

RESEARCH ARTICLE

Open Access



Oncogenic KRAS alters splicing factor phosphorylation and alternative splicing in lung cancer

April Lo^{1,2}, Maria McSharry¹ and Alice H. Berger^{1,2,3*}

Abstract

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 cells. 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 variant 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, the Ras signaling pathway can promote post-transcriptional splicing changes that may contribute to oncogenic processes.

Keywords: KRAS, Splicing, Lung cancer, Gene expression, Proteomics

*Correspondence: ahberger@fredhutch.org

³ Herbold Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, USA

Full list of author information is available at the end of the article



Background

Lung cancers are the leading cause of cancer death worldwide, with lung adenocarcinomas accounting for over 40% of lung cancers [1]. Lung adenocarcinomas exhibit a high somatic mutational burden, with a median of 8.7 exonic somatic mutations per megabase in the tumor genome [2]. 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 to modulate protein abundance. In a phenomenon known as alternative splicing, cells can 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 dysregulation is widespread in many cancers including lung cancer [3]. One mechanism by which splicing can be altered in cancer is *cis*-acting splice site mutations which disrupt the recognition of exon splice sites [4]. For instance, a mutation in a splice site of the tyrosine kinase gene *MET* leads to exon 14 skipping and overexpression of *MET* protein [5, 6]. Notably, lung cancers harboring this mutation confer clinical sensitivity to *MET* inhibitors [7, 8]. RNA splicing can also be altered in lung tumors by *trans*-acting mutations in splicing factor genes including *U2AF1*, *RBM10*, and *SF3B1* [9–11]. 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 [12]. 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 [13]. SR proteins are a family of splicing factors that plays a large role in splice site selection [14]. The function of SR proteins as recruiters of the spliceosome and catalyzers of the splicing reaction is heavily dependent on their phosphorylation state [15]. As such, these proteins are regulated by kinases including SRPK1 and Clk/Sty [16], and changes in their activity are associated with advanced forms of lung adenocarcinomas [17]. In particular, previous studies show that alternative splicing can be disrupted through the AKT-SRPK-SR protein axis or the TGF β -CLK/sty-SR axis [18–20].

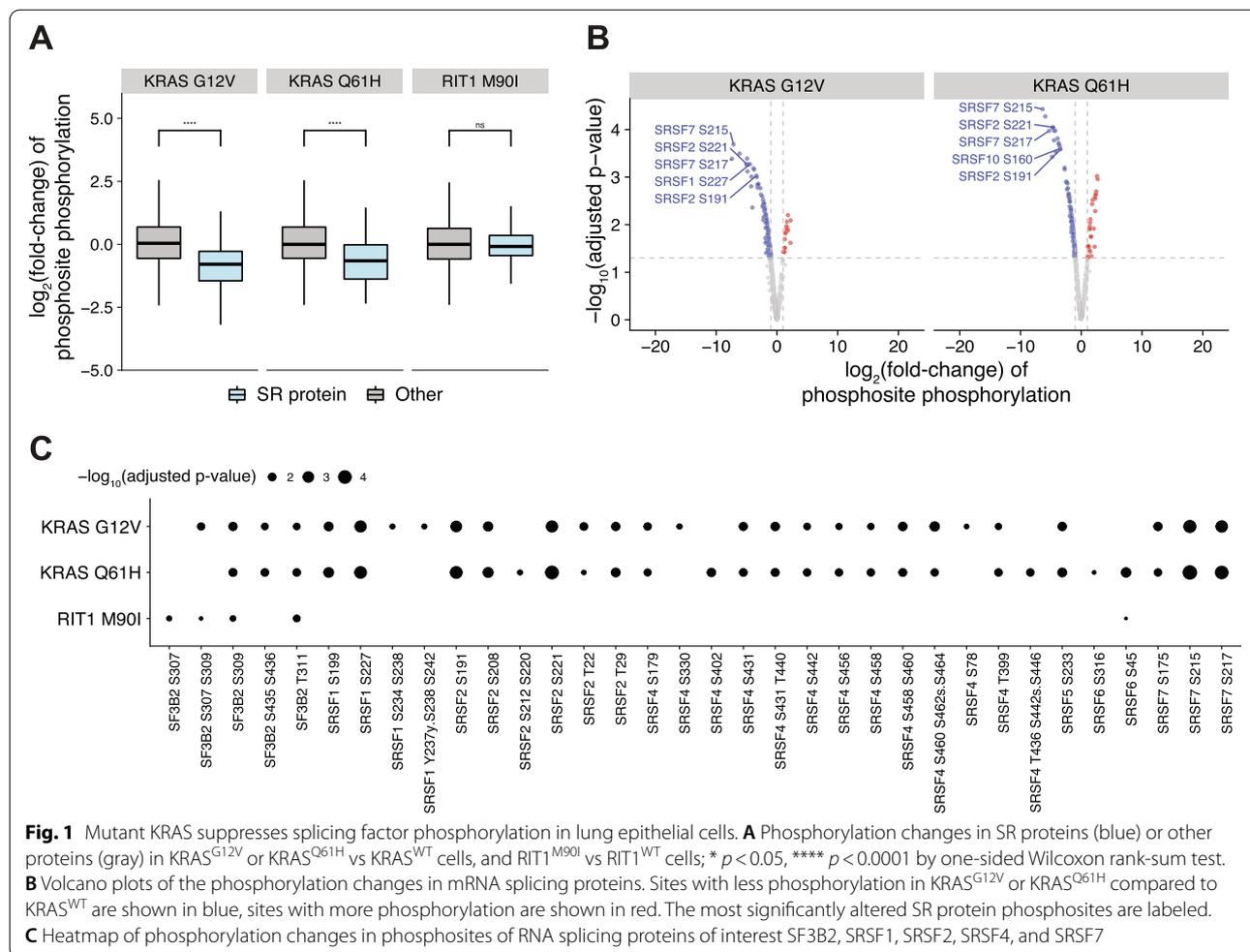
The AKT-SRPK signaling pathway is one of many arms of the receptor tyrosine kinase (RTK) and Ras signaling cascade, a network that regulates many cellular processes and is often altered in cancers [21]. Mutations in *KRAS* are seen in ~27% of lung adenocarcinomas, and in total, ~76% of tumors harbor mutations in RTK-Ras-Raf pathway genes including the tyrosine kinases *EGFR* and *MET* [2, 11, 22]. 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 [23]. Inhibitors targeting *EGFR* or other RTKs are effective only for a limited time, with acquired resistance a nearly universal occurrence [24, 25]. 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 [26]. Moreover, oligonucleotide therapies for aberrant splicing have proven effective in treating genetic disorders and could be used in precision oncology as well [27].

