Title

Splicing Regulation of MYC Dependent Exons in Prostate Cancer

Author

Chen, Xinyuan

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Splicing Regulation of MYC Dependent Exons in Prostate Cancer

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular Biology

by

Xinyuan Chen

2022

ABSTRACT OF THE DISSERTATION

Splicing Regulation of MYC Dependent Exons

in Prostate Cancer

by

Xinyuan Chen

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2022

Professor Douglas L. Black, Chair

The Myc proto-oncogene contributes to the pathogenesis of more than 50 percent of human cancers including prostate cancers. Malignant transformation by Myc is known to transcriptionally upregulate the core pre-mRNA splicing machinery and cause mis-regulation of alternative splicing. However, our understanding of how changes of splice isoforms affect the cancer phenotype is limited and how these changes are directed by Myc remains largely unknown. Here, we performed a signaling pathway-guided differential splicing analysis to identify Myc dependent splicing events, including an HRAS cassette exon that is repressed by Myc transformation across multiple tumor types. To molecularly dissect the regulation of this HRAS cassette exon, we used antisense oligonucleotide tiling to identify intronic splicing enhancers and silencers in the HRAS flanking introns. RNA binding motif prediction indicated the presence of multiple binding sites for hnRNP H and hnRNP F within these cis-regulatory

elements. Using siRNA knockdown and cDNA expression, we found that both hnRNP H and F activate the HRAS cassette exon. Mutagenesis and targeted RNA immunoprecipitation implicate two downstream intronic G-rich elements in this H/F mediated splicing activation. Bioinformatic analyses of ENCODE RNAseq datasets also confirmed hnRNP H regulation of HRAS splicing. Consistent with the observed effects of hnRNP H on HRAS splicing, analyses of RNAseq datasets across multiple cancers showed a negative correlation of hnRNP H gene expression with Myc hallmark enrichment. Loss of hnRNP H/F altered the cell cycle progression and induced apoptosis in prostate cancer cells. In the second part of this dissertation, we again applied antisense oligonucleotide tiling to dissect the cis-regulatory elements of another MYC-dependent splicing event, SRSF3 cassette exon, whose inclusion is repressed by Myc transformation in prostate cancer. We identified multiple exonic splicing enhancers as well as downstream intronic splicing silencers. Motif-based RNA binding protein prediction suggested the presence of Sam68 binding motifs in the intronic splicing silencers. siRNA knockdown of Sam68 confirmed it functions as a SRSF3 splicing repressor. Lastly, using RNA-seq data generated from an in vitro model of prostate cancer with inducible Myc expression, we performed RNA binding protein-focused differential gene expression analysis and found majority of differentially expressed RNA binding proteins are downregulated with the depletion of MYC oncoprotein. CRISPRi screening of top Myc-responsive RNA binding proteins suggested that HRAS and SRSF3 are differentially spliced in the loss of RBM28, SNRPD1, EIF3B, and PINX1. Collectively, our results reveal mechanisms for the Myc-dependent regulation of splicing, and point to new possible therapeutic targets in advanced prostate cancers.

The dissertation of Xinyuan Chen is approved.

Heather R. Christofk

Robert D. Damoiseaux

Tracy L. Johnson

Owen N. Witte

Douglas L. Black, Committee Chair

University of California, Los Angeles

2022

DEDICATION

To my mom and dad, for your constant love and support.

To my husband Qin An, for always being there for me,

without whom I would not have been able to reach this far in my PhD journey.

To my friends, for all those laughs.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to first acknowledge my mentor, Doug Black, for his kindness, patience, continuous guidance and tremendous supports of my graduate work, that have made my research possible and shaped me into a scientist. I would like to thank all the Black Lab members, Andrey Damianov, Julia Nikolic, Chia-ho Lin, Wen Xiao, Nazim Mohammad, Kyu-Hyeon Yeom, Xiaojun Wang, Hanyoung Lim, Nivedita Damodaren, Parham Peyda, Daniel Acre, Harry Taegyun Yang, for their help and supports during the past five years. Furthermore, I am very grateful to my committee, Heather Christofk, Robert Damoiseaux, Tracy Johnson, and Owen Witte, for scientific guidance and suggestions throughout my graduate study.

Many thanks to my collaborators for their help and contributions to my thesis project. To Yi Xing, Harry Taegyun Yang, Beatrice Zhang, for bioinformatic help, and constructive discussion and comments on the manuscript; to Frank Rigo, for his help with ASO tiling methods; to John Phillips, who initiated and greatly impacted this project, and motivated me to continue my cancer biology research.

Chapters 2 is a manuscript that is in preparation:

Xinyuan Chen, Harry Taegyun Yang, Beatrice Zhang, John W. Phillips, Donghui Cheng, Frank Rigo, Owen N. Witte, Yi Xing, Douglas L. Black. *HnRNP H/F regulation of a MYC dependent HRAS exon in Prostate Cancer.* Manuscript in preparation.

This work was supported by Whitcome Pre-doctoral Fellowship and Warsaw Fellowship at University of California, Los Angeles.

VITA

EDUCATION

2016 B.S. in Microbiology, Immunology, and Molecular Genetics
University of California, Los Angeles

PUBLICATIONS

Xinyuan Chen, Harry Taegyun Yang, Beatrice Zhang, John W. Phillips, Donghui Cheng, Frank Rigo, Owen N. Witte, Yi Xing, Douglas L. Black. *HnRNP H/F regulation of a MYC dependent HRAS exon in Prostate Cancer.* Manuscript in preparation.

Yu Zhang, C Jake Harris, Qikun Liu, Wanlu Liu, Israel Ausin, Yanping Long, Lidan Xiao, Li Feng, Xu Chen, Yubin Xie, Xinyuan Chen, Lingyu Zhan, Suhua Feng, Jingyi Jessica Li, Haifeng Wang, Jixian Zhai, Steven E Jacobsen (2018). *Large-scale comparative epigenomics reveals hierarchical regulation of non-CG methylation in Arabidopsis*. Proceedings of the National Academy of Sciences 115(5): E1069-E1074.

Hsuan Yu Kuo, Elise L Jacobsen, Yanping Long, Xinyuan Chen, Jixian Zhai (2017). *Characteristics and processing of Pol IV-dependent transcripts in Arabidopsis.* Journal of Genetics and Genomics 44(1): 3-6.

Jixian Zhai, Sylvain Bischof, Haifeng Wang, Suhua Feng, Tzuu-fen Lee, Chong Teng, Xinyuan Chen, Soo Young Park, Linshan Liu, Javier Gallego-Bartolome, Wanlu Liu, Ian R Henderson, Blake C Meyers, Israel Ausin, Steven E Jacobsen (2015). *A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis.* Cell 163(2): 445-455.

PRESENTATIONS

2022 The 27th Annual Meeting of the RNA Society. Boulder, Colorado.
Poster: Dissecting HRAS Splicing Regulation in Myc-driven Prostate Cancer.

2019 Eukaryotic mRNA processing. Cold Spring Harbor Laboratory, New York.
Poster: Genome-wide CRISPRi/a-based screens for identifying splicing regulators in Myc-driven prostate cancer.

FELLOWSHIPS AND AWARDS

2022-2023 UCLA Warsaw Fellowship
2022 Poster Award at RNA 2022 (the 27th Annual Meeting of the RNA Society)
2021-2022 UCLA John W. Phillips Legacy Award
2019-2020 UCLA Whitcome Pre-doctoral Fellowship

Chapter 1

Introduction to splicing dysregulation in cancer

REGULATION OF ALTERNATIVE SPLICING

Precursor messenger RNA (pre-mRNA) splicing is an essential post-transcriptional gene regulation process, by removal of non-coding introns on nascent RNA and merging of exons to produce mature mRNA for protein translation. This process requires the assembly of a large catalytic ribonucleoprotein complex termed spliceosome, consisting of five small nuclear ribonucleoproteins (U snRNPs), U1, U2, U4, U5 and U6, facilitated by numerous auxiliary proteins (Wahl et al., 2009; Wilkinson et al., 2020). U1 snRNP recognizes and binds to the upstream exon and intron boundary (5' splice site), while U2 snRNP interacts with the branch point adjacent to the intron and downstream exon boundary (3' splice site), with its auxiliary factors (U2AFs) recognize and bind to 3' splice site. U4, U5 and U6 tri-snRNP then assembles and transforms spliceosome into an active conformation that excises introns.

Exons are constitutively or alternatively spliced. Almost all human genes undergo alternative splicing, resulting in an average of three or more distinct mRNA isoforms per gene, to generate proteome diversity and functional complexity. Splice site choice is regulated by additional RNA-protein and protein-protein interactions (Black, 2003; Fu and Ares, 2014; Lee and Rio, 2015). It involves short sequence motifs on pre-mRNA, known as cis-regulatory elements, and proteins that bind to these RNA-motifs, known as trans-acting splicing factors. Two major families of splicing factors involved in alternative splicing mechanism are serine/arginine-rich (SR) proteins and heterogenous nuclear ribonucleoprotein (hnRNPs) that mediates sequence-specific splicing regulation. Generally, SR proteins act positively by binding to enhancer elements, including exonic or intronic splicing enhancers (ESE or ISE), and recruit spliceosome to activate splicing. SR proteins are antagonized by hnRNPs, which bind to silencers, including exonic or intronic splicing silencers (ESS or ISS), and repress the splicing. Nevertheless, studies have showed that these regulators could mediate opposing outcomes, depending on their binding position on pre-mRNA. SR proteins could repress splicing when

bound to introns, while in some cases, hnRNPs could activate splicing in a context-dependent manner (Erkelenz et al., 2013; Huelga et al., 2012; Pandit et al., 2013). These splicing factors can cooperate and compete to regulate specific splicing events, in a concentration dependent manner. SR proteins also autoregulate their expression by coupling splicing with nonsense-mediated mRNA decay (NMD) (Lareau et al., 2007; Ni et al., 2007). Besides SR proteins and hnRNPs, many other RNA binding regulatory proteins are involved in alternative splicing regulation, such as RBFOX proteins, signal transduction and activation of RNA (STAR) proteins, and many others.

Recent high-throughput sequencing studies have further defined genome-wide differential alternative splicing patterns, including cassette exon, alternative 5' or 3' splice site, mutually exclusive exons, alternative first or last exon, and intron retention (Park et al., 2018). Among those, cassette exon, also termed skipped exon (SE), where the alternative exon could be either skipped or included, is the most frequently detected and well-defined type of alternative splicing. Using RNA-seq data, percent spliced-in (PSI or Ψ) is the widely used metrics to quantify the inclusion level of a particular alternative exon (Katz et al., 2010). It is quantified based the inclusion-specific reads divided by all reads.

SPLICING DYSREGULATION IN CANCER

Dysregulation of splicing contributes to oncogenic transformation, tumor progression, and conveys resistance to cancer therapy. Cancer-associated splicing is found altered by a variety of means (Anczuków and Krainer, 2016; Cherry and Lynch, 2020; Dvinge et al., 2016; Escobar-Hoyos et al., 2019; Urbanski et al., 2021). Disease-causing mutations often happen at splice sites or other cis-elements, and in genes encoding core spliceosomal proteins and other splicing factors acting in trans. Unmutated splicing factors are found overexpressed or under expressed in cancer cells. Other dysregulation involves changes in the post-translational modifications of splicing factors that alters their activity in tumor cells, as well as indirect effects

through transcription or epigenetic modifications that contributes to the cancer-associated splicing.

Recurrent somatic mutations of core spliceosomal proteins and other splicing factors are commonly found in hematological malignancies (Dvinge et al., 2016). Most frequently detected mutations include SF3B1 mutations, a U2 snRNP component facilitating branch point recognition, U2AF1 mutations affecting 3' splice site recognition, splicing factor SRSF2, and ZRSR2, a component of the minor spliceosome that excises non-canonical splice sites. These recurrent mutations are distributed across blood cancers such as acute myeloid leukemia (AML), chronic lymphocytic leukemia (CLL), refractory anemia with ring sideroblasts (RARS) and others. Interestingly, mutations in splicing factors are less common in solid tumors (Anczuków and Krainer, 2016; Urbanski et al., 2018). However, solid tumors do exhibit global splicing changes compared with normal tissues, accompanied with copy number variation of genes encoding splicing factors and altered splicing factor expression levels that may be driving oncogenic signaling. In this dissertation, we focus on the abnormal expression of unmutated splicing factors that potentially function as proto-oncogenes or tumor suppressors.

A number of SR proteins are found overexpressed or sometimes under expressed in a variety of solid tumors. SRSF1 (ASF/SF2) is a proto-oncogene frequently upregulated in lung, colon, breast, thyroid, and kidney cancers (Anczuków and Krainer, 2016; Urbanski et al., 2018). SRSF1 promote mammary epithelial cell transformation in both in vitro and in vivo models of breast cancer by regulating apoptosis and proliferation (Anczuków et al., 2012). It also drives tumorigenicity of small cell lung cancer and potentially reduces chemo-sensitivity (Jiang et al., 2016). Splicing factor SRSF3 (SRp20) is found overexpressed in many tumors including bladder, breast, colon, liver, lung, oral and others. SRSF3 promotes cancer cell proliferation, in part by regulating ILF3 alternative splicing (Jia et al., 2019). It forms a complex with TDP43 and alters splicing patterns in triple-negative breast cancer (Ke et al., 2018). Other SR proteins

altered in cancer involve SRSF4, SRSF5, SRSF6, SRSF7, and TRA2β. Interestingly, SRSF1, SRSF3, and SRSF4 expression are downregulated in AML and correlated with caspase-8 splicing, in contrast to their overexpression in solid tumors (Liu et al., 2012).

Expression of hnRNPs are also found altered in cancers. Two genes of the hnRNP A/B family, hnRNP A1 and the closely related hnRNP A2/B1, are both upregulated in colon and lung cancers, with A2B1 also overexpressed in brain and breast cancers (Anczuków and Krainer, 2016; Urbanski et al., 2018). hnRNP A2/B1 was reported as an oncogene in glioblastoma by regulating the splicing of tumor suppressors and oncogene RON (Golan-Gerstl et al., 2011). Moreover, PTBP1 (hnRNP I) is overexpressed in breast, brain colon and ovarian cancers. Loss of PTBP1 reduces tumor cell proliferation and invasiveness in ovarian cancer (He et al., 2007). Other hnRNPs including hnRNP K, M, H1 and C also have recurrent expression alterations in human cancers.

The transcription factor Myc has been linked to splicing mis-regulation in cancer. Myc gene family includes MYC (c-Myc), MYCL (l-Myc) and MYCN (n-Myc). Myc forms a heterodimer with Max, interacts with other cofactors, binds to the E-box motif near the transcription start sites of targeted genes, activates transcription or mediates gene repression, and regulates cell growth and proliferation (Dang, 2012). MYC is a extensively studied proto-oncogene commonly found amplified and contributes to the initiation and maintenance of many human cancers. MYC amplification results in transcriptional dysregulation, including of genes encoding components of the core splicing machinery and splicing regulatory proteins. These changes in the levels of spliceosomal components drive cancer associated changes in splicing, and MYC transformed cells have been shown to be unusually sensitive to splicing inhibition (Hsu et al., 2015; Koh et al., 2015). MYC transformation also enhances expression of multiple RNA binding regulators of splicing, leading to cancer associated changes in alternative splicing programs (Escobar-Hoyos et al., 2019). MYC upregulates the transcription of hnRNP A1, A2 and PTBP1 to deregulate

pyruvate kinase mRNA splicing in glioma (David et al., 2010). MYC upregulates SRSF1 expression, which in turn regulates the splicing of signaling kinase MKNK2 and transcription factor TEAD (Das et al., 2012). MYC also transcriptionally targets hnRNP H in controlling ARAF kinase splicing (Rauch et al., 2011).

Analyses of accumulating genome and transcriptomic data have defined splicing signatures in cancer tissues in relative to normal cells. In an earlier study, we developed a pathway-guided transcriptomic analysis of prostate cancer using 876 RNA-seq datasets from cells ranging from normal prostatic tissue to aggressive prostate cancer (Phillips et al., 2020). This identified a program of 1039 cassette exons whose splicing correlated with Myc signaling during cancer progression. Myc-correlated exons were enriched in genes encoding splicing regulatory proteins and core spliceosomal components, as well as others. RNA-seq analyses of an in vitro prostate cancer model with constitutive AKT overexpression and doxycycline-inducible MYC expression further define these Myc-correlated exons to be MYC-dependent. The splicing of HRAS exon 5 and SRSF3 exon 4 was found to be particularly responsive to Myc activity (Phillips et al., 2020; Urbanski et al., 2021). In this dissertation, we sought to dissect the splicing regulation of these MYC-dependent alternative exons and elucidate the connections between splice isoform choices and MYC oncogenic signaling.

REFERENCES

Anczuków, O., Krainer, A.R., 2016. Splicing-factor alterations in cancers. RNA 22, 1285–1301. https://doi.org/10.1261/rna.057919.116

Anczuków, O., Rosenberg, A.Z., Akerman, M., Das, S., Zhan, L., Karni, R., Muthuswamy, S.K., Krainer, A.R., 2012. The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. Nat Struct Mol Biol 19, 220–228. https://doi.org/10.1038/nsmb.2207

Black, D.L., 2003. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 72, 291–336. https://doi.org/10.1146/annurev.biochem.72.121801.161720

Cherry, S., Lynch, K.W., 2020. Alternative splicing and cancer: insights, opportunities, and challenges from an expanding view of the transcriptome. Genes Dev 34, 1005–1016. https://doi.org/10.1101/gad.338962.120

Dang, C.V., 2012. MYC on the Path to Cancer. Cell 149, 22–35. https://doi.org/10.1016/j.cell.2012.03.003

Das, S., Anczuków, O., Akerman, M., Krainer, A.R., 2012. Oncogenic Splicing Factor SRSF1 Is a Critical Transcriptional Target of MYC. Cell Reports 1, 110–117. https://doi.org/10.1016/j.celrep.2011.12.001

David, C.J., Chen, M., Assanah, M., Canoll, P., Manley, J.L., 2010. HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. Nature 463, 364–368. https://doi.org/10.1038/nature08697

Dvinge, H., Kim, E., Abdel-Wahab, O., Bradley, R.K., 2016. RNA splicing factors as oncoproteins and tumour suppressors. Nat Rev Cancer 16, 413–430. https://doi.org/10.1038/nrc.2016.51

Erkelenz, S., Mueller, W.F., Evans, M.S., Busch, A., Schöneweis, K., Hertel, K.J., Schaal, H., 2013. Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. RNA 19, 96–102. https://doi.org/10.1261/rna.037044.112

Escobar-Hoyos, L., Knorr, K., Abdel-Wahab, O., 2019. Aberrant RNA Splicing in Cancer. Annu. Rev. Cancer Biol. 3, 167–185. https://doi.org/10.1146/annurev-cancerbio-030617-050407

Fu, X.-D., Ares, M., 2014. Context-dependent control of alternative splicing by RNA-binding proteins. Nat Rev Genet 15, 689–701. https://doi.org/10.1038/nrg3778

Golan-Gerstl, R., Cohen, M., Shilo, A., Suh, S.-S., Bakàcs, A., Coppola, L., Karni, R., 2011. Splicing factor hnRNP A2/B1 regulates tumor suppressor gene splicing and is an oncogenic driver in glioblastoma. Cancer Res 71, 4464–4472. https://doi.org/10.1158/0008-5472.CAN-10-4410

He, X., Pool, M., Darcy, K.M., Lim, S.B., Auersperg, N., Coon, J.S., Beck, W.T., 2007. Knockdown of polypyrimidine tract-binding protein suppresses ovarian tumor cell growth and invasiveness in vitro. Oncogene 26, 4961–4968. https://doi.org/10.1038/sj.onc.1210307

Hsu, T.Y.-T., Simon, L.M., Neill, N.J., Marcotte, R., Sayad, A., Bland, C.S., Echeverria, G.V., Sun, T., Kurley, S.J., Tyagi, S., Karlin, K.L., Dominguez-Vidaña, R., Hartman, J.D.,

Renwick, A., Scorsone, K., Bernardi, R.J., Skinner, S.O., Jain, A., Orellana, M., Lagisetti, C., Golding, I., Jung, S.Y., Neilson, J.R., Zhang, X.H.-F., Cooper, T.A., Webb, T.R., Neel, B.G., Shaw, C.A., Westbrook, T.F., 2015. The spliceosome is a therapeutic vulnerability in MYC-driven cancer. Nature 525, 384–388. https://doi.org/10.1038/nature14985

Huelga, S.C., Vu, A.Q., Arnold, J.D., Liang, T.Y., Liu, P.P., Yan, B.Y., Donohue, J.P., Shiue, L., Hoon, S., Brenner, S., Ares, M., Yeo, G.W., 2012. Integrative Genome-wide Analysis Reveals Cooperative Regulation of Alternative Splicing by hnRNP Proteins. Cell Reports 1, 167–178. https://doi.org/10.1016/j.celrep.2012.02.001

Jia, R., Ajiro, M., Yu, L., McCoy, P., Zheng, Z.-M., 2019. Oncogenic splicing factor SRSF3 regulates ILF3 alternative splicing to promote cancer cell proliferation and transformation. RNA 25, 630–644. https://doi.org/10.1261/rna.068619.118

Jiang, L., Huang, J., Higgs, B.W., Hu, Z., Xiao, Z., Yao, X., Conley, S., Zhong, H., Liu, Z., Brohawn, P., Shen, D., Wu, S., Ge, X., Jiang, Y., Zhao, Y., Lou, Y., Morehouse, C., Zhu, W., Sebastian, Y., Czapiga, M., Oganesyan, V., Fu, H., Niu, Y., Zhang, W., Streicher, K., Tice, D., Zhao, H., Zhu, M., Xu, L., Herbst, R., Su, X., Gu, Y., Li, S., Huang, L., Gu, J., Han, B., Jallal, B., Shen, H., Yao, Y., 2016. Genomic Landscape Survey Identifies SRSF1 as a Key Oncodriver in Small Cell Lung Cancer. PLOS Genetics 12, e1005895. https://doi.org/10.1371/journal.pgen.1005895

Katz, Y., Wang, E.T., Airoldi, E.M., Burge, C.B., 2010. Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat Methods 7, 1009–1015. https://doi.org/10.1038/nmeth.1528

Ke, H., Zhao, L., Zhang, H., Feng, X., Xu, H., Hao, J., Wang, S., Yang, Q., Zou, L., Su, X., Wang, L., Wu, C., Wang, Y., Nie, J., Jiao, B., 2018. Loss of TDP43 inhibits progression of triple-negative breast cancer in coordination with SRSF3. Proceedings of the National Academy of Sciences 115, E3426–E3435. https://doi.org/10.1073/pnas.1714573115

Koh, C.M., Bezzi, M., Low, D.H.P., Ang, W.X., Teo, S.X., Gay, F.P.H., Al-Haddawi, M., Tan, S.Y., Osato, M., Sabò, A., Amati, B., Wee, K.B., Guccione, E., 2015. MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. Nature 523, 96–100. https://doi.org/10.1038/nature14351

Lareau, L.F., Inada, M., Green, R.E., Wengrod, J.C., Brenner, S.E., 2007. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature 446, 926–929. https://doi.org/10.1038/nature05676

Lee, Y., Rio, D.C., 2015. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu Rev Biochem 84, 291–323. https://doi.org/10.1146/annurev-biochem-060614-034316

Liu, J., Huang, B., Xiao, Y., Xiong, H.-M., Li, J., Feng, D.-Q., Chen, X.-M., Zhang, H.-B., Wang, X.-Z., 2012. Aberrant Expression of Splicing Factors in Newly Diagnosed Acute Myeloid Leukemia. ORT 35, 335–340. https://doi.org/10.1159/000338941

Ni, J.Z., Grate, L., Donohue, J.P., Preston, C., Nobida, N., O'Brien, G., Shiue, L., Clark, T.A., Blume, J.E., Ares, M., 2007. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev 21, 708–718. https://doi.org/10.1101/gad.1525507

Pandit, S., Zhou, Y., Shiue, L., Coutinho-Mansfield, G., Li, H., Qiu, J., Huang, J., Yeo, G.W., Ares, M., Fu, X.-D., 2013. Genome-wide analysis reveals SR protein cooperation and

competition in regulated splicing. Mol Cell 50, 223–235. https://doi.org/10.1016/j.molcel.2013.03.001

Park, E., Pan, Z., Zhang, Z., Lin, L., Xing, Y., 2018. The Expanding Landscape of Alternative Splicing Variation in Human Populations. The American Journal of Human Genetics 102, 11–26. https://doi.org/10.1016/j.ajhg.2017.11.002

Phillips, J.W., Pan, Y., Tsai, B.L., Xie, Z., Demirdjian, L., Xiao, W., Yang, H.T., Zhang, Y., Lin, C.H., Cheng, D., Hu, Q., Liu, S., Black, D.L., Witte, O.N., Xing, Y., 2020. Pathway-guided analysis identifies Myc-dependent alternative pre-mRNA splicing in aggressive prostate cancers. Proc. Natl. Acad. Sci. U.S.A. 117, 5269–5279. https://doi.org/10.1073/pnas.1915975117

Rauch, J., Moran-Jones, K., Albrecht, V., Schwarzl, T., Hunter, K., Gires, O., Kolch, W., 2011. c-Myc Regulates RNA Splicing of the A-Raf Kinase and Its Activation of the ERK Pathway. Cancer Research 71, 4664–4674. https://doi.org/10.1158/0008-5472.CAN-10-4447

Sebestyén, E., Singh, B., Miñana, B., Pagès, A., Mateo, F., Pujana, M.A., Valcárcel, J., Eyras, E., 2016. Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. Genome Res. 26, 732–744. https://doi.org/10.1101/gr.199935.115

Urbanski, L., Brugiolo, M., Park, S., Angarola, B., Leclair, N.K., Palmer, P., Sahu, S.K., Anczuków, O., 2021. MYC regulates a pan-cancer network of co-expressed oncogenic splicing factors (preprint). Cancer Biology. https://doi.org/10.1101/2021.11.24.469558

Urbanski, L.M., Leclair, N., Anczuków, O., 2018. Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics. WIREs RNA 9, e1476. https://doi.org/10.1002/wrna.1476

Wahl, M.C., Will, C.L., Lührmann, R., 2009. The spliceosome: design principles of a dynamic RNP machine. Cell 136, 701–718. https://doi.org/10.1016/j.cell.2009.02.009

Wilkinson, M.E., Charenton, C., Nagai, K., 2020. RNA Splicing by the Spliceosome. Annu Rev Biochem 89, 359–388. https://doi.org/10.1146/annurev-biochem-091719-064225

Chapter 2

HnRNP H/F regulation of a MYC dependent HRAS exon in Prostate Cancer

This chapter is adapted from the following manuscript in preparation with minor changes:

HnRNP H/F regulation of a MYC dependent HRAS exon in Prostate Cancer

Xinyuan Chen[1], Harry Taegyun Yang[2], Beatrice Zhang[3], John W. Phillips[4], Donghui Cheng[4], Frank Rigo[5], Owen N. Witte[4,6,7,8,9], Yi Xing[3,10], Douglas L. Black[4,7,8,9]

[1]Molecular Biology Interdepartmental Doctoral Program, University of California, Los Angeles, Los Angeles, California 90095, USA.

[2]Bioinformatics Interdepartmental Graduate Program, University of California, Los Angeles, California 90095, USA.

[3]Center for Computational and Genomic Medicine, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA.

[4]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, Los Angeles, California 90095, USA.

[5]Ionis Pharmaceuticals, Inc., 2855 Gazelle Ct., Carlsbad, California 92010, USA.

[6]Department of Molecular and Medical Pharmacology, University of California, Los Angeles, California 90095, USA.

[7]Jonsson Comprehensive Cancer Center, University of California, Los Angeles, California 90095, USA.

[8]Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, California 90095, USA.

[9]Molecular Biology Institute, University of California, Los Angeles, California 90095, USA.

[10]Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

Correspondence: dougb@microbio.ucla.edu; xingyi@chop.edu

ABSTRACT

The Myc proto-oncogene contributes to the pathogenesis of more than 50 percent of human cancers including prostate cancers. Malignant transformation by Myc is known to transcriptionally upregulate the core pre-mRNA splicing machinery and cause mis-regulation of alternative splicing. However, our understanding of how changes of splice isoforms affect the cancer phenotype is limited and how these changes are directed by Myc remains largely unknown. Here, we performed a signaling pathway-guided differential splicing analysis to identify Myc dependent splicing events, including an HRAS cassette exon that is repressed by Myc transformation across multiple tumor types. To molecularly dissect the regulation of this HRAS cassette exon, we used antisense oligonucleotide tiling to identify intronic splicing enhancers and silencers in the HRAS flanking introns. RNA binding motif prediction indicated the presence of multiple binding sites for hnRNP H and hnRNP F within these cis-regulatory elements. Using siRNA knockdown and cDNA expression, we found that both hnRNP H and F activate the HRAS cassette exon. Mutagenesis and targeted RNA immunoprecipitation implicate two downstream intronic G-rich elements in this H/F mediated splicing activation. Bioinformatic analyses of ENCODE RNA-seq datasets also confirmed hnRNP H regulation of HRAS splicing. Consistent with the observed effects of hnRNP H on HRAS splicing, analyses of RNA-seq datasets across multiple cancers showed a negative correlation of hnRNP H gene expression with Myc hallmark enrichment. Interestingly, hnRNP F expression exhibited a positive correlation with Myc transformation and thus was not consistent with the observed effects of hnRNP F. Loss of hnRNP H/F altered the cell cycle progression and induced apoptosis in prostate cancer cells. Collectively, our results reveal mechanisms for the Myc-dependent regulation of splicing, and point to new possible therapeutic targets in advanced prostate cancers.

