheng, Sydney P. Alibeckoff, et al. 2020. "MAX Functions as a Tumor Suppressor and Rewires Metabolism in Small Cell Lung Cancer." *Cancer Cell* 38 (1): 97-114.e7.

Barbie, David A., Pablo Tamayo, Jesse S. Boehm, So Young Kim, Susan E. Moody, Ian F. Dunn, Anna C. Schinzel, et al. 2009. "Systematic RNA Interference Reveals That Oncogenic KRAS-Driven Cancers Require TBK1." *Nature* 462 (7269): 108–12.

Bechara, Elias G., Endre Sebestyén, Isabella Bernardis, Eduardo Eyras, and Juan Valcárcel. 2013. "RBM5, 6, and 10 Differentially Regulate NUMB Alternative Splicing to Control Cancer Cell Proliferation." *Molecular Cell* 52 (5): 720–33.

Berger, A. H., M. Imielinski, F. Duke, J. Wala, N. Kaplan, G-X Shi, D. A. Andres, and M. Meyerson. 2014. "Oncogenic RIT1 Mutations in Lung Adenocarcinoma." *Oncogene* 33 (35): 4418–23.

Berger, Alice H., Angela N. Brooks, Xiaoyun Wu, Yashaswi Shrestha, Candace Chouinard, Federica Piccioni, Mukta Bagul, et al. 2016. "High-Throughput Phenotyping of Lung Cancer Somatic Mutations." *Cancer Cell* 30 (2): 214–28.

Blaustein, Matías, Federico Pelisch, Tamara Tanos, Manuel J. Muñoz, Diego Wengier, Leandro Quadrana, Jeremy R. Sanford, et al. 2005. "Concerted Regulation of Nuclear and Cytoplasmic Activities of SR Proteins by AKT." *Nature Structural & Molecular Biology* 12 (12): 1037–44.

Bradley, Todd, Malcolm E. Cook, and Marco Blanchette. 2015. "SR Proteins Control a Complex Network of RNA-Processing Events." *RNA* 21 (1): 75–92.

Brooks, Angela N., Peter S. Choi, Luc De Waal, Tanaz Sharifnia, Marcin Imielinski, Gordon Saksena, Pedamallu Chandra Sekhar, et al. 2014. "A Pan-Cancer Analysis of Transcriptome Changes Associated with Somatic Mutations in U2AF1 Reveals

Commonly Altered Splicing Events." *PloS One* 9 (1). https://doi.org/10.1371/journal.pone.0087361.

Cai, Weikang, Shaun W. Carlson, Jennifer M. Brelsfoard, Catherine E. Mannon, Carole L. Moncman, Kathryn E. Saatman, and Douglas A. Andres. 2012. "Rit GTPase Signaling Promotes Immature Hippocampal Neuronal Survival." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 32 (29): 9887–97.

Campbell, Joshua D., Anton Alexandrov, Jaegil Kim, Jeremiah Wala, Alice H. Berger, Chandra Sekhar Pedamallu, Sachet A. Shukla, et al. 2016. "Distinct Patterns of Somatic Genome Alterations in Lung Adenocarcinomas and Squamous Cell Carcinomas." *Nature Genetics* 48 (6): 607–16.

Cancer Genome Atlas Research Network. 2014. "Comprehensive Molecular Profiling of Lung Adenocarcinoma." *Nature* 511 (7511): 543–50.

Cancer Genome Atlas Research Network, John N. Weinstein, Eric A. Collisson, Gordon B. Mills, Kenna R. Mills Shaw, Brad A. Ozenberger, Kyle Ellrott, Ilya Shmulevich, Chris Sander, and Joshua M. Stuart. 2013. "The Cancer Genome Atlas Pan-Cancer Analysis Project." *Nature Genetics* 45 (10): 1113–20.

Canon, Jude, Karen Rex, Anne Y. Saiki, Christopher Mohr, Keegan Cooke, Dhanashri Bagal, Kevin Gaida, et al. 2019. "The Clinical KRAS(G12C) Inhibitor AMG 510 Drives Anti-Tumour Immunity." *Nature*, October. https://doi.org/10.1038/s41586-019-1694-1.

Casado, Pedro, Juan-Carlos Rodriguez-Prados, Sabina C. Cosulich, Sylvie Guichard, Bart Vanhaesebroeck, Simon Joel, and Pedro R. Cutillas. 2013. "Kinase-Substrate Enrichment Analysis Provides Insights into the Heterogeneity of Signaling Pathway Activation in Leukemia Cells." *Science Signaling* 6 (268): rs6.

Casalino, Laura, Dario De Cesare, and Pasquale Verde. 2003. "Accumulation of Fra-1 in Ras-Transformed Cells Depends on Both Transcriptional Autoregulation and MEK-Dependent Posttranslational Stabilization." *Molecular and Cellular Biology* 23 (12): 4401–15.