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 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.

Results

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

The process of mRNA splicing is known to involve regulation of phosphorylation states [28, 29]. 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 [30]. As expected, phosphorylation of Ras pathway proteins was significantly altered in cells expressing *KRAS* or *RIT1* variants (Fig. 1 A). Interestingly, in cells expressing *KRAS*^{G12V} or *KRAS*^{Q61H} variants, we also observed a marked change in the phosphorylation of splicing factor proteins (Fig. 1 A) with several SR proteins showing decreased phosphorylation state (Fig. 1 B). As each phosphosite was normalized to total protein level, and total protein level of these SR proteins was unchanged (Additional file 1: Supplemental Fig. 1A), the decreased SR protein phosphorylation was likely due to differences in phosphorylation state itself. The phosphosites with the greatest decrease

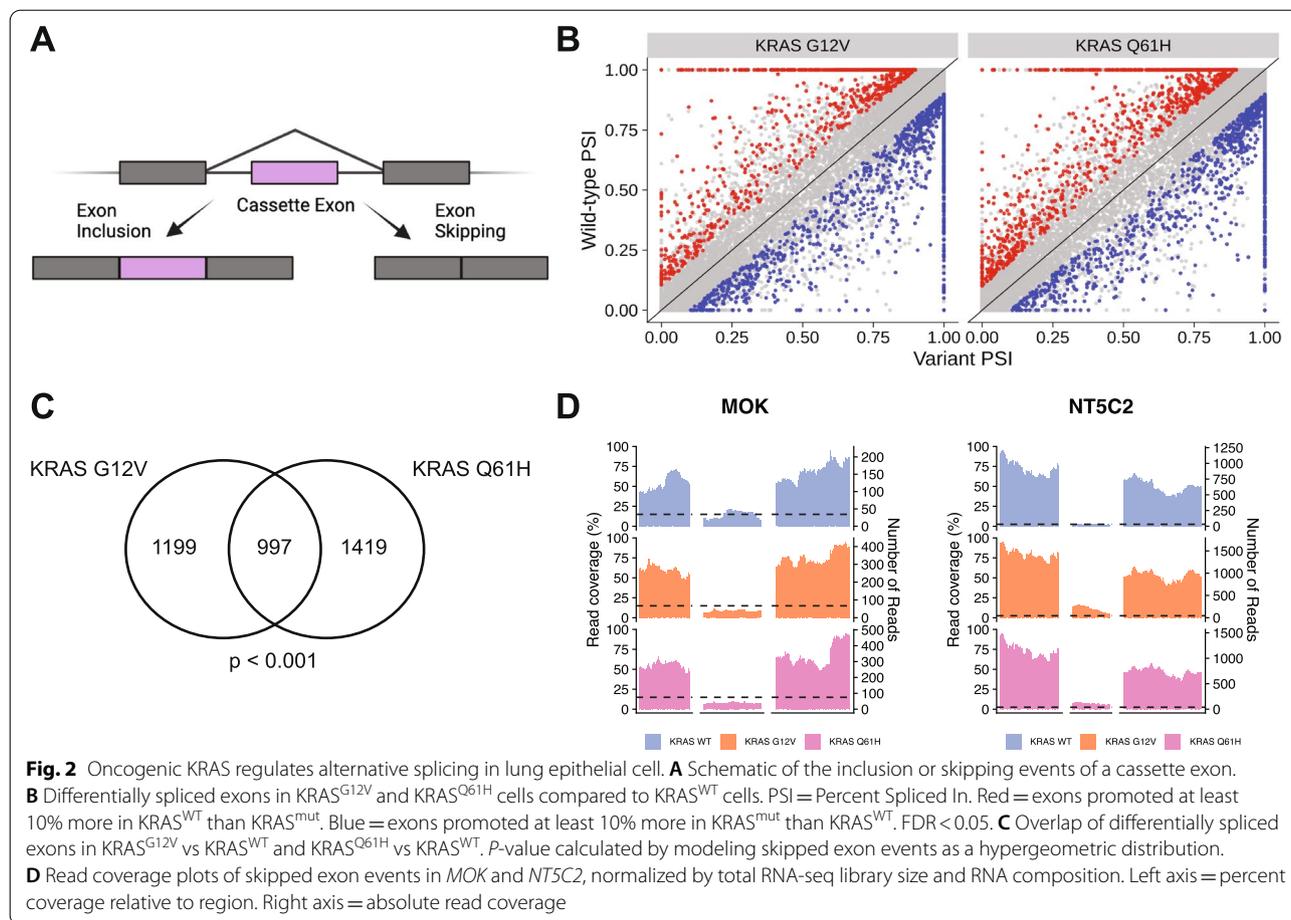


in phosphorylation levels (Fig. 1 B) occurred in proteins known to be precisely regulated by phosphorylation: SRSF7, SRSF1, and SRSF2 [28, 31, 32]. Notably, alteration of SRSF7, SRSF1, and SRSF2 protein phosphorylation occurred predominantly in KRAS^{mut} cells compared to KRAS^{WT} cells, and not in the RIT1^{M90I} cells compared to RIT1^{WT} cells (Fig. 1 C, Additional file 1: Supplemental Fig. 1B). These data suggest that KRAS may uniquely alter SR protein phosphorylation state in a manner distinct from another RAS-family GTPase, RIT1.