INTRODUCTION

Changes in pre-mRNA splicing have emerged as important contributors to the cancer phenotype. Core splicing components including the U snRNPs assemble onto nascent RNAs to form the catalytic spliceosome that will excise each intron (Mc et al., 2009; Wilkinson et al., 2020). This assembly is regulated by RNA binding proteins that bind to the pre-mRNA at cis-regulatory elements to direct splice site choices and create alternatively spliced mRNA isoforms (Black, 2003; Fu and Ares, 2014; Lee and Rio, 2015). These regulators of splicing are very diverse and each alternative splicing event is regulated by multiple factors that can act either positively or negatively on the selection of a particular isoform. Aberrant splicing in cancer can result from mutations in core spliceosomal components that give rise to aberrant mRNAs, or from altered expression and modulation of regulatory RNA binding proteins that shift the production of particular mRNA isoforms (Dvinge et al., 2016; Escobar-Hoyos et al., 2019). These changes in isoforms can affect many aspects of the tumor phenotype including cellular growth control and cell cycle progression, suppression of apoptosis, response to hormones and growth factors, loss of cellular differentiation, metastasis, angiogenesis and drug resistance (Urbanski et al., 2018). Splicing in tumor cells also appears to be more error prone, producing mRNAs that are not normally produced elsewhere and which provide appealing targets for immunotherapies (Pan et al., 2021; Smith et al., 2019; Wang and Aifantis, 2020).

Deregulation of the Myc proto-oncogene contributes to many cancers. Myc is a DNA binding protein that interacts across the genome resulting in a broad deregulation of transcription, including of genes encoding components of the core splicing machinery. These changes in the levels of spliceosomal components drive cancer associated changes in splicing, and Myc transformed cells have been shown to be unusually sensitive to splicing inhibition (Hsu et al., 2015; Koh et al., 2015). Myc transformation also enhances expression of multiple RNA binding regulators of splicing, leading to cancer associated changes in alternative splicing programs

(Escobar-Hoyos et al., 2019). In prostate cancer, changes in the expression and function of multiple splicing regulators, including SAM68, hnRNP L, and others have been shown to contribute to the cellular phenotype (Caggiano et al., 2019; Fei et al., 2017; Ho et al., 2021).

One family of RNA binding proteins implicated in a variety of aspects of cancer are the heterogenous nuclear ribonucleoproteins H and F (hnRNP H/F). HnRNP H is encoded on three genes H1, H2, H3, and hnRNP F on one gene. HnRNPs H and F bind to G-run motifs GGG and GGGG, that act to enhance splicing of alternative exons when present downstream and to repress splicing when present upstream or within the exon (Caputi and Zahler, 2001; Carlo et al., 1996; Chou et al., 1999; Garneau et al., 2005; Martinez-Contreras et al., 2006; Matunis et al., 1994; Min et al., 1995; Modafferi and Black, 1999, 1997; Schaub et al., 2007; Xiao et al., 2009). Activation of splicing by G-run elements is strongly affected by regulatory elements in the upstream exon and its 3' splice site, indicating that hnRNP H/F activity requires additional cofactors (Modafferi and Black, 1999, 1997). The two proteins have similar effects on splicing, but slightly different binding specificities and can differ in their activities on particular target exons (Huelga et al., 2012). They can also form a heterodimer that may allow them to coordinately affect some targets (Min et al., 1995; Chou et al., 1999; Schaub et al., 2007). HnRNP H was found to be upregulated in glioma (LeFave et al., 2011), as well as colon cancer, and head and neck cancers (Rauch et al., 2010, 2011). Oncogenic splicing switches driven by hnRNP H include targets such as IG20/MADD in glioma (LeFave et al., 2011), TCF3 in lymphoma (Yamazaki et al., 2018), Mcl-1 and HER2 in breast cancer (Gautrey et al., 2015; Tyson-Capper and Gautrey, 2018), KHK in hepatocellular carcinoma (Li et al., 2016), and A-Raf in colon and head and neck cancers (Rauch et al., 2010). hnRNP H also regulates alternative splicing of the oncogenic fusion transcript EWS-FLI1 (Vo et al., 2022), and the RON protooncogene (Braun et al., 2018), and may alter translation in glioblastoma (Herviou et al., 2020). HnRNP F is less studied in the context of cancer cells but has

been shown to be needed for the productive splicing of Sam68 in Prostate cancer (Caggiano et al., 2019).

The availability of whole-transcriptome sequencing data across cancers has enabled the definition of splicing signatures in cancer tissues compared to normal cells. In an earlier study, we developed a pathway-guided transcriptomic analysis of prostate cancer using 876 RNA-seq datasets from cells ranging from normal prostatic tissue to aggressive prostate cancer (Phillips et al., 2020). This identified a program of 1039 cassette exons whose splicing correlated with Myc signaling during cancer progression. Myc-correlated exons were enriched in genes encoding splicing regulatory proteins and core spliceosomal components, as well as other cellular functions. The splicing of HRAS exon 5 was found to be particularly responsive to Myc activity, and the correlation between HRAS splicing and Myc activation is found in other tumor types (Phillips et al., 2020; Urbanski et al., 2021). HRAS belongs to the Ras oncogene family, regulates cell division, and is involved in multiple signaling transduction pathways. HRAS exon 5 affects overall expression from the gene, such that its inclusion leads to premature translation termination, and nonsense-mediated decay of the HRAS mRNA (Cohen et al., 1989). Transcripts that escape NMD encode a C-terminal truncated p19 Ras protein with distinct functions from the canonical p21 Ras protein (Guil et al., 2003a; Camats et al., 2009). High Myc levels lead to reduced exon 5 splicing and potentially higher levels of p21 HRAS protein.

Here we report that large scale bioinformatic analyses of splicing and Myc expression confirms the correlation of HRAS exon 5 repression with myc signature score in prostate cancer and across many tumor types. To obtain mechanistic links between Myc oncogenic transformation and splice isoform choices we dissected the regulation of HRAS exon 5 splicing. We utilized antisense oligonucleotide tiling to identify several intronic splicing enhancers and silencers adjacent to the exon. RNA binding motif enrichments indicated the presence of many hnRNP H/F binding sites within these cis-regulatory regions. We found that both hnRNP H and F activate

HRAS exon 5 splicing and this activation required G_4 and G_3 elements in the downstream intron. Bioinformatic analyses of ENCODE RNA-seq datasets confirm hnRNP H regulation of HRAS exon 5 and indicated that it is one of many exons regulated by both MYC and H/F. Additional pan-cancer bioinformatic analyses correlate the downregulation of hnRNP H expression and the upregulation of hnRNP F with Myc hallmark score. This transcriptional regulation was confirmed using published nascent RNA-seq datasets from osteoblasts carrying an inducible MYC. Loss of hnRNP H/F resulted in cell cycle arrest and induced apoptosis in prostate cancer cells. Taken together, our results reveal mechanisms by which Myc alters splicing regulation.

RESULTS

Pathway Enrichment-Guided Activity Study of Alternative Splicing (PEGASAS) identifies HRAS exon 5 as repressed by Myc transformation across multiple tumor types

The protooncogene HRAS contains a conserved poison exon (exon 5) that alters both its expression and function. The exon-skipped isoform encodes the full length functional p21 HRAS protein, while the exon-included isoform contains a premature termination codon (PTC) that triggers the nonsense-mediated decay (NMD) of the HRAS transcripts. Transcripts that escape from NMD are translated into a C-terminal truncated p19 HRAS protein (Cohen et al., 1989; Guil et al., 2003a) (**Figure 1A**). HRAS p21 and p19 share most of the N-terminal G domain that mediates GTP hydrolysis (Pálfy et al., 2020; Simanshu et al., 2017). However, p19 lacks the last 16 amino acids of the allosteric lobe, and is reported not to bind GTP (Guil et al., 2003a) (**Figure 1B**). HRAS p21 also has a C-terminal hypervariable region (HVR) that is responsible for membrane binding and trafficking (Simanshu et al., 2017). HRAS p19 lacks this C-terminal domain, and in its place a C-terminal 20 amino acid sequence that is conserved across species, but whose function is not known. Several bioinformatic studies have connected HRAS exon 5 splicing with Myc transformation. HRAS exon 5 inclusion was found to be anti-correlated with Myc activity across prostate and breast cancers (Phillips et al., 2020). Greater skipping of this exon

was also seen in Myc-active tumors in a pan-cancer analysis that implicated a network of SR proteins in its regulation (Urbanski et al., 2021).

To broadly assess HRAS splicing changes in response to Myc signaling pathway activation in tumors, we used the computational framework PEGASAS (Phillips et al., 2020) to analyze RNA-seq data compiled from 9,490 samples from The Cancer Genome Atlas (TCGA) and 5,862 tissue matched samples from the Genotype-Tissue Expression project (GTEx) (Lonsdale et al., 2013; Weinstein et al., 2013). Briefly, gene expression and exon inclusion (Percent spliced in, PSI) values were computed for all genes and exons for each sample. Myc activity scores were calculated using the hallmark gene set Myc Targets V2 from the Molecular Signatures Database (MSigDB) (Liberzon et al., 2015). Myc activity scores were then correlated with all exon PSI values across all the datasets. Myc activity was seen to increase with disease progression from normal tissue (gray), to tumor adjacent benign tissue (orange), and to primary tumor (light blue) and more malignant disease stages. Inclusion of HRAS exon 5 is found to negatively correlate with Myc hallmark enrichment in the majority of 27 tumor types (**Figure 1C**). In addition to the correlation that we observed previously in prostate adenocarcinoma (PRAD), other epithelial cell cancers, such as colon adenocarcinoma (COAD) and lung squamous cell carcinoma (LUSC) showed particularly strong correlations (**Figure 1D**).

ASO tiling reveals splicing enhancers and silencers controlling HRAS exon 5

As a first step in delineating the regulatory elements in the region of exon 5, we applied antisense oligonucleotides (ASOs) that base pair and potentially block RNA regulatory elements (Rinaldi and Wood, 2018). We designed and synthesized 22 ASOs that tiled across the highly conserved sequences of the HRAS exon 5 region, including 94 nucleotides of intron 4, the 82nt exon, and 200nt of intron 5 (**Figure 2A**). These ASOs were 18nt in length, non-overlapping, and had a uniform phosphorothioate backbone chemistry and methoxyethyl modifications at the 2’ ribose position (2’MOE-PS). Each ASO, and a non-targeting control (NTC) ASO, was transfected

into HEK293 cells. 24 hours after transfection, RNA was isolated and HRAS splicing was measured by semi-quantitative RT-PCR. Comparing the exon 5 percent-spliced in (PSI) value in the presence of each ASO to the NTC, we identified ASO's that decrease exon 5 splicing and others that increase it (**Figure 2B**). As expected, ASOs targeting the 3' and 5' splice sites strongly inhibited splicing. ASOs targeting the body of the exon also induced exon skipping, indicating the presence of exonic splicing enhancers within the HRAS exon. ASOs targeting downstream intron 5 showed diverse effects with some increasing and others decreasing exon 5 splicing, suggesting the presence of multiple intronic splicing silencers and enhancers. ASOs I5-1, I5-4, and I5-7 increased exon inclusion and may block intronic splicing silencer elements (ISSs). In contrast, ASOs I5-3, I5-5**,** I5-9, and I5-11 all induced exon skipping compared to controls and suggest the presence of splicing enhancers in this region.

We next constructed an HRAS minigene reporter by cloning the genomic region spanning exon 5, including the flanking introns and portions of exons 4 and 6, into the pcDNA3.1(+) expression vector (**Figure S1A**). We introduced an in-frame ATG start codon downstream from the CMV promoter, and a TGA stop codon upstream of the BgH polyadenylation site to reduce nonsense mediated decay of the product mRNA. In HEK293 cells, transcripts from the minigene show 18.9% exon 5 inclusion compared to 8.0% inclusion in the endogenous HRAS transcripts (**Figure S1B-C**). In co-transfection experiments we found that ASOs had pronounced effects on splicing the HRAS minigene. With some exceptions, the ASO induced minigene splicing changes were consistent with those on endogenous HRAS. ASO's targeting the exon were strongly inhibitory. Enhanced splicing, indicating the position of an intronic splicing silencer, was seen with I5-1 and I5-7. The adjacent I5-6 ASO also enhanced splicing of the minigene transcripts, but had more limited effect on the endogenous transcripts. Splicing inhibition due to the likely blocking of splicing enhancers was seen with ASO's I5-5, and I5-8 to I5-11, often with stronger effects on the minigene RNA than seen on the endogenous RNA (**Figure 2B, S1C**). ASO I4-1 in intron 4

generated a new band in the RT-PCR likely due to a cryptic splice site in the minigene construct. The weaker effects of the ASO's on the endogenous RNA could result from a number of factors. Different rates of transcription for the native and transfected genes could affect the ability of ASO's to act on the RNA. Also mature endogenous RNA is already present at the time of ASO transfection and may not turn over completely during the assay of the ASO's. In contrast, the minigene is co-transfected with the ASO's so all RNA is processed in their presence. A previous study identified a splicing silencer element called rasISS1 in HRAS intron 5 (Guil et al., 2003b) (**Figure 4A**). The inhibitory sequences of rasISS1 maps to the region covered by our ASOs I5-1 and I5-2. Our data also indicates a silencer in the I5-1 region. The limited effect we observe from I5-2 may be due to the secondary structure proposed for this region interfering with targeting by antisense oligos. We have focused on regulatory elements where ASO's inhibited splicing of the endogenous HRAS exon indicating intronic enhancer elements (I5-3, I5-5, and I5-8).

To identify trans-acting splicing factors that potentially bind the HRAS splicing regulatory elements, we examined the sequences surrounding HRAS exon 5 with motif finding tools SpliceAid2 and RBPmap (Paz et al., 2014; Piva et al., 2012). Within the region targeted by the ASO's downstream of exon 5, many binding motifs for hnRNP H and its paralog hnRNP F were identified, as well as for motifs for hnRNP A1, SRSF5, and other proteins (**Figure 2C, S2A**). Previous work identified hnRNP A1 as repressor of exon 5 that binds to the rasISS1 element (Guil et al., 2003b). The SR proteins SRSF2 and SRSF5 were also identified as factors that correlate with exon 5 activation.

hnRNPs H and F activate HRAS exon splicing

Heterogenous nuclear ribonucleoproteins H and F (H1, H2, H3 and F) are paralogous splicing factors that activate splicing when bound downstream of alternative exons. H and F both bind to G-run motifs GGG and GGGG, although their specificity has slight differences (Caputi and Zahler, 2001; Carlo et al., 1996; Chou et al., 1999; Garneau et al., 2005; Martinez-Contreras et

al., 2006; Matunis et al., 1994; Min et al., 1995; Modafferi and Black, 1999, 1997; Schaub et al., 2007; Xiao et al., 2009). HnRNP H3 is a pseudogene that lacks the N-terminal RNA Recognition motif (Mahé et al., 1997; Honoré, 2000). HnRNP's H1 and H2 are highly similar in peptide sequence (Honoré et al., 1995), but differ in their relative expression, with H1 more highly expressed in most cell types including HEK293 cells and a Myc dependent prostate cancer model (Lonsdale et al., 2013; Phillips et al., 2020). (Honoré et al., 1995; Mahé et al., 1997). In this study, we focus on hnRNP H1 and F.

To assess the effects of hnRNPs F and H on HRAS splicing, we performed siRNA-mediated knockdown of these factors individually and together in HEK293 cells. Immunoblot confirmed that the siRNAs depleted hnRNP F by 92.7% and H1 by 69.6% when they were targeted individually, and by 88.9% and 68.1% when hnRNPs F and H were targeted together (**Figure 3A**). Depletion of F or H individually led to a modest increase in the other factor, an apparent cross-regulation that is commonly observed for pairs of related RNA binding proteins. 72 hours after introduction of the siRNAs, we assayed the splicing of exon 5 in the endogenous HRAS transcripts by RT-PCR. Depletion of either hnRNP F or hnRNP H decreased exon 5 splicing, with a stronger effect seen with the loss of hnRNP H, despite its less complete depletion. Depletion of both H and F resulted in even greater exon skipping. Thus, both H and F act to enhance HRAS splicing (**Figure 3B**). H/F knockdown resulted in an upregulation of the HRAS p21 protein isoform as seen by immunoblot, consistent with the observed splicing changes (**Figure 3A**). To rule out off-target effects of the siRNAs, we re-expressed 6xHis tagged siRNA-resistant HNRNPH1 or HNRNPF cDNAs, after siRNA depletion of the endogenous transcripts (**Figure 3C-D**). Immunoblots confirmed the expression of recombinant hnRNP H or F at levels comparable to the endogenous proteins. Re-expression of either hnRNP H or F stimulated splicing of HRAS exon 5 in both minigene and endogenous RNAs validating their roles as splicing activators and ruling out possible off target effects of the siRNAs.

G_3 and G_4 elements in the downstream intronic splicing enhancer mediate the hnRNP H/F dependent enhancement of HRAS splicing.

HnRNPs H and F bind to motifs containing runs of three or four G nucleotides which act as splicing enhancers when found in a downstream intron (Chou et al., 1999; Schaub et al., 2007). There are ten G-runs downstream of exon 5, of which four were identified by the ASO tiling as enhancer elements (**Figure 4A**). We constructed a series of minigene reporters carrying single mutations at each of these four G-run elements, a double mutation of the neighboring G1 and G2 runs, and a mutation of the four G-runs (G1,G2,G3,G4). After transfection into HEK293 cells, the splicing of each of these constructs was compared to the wildtype clone by RT-PCR (**Figure 4B**). We observed small decreases in exon 5 splicing when either G1 or G2 was mutated, and minimal splicing changes resulting from G3 or G4 mutations. In the constructs carrying single G-run mutations the expression of recombinant F or H enhanced exon inclusion. The double mutation of both G1 and G2 resulted in a nearly complete loss of exon 5 splicing and a similar effect was seen when all four G runs were mutated together. Notably, for these mtG1G2 and mtG1G2G3G4 constructs, the overexpression of H or F could no longer rescue exon 5 splicing. We also tested mutations in other G-runs downstream of exon 5 but where the blocking ASO either indicated the presence of a splicing silencer or had minimal effect (G-runs located in the I5-1, I5-4 and I5-6 targeted regions, Figure 4A). The splicing changes induced by these mutations were mostly consistent with the ASO data, except one G-run in the I5-1 region whose mutation indicated that it also acted as an enhancer but the ASO was apparently blocking a silencer (Supplementary table 5). This region is thus complex and likely contains multiple regulatory elements, one of which might be an additional hnRNP H/F dependent enhancer. Altogether the results indicate that multiple G-runs within HRAS intron 5 act as hnRNP H/F dependent splicing enhancers, with individual elements acting redundantly and with G1 and G2 having the strongest effects.

To assess the interactions of hnRNPs H and F with HRAS intron 5, we assayed for the presence of HRAS RNAs in hnRNP H/F immunoprecipitates. HEK293 cells were transfected with the wildtype HRAS minigene or that carrying the G1G2 mutation. These were examined in cells expressing 3xFlag tagged hnRNP H or 6xHis tagged hnRNP F, or in untransfected cells to assay endogenous H and F protein. Ectopically expressed H and F were immunoprecipitated with anti-Flag and anti-His antibodies respectively, and endogenous hnRNPs H and F with antibodies reactive with each protein. The specificity of the pull downs and yield of the immunoprecipitations were monitored by immunoblot (**Figure 4C**). RT-PCR analysis of the immunoprecipitated RNA identified HRAS pre-mRNA associated with the H and F immunoprecipitates but not the control IgG. Primer pairs were designed to amplify either the exon4-intron4 junction or a segment of intron 5 neighboring the G-runs. These yielded the expected products from the minigene reporter whose unspliced products and much more abundant than the nascent endogenous HRAS RNA, while a minus reverse transcriptase control confirmed the absence of minigene DNA (**Figure S3A**). To quantify the amounts of mutant and wildtype RNAs in each pulldown, we performed qRT-PCR, normalizing the amount of RNA in the precipitates to the input. We found that the binding of HRAS RNA to hnRNPs H and F was substantially reduced by G1G2 mutation. This was seen in immunoprecipitates of both the exogenous and endogenous proteins (**Figure 4D**). The Intron 5 fragment was more abundant in the immunoprecipitates than the exon 4 - intron 4 fragment and showed the biggest reduction in binding with the G1G2 mutation. However, reduced binding was also observed with the exon-intron fragment. Taken together, the data indicate hnRNPs H and F interact with the G-runs in HRAS intron 5 to activate exon 5 splicing.

HNRNPH gene expression is repressed by Myc activation across multiple tumor types

To further assess the association of hnRNPs H and F with Myc, we performed a correlation analysis of Myc activity score and normalized splicing factor expression across tumor types. We used DEseq2 to normalize the read counts of 220 genes of known or putative splicing factors

(Han et al., 2013) in the TCGA and GTEx samples analyzed above. We then correlated splicing factor expression with Myc activity scores computed from PEGASAS. HNRNPH1 exhibited a clear negative correlation with Myc activation in the majority of 27 tumor types, including prostate cancer (PRAD, **Figure 5A-B**). In contrast, HNRNPF exhibited a positive correlation with Myc activation in almost all tumor types (**Figure S4A-B**). The negative correlation between HNRNPH1 and Myc was found in multiple epithelial cancers such as breast invasive carcinoma (BRCA), colon adenocarcinoma (COAD), lung squamous cell carcinoma (LUSC) and Ovarian serous cystadenocarcinoma (OV). Interestingly, the correlations were reversed in the acute myeloid leukemia (LAML) samples, in these tumors hnRNP H1 was positively correlated with Myc and hnRNP F negatively correlated.

To look more directly at transcription of the hnRNP H1 and hnRNP F loci, we examined datasets of nascent RNA-seq generated with 4-Thio-U labeling under conditions of normal endogenous Myc, Myc depletion, and Myc overexpression in U2OS osteoblastoma cells carrying an inducible Myc gene (Baluapuri et al., 2019). These nascent RNA data indicated increased HNRNPF transcription and decreased HNRNPH1 transcription upon MYC overexpression (MYC_HIGH) compared with normal Myc levels (MYC_ON) or with MYC-depletion (MYC_OFF) (**Figure S5A-B**) (Baluapuri et al., 2019). These data are consistent with the pan-cancer analysis showing negative correlation of hnRNP H RNA with Myc status and positive correlation of hnRNP F RNA. They also indicate that the change in steady state hnRNP H RNA with MYC overexpression results from reduced transcription of HNRNPH1. Overall, the data are consistent with a model that down regulation of hnRNP H1 by Myc contributes to reduced HRAS exon 5 splicing in prostate and other epithelial tumors.

The HRAS exon is one of many exons controlled by both MYC and hnRNP H/F activity

To identify additional exons regulated by hnRNPs H and F, we analyzed RNA-seq datasets from the ENCODE project (Dunham et al., 2012). Using rMATS-turbo, we compared

RNA from HepG2 cells after HNRNPH or HNRNPF knockout to RNA from the non-targeted control cells (**Figure 6A, S6A**) (Shen et al., 2014). This identified 2190 and 1516 changes in skipped exons (SE) after H and F knockout respectively. HnRNP H can act either as a splicing repressor or activator depending on its binding location (Caputi and Zahler, 2001; Carlo et al., 1996; Chou et al., 1999; Garneau et al., 2005; Martinez-Contreras et al., 2006; Modafferi and Black, 1999, 1997; Schaub et al., 2007; Xiao et al., 2009). Consistent with this, we observed that approximately 50% of the exons showed reduced splicing upon hnRNP H knockout, indicating the protein acted to enhance their splicing. These included HRAS exon 5 which exhibited a PSI of 0.11 in the HNRNPH1 knockout samples, compared with 0.27 in the non-targeted control (**Figure 6B**). Similar overall results were obtained with the hnRNP F knockout, except that HRAS exon 5 showed no significant change upon hnRNP F knockout compared with the non-targeted control. This limited effect might result from the lower expression of hnRNP F yielding a smaller effect on the PSI after knockout (**Figure S6A-B**). To further validate the ENCODE findings, we knocked down hnRNPs H and F by siRNA in HepG2 cells and performed RT-PCR on the endogenous HRAS transcripts. This confirmed that the loss of either H or F alone reduce HRAS exon inclusion, and the double knockdown has a stronger effect (**Figure S6C**). Interestingly, in these cells the double depletion of H and F also slightly reduced MYC expression as seen by immunoblot (see below). The data for HepG2 and HEK293 both confirm that hnRNP H1 is an activator of HRAS exon 5 splicing (Figure 3B).

Analyses of the other ENCODE line, K562, also showed that HRAS exon 5 was positively regulated by hnRNP H1 (data not shown). However, in these cells hnRNP F gave different results exhibiting an increase in exon 5 splicing after hnRNP F knockout (data not shown). Presumably K562 expresses a different complement of splicing regulators that allow hnRNP F to act differently on some of its target RNAs. Notably, in our pan-cancer gene expression analysis (Figure 5, S4), we found both hnRNP H and F are under different transcriptional control by MYC in Acute Myeloid

Leukemia (LAML), compared to most other tumor types. Thus, the lymphoblastic K562 cells derived from a chronic myelogenous leukemia may behave more similarly to LAML than to HepG2.

To examine whether other exons that are regulated by hnRNPs H and F are also regulated by Myc, similarly to HRAS, we analyzed RNA-seq data from a prostate cancer model carrying a doxycycline-inducible MYC gene (Phillips et al., 2020). This identified 2591 differentially spliced SE events between the MYC-on and MYC-off conditions. These MYC-dependent SE events overlapped with several hundred of the hnRNP H1 or F -dependent SE events from HepG2 (**Figure 6C**). A hypergeometric test comparing these exon sets yielded a p-value of $9.53xE^{-141}$ indicating a statistically significant fraction of the splicing events regulated by MYC overlap with those controlled by hnRNP H1 and/or hnRNP F. Gene ontology analyses (PANTHER) (Thomas et al., 2022) indicate that the genes containing splicing events in the three way overlap are enriched for genes involved in metabolic processes, with mRNA metabolic processes as the top term (**Figure 6D**). It appears that there are multiple exons controlled by both MYC and hnRNPs H and F, and that HRAS exon 5 is just one example.

hnRNP H/F are required for cell proliferation in prostate cancer cell lines

To evaluate effects of splicing factors hnRNP H and F on the growth of Myc transformed cells, we knocked down their expression in two prostate cancer cell lines. PC3 is an advanced adenocarcinoma cell line with high metastatic potential, while DU145 cells derive from a prostate carcinoma with moderate metastatic potential. PC3 and DU145 cells transfected with siRNAs targeting hnRNP's F, H, or both showed >75% depletion of each factor. The combined depletion of hnRNP H and F induced HRAS exon 5 skipping as seen previously (**Figure 7A-B**). The knockdown of H or F or both also resulted in a small reduction in MYC in the PC3 cells. The double knockdown also showed an increase in cPARP protein indicating the induction of cell apoptosis (**Figure 7A**), and it was apparent that the cultures had stopped proliferating after the hnRNP H/F depletion. To evaluate how loss of hnRNP H and F affected cell growth, we performed

flow cytometry of propidium iodide stained PC3 cells. Depletion of hnRNP F or hnRNP H, and particularly the double knockdown, reduced the number of cells in the G1 phase and increased cells in G2 compared with control cells (**Figure 7C-D**), indicative of a mitotic block. HnRNP's H and F are thus needed for proper mitotic progression in the prostate cancer cells.