Castel, Pau, Alice Cheng, Antonio Cuevas-Navarro, David B. Everman, Alex G. Papageorge, Dhirendra K. Simanshu, Alexandra Tankka, Jacqueline Galeas, Anatoly Urisman, and Frank McCormick. 2019. "RIT1 Oncoproteins Escape LZTR1-Mediated Proteolysis." *Science* 363 (6432): 1226–30.

Castel, Pau, Ann Holtz-Morris, Yongwon Kwon, Bernhard P. Suter, and Frank McCormick. 2020. "DoMY-Seq: A Yeast Two-Hybrid-Based Technique for Precision Mapping of Protein-Protein Interaction Motifs." *The Journal of Biological Chemistry*, November. https://doi.org/10.1074/jbc.RA120.014284.

Chen, Haiquan, Jian Carrot-Zhang, Yue Zhao, Haichuan Hu, Samuel S. Freeman, Su Yu, Gavin Ha, et al. 2019. "Genomic and Immune Profiling of Pre-Invasive Lung Adenocarcinoma." *Nature Communications* 10 (1): 5472.

Cherniack, Andrew D., Hui Shen, Vonn Walter, Chip Stewart, Bradley A. Murray, Reanne Bowlby, Xin Hu, et al. 2017. "Integrated Molecular Characterization of Uterine Carcinosarcoma." *Cancer Cell* 31 (3): 411–23.

Cho, Suhyung, Amy Hoang, Rahul Sinha, Xiang-Yang Zhong, Xiang-Dong Fu, Adrian R. Krainer, and Gourisankar Ghosh. 2011. "Interaction between the RNA Binding Domains of Ser-Arg Splicing Factor 1 and U1-70K SnRNP Protein Determines Early Spliceosome Assembly." *Proceedings of the National Academy of Sciences of the United States of America* 108 (20): 8233–38.

Climente-González, Héctor, Eduard Porta-Pardo, Adam Godzik, and Eduardo Eyras. 2017. "The Functional Impact of Alternative Splicing in Cancer." *Cell Reports* 20 (9): 2215–26.

Cook, Joshua H., Giorgio E. M. Melloni, Doga C. Gulhan, Peter J. Park, and Kevin M. Haigis. 2021. "The Origins and Genetic Interactions of KRAS Mutations Are Allele- and Tissue-Specific." *Nature Communications* 12 (1): 1808.

Daniluk, Jaroslaw, Yan Liu, Defeng Deng, Jun Chu, Haojie Huang, Sebastian Gaiser, Zobeida Cruz-Monserrate, Huamin Wang, Baoan Ji, and Craig D. Logsdon. 2012. "An NF-KB Pathway-Mediated Positive Feedback Loop Amplifies Ras Activity to Pathological Levels in Mice." *The Journal of Clinical Investigation* 122 (4): 1519–28.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29 (1): 15–21.

Drilon, Alexander, Jeffrey W. Clark, Jared Weiss, Sai-Hong Ignatius Ou, D. Ross Camidge, Benjamin J. Solomon, Gregory A. Otterson, et al. 2020. "Antitumor Activity of Crizotinib in Lung Cancers Harboring a MET Exon 14 Alteration." *Nature Medicine* 26 (1): 47–51.

Drosten, Matthias, and Mariano Barbacid. 2020. "Targeting the MAPK Pathway in KRAS-Driven Tumors." *Cancer Cell* 37 (4): 543–50.

Dvinge, Heidi, and Robert K. Bradley. 2015. "Widespread Intron Retention Diversifies Most Cancer Transcriptomes." *Genome Medicine* 7 (1): 1–13.

Dvinge, Heidi, Eunhee Kim, Omar Abdel-Wahab, and Robert K. Bradley. 2016. "RNA Splicing Factors as Oncoproteins and Tumour Suppressors." *Nature Reviews. Cancer* 16 (7): 413–30.

Elgort, Marc G., John M. O'Shea, Yike Jiang, and Donald E. Ayer. 2010. "Transcriptional and Translational Downregulation of Thioredoxin Interacting Protein Is Required for Metabolic Reprogramming during G(1)." *Genes & Cancer* 1 (9): 893–907.

Emil F. Pai, Wolfgang Kabsch, Ute Krengel, Kenneth C. Holmes, Jacob John & Alfred Wittinghofer. 1989. "Structure of the Guanine-Nucleotide-Binding Domain of the Ha-Ras Oncogene P21 in the Triphosphate Conformation." *Nature* 2: 1105–11.

Escobar-Hoyos, Luisa F., Alex Penson, Ram Kannan, Hana Cho, Chun-Hao Pan, Rohit K. Singh, Lisa H. Apken, et al. 2020. "Altered RNA Splicing by Mutant P53 Activates Oncogenic RAS Signaling in Pancreatic Cancer." *Cancer Cell* 38 (2): 198-211.e8.