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 [33]. Alternative splicing of cassette exons (Fig. 2 A) is tightly

regulated by SR proteins [13, 34] and these events are typically the most prevalent form of alternative splicing detected in cancer and development [35, 36]. The majority of splicing events observed in AALE cells were cassette exon events (70.3%, Additional file 2: Supplemental Fig. 2A), so we focused on cassette exon events for subsequent analyses. We observed that many cassette exons are more skipped or more included in KRAS^{G12V} and KRAS^{Q61H} cells compared to KRAS^{WT} cells (Fig. 2 B). 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 $p < 0.0001$) (Fig. 2 C, Additional file 3: Supplementary Table 1). This similarity between oncogenic KRAS variants holds true when looking at included and excluded exons separately (Additional file 2: Supplemental Fig. 2B). 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 [37, 38] (Fig. 2 D). These shared cassette exon events point to a change in splicing regulation driven by activated KRAS.

Generating a compendium of transcriptome signatures of lung cancer oncogenes

To further study how KRAS and related oncogenes regulate alternative splicing, we performed a large-scale perturbation screen in A549 lung adenocarcinoma cells where the phenotypic readout is RNA sequencing. Isogenic A549 cell lines were generated by lentiviral transduction in 384 well format as described previously [39] with multiple biological replicates per variant. Four days after transduction, cells were lysed for transcriptome analysis using the SmartSeq low input method (Fig. 3 A). 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 [39]. As our previous work quantified gene expression using the probe-based L1000 assay, here we instead performed

high-throughput RNA sequencing to enable discovery of alternative splicing patterns. In total, we generated whole transcriptome profiles for 374 unique replicates of arrayed isogenic cell lines expressing 79 different lentiviral constructs. These included 4 control constructs and 75 vectors for expression of wild-type and/or variant ORFs corresponding to 28 genes (GSE207511, Table 1). For each replicate used in subsequent analyses, we confirmed overexpression of the allele by direct analysis of the RNA-seq data (Additional file 4: Supplemental Fig. 3A-E).

Before using this large-scale dataset for splicing analyses, we first performed biological quality checks to ensure the transcriptomic signatures identified matched known biology. 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 (Fig. 3 B). Expression profiles also

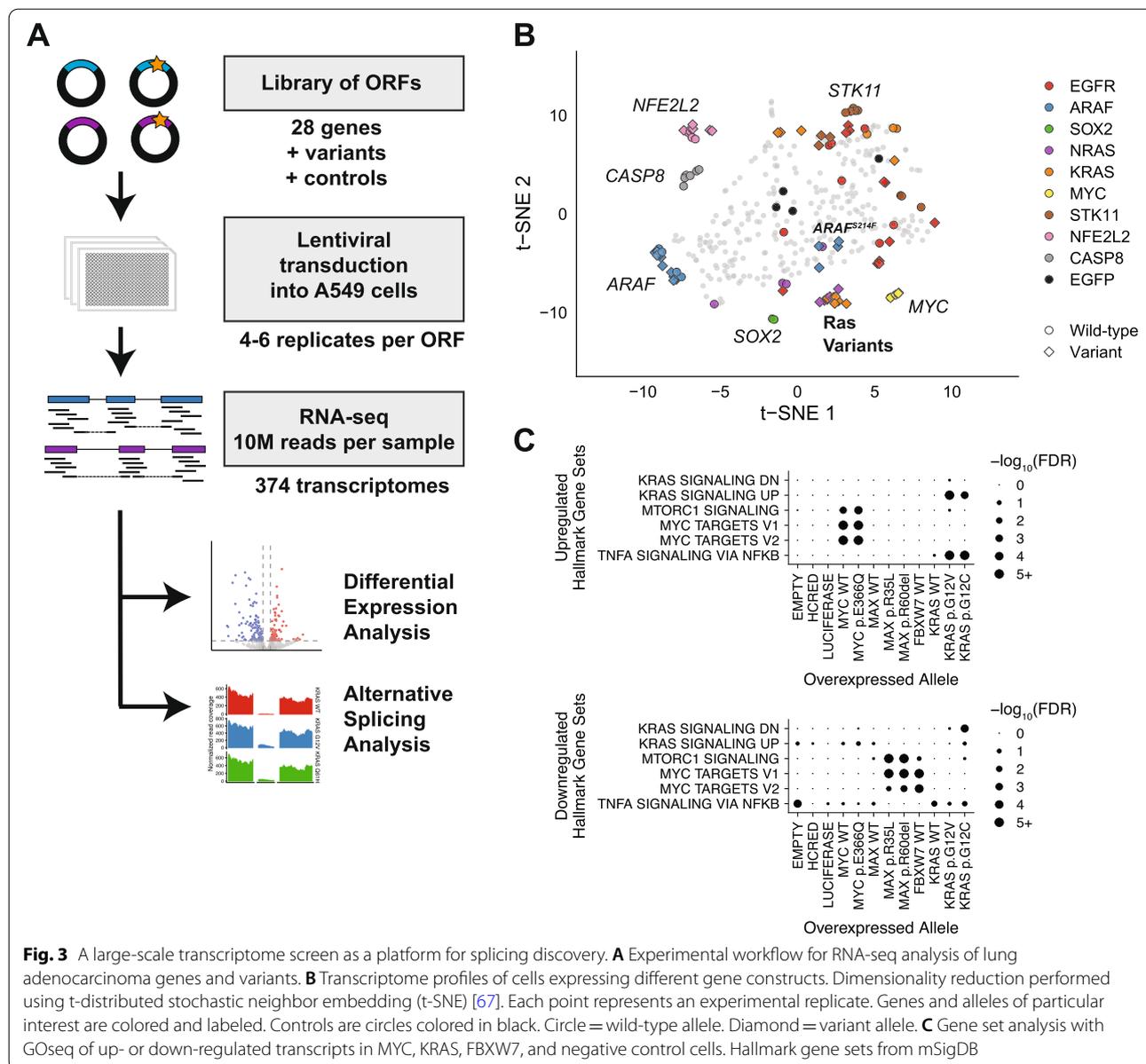


Fig. 3 A large-scale transcriptome screen as a platform for splicing discovery. **A** Experimental workflow for RNA-seq analysis of lung adenocarcinoma genes and variants. **B** Transcriptome profiles of cells expressing different gene constructs. Dimensionality reduction performed using t-distributed stochastic neighbor embedding (t-SNE) [67]. 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

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 [39] (Fig. 3 B).