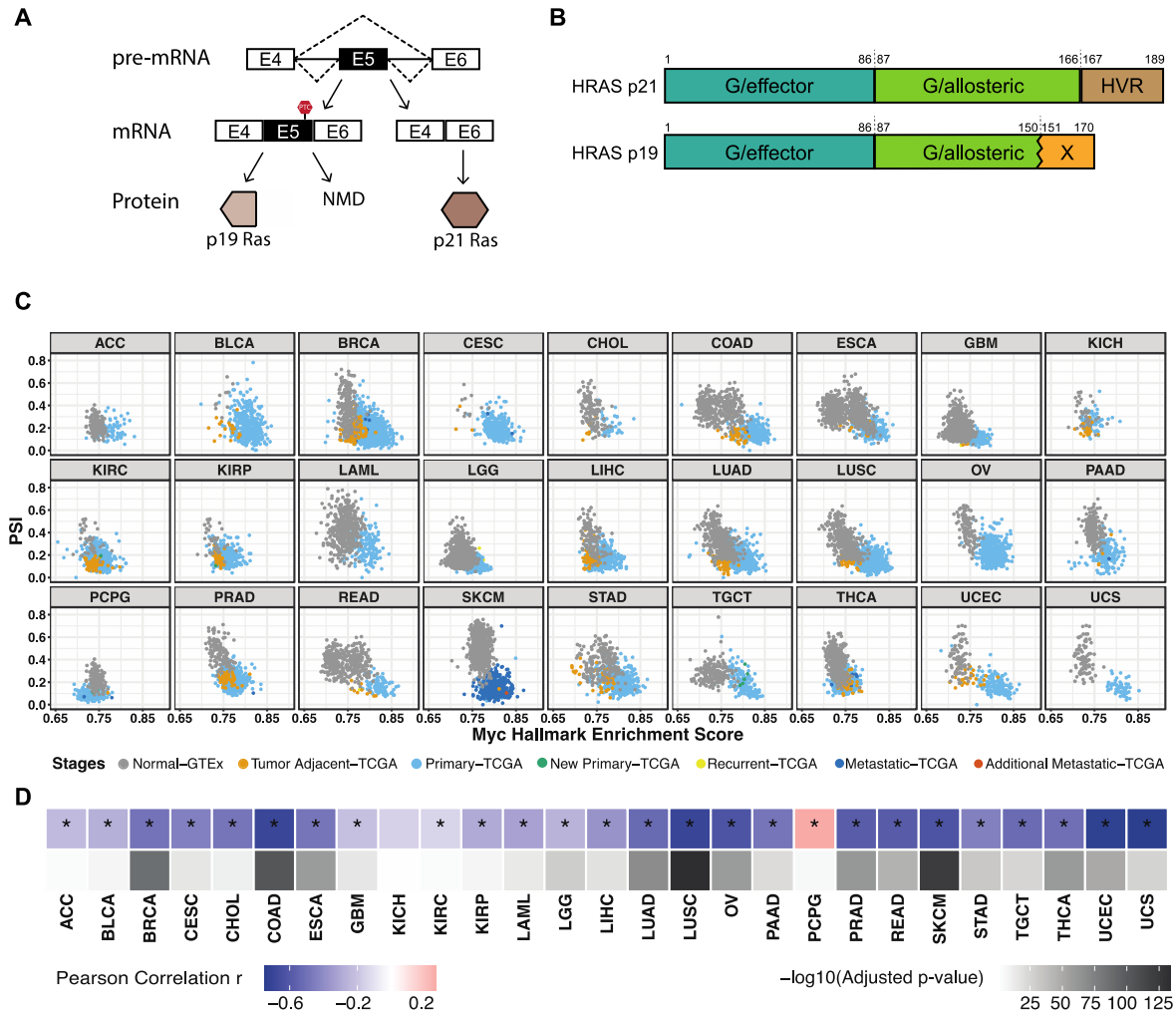

Figure 2-1. Pan-cancer analysis indicates that HRAS exon 5 is repressed by Myc activation cross multiple tumor types. (A) Diagram of HRAS pre-mRNA alternative splicing. NMD, nonsense-mediated decay. **(B)** Domain diagram of p21 and p19 HRAS isoforms (Pálfy et al., 2020). G/effector, G domain/effector lobe; G/allosteric, G domain/allosteric lobe; HVR, hypervariable region. **(C)** Scatterplot matrix showing the correlation of HRAS exon 5 PSI with Myc hallmark enrichment score across multiple tumor types. ACC, Adrenocortical carcinoma; BLCA, Bladder Urothelial Carcinoma; BRCA, Breast invasive carcinoma; CESC, Cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, Cholangiocarcinoma; COAD, Colon adenocarcinoma; ESCA, Esophageal carcinoma; GBM, Glioblastoma multiforme;

Figure 2-1. Continued.

KICH, Kidney Chromophobe; KIRC, Kidney renal clear cell carcinoma; KIRP, Kidney renal papillary cell carcinoma; LAML, Acute Myeloid Leukemia; LGG, Brain Lower Grade Glioma; LIHC, Liver hepatocellular carcinoma; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; OV, Ovarian serous cystadenocarcinoma; PAAD, Pancreatic adenocarcinoma; PCPG, Pheochromocytoma and Paraganglioma; PRAD, Prostate adenocarcinoma; READ, Rectum adenocarcinoma; SKCM, Skin Cutaneous Melanoma; STAD, Stomach adenocarcinoma; TGCT, Testicular Germ Cell Tumors; THCA, Thyroid carcinoma; UCEC, Uterine Corpus Endometrial Carcinoma; UCS, Uterine Carcinosarcoma. **(D)** Heatmap summarizing the Pearson correlation coefficient for PSI vs Myc score in each tumor type, accompanied with corresponding adjusted p-value. *, tumor types with statistically significant adjusted p-value (< 0.05).

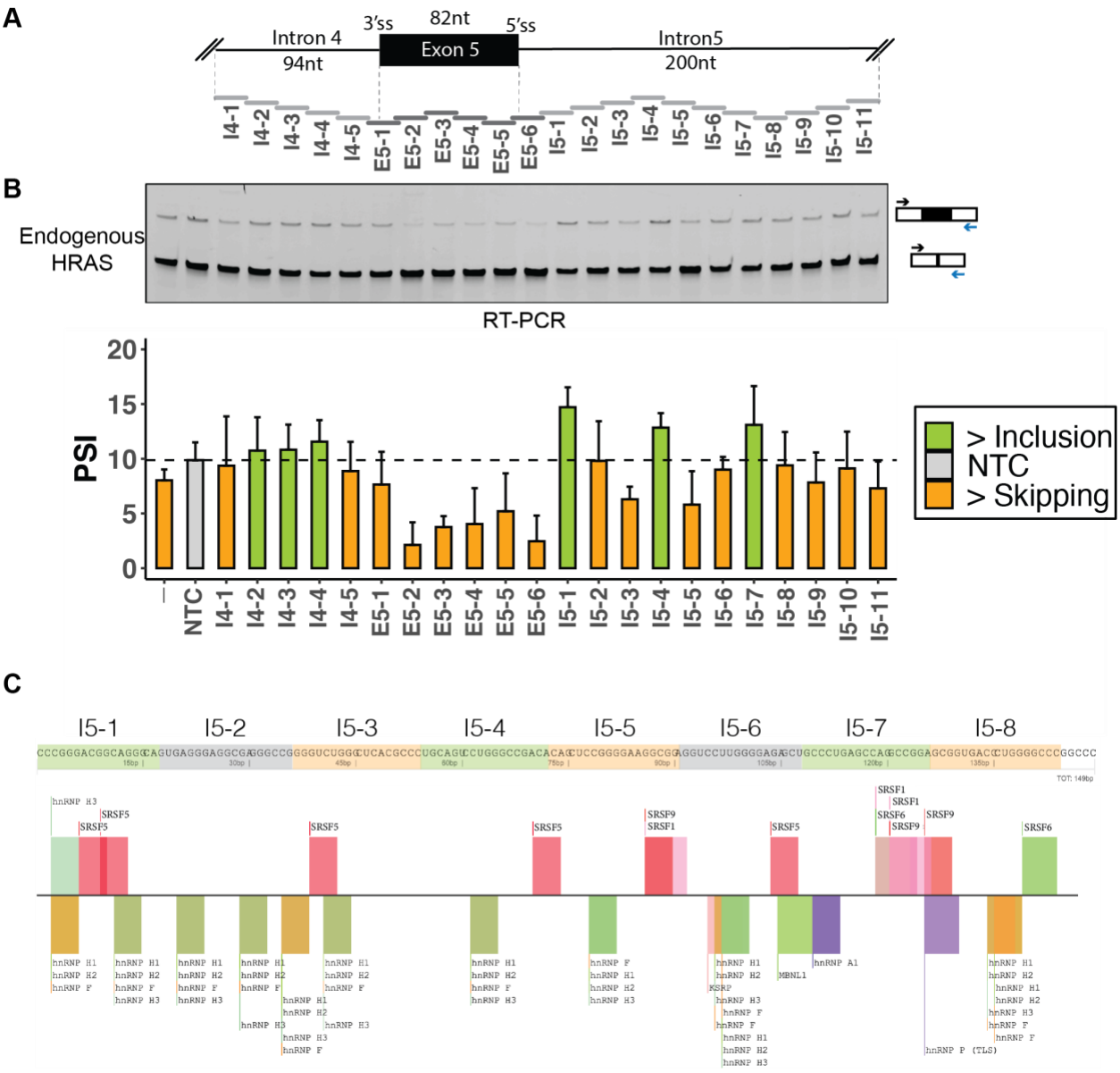

Figure 2-2. Identification of splicing cis-regulatory elements controlling HRAS exon 5. (A) Schematic of ASO tiling across HRAS exon 5 and its partial flanking introns, including 376-nucleotide region tiled by non-overlapping 18-mer ASOs. Each horizontal bar represents an ASO. **(B)** Semi-quantitative RT-PCR analyses showing the effects of ASOs on endogenous HRAS splicing. The arrows indicate RT-PCR primers for assay of exon 5 in the endogenous HRAS transcripts. The bar graph presents the quantification of the RT-PCR calculated as percent-spliced in (PSI) (gray: control; orange: more skipping; green: more inclusion).

Figure 2-2. Continued.

Each bar represents the mean value +/- SD of three biological replicates. NTC, non-targeting control. -, no-ASO mock control.

(C) SpliceAid2 snapshot of predicted splicing factor binding sites on HRAS intron 5. The sequences targeted by each ASO are highlighted in orange (more skipping), green (more inclusion), or gray (no change).

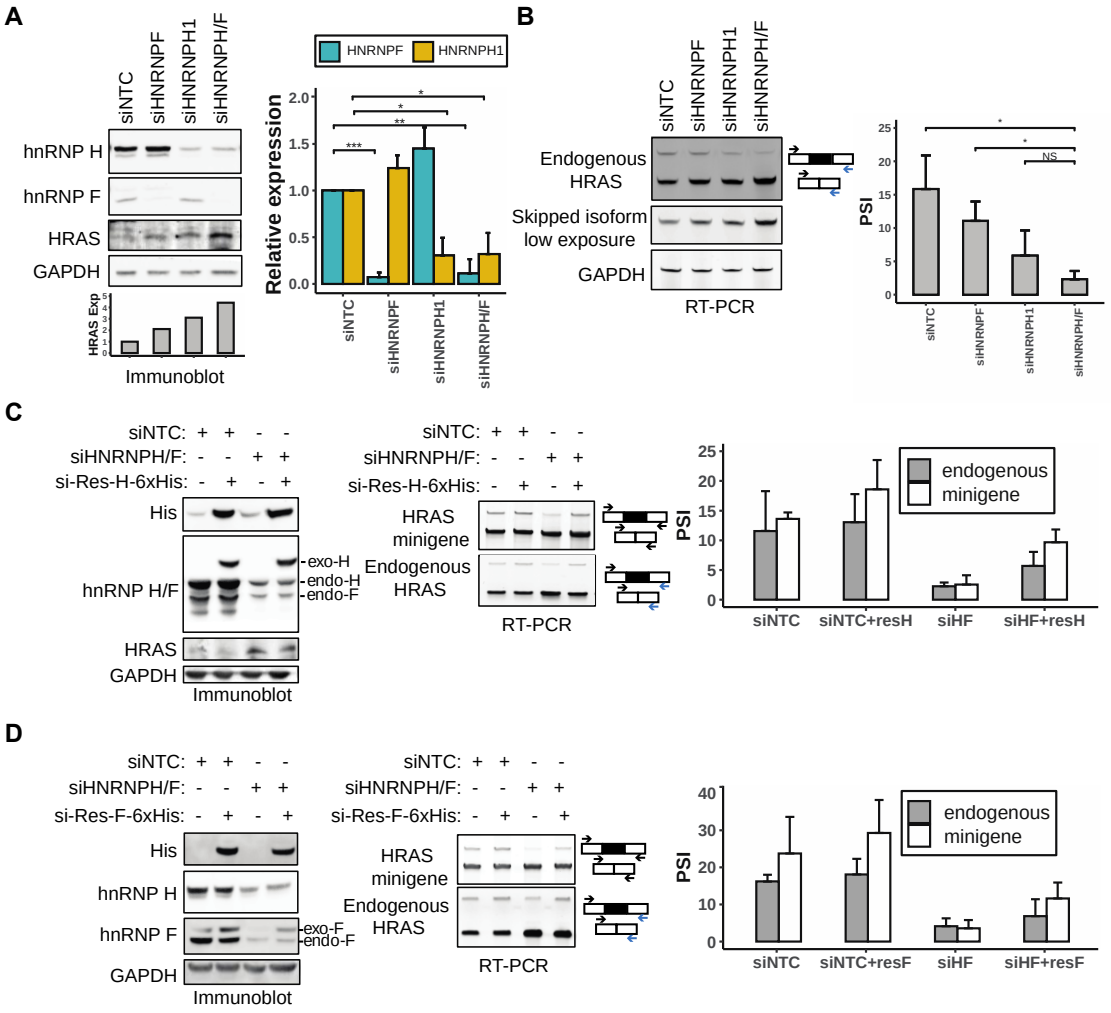

Figure 2-3. hnRNP H and F activate HRAS exon 5 splicing. (A) Immunoblot showing the expression of hnRNP H, hnRNP F, and HRAS proteins in HEK293 cells, after transfection with control, HNRNPF, HNRNPH1, or HNRNPH/F siRNAs. GAPDH was used as a loading control. The bar graph (right) shows the quantification of hnRNP H and F proteins in response to each siRNA perturbation. The grayscale bar graph (bottom) shows the quantification of HRAS p21 protein expression in response to each siRNA perturbation. NS: $p >= 0.05$; *: $p < 0.05$;**: $p < 0.01$; ***: $p < 0.001$ (Student's t-test).

Figure 2-3. Continued.

(B) RT-PCR analysis endogenous HRAS splicing after siRNA knockdown of hnRNP's H and F. The bar graph (right) shows the quantification of the RT-PCR results. **(C-D)** Rescue of hnRNP H/F expression after knockdown in HEK293 cells. Immunoblot of hnRNP H, hnRNP F, His tagged rescue protein, and HRAS in HEK293 cells. Cells were transfected with control or HNRNPH/F siRNAs followed by transfection with siRNA-resistant C-terminal 6xHis tagged hnRNP's H (C) or F (D). RT-PCR of minigene and endogenous HRAS splicing in each experimental condition is quantified in the bar graphs (right). Each bar represents the mean +/- SD of three biological replicates.

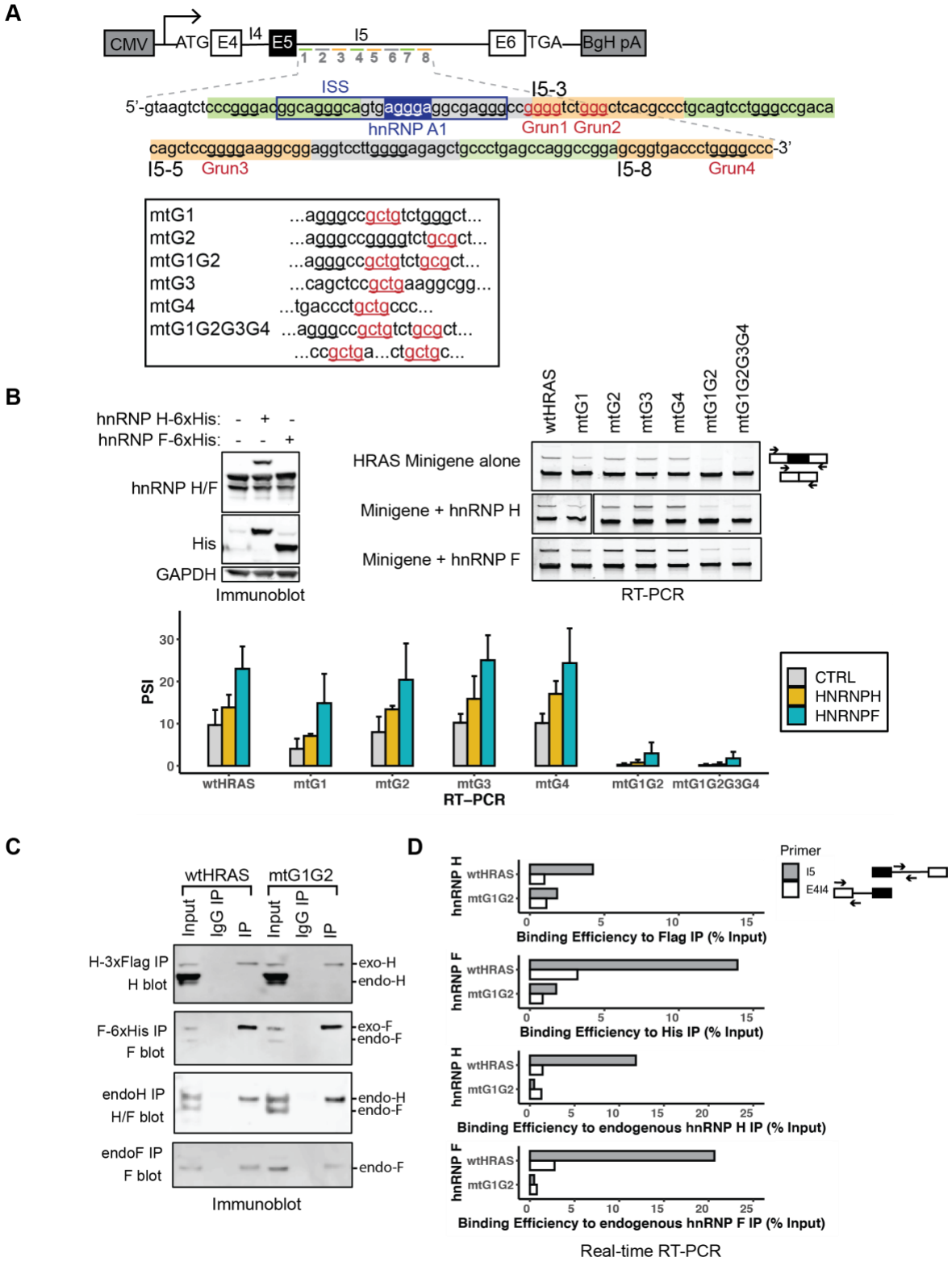

Figure 2-4. hnRNP H and hnRNP F modulate HRAS exon 5 splicing through G-run elements within the downstream ISE.

Figure 2-4. Continued.

(A) Diagram of HRAS minigene reporters carrying mutations at putative hnRNP H/F binding motifs. Intron 5 nucleotide sequences targeted by ASO's I5-1 to I5-8 are shown in orange as potential enhancers and green as potential suppressers. G-runs within this intron 5 region is underlined, and those that are within the enhancer regions and mutated are labeled in red and numbered. Mutations in these elements are indicated below, with the mutated sequences underlined and labeled in red. The known silencer element rasISS1 is boxed in navy and its putative hnRNP A1 binding motif is highlighted (Guil et al., 2003b). **(B)** Immunoblot showing the expression of endogenous hnRNP's H and F and His-tagged ectopic proteins. GAPDH was used as the loading control. RT-PCR analyses showing the splicing changes of wild type and mutant minigene reporters after transfection of control (top), hnRNP H (middle), and hnRNP F (bottom) expression plasmids (wt, wild type; mt, mutant). The bar graph presents the quantification of RT-PCR results, with the mean +/- SD of three biological replicates. **(C)** RNA immunoprecipitation assayed by real-time RT-PCR of wildtype and G1G2 mutant HRAS intron 5 RNA bound by hnRNP H and hnRNP F proteins. HEK293 cells transfected with wtHRAS or mtG1G2 minigenes were immunoprecipitated using anti-epitope tag (Flag or His) antibodies, anti-endogenous protein (hnRNP H or hnRNP F) antibodies, or non-immune IgG. The immunoblot shows the recovery of hnRNP H, hnRNP F or both in the IP's. **(D)** The bar graphs (right) of real-time RT-PCR quantification of recovered RNA using two sets of primers: the I5 product targets an intronic region neighboring the mutation and the E4I4 product targets the sequence immediately upstream.

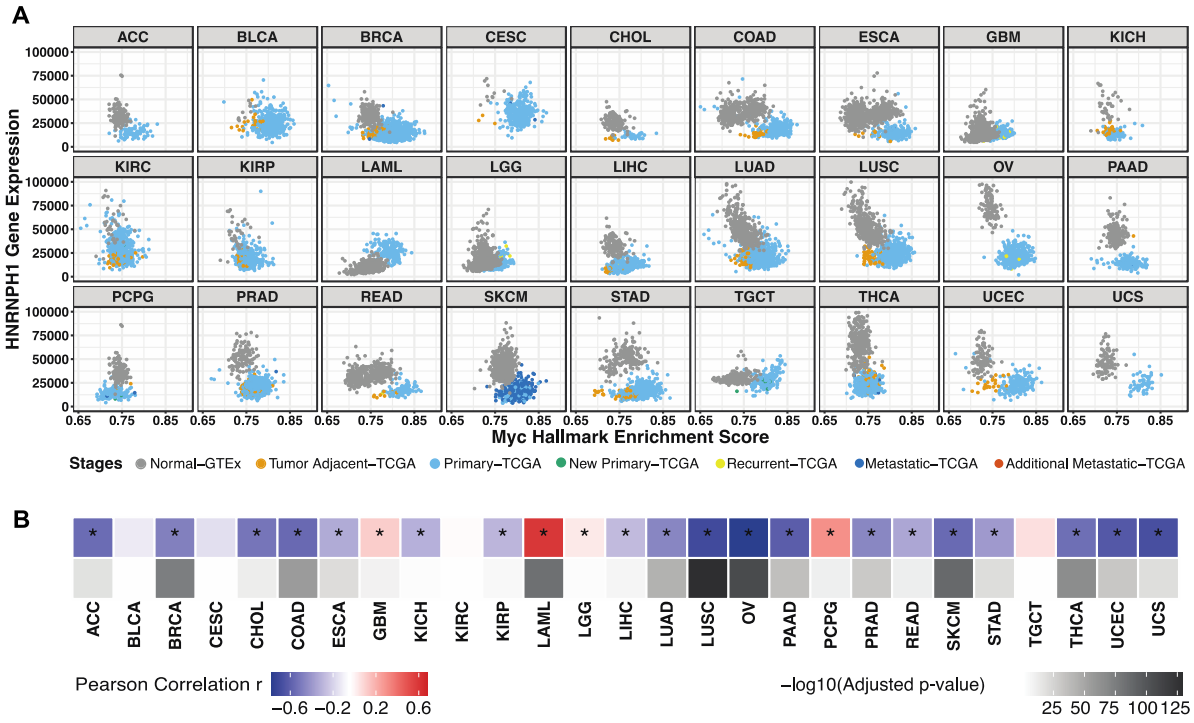

Figure 2-5. HNRNPH1 gene expression is downregulated with Myc activation across tumor types. (A) Scatterplot matrix showing the correlation of normalized HNRNPH1 versus Myc hallmark enrichment score across the disease spectrum in multiple tumor types. **(B)** Heatmap summarizing the Pearson correlation coefficient of Myc vs hnRNP H expression for each tumor type, accompanied with its adjusted p-value. *, tumor types with statistically significant adjusted p-value (< 0.05).

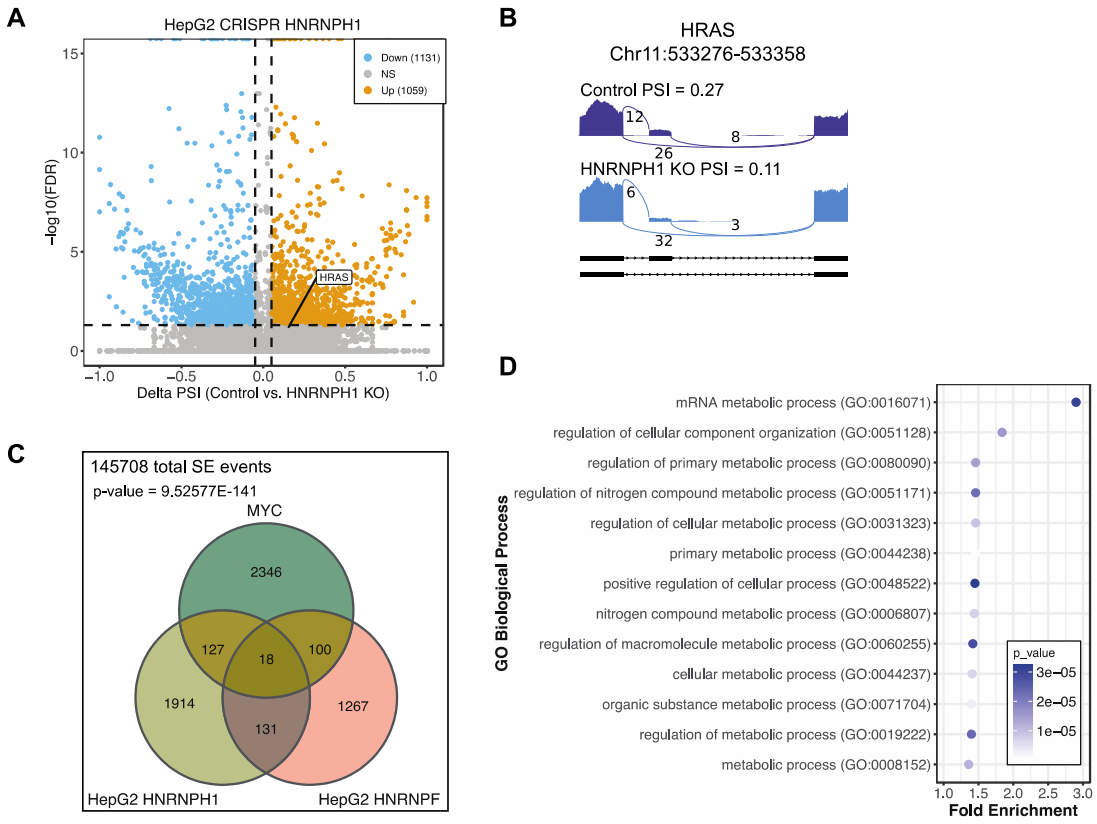

Figure 2-6. HRAS exon 5 is one of many exons controlled by both MYC and hnRNP's H and F. (A) Scatterplot showing the SE (Skipped Exon) events detected by RNA-seq from the ENCODE project in HepG2 cells after HNRNPH1 CRISPR knockout. Significant SE events were filtered by junction reads per event > 10, |deltaPSI| > 0.05, and FDR < 0.05. **(B)** Sashimi plot showing the PSI of HRAS exon 5 in control cells and after HNRNPH1 KO. **(C)** Venn diagram showing the overlapping significant SE events across three datasets: MYC on vs. MYC off, HNRNPH1 KO vs Control, and HNRNPF KO vs Control. The numbers reflect overlapping events showing changes in each comparison without considering the direction of the changes. **(D)** Gene ontology analyses using PANTHER for genes represented by SE events between MYC and HNRNPH, between MYC and HNRNPF, or between both comparisons.

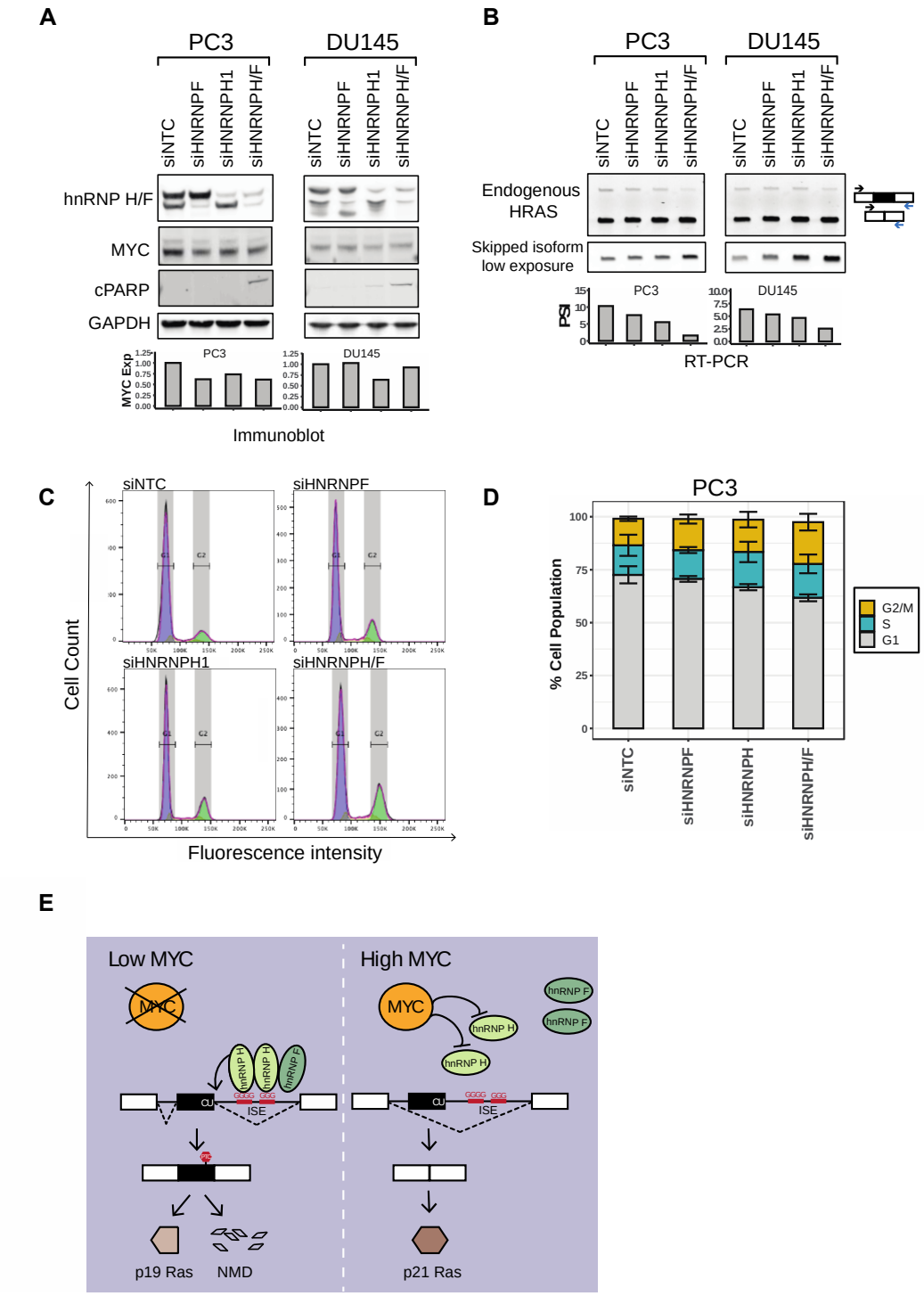

Figure 2-7. hnRNP's H and F are required for cell proliferation in prostate cancer cell lines.