Fang, Zhenhao, Christopher B. Marshall, Jiani C. Yin, Mohammad T. Mazhab-Jafari, Geneviève M. C. Gasmi-Seabrook, Matthew J. Smith, Tadateru Nishikawa, Yang Xu, Benjamin G. Neel, and Mitsuhiko Ikura. 2016. "Biochemical Classification of Disease-Associated Mutants of RAS-like Protein Expressed in Many Tissues (RIT1)." *The Journal of Biological Chemistry* 291 (30): 15641–52.

Finkel, Richard S., Claudia A. Chiriboga, Jiri Vajsar, John W. Day, Jacqueline Montes, Darryl C. De Vivo, Mason Yamashita, et al. 2016. "Treatment of Infantile-Onset Spinal Muscular Atrophy with Nusinersen: A Phase 2, Open-Label, Dose-Escalation Study." *The Lancet* 388 (10063): 3017–26.

Frampton, Garrett M., Siraj M. Ali, Mark Rosenzweig, Juliann Chmielecki, Xinyuan Lu, Todd M. Bauer, Mikhail Akimov, et al. 2015. "Activation of MET via Diverse Exon 14 Splicing Alterations Occurs in Multiple Tumor Types and Confers Clinical Sensitivity to MET Inhibitors." *Cancer Discovery* 5 (8): 850–59.

Gómez-Seguí, I., H. Makishima, A. Jerez, K. Yoshida, B. Przychodzen, S. Miyano, Y. Shiraishi, et al. 2013. “Novel Recurrent Mutations in the RAS-like GTP-Binding Gene RIT1 in Myeloid Malignancies.” *Leukemia* 27 (9): 1943–46.

Gout, Stephanie, Elisabeth Brambilla, Asma Boudria, Romain Drissi, Sylvie Lantuejoul, Sylvie Gazzeri, and Beatrice Eymin. 2012. “Abnormal Expression of the Pre-MRNA Splicing Regulators SRSF1, SRSF2, SRPK1 and SRPK2 in Non Small Cell Lung Carcinoma.” *PloS One* 7 (10): e46539.

Han, Hong, Ulrich Braunschweig, Thomas Gonatopoulos-Pournatzis, Robert J. Weatheritt, Calley L. Hirsch, Kevin C. H. Ha, Ernest Radovani, et al. 2017. “Multilayered Control of Alternative Splicing Regulatory Networks by Transcription Factors.” *Molecular Cell* 65 (3): 539-553.e7.

Havens, Mallory A., and Michelle L. Hastings. 2016. “Splice-Switching Antisense Oligonucleotides as Therapeutic Drugs.” *Nucleic Acids Research* 44 (14): 6549–63.

Hobbs, G. Aaron, Channing J. Der, and Kent L. Rossman. 2016. “RAS Isoforms and Mutations in Cancer at a Glance.” *Journal of Cell Science* 129 (7): 1287–92.

Hornbeck, P. V., J. M. Kornhauser, S. Tkachev, B. Zhang, E. Skrzypek, B. Murray, V. Latham, and M. Sullivan. 2011. “PhosphoSitePlus: A Comprehensive Resource for Investigating the Structure and Function of Experimentally Determined Post-Translational Modifications in Man and Mouse.” *Nucleic Acids Research* 40 (D1): D261–70.

Imielinski, Marcin, Alice H. Berger, Peter S. Hammerman, Bryan Hernandez, Trevor J. Pugh, Eran Hodis, Jeonghee Cho, et al. 2012. “Mapping the Hallmarks of Lung Adenocarcinoma with Massively Parallel Sequencing.” *Cell*, no. 150: 1107–20.

Jayasinghe, Reyka G., Song Cao, Qingsong Gao, Michael C. Wendl, Nam Sy Vo, Sheila M. Reynolds, Yanyan Zhao, et al. 2018. “Systematic Analysis of Splice-Site-Creating Mutations in Cancer.” *Cell Reports* 23 (1): 270-281.e3.

Jeanson, Arnaud, Pascale Tomasini, Maxime Souquet-Bressand, Nicolas Brandone, Mohamed Boucekine, Mathieu Grangeon, Solène Chaleat, et al. 2019. “Efficacy of Immune Checkpoint Inhibitors in KRAS-Mutant Non-Small Cell Lung Cancer (NSCLC).” *Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer* 14 (6): 1095–1101.

Kahles, André, Kjong-Van Lehmann, Nora C. Toussaint, Matthias Hüser, Stefan G. Stark, Timo Sachsenberg, Oliver Stegle, et al. 2018. “Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients.” *Cancer Cell* 0 (0). https://doi.org/10.1016/j.ccell.2018.07.001.

Karlsson, Kasper, Moritz Przybilla, Hang Xu, Eran Kotler, Kremena Karagyozova, Alexandra Sockell, Katherine Liu, et al. 2022. “Experimental Evolution in *TP53* Deficient Human Gastric Organoids Recapitulates Tumorigenesis.” *BioRxiv*. https://doi.org/10.1101/2022.04.09.487529.

