AGO2 overexpression exhibits oncogenic functions in Kras^{G12D}-associated mouse tumor models
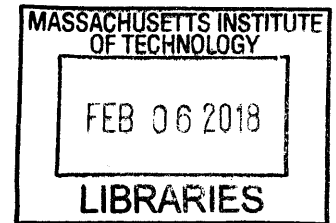
By
Kevin K. Thai

B.S. Bioengineering, University of California, San Diego (2009)

Submitted to the Department of Biology
In partial fulfillment of the requirements for the degree of

Doctor of Philosophy
at the
Massachusetts Institute of Technology

February 2018

Signature of Author:_ **Signature redacted**

Kevin K. Thai
Department of Biology
November 27, 2017

Certified by:_ **Signature redacted**

Tyler E. Jacks
Department of Biology
Thesis Supervisor

Certified by:_ **Signature redacted**

Phillip A. Sharp
Department of Biology
Thesis Supervisor

Accepted by:__ **Signature redacted**

Amy Keating
Department of Biology
Co-Chair, Biology Graduate Committee

Curriculum Vitae
Kevin K. Thai

Education

Massachusetts Institute of Technology, Ph.D. in Biology (2012-present)
Thesis advisors: Dr. Tyler E. Jacks and Dr. Phillip A. Sharp

University of California, San Diego, B.S. in Bioengineering (2005-2009)

Skyline High School (2001-2005)

Research Experience

Ph.D. Dissertation (Biology) – Massachusetts Institute of Technology
Laboratories of Dr. Tyler E. Jacks and Dr. Phillip A. Sharp
Investigating the role of Argonaute-2 in tumor progression using mouse models

Industry Internship (CBSS) – Amgen, Cambridge, MA
Mentor: Qinghong Yan
Characterized transcriptional changes during hepatocyte differentiation using human iPSCs

Laboratory Technician (Biomicro Center) – Massachusetts Institute of Technology
Supervisor: Dr. Stuart Levine
Provided assistance to scientific community in designing and performing experiments to generate high-throughput data using Illumina sequencing platform

Research Technician – Massachusetts Institute of Technology
Laboratory of Dr. Laurie A. Boyer
Profiled the epigenetic landscape during cardiomyocyte differentiation and investigated the role of H2A.Z and Polycomb Group Proteins in chromatin structure using embryonic stem cells

Research Technician – University of California, San Diego
Laboratory of Dr. Trey Ideker
Characterized global transcriptional changes due to genotoxic stresses in yeast and engineered plasmids to constitutively express epitope-tagged proteins

Researcher's Assistant – Scripps Institute of Oceanography
Laboratory of Dr. Dan Cayan
Analyzed climate change data sets and participated in field work

Teaching Experience

MIT 7.02 – Introduction to Experimental Biology & Communication (Fall 2015)

MIT 7.013 – Introduction to Biology (Spring 2014)

Publications
A novel mouse model of ABC-like DLBCL initiated by oncogenic KRAS and overexpression of AGO2
(PNAS – Manuscript ready for submission)

Dicer loss and recover induces an oncogenic switch driven by transcriptional activation of the oncofetal Imp1-3 family
(Genes & Development 2017 - Published)

H2AZ acidic patch couples chromatin dynamics to regulation of gene expression programs during ESC differentiation
(PLOS Genetics 2013 - Published)

Sox2 co-occupies distal enhancer elements with distinct POU factors in ESCs and NPCs to specify cell state
(PLOS Genetics 2013 - Published)

DNA replication stress differentially regulates G1/S Genes via Rad53-dependent inactivation of Nrm1
(EMBO 2012 - Published)

Fellowships, Honors and Awards
School of Science Ludwig Fund for Cancer Research (2015)

Provost Honors (2005-2009)

USTA Michael Cheng Tennis Scholarship (2005-2009)

Gray Davis Academic Scholarship (2005)

Presentations
International Conference on Malignant Lymphomas (Poster 2017)

Koch Institute Retreat (Poster 2017)

Acknowledgements

To Phil and Tyler, I can't thank you guys enough for taking a chance on me as a graduate student. When I look back at the early parts of this journey, I can't help but think about how overwhelming it was learning how to find balance to become a successful researcher in two well-established labs. This has been without a doubt a very humbling experience working for two such accomplished people that I respect and admire. There are many ups and downs that graduate students face but I am so grateful that you two have pushed me out of my comfort zone and made me work hard to find my true potential. Today, I can honestly say that I'm a better critical thinker, most independent that I've ever felt, and believe that I have developed many tools that will make me become a successful scientist. I also appreciate our personal conversations about life and career outside of grad school. It has put a lot into perspective for me and given me confidence in terms of the path that I've chosen after this program.

To my committee members Michael Hemann and Matthew Vander Heiden, I would like to thank you for your insight and guidance in and out of our committee meetings. Your constructive criticisms and advice played a huge role in my growth as a scientist and my ability to produce a body of work that I'm very proud of.

To members of the Sharp and Jacks labs, thank you for providing a work environment that I love coming to everyday and giving me the "balance" that I needed. Both labs have been instrumental in providing me with incredible scientific resources that I would be lost without. My residency in the Sharp lab was a wonderful experience largely due to Courtney, an incredibly genuine person and the best baymate and tennis fan that I could have asked for. Tuomas, Tim and Jesse, I can't thank you enough for making me fall in love with both labs from the get-go. You have showed me how to become a well-rounded scientist by working hard and thinking critically but also enjoying science and sharing all the grief it has to offer over a nice cold brewsky. Garg and Whipple, thank you for your endless support, tennis matches and believing that I'm the happiest person in lab ever (it's easy when I have incredible labmates!). To Leah, Leny, and Nik, thank you for not only our scientific discussions and help with

4

experimental procedures but also your friendship and life advice that have been instrumental in my success as a grad student.

To Stuart Levine, I am very thankful for your continued support and encouragement over the years. When you first hired me, I was fresh out of college and was delighted that I could pay rent. I also wasn't sure if graduate school was right for me but you gave me the freedom to explore my curiosity and believed in my potential every step of the way. Thank you for helping to guide me to where I am today.