Figure 2-7. Continued.

(A) The effects of hnRNP H and hnRNP F knockdown on the expression of Myc and the apoptosis marker cPARP in the prostate cancer cell lines, PC3 and DU145. Immunoblot showing the expression levels of hnRNP's H and F, MYC and cPARP in cells transfected with control, HNRNPF, HNRNPH1, or HNRNPH/F siRNAs. GAPDH was used as a loading control. The bar graphs present the quantification of MYC expression in response to each siRNA perturbation. **(B)** RT-PCR analysis showing the splicing changes of endogenous HRAS transcripts in response to each siRNA perturbation. The bar graphs show the quantification of HRAS exon 5 PSI. **(C)** Cell cycle analysis by FACS of PC3 cells transfected with control, HNRNPF, HNRNPH1, or HNRNPH/F siRNAs, and stained with propidium iodide. **(D)** Stacked bar plot showing the quantification of cells in the G1, S, and G2/M phases. Bars present the mean +/- SD of three biological replicates**. (E)** Model of hnRNP H regulation of HRAS exon alternative splicing under low MYC versus high MYC conditions

DISCUSSION

We found that Myc transformation influences the splicing of HRAS protooncogene transcripts to decrease exon 5 inclusion and allow greater production of the full-length isoform. We identified clusters of positive and negative splicing regulatory elements in the sequence encompassing exon 5, including splicing enhancers in the intron downstream. We showed that these G-run enhancers bind to the splicing regulators hnRNP H and hnRNP F and are required for activating splicing of exon 5. HnRNP H1 expression was found to be anticorrelated with Myc score across many tumor types including lung, breast, and prostate, consistent with the repression of exon 5 by Myc. HnRNP F showed the opposite correlation. We found that HRAS exon 5 is one of a group of Myc regulated exons that are also regulated by hnRNP's H and F, which are essential for the growth of prostate cancer cells. These results are summarized in the model shown in Figure 7E, where the activation of MYC downregulates hnRNP H, leading to decreased HRAS exon 5 splicing and increased expression of the full-length p21 HRAS oncoprotein.

Human HRAS exon 5 was originally named IDX and shown to carry an in-frame premature termination codon (PTC) to create a truncated protein (Cohen et al., 1989). Human mutations that reduce IDX inclusion were shown to increase the activity of full length p21 HRAS (Cohen et al., 1989; Cohen and Levinson, 1988). HRAS exon 5 containing transcripts were later shown to undergo nonsense mediated mRNA decay (NMD) in the cytosol, with the undegraded portion found largely in the nucleus (Barbier et al., 2007). We also found that cycloheximide treatment does increase this isoform consistent with its loss to NMD (data not shown). Thus, one role of the exon seems to be the modulation of p21 HRAS levels and activity. However, transcripts that escape from NMD are potentially translated into the C-terminal truncated p19 HRAS protein, reported to have a distinct function from the full length p21 HRAS protein (Guil et al., 2003a; Camats et al., 2009). Exon 5 and its flanking introns are conserved in mammals, although the

mouse gene does not contain the stop codon with the exon 5 included isoform predicted to terminate in exon 6. The C-terminal peptides of the truncated isoforms in the human and mouse are quite similar, supporting the idea that p19 is a functional protein variant of HRAS.

We have not detected this p19 isoform by immunoblot in our system. However, p19 was detected in the nucleus and cytoplasm of HeLa cells using an antibody specifically targeting its divergent C-terminus (Guil et al., 2003a). This p19 form failed to bind to the known p21 interactors Raf1 and Rin1(Camats et al., 2009; Guil et al., 2003a). When overexpressed in different settings, p19 has been found to bind a variety of other proteins including p73, MDM2, Neuron-Specific Enolase, and RACK, and to have varying effects on cell growth and physiology (Camats et al., 2009; García-Cruz et al., 2015; Guil et al., 2003a; Jang et al., 2010; Jeong et al., 2006). These findings indicate that p19 likely serves a separate role from p21 and that modulation of HRAS splicing will alter p19 function in addition to changing p21 activity.

Earlier studies identified a regulatory element called rasISS1 in the intron downstream of HRAS exon 5 that acted to silence exon 5 splicing (Guil et al., 2003b). This inhibition was also observed in an in vitro splicing system, where it required the protein hnRNP A1, and was counteracted by the SR proteins, SRSF2 (SC35) and SRSF5 (SRp40). The RNA binding proteins FUS/TLS and hnRNP H, and the RNA helicase p68 (DDX5) were also found to bind rasISS1, and depletion of p68 in vivo led to increased exon 5 splicing, indicating that it might counteract splicing repression by rasISS1. The rasISS1 was predicted to form a basepaired stem with exon 5 that may inhibit its splicing, and p68 was shown to unwind the exon 5 – rasISS1 stem in vitro. Knockdown of FUS/TLS or hnRNP H downregulated the abundance of p19 HRAS protein in vivo (Camats et al., 2008). These findings are consistent with our analyses and indicate that HRAS exon 5 is modulated by a combination of positive and negative acting factors as seen with most alternative exons.

The ASO tiling approach allowed us to map regulatory elements more comprehensively across the exon 5 region. ASO's I5-1 and I5-2 target the inhibitory sequence of rasISS1 element and the activation of splicing by I5-1 indicates the presence of a silencer. The limited effect of ASO I5-2 may come from the ASO not being able to disrupt the proposed secondary structure. Our analysis identified enhancer elements near the 3' end of the ISS and downstream that contain G-runs, and which bind hnRNP's H and F. We find that these proteins strongly activate splicing despite the presence of the rasISS1, and this activation requires the G-runs in the I5-3 region. The ASO tiling also indicated the presence of multiple exonic enhancers within exon 5. Some of these may recruit SRSF2 and SRSF5 whose activity on the exon was previously reported (Guil et al., 2003b). We confirmed SRSF5 as splicing activator through transient overexpression (data not shown). The ASO tiling can be refined and in future work we can more precisely delineate the cis-elements using overlapping oligos to identify those that most strongly shift exon 5 splicing and the expression of the p21 HRAS oncoprotein. These can then be tested for effects on tumor growth. Other studies examining the programs of splicing regulation in cancer have also identified HRAS as a Myc dependent exon. Several SR proteins are altered sin expression in response to Myc (Anczuków et al., 2012, 2015). This group recently reported that some of these SR proteins, particularly SRSF2, act to repress HRAS cassette exon splicing in Myc-active tumors (Urbanski et al., 2021).

HnRNP H has been connected to other aspects of Ras signaling. Studies of the ARaf kinase found that splicing to create its full length isoform required hnRNP H (Rauch et al., 2010, 2011). This isoform inhibits apoptosis in tumor cells through interaction with the MST2 kinase. The short ARaf isoform expressed in low hnRNP H conditions can act as a dominant negative protein to suppress Ras activation and oncogenic transformation. These studies found that high Myc correlated with high hnRNP H expression in HeLa and several other tumor cell lines, the opposite of the correlation we observe in most tumors in the TGCA database. It will be interesting

to investigate whether these different results arise for differences between cell lines and primary tumors, differences between tumor types, or some other difference between the systems. Another earlier study found that hnRNP F affected cell proliferation through interactions with mTOR and the S6 Kinase 2 pathway (Goh et al., 2010). We found that both hnRNP H and F are required for prostate cancer cell growth. It will be interesting to assess the signaling pathways involved in these effects on cell proliferation control, and whether the Ras-MEK-ERK or Ras-PI3K-AKT pathways are involved.

There are several findings within our data that remain unexplained. One question is the apparent upregulation of hnRNP F by Myc. Since hnRNP F also seems to stimulate HRAS exon 5, one would expect it to go down with increased Myc. It is possible that other factors, of the many affecting exon 5, are counteracting the effect of hnRNP F. It should also be noted that Myc correlations with gene expression are only measuring RNA and it is possible the proteins encoded by these mRNAs are behaving differently. It is also possible that modifications of the hnRNP F protein could alter its activity. Another question regards the requirement for hnRNP H in the growth of Myc transformed cells. Given that reduced splicing of HRAS exon 5 resulting from reduced hnRNP H is conducive to growth, it appears that some level of hnRNP H is still required. What mRNA isoforms are responsible this hnRNP H dependence will be interesting to investigate.

MATERIALS AND METHODS

Construction of minigene reporters and cDNA expression vectors

The HRAS minigene reporter was constructed by PCR amplifying a ~900bp region spanning exon 4 to exon 6 from human genomic DNA isolated from HEK293 cells using Phusion high-fidelity DNA polymerase (NEB). The PCR fragment were inserted into pcDNA3.1(+) vector through restriction sites BamHI/EcoRI. G-run point mutations were introduced using site-directed mutagenesis. Construction of the C-terminal 6xHistidine tagged hnRNP H and hnRNP F cDNA constructs, and siRNA-resistant hnRNP H construct was described previously (Nazim et al., 2016). Mutations to create the siRNA-resistant hnRNP F construct were introduced by site-directed mutagenesis. hnRNP H coding sequences were subcloned into C-terminal p3xFLAG CMV vector. All the constructs were confirmed by sequencing.

Cell culture, plasmid and siRNA transfections

HEK293, HepG2, PC3, DU145 cells were maintained in DMEM, EMEM, F12K, RPMI1640, respectively, all supplemented with 10% fetal bovine serum at 37C in 5% CO2. Cells were plated the day before transfection in a 6-well culture plate. Plasmids (3ug per well) were transfected into HEK293 cells using lipofectamine 2000 (Invitrogen), according to manufacturer's instructions. Cells were harvested 48hrs post transfection for RT-PCR and immunoblotting analyses. siRNAs were transfected into HEK293 (10nM) or HepG2, PC3 and DU145 (20nM) using lipofectamine RNAiMAX (Invitrogen), according to manufacturer's instructions. 24hrs after transfection, siRNAs were transfected a second time. 72hrs after the first siRNA transfection, cells were harvested for RT-PCR and immunoblotting analyses. A list of siRNA sequences is presented in the supplemental table S3.

ASO transfections

ASOs were provided by IONIS Pharmaceuticals and have uniform phosphorothioate backbone chemistry with modified 2' methoxyethyl sugars. HEK293 cells were plated the day before transfection in 12-well culture plates. ASOs (100nM) were transfected into cells using lipofectamine 2000 (Invitrogen). Cells were harvested 24 hrs after transfection for RT-PCR analyses. A list of ASO sequences is presented in the supplemental table S1.

RNA isolation, RT-PCR, RT-qPCR

Total RNA was isolated using TRIzol Reagent (Life Technologies), followed by DNase I treatment (Invitrogen Ambion), and then extracted again with acid phenol chloroform (Invitrogen Ambion). 500ng – 1ug of total RNA was subjected to cDNA synthesis using SuperScript III reverse transcriptase (Thermo Fisher Scientific), primed by random hexamers. PCR was conducted using GoTaq, 18 cycles for minigene reporters and 24 cycles for endogenous RNAs. PCR products were run on 5% PAGE, stained with SYBRGold (Thermo Fisher Scientific) (unless otherwise specified), and scanned with a Typhoon imager (GE healthcare). PCR band intensities were quantified using ImageJ. RT-qPCR was conducted with SensiFast SYBR lo-ROX Kit (Bioline) on a Quant Studio 6 Flex real-time PCR system (Thermo Fisher Scientific). RT-PCR primers are listed in the supplemental table S2.

Immunoblotting

Proteins were extracted using RIPA lysis buffer (Thermo Fisher Scientific) supplemented with Benzonase Nuclease (Sigma Aldrich) and protease inhibitor cocktail (Sigma Aldrich). Protein concentrations were quantified by BCA protein assay kit (Pierce Biotechnology). Cell lysates were prepared with SDS loading buffer (50 mM Tris·Cl, pH 6.8, 0.05% bromophenol blue, 10% glycerol, 2% SDS, and 0.1 M DTT) and proteins were separated on SDS-PAGE gels and transferred to PVDF membranes. Antibodies used in this study were listed in the supplemental table S4.

Western blot images were scanned on a Typhoon imager (GE healthcare) and band intensities were quantified using ImageJ.

RNA Immunoprecipitation

Antibodies targeting proteins of interest or IgG isotype control (5ug) were incubated with 20ul of Dynabeads Protein G (Thermo Fisher Scientific) in buffer WB150 (20 mM HEPES-KOH pH 7.5, 150 mM NaCl, 0.1% Triton-X100) at 4C for 2 hrs. HEK293 cells were harvested and sonicated in cold lysis buffer (20 mM HEPES-KOH pH 7.5, 150 mM NaCl, 0.5% Triton X-100, 0.05% SDS, 1 mM EDTA, 0.5 mM DTT, 1x protease inhibitors, 100U/ml RNaseOUT). After centrifugation at 20,000g for 10min at 4C, the supernatant was incubated with antibody conjugated beads at 4C for 3 hrs. The beads were washed five times with buffer WB150 and RNA was extracted with TRIzol.

Cell cycle analysis

PC3 cells were plated the day before transfection in a 6-well culture plate. siRNA were transfected as described above. 72hrs after transfection, cells were trypsinyzed and washed with PBS. These cell suspensions were fixed with 70% ice-cold ethanol overnight at -20C. Cells were stained with propidium iodide at 75 ug/ml and incubated at 37C for 30min, followed by FACS analysis.

RNA-seq data processing, gene expression and splicing analysis for cell lines

The fastq files of HNRNPH1 and HNRNPF CRISPR RNA-seq were downloaded from ENCODE Consortium (Dunham et al., 2012) (https://www.encodeproject.org/) (K562 HNRNPF shRNA KD, ENCODE accession no. ENCSR392HSJ; K562 HNRNPH1 CRISPR KO, ENCSR354RSR; HepG2 HNRNPF CRISPR KO, ENCSR599NNK; HepG2 HNRNPH1 CRISPR KO ENCSR094HEU). The RNA-seq fastq files of Myc cell lines were downloaded from Gene Expression Omnibus (accession no. GSE141633) via fastq-dump in SRA toolkit. The raw sequencing reads were aligned to human reference genome and gene annotation (GRCh37,

GENCODE release 26) using STAR (v2.7.3a) (Dobin et al., 2013). Alternative splicing in the samples was quantified as Percent Spliced In (PSI, ψ) using rMATS-turbo (v4.1.1) (Shen et al., 2014) by using the same gene annotation as the alignment step. Skipped exon (SE) events with insufficient average coverage, $IC + SC \leq 10$, were excluded from further analysis. Significant SE events were identified using $|\Delta\psi| > 0.05 \; and \; FDR < 0.05$, calculated using PAIRADISE (Demirdjian et al., 2020) with the equal variance option. Gene Ontology analysis was done using PANTHER (Thomas et al., 2022), p-value was computed using fisher's exact test, and Benjamini-Hochberg procedure was used to control for FDR < 0.05.

RNA-seq data compilation and processing for the pan-cancer analysis

RNA-seq data of tissue samples were compiled from two public domains: TCGA (Weinstein et al., 2013), GTEx (Lonsdale et al., 2013). Tumor tissue samples were obtained from

TCGA while normal tissue was obtained from GTEx. The TCGA RNA-seq fastq files were

downloaded from GDC through their GDC Data Transfer Tool Client (Grossman et al., 2016)

while the GTEx files were downloaded from dbGAP (Mailman et al., 2007) through fastq-dump

from the SRA toolkit. The TCGA and GTEx samples were matched by tissue origin (Robinson et al., 2019). The fastq files were mapped by STAR 2.5.3a (Dobin et al., 2013). The STAR 2-pass function was enabled to improve the accuracy of the alignment. The genome annotation file was obtained from GENCODE V26 (Harrow et al., 2012) under human genome version hg19. Gene expression quantification and alternative splicing quantification were done using Cufflinks v2.2.1 (Trapnell et al., 2012) and rMATS 4.1.0 (Shen et al., 2014) respectively. Percent Spliced In (PSI) ratio was used as a statistic to quantify alternative splicing events from the rMATS output. Splicing events were filtered for splice junction reads >= 10 (otherwise the sample's PSI value is marked with 'NA'), PSI range (difference between maxPSI and minPSI) > 5%, and for mean skipping or

inclusion values over 5% across the entire dataset, with fewer than 50% of samples missing values.

Pathway Enrichment-Guided Activity Study of Alternative Splicing for the pan-cancer analysis

The PEGASAS pipeline (Phillips et al., 2020) was used to calculate pathway activity scores and identify alternative splicing events that correlate with the signaling pathway of interest. The Myc Targets V2 hallmark gene signature list obtained from The Molecular Signatures Database (MSigDB) (Liberzon et al., 2011) was used for pathway activity score calculation. Myc pathway scores were calculated using cufflinks expression outputs. Alternative splicing events (SE) correlated with the Myc signaling pathway were identified through the PEGASAS correlation step. A correlation permutation test was used to acquire empirical p-values, by permuting the pathway scores 5,000 times. Highly correlated events were defined as events with an empirical p-value < 0.0002 and with a |Pearson correlation coefficient| > 0.3.

Gene expression values for RBP correlation

FeatureCounts v2.0.1 was used to quantify reads from aligned BAM files (Liao et al., 2014). DESeq2 v1.26.0 was used to normalize gene expression (Love et al., 2014). The normalized gene expression values for HNRNPH1 and HNRNPF along with 218 other splicing factors are correlated with Myc scores (Han et al., 2013).

REFERENCES

Anczuków, O., Akerman, M., Cléry, A., Wu, J., Shen, C., Shirole, N.H., Raimer, A., Sun, S., Jensen, M.A., Hua, Y., Allain, F.H.-T., Krainer, A.R., 2015. SRSF1-Regulated Alternative Splicing in Breast Cancer. Molecular Cell 60, 105–117. https://doi.org/10.1016/j.molcel.2015.09.005

Anczuków, O., Rosenberg, A.Z., Akerman, M., Das, S., Zhan, L., Karni, R., Muthuswamy, S.K., Krainer, A.R., 2012. The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. Nat Struct Mol Biol 19, 220–228. https://doi.org/10.1038/nsmb.2207

Baluapuri, A., Hofstetter, J., Dudvarski Stankovic, N., Endres, T., Bhandare, P., Vos, S.M., Adhikari, B., Schwarz, J.D., Narain, A., Vogt, M., Wang, S.-Y., Düster, R., Jung, L.A., Vanselow, J.T., Wiegering, A., Geyer, M., Maric, H.M., Gallant, P., Walz, S., Schlosser, A., Cramer, P., Eilers, M., Wolf, E., 2019. MYC Recruits SPT5 to RNA Polymerase II to Promote Processive Transcription Elongation. Molecular Cell 74, 674-687.e11. https://doi.org/10.1016/j.molcel.2019.02.031

Barbier, J., Dutertre, M., Bittencourt, D., Sanchez, G., Gratadou, L., de la Grange, P., Auboeuf, D., 2007. Regulation of H-ras Splice Variant Expression by Cross Talk between the p53 and Nonsense-Mediated mRNA Decay Pathways. Mol Cell Biol 27, 7315–7333. https://doi.org/10.1128/MCB.00272-07

Black, D.L., 2003. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 72, 291–336. https://doi.org/10.1146/annurev.biochem.72.121801.161720

Braun, S., Enculescu, M., Setty, S.T., Cortés-López, M., de Almeida, B.P., Sutandy, F.X.R., Schulz, L., Busch, A., Seiler, M., Ebersberger, S., Barbosa-Morais, N.L., Legewie, S., König, J., Zarnack, K., 2018. Decoding a cancer-relevant splicing decision in the RON proto-oncogene using high-throughput mutagenesis. Nat Commun 9, 3315. https://doi.org/10.1038/s41467-018-05748-7

Caggiano, C., Pieraccioli, M., Panzeri, V., Sette, C., Bielli, P., 2019. c-MYC empowers transcription and productive splicing of the oncogenic splicing factor Sam68 in cancer. Nucleic Acids Research 47, 6160–6171. https://doi.org/10.1093/nar/gkz344

Camats, M., Guil, S., Kokolo, M., Bach-Elias, M., 2008. P68 RNA Helicase (DDX5) Alters Activity of Cis- and Trans-Acting Factors of the Alternative Splicing of H-Ras. PLoS ONE 3, e2926. https://doi.org/10.1371/journal.pone.0002926

Camats, M., Kokolo, M., Heesom, K.J., Ladomery, M., Bach-Elias, M., 2009. P19 H-Ras Induces G1/S Phase Delay Maintaining Cells in a Reversible Quiescence State. PLoS ONE 4, e8513. https://doi.org/10.1371/journal.pone.0008513

Caputi, M., Zahler, A.M., 2001. Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. J Biol Chem 276, 43850–43859. https://doi.org/10.1074/jbc.M102861200

Carlo, T., Sterner, D.A., Berget, S.M., 1996. An intron splicing enhancer containing a G-rich repeat facilitates inclusion of a vertebrate micro-exon. RNA 2, 342–353.

Chou, M.-Y., Rooke, N., Turck, C.W., Black, D.L., 1999. hnRNP H Is a Component of a Splicing Enhancer Complex That Activates a c- *src* Alternative Exon in Neuronal Cells. Mol Cell Biol 19, 69–77. https://doi.org/10.1128/MCB.19.1.69

Cohen, J.B., Broz, S.D., Levinson, A.D., 1989. Expression of the H-ras proto-oncogene is controlled by alternative splicing. Cell 58, 461–472. https://doi.org/10.1016/0092-8674(89)90427-3

Cohen, J.B., Levinson, A.D., 1988. A point mutation in the last intron responsible for increased expression and transforming activity of the c-Ha-ras oncogene. Nature 334, 119–124. https://doi.org/10.1038/334119a0

Demirdjian, L., Xu, Y., Bahrami-Samani, E., Pan, Y., Stein, S., Xie, Z., Park, E., Wu, Y.N., Xing, Y., 2020. Detecting Allele-Specific Alternative Splicing from Population-Scale RNA-Seq Data. The American Journal of Human Genetics 107, 461–472. https://doi.org/10.1016/j.ajhg.2020.07.005

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., et al., 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74. https://doi.org/10.1038/nature11247

Dvinge, H., Kim, E., Abdel-Wahab, O., Bradley, R.K., 2016. RNA splicing factors as oncoproteins and tumour suppressors. Nat Rev Cancer 16, 413–430. https://doi.org/10.1038/nrc.2016.51

Escobar-Hoyos, L., Knorr, K., Abdel-Wahab, O., 2019. Aberrant RNA Splicing in Cancer. Annu. Rev. Cancer Biol. 3, 167–185. https://doi.org/10.1146/annurev-cancerbio-030617-050407

Fei, T., Chen, Y., Xiao, T., Li, W., Cato, L., Zhang, P., Cotter, M.B., Bowden, M., Lis, R.T., Zhao, S.G., Wu, Q., Feng, F.Y., Loda, M., He, H.H., Liu, X.S., Brown, M., 2017. Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing. Proc. Natl. Acad. Sci. U.S.A. 114. https://doi.org/10.1073/pnas.1617467114

Fu, X.-D., Ares, M., 2014. Context-dependent control of alternative splicing by RNA-binding proteins. Nat Rev Genet 15, 689–701. https://doi.org/10.1038/nrg3778

García-Cruz, R., Camats, M., Calin, G.A., Liu, C.-G., Volinia, S., Taccioli, C., Croce, C.M., Bach-Elias, M., 2015. The role of p19 and p21 H-Ras proteins and mutants in miRNA expression in cancer and a Costello syndrome cell model. BMC Med Genet 16, 46. https://doi.org/10.1186/s12881-015-0184-z

Garneau, D., Revil, T., Fisette, J.-F., Chabot, B., 2005. Heterogeneous Nuclear Ribonucleoprotein F/H Proteins Modulate the Alternative Splicing of the Apoptotic Mediator Bcl-x. Journal of Biological Chemistry 280, 22641–22650. https://doi.org/10.1074/jbc.M501070200

Gautrey, H., Jackson, C., Dittrich, A.-L., Browell, D., Lennard, T., Tyson-Capper, A., 2015. SRSF3 and hnRNP H1 regulate a splicing hotspot of HER2 in breast cancer cells. RNA Biol 12, 1139–1151. https://doi.org/10.1080/15476286.2015.1076610

Goh, E.T.H., Pardo, O.E., Michael, N., Niewiarowski, A., Totty, N., Volkova, D., Tsaneva, I.R., Seckl, M.J., Gout, I., 2010. Involvement of heterogeneous ribonucleoprotein F in the

regulation of cell proliferation via the mammalian target of rapamycin/S6 kinase 2 pathway. J Biol Chem 285, 17065–17076. https://doi.org/10.1074/jbc.M109.078782

Grossman, R.L., Heath, A.P., Ferretti, V., Varmus, H.E., Lowy, D.R., Kibbe, W.A., Staudt, L.M., 2016. Toward a Shared Vision for Cancer Genomic Data. New England Journal of Medicine 375, 1109–1112. https://doi.org/10.1056/NEJMp1607591

Guil, S., de La Iglesia, N., Fernández-Larrea, J., Cifuentes, D., Ferrer, J.C., Guinovart, J.J., Bach-Elias, M., 2003a. Alternative splicing of the human proto-oncogene c-H-ras renders a new Ras family protein that trafficks to cytoplasm and nucleus. Cancer Res 63, 5178–5187.