Keenan, Alexandra B., Denis Torre, Alexander Lachmann, Ariel K. Leong, Megan L. Wojciechowicz, Vivian Utti, Kathleen M. Jagodnik, Eryk Kropiwnicki, Zichen Wang, and Avi Ma’ayan. 2019. “ChEA3: Transcription Factor Enrichment Analysis by Orthogonal Omics Integration.” *Nucleic Acids Research* 47 (W1): W212–24.

Kim, Dongsung, Jenny Yaohua Xue, and Piro Lito. 2020. “Targeting KRAS(G12C): From Inhibitory Mechanism to Modulation of Antitumor Effects in Patients.” *Cell* 183 (4): 850–59.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, et al. 2001. “Initial Sequencing and Analysis of the Human Genome.” *Nature* 409 (6822): 860–921.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, et al. 2007. “Clustal W and Clustal X Version 2.0.” *Bioinformatics* 23 (21): 2947–48.

Lee, Stanley Chun-Wei, Heidi Dvinge, Eunhee Kim, Hana Cho, Jean-Baptiste Micol, Young Rock Chung, Benjamin H. Durham, et al. 2016. “Modulation of Splicing Catalysis for Therapeutic Targeting of Leukemia with Mutations in Genes Encoding Spliceosomal Proteins.” *Nature Medicine* 22 (6): 672–78.

Levin, Arthur A. 2019. “Treating Disease at the RNA Level with Oligonucleotides.” *The New England Journal of Medicine* 380 (1): 57–70.

Li, Yuan, Nan Sun, Zhiliang Lu, Shouguo Sun, Jianbing Huang, Zhaoli Chen, and Jie He. 2017. “Prognostic Alternative MRNA Splicing Signature in Non-Small Cell Lung Cancer.” *Cancer Letters* 393: 40–51.

Liao, Yang, Gordon K. Smyth, and Wei Shi. 2014. “FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features.” *Bioinformatics* 30 (7): 923–30.

Liberzon, Arthur, Chet Birger, Helga Thorvaldsdóttir, Mahmoud Ghandi, Jill P. Mesirov, and Pablo Tamayo. 2015. “The Molecular Signatures Database (MSigDB) Hallmark Gene Set Collection.” *Cell Systems* 1 (6): 417–25.

Linding, Rune, Lars Juhl Jensen, Gerard J. Ostheimer, Marcel A. T. M. van Vugt, Claus Jørgensen, Ioana M. Miron, Francesca Diella, et al. 2007. “Systematic Discovery of in Vivo Phosphorylation Networks.” *Cell* 129 (7): 1415–26.

Liu, Yansheng, Andreas Beyer, and Ruedi Aebersold. 2016. “On the Dependency of Cellular Protein Levels on MRNA Abundance.” *Cell* 165 (3): 535–50.

Lo, April, Kristin Holmes, Shriya Kamlapurkar, Filip Mundt, Sitapriya Moorthi, Iris Fung, Shaunt Fereshetian, et al. 2021. “Multiomic Characterization of Oncogenic Signaling Mediated by Wild-Type and Mutant RIT1.” *Science Signaling* 14 (711): eabc4520.

Lu, Xinyuan, Nir Peled, John Greer, Wei Wu, Peter Choi, Alice H. Berger, Sergio Wong, et al. 2017. “MET Exon 14 Mutation Encodes an Actionable Therapeutic Target in Lung Adenocarcinoma.” *Cancer Research* 77 (16): 4498–4505.

Lundberg, Ante S., Scott H. Randell, Sheila A. Stewart, Brian Elenbaas, Kimberly A. Hartwell, Mary W. Brooks, Mark D. Fleming, et al. 2002. “Immortalization and Transformation of Primary Human Airway Epithelial Cells by Gene Transfer.” *Oncogene* 21 (29): 4577–86.

Lundberg, Emma, Linn Fagerberg, Daniel Klevebring, Ivan Matic, Tamar Geiger, Juergen Cox, Cajsa Algenäs, Joakim Lundeberg, Matthias Mann, and Mathias Uhlen. 2010. “Defining the Transcriptome and Proteome in Three Functionally Different Human Cell Lines.” *Molecular Systems Biology* 6 (December): 450.

Mechta, F., D. Lallemand, C. M. Pfarr, and M. Yaniv. 1997. “Transformation by Ras Modifies AP1 Composition and Activity.” *Oncogene* 14 (7): 837–47.

Mermoud, J. E., P. Cohen, and A. I. Lamond. 1992. “Ser/Thr-Specific Protein Phosphatases Are Required for Both Catalytic Steps of Pre-MRNA Splicing.” *Nucleic Acids Research* 20 (20): 5263–69.

Mertins, Philipp, Lauren C. Tang, Karsten Krug, David J. Clark, Marina A. Gritsenko, Lijun Chen, Karl R. Clauser, et al. 2018. “Reproducible Workflow for Multiplexed Deep-Scale

Proteome and Phosphoproteome Analysis of Tumor Tissues by Liquid Chromatography-Mass Spectrometry." *Nature Protocols* 13 (7): 1632–61.