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

(Additional file 4: Supplemental Fig. 3F). Gene set analysis further showed that the transcripts with increased expression in *MYC*-overexpressing cells include known *MYC* target genes (Fig. 3 C). 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, suggesting these mutants can function as dominant negative mutations [40, 41] (Fig. 3 C). These findings demonstrate that this large-scale gene expression profiling dataset can recapitulate known biology, supporting the utility of these data for further discovery.

Table 1 Alleles expressed in transcriptome screen

Gene	Alleles
ARAF	WT, p.S214C_p.D429A, p.S214F, p.S214F_p.D429A, p.V145L
AURKB	WT
CASP8	WT
CTNNB1	WT, p.S33N, p.S37C
EGFR	WT, p.S645C, p.T790M_p.L858R
ETV1	WT
FBXW7	WT, p.V464E
KEAP1	WT, p.G603W
KRAS	WT, p.G12C, p.G12V
MAX	WT, p.R35L, p.R60del
MET	WT
MYC	WT, p.E366Q
NFE2L2	WT, p.T80K
NRAS	WT, p.Q61L
RALA	WT
RB1	WT, p.A722S, p.A772S, p.E398L, p.Y756C
RBM10	p.W658Y
RBM34	WT
RBM45	WT, p.D434Y, p.M126I
RIT1	WT, p.F82L, p.T76insTDLT
SMAD4	p.D493N, p.G386V, p.P215S, p.P356R, p.R361H, p.V492F
SOX2	WT
STK11	WT, p.D194Y
TP53	p.P278A, p.P278H, p.P278T
TPK1	WT, p.E81Q, p.G48C, p.K111M, p.L185I, p.P152T, p.T205S, p.T213S
U2AF1	WT
ZBTB24	WT, p.G568C, p.G693S, p.K123M, p.M78I, p.Q669R
ZNF597	WT, p.L283V, p.V21L
EGFP	N/A (Control vector)
HCRED	N/A (Control vector)
LUCIFERASE	N/A (Control vector)
EMPTY	N/A (Parental cells with no vector transduction)

One potentially confounding factor for this system is the natural endogenous KRAS^{G12S} variant found in A549 cells. To verify in this background that we were observing the activity of activated KRAS from ectopic expression of KRAS variants, we first ensured that KRAS was overexpressed by a significant amount. We defined significant allele overexpression as two standard deviations above endogenous KRAS expression (Additional file 4: Supplemental Fig. 3A). 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^{G12V} cells was the Nf-κB signaling

pathway (Fig. 3 C), which oncogenic KRAS is known to activate [42, 43]. 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 [39].

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 (Fig. 2) and the large-scale screen in A549 cells (Fig. 3), we performed differential splicing analysis to first compare replicates with ORF expression against replicates expressing a control vector. In the A549 screen, we observed a mean of 456 differentially spliced events per overexpressed wild-type or variant allele and fewer than 1000 events in all comparisons except overexpression of wild-type RBM45, which induced 1496 total differential alternative splicing events compared to vector control (Additional file 5: Supplemental Fig. 4A-B, z-score = 4.51). RBM45 is an RNA binding protein involved in RNA processing and splicing and is associated with neurodegenerative disease and response to DNA damage [44, 45]. 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, and MYC (Additional file 5: Supplemental Fig. 4A). 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 [46].

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, Fig. 4 A). 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 (Additional file 5: Supplemental Fig. 4C). Similarly, KRAS^{G12V} cells differentially splice 409 events compared to KRAS^{WT} cells including 232 cassette exons (Fig. 4 A). Of 7568 differentially spliced events observed in total, 5233 were observed only once in the experiment,

however, 2335 events were shared between at least two variant alleles in the screen (Additional file 5: Supplemental Fig. 4D). Since G12V, G12C, and Q61H are all KRAS activating variants that promote deregulated downstream MAPK signaling [47], we took advantage of our parallel analyses of multiple KRAS variants across both AALE and A549 cellular contexts to identify cassette exon splicing variants that are 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 (Fig. 4 B, Additional file 6: Supplementary Table 2). The approximate degree of splicing change (percent spliced in) induced by KRAS^{G12V} was similar in A549 and AALE cells (Fig. 4 B; Pearson's coefficient = 0.71, p -value = 0.020), and there was no observable correlation between the degree of splicing change and changes in overall gene expression (Additional file 5: Supplemental Fig. 4E).

As our phosphoproteomic analysis indicated changes in the activity of SR proteins, we asked if SR 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* (Additional file 7: Supplemental Fig. 5A). 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 possible contribution to the alternative splicing changes we observe is phosphorylation based.

Of the 10 splice variants identified, we focused on the isoforms of the transcription factor Myc-associated zinc finger protein (*MAZ*, ENSG00000103495.9) due to its previously characterized isoforms and its role in increasing the expression of *KRAS* and *HRAS* [48]. The major mRNA isoform *MAZ-1* (ENST00000322945.11) is composed of five exons whereas one alternative isoform *MAZ-2* (ENST00000219782.10) has six exons, with the inclusion of exon V' (Fig. 4 C). This additional exon includes an alternative early stop codon, leading to an mRNA with an alternative 3' end. This alternative

splicing of *MAZ* therefore leads to two distinct protein isoforms with different C termini, which have been shown to have different DNA-binding properties [49]. In both AALE and A549 cells, expression of KRAS^{mut} resulted in an increase in exon V' inclusion relative to cells expressing KRAS WT (*MAZ-2* isoform; Δ PSI = .14 to .34, FDR < .01) (Fig. 4 D). These alternatively spliced isoforms were validated through qRT-PCR, using primers specific to the exon IV to V junction or the exon IV to V' junction (Fig. 4 C, Methods). This quantitative experiment showed an increased *MAZ-2* to *MAZ-1* isoform ratio in *KRAS*-mutant overexpressing AALE cells (Fig. 4 E, Additional file 7: Supplemental Fig. 5B). 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 uniquely altered downstream of KRAS (Additional file 7: Supplemental Fig. 5C).