Guil, S., Gattoni, R., Carrascal, M., Abián, J., Stévenin, J., Bach-Elias, M., 2003b. Roles of hnRNP A1, SR Proteins, and p68 Helicase in c-H- *ras* Alternative Splicing Regulation. Mol Cell Biol 23, 2927–2941. https://doi.org/10.1128/MCB.23.8.2927-2941.2003

Han, H., Irimia, M., Ross, P.J., Sung, H.-K., Alipanahi, B., David, L., Golipour, A., Gabut, M., Michael, I.P., Nachman, E.N., Wang, E., Trcka, D., Thompson, T., O'Hanlon, D., Slobodeniuc, V., Barbosa-Morais, N.L., Burge, C.B., Moffat, J., Frey, B.J., Nagy, A., Ellis, J., Wrana, J.L., Blencowe, B.J., 2013. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature 498, 241–245. https://doi.org/10.1038/nature12270

Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., Barnes, I., Bignell, A., Boychenko, V., Hunt, T., Kay, M., Mukherjee, G., Rajan, J., Despacio-Reyes, G., Saunders, G., Steward, C., Harte, R., Lin, M., Howald, C., Tanzer, A., Derrien, T., Chrast, J., Walters, N., Balasubramanian, S., Pei, B., Tress, M., Rodriguez, J.M., Ezkurdia, I., Baren, J. van, Brent, M., Haussler, D., Kellis, M., Valencia, A., Reymond, A., Gerstein, M., Guigó, R., Hubbard, T.J., 2012. GENCODE: The reference human genome annotation for The ENCODE Project. Genome Res. 22, 1760–1774. https://doi.org/10.1101/gr.135350.111

Herviou, P., Le Bras, M., Dumas, L., Hieblot, C., Gilhodes, J., Cioci, G., Hugnot, J.-P., Ameadan, A., Guillonneau, F., Dassi, E., Cammas, A., Millevoi, S., 2020. hnRNP H/F drive RNA G-quadruplex-mediated translation linked to genomic instability and therapy resistance in glioblastoma. Nat Commun 11, 2661. https://doi.org/10.1038/s41467-020-16168-x

Ho, J.S., Di Tullio, F., Schwarz, M., Low, D., Incarnato, D., Gay, F., Tabaglio, T., Zhang, J., Wollmann, H., Chen, L., An, O., Chan, T.H.M., Hall Hickman, A., Zheng, S., Roudko, V., Chen, S., Karz, A., Ahmed, M., He, H.H., Greenbaum, B.D., Oliviero, S., Serresi, M., Gargiulo, G., Mann, K.M., Hernando, E., Mulholland, D., Marazzi, I., Wee, D.K.B., Guccione, E., 2021. HNRNPM controls circRNA biogenesis and splicing fidelity to sustain cancer cell fitness. eLife 10, e59654. https://doi.org/10.7554/eLife.59654

Honoré, B., 2000. The hnRNP 2H9 gene, which is involved in the splicing reaction, is a multiply spliced gene. Biochim Biophys Acta 1492, 108–119. https://doi.org/10.1016/s0167-4781(00)00092-0

Honoré, B., Rasmussen, H.H., Vorum, H., Dejgaard, K., Liu, X., Gromov, P., Madsen, P., Gesser, B., Tommerup, N., Celis, J.E., 1995. Heterogeneous nuclear ribonucleoproteins H, H', and F are members of a ubiquitously expressed subfamily of related but distinct proteins encoded by genes mapping to different chromosomes. J Biol Chem 270, 28780–28789. https://doi.org/10.1074/jbc.270.48.28780

Hsu, T.Y.-T., Simon, L.M., Neill, N.J., Marcotte, R., Sayad, A., Bland, C.S., Echeverria, G.V., Sun, T., Kurley, S.J., Tyagi, S., Karlin, K.L., Dominguez-Vidaña, R., Hartman, J.D., Renwick,

A., Scorsone, K., Bernardi, R.J., Skinner, S.O., Jain, A., Orellana, M., Lagisetti, C., Golding, I., Jung, S.Y., Neilson, J.R., Zhang, X.H.-F., Cooper, T.A., Webb, T.R., Neel, B.G., Shaw, C.A., Westbrook, T.F., 2015. The spliceosome is a therapeutic vulnerability in MYC-driven cancer. Nature 525, 384–388. https://doi.org/10.1038/nature14985

Huelga, S.C., Vu, A.Q., Arnold, J.D., Liang, T.Y., Liu, P.P., Yan, B.Y., Donohue, J.P., Shiue, L., Hoon, S., Brenner, S., Ares, M., Yeo, G.W., 2012. Integrative Genome-wide Analysis Reveals Cooperative Regulation of Alternative Splicing by hnRNP Proteins. Cell Reports 1, 167–178. https://doi.org/10.1016/j.celrep.2012.02.001

Jang, S.-M., Kim, J.-W., Kim, C.-H., Kim, D., Rhee, S., Choi, K.-H., 2010. p19ras Represses proliferation of non-small cell lung cancer possibly through interaction with Neuron-Specific Enolase (NSE). Cancer Letters 289, 91–98. https://doi.org/10.1016/j.canlet.2009.08.005

Jeong, M.-H., Bae, J., Kim, W.-H., Yoo, S.-M., Kim, J.-W., Song, P.I., Choi, K.-H., 2006. p19 Interacts with and Activates p73 by Involving the MDM2 Protein. Journal of Biological Chemistry 281, 8707–8715. https://doi.org/10.1074/jbc.M513853200

Koh, C.M., Bezzi, M., Low, D.H.P., Ang, W.X., Teo, S.X., Gay, F.P.H., Al-Haddawi, M., Tan, S.Y., Osato, M., Sabò, A., Amati, B., Wee, K.B., Guccione, E., 2015. MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. Nature 523, 96–100. https://doi.org/10.1038/nature14351

Lee, Y., Rio, D.C., 2015. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu Rev Biochem 84, 291–323. https://doi.org/10.1146/annurev-biochem-060614-034316

LeFave, C.V., Squatrito, M., Vorlova, S., Rocco, G.L., Brennan, C.W., Holland, E.C., Pan, Y.-X., Cartegni, L., 2011. Splicing factor hnRNPH drives an oncogenic splicing switch in gliomas: hnRNPH-dependent splicing switch in gliomas. The EMBO Journal 30, 4084–4097. https://doi.org/10.1038/emboj.2011.259

Li, X., Qian, X., Peng, L.-X., Jiang, Y., Hawke, D.H., Zheng, Y., Xia, Y., Lee, J.-H., Cote, G., Wang, H., Wang, L., Qian, C.-N., Lu, Z., 2016. A splicing switch from ketohexokinase-C to ketohexokinase-A drives hepatocellular carcinoma formation. Nat Cell Biol 18, 561–571. https://doi.org/10.1038/ncb3338

Liao, Y., Smyth, G.K., Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. https://doi.org/10.1093/bioinformatics/btt656

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., Tamayo, P., 2015. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Systems 1, 417–425. https://doi.org/10.1016/j.cels.2015.12.004

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

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., et al., 2013. The Genotype-Tissue Expression (GTEx) project. Nat Genet 45, 580–585. https://doi.org/10.1038/ng.2653

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

Mahé, D., Mähl, P., Gattoni, R., Fischer, N., Mattei, M.G., Stévenin, J., Fuchs, J.P., 1997. Cloning of human 2H9 heterogeneous nuclear ribonucleoproteins. Relation with splicing and early heat shock-induced splicing arrest. J Biol Chem 272, 1827–1836. https://doi.org/10.1074/jbc.272.3.1827

Mailman, M.D., Feolo, M., Jin, Y., Kimura, M., Tryka, K., Bagoutdinov, R., Hao, L., Kiang, A., Paschall, J., Phan, L., Popova, N., Pretel, S., Ziyabari, L., Lee, M., Shao, Y., Wang, Z.Y., Sirotkin, K., Ward, M., Kholodov, M., Zbicz, K., Beck, J., Kimelman, M., Shevelev, S., Preuss, D., Yaschenko, E., Graeff, A., Ostell, J., Sherry, S.T., 2007. The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 39, 1181–1186. https://doi.org/10.1038/ng1007-1181

Martinez-Contreras, R., Fisette, J.-F., Nasim, F.H., Madden, R., Cordeau, M., Chabot, B., 2006. Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. PLoS Biol 4, e21. https://doi.org/10.1371/journal.pbio.0040021

Matunis, M.J., Xing, J., Dreyfuss, G., 1994. The hnRNP F protein: unique primary structure, nucleic acid-binding properties, and subcellular localization. Nucleic Acids Res 22, 1059–1067.

Mc, W., Cl, W., R, L., 2009. The spliceosome: design principles of a dynamic RNP machine. Cell 136. https://doi.org/10.1016/j.cell.2009.02.009

Min, H., Chan, R.C., Black, D.L., 1995. The generally expressed hnRNP F is involved in a neural-specific pre-mRNA splicing event. Genes Dev 9, 2659–2671. https://doi.org/10.1101/gad.9.21.2659

Modafferi, E.F., Black, D.L., 1999. Combinatorial control of a neuron-specific exon. RNA 5, 687–706. https://doi.org/10.1017/s1355838299990155

Modafferi, E.F., Black, D.L., 1997. A complex intronic splicing enhancer from the c-src pre-mRNA activates inclusion of a heterologous exon. Mol Cell Biol 17, 6537–6545. https://doi.org/10.1128/MCB.17.11.6537

Nazim, M., Masuda, A., Rahman, M.A., Nasrin, F., Takeda, J., Ohe, K., Ohkawara, B., Ito, M., Ohno, K., 2016. Competitive regulation of alternative splicing and alternative polyadenylation by hnRNP H and CstF64 determines acetylcholinesterase isoforms. Nucleic Acids Res gkw823. https://doi.org/10.1093/nar/gkw823

Pálfy, G., Menyhárd, D.K., Perczel, A., 2020. Dynamically encoded reactivity of Ras enzymes: opening new frontiers for drug discovery. Cancer Metastasis Rev 39, 1075–1089. https://doi.org/10.1007/s10555-020-09917-3

Pan, Y., Kadash-Edmondson, K.E., Wang, R., Phillips, J., Liu, S., Ribas, A., Aplenc, R., Witte, O.N., Xing, Y., 2021. RNA Dysregulation: An Expanding Source of Cancer Immunotherapy Targets. Trends Pharmacol Sci 42, 268–282. https://doi.org/10.1016/j.tips.2021.01.006

Paz, I., Kosti, I., Ares, M., Cline, M., Mandel-Gutfreund, Y., 2014. RBPmap: a web server for mapping binding sites of RNA-binding proteins. Nucleic Acids Res 42, W361–W367. https://doi.org/10.1093/nar/gku406

Phillips, J.W., Pan, Y., Tsai, B.L., Xie, Z., Demirdjian, L., Xiao, W., Yang, H.T., Zhang, Y., Lin, C.H., Cheng, D., Hu, Q., Liu, S., Black, D.L., Witte, O.N., Xing, Y., 2020. Pathway-guided analysis identifies Myc-dependent alternative pre-mRNA splicing in aggressive prostate cancers. Proc. Natl. Acad. Sci. U.S.A. 117, 5269–5279. https://doi.org/10.1073/pnas.1915975117

Piva, F., Giulietti, M., Burini, A.B., Principato, G., 2012. SpliceAid 2: A database of human splicing factors expression data and RNA target motifs. Human Mutation 33, 81–85. https://doi.org/10.1002/humu.21609

Rauch, J., Moran-Jones, K., Albrecht, V., Schwarzl, T., Hunter, K., Gires, O., Kolch, W., 2011. c-Myc Regulates RNA Splicing of the A-Raf Kinase and Its Activation of the ERK Pathway. Cancer Research 71, 4664–4674. https://doi.org/10.1158/0008-5472.CAN-10-4447

Rauch, J., O'Neill, E., Mack, B., Matthias, C., Munz, M., Kolch, W., Gires, O., 2010. Heterogeneous Nuclear Ribonucleoprotein H Blocks MST2-Mediated Apoptosis in Cancer Cells by Regulating *a-raf* Transcription. Cancer Research 70, 1679–1688. https://doi.org/10.1158/0008-5472.CAN-09-2740

Rinaldi, C., Wood, M.J.A., 2018. Antisense oligonucleotides: the next frontier for treatment of neurological disorders. Nat Rev Neurol 14, 9–21. https://doi.org/10.1038/nrneurol.2017.148

Robinson, J.L., Feizi, A., Uhlén, M., Nielsen, J., 2019. A Systematic Investigation of the Malignant Functions and Diagnostic Potential of the Cancer Secretome. Cell Reports 26, 2622-2635.e5. https://doi.org/10.1016/j.celrep.2019.02.025

Schaub, M.C., Lopez, S.R., Caputi, M., 2007. Members of the heterogeneous nuclear ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by promoting formation of ATP-dependent spliceosomal complexes. J Biol Chem 282, 13617–13626. https://doi.org/10.1074/jbc.M700774200

Shen, S., Park, J.W., Lu, Z., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., Xing, Y., 2014. rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl. Acad. Sci. U.S.A. 111. https://doi.org/10.1073/pnas.1419161111

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

Smith, C.C., Selitsky, S.R., Chai, S., Armistead, P.M., Vincent, B.G., Serody, J.S., 2019. Alternative tumour-specific antigens. Nat Rev Cancer 19, 465–478. https://doi.org/10.1038/s41568-019-0162-4

Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.-P., Mi, H., 2022. PANTHER: Making genome-scale phylogenetics accessible to all. Protein Science 31, 8–22. https://doi.org/10.1002/pro.4218

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7, 562–578. https://doi.org/10.1038/nprot.2012.016

Tyson-Capper, A., Gautrey, H., 2018. Regulation of Mcl-1 alternative splicing by hnRNP F, H1 and K in breast cancer cells. RNA Biol 15, 1448–1457. https://doi.org/10.1080/15476286.2018.1551692

Urbanski, L., Brugiolo, M., Park, S., Angarola, B., Leclair, N.K., Palmer, P., Sahu, S.K., Anczuków, O., 2021. MYC regulates a pan-cancer network of co-expressed oncogenic splicing factors (preprint). Cancer Biology. https://doi.org/10.1101/2021.11.24.469558

Urbanski, L.M., Leclair, N., Anczuków, O., 2018. Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics. Wiley Interdiscip Rev RNA 9, e1476. https://doi.org/10.1002/wrna.1476

Vo, T., Brownmiller, T., Hall, K., Jones, T.L., Choudhari, S., Grammatikakis, I., Ludwig, K.R., Caplen, N.J., 2022. HNRNPH1 destabilizes the G-quadruplex structures formed by G-rich RNA sequences that regulate the alternative splicing of an oncogenic fusion transcript. Nucleic Acids Research 50, 6474–6496. https://doi.org/10.1093/nar/gkac409

Wang, E., Aifantis, I., 2020. RNA Splicing and Cancer. Trends in Cancer 6, 631–644. https://doi.org/10.1016/j.trecan.2020.04.011

Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R.M., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., Stuart, J.M., 2013. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 45, 1113–1120. https://doi.org/10.1038/ng.2764

Wilkinson, M.E., Charenton, C., Nagai, K., 2020. RNA Splicing by the Spliceosome. Annu Rev Biochem 89, 359–388. https://doi.org/10.1146/annurev-biochem-091719-064225

Xiao, X., Wang, Z., Jang, M., Nutiu, R., Wang, E.T., Burge, C.B., 2009. Splice site strength-dependent activity and genetic buffering by poly-G runs. Nat Struct Mol Biol 16, 1094–1100. https://doi.org/10.1038/nsmb.1661

Yamazaki, T., Liu, L., Lazarev, D., Al-Zain, A., Fomin, V., Yeung, P.L., Chambers, S.M., Lu, C.-W., Studer, L., Manley, J.L., 2018. TCF3 alternative splicing controlled by hnRNP H/F regulates E-cadherin expression and hESC pluripotency. Genes Dev. 32, 1161–1174. https://doi.org/10.1101/gad.316984.118

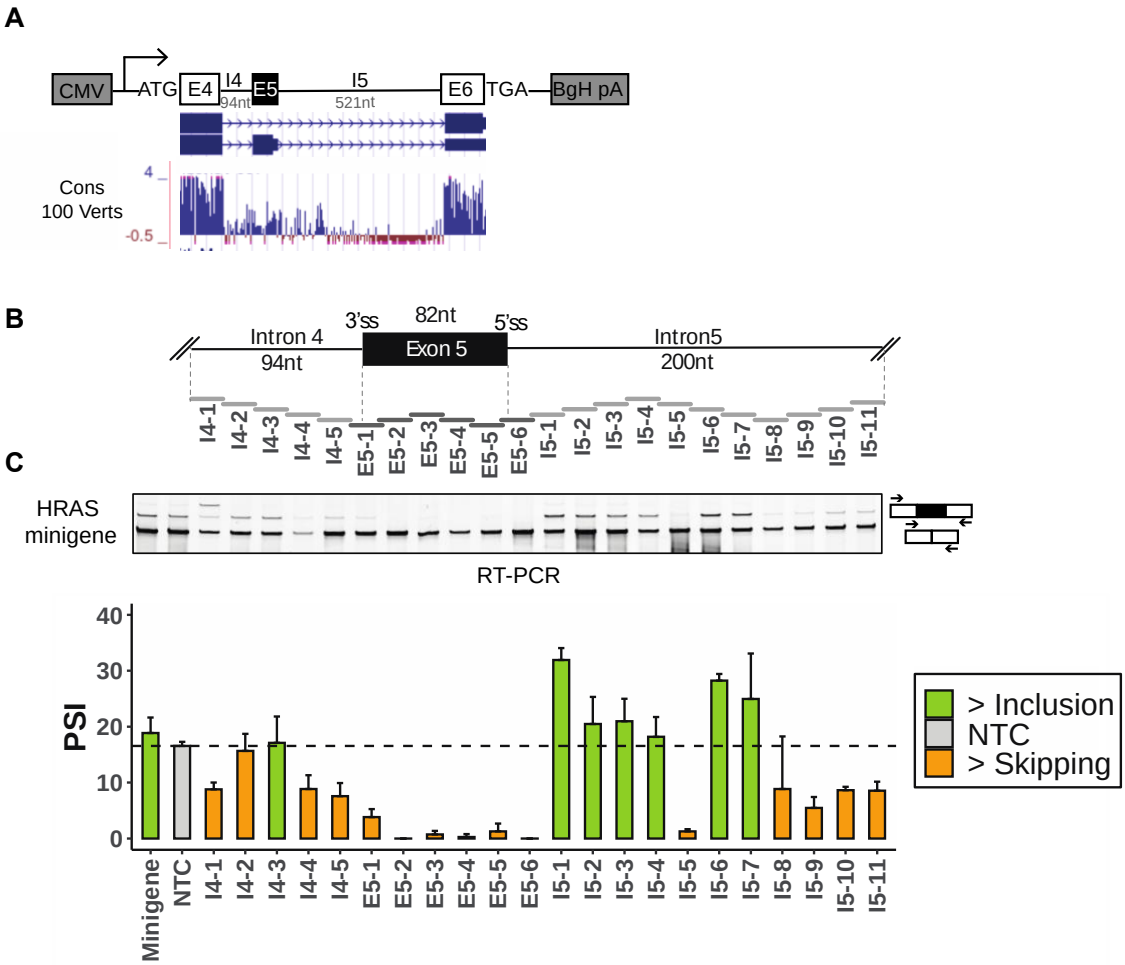

Figure 2-S1. Identification of splicing cis-regulatory elements with HRAS minigene reporter. **(A)** Diagram of HRAS minigene reporter with a schematic representation of gene structure and conservation. **(B)** Schematic of ASO tiling across HRAS cassette exon and its partial flanking introns on minigene reporter. A 376-nucleotide region tiled by non-overlapping 18-mer ASOs. Each horizontal bar represents an ASO. (C) Semi-quantitative RT-PCR analyses showing the effects of ASOs on HRAS minigene reporter. The arrows indicate the locations of minigene-specific primers. The bar graph shows the quantification of RT-PCR results calculated as percent-spliced-in (PSI) (gray: control; orange: more skipping; green: more inclusion). Each bar shows the mean PSI +/- SD of three biological replicates. NTC, non-targeting control. Minigene, no-ASO mock control.

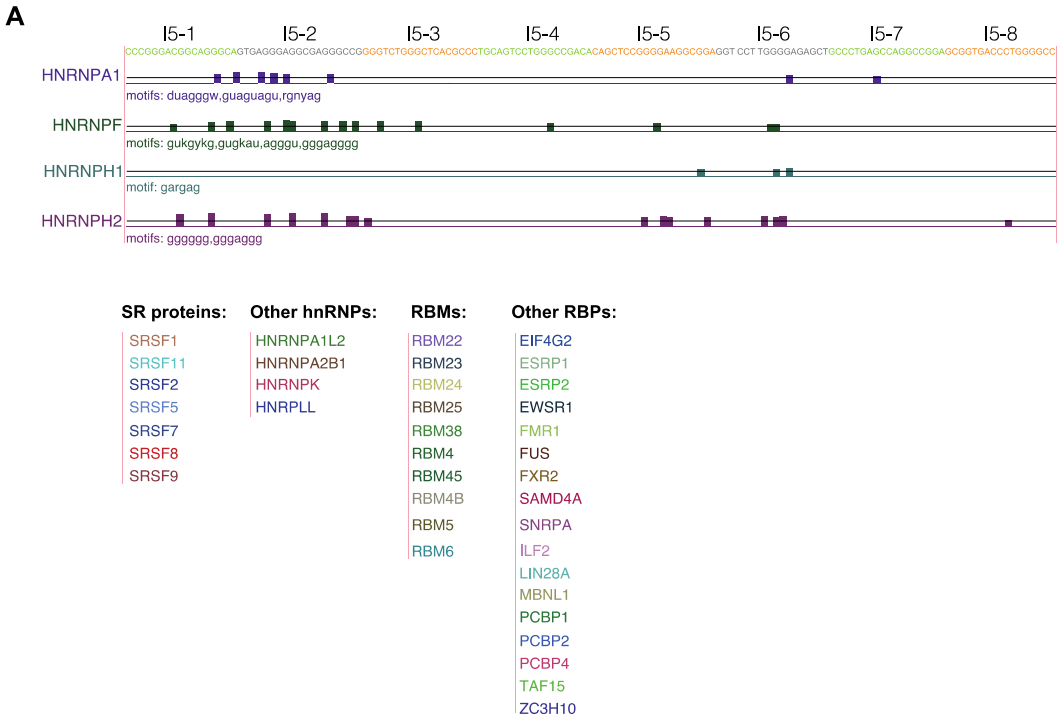

Figure 2-S2. RBPmap identifies binding motifs of potential splicing regulators in the HRAS intron 5 sequence. (A) UCSC genome browser visualization of representative RBPmap-predicted splicing factor binding sites on HRAS intron 5. The sequences targeted by each ASO are highlighted in orange (more skipping), green (more inclusion), gray (no change). Below are lists summarizing other potential splicing factors identified by the program, including SR proteins, other hnRNPs, RBMs, and others. Note that the motif listed for hnRNP H in this analyses is only one of a set of known hnRNP H motifs, which also include those listed for hnRNP F (Dominguez et al., 2018; Huelga et al., 2012; Uren et al., 2016).

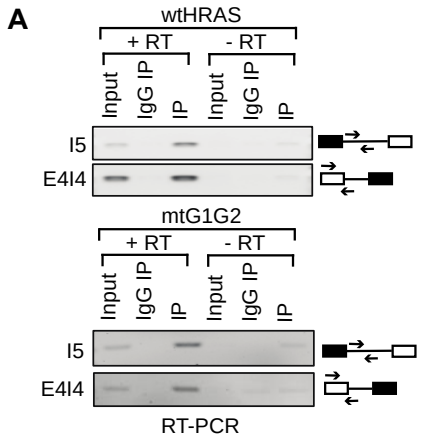

Figure 2-S3. Immunoprecipitated RNA is free of minigene DNA contamination. (A) Standard RT-PCR of immunoprecipitated RNA using primer pairs amplifying a segment of intron5 or the exon4-intron4 junction, in the presence [+] or absence [-] of reverse transcriptase. Immunoprecipitated RNA (80%) was amplified with 5 more PCR cycles than the Input control (5%). wt, wild type; mt, mutant.

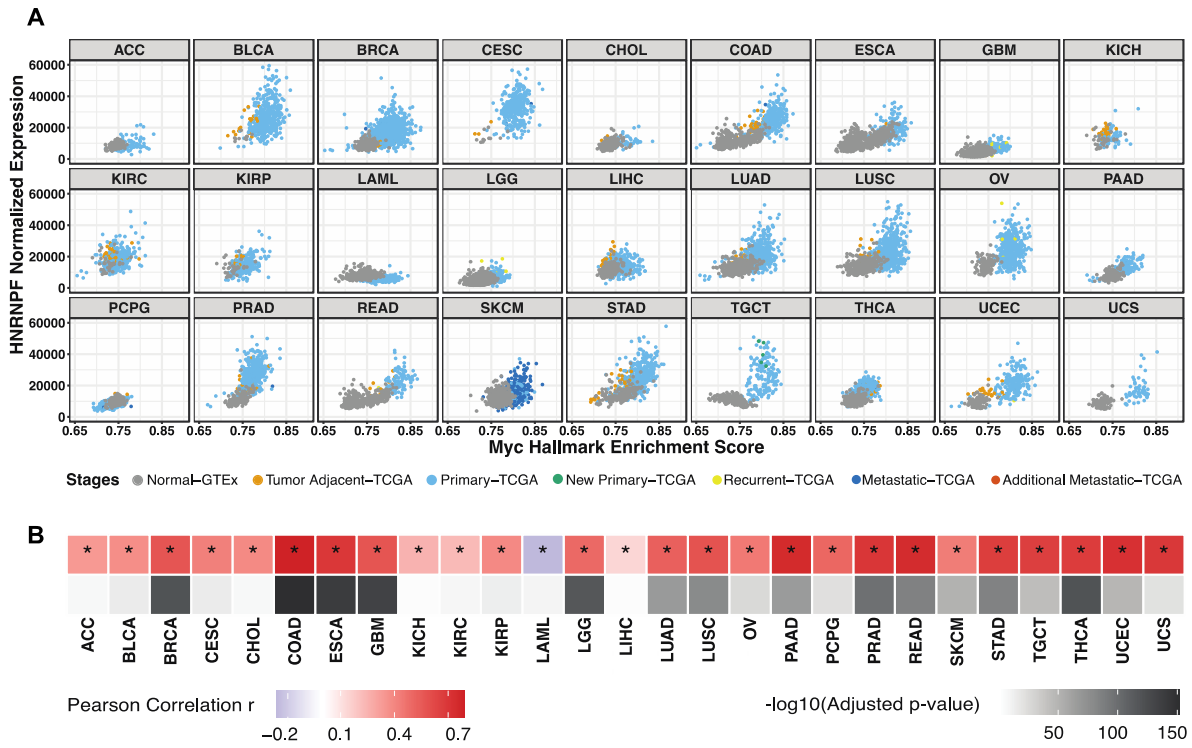

Figure 2-S4. HNRNPF gene expression is upregulated with Myc activation across tumor types. (A) Scatterplot matrix showing the correlation of normalized HNRNPF RNA level versus Myc hallmark enrichment score spanning the disease spectrum in multiple tumor types. **(B)** Heatmap summarizing the Pearson correlation coefficient for each tumor type, accompanied with its corresponding adjusted p-value. *, tumor types with statistically significant adjusted p-value (< 0.05).

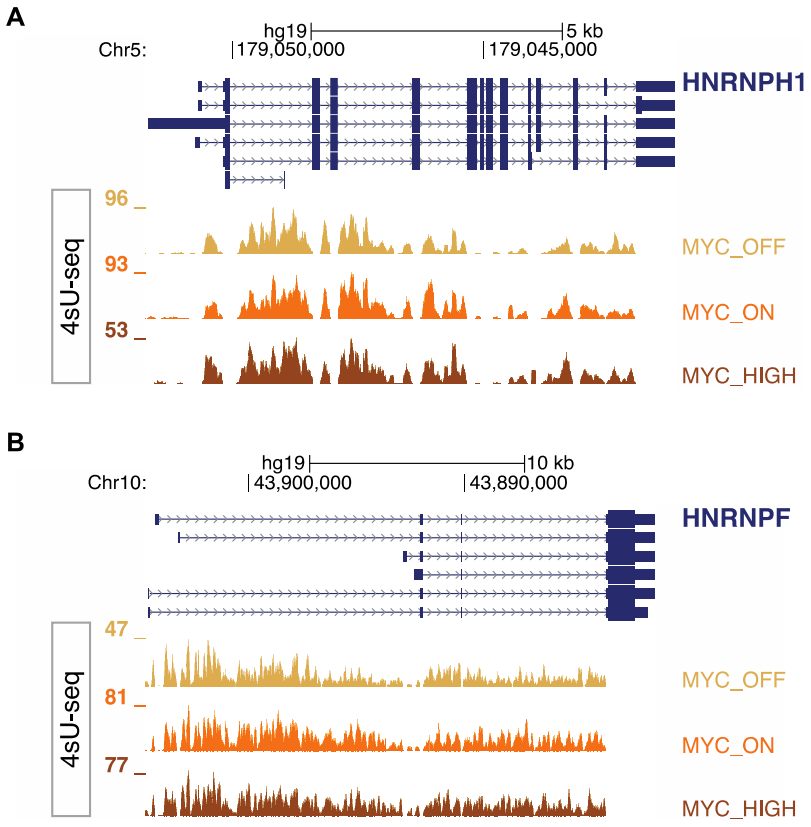

Figure 2-S5. HNRNPH1 transcription is reduced and HNRNPF transcription is increased at high MYC levels in an osteoblastoma model. UCSC genome browser track visualization of 4sU-seq datasets on the HNRNPH1 locus **(A)** and the HNRNPF locus **(B)** in an engineered U2OS cell line carrying an inducible MYC gene (Baluapuri et al., 2019). Data was generated under three different experimental conditions, MYC_OFF (siRNA-mediated MYC knockdown), MYC_ON (concurrent siRNA knockdown and induction of MYC which yields expression comparable to endogenous levels), and MYC_HIGH (induction of MYC overexpression to levels seen in cancer cells).

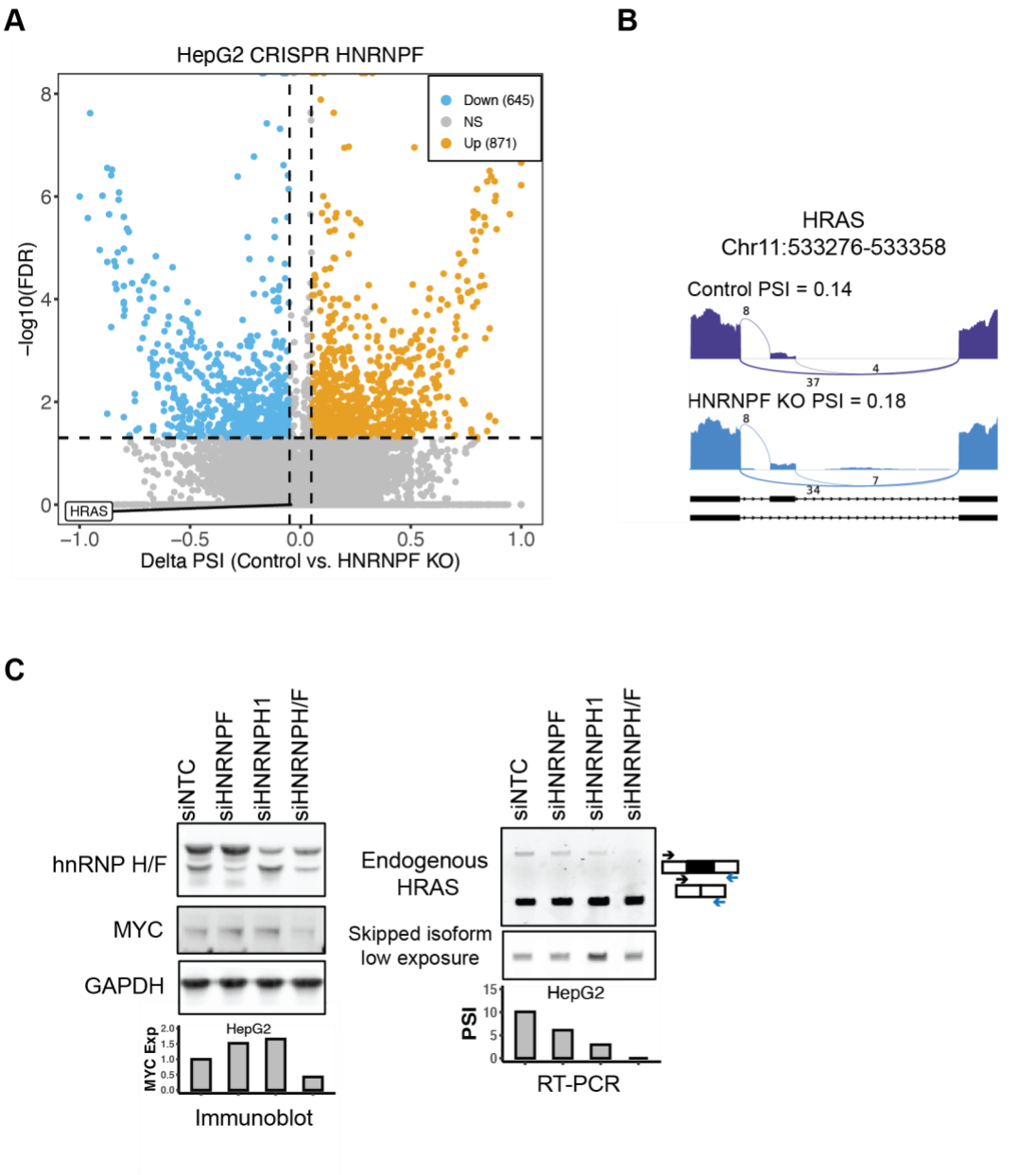

Figure 2-S6. Differential splicing analysis after HNRNPF depletion from HepG2 cells. (A) Scatterplot showing the SE events detected after HNRNPF CRISPR knockout in HepG2 cells compared to control. RNA-seq data from ENCODE was analyzed. Significant SE events were filtered by junction reads per event > 10, |deltaPSI| > 0.05, and FDR < 0.05. **(B)** Sashimi plot showing the PSI of HRAS exon 5 in hnRNPF KO and control cells. KO, knockout; NS, non-significant.