To Laurie Boyer, you took a chance on me and gave me the first glimpse of the life of a graduate student. You believed in my ability to succeed early on, allowed me to tackle independent projects and been a source of nonstop support. Your mentorship prepared me well for grad school and made me feel like I belonged at MIT and for that I cannot thank you enough.

To my long time friends from home: Eric, Tedde, Steve, Andrew, Ken and Jenn, you guys have been with me through thick and thin. Yes, I know it's a bit cliché. I can't even begin to describe what your friendships have meant to me. We may not see each other as much as we would like but when we're together, it feels like home for me. You guys have been with me since the very start of this journey and have helped shaped me into the person that I am today. And of course, thank you for the much needed ratchet fun times!

To my core Boston friends: Ian, Dan, Aurelia and Kari, you guys have been my family away from home. Your successes and support have inspired and encouraged me to become the best person that I could be. Thank you for your incredible friendship and being my source of stress relief. I've enjoyed all the wonderful trips we've taken together and frequent dinner and game nights we've had over the years. I've been able to survive grad school knowing that you guys were always there for me. Paul, you've been one of the closest friends I've had. It's great to have a friend that enjoys sports and whiskey as much as I do, and even better that you make me feel normal with my dietary restriction. Brian and Lilly, thank you guys for being genuinely good peeps. I am lucky to have you guys in my life. Both of you inspire me to be better.

Lastly, to my family, I can't thank you enough for the love and support you have each given me. Grandpa and grandma, I could not have done this without either one of you. Your unconditional love and belief in me means more to me than you'll ever know. Thank you for always believing in me, helping me to set high goals and pushing me to be the best person I could be. I hope I've made you both proud. To my parents, thank you for working so hard to provide for our family and instilling me with the work ethic that has helped me get to where I am today. To Allison, thanks for being the best sibling I could ask for. Thanks for being an ear for me to vent to, for understanding the choices that I've made and being someone I always look forward to hanging out with when I visit home. To my uncles Neil and John, I've always looked up to you guys as my big brothers. You guys were one of the first in the family to get a higher education and showed me the possibilities with a strong work ethic and having strong core values. You've always made me feel like I could achieve anything and have given me so much confidence as a person.

Table of Contents

Abstract

Cancer is a disease of normal healthy cells that have accumulated genetic aberrations that contribute to uncontrolled cell divisions. Generally, cancer cells have acquired gain of function mutations in oncogenes that positively promote cell proliferation and growth. Simultaneously, mutations in tumor suppressor genes are frequently detected, allowing cells to evade cell cycle checkpoints, resulting in the inhibition of cell death signals. Therefore, identifying genetic abnormalities that promote tumor initiation and progression is imperative in the development of targeted therapeutics. This thesis focuses on the role of Argonaute-2 in promoting cellular transformation in mouse model systems, highlighting novel oncogenic functions associated with AGO2 overexpression. In short, we have determined that AGO2 overexpression promotes metastasis in an autochthonous mouse model of non-small cell lung cancer while elevated AGO2 levels in B cells contribute to the initiation and maintenance of activated B cell-like diffuse large B cell lymphoma (ABC-like DLBCL), both in the context of KRAS activation and *Tp53* deletion.

Chapter 1

Introduction

In the following introduction, I will summarize our current understanding of the function of the family of mammalian Argonaute proteins as it relates to microRNA activity. Furthermore, I will describe the involvement of AGO2 in tumorigenesis via both microRNA dependent and independent processes. This will build a framework for understanding how the interaction between AGO2 and KRAS may enhance pro-survival signals to promote cellular transformation and highlight the dynamic role of AGO2 in cancer.

Discovery of Argonaute proteins

Argonaute (AGO) proteins are evolutionarily conserved and play a key role in various RNA silencing pathways. These proteins can be found in a diverse spectrum of species, including plants, fungi, yeast, flies and mammals. Though highly conserved, gene duplications during evolution has been responsible for high degrees of variability in terms of *Ago* homologs detected amongst species, ranging from 1 *Ago* gene in fission yeast (*Schizosaccharomyces pombe*) to 27 in the nematode worm (*Caenorhabditis elegans*) (Höck and Meister, 2008). Many of these Argonaute genes have gained specialized functions associated with RNA silencing, such as heterochromatin silencing, post-transcriptional silencing and translational repression (Yigit et al., 2006; Hutvágner et al., 2008). Furthermore, these Argonaute proteins can be phylogenetically divided into two subclasses, Ago and Piwi, where the Piwi clade is restricted to germline (Carmell MA, 2002; Peters and Meister, 2007).

AGO proteins were initially discovered through genetic studies in

Arabidopsis Thaliana, where mutants were screened to identify genes that

affected the premature development of adult leaf traits (Hunter C, 2003). Prior to

this study, it was established that short non-coding RNAs were transcribed and

participated in antisense RNA-RNA interactions to facilitate post-transcriptional

gene silencing. The Ambros and Ruvkun groups found that these short 21-

nucleotide RNAs, later coined as microRNAs (miRNA), were important in

regulating gene expression of developmental genes in *C. elegans*, such as lin-

14, to facilitate proper transition from larval to adult stages by 3'UTR dependent

mRNA repression (Lee RC, 1993; Reinhart et al., 2000).

Further studies of miRNA activity revealed that it is loaded into AGO,

which is part of a ribonucleoprotein protein complex now known as RNA-induced

silencing complex (RISC), that directs gene silencing dependent on the degree of

complimentary sequence between miRNAs and their target mRNAs (Hutvágner

and Zamore, 2002). Together, the discovery of AGO proteins and miRNAs

revealed a novel mechanism of fine-tuning gene expression programs important

for many cellular processes, including differentiation, that was more of a general

phenomenon than initially appreciated.

Structure of Argonaute proteins

AGO proteins are one of the essential components of RISC, where they

function to mediate gene silencing through either mRNA decay and/or

translational repression (Bartel DP, 2009; Ghildiyal and Zamore, 2009). In the

canonical RNA interference (RNAi) pathway, small RNA duplexes are processed

from longer RNA transcripts where they are loaded into AGO that is held in an

open conformation by heat shock protein 90 (HSP90) -mediated ATP hydrolysis.