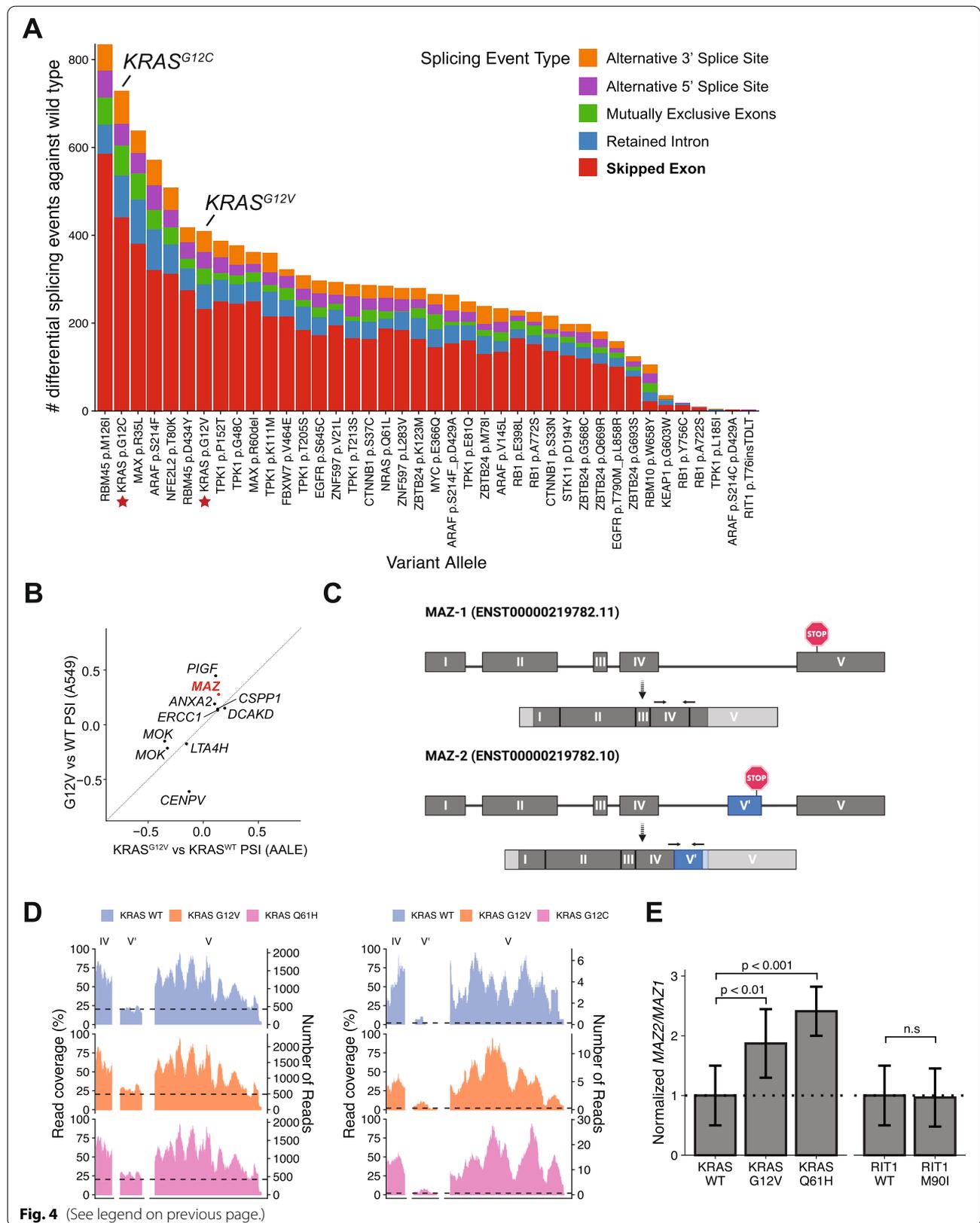
Discussion

Here we asked whether oncogenic signaling by KRAS and other oncogenes in lung cancer can perturb alternative splicing. Mis-splicing is pervasive in human cancer [50] and cannot be fully explained by cis-acting splice site mutations or trans-acting somatic mutations in splicing factors. Our study was motivated by our initial observation that SR protein phosphorylation is significantly suppressed upon expression of oncogenic KRAS in human lung epithelial cells (Fig. 1). Expression of a related GTPase, RIT1, did not perturb splicing factor phosphorylation, indicating this function of mutant KRAS is specific.

SR proteins are critical determinants of splice site selection in both constitutive and alternative RNA splicing. SR proteins recognize and bind to exon splicing enhancers (ESEs) to recruit the spliceosome [13]. Therefore, to modulate isoform abundance, both the abundance and activity of SR proteins and the presence of ESEs in the alternate exons are important. One way that SR protein activity is regulated is by the protein kinases Clk1 and SRPK1 [29]. To determine if the dephosphorylation of SR proteins in KRAS-mutant cells was also associated with changes in splicing, we performed a series of sequencing-based splicing assays.

(See figure on next page.)

Fig. 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. Δ PSI > 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). Δ PSI = Change in Percent Spliced In. **C** Schematic of *MAZ-1* and *MAZ-2* isoforms with labeled active stop codons. Shown are the pre-mRNA structures and the spliced mRNA with untranslated regions grayed out. Arrows above the spliced mRNAs indicate qRT-PCR primer binding sites. **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** Quantitative levels of *MAZ-2* relative to *MAZ-1* isoforms in AALE cells overexpressing KRAS or RIT1 variants, normalized to KRAS^{WT} or RIT1^{WT}, as measured by qRT-PCR. Error bars = standard error



First, we quantified cassette exon splicing using next-generation sequencing in AALE, the same human lung epithelial cell line used for the proteomics assay. Second, we performed a similar sequencing-based assay in A549 cells engineered to express either wild-type or mutant KRAS, or other lung cancer-associated variants. We identified many consistently differentially spliced transcripts in both A549 and AALE. Among the most recurrently differentially spliced transcripts in these multiple lung cell contexts was alternative splicing of *MAZ*, which we validated by quantitative RT-PCR. *MAZ* encodes a transcription factor which has been shown to regulate expression of *KRAS* and *HRAS* [49]. *MAZ* is a member of the *MAZ*-like family of transcription factors which also includes *VEZFI* [51]. Both *MAZ* and *VEZFI* can influence splicing themselves through RNA pol II elongation pausing which can favor alternative exon inclusion [51].