Figure 2-S6. Continued.

(C) The effects of hnRNP H/F knockdown on HRAS splicing in HepG2 cells. Immunoblot showing the protein levels of hnRNP H/F and MYC in HepG2 cells after transfection with control, HNRNPF, HNRNPH1, or HNRNPH/F siRNAs. GAPDH was used as a loading control.

The bar graph (left) shows the quantification of MYC expression in response to each siRNA perturbation. RT-PCR analysis of the splicing of endogenous HRAS transcripts after siRNA treatment. The bar graph (right) shows the quantification of HRAS PSI.

SUPPLEMENTARY METHODS

Nascent RNA-seq data processing and visualization

The bedGraph files of nascent 4sU-seq in U2OS cells were downloaded from Gene Expression Omnibus (accession no. GSE115365) (Baluapuri et al., 2019). The bigwig tracks were generated using bedGraphToBigWig (Kent et al., 2010) and visualized on the USCS genome browser (**http://genome.ucsc.edu**) (Kent et al., 2002).

Table 2-S1: List of ASOs

Name	Sequence 5'-3'
NTC	TCTATTAGCTAACAAATA
I4-1	GGTGGAGAGCTGCCTCAC
I4-2	GTCCCTGGCTAGCTGTGG
I4-3	GGGCGGGTCCCTGGCTAG
I4-4	TGCTGCTCCCTGGCTGGG
I4-5	AGGGAGAGGGTCAGTGAG
E5-1	AGCGGCTGCCCTGTGTCA
E5-2	AGCTGGAGCTAGAGCCAG
E5-3	GGTCCCAGAGGGTCCCGG
E5-4	GGTCACATGGGTCCCGGG
E5-5	GGGTCACATGGGTCCCGG
E5-6	AGACTTACAGCGCGAGGG
I5-1	CTGCCCTGCCGTCCCGGG
I5-2	CGGCCCTCGCCTCCCTCA
I5-3	GGGCGTGAGCCCAGACCC
I5-4	TGTCGGCCCAGGACTGCA
I5-5	CCGCCTTCCCCGGAGCTG
I5-6	AGCTCTCCCCAAGGACCT

I5-7	TCCGGCCTGGCTCAGGGC
I5-8	GGGCCCCAGGGTCACCGC
I5-9	GGGCCGGGCCCCAGGGTC
I5-10	CCCGTGGGACACTCTGGG
I5-11	CTCAGAACCAACAGGTGC

Table 2-S2: Lists of primers

Name	**Sequence 5'-3'**
T7-Fwd	CGACTCACTATAGGGAGACCCA
BgH-Rev	AACAACAGATGGCTGGCAAC
HRAS-Fwd	GAC TCG GAT GAC GTG CCC ATG
Cy5-HRAS-Rev	5'Cy5 - AGG AGG GTT CAG CTT CCG C
GAPDH-Fwd	GGA GCG AGA TCC CTC CAA AAT
GAPDH-Rev	GGC TGT TGT CAT ACT TCT CAT GG
HRAS E4I4 FWD	GCA CGC ACT GTG GAA TCT
HRAS E4I4 REV	GGA GAG CTG CCT CAC CT
HRAS I5 FWD	GGG CAC CTG TTG GTT CT
HRAS I5 REV	CCT CCT GAA CTC CAG GTC T

Table 2-S3: Lists of siRNAs

Name	**Sequence 5'-3'**	**Catalog Number/Reference**
siNTC	Silencer Select Negative Control #1	Thermo Fisher 4390843
siHNRNPF	GCGACCGAGAACGACAUUU	(Nazim et al., 2016)
siHNRNPH1	GAAGCAUACUGGUCCAAAU	Thermo Fisher s6728
siHNRNPH/F	GGAAGAAAUUGUUCAGUUC	(Garneau et al., 2005)

Table 2-S4: Lists of antibodies

Name	Manufacturer	Catalog Number
Anti-hnRNP H	Lab made	N/A
Anti-hnRNP F	Lab made	N/A
Anti-hnRNP H/F	Santa Cruz Biotechnology	sc-32310
Anti-His	MBL International	D291-3
Anti-GAPDH	Santa Cruz Biotechnology	sc-47724
AntI-HRAS	Abcam	ab86696
Anti-Flag	Sigma Aldrich	F3165
Mouse IgG isotype control	Invitrogen	14471485
Rabbit IgG isotype control	Thermo Fisher	02-6102
Anti-hnRNP H (IP)	Abcam	ab10374
Anti-hnRNP F (IP)	Sigma Aldrich	04-1462
Anti-MYC	Abcam	ab32072
Anti-cPARP	Cell signaling	5625S

Table 2-S5: Splicing changes of other G-run mutations

ASO Target	Mutations	Splicing changes	Consistency with ASO hits
I5-1 first G-run	ggg to gcg	No change	NA
I5-1 second G-run	ggg to gcg	Exon skipping	No
I5-4	ggg to gcg	Slight exon inclusion	Yes
I5-6	gggg to gctg	Exon inclusion	Yes with minigene data; No with endogenous data

REFERENCES

Baluapuri, A., Hofstetter, J., Dudvarski Stankovic, N., Endres, T., Bhandare, P., Vos, S.M., Adhikari, B., Schwarz, J.D., Narain, A., Vogt, M., Wang, S.-Y., Düster, R., Jung, L.A., Vanselow, J.T., Wiegering, A., Geyer, M., Maric, H.M., Gallant, P., Walz, S., Schlosser, A., Cramer, P., Eilers, M., Wolf, E., 2019. MYC Recruits SPT5 to RNA Polymerase II to Promote Processive Transcription Elongation. Molecular Cell 74, 674-687.e11. https://doi.org/10.1016/j.molcel.2019.02.031

Dominguez, D., Freese, P., Alexis, M.S., Su, A., Hochman, M., Palden, T., Bazile, C., Lambert, N.J., Van Nostrand, E.L., Pratt, G.A., Yeo, G.W., Graveley, B.R., Burge, C.B., 2018. Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. Molecular Cell 70, 854-867.e9. https://doi.org/10.1016/j.molcel.2018.05.001

Garneau, D., Revil, T., Fisette, J.-F., Chabot, B., 2005. Heterogeneous Nuclear Ribonucleoprotein F/H Proteins Modulate the Alternative Splicing of the Apoptotic Mediator Bcl-x. Journal of Biological Chemistry 280, 22641–22650. https://doi.org/10.1074/jbc.M501070200

Huelga, S.C., Vu, A.Q., Arnold, J.D., Liang, T.Y., Liu, P.P., Yan, B.Y., Donohue, J.P., Shiue, L., Hoon, S., Brenner, S., Ares, M., Yeo, G.W., 2012. Integrative Genome-wide Analysis Reveals Cooperative Regulation of Alternative Splicing by hnRNP Proteins. Cell Reports 1, 167–178. https://doi.org/10.1016/j.celrep.2012.02.001

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D., 2002. The human genome browser at UCSC. Genome Res 12, 996–1006. https://doi.org/10.1101/gr.229102

Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., Karolchik, D., 2010. BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics 26, 2204–2207. https://doi.org/10.1093/bioinformatics/btq351

Nazim, M., Masuda, A., Rahman, M.A., Nasrin, F., Takeda, J., Ohe, K., Ohkawara, B., Ito, M., Ohno, K., 2016. Competitive regulation of alternative splicing and alternative polyadenylation by hnRNP H and CstF64 determines acetylcholinesterase isoforms. Nucleic Acids Res gkw823. https://doi.org/10.1093/nar/gkw823

Uren, P.J., Bahrami-Samani, E., de Araujo, P.R., Vogel, C., Qiao, M., Burns, S.C., Smith, A.D., Penalva, L.O.F., 2016. High-throughput analyses of hnRNP H1 dissects its multi-functional aspect. RNA Biol 13, 400–411. https://doi.org/10.1080/15476286.2015.1138030

Chapter 3

Dissecting splicing regulation of the SRSF3 cassette exon

INTRODUCTION

Serine and arginine-rich splicing factor 3 (SRSF3, also known as SRp20) belongs to the highly conserved SR family of splicing factors that regulate both constitutive and alternative splicing (Corbo et al., 2013). SRSF3 expression is autoregulated by its splicing (Jumaa and Nielsen, 1997). E4-skipped transcripts are translated into full-length functional SRSF3 protein, with a N-terminal RNA recognition motif (RRM) mainly mediating RNA interaction, and C-terminal arginine and serine-rich (RS) domain responsible to interact with other splicing factors (**Figure 3-1A**). Overexpressed SRSF3 protein binds to its own pre-mRNA transcripts and activates the inclusion of E4 by enhancing the recognition of a weak splice donor site of E4. The inclusion of E4 brings in a PTC and leads to either NMD, or a truncated protein lacking of C-terminal RS domain but with implicated functions (Jiménez et al., 2019; Kano et al., 2014), and thus downregulates full-length functional SRSF3 protein level. This autoregulation activity is antagonized by another SR protein, SRSF1 (ASF/SF2), through RNA binding domain rather than RS domain (Jumaa and Nielsen, 2000, 1997). SRSF3 E4 and its flanking introns are ultra-conserved across mammalian species, suggesting the functional importance and potential cancer relevance of this poison cassette exon (Bejerano et al., 2004; Thomas et al., 2020) (**Figure 3-1B**). SRSF3 expression is also cell cycle dependent, regulated at least partially through its own splicing (Jumaa et al., 1997).

SRSF3 is frequently found upregulated in human cancers including breast, ovary, lung, colorectal and other cancer types (Anczuków and Krainer, 2016; Rahman et al., 2020; Zhou et al., 2020). SRSF3 functions as a proto-oncogene by regulating downstream splicing targets or coordinating with other oncogenic factors. SRSF3 globally affects splicing and gene expression in osteosarcoma U2OS cells (Ajiro et al., 2016), regulates ILF3 splicing to enhance cell proliferation and oncogenic transformation (Jia et al., 2019), cooperates with TDP43 to promote disease progression in triple negative breast cancer (Ke et al., 2018), regulates lncRNA ANRIL

splicing to promote drug resistance in pancreatic cancers (Wang et al., 2022), and promotes glioma tumorigenicity (Song et al., 2019). SRSF3 has an emerging role as a target in cancer therapeutics and recent efforts include using splice-switch ASO to block an ESS such that reduce SRSF3 expression in oral squamous carcinoma cells and breast cancer cells (Guo et al., 2018; Sun et al., 2019).

Src-associated substrate during mitosis of 68kDa (Sam68), also known as KH domain-containing RNA binding signaling transduction associated 1 (KHDRBS1), belongs to signaling transduction and activation of RNA (STAR) family of RNA binding protein (Lukong and Richard, 2003). Sam68 contains a GRP33/SAM68/GLD-1(GSG) domain consisting of a central heterogenous nuclear ribonucleoprotein K (hnRNP K) homology (KH) domain flanked by conserved N-terminal and C-terminal segments. GSG domain of Sam68 binds to A/U-rich RNA targets, specifically with UAAA or UUUA binding motifs (Lin et al., 1997). Sam68 is involved in multiple biological functions including mitosis and cell cycle regulation, signal transduction, transcription, alternative splicing, and nuclear export (Lukong and Richard, 2003). Increasing evidences suggest the Sam68 involvement in cancer onset and progression, particularly endocrine cancers such as prostate and breast cancers (Bielli et al., 2011). Sam68 is phosphorylated and activated by RAS/MAPK signaling pathway (Matter et al., 2002), and is known to regulate the splicing of proto-oncogene CCND1 (Paronetto et al., 2010), cell surface marker CD44 (Matter et al., 2002), and the transcription of androgen receptor (AR) (Rajan et al., 2008) in prostate cancer cells.

We previously identified that besides the HRAS exon, splicing of the SRSF3 cassette exon, is also particularly responsive to Myc activation in prostate and breast cancers (Phillips et al., 2020). However, the connection between its splicing mechanism and Myc oncogenic transformation remains unknown. Here, we applied ASO tiling to identify the cis-elements regulating SRSF3 cassette exon splicing. Motif based prediction suggested the presence of

multiple Sam68 binding sites in the downstream intron identified as potential ISS by ASO tiling. Using siRNA knockdown of KHDRBS1, we found Sam68 represses SRSF3 splicing. Together, our results reveal mechanism of Sam68 regulation of SRSF3 splicing.

RESULTS

ASO tiling reveals cis-regulatory elements controlling SRSF3 cassette exon splicing

SRSF3 contains a Myc-dependent cassette exon, the skipping of which is correlated with Myc hallmark enrichment in multiple cancers (Phillips et al., 2020). To comprehensive dissect the cis-regulatory elements of this exonic and intronic region, we constructed a SRSF3 minigene reporter by cloning E4, partial upstream E3 and downstream E5, as well as intervening intron into a pcDNA3.1(+) expression vector (**Figure 3-1B**). We utilized ASO tiling combined with minigene reporter to walk across the entire SRSF3 E4 (456nt) and its partial flanking introns (200nt each side) (**Figure 3-2A**). We co-transfected the minigene reporter with each ASO into HEK293 cells, and compared the splicing changes induced by each ASO with minigene only control (-) and non-targeting control ASO, by semi-quantitative RT-PCR (**Figure 3-2B**). Noticeably, multiple oligos targeting the exon, including ASO_095, ASO_112-115, ASO_121-114, ASO_113-119, and ASO128-126, induced strong exon skipping, suggesting the presence of numerous exonic splicing enhancers. Some of these enhancer regions might be SRSF3 binding sites that activate the inclusion of exon 4. As expected, ASO targeting the 5' splice site, ASO_122, induced a cryptic exon, which is commonly seen in minigene reporters. In addition, ASO_138-139 and ASO_133, targeting downstream intron 4, resulted in exon inclusion that suggested the presence of intronic splicing enhancers. To validate our ASO tiling results, we transfected ASO into a prostate cancer cell line (LNCaP) and performed RT-PCR to analyze the splicing changes of endogenous SRSF3 transcripts (**Figure 3-S1A-B**). The default PSI value for endogenous SRSF3 exon 4 is 6.8%, which is fairly low compared with 24.7% in SRSF3 minigene. Nevertheless, ASO_121 and ASO_126 still induced exon skipping, that is consistent with the minigene data, further suggesting

the presence of enhancers at these two regions. More importantly, ASOs targeting the downstream intron 4, including ASO130-138-130-133, strongly induced exon inclusion, which confirmed the minigene data and suggested the presence of intronic splicing silencers. Together, our minigene and endogenous ASO results suggested the presence of multiple exonic splicing enhancers as well as putative downstream intronic silencers.

Sam68 represses SRSF3 exon 4 splicing

To identify the trans-acting splicing factors that represses SRSF3 exon4 splicing in the downstream intron regions, we performed motif-based RBP prediction using SpliceAid2 (Piva et al., 2012). Noticeable, this program identified multiped binding motifs for Sam68 in the region predicted as silencers by ASO138-139, and ASO_133 (**Figure 3-3A**). We then performed siRNA-mediated knockdown of KHDRBS1 in HEK293 cells and analyzed the splicing change of endogenous SRSF3 by RT-PCR. The immunoblot confirmed the loss of Sam68 protein using siRNA targeting KHDRBS1 compared with siRNA non-targeting control, and RT-PCR results confirmed the depletion of Sam68 led to the inclusion of SRSF3 cassette exon (**Figure 3-3B**). Collectively, our data indicated that Sam68 acts as a SRSF3 splicing repressor.

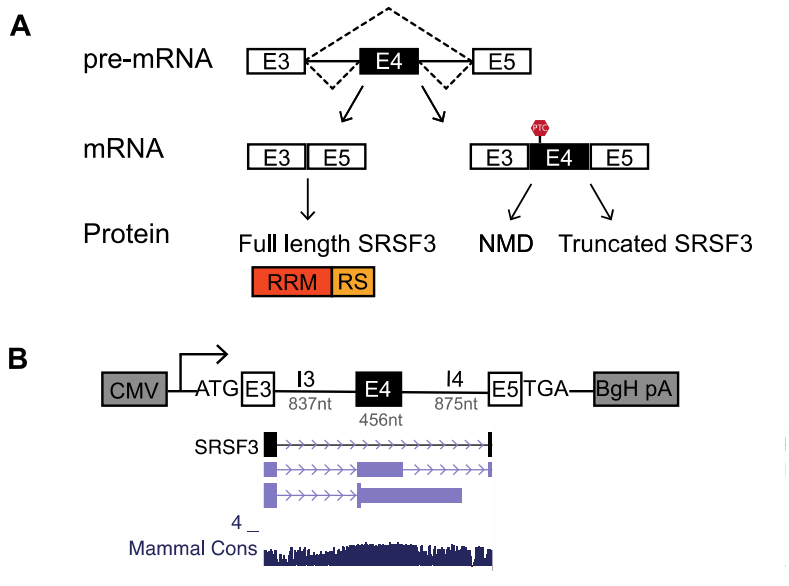

Figure 3-1. SRSF3 splicing and construction of its minigene reporter. (A) Diagram of SRSF3 pre-mRNA alternative splicing. NMD, nonsense-mediated decay; RRM, RNA recognition motif; RS, arginine- and serine-rich domain. **(B)** Diagram of SRSF3 minigene reporter with a schematic representation of gene structure and conservation.

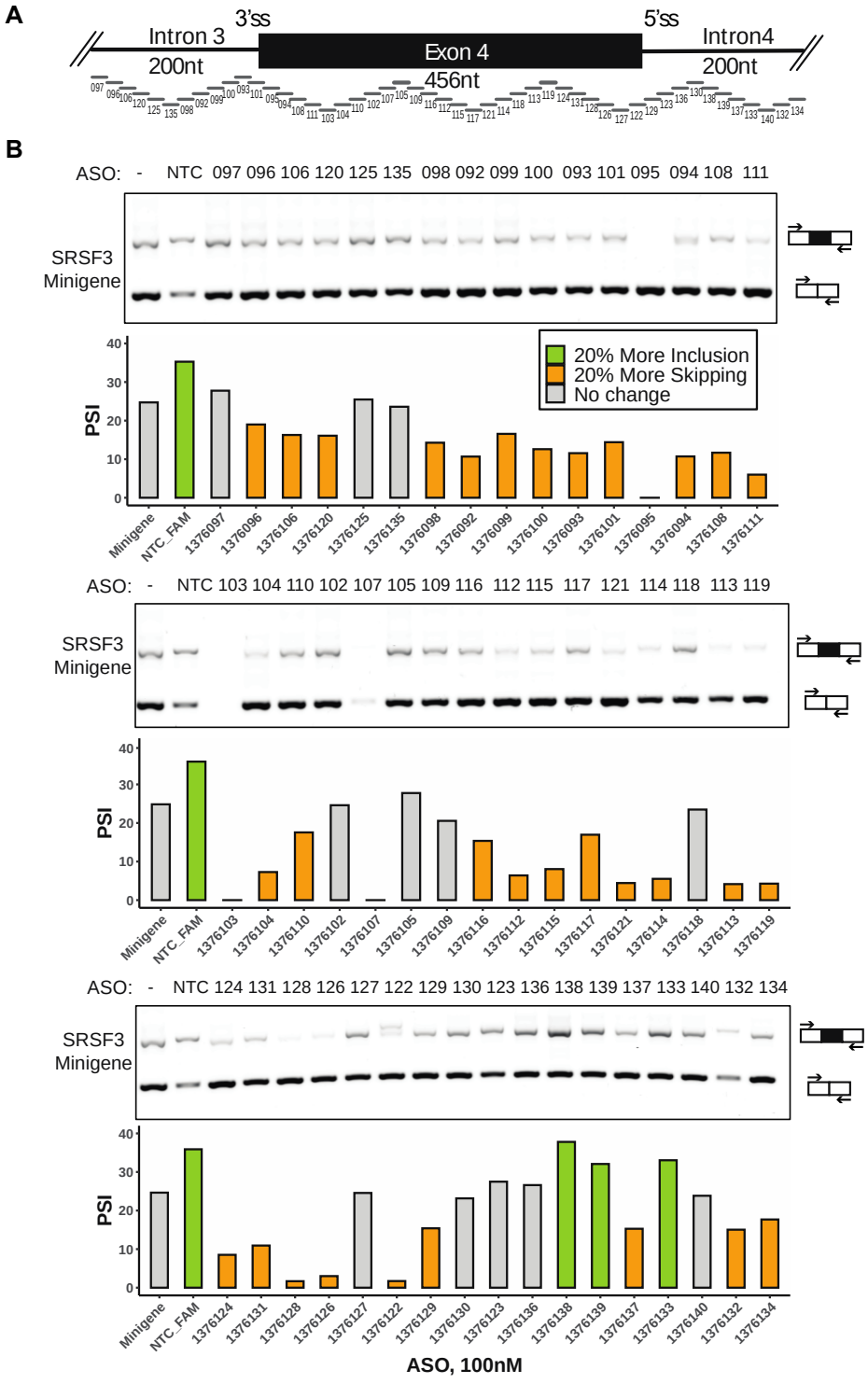

Figure 3-2. Identification of splicing cis-regulatory elements controlling SRSF3 exon 4 with the minigene reporter.

Figure 3-2. Continued.

(A) Schematic of ASO tiling across SRSF3 exon 4 and its partial flanking introns, including 856-nucleotide region tiled by non-overlapping 18-mer ASOs. Each horizontal bar represents an ASO. **(B)** Semi-quantitative RT-PCR analyses showing the effects of ASOs on SRSF3 minigene splicing in HEK293 cells. The bar graphs present the quantification of the RT-PCR calculated as PSI (gray: control or less than 20% change); orange: 20% more skipping compared to control; green: 20% more inclusion compared to control). NTC_FAM, non-targeting control labeled with 3' FAM fluorophore. -, minigene only control.

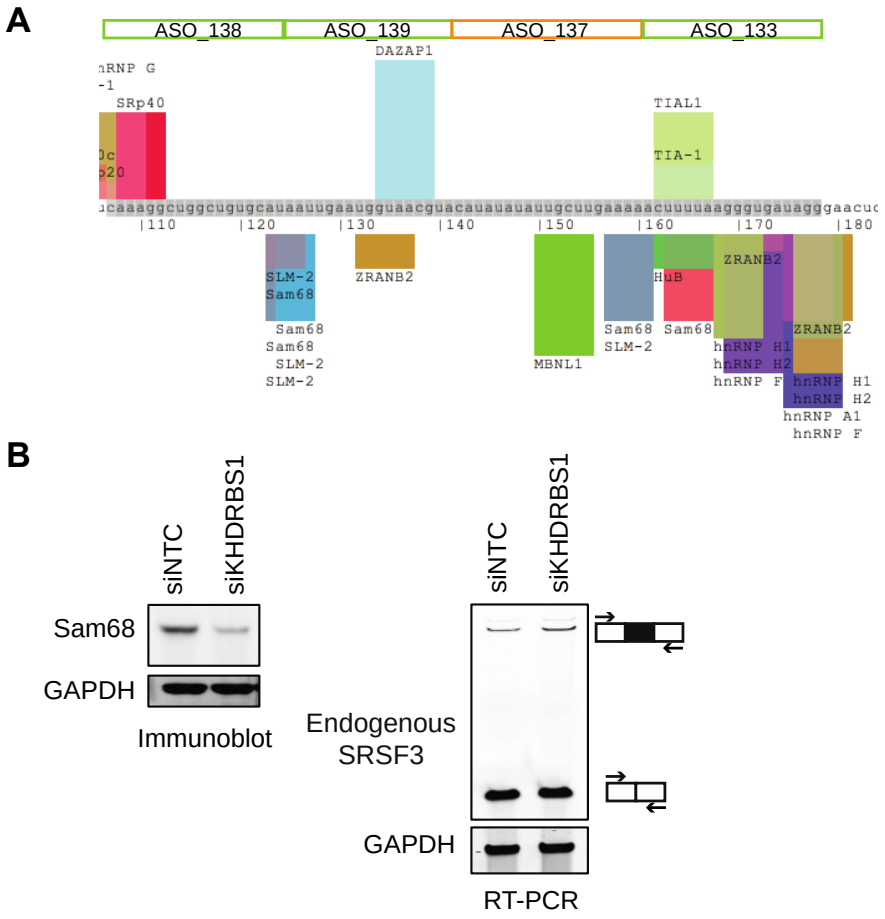

Figure 3-3. Sam68 represses SRSF3 exon 4 splicing. (A) SpliceAid2 snapshot of predicted splicing factor binding sites on SRSF3 intron 4. The sequences targeted by each ASO are highlighted in orange (20% more skipping), green (20% more inclusion). **(B)** Immunoblot showing the expression of Sam68 protein in HEK293 cells, after transfection with control, KHDRBS1 siRNAs. GAPDH was used as a loading control. RT-PCR analysis of endogenous SRSF3 splicing after siRNA knockdown of KHDRBS1 in HEK293 cells.

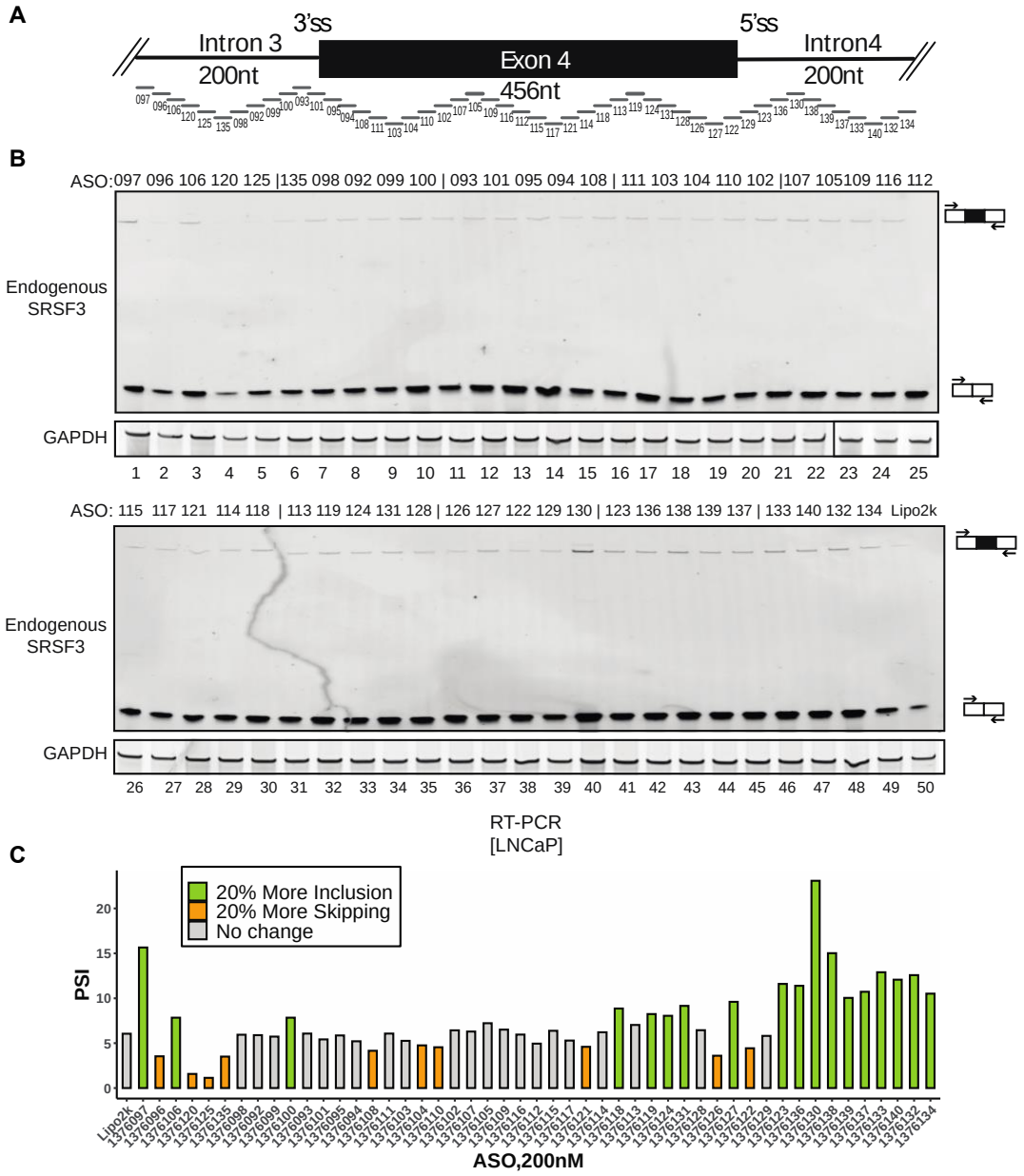

Figure 3-S1. Identification of splicing cis-regulatory elements controlling endogenous SRSF3 exon 4 in prostate cancer cells. **(A)** Schematic of ASO tiling across SRSF3 exon 4 and its partial flanking introns. **(B)** Semi-quantitative RT-PCR analyses showing the effects of ASOs on endogenous SRSF3 splicing in LNCaP cells. The bar graphs present the quantification

Figure 3-S1. Continued.

of the RT-PCR calculated as PSI (gray: control or less than 20% change); orange: 20% more skipping compared to control; green: 20% more inclusion compared to control).