Once single-stranded RNA (ssRNA) is loaded into AGO, HSP90 and other co-

chaperones leave RISC, resulting in a closed AGO conformation that is activated

for target repression (Meister G, 2013).

In the mammalian system, there are four Argonaute proteins in the Ago

clade. These 99 kDa proteins contain four distinct domains: N-terminal, MID,

PAZ, and PIWI domains (Sasaki T, 2003; Jinek M, 2009). They generally function

sequentially to guide small RNAs to their targets to promote RNA silencing.

The Joshua-Tor group was the first to crystallize the AGO protein from

Pyrococcus furiosus (Song J, 2004). This revealed a bilobal architecture where

the N-terminal and PAZ domain are linked to the MID and PIWI domains. Within

the PAZ domain, there are conserved aromatic residues that are known to bind

to 2nt 3' overhang of small RNAs. The 5' end of guide RNAs was later

discovered to bind to the MID domain through stacking interactions with a

conserved tyrosine residue in addition to an extensive hydrogen bond network

(Miyoshi T, 2016).

The tertiary fold of the PIWI domain resembles the structure of the RNase

H enzyme. However, of the four mammalian AGO proteins, only AGO2 exhibits

endonuclease activity. AGO2 contains the catalytic tetrad DEDX, where X is

usually an aspartic acid or histidine, that is responsible for target mRNA

cleavage, referred to as "slicing" (Swarts et al. 2014). Intriguingly, restoration of

this DEDX tetrad in other AGO proteins is not sufficient to induce slicing activity. Recent studies suggests that the structural components of both the N-terminal and PIWI domains contribute to slicing activity, as replacing the AGO1 N-terminal domain with the AGO2 N-terminal domain activated AGO1 endonuclease activity (Faehnle CR, 2013; Hauptmann J, 2014).

Role of Argonaute protein in microRNA processing and activity

A large percentage of genes in the human genome code for functional proteins, referred to as protein-coding genes (Lander E, 2011). However, a fraction of the genome contains ribosomal RNA genes and regulatory elements, such as promoters and enhancers (Kobayashi T, 2011; Kellis et al., 2014). However, in the last decade, we have come to appreciate that a significant fraction of the genome code for functional RNAs that are not translated into protein, termed non-coding RNAs (ncRNA). This class of RNAs includes long non-coding RNAs (lncRNAs) and miRNAs (Cech and Steitz, 2014).

Well over 2,000 miRNAs have been identified in the human genome that may contribute to coordinating gene expression programs to regulate numerous biological processes (Griffiths-Jones S, 2008). Over 60% of the protein-coding genes of the human genome contain at least one predicted miRNA binding site (Friedman RC, 2009). The biogenesis of miRNAs is a multistep process, where multiple enzymes work sequentially to process long miRNA precursor transcripts into a mature ~22nt guide RNA. These mature miRNAs are loaded into Argonaute, the most downstream component of the RNAi pathway, to activate

target gene silencing via mRNA degradation and/or translational repression (Ha and Kim, 2014).

In the canonical miRNA biogenesis pathway, miRNAs are typically transcribed by RNA Polymerase II (RNAP II) from various regions of the genome, including intergenic/intragenic regions and polycistronic clusters (Saini HK, 2007; Axtell MJ, 2011). These small RNAs (sRNA) are initially transcribed into primary miRNAs (pri-miRNA), where they are 3' poly-adenylated and 5' capped, similar to protein-coding genes. The pri-miRNAs are able to fold upon itself due to extensive complementary sequences that results in a long hairpin structure. These hairpins are processed by the nuclear "microprocessor" complex, minimally composed of DROSHA, an RNase III enzyme, and DGCR8, an RNA-binding protein that guides Drosha-mediated cleavage, into shorter hairpin structures (pre-miRNAs) that are ~70 nucleotides in length (Denli AM, 2004; Kim VN, 2009).