Mok, Tony S., Yi-Long Wu, Myung-Ju Ahn, Marina C. Garassino, Hye R. Kim, Suresh S. Ramalingam, Frances A. Shepherd, et al. 2017. "Osimertinib or Platinum-Pemetrexed in EGFR T790M-Positive Lung Cancer." *The New England Journal of Medicine* 376 (7): 629–40.

Moore, Amanda R., and Shiva Malek. 2021. "The Promise and Peril of KRAS G12C Inhibitors." *Cancer Cell*. Elsevier BV.

Mukhopadhyay, Suman, Matthew G. Vander Heiden, and Frank McCormick. 2021. "The Metabolic Landscape of RAS-Driven Cancers from Biology to Therapy." *Nature Cancer* 2 (3): 271–83.

Muñoz, Úrsula, Juan E. Puche, Rebekka Hannivoort, Ursula E. Lang, Michal Cohen-Naftaly, and Scott L. Friedman. 2012. "Hepatocyte Growth Factor Enhances Alternative Splicing of the Kruppel-like Factor 6 (KLF6) Tumor Suppressor to Promote Growth through SRSF1." *Molecular Cancer Research: MCR* 10 (9): 1216–27.

Nakano, H., F. Yamamoto, C. Neville, D. Evans, T. Mizuno, and M. Perucho. 1984. "Isolation of Transforming Sequences of Two Human Lung Carcinomas: Structural and Functional Analysis of the Activated c-K-Ras Oncogenes." *Proceedings of the National Academy of Sciences of the United States of America* 81 (1): 71–75.

Naro, Chiara, and Claudio Sette. 2013. "Phosphorylation-Mediated Regulation of Alternative Splicing in Cancer." *International Journal of Cell Biology* 2013 (August): 151839.

Nattestad, Maria, Sara Goodwin, Karen Ng, Timour Baslan, Fritz J. Sedlazeck, Philipp Rescheneder, Tyler Garvin, et al. 2018. "Complex Rearrangements and Oncogene Amplifications Revealed by Long-Read DNA and RNA Sequencing of a Breast Cancer Cell Line." *Genome Research* 28 (8): 1126–35.

Ngo, Jacky Chi Ki, Sutapa Chakrabarti, Jian-Hua Ding, Adolfo Velazquez-Dones, Brad Nolen, Brandon E. Aubol, Joseph A. Adams, Xiang-Dong Fu, and Gourisankar Ghosh. 2005. "Interplay between SRPK and Clk/Sty Kinases in Phosphorylation of the Splicing Factor ASF/SF2 Is Regulated by a Docking Motif in ASF/SF2." *Molecular Cell* 20 (1): 77–89.

North, Khrystyna, Salima Benbarche, Bo Liu, Joseph Pangallo, Sisi Chen, Maximilian Stahl, Jan Philipp Bewersdorf, et al. 2022. "Synthetic Introns Enable Splicing Factor Mutation-Dependent Targeting of Cancer Cells." *Nature Biotechnology*, March. https://doi.org/10.1038/s41587-022-01224-2.

Nowak, Dawid G., Jeanette Woolard, Elianna Mohamed Amin, Olga Konopatskaya, Moin A. Saleem, Amanda J. Churchill, Michael R. Ladomery, Steven J. Harper, and David O. Bates. 2008. "Expression of Pro- and Anti-Angiogenic Isoforms of VEGF Is Differentially Regulated by Splicing and Growth Factors." *Journal of Cell Science* 121 (Pt 20): 3487–95.

Oltean, Sebastian, Melissa Gammons, Richard Hulse, Maryam Hamdollah-Zadeh, Athina Mavrou, Lucy Donaldson, Andrew H. Salmon, Steve J. Harper, Michael R. Ladomery, and David O. Bates. 2012. "SRPK1 Inhibition in Vivo: Modulation of VEGF Splicing and Potential Treatment for Multiple Diseases." *Biochemical Society Transactions* 40 (4): 831–35.

Onozato, Ryoichi, Takayuki Kosaka, Hiroyuki Kuwano, Yoshitaka Sekido, Yasushi Yatabe, and Tetsuya Mitsudomi. 2009. "Activation of MET by Gene Amplification or by Splice Mutations Deleting the Juxtamembrane Domain in Primary Resected Lung Cancers."

Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer 4 (1): 5–11.

Paez, J. Guillermo, Pasi A. Jänne, Jeffrey C. Lee, Sean Tracy, Heidi Greulich, Stacey Gabriel, Paula Herman, et al. 2004. "EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy." *Science (New York, N.Y.)* 304 (5676): 1497–1500.

Pao, William, Vincent Miller, Maureen Zakowski, Jennifer Doherty, Katerina Politi, Inderpal Sarkaria, Bhuvanesh Singh, et al. 2004. "EGF Receptor Gene Mutations Are Common in Lung Cancers from 'Never Smokers' and Are Associated with Sensitivity of Tumors to Gefitinib and Erlotinib." *Proceedings of the National Academy of Sciences of the United States of America* 101 (36): 13306–11.