DISCUSSION

Our ASO-tiling analysis identified multiple cis-regulatory elements in SRSF3 E4 and its flanking introns, and we further confirmed that Sam68 repressed SRSF3 E4 inclusion, potentially through a downstream ISS.

Earlier studies found PTBP1/2 bind to an ESS in SRSF3 E4 to repress its splicing, and a blocking of this ESS by an ASO named SR-3 could promote E4 inclusion and reduce SRSF3 expression (Guo et al., 2018, 2015). This ESS is spanned by our ASO_126/127/122, with SR-3 sequences largely overlapped with our ASO_127 sequences. Consistent with their results, our ASO_127 induced E4 inclusion in the endogenous ASO data (Figure 3-S1B), but the ASO_126/122 targeting neighboring sequences induced strong exon skipping in both minigene and endogenous data (Figure 3-2B, 3-S1B), suggesting the presence of multiple ESE within this exonic region. It would be interesting to investigate other splicing factors besides PTBP1/2 that regulates SRSF3 E4 splicing through this complex regulatory sequences.

The autoregulation of SRSF3 makes it a challenging target to dissect by ASO tiling. ASOs that induce E4 skipping would presumably induce SRSF3 overexpression, which in turn activates the inclusion of E4 in order to downregulate SRSF3 expression. This negative feedback might explain that ASO tiling only induced relatively minor effects in SRSF3 E4 analysis. The E4 is largely skipped in semi-quantitative RT-PCR results, which is partially resulted from the NMD of E4-included endogenous transcripts, as well as the potential PCR bias towards the shorter amplicon (skipping isoform, 134bp) in relative to the longer one (included isoform, 590bp). It challenged the quantification of endogenous PSI and the identification of potential splicing enhancer regions.

One of the RBPs found highly relevant to prostate cancer is Sam68. It is upregulated in many cancers including prostate cancer and it promotes prostate cancer cell proliferation and

survival (Busà et al., 2007). The downregulation of Sam68 results in delay in cell cycle progression and reduced proliferation in LNCaP cells. An earlier study showed MYC controls the transcription and production splicing of Sam68 in prostate cancer, and the its splicing is regulated by hnRNP H/F (Caggiano et al., 2019). In Chapter 2 of this dissertation, our HRAS study also identified hnRNP H/F as important splicing regulators in Myc-driven prostate cancer. It would be interesting to further study the connection between MYC, Sam68 and hnRNP H/F, to understand if MYC empowers hnRNPH/F to function as 'master' splicing regulators in prostate cancer.

METHODS

Cell culture, cloning, siRNA/ASO transfection, RT-PCR, and immunoblotting were performed as described in Chapter 2. ASOs were transfected into LNCaP cells at 200nM instead of 100nM concentration. A list of SRSF3 ASO sequences is presented in the supplemental table S3-1. Antibody specifically used in this Chapter includes: Anti-SAM68 (Bethyl Laboratories A302-110A); siRNA specifically used in this Chapter includes: siKHDRBS1 5' GAAGAUUCUUGGACCACAAtt-3' (Thermo Fisher s20951); RT-PCR primers for endogenous SRSF3 transcripts: Forward 5'-GAGTGGAACTGTCGAATGGTG-3'; Reverse 5'-GGCTGCGAGAGAAGCTTC-3'.

Table S3-1: List of ASOs

ID#	Sequence 5'-3'
1376106	GGGCAAGATTTTGAACTC
1376093	GGGTTCATTTTTATTAAT
1376110	GGGTCATGTGAAACGACA
1376117	AAATAAACAGCTGAGCTG
1376128	ACTGACGGCCAGGAAGAA
1376138	ATGCACAGCCAGCCTTTG
1376120	TTCAAAATAGCAAAAGGG
1376101	TGGTGACTCTGTAACGGG
1376102	CTGACGACTGGCCAGCCT
1376121	GTTGTGAAGACATTTGGT
1376126	GTGAGGTGAGGCTGCCGG
1376139	GTACGTTACCATTCAATT
1376125	AACAGTGTGCTGTTGTTT
1376095	TGGTGAGAAGAGACATGA
1376107	ACCAAAGGGCGGTGCAAC
1376114	AAGGCTGCTGCTGTAGTT
1376127	TTCAACAAGCTAGAAATG
1376137	TTTTCAAGCAATATATAT
1376135	CTCTTTATTATGATGGGC
1376094	TAATGCAGATTCAGAGGG
1376105	AAAACAGCATGCTCGGGA
1376118	CTGCTTTTTTGTTAGCCG
1376122	AACTTACTAGTTTTGGGT
1376133	CCTATCACCCTTAAAAGT
1376098	TCTATTAGCTAACAAATA
1376108	AAGGGCTAGTTGACTGGC
1376109	TGGAGAGAAGGCTGAGAG

1376113	GGGTGTTGTGGATTTTTC
1376129	AACTGTATAAGCAGGAAA
1376140	TAGTTACATTACGAGTTC
1376092	AAGCCATCAGTACAACCA
1376111	CTGGTCACATGACGCTGA
1376116	TGCCGATTTGGTTAAGGT
1376119	TTAGTTGGTTGGCGAAGG
1376130	TTATTTTAACCAGCAGTA
1376132	AAGTTTCACTTGAAGGCC
1376097	CTCCACTCACCTGTAAGA
1376099	ACAAAAAAAATGAAAAAC
1376103	GGGCGCGCTGGTCACATG
1376112	GGCGGTCGAGGTGGCTGC
1376124	CCAGATGTTGCGTTGGAT
1376123	CTTTAAGCCGCTTACTCC
1376134	TGCACATGCAAATAAGTT
1376096	TCCTGTCTCCTAGAAGGA
1376100	ATAGATGGACCAAAAAGC
1376104	CCAGCCAAGCTGAATGGG
1376115	ATTGGCCAGGAATGTGTG
1376131	TTTGCTGAAAAGGTTTTG
1376136	ATCCAGAAAAAAGAATTA
NTC-FAM	CCAGGATGGGCACCACCC

REFERENCES

Ajiro, M., Jia, R., Yang, Y., Zhu, J., Zheng, Z.-M., 2016. A genome landscape of SRSF3-regulated splicing events and gene expression in human osteosarcoma U2OS cells. Nucleic Acids Res 44, 1854–1870. https://doi.org/10.1093/nar/gkv1500

Anczuków, O., Krainer, A.R., 2016. Splicing-factor alterations in cancers. RNA 22, 1285–1301. https://doi.org/10.1261/rna.057919.116

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., Haussler, D., 2004. Ultraconserved Elements in the Human Genome. Science 304, 1321–1325. https://doi.org/10.1126/science.1098119

Bielli, P., Busà, R., Paronetto, M.P., Sette, C., 2011. The RNA-binding protein Sam68 is a multifunctional player in human cancer. Endocr Relat Cancer 18, R91–R102. https://doi.org/10.1530/ERC-11-0041

Busà, R., Paronetto, M.P., Farini, D., Pierantozzi, E., Botti, F., Angelini, D.F., Attisani, F., Vespasiani, G., Sette, C., 2007. The RNA-binding protein Sam68 contributes to proliferation and survival of human prostate cancer cells. Oncogene 26, 4372–4382. https://doi.org/10.1038/sj.onc.1210224

Caggiano, C., Pieraccioli, M., Panzeri, V., Sette, C., Bielli, P., 2019. c-MYC empowers transcription and productive splicing of the oncogenic splicing factor Sam68 in cancer. Nucleic Acids Research 47, 6160–6171. https://doi.org/10.1093/nar/gkz344

Corbo, C., Orrù, S., Salvatore, F., 2013. SRp20: An overview of its role in human diseases. Biochemical and Biophysical Research Communications 436, 1–5. https://doi.org/10.1016/j.bbrc.2013.05.027

Guo, J., Che, X., Wang, X., Jia, R., 2018. Inhibition of the expression of oncogene SRSF3 by blocking an exonic splicing suppressor with antisense oligonucleotides. RSC Adv. 8, 7159–7163. https://doi.org/10.1039/C7RA11267J

Guo, J., Jia, J., Jia, R., 2015. PTBP1 and PTBP2 impaired autoregulation of SRSF3 in cancer cells. Sci Rep 5, 14548. https://doi.org/10.1038/srep14548

Jia, R., Ajiro, M., Yu, L., McCoy, P., Zheng, Z.-M., 2019. Oncogenic splicing factor SRSF3 regulates ILF3 alternative splicing to promote cancer cell proliferation and transformation. RNA 25, 630–644. https://doi.org/10.1261/rna.068619.118

Jiménez, M., Urtasun, R., Elizalde, M., Azkona, M., Latasa, M.U., Uriarte, I., Arechederra, M., Alignani, D., Bárcena-Varela, M., Álvarez-Sola, G., Colyn, L., Santamaría, E., Sangro, B., Rodriguez-Ortigosa, C., Fernández-Barrena, M.G., Ávila, M.A., Berasain, C., 2019. Splicing events in the control of genome integrity: role of SLU7 and truncated SRSF3 proteins. Nucleic Acids Research 47, 3450–3466. https://doi.org/10.1093/nar/gkz014

Jumaa, H., Guénet, J.L., Nielsen, P.J., 1997. Regulated expression and RNA processing of transcripts from the Srp20 splicing factor gene during the cell cycle. Mol Cell Biol 17, 3116–3124. https://doi.org/10.1128/MCB.17.6.3116

Jumaa, H., Nielsen, P.J., 2000. Regulation of SRp20 exon 4 splicing. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 1494, 137–143. https://doi.org/10.1016/S0167-4781(00)00233-5

Jumaa, H., Nielsen, P.J., 1997. The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. EMBO J 16, 5077–5085. https://doi.org/10.1093/emboj/16.16.5077

Kano, S., Nishida, K., Kurebe, H., Nishiyama, C., Kita, K., Akaike, Y., Kajita, K., Kurokawa, K., Masuda, K., Kuwano, Y., Tanahashi, T., Rokutan, K., 2014. Oxidative stress-inducible truncated serine/arginine-rich splicing factor 3 regulates interleukin-8 production in human colon cancer cells. American Journal of Physiology-Cell Physiology 306, C250–C262. https://doi.org/10.1152/ajpcell.00091.2013

Ke, H., Zhao, L., Zhang, H., Feng, X., Xu, H., Hao, J., Wang, S., Yang, Q., Zou, L., Su, X., Wang, L., Wu, C., Wang, Y., Nie, J., Jiao, B., 2018. Loss of TDP43 inhibits progression of triple-negative breast cancer in coordination with SRSF3. Proceedings of the National Academy of Sciences 115, E3426–E3435. https://doi.org/10.1073/pnas.1714573115

Lin, Q., Taylor, S.J., Shalloway, D., 1997. Specificity and Determinants of Sam68 RNA Binding: IMPLICATIONS FOR THE BIOLOGICAL FUNCTION OF K HOMOLOGY DOMAINS*. Journal of Biological Chemistry 272, 27274–27280. https://doi.org/10.1074/jbc.272.43.27274

Lukong, K.E., Richard, S., 2003. Sam68, the KH domain-containing superSTAR. Biochim Biophys Acta 1653, 73–86. https://doi.org/10.1016/j.bbcan.2003.09.001

Matter, N., Herrlich, P., König, H., 2002. Signal-dependent regulation of splicing via phosphorylation of Sam68. Nature 420, 691–695. https://doi.org/10.1038/nature01153

Paronetto, M.P., Cappellari, M., Busà, R., Pedrotti, S., Vitali, R., Comstock, C., Hyslop, T., Knudsen, K.E., Sette, C., 2010. Alternative Splicing of the Cyclin D1 Proto-Oncogene Is Regulated by the RNA-Binding Protein Sam68. Cancer Research 70, 229–239. https://doi.org/10.1158/0008-5472.CAN-09-2788

Phillips, J.W., Pan, Y., Tsai, B.L., Xie, Z., Demirdjian, L., Xiao, W., Yang, H.T., Zhang, Y., Lin, C.H., Cheng, D., Hu, Q., Liu, S., Black, D.L., Witte, O.N., Xing, Y., 2020. Pathway-guided analysis identifies Myc-dependent alternative pre-mRNA splicing in aggressive prostate cancers. Proc. Natl. Acad. Sci. U.S.A. 117, 5269–5279. https://doi.org/10.1073/pnas.1915975117

Piva, F., Giulietti, M., Burini, A.B., Principato, G., 2012. SpliceAid 2: A database of human splicing factors expression data and RNA target motifs. Human Mutation 33, 81–85. https://doi.org/10.1002/humu.21609

Rahman, M.A., Krainer, A.R., Abdel-Wahab, O., 2020. SnapShot: Splicing Alterations in Cancer. Cell 180, 208-208.e1. https://doi.org/10.1016/j.cell.2019.12.011

Rajan, P., Gaughan, L., Dalgliesh, C., El-Sherif, A., Robson, C., Leung, H., Elliott, D., 2008. The RNA-binding and adaptor protein Sam68 modulates signal-dependent splicing and transcriptional activity of the androgen receptor. The Journal of Pathology 215, 67–77. https://doi.org/10.1002/path.2324

Song, X., Wan, X., Huang, T., Zeng, C., Sastry, N., Wu, B., James, C.D., Horbinski, C., Nakano, I., Zhang, W., Hu, B., Cheng, S.-Y., 2019. SRSF3-Regulated RNA Alternative Splicing Promotes Glioblastoma Tumorigenicity by Affecting Multiple Cellular Processes. Cancer Res 79, 5288–5301. https://doi.org/10.1158/0008-5472.CAN-19-1504

Sun, Y., Yan, L., Guo, J., Shao, J., Jia, R., 2019. Downregulation of SRSF3 by antisense oligonucleotides sensitizes oral squamous cell carcinoma and breast cancer cells to paclitaxel treatment. Cancer Chemother Pharmacol 84, 1133–1143. https://doi.org/10.1007/s00280-019-03945-9

Thomas, J.D., Polaski, J.T., Feng, Q., De Neef, E.J., Hoppe, E.R., McSharry, M.V., Pangallo, J., Gabel, A.M., Belleville, A.E., Watson, J., Nkinsi, N.T., Berger, A.H., Bradley, R.K., 2020. RNA isoform screens uncover the essentiality and tumor suppressor activity of ultraconserved poison exons. Nat Genet 52, 84–94. https://doi.org/10.1038/s41588-019-0555-z

Wang, Z.-W., Pan, J.-J., Hu, J.-F., Zhang, J.-Q., Huang, L., Huang, Y., Liao, C.-Y., Yang, C., Chen, Z.-W., Wang, Y.-D., Shen, B.-Y., Tian, Y.-F., Chen, S., 2022. SRSF3-mediated regulation of N6-methyladenosine modification-related lncRNA ANRIL splicing promotes resistance of pancreatic cancer to gemcitabine. Cell Reports 39. https://doi.org/10.1016/j.celrep.2022.110813

Zhou, Z., Gong, Q., Lin, Z., Wang, Y., Li, M., Wang, L., Ding, H., Li, P., 2020. Emerging Roles of SRSF3 as a Therapeutic Target for Cancer. Front. Oncol. 10, 577636. https://doi.org/10.3389/fonc.2020.577636

Chapter 4

CRISPRi screening of trans-acting splicing regulators of MYC-dependent exons

INTRODUCTION

RNA-guided CRISPR (clustered regularly interspaced short palindromic repeats)-Cas nuclease system is a revolutionary and widely used technology for genetic engineering and functional genomics (Doench, 2018; Knott and Doudna, 2018; Shalem et al., 2015). CRISPR-Cas system is repurposed from bacterial adaptive immunity (Barrangou et al., 2007; Brouns et al., 2008; Marraffini and Sontheimer, 2008). The basic CRISPR-Cas9 system requires a customizable 20-nt single guide RNA (sgRNA) that recognize targeted DNA sequences known as protospacer via base-pairing, and recruits the Cas9 nuclease that makes a double strand break (DSB) three nucleotides upstream of protospacer associated motif (PAM). Beyond its extensive usage in CRISPR knockout (CRISPRko) genome editing, CRISPR-Cas9 system has been redesigned to modulate transcription, either transcriptionally repressing (CRISPR interference) or activating (CRISPR activation) genes (Dominguez et al., 2016; Gilbert et al., 2014, 2013; Kampmann, 2018; Konermann et al., 2015). Catalytically inactive Cas9 protein, termed dCas9, is mutated in its nuclease domains and lacing of endonucleolytic activity. dCas9 protein binds but does not cleave at targeted genomic loci. Directed by sgRNA, dCas9 fused with a variety of effector domain, such as transcriptional repression KRAB domain, or transcriptional activation VP64 domain, recruits transcriptional silencing or activating factors when it binds to transcriptional start sites (TSS), and achieves a sequence-specific control of gene expression.

CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa) screening enables efficient transcriptional control at endogenous genomic loci on a genome-wide scale, to facilitate understanding of complex gene functions and genetic interactions in diseases. CRISPRi/a screens have identified long non-coding RNA targets affecting cancer cell growth or as potential cancer therapeutics (Esposito et al., 2019; Liu et al., 2020, 2017). Combined CRISPRi/a chemical genetic screen approaches identified targets and revealed mechanisms for small

molecular drugs (Jost et al., 2017; Jost and Weissman, 2018). sgRNA optimization using machine learning approaches enables the design of compact but highly-active genome-wide scale CRISPRi/a libraries (Horlbeck et al., 2016). Genome-wide CRISPRko fitness screening has potential systematic bias towards false negatives (Dede and Hart, 2022). We reason that a certain advantage of CRISPRi/a screening over the canonical CRISPRko screening would be to interrogate the function of essential genes, including most of the splicing factors. CRISPRi/a could potentially detect potential false negatives in CRISPRko screening – missed hits resulted from induced cell apoptosis after knockout of essential genes.

In previous chapters of this dissertation, we dissected cis-regulatory elements of HRAS and SRSF3 cassette exon splicing. However, it remains unclear what potential trans-acting splicing factors might function as 'master' regulators controlling splice isoform choice in Myc-driven prostate cancer. Here, we performed RBP-focused differentially gene expression analysis on RNA-seq data generated from an engineered prostate cancer cell model with inducible MYC expression. Multiple RBPs exhibited reduced mRNA expression in response to MYC depletion. We then performed a pilot CRISPRi screening on the top MYC-responsive RBPs, along with ASO hits identified in Chapter 2 and Chapter 3. CRISPRi produced a moderate level of transcriptional repression that vary across RBP targets. RT-PCR analyses of HRAS and SRSF3 cassette exon splicing suggested these two MYC-dependent splicing events are most responsive to the knockdown of RBM28, SNRPD1, EIF3B, and PINX1.

RESULTS

Multiple RNA binding proteins show reduced mRNA expression upon MYC depletion in a prostate cancer model

Myc has widespread effects on gene regulation including splicing. To assess which RBPs are especially affected by Myc as CRISPRi candidates for our pilot screening, we

performed differential gene expression analysis of the RNA-seq data from three cell lines derived from normal human prostate epithelium that have been transformed with AKT and with Myc under the control of a doxycycline inducible promoter (Phillips et al., 2020). In previous work, these cells were shown to have the gene expression signature of late neuroendocrine stage prostate cancer. Comparing cells plus and minus Dox to modulate the Myc levels, we used DEseq2 (Love et al., 2014) to identify differentially expressed genes and then overlapped these with a list of 454 well-known RBP compiled in earlier studies (Dominguez et al., 2018; Han et al., 2013; Yang et al., 2016). We found 177 out of 454 RBPs were differentially expressed between the MYC off and MYC on states. Among these 159 were downregulated upon Myc depletion and only 18 were upregulated (**Figure 4-1A**). This is in keeping with results indicating that Myc drives a general up regulation of gene expression across the genome (Lin et al., 2012; Nie et al., 2012). The top ten downregulated RBPs sorted by adjusted p-value are shown in a volcano plot (**Figure 4-1B**). These included hnRNP F, while the change hnRNP H1 showed limited significance (**Figure 4-1B**). In this prostate cancer model with Myc overexpression, HNRNPH1 is the most abundance while HNRNPF are less expressed, with mean TPMs ~1100 and ~300, respectively, but hnRNP H2 or H3 are less expressed. Other RNA binding proteins exhibiting large changes in mRNA expression upon Myc depletion included nucleolar factors Nucleolin and RBM28 involved in ribosome biogenesis, MAGOH, a component of the exon junction complex, PINX1 involved in telomerase activity inhibition, NUP35, a part of the nuclear pore complex, EIF3B involved in translation initiation, and PLD6 that presents phospholipase and nuclease activities. Proteins involved in splicing included SNRPD1 and HNRNPAB. On the contrary, a few RBPs showing increases in mRNA expression upon Myc depletion included mRNA decay activator protein ZFP36, polyadenylation factors CPSF4L and CPEB4, and RBM5, a component of the spliceosome A complex and has been reported as a tumor suppressor (Kotlajich and Hertel, 2008).

To assess whether Myc depletion induced changes in protein levels, similar to the changes in mRNA, we performed MYC withdrawal for 16 hours on the three patient-derived ICA lines used in the RNA-seq analysis and analyzed their cell lysates by immunoblot. We confirmed the dramatic downregulation of MYC protein levels (**Figure 4-1C**). However, the changes in RNA binding proteins were smaller and less consistent that was seen by RNA-seq analysis, with the clearest change being a 30% drop in RBM28 protein across three cell lines ($p < 0.05$). hnRNPs H and F showed modest decreases in protein levels that did not have significant p values (>0.05). The lack of correlation between the RNA and protein measurements may result from a longer half life of the proteins compared to the RNA after Myc depletion (Mittal et al., 2009). The Myc mRNA and protein are known to have a short half life of nearly 20 to 30 minutes (Dani et al., 1984; Hann and Eisenman, 1984), whereas the RBPs may be much more stable.

CRISPRi knockdown of RBP targets produce a moderate level of transcriptional repression that varies across targets

We implemented CRISPR interference system (Gilbert et al., 2014; Horlbeck et al., 2016) to screen top Myc-responsive RBPs for potential trans-acting splicing regulators of two Myc-dependent exons. Sequences encoding a transcriptional repression domain, KRAB domain, fused to a catalytic inactive Cas9 protein, with a mCherry marker was subcloned into a pcDNA5-FRT/TO vector (**Figure 4-S1A**). pCRISPRia-v2 plasmid expresses sgRNA under the control of mU6 promoter, with TagBPF and puromycin-resistance as the selection marker (**Figure 4-S1B**). The 20-nt sgRNA protospacer is flanked by two restriction sites, BstXI and BlpI. We annealed a 33-nt forward oligo with a 40-nt reverse oligo encoding protospacer sequences, with 5' and 3' restrictive site overhangs, and ligated annealed sgRNA sequences with double digested pCRISPRia-v2 backbone. This method allowed us to clone 5 sgRNAs for each RBP targets in a relatively efficient and cost-effective manner. We cloned sgRNA targeting 17 genes

in total, along with non-targeting sgRNAs as control. It included 9 of the most Myc-responsive RBPs including HNRNPF, and RBP targets from our ASO hits including 3 HNRNPH paralogs and KHDRBS1. We also incorporated known regulators of HRAS and SRSF3 cassette exon splicing as positive controls for our pilot screening, including SRSF2, PTB, SRSF3 itself and MYC.

We generated a stable HEK293 Flp-In cell line with Doxycycline-inducible expression of dCas9-KRAB fusion protein (**Figure 4-2A**). Flp-In cells were seeded on Day 0 and 5 sgRNAs targeting each gene were co-transfected on Day 1. 24hr post-transfection, we started selection of transfected cells using puromycin and induction of dCas9 expression using doxycycline. We kept culturing the cells and lysed on Day 5 and proceeded with real-time RT-PCR to assay the knockdown efficiency. CRISPRi yielded nearly 70% knockdown in 7 out of 17 targets (**Figure 4-2B**). Another 7 targets exhibited moderate levels of knockdown ranging from 30-60%. For the remaining 3 targets, CRISPRi resulted in slight or even large transcription activation instead of repression. The variation across targets might arise from a number of reasons, such as the differences in transcription rates, abundance of basal level mRNA, sgRNA quality, and indirect effects from potential off-target effects of CRISPR.

We then analyzed the splicing changes of HRAS and SRSF3 exons in responsive to each target by semi-quantitative RT-PCR (**Figure 4-2C**). CRISPRi of SRSF2 resulted in more skipping of HRAS exon (PSI = 7.6) compared with non-targeting control (PSI = 11.5), which is consistent with previous work suggesting SRSF2 functions as HRAS splicing activator (Guil et al., 2003). SRSF3 is known to autoregulate its splicing by inducing poison exon inclusion (Jumaa and Nielsen, 1997), and the downregulation of SRSF3 by 54% resulted in exon skipping (PSI = 1.9) compared with control (PSI = 6.1). Downregulation of MYC induced the inclusion of both HRAS and SRSF3 exons, which is also consistent with previous findings (Phillips et al., 2020; Urbanski et al., 2021). PTBP1 was shown to repress SRSF3 exon inclusion (Guo et al.,

2015), and our analysis suggested slightly more inclusion when PTB was knockdown by 70%. Together, these positive controls suggested our CRISPRi system exhibited reliable RT-PCR readouts when the knockdown worked effectively. sgRNAs targeting RBM28 and SNRPD1 resulted in 81% knockdown, and the downregulation of each factor induced clear exon inclusion of both HRAS (RBM28 PSI = 17.9, SNRPD1 PSI = 15.4) and SRSF3 (RBM28 PSI = 6.9, SNRPD1 PSI = 15.4) exons, compared with controls (NT = 11.5 and 2.7 respectively). In addition, EIF3B also induced more exon inclusion even though it mainly function as translation initiation factor. Knockdown of PINX1, which involved in telomerase activity inhibition, induced skipping of both exons. Taken together, HRAS and SRSF3 cassette exon splicing are most responsive to the loss of RBM28, SNRPD1, EIF3B, and PINX1, and these RBPs would be interesting to investigate in future.

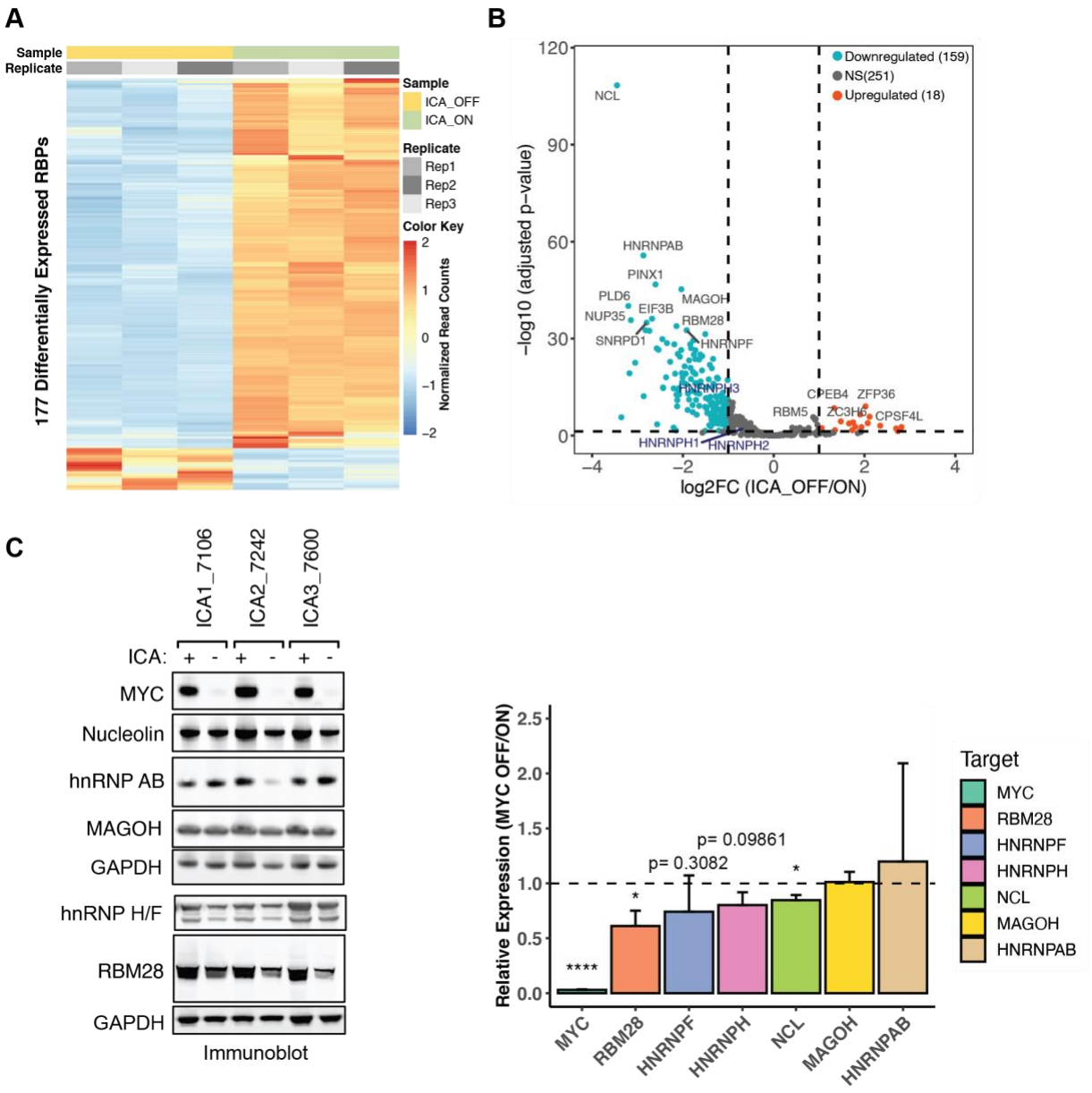

Figure 4-1. RNA binding proteins are largely downregulated with loss of MYC. (A) Hierarchical clustering of 177 differentially expressed RBPs when MYC Off vs. MYC On in three patient-derived prostate cancer lines. Each column represents biological replicate in MYC on or off condition, and each row represents the gene expression of a specific RBP. Significant genes were filtered by read count > 10, adjusted p-value < 0.05, and |log2FC| > 1. ICA, inducible c-Myc and constitutive Akt.