The pre-miRNAs contain a 5' phosphate and '3 2nt overhang as a result of DROSHA-mediated cleavage (Han et al., 2006). This hairpin structure is recognized and exported out of the nucleus by Exportin-5 and RanGTP (Melo et al., 2010). Once exported to the cytoplasm, DICER, another RNase III enzyme, further processes these pre-miRNAs into sRNA duplexes that are ~22nt in length (Hutvágner et al., 2001). TRBP, an RNA-binding protein, interacts with DICER to accurately select the strand that will get preferentially loaded into activated Argonaute, known as the guide strand. The preferential strand selection is based off of thermodynamic stability and 5' end structure of the sRNA duplex (Gu et al.
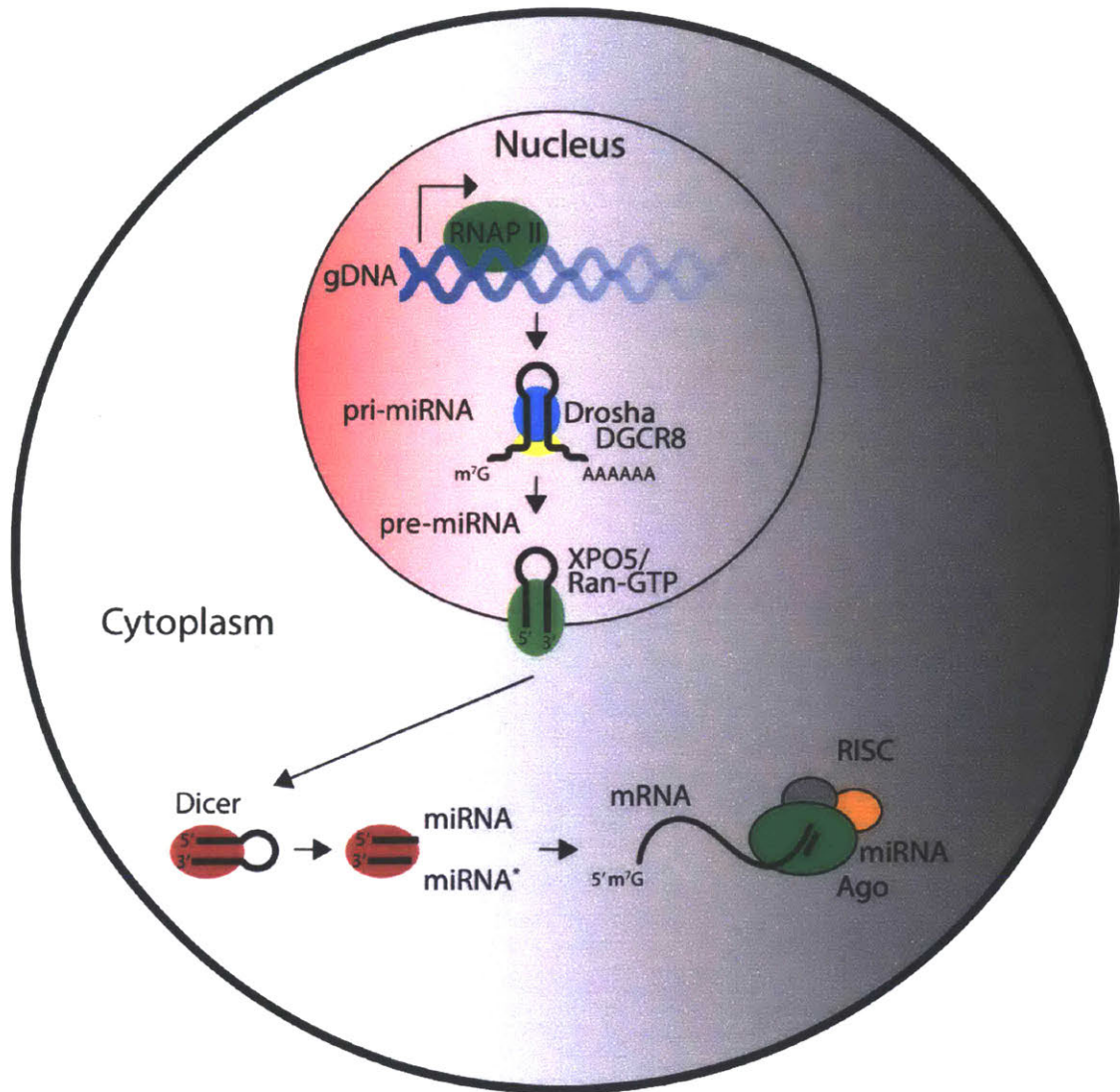
Figure 1: Canonical microRNA biogenesis pathway. Pri-microRNA transcripts are transcribed by RNAP II and are initially processed by Drosha and DGCR8 to produce a ~70nt hairpin structure with a 2nt 3' overhang. This hairpin in exported out of the nucleus by Exportin 5 and is further processed by Dicer to produce a 22nt miRNA duplex, where 1 strand is preferentially loaded into Argonaute to repress target expression.

2001) (Figure 1). Though the mechanism of sRNA unwinding in RISC is still not fully understood, it has been demonstrated that the N-terminal domain of AGO1 and AGO2 display helicase-like activity to facilitate dissociation of the sRNA duplex. This occurs through a conformational change that is induced in the structure of Ago upon transfer of the sRNA duplex, where the N-terminal domain "wedges" between the 2 sRNA strands to initiate unwinding (Wang et al., 2009; Kwak and Tomari, 2012).

Once the guide strand is loaded into an Argonaute protein, the passenger strand (miRNA*) is released for C3PO-mediated degradation and the ribonucleoprotein complex is activated to silence target mRNA transcripts (Liu et al., 2009; Ha and Kim, 2014). Nucleotide positions 2-7, known as the "seed region" of the mature miRNA, forms a helical structure between the MID and PAZ domains of Argonaute proteins that allows for efficient scanning of target sites, which are located primarily at the 3' untranslated region (UTR) of cellular mRNAs (Swarts et al., 2014). Notably, perfect pairing between the miRNA seed region and target site is necessary to promote miRNA binding and activity (Friedman RC, 2009).

The degree of complementation between the mature miRNA and target sites determines the mode of RNA silencing. Perfect pairing results in AGO2-mediated endonucleolytic cleavage of the mRNA transcript, which is dependent on the Mg^{2+} cofactor, while translational repression is favored from imperfect pairing (Schwarz DS, 2004; Pasquinelli AE, 2012). Translational repression is initiated by recruitment of the GW182/TNRC6 proteins to the tryptophan-binding

pockets of AGO. This interaction results in the recruitment of the CCR4-NOT deadenylase complex that destabilizes the mRNA transcript via deadenylation. The mRNA is immediately decapped by DCP2 and is subjected to 5' to 3' degradation by the XRN1 exonuclease (Rehwinkel J, 2005; Jonas S, 2015).

In an alternative model, there is evidence that AGO competes with eIF4E, a translation initiation factor, to repress translation. AGO contains sequence homology to the 5' cap binding sequence of eIF4E and competition between AGO and eIF4E to bind to the m^7G cap structure results in inhibition of ribosome recruitment to initiate translation (Kiriakidou et al., 2007). It's currently unclear how significant the contributions of dislodging eIF4E binding is in the context of mRNA destabilization and how it is coupled to GW182-mediated recruitment of mRNA destabilization complexes.

Regulation of Argonaute proteins

Our current understanding of Argonaute regulation involves AGO2 post-translational modifications and miRNA-dependent stability. The stability of AGO2 is largely dependent on hydroxylation of proline 700 by type I collagen prolyl-4-hydroxylase (C-P4H). The physical interaction between AGO2 and C-P4H was confirmed through mass spectrometry analysis. Missense mutations at proline 700 of AGO2 or genetic ablation of C-P4H in human U2OS cells resulted in reduced AGO2 stability and impaired RISC activity (Qi et al., 2008)

Furthermore, AGO2 turnover has been linked to LIN-41, a heterochronic gene associated with translational regulation in mammals and C. elegans

(Worringer et al., 2014). LIN-41 was demonstrated to function as an E3 ubiquitin ligase and colocalized with AGO2 to P-bodies. *In vitro* studies revealed LIN-41 mediated ubiquitylation of AGO2 in embryonic stem cells (ESCs), which resulted in AGO2 instability. This was confirmed by ectopic expression of LIN-41, which reduced endogenous AGO2 expression, while AGO2 upregulation was detected in LIN-41 loss-of-function mutants. Consequently, miRNA activity was suppressed upon LIN-41 overexpression (Rybak et al., 2009).