Parikh, Hemang, Emma Carlsson, William A. Chutkow, Lovisa E. Johansson, Heidi Storgaard, Pernille Poulsen, Richa Saxena, et al. 2007. "TXNIP Regulates Peripheral Glucose Metabolism in Humans." *PLoS Medicine* 4 (5): e158.

Parrish, Phoebe C. R., James D. Thomas, Austin M. Gabel, Shriya Kamlapurkar, Robert K. Bradley, and Alice H. Berger. 2021. "Discovery of Synthetic Lethal and Tumor Suppressor Paralog Pairs in the Human Genome." *Cell Reports* 36 (9): 109597.

Passaro, Antonio, Pasi A. Jänne, Tony Mok, and Solange Peters. 2021. "Overcoming Therapy Resistance in EGFR -Mutant Lung Cancer." *Nature Cancer* 2 (4): 377–91.

"Picard Tools - By Broad Institute." n.d. Accessed December 16, 2019. http://broadinstitute.github.io/picard.

Qian, Jun, Qin Chen, Dong-Ming Yao, Lei Yang, Jing Yang, Xiang-Mei Wen, Ying-Ying Zhang, et al. 2015. "MOK Overexpression Is Associated with Promoter Hypomethylation in Patients with Acute Myeloid Leukemia." *International Journal of Clinical and Experimental Pathology* 8 (1): 127–36.

Ray, Bimal K., Ryan Murphy, Papiya Ray, and Alpana Ray. 2002. "SAF-2, a Splice Variant of SAF-1, Acts as a Negative Regulator of Transcription." *The Journal of Biological Chemistry* 277 (48): 46822–30.

Robinson, Mark D., Davis J. McCarthy, and Gordon K. Smyth. 2009. "EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data." *Bioinformatics* 26 (1): 139–40.

Robinson, Mark D., and Alicia Oshlack. 2010. "A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data." *Genome Biology* 11 (3): R25.

Rodenhuis, S., M. L. van de Wetering, W. J. Mooi, S. G. Evers, N. van Zandwijk, and J. L. Bos. 1987. "Mutational Activation of the K-Ras Oncogene. A Possible Pathogenetic Factor in Adenocarcinoma of the Lung." *The New England Journal of Medicine* 317 (15): 929–35.

Rotow, Julia, and Trever G. Bivona. 2017. "Understanding and Targeting Resistance Mechanisms in NSCLC." *Nature Reviews. Cancer* 17 (11): 637–58.

Rusyn, Elena V., Evangeline R. Reynolds, Haipeng Shao, Theresa M. Grana, Tung O. Chan, Douglas A. Andres, and Adrienne D. Cox. 2000. "Rit, a Non-Lipid-ModiÆed Ras-Related Protein, Transforms NIH3T3 Cells without Activating the ERK, JNK, P38 MAPK or PI3K/Akt Pathways." *Oncogene* 19: 4685–94.

Sakabe, Kaoru, Hidemi Teramoto, Muriel Zohar, Babak Behbahani, Hiroshi Miyazaki, Hiroki Chikumi, and J. Silvio Gutkind. 2002. "Potent Transforming Activity of the Small GTP-Binding Protein Rit in NIH 3T3 Cells: Evidence for a Role of a P38gamma-Dependent Signaling Pathway." *FEBS Letters* 511 (1–3): 15–20.

Scotti, Marina M., and Maurice S. Swanson. 2016. “RNA Mis-Splicing in Disease.” *Nature Reviews. Genetics* 17 (1): 19–32.

Seiler, Michael, Shouyong Peng, Anant A. Agrawal, James Palacino, Teng Teng, Ping Zhu, Peter G. Smith, Cancer Genome Atlas Research Network, Silvia Buonamici, and Lihua Yu. 2018. “Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types.” *Cell Reports* 23 (1): 282-296.e4.

Shao, H., and D. A. Andres. 2000. “A Novel RalGEF-like Protein, RGL3, as a Candidate Effector for Rit and Ras.” *The Journal of Biological Chemistry* 275 (35): 26914–24.

Shen, Shihao, Juw Won Park, Zhi-Xiang Lu, Lan Lin, Michael D. Henry, Ying Nian Wu, Qing Zhou, and Yi Xing. 2014. “RMATS: Robust and Flexible Detection of Differential Alternative Splicing from Replicate RNA-Seq Data.” *Proceedings of the National Academy of Sciences* 111 (51): E5593–5601.

Shepard, Peter J., and Klemens J. Hertel. 2009. “The SR Protein Family.” *Genome Biology* 10 (10): 242.

Shi, G. X., and D. A. Andres. 2005. “Rit Contributes to Nerve Growth Factor-Induced Neuronal Differentiation via Activation of B-Raf-Extracellular Signal-Regulated Kinase and P38 Mitogen-Activated Protein Kinase Cascades.” *Molecular and Cellular Biology* 25 (2): 830–46.