Here we observed alternative splicing of *MAZ* mRNA itself, with mutant KRAS promoting inclusion of the minor *MAZ-2* isoform with an alternative 3' end. Typically, loss of SR proteins is thought to favor splicing of the constitutive exons, but prior work has established that SR protein loss can also lead to exon inclusion of alternative exons [52, 53]. The regulation of exon skipping/inclusion by SR proteins is complex and appears to reflect the interaction and balance among multiple SR proteins at ESEs within the alternative exon and in flanking constitutive exons [13, 14]. However, one limitation of our study is that we did not directly link the activity of SR proteins to the observed splicing changes. Future work should be directed at modifying the activity or localization of SR proteins directly and assessing their effect on alternative splicing.

More broadly, our work contributes to efforts to understand how alternative splicing becomes dysregulated in cancer [54, 55]. By studying RNA, protein, and phosphorylation levels, we interrogated RNA splicing regulation driven by multiple oncogenes and their tumor-associated variants. Further studies of these alternative splicing events could 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 [56, 57]. Similarly, existing small molecule inhibitors may be used to correct splicing factor dysregulation; For example, inhibition of SRPK1 can be used to modulate SR protein activity [58].

Alternative isoforms in cancers also introduce neoepitopes specific to the tumor [3]. 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 [59]. Our study is limited to effects in vitro isolated from the immune environment, and so 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 [60].

Conclusions

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 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.

Methods

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 [30]. Briefly, isogenic AALE cells were generated by transduction with lentiviral vectors encoding wild-type KRAS, KRAS^{G12V}, KRAS^{Q61H}, or wild-type RIT1 or RIT1^{M90I}. 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.

RNA sequencing

RNA-seq libraries of KRAS- and RIT1-mutant AALE cells were previously described [30]. 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 50bp 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 [39]. 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. Ninety-six 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 [61]. RNA was isolated from 9 µl lysate and 1 µl of ERCC 1:1000 Spike in control (Thermo Fisher Scientific) 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: 3 min 98 °C; 20 cycles of 15 s at 98 °C, 15 s 67 °C, 6 min at 72 °C; final extension of 5 min 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 75bp paired end reads per replicate. 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 numbers GSE183670 and GSE207511.

Differential gene expression analysis

Transcripts were quantified with featureCounts from the Subread package v.1.5.3 [62] using RefSeq gene annotations. Differential expression was then determined using edgeR v.3.30.3 [63], normalizing transcript counts by library size and RNA composition scale factors computed by using trimmed mean of M-values (TMM) between sample pairs [64]. 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 across 4 controls and 75 experimental alleles, each profile quantifying 12,543 genes. The analysis pipeline is available as a Snakemake workflow at <https://github.com/bergerbio/RNA-splicing-screen>.

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 genome v.19 [33]. 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>).

Gene set analysis

Analysis of enrichment of KRAS signaling in differential RNA expression profiles was performed in R with Goseq [65]. KRAS signaling gene sets were taken from MSigDB hallmark gene sets [66].

Quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Total RNA was extracted using Direct-zol RNA Mini-prep 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 TGA GGC AGC TTT CGC on exon IV; Reverse: TCA CCA GTA CCT TTG TTG CA spanning exon IV and V;

ENST00000219782.11) and with exon V' (Forward: AGC TCT GCA ACA AAG GCT TC spanning exon IV and V'; Reverse: GGG CAG GGG TCT TGC A on exon V'; ENST00000219782.10). 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.

Abbreviations

RTK: Receptor tyrosine kinase; LC-MS/MS: Liquid chromatography and tandem mass spectrometry; ORF: Open reading frame; ESE: Exon splicing enhancer; PSI: Percent spliced in.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12885-022-10311-1>.

Additional file 1: Supplementary Fig. 1. 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).

Additional file 2: Supplementary Fig. 2. A) Number of differentially spliced events vs WT grouped by alternative splicing event type (A3SS = Alternative 3' splice site; A5SS = Alternative 5' splice site; MXE = Mutually exclusive exon; RI = Retained intron; SE = Skipped exon). Labeled percentages describe proportion of the group out of all differentially spliced events. B) Overlap of differentially spliced exons either more included (top) or more excluded (bottom) in KRAS^{G12V} vs KRAS^{WT} and KRAS^{Q61H} vs KRAS^{WT}, *P*-values calculated by modeling events as a hypergeometric distributions.

Additional file 3: Supplementary Table 1. (TSV 13 kb)

Additional file 4: Supplementary Fig. 3. 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*.

Additional file 5: Supplementary Fig. 4. 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) Number of differential alternative splicing (AS) events unique to one variant allele or overlapping between multiple alleles. E) Comparing change in Percent Spliced In (Δ PSI) of cassette exons and the differential expression of the corresponding gene transcript.

Additional file 6: Supplementary Table 2. (TSV 420 kb).

Additional file 7: Supplementary Fig. 5. A) mRNA expression of splicing factors *SF3B1*, *SRSF1*, and *SRSF7*. B) End-point PCR gel detecting *MAZ-1* and *MAZ-2* isoforms in AALE cells overexpressing vector, *KRAS*, or RIT1 variants. C) Differential splicing of *MAZ* exon V' in variant alleles compared to respective wild-type alleles.

Acknowledgements

We thank Robert 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 Mertins, 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.

Authors' contributions

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.

Funding

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.

Availability of data and materials

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 GSE146479, GSE183670, and GSE207511.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

AHB reports consulting fees from Puma Biotechnology.

Author details

¹Human Biology Division, Fred Hutchinson Cancer Center, Seattle, WA, USA.

²Department of Genome Sciences, University of Washington, Seattle, WA, USA.

³Herbold Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, USA.