Figure 4-1. Continued.

(B) The volcano plot showing differentially expressed RBPs when MYC off vs. MYC On, with labeling of top ten most MYC-downregulated and top five most MYC-upregulated RBPs sorted by adjusted p-value (black). HNRNPH paralogs were labeled in navy.

(C) Immunoblot showing the protein expression of some top MYC-regulated RBPs in MYC On vs. Off. GAPDH was used as the loading control. The bar graph represents the quantification of immunoblot. Data represent mean +/- SD of three biological replicates. *: $p < 0.05$;****: $p < 0.0001$ (Student's t-test).

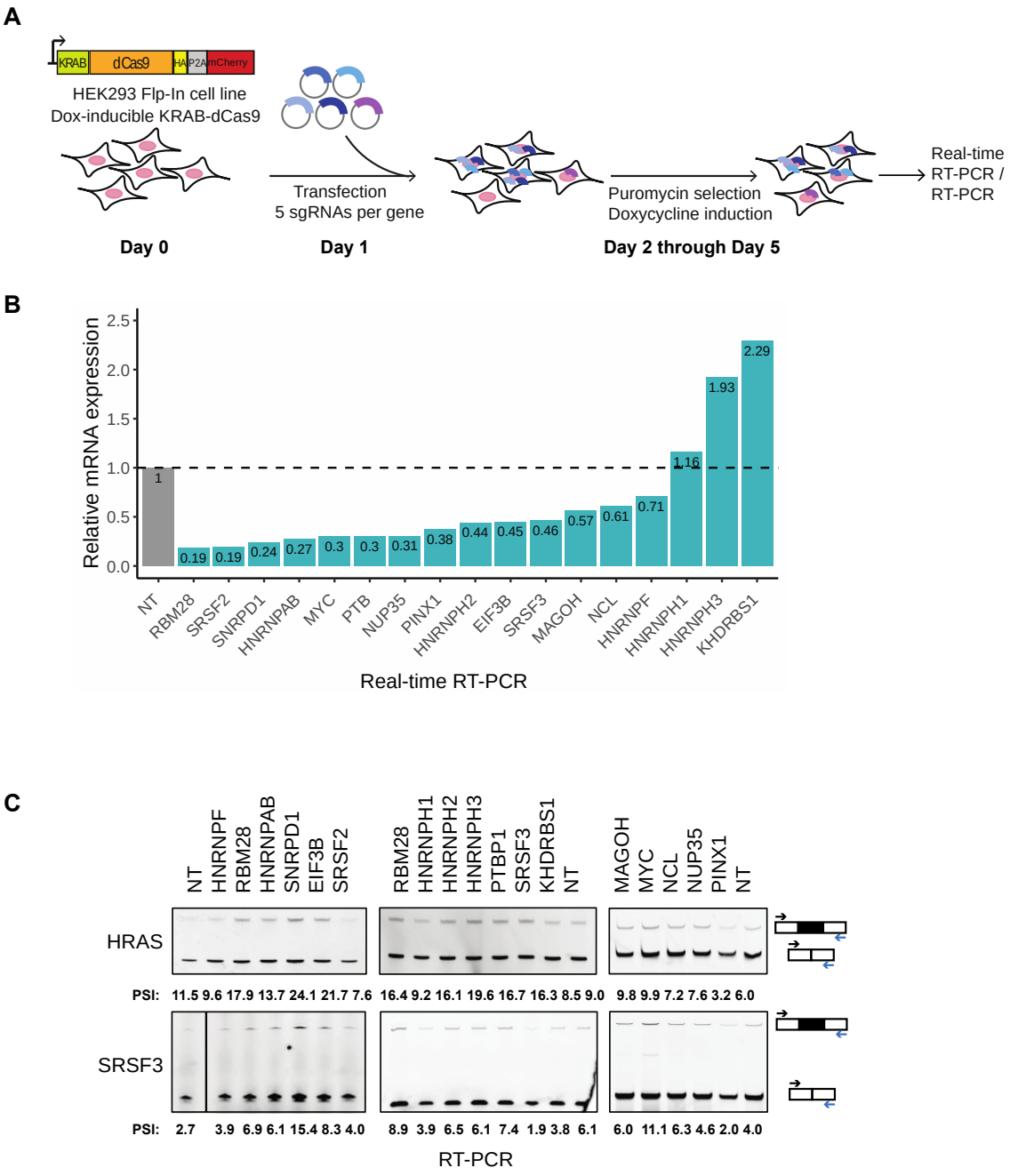

Figure 4-2. CRISPRi induces a moderate level of transcriptional repression that varies across targets. (A) CRISPRi screening work flow. **(B)** Bar plot showing the real-time RT-PCR analyses of RBP mRNA levels in response to the sgRNA-mediated knockdown of each target. NT, non-targeting. **(C)** RT-PCR analyses of HRAS and SRSF3 cassette exon splicing in response to the knockdown of each RBP targets.

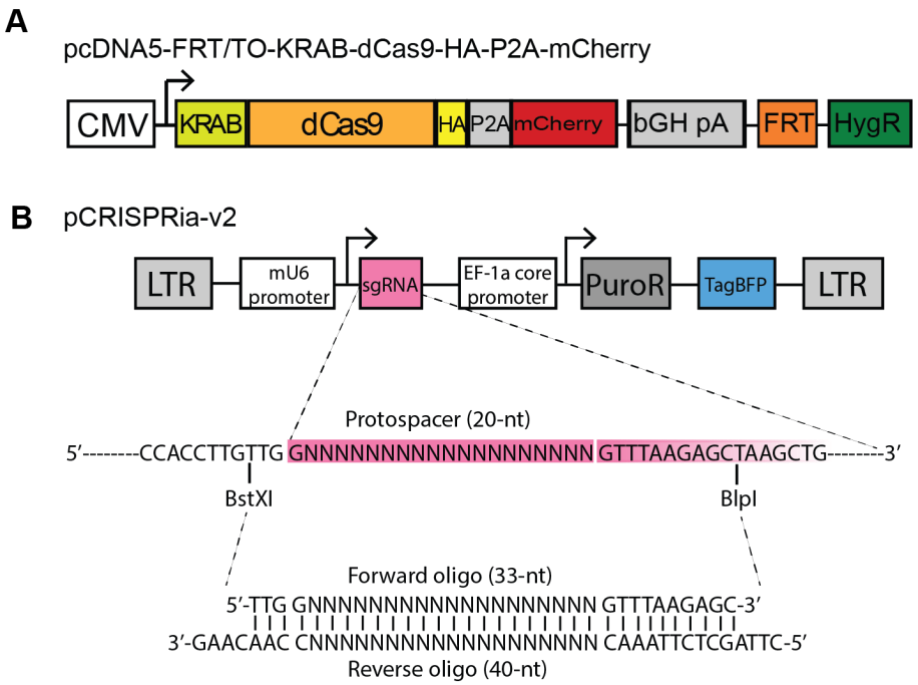

Figure 4-S1. CRISPRi system for transcription repression of RBP targets. (A) Diagram of the construct encoding KRAB-dCas9 protein. **(B)** Diagram of pCRISPRia-v2 construct and sgRNA cloning method.

DISCUSSION

We found MYC depletion largely downregulated multiple RBP mRNA expression in an engineered prostate cancer cell model, with a handful RBPs being upregulated. We implemented a CRISPRi screening system and produced CRISPRi-mediated knockdown of top MYC-responsive RBPs as well as pre-identified HRAS and SRSF3 splicing regulators. Monitored by real-time RT-PCR, CRISPRi yielded moderate level of transcriptional silencing, the effectiveness of which is target-dependent. Semi-quantitative RT-PCR analyses on the splicing of the HRAS and SRSF3 cassette exons suggested the exon usage is most responsive to the downregulation of several RBPs with diverse functions in RNA processing.

A previous study found that majority of differentially expressed splicing factors were upregulated in Myc-active breast tumors, when compared with Myc-inactive breast tumors, in both TCGA datasets and engineered breast cancer cell models (Urbanski et al., 2021). This result is in line with our RBP-focused DEG analysis in Myc-driven prostate cancer cells. Another group performed an RBP domain focused CRISPR knockout screening and revealed RBP dependency in acute myeloid leukemia (AML) (Wang et al., 2019). They further identified RMB39 as an essential splicing factors controlling the mis-splicing of HOXA9 target genes, and pharmacologic inhibition of RBM39 showed efficacy in AML. Following their path, a future direction would be narrowing down our CRISPRi hits, investigating its splicing regulation mechanism, and testing their essentiality in prostate cancer cells. It has also been shown that dysregulation of a translation factor, eIF4E, reprogramed splicing in AML, potentially through eIF4E-mediated increased production of spliceosomal components including SF3B1, U2AF1 and others (Ghram et al., 2022). Interestingly, we also identified a MYC-responsive translation initiation factor, eIF3B, repressed the inclusion of both HRAS and SRSF3 cassette exons. We speculated that given these exon-included transcripts are NMD targets, the perturbation of a translation initiation factor might affect NMD or RNA turnover. It would be interesting to

investigate the mechanism of eIF3B in controlling alternative splicing in Myc-driven prostate cancer.

Our CRISPRi system required improvements from various aspects. We would need to optimize sgRNA sequences to ensure a more stable and uniform knockdown efficiency across targets. We would then expand the scope from our pilot screening to incorporate all canonical RBPs, which would require the transformation of current screening method into an array-based automated platform. Currently, CRISPR-based screening methods are largely in the pooled format and rely on cell growth and drug resistance as readouts, which might introduce false negatives [Ref]. By improving our current method, we hope to establish a reliable array-based platform enabling more sensitive and specific readouts.

METHODS

Cell culture

HEK293 Flp-In cells were maintained in DMEM supplemented with 10% fetal bovine serum. Engineered prostate cancer cells were maintained in stem cell media described in previous study, supplemented with 1ug/ml doxycycline to maintain Myc expression, and Myc withdrawal experiments were done in accordance with previous work (Phillips et al., 2020).

Immunoblotting

Cell lysates were prepared and blotted as described in Chapter 2. Antibodies specifically used in this Chapter include: Anti-Nucleolin (Abcam ab22758), Anti-hnRNP AB (Abcam ab199724), Anti-MAGOH (nvitrogen PA5-80788), Anti-RBM28 (Bethyl Laboratories A303-925A).

RNA-seq data analysis

The RNA-seq fastq files of Myc cell lines (Phillips et al., 2020) were downloaded from Gene Expression Omnibus (accession no. GSE141633) via fastq-dump in SRA toolkit. Alignment was

done by STAR 2.7.6a (Dobin et al., 2013). The genome annotation file was downloaded from GENCODE V38 under human genome version hg38 (GRCh38) (Harrow et al., 2012). Differential gene expression analysis was done using DEseq2 (Love et al., 2014, p. 2) with filtering read count > 10, adjusted p-value < 0.05, and |log2FC| > 1.

CRISPRi

pHR-SFFV-KRAB-dCas9-P2A-mCherry construct (Addgene 60954) (Gilbert et al., 2014) was subcloned into a pcDNA5-FRT/TO backbone. Stable and inducible HEK293 Flp-In cells expressing KRAB-dCas9 was generated according to manufacturer's instructions (Invitrogen K6500-01). Forward and reverse sgRNA oligo sequences were downloaded (Horlbeck et al., 2016) and synthesized (IDT). Annealed oligos with 5' and 3' overhangs were ligated into BstXI/ BlpI digested pCRISPRia-v2 vector (Addgene 84832) (Horlbeck et al., 2016). All of the sgRNA constructs were confirmed by sequencing. KRAB-dCas9 293 Flp-In cells were seeded in 6-well plates on Day 0. On Day 1, five sgRNAs targeting each gene were co-transfected using lipofectamine 2000 (Invitrogen), according to manufacturer's instructions. 24hrs post-transfection, cells were trypsinized and reseeded in 10-cm dishes in complete growth medium supplemented with 1ug/ml doxycycline for dCas9 induction and 3ug/ml puromycin for selecting transfected cells. Cells were maintained in doxycycline- and puromycin-supplemented medium until Day 5 (96hrs post-transfection) and lysed with TRIzol reagent (Life Technologies). RNA isolation, real-time and semi-quantitative RT-PCR were performed as described in Chapter 2. Real-time RT-PCR primers are listed in the supplemental table S4-1.

Table S4-1: List of real-time RT-PCR primers

Target	Sequence 5'-3'
MYC-FWD	GGCTCCTGGCAAAAGGTCA
MYC-REV	CTGCGTAGTTGTGCTGATGT
MAGOH-FWD	TTTATCTGCGTTACTACGTGGGG
MAGOH-REV	CGTCCGGTCGAAACTCAAACT
NCL-FWD	GCACCTGGAAAACGAAAGAAGG
NCL-REV	GAAAGCCGTAGTCGGTTCTGT
NUP35-FWD	GGCCCTTCAGTAGGAGTAATGG
NUP35-REV	CTGGTGGAGCGCCACTTTTA
PINX1-FWD	CCAGCGGATGCTAGAGAAGAT
PINX1-REV	CGAGTCCCAGGTGGTTATTTTTC
HNRNPF-Fwd	CTGCTCTGTTGAGGACGTG
HNRNPF-Rev	CCTGCCCTCTCTAGTGTAGATG
EIF3B-Fwd	GACCGCACTTCCATATTCTGG
EIF3B-Rev	GCACATACGTCTCTGTCCATCT
HNRNPAB-Fwd	AGAAGAAGAACCCGTGAAGAAG
HNRNPAB-Rev	CATACTGCTGCTGCTGATAGA
PTBP1-Fwd	GACGGCATTGTCCCAGATATAG
PTBP1-Rev	CCGAGTTGCTGCTCATGATA
RBM28-Fwd	TTATCGTCCGAGCAGGTATCTA
RBM28-Rev	GCTCCTTTAGAAGGTCCCAATAA
SNRPD1-Fwd	CCAGACAGTTTACCTCTGGATAC
SNRPD1-Rev	CTCTTCCTCTTCCTCTTCCTCT
SRSF2-Fwd	AGGTGGACAACCTGACCTA
SRSF2-Rev	TGTCGTGAAAGCGAACGAA
HNRNPH1-Fwd	GTCCAAATAGTCCTGACACGG
HNRNPH1-Rev	TCCACCGGCAATGTTATCCC

HNRNPH2-Fwd	CGAACCCACTATGATCCCCCT
HNRNPH2-Rev	GCCTCATCCTTTCAAACCCAG
HNRNPH3-Fwd	TGGGAATGATGGCTTTGATGAC
HNRNPH3-Rev	GCACGAAAAGGCAACCCTCT
KHDRBS1-Fwd	GGAGCCAGAGAACAAGTACCT
KHDRBS1-Rev	CATGGCGTGAGTGAAGGAC
SRSF3-Fwd	TGGCAACAAGACGGAATTGGA
SRSF3-Rev	CAAAGCCGGGTGGGTTTCTA

REFERENCES

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315, 1709–1712. https://doi.org/10.1126/science.1138140

Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E.V., van der Oost, J., 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960–964. https://doi.org/10.1126/science.1159689

Dani, C., Blanchard, J.M., Piechaczyk, M., El Sabouty, S., Marty, L., Jeanteur, P., 1984. Extreme instability of myc mRNA in normal and transformed human cells. Proc Natl Acad Sci U S A 81, 7046–7050. https://doi.org/10.1073/pnas.81.22.7046

Dede, M., Hart, T., 2022. Recovering false negatives in CRISPR fitness screens with JLOE (preprint). Bioinformatics. https://doi.org/10.1101/2022.03.22.485359

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635

Doench, J.G., 2018. Am I ready for CRISPR? A user's guide to genetic screens. Nat Rev Genet 19, 67–80. https://doi.org/10.1038/nrg.2017.97

Dominguez, A.A., Lim, W.A., Qi, L.S., 2016. Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation. Nat Rev Mol Cell Biol 17, 5–15. https://doi.org/10.1038/nrm.2015.2

Dominguez, D., Freese, P., Alexis, M.S., Su, A., Hochman, M., Palden, T., Bazile, C., Lambert, N.J., Van Nostrand, E.L., Pratt, G.A., Yeo, G.W., Graveley, B.R., Burge, C.B., 2018. Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. Molecular Cell 70, 854-867.e9. https://doi.org/10.1016/j.molcel.2018.05.001

Esposito, R., Bosch, N., Lanzós, A., Polidori, T., Pulido-Quetglas, C., Johnson, R., 2019. Hacking the Cancer Genome: Profiling Therapeutically Actionable Long Non-coding RNAs Using CRISPR-Cas9 Screening. Cancer Cell 35, 545–557. https://doi.org/10.1016/j.ccell.2019.01.019

Ghram, M., Morris, G., Culjkovic-Kraljacic, B., Mars, J.-C., Gendron, P., Skrabanek, L., Revuelta, M.V., Cerchietti, L., Guzman, M.L., Borden, K.L., 2022. The eukaryotic translation initiation factor eIF4E reprogrammes the splicing machinery and drives alternative splicing. https://doi.org/10.1101/2021.12.07.471666

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

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell 154, 442–451. https://doi.org/10.1016/j.cell.2013.06.044

Guil, S., Gattoni, R., Carrascal, M., Abián, J., Stévenin, J., Bach-Elias, M., 2003. Roles of hnRNP A1, SR Proteins, and p68 Helicase in c-H- *ras* Alternative Splicing Regulation. Mol Cell Biol 23, 2927–2941. https://doi.org/10.1128/MCB.23.8.2927-2941.2003

Guo, J., Jia, J., Jia, R., 2015. PTBP1 and PTBP2 impaired autoregulation of SRSF3 in cancer cells. Sci Rep 5, 14548. https://doi.org/10.1038/srep14548

Han, H., Irimia, M., Ross, P.J., Sung, H.-K., Alipanahi, B., David, L., Golipour, A., Gabut, M., Michael, I.P., Nachman, E.N., Wang, E., Trcka, D., Thompson, T., O'Hanlon, D., Slobodeniuc, V., Barbosa-Morais, N.L., Burge, C.B., Moffat, J., Frey, B.J., Nagy, A., Ellis, J., Wrana, J.L., Blencowe, B.J., 2013. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature 498, 241–245. https://doi.org/10.1038/nature12270

Hann, S.R., Eisenman, R.N., 1984. Proteins encoded by the human c-myc oncogene: differential expression in neoplastic cells. Mol Cell Biol 4, 2486–2497. https://doi.org/10.1128/mcb.4.11.2486-2497.1984

Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., Barnes, I., Bignell, A., Boychenko, V., Hunt, T., Kay, M., Mukherjee, G., Rajan, J., Despacio-Reyes, G., Saunders, G., Steward, C., Harte, R., Lin, M., Howald, C., Tanzer, A., Derrien, T., Chrast, J., Walters, N., Balasubramanian, S., Pei, B., Tress, M., Rodriguez, J.M., Ezkurdia, I., Baren, J. van, Brent, M., Haussler, D., Kellis, M., Valencia, A., Reymond, A., Gerstein, M., Guigó, R., Hubbard, T.J., 2012. GENCODE: The reference human genome annotation for The ENCODE Project. Genome Res. 22, 1760–1774. https://doi.org/10.1101/gr.135350.111

Horlbeck, M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y., Fields, A.P., Park, C.Y., Corn, J.E., Kampmann, M., Weissman, J.S., 2016. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. eLife 5, e19760. https://doi.org/10.7554/eLife.19760

Jost, M., Chen, Y., Gilbert, L.A., Horlbeck, M.A., Krenning, L., Menchon, G., Rai, A., Cho, M.Y., Stern, J.J., Prota, A.E., Kampmann, M., Akhmanova, A., Steinmetz, M.O., Tanenbaum, M.E., Weissman, J.S., 2017. Combined CRISPRi/a-Based Chemical Genetic Screens Reveal that Rigosertib Is a Microtubule-Destabilizing Agent. Molecular Cell 68, 210-223.e6. https://doi.org/10.1016/j.molcel.2017.09.012

Jost, M., Weissman, J.S., 2018. CRISPR Approaches to Small Molecule Target Identification. ACS Chem. Biol. 13, 366–375. https://doi.org/10.1021/acschembio.7b00965

Jumaa, H., Nielsen, P.J., 1997. The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. EMBO J 16, 5077–5085. https://doi.org/10.1093/emboj/16.16.5077

Kampmann, M., 2018. CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. ACS Chem. Biol. 13, 406–416. https://doi.org/10.1021/acschembio.7b00657

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

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., Nureki, O., Zhang, F., 2015. Genome-

scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517, 583–588. https://doi.org/10.1038/nature14136

Kotlajich, M.V., Hertel, K.J., 2008. Death by Splicing: Tumor Suppressor RBM5 Freezes Splice-Site Pairing. Molecular Cell 32, 162–164. https://doi.org/10.1016/j.molcel.2008.10.008

Lin, C.Y., Lovén, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., Young, R.A., 2012. Transcriptional Amplification in Tumor Cells with Elevated c-Myc. Cell 151, 56–67. https://doi.org/10.1016/j.cell.2012.08.026

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

Liu, S.J., Malatesta, M., Lien, B.V., Saha, P., Thombare, S.S., Hong, S.J., Pedraza, L., Koontz, M., Seo, K., Horlbeck, M.A., He, D., Birk, H.S., Jain, M., Olsen, H.E., Akeson, M., Weissman, J.S., Monje, M., Gupta, N., Raleigh, D.R., Ullian, E.M., Lim, D.A., 2020. CRISPRi-based radiation modifier screen identifies long non-coding RNA therapeutic targets in glioma. Genome Biol 21, 83. https://doi.org/10.1186/s13059-020-01995-4

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

Marraffini, L.A., Sontheimer, E.J., 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 322, 1843–1845. https://doi.org/10.1126/science.1165771

Mittal, N., Roy, N., Babu, M.M., Janga, S.C., 2009. Dissecting the expression dynamics of RNA-binding proteins in posttranscriptional regulatory networks. Proceedings of the National Academy of Sciences 106, 20300–20305. https://doi.org/10.1073/pnas.0906940106

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., Zhao, K., Levens, D., 2012. c-Myc Is a Universal Amplifier of Expressed Genes in Lymphocytes and Embryonic Stem Cells. Cell 151, 68–79. https://doi.org/10.1016/j.cell.2012.08.033

Phillips, J.W., Pan, Y., Tsai, B.L., Xie, Z., Demirdjian, L., Xiao, W., Yang, H.T., Zhang, Y., Lin, C.H., Cheng, D., Hu, Q., Liu, S., Black, D.L., Witte, O.N., Xing, Y., 2020. Pathway-guided analysis identifies Myc-dependent alternative pre-mRNA splicing in aggressive prostate cancers. Proc. Natl. Acad. Sci. U.S.A. 117, 5269–5279. https://doi.org/10.1073/pnas.1915975117

Shalem, O., Sanjana, N.E., Zhang, F., 2015. High-throughput functional genomics using CRISPR–Cas9. Nat Rev Genet 16, 299–311. https://doi.org/10.1038/nrg3899

Urbanski, L., Brugiolo, M., Park, S., Angarola, B., Leclair, N.K., Palmer, P., Sahu, S.K., Anczuków, O., 2021. MYC regulates a pan-cancer network of co-expressed oncogenic splicing factors (preprint). Cancer Biology. https://doi.org/10.1101/2021.11.24.469558

Wang, E., Lu, S.X., Pastore, A., Chen, X., Imig, J., Chun-Wei Lee, S., Hockemeyer, K., Ghebrechristos, Y.E., Yoshimi, A., Inoue, D., Ki, M., Cho, H., Bitner, L., Kloetgen, A., Lin, K.-T., Uehara, T., Owa, T., Tibes, R., Krainer, A.R., Abdel-Wahab, O., Aifantis, I., 2019.

Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. Cancer Cell 35, 369-384.e7. https://doi.org/10.1016/j.ccell.2019.01.010

Yang, Y., Park, J.W., Bebee, T.W., Warzecha, C.C., Guo, Y., Shang, X., Xing, Y., Carstens, R.P., 2016. Determination of a Comprehensive Alternative Splicing Regulatory Network and Combinatorial Regulation by Key Factors during the Epithelial-to-Mesenchymal Transition. Molecular and Cellular Biology 36, 1704–1719. https://doi.org/10.1128/MCB.00019-16

Chapter 5

Conclusions

Myc transformation affects the splicing of proto-oncogene HRAS and splicing regulatory protein SRSF3 in many cancers including prostate cancer (Phillips et al., 2020; Urbanski et al., 2021). The HRAS and SRSF3 cassette exons are more skipped in Myc-driven cancer cells in relative to normal tissues. Exon skipping leads to the production of more full-length functional HRAS and SRSF3 proteins(Cohen et al., 1989; Jumaa and Nielsen, 1997). The inclusion of these poison exons results in nonsense-mediated mRNA decay of exon-included transcripts and the production of truncated proteins (Guil et al., 2003; Jumaa and Nielsen, 1997).

Using the antisense oligonucleotide tiling approach, we identified clusters of cis-regulatory elements within the HRAS and SRSF3 cassette exons, as well as their flanking introns. Motif-based RBP prediction suggested hnRNP H/F potentially regulated HRAS splicing. We further confirmed hnRNP H/F activated HRAS splicing through G-runs mapped to the downstream intronic splicing enhancers. hnRNP H expression was found negatively correlated with Myc scores across multiple cancer types, while hnRNP F showed positive correlation with Myc. Thus, the activation of MYC downregulates hnRNP H expression, resulted in the skipping of HRAS cassette exon and increased p21 HRAS oncoprotein expression. Additionally, we identified Sam 68 as SRSF3 splicing repressor, potentially through the silencers in the downstream intron. Lastly, we perform RBP-focused DEG analysis on RNA-seq data generated from an in vitro prostate cancer model carrying inducible Myc expression. We found most of the differentially expressed RBPs were downregulated in response to Myc depletion. CRISPRi-mediated transcriptional repression identified several RBPs that modulated both HRAS and SRSF3 alternatively splicing. Together, our work reveal mechanisms connecting Myc transformation to splicing regulation.

One interesting future direction would be to investigate the functional consequences of full-length HRAS p21 versus truncated HRAS p19 isoforms, whether and how they contribute to cancer phenotypes. Mutations of RAS proto-oncogene have been extensively studied in the past decades (Moore et al., 2020), but it remain unknown if overexpression of unmutated RAS would

control to cancer cell proliferation. It would be interesting to study how HRAS overexpression alters its downstream signaling pathways, such as RAS-MEK-ERK and RAS-PI3K-AKT pathways.

REFERENCES

Cohen, J.B., Broz, S.D., Levinson, A.D., 1989. Expression of the H-ras proto-oncogene is controlled by alternative splicing. Cell 58, 461–472. https://doi.org/10.1016/0092-8674(89)90427-3

Guil, S., de La Iglesia, N., Fernández-Larrea, J., Cifuentes, D., Ferrer, J.C., Guinovart, J.J., Bach-Elias, M., 2003. Alternative splicing of the human proto-oncogene c-H-ras renders a new Ras family protein that trafficks to cytoplasm and nucleus. Cancer Res 63, 5178–5187.

Jumaa, H., Nielsen, P.J., 1997. The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. EMBO J 16, 5077–5085. https://doi.org/10.1093/emboj/16.16.5077

Moore, A.R., Rosenberg, S.C., McCormick, F., Malek, S., 2020. RAS-targeted therapies: is the undruggable drugged? Nat Rev Drug Discov 19, 533–552. https://doi.org/10.1038/s41573-020-0068-6

Phillips, J.W., Pan, Y., Tsai, B.L., Xie, Z., Demirdjian, L., Xiao, W., Yang, H.T., Zhang, Y., Lin, C.H., Cheng, D., Hu, Q., Liu, S., Black, D.L., Witte, O.N., Xing, Y., 2020. Pathway-guided analysis identifies Myc-dependent alternative pre-mRNA splicing in aggressive prostate cancers. Proc. Natl. Acad. Sci. U.S.A. 117, 5269–5279. https://doi.org/10.1073/pnas.1915975117

Urbanski, L., Brugiolo, M., Park, S., Angarola, B., Leclair, N.K., Palmer, P., Sahu, S.K., Anczuków, O., 2021. MYC regulates a pan-cancer network of co-expressed oncogenic splicing factors (preprint). Cancer Biology. https://doi.org/10.1101/2021.11.24.469558