MAPK signaling has also been implicated in the regulation of AGO2 function. Phosphorylation of AGO2 at serine 387 is associated with P-bodies localization under normal cellular conditions. Upon activation of p38 MAPK with either sodium arsenite or anisomycin, phosphorylation of S387-AGO2 was significantly induced. However, inhibition of p38 MAPK reduced AGO2 localization to P-bodies and disrupted RISC activity (Zeng Y, 2008). Additionally, stimulation of the EGFR/MAPK signaling pathway was demonstrated to elevate total AGO2 levels in breast cancer cells. In this system, treatment with siEGFR or the U1026 MEK inhibitor significantly reduced AGO2 expression (Adams BD, 2009). Although not tested, this suggests that MAPK mediates phosphorylation of S387-AGO2, resulting in AGO2 P-bodies localization, rationalizing increased AGO2 stability in MAPK activated transformed cells.

Finally, as Argonaute proteins are essential in promoting miRNA activity, it is not surprising that the stability of AGO and miRNAs are interdependent. The half-life of miRNAs was measured in conditions inhibiting transcription, such as treatment with Actinomycin D, using AGO2 depleted cells and compared to wild

type cells. As expected, the half-life was significantly reduced in AGO2 depleted cells but was stabilized upon AGO2 reconstitution (Winter and Diederichs, 2011). Conversely, depleting mature miRNAs in DICER knockout cells destabilized AGO2 expression (Zamudio JR, 2014). Ectopic expression of miR-451, which matures in a DICER-independent mechanism, rescued AGO2 stability, which was not detected using DICER-dependent miRNAs, such as miR-155. This interdependent relationship is not exclusive to AGO2. Using mammalian and *drosophila* systems, knocking out miRNA biogenesis factors, including DICER, results in reduced AGO1 expression as well. Ectopic expression of DICER-dependent miRNAs was demonstrated to rescue AGO1 stability (Smibert P, 2013).

Differences among Argonautes

All mammalian Argonautes are well conserved and have diverged from a single common ancestral gene (Valdmanis et al., 2011). Though these four AGO orthologues are thought to share redundant functions, there are distinct features of AGO2 that are unique from other AGO proteins. At the genetic level, AGO1, AGO3, and AGO4 are arranged in tandem on chromosome 1 in humans while AGO2 is located on chromosome 8. In mice, there's a similar arrangement where AGO1, AGO3, and AGO3 are in tandem on chromosome 4 while AGO2 is on chromosome 15.

There is a high degree of sequence conservation amongst the Argonaute proteins at both the nucleic acid and amino acid level, but the subtle differences

contribute largely to their differences in function and expression (Swarts et al., 2014). For example, AGO1 and AGO2 are highly expressed in most mammalian tissue and exhibit no tissue specificity associated with their expression pattern. Conversely, AGO3 and AGO4 are expressed at significantly lower levels than AGO1 and AGO2 at the mRNA level. Furthermore, AGO4 expression appears to be restricted to germline, suggesting that there are differences in mature miRNA affinity for this particular ortholog that regulate meiosis (Modzelewski AJ, 2012). Important to note is that the study of individual AGO proteins using biochemical techniques has been a challenge, largely owing to the high degree of conservation at the protein level. This has resulted in limited commercially available antibodies with high specificity to individual AGO proteins.

All AGO proteins are able to bind mature miRNAs with partial complimentary sequences to their targets to initiate translational repression. However, there's a higher affinity for mature sRNAs with perfect complimentary sequences to their target binding site associated with AGO2 (Meister et al., 2004). As previously alluded, only AGO2 possesses endonuclease activity that can cleave target mRNAs from perfect pairing between small RNAs and their target binding site. Additionally, there are miRNAs that are processed independent of DICER and are reliant on AGO2 catalysis for processing. The best characterized of this limited subset is miR-451, which is involved in erythropoiesis. Once exported out of the nucleus, the miR-451 pre-miRNA hairpin structure is loaded into AGO2 and is processed into a 22nt duplex to activate RISC activity (Cheloufi S, 2010; Frohn et al., 2012).

Localization of AGO2

Argonaute proteins have been well characterized for their cytoplasmic functions in promoting the activity of miRNAs processed from the canonical pathway. Phosphorylation of AGO2 at serine-387 affects its subcellular localization to processing bodies (P-bodies), which are distinct foci associated with mRNA turnover. Here, other RISC factors, such as GW182 and decapping enzymes, are concentrated to facilitate miRNA-mediated target repression (Parker and Sheth, 2007; Zeng Y, 2008). Moreover, phosphorylation of the highly conserved tyrosine residue Y529 is strongly associated with P-bodies localization as well. It has been demonstrated that mutating this residue to inhibit phosphorylation results in a diffused cytoplasmic AGO expression pattern (Rüdel et al., 2010).

In plants, it is well established that AGO proteins promote cytoplasmic RISC activity but a subset are also shuttled into the nucleus to engage in transcriptional silencing (Meister G, 2013). For example, heterochromatic siRNAs are loaded into cytoplasmic AGO4 prior to nuclear localization, where it engages in epigenetic silencing via RNA-directed DNA methylation (Ye et al., 2012; Schraivogel ad Meister, 2014). Nuclear AGO localization has also been characterized in other eukaryotic organisms, such as flies and worms, but there has been considerable debate about the nuclear presence and function of AGO proteins in mammalian cells.