Received: 25 May 2022 Accepted: 10 November 2022

Published online: 16 December 2022

References

1. American Cancer Society. Cancer Facts & Figures 2021. 2021. <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2021.html>.
2. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet*. 2016;48:607–16.
3. Kahles A, Lehmann K-V, Toussaint NC, Hüser M, Stark SG, Sachsenberg T, et al. Comprehensive analysis of alternative splicing across tumors from 8,705 patients. *Cancer Cell* 2018;0.
4. Jayasinghe RG, Cao S, Gao Q, Wendl MC, Vo NS, Reynolds SM, et al. Systematic Analysis of Splice-Site-Creating Mutations in Cancer. *Cell Rep*. 2018;23:270–281.e3.
5. Lu X, Peled N, Greer J, Wu W, Choi P, Berger AH, et al. MET exon 14 mutation encodes an actionable therapeutic target in lung adenocarcinoma. *Cancer Res*. 2017;77:4498–505.
6. Onozato R, Kosaka T, Kuwano H, Sekido Y, Yatabe Y, Mitsudomi T. Activation of MET by gene amplification or by splice mutations deleting the

- juxtamembrane domain in primary resected lung cancers. *J Thorac Oncol.* 2009;4:5–11.
7. Frampton GM, Ali SM, Rosenzweig M, Chmielecki J, Lu X, Bauer TM, et al. Activation of MET via diverse exon 14 splicing alterations occurs in multiple tumor types and confers clinical sensitivity to MET inhibitors. *Cancer Discov.* 2015;5:850–9.
 8. Drilon A, Clark JW, Weiss J, Ou S-HI, Camidge DR, Solomon BJ, et al. Antitumor activity of crizotinib in lung cancers harboring a MET exon 14 alteration. *Nat Med.* 2020;26:47–51.
 9. Brooks AN, Choi PS, De Waal L, Sharifnia T, Imielinski M, Saksena G, et al. A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events. *PLoS One.* 2014;9.
 10. Zhao J, Sun Y, Huang Y, Song F, Huang Z, Bao Y, et al. Functional analysis reveals that RBM10 mutations contribute to lung adenocarcinoma pathogenesis by deregulating splicing. *Sci Rep.* December 2016;2017(7):1–11.
 11. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell.* 2012;150:1107–20.
 12. Li Y, Sun N, Lu Z, Sun S, Huang J, Chen Z, et al. Prognostic alternative mRNA splicing signature in non-small cell lung cancer. *Cancer Lett.* 2017;393:40–51.
 13. Zhou Z, Fu X-D. Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma.* 2013;122:191–207.
 14. Bradley T, Cook ME, Blanchette M. SR proteins control a complex network of RNA-processing events. *RNA.* 2015;21:75–92.
 15. Mermoud JE, Cohen P, Lamond AI. Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. *Nucleic Acids Res.* 1992;20:5263–9.
 16. Shepard PJ, Hertel KJ. The SR protein family. *Genome Biol.* 2009;10:242.
 17. Gout S, Brambilla E, Boudria A, Drissi R, Lantuejoul S, Gazzeri S, et al. Abnormal expression of the pre-mRNA splicing regulators SRSF1, SRSF2, SRPK1 and SRPK2 in non small cell lung carcinoma. *PLoS One.* 2012;7:e46539.
 18. Zhou Z, Qiu J, Liu W, Zhou Y, Plocinik RM, Li H, et al. The Akt-SRPK-SR axis constitutes a major pathway in transducing EGF signaling to regulate alternative splicing in the nucleus. *Mol Cell.* 2012;47:422–33.
 19. Nowak DG, Woolard J, Amin EM, Konopatskaya O, Saleem MA, Churchill AJ, et al. Expression of pro- and anti-angiogenic isoforms of VEGF is differentially regulated by splicing and growth factors. *J Cell Sci.* 2008;121(Pt 20):3487–95.
 20. Blaustein M, Pelisch F, Tanos T, Muñoz MJ, Wengier D, Quadrana L, et al. Concerted regulation of nuclear and cytoplasmic activities of SR proteins by AKT. *Nat Struct Mol Biol.* 2005;12:1037–44.
 21. Mukhopadhyay S, Vander Heiden MG, McCormick F. The metabolic landscape of RAS-driven cancers from biology to therapy. *Nature Cancer.* 2021;2:271–83.
 22. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* 2014;511:543–50.
 23. Canon J, Rex K, Saiki AY, Mohr C, Cooke K, Bagal D, et al. The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature.* 2019. <https://doi.org/10.1038/s41586-019-1694-1>.
 24. Passaro A, Jänne PA, Mok T, Peters S. Overcoming therapy resistance in EGFR-mutant lung cancer. *Nature Cancer.* 2021;2:377–91.
 25. Drosten M, Barbacid M. Targeting the MAPK pathway in KRAS-driven tumors. *Cancer Cell.* 2020;37:543–50.
 26. da Silva MR, Moreira GA, Gonçalves da Silva RA, de Almeida Alves Barbosa É, Pais Siqueira R, Teixeira RR, et al. Splicing regulators and their roles in cancer biology and therapy *Biomed Res Int* 2015;2015:150514.
 27. Finkel RS, Chiriboga CA, Vajsar J, Day JW, Montes J, De Vivo DC, et al. Treatment of infantile-onset spinal muscular atrophy with nusinersen: a phase 2, open-label, dose-escalation study. *Lancet.* 2016;388:3017–26.
 28. Cho S, Hoang A, Sinha R, Zhong X-Y, Fu X-D, Krainer AR, et al. Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly. *Proc Natl Acad Sci U S A.* 2011;108:8233–8.
 29. Aubol BE, Wu G, Keshwani MM, Movassat M, Fattat L, Hertel KJ, et al. Release of SR proteins from CLK1 by SRPK1: a symbiotic kinase system for phosphorylation control of pre-mRNA splicing. *Mol Cell.* 2016;63:218–28.
 30. Lo A, Holmes K, Kamapurkar S, Mundt F, Moorthi S, Fung I, et al. Multi-omic characterization of oncogenic signaling mediated by wild-type and mutant RIT1. *Sci Signal.* 2021;14:eabc4520.
 31. Aubol BE, Chakrabarti S, Ngo J, Shaffer J, Nolen B, Fu X-D, et al. Processive phosphorylation of alternative splicing factor/splicing factor 2. *Proc Natl Acad Sci U S A.* 2003;100:12601–6.
 32. Naro C, Sette C. Phosphorylation-mediated regulation of alternative splicing in cancer. *Int J Cell Biol.* 2013;2013:151839.
 33. Shen S, Park JW, Lu Z-X, Lin L, Henry MD, Wu YN, et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci.* 2014;111:E5593–601.
 34. Zheng X, Peng Q, Wang L, Zhang X, Huang L, Wang J, Qin Z. Serine/arginine-rich splicing factors: the bridge linking alternative splicing and cancer. *Int J Biol Sci.* 2020;16:2442–53.
 35. Scotti MM, Swanson MS. RNA mis-splicing in disease. *Nat Rev Genet.* 2016;17:19–32.
 36. Singh RK, Cooper TA. Pre-mRNA splicing in disease and therapeutics. *Trends Mol Med.* 2012;18:472–82.
 37. Qian J, Chen Q, Yao D-M, Yang L, Yang J, Wen X-M, Zhang Y-Y, Chai H-Y, Ma J-C, Deng Z-Q, et al. MOK overexpression is associated with promoter hypomethylation in patients with acute myeloid leukemia. *Int J Clin Exp Pathol.* 2015;8:127–36.
 38. Tzoneva G, Perez-Garcia A, Carpenter Z, Khiabani H, Tosello V, Allegretta M, Paietta E, Racevskis J, Rowe JM, Tallman MS, et al. Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. *Nat Med.* 2013;19:368–71.
 39. Berger AH, Brooks AN, Wu X, Shrestha Y, Chouinard C, Piccioni F, et al. High-throughput phenotyping of lung Cancer somatic mutations. *Cancer Cell.* 2016;30:214–28.
 40. Yeh C-H, Bellon M, Nicot C. FBXW7: a critical tumor suppressor of human cancers. *Mol Cancer.* 2018;17:115.
 41. Augert A, Mathsyaraja H, Ibrahim AH, Freie B, Geuenich MJ, Cheng P-F, et al. MAX Functions as a Tumor Suppressor and Rewires Metabolism in Small Cell Lung Cancer. *Cancer Cell.* 2020;38:97–114.e7.
 42. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature.* 2009;462:108–12.
 43. Daniluk J, Liu Y, Deng D, Chu J, Huang H, Gaiser S, et al. An NF- κ B pathway-mediated positive feedback loop amplifies Ras activity to pathological levels in mice. *J Clin Invest.* 2012;122:1519–28.
 44. Bakkar N, Kousari A, Kovalik T, Li Y, Bowser R. RBM45 modulates the antioxidant response in amyotrophic lateral sclerosis through interactions with KEAP1. *Mol Cell Biol.* 2015;35:2385–99.
 45. Gong J, Huang M, Wang F, Ma X, Liu H, Tu Y, et al. RBM45 competes with HDAC1 for binding to FUS in response to DNA damage. *Nucleic Acids Res.* 2017;45:12862–76.
 46. Han H, Braunschweig U, Gonatopoulos-Pournatzis T, Weatheritt RJ, Hirsch CL, Ha KCH, et al. Multilayered Control of Alternative Splicing Regulatory Networks by Transcription Factors. *Mol Cell.* 2017;65:539–553.e7.
 47. Cook JH, Melloni GEM, Gulhan DC, Park PJ, Haigis KM. The origins and genetic interactions of KRAS mutations are allele- and tissue-specific. *Nat Commun.* 2021;12:1808.
 48. Yang Q, Lang C, Wu Z, Dai Y, He S, Guo W, et al. MAZ promotes prostate cancer bone metastasis through transcriptionally activating the KRas-dependent RalGEFs pathway. *J Exp Clin Cancer Res.* 2019;38:391.
 49. Ray BK, Murphy R, Ray P, Ray A. SAF-2, a splice variant of SAF-1, acts as a negative regulator of transcription. *J Biol Chem.* 2002;277:46822–30.
 50. Climente-González H, Porta-Pardo E, Godzik A, Eyra E. The functional impact of alternative splicing in Cancer. *Cell Rep.* 2017;20:2215–26.
 51. Rambout X, Dequiedt F, Maquat LE. Beyond transcription: roles of transcription factors in pre-mRNA splicing. *Chem Rev.* 2018;118:4339–64.
 52. Han J, Ding J-H, Byeon CW, Kim JH, Hertel KJ, Jeong S, et al. SR proteins induce alternative exon skipping through their activities on the flanking constitutive exons. *Mol Cell Biol.* 2011;31:793–802.
 53. Xu X, Yang D, Ding J-H, Wang W, Chu P-H, Dalton ND, et al. ASF/SF2-regulated CaMKII δ alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. *Cell.* 2005;120:59–72.
 54. Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. *Nat Rev Cancer.* 2016;16:413–30.