Silva, Maria Roméria da, Gabriela Alves Moreira, Ronni Anderson Gonçalves da Silva, Éverton de Almeida Alves Barbosa, Raoni Pais Siqueira, Róbson Ricardo Teixera, Márcia Rogéria Almeida, Abelardo Silva Júnior, Juliana Lopes Rangel Fietto, and Gustavo Costa Bressan. 2015. “Splicing Regulators and Their Roles in Cancer Biology and Therapy.” *BioMed Research International* 2015 (July): 150514.

Singh, Ravi K., and Thomas A. Cooper. 2012. “Pre-MRNA Splicing in Disease and Therapeutics.” *Trends in Molecular Medicine* 18 (8): 472–82.

Skoulidis, Ferdinandos, Bob T. Li, Grace K. Dy, Timothy J. Price, Gerald S. Falchook, Jürgen Wolf, Antoine Italiano, et al. 2021. “Sotorasib for Lung Cancers with KRAS p.G12C Mutation.” *The New England Journal of Medicine* 384 (25): 2371–81.

Smart, Alicia C., Claire A. Margolis, Harold Pimentel, Meng Xiao He, Diana Miao, Dennis Adeegbe, Tim Fugmann, Kwok-Kin Wong, and Eliezer M. Van Allen. 2018. “Intron Retention Is a Source of Neoepitopes in Cancer.” *Nature Biotechnology*, August. https://doi.org/10.1038/nbt.4239.

Smith, Matthew J., Benjamin G. Neel, and Mitsuhiko Ikura. 2013. “NMR-Based Functional Profiling of RASopathies and Oncogenic RAS Mutations.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (12): 4574–79.

Song, Zhen, Tengfei Liu, Jing Chen, Chao Ge, Fangyu Zhao, Miaoxin Zhu, Taoyang Chen, et al. 2019. “HIF-1α-Induced RIT1 Promotes Liver Cancer Growth and Metastasis and Its Deficiency Increases Sensitivity to Sorafenib.” *Cancer Letters* 460 (September): 96–107.

Soria, Jean-Charles, Yuichiro Ohe, Johan Vansteenkiste, Thanyanan Reungwetwattana, Busyamas Chewaskulyong, Ki Hyeong Lee, Arunee Dechaphunkul, et al. 2018. “Osimertinib in Untreated EGFR-Mutated Advanced Non-Small-Cell Lung Cancer.” *The New England Journal of Medicine* 378 (2): 113–25.

Subramanian, Aravind, Pablo Tamayo, Vamsi K. Mootha, Sayan Mukherjee, Benjamin L. Ebert, Michael A. Gillette, Amanda Paulovich, et al. 2005. “Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles.”

Proceedings of the National Academy of Sciences of the United States of America 102 (43): 15545–50.

TCGA. 2014. "Comprehensive Molecular Profiling of Lung Adenocarcinoma." *Nature* 511 (7511): 543–50.

Thomas, James D., Jacob T. Polaski, Qing Feng, Emma J. De Neef, Emma R. Hoppe, Maria V. McSharry, Joseph Pangallo, et al. 2020. "RNA Isoform Screens Uncover the Essentiality and Tumor-Suppressor Activity of Ultraconserved Poison Exons." *Nature Genetics*, January, 1–11.

Trahey, M., and F. McCormick. 1987. "A Cytoplasmic Protein Stimulates Normal N-Ras P21 GTPase, but Does Not Affect Oncogenic Mutants." *Science* 238 (4826): 542–45.

Trombetta, John J., David Gennert, Diana Lu, Rahul Satija, Alex K. Shalek, and Aviv Regev. 2014. "Preparation of Single-Cell RNA-Seq Libraries for next Generation Sequencing." *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]* 107 (1): 4.22.1-17.

Tzoneva, Gannie, Arianne Perez-Garcia, Zachary Carpenter, Hossein Khiabanian, Valeria Tosello, Maddalena Allegretta, Elisabeth Paietta, et al. 2013. "Activating Mutations in the NT5C2 Nucleotidase Gene Drive Chemotherapy Resistance in Relapsed ALL." *Nature Medicine* 19 (3): 368–71.

Wang, Liguo, Shengqin Wang, and Wei Li. 2012. "RSeQC: Quality Control of RNA-Seq Experiments." *Bioinformatics* 28 (16): 2184–85.

Wiredja, Danica D., Mehmet Koyutürk, and Mark R. Chance. 2017. "The KSEA App: A Web-Based Tool for Kinase Activity Inference from Quantitative Phosphoproteomics." *Bioinformatics* , June. https://doi.org/10.1093/bioinformatics/btx415.

Xu, Fengjuan, Su'an Sun, Shilan Yan, Hongling Guo, Miao Dai, and Yincheng Teng. 2015. "Elevated Expression of RIT1 Correlates with Poor Prognosis in Endometrial Cancer." *International Journal of Clinical and Experimental Pathology* 8 (9): 10315–24.