The UI-Tei group was one of the first to detect the nuclear shuttling of AGO proteins in human cells (Nishi K, 2013). In this study, they found that the

human GW182 paralog, TNRC6A, has both a nuclear export signal (NES) in addition to a nuclear localization signal (NLS). Through its cytoplasmic interaction with AGO proteins, TNRC6A facilitates shuttling of sRNA-bound AGO to the nucleus where it exhibits gene silencing activity. In subsequent studies using HeLa cells, many other RNAi-associated factors were also found in the nucleus through immunofluorescence microscopy and subcellular fractionation assays. These factors, in addition to AGO2 and GW182, included DICER and TRBP. Size exclusion chromatography and coimmunoprecipitation (coIP) from nuclear extracts confirmed that these nuclear RNAi factors form a RISC-like complex and have been demonstrated to silence nuclear RNA transcripts. Moreover, nuclear AGO2 appeared to be deficient in the loading of small RNA duplexes, consistent with prior observations of cytoplasmic sRNA loading. There are also RNAi-associated factors that were not detected in the nuclear extracts, such as HSP90, a chaperone that assists in loading of sRNA duplexes into RISC. This suggests that, in general, RISC factors that are associated with sRNA loading into AGO are restricted to the cytoplasmic space but the minimal components necessary to activate RISC activity are present in both the nucleus and cytoplasm (Gagnon KT, 2014).

Recently, AGO2 association with various intracellular organelles, such as the mitochondria and endoplasmic reticulum membrane, has been detected in human cells. Localization to these organelles has revealed novel AGO2 functions that are currently not well characterized. In the mitochondria, an increase in mitochondrial protein synthesis was evident during muscle cell differentiation.

Here, a significant increase in miR-1 and AGO2 can be detected in the mitochondria during myogenesis. Surprisingly, miR-1 and AGO2 stimulate, as opposed to repress, translation of mitochondrial transcripts using a mechanism that is not well understood. The lack of mitochondrial GW182 and differences in the structure of mitochondrial RNA transcripts may provide insights into the positive, rather than repressive, function of AGO2 in mitochondrial translation (Zhang et al., 2014). Additionally, the interaction between AGO2 and oncogenic factors at the endoplasmic reticulum membrane has been detected in various human cancer cells (Shankar et al., 2016). This study demonstrated novel oncogenic AGO2 functions that are still mechanistically unclear and requires further investigation to characterize the role of AGO2 in malignant cells, which will be discussed in subsequent sections.

AGO2 and microRNAs in cancer

Initial implications of AGO2 promoting tumor progression were in studies of colon cancer in addition to head and neck squamous cell carcinoma (HNSCC) in 2010 (Li L, 2010; Chang SS, 2010). In the study of colon cancer, AGO2 was found to be upregulated in advanced colon tumors compared to adjacent normal tissue. Here, a positive correlation between AGO2 expression and metastasis was detected from immunophenotypic profiling of human tissue. In a separate study, AGO2 DNA amplification and overexpression were detected from human HNSCC tissue samples using comparative genomic hybridization (CGH) arrays and mRNA microarrays, respectively. AGO2 was identified to have a functional

role in the proliferation of HNSCC cell lines, demonstrated by siRNA-mediated knockdown of AGO2, resulting in reduced proliferation rates of the cell lines used. It is unclear from these studies whether modulations in miRNA activity, as a consequence of AGO2 knockdown or overexpression, is the cause of increased aggressiveness in these cancer types.

The mechanism of AGO2 promoting metastasis was further explored from the Han group, where AGO2 overexpression was detected from tumors derived from hepatocellular carcinoma (HCC) patients (Cheng N, 2013). From *in vitro* assays, HCC cell lines exhibited increased cellular proliferation and colony formation as a consequence of ectopic AGO2 expression. Interestingly, although knockdown of AGO2 suppressed colony formation, there was no effect detected with knockdown of DICER. This suggests a miRNA-independent role of AGO2 in the progression of HCC.

Furthermore, transplantation of AGO2 overexpressed HCC cell lines through subcutaneous and tail-vein injections in immunocompromised mice resulted in increased tumor mass and metastases, respectively. Consistent with this, chromatin immunoprecipitation (ChIP) revealed AGO2 was bound to the promoter of focal adhesion kinase (FAK), a cytosolic protein involved in promoting cell motility and metastasis. Here, AGO2 occupancy positively correlated with an increase in FAK transcription. This was the first identified mechanism that directly linked a miRNA-independent functional role of AGO2 in promoting tumor progression.

As perturbations in AGO2 expression affects pathological conditions, it is reasonable to hypothesize that as a consequence miRNA maturation and activity would be affected. To date, the function of many miRNAs have been well characterized with respect to their role in development and cellular processes, such as proliferation and differentiation. A large subset of these miRNAs are dsyregulated in cancer, where there's a selective pressure to promote activity of oncogenic miRNAs and suppress those that stimulate cell cycle arrest and apoptosis, such as the let-7 family of miRNAs (Hayes J, 2014).

In an early miRNA profiling study of a large subset of human cancers using a bead-based flow cytometric approach, the expression of 217 miRNAs were characterized from tumor tissue and compared to normal tissue (Lu et al., 2015). There were two major revelations from this study. First, hierarchical clustering of miRNA expression from all samples revealed that samples were arranged in a very ordered fashion, where samples were grouped based off of their developmental lineage. As an example, samples derived from colon, liver, pancreas and stomach cancers clustered together, which were tissues of embryonic endoderm origin. Second, in comparisons between miRNA expression from tumor and normal tissues, it was evident that a general down-regulation of miRNA expression was detected from tumor tissues compared to their normal counterpart. Although this analysis suggests that miRNA downregulation is common to most cancers, characterization of global miRNA levels must be conducted in a diverse array of cancers to confirm this theory.

Contrary to this study, miRNA expression was profiled from multiple myeloma (MM) patients using whole genome microarray technology and revealed a global increase in miRNA expression when compared to samples derived from healthy patients (Zhou et al., 2010). As MM is a malignancy of atypical plasma cells, this study isolated plasma cells from healthy and MM patients based on CD138 expression, a plasma cell-specific marker, for subsequent experiments. In addition to elevated miRNA levels, AGO2 overexpression was also detected from tumor cells, which was driven by gains in *AGO2* DNA copy number. Knockdown of AGO2 expression in MM cell lines using shRNAs resulted in loss of cell viability and increased apoptosis. Moreover, knockdown of DICER similarly resulted in a decrease in cell viability, suggesting that these tumor cells with elevated AGO2 expression are sensitive to loss of miRNA activity.