55. Lee SC-W, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. *Nat Med.* 2016;22:976–86.
56. Havens MA, Hastings ML. Splice-switching antisense oligonucleotides as therapeutic drugs. *Nucleic Acids Res.* 2016;44:6549–63.
57. Levin AA. Treating disease at the RNA level with oligonucleotides. *N Engl J Med.* 2019;380:57–70.
58. Oltean S, Gammons M, Hulse R, Hamdollah-Zadeh M, Mavrou A, Donaldson L, et al. SRPK1 inhibition in vivo: modulation of VEGF splicing and potential treatment for multiple diseases. *Biochem Soc Trans.* 2012;40:831–5.
59. Jeanson A, Tomasini P, Souquet-Bressand M, Brandone N, Boucekine M, Grangeon M, et al. Efficacy of immune checkpoint inhibitors in KRAS-mutant non-small cell lung Cancer (NSCLC). *J Thorac Oncol.* 2019;14:1095–101.
60. Smart AC, Margolis CA, Pimentel H, He MX, Miao D, Adeegbe D, et al. Intron retention is a source of neoepitopes in cancer. *Nat Biotechnol.* 2018. <https://doi.org/10.1038/nbt.4239>.
61. Trombetta JJ, Gennert D, Lu D, Satija R, Shalek AK, Regev A. Preparation of single-cell RNA-Seq libraries for next generation sequencing. *Curr Protoc Mol Biol.* 2014;107:4.22.1–17.
62. Liao Y, Smyth GK, Shi W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30:923–30.
63. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2009;26:139–40.
64. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010;11:R25.
65. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 2010;11:R14.
66. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* 2015;1:417–25.
67. Van Der Maaten LJP, Hinton GE. Visualizing high-dimensional data using t-sne. *J Mach Learn Res.* 2008;9:2579–605.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