Yang, Qing, Chuandong Lang, Zhengquan Wu, Yuhu Dai, Shaofu He, Wei Guo, Shuai Huang, Hong Du, Dong Ren, and Xinsheng Peng. 2019. "MAZ Promotes Prostate Cancer Bone Metastasis through Transcriptionally Activating the KRas-Dependent RalGEFs Pathway." *Journal of Experimental & Clinical Cancer Research: CR* 38 (1): 391.

Ye, Xin, and Robert A. Weinberg. 2015. "Epithelial-Mesenchymal Plasticity: A Central Regulator of Cancer Progression." *Trends in Cell Biology* 25 (11): 675–86.

Ye, Zhizhou, and Donald E. Ayer. 2018. "Ras Suppresses TXNIP Expression by Restricting Ribosome Translocation." *Molecular and Cellular Biology* 38 (20). https://doi.org/10.1128/MCB.00178-18.

Yea, Steven, Goutham Narla, Xiao Zhao, Rakhi Garg, Sigal Tal-Kremer, Eldad Hod, Augusto Villanueva, et al. 2008. "Ras Promotes Growth by Alternative Splicing-Mediated Inactivation of the KLF6 Tumor Suppressor in Hepatocellular Carcinoma." *Gastroenterology* 134 (5): 1521–31.

Yeh, Chien-Hung, Marcia Bellon, and Christophe Nicot. 2018. "FBXW7: A Critical Tumor Suppressor of Human Cancers." *Molecular Cancer* 17 (1): 115.

Yoshihara, Eiji, So Masaki, Yoshiyuki Matsuo, Zhe Chen, Hai Tian, and Junji Yodoi. 2014. "Thioredoxin/Txnip: Redoxisome, as a Redox Switch for the Pathogenesis of Diseases." *Frontiers in Immunology* 4 (January): 514.

Young, Matthew D., Matthew J. Wakefield, Gordon K. Smyth, and Alicia Oshlack. 2010. “Gene Ontology Analysis for RNA-Seq: Accounting for Selection Bias.” *Genome Biology* 11 (2): R14.

Young, Matthew R., and Nancy H. Colburn. 2006. “Fra-1 a Target for Cancer Prevention or Intervention.” *Gene* 379 (September): 1–11.

Zhao, Jiawei, Yue Sun, Yin Huang, Fan Song, Zengshu Huang, Yufang Bao, Ji Zuo, et al. 2017. “Functional Analysis Reveals That RBM10 Mutations Contribute to Lung Adenocarcinoma Pathogenesis by Deregulating Splicing.” *Scientific Reports* 7 (December 2016): 1–11.

Zhao, Min, Yining Liu, Chong Zheng, and Hong Qu. 2019. “DbEMT 2.0: An Updated Database for Epithelial-Mesenchymal Transition Genes with Experimentally Verified Information and Precalculated Regulation Information for Cancer Metastasis.” *Journal of Genetics and Genomics = Yi Chuan Xue Bao* 46 (12): 595–97.

Zhao, Tongtong, Zachary D. Chiang, Julia W. Morriss, Lindsay M. LaFave, Evan M. Murray, Isabella Del Priore, Kevin Meli, et al. 2021. “Spatial Genomics Enables Multi-Modal Study of Clonal Heterogeneity in Tissues.” *Nature*, December, 1–7.

Zheng, Xiang, Qiu Peng, Lujuan Wang, Xuemei Zhang, Lili Huang, Jia Wang, and Zailong Qin. 2020. “Serine/Arginine-Rich Splicing Factors: The Bridge Linking Alternative Splicing and Cancer.” *International Journal of Biological Sciences* 16 (13): 2442–53.

Zhou, Zhihong, and Xiang-Dong Fu. 2013. “Regulation of Splicing by SR Proteins and SR Protein-Specific Kinases.” *Chromosoma* 122 (3): 191–207.

Zhou, Zhihong, Jinsong Qiu, Wen Liu, Yu Zhou, Ryan M. Plocinik, Hairi Li, Qidong Hu, et al. 2012. “The Akt-SRPK-SR Axis Constitutes a Major Pathway in Transducing EGF Signaling to Regulate Alternative Splicing in the Nucleus.” *Molecular Cell* 47 (3): 422–33.

VITA

April Lo, also known as Fong Shun Lo, was born in Hong Kong where she grew up and attended Shatin Junior School and Shatin College. At age 13, she moved to the suburbs of Worcester, Massachusetts, and attended Shrewsbury High School and Bancroft School. Afterwards, she matriculated to The Johns Hopkins University where she graduated in 2016 with a B.S. in Biomedical Engineering and Applied Mathematics and Statistics. As an undergraduate at Johns Hopkins, she worked in the lab of Rachel Karchin studying somatic variants and clonal evolution in cancer. April began her PhD in Genome Sciences at the University of Washington in 2016. Her graduate work with Alice Berger at the Fred Hutchinson Cancer Center focuses on using functional genomics and multi-omic profiling to study post-transcriptional regulation in lung cancer cells. When she is not doing analysis or experiments, she enjoys knitting, writing, and playing music.