Recently, post-translational modifications of AGO2 was correlated with increased metastasis and poor overall survival in breast cancer patients (Shen et al., 2013). This was mediated through epidermal growth factor receptor (EGFR), an oncogene involved in activating signaling cascades associated with cell growth and proliferation. EGFR is frequently misregulated in breast cancer patients, in addition to many other cancer types, where its increased activity can result from gene amplifications, oncogenic mutations or increased transcriptional activity (Sharma SV, 2007). Intriguingly, EGFR is commonly upregulated in hypoxic conditions, where it facilitates activation of hypoxia-inducible factors (HIF) to promote angiogenesis (Salomon DS, 1995; Franovic et al., 2007).

The Hung group detected an increase in endocytosis of EGFR into late endosomal vesicles concentrated in the poorly vascularized core of malignant breast tumors. This resulted in EGFR-mediated phosphorylation of tyrosine residue Y393 of AGO2 prior to lysosomal degradation. This accumulation of p-Y393-AGO2 resulted in reduced DICER-AGO2 interaction. As a consequence, long-loop structured precursor miRNAs, frequently associated with tumor suppressive properties, are not processed into mature miRNAs, contributing to the aggressive nature of this malignancy.

The phosphorylation of Y393-AGO2 has also been positively correlated with oncogenic RAS-induced oxidative stress (Yang et al., 2014). In this context, Protein tyrosine phosphatase 1B (PTP1B) was found to regulate the dephosphorylation of Y383-AGO2, but PTP1B inactivation via H-Ras-induced oxidation led to accumulation of p-Y393-AGO2 and reduced AGO2-DICER interaction, consistent with previous studies. Although untested, it is likely that through HIF repression of PTP1B transcription, there is loss of p-Y393-AGO2 regulation as a result of EGFR upregulation (Suwaki et al., 2011).

Consistent with these findings, AGO2 was previous identified as an essential factor during angiogenesis where it upregulates and stabilizes the expression of pro-angiogenic miRNAs, such as the miR-17/92 cluster of miRNAs, which are often overexpressed in a variety of cancer types to suppress apoptosis (Concepcion CP, 2012; Wu et al., 2014). Overexpressing AGO2 in human umbilical vein endothelial cells resulted in increased cell migration and tube formation *in vitro*. Furthermore, overexpressing AGO2 in multiple myeloma cell

lines *in vivo* accelerated blood vessel formation through stabilization of the pro-angiogenic let-7 family members and miR-17/92 miRNAs while downregulating miR-145, a tumor suppressor miRNA. As angiogenesis is a hallmark of tumor progression, it is conceivable that AGO2 upregulation could enhance tumor progression in various tumor models.

Together, these findings suggest that the consequences of AGO2 overexpression in transformed cells are diverse, which may reflect the mechanism of cellular transformation unique to each cancer type.

Relationship between AGO2 and KRAS

Recent investigation into the protein interactome of RAS-driven cancers revealed a surprising recurrent KRAS interacting protein, AGO2 (Shankar et al., 2016). In this study, the Chinnaiyan group utilized various human pancreatic and lung cancer cell lines to immunoprecipitate RAS, followed by mass spectrometry analysis. This demonstrated that unique AGO2 peptides were detected in all tested cell lines, independent of *KRAS* mutation status, and was validated through coIPs. To determine where KRAS and AGO2 co-localized intracellularly, immunofluorescence microscopy and fractionation assays were performed. Sucrose gradients demonstrated that mutant KRAS co-sedimented with AGO2 and microscopy revealed this complex was concentrated at the endoplasmic reticulum. Moreover, this association was observed at other intracellular perinuclear regions of the cells but the functional differences were not characterized.

The direct association between KRAS and AGO2 occurred at the switch II domain of KRAS, a region where effector molecules bind, and the N-terminal domain AGO2 (Figure 2). More specifically, critical residues that facilitated this interaction include Y64-KRAS, K112-AGO2, and E114-AGO2. Through oncogenic Kras-Ago2 complex formation, a small RNA duplex unwinding assay detected minimal single-stranded RNA molecules, indicative of decreased RISC activity. This suggests that Kras-Ago2 complex formation inhibits miRNA processing. Coincidentally, reduced RISC activity correlated with elevated oncogenic Kras expression and increased PI3K signaling (Figure 5A).

Additionally, this interaction demonstrated promotion of cellular transformation in NIH 3T3 cells, an immortalized mouse fibroblast cell line, from *in vitro* foci formation assays and *in vivo* transplantations to assess tumor growth. Mutations in the critical binding residues of either KRAS or AGO2 reduced the oncogenic potential of KRAS, as observed by impediments in foci formation, supporting the notion that this association promotes cellular transformation. Previously, Y64-KRAS was identified as a critical, but not necessary, residue in activating PI3K signaling to initiate T-lineage acute lymphoblastic leukemia (T-ALL) in mice. This was demonstrated by transducing bone marrow derived T-cells with various Kras mutants (Shieh et al., 2013). Disrupting the interaction between KRAS^{G12D} and effector molecules at this tyrosine residue using second site mutations resulted in a selective pressure for cells to activate PI3K signaling cascades through alternative mechanisms, such as downregulation of PTEN, a negative regulator of PI3K (Stambolic et al., 1998).
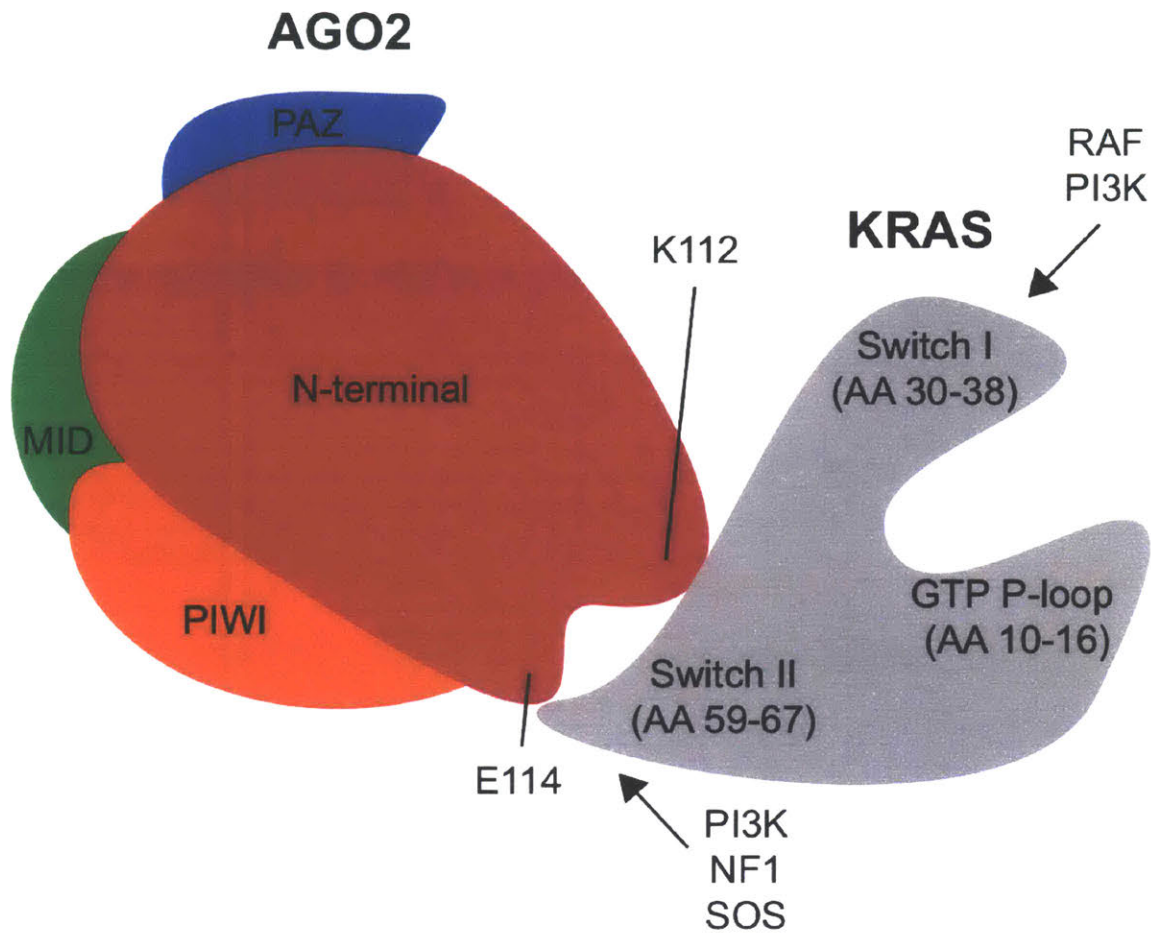
Figure 2: Schematic displaying the regions critical for the physical interaction between human AGO2 and KRAS. Notably, the switch II domain of KRAS binds to other effector molecules, such as PI3K, to activate signaling cascades involved in cell growth and proliferation.

Interestingly, sequencing tumors harboring Y64G-KRAS^{G12D} mutations from this model revealed a recurrent *de novo* insertion in close proximity to the mutated tyrosine residue, suggestive of a novel mechanism to circumvent the inhibition of effector molecule association with KRAS to activate pro-tumorigenic signaling pathways. It is important to note that the initiation of T-ALL was not detected in conditions where only KRAS^{WT} or KRAS^{Y64G} were ectopically expressed.

Discovery/Function of Kras

KRAS is a proto-oncogene that was first discovered in 1967 from studies of acutely transforming retroviruses isolated from rats (Kirsten and Mayer, 1967). Isolation of the Kirsten rat sarcoma virus demonstrated to transform cells in culture and induced initiation of sarcomas in infected animals. It was later determined that the gene responsible for the transforming properties of the sarcoma virus was KRAS, which encodes a 21 kDa GTPase (Shih TY, 1979).

Similar to the other RAS subfamily members, HRAS and NRAS, the KRAS mammalian homolog encodes a GTPase that functions as an early regulator in many signaling transduction pathways controlling cell proliferation, differentiation, growth and more (Dhillon AS, 2007; Pylayeva-Gupta Y, 2011). Genetic alterations in the RAS genes resulting in their constitutive activation are common in many human diseases, such as neurological disorders, but are most frequent in cancer. These three RAS genes encode four distinct Ras proteins, where KRAS4A and KRAS4B are alternative splice variants of the KRAS gene. Although RAS is expressed in all cell lineages, there are differences in isoform

expression that is attributed by tissue-specificity and developmental staging (Castellano and Santos, 2011).

KRAS functions as binary switch that is either active when bound to GTP or inactive when bound to GDP. The conversation between the two states is regulated by the recruitment of guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs) that either activate or inactive KRAS, respectively (Milburn et al., 1990). Various effector molecules, such as phosphoinositide 3-kinase (PI3K) or RAF kinases, bind to either the switch I (amino acids 32-38) or switch II (amino acids 60-67) domains of activated KRAS to initiate downstream signaling cascades (Chen et al., 2013; DK, 2017).

Though KRAS is localized to many intracellular membrane surfaces, it is primarily concentrated at the plasma membrane during steady state conditions (Schmick et al., 2014). This is the result of the farnesylated and positively charged basic residues of the C-terminus of KRAS that increases its affinity for the negatively charged phospholipid heads of the plasma membrane. Here, signals from various cell surface receptors, such as receptor tyrosine kinases (RTKs), tightly regulate KRAS activity in normal physiological conditions, where KRAS serves as a key component in linking extracellular signals to a complex intracellular network of signal transduction pathways with extensive crosstalk (Margolis and Skolnick, 1994; Hancock and Parton, 2005).

Kras-associated signaling cascades

A large subset of effector molecules have been identified that bind to KRAS to activate various signaling cascades that regulate important cellular processes, such as cell cycle entry/progression and cellular differentiation. The mitogen-activated protein kinases (MAPK) are prominent in RAS-mediated signaling, where there are six distinct mammalian MAPK sets that are activated by different stimuli and regulate cellular processes through distinct mechanisms (Zhang and Liu, 2002).

One of the best characterized MAPK signaling pathway is the ERK1/2 pathway, which is estimated to be perturbed in ~1/3 of human cancers (Figure 3). Various stimuli in the form of growth factors and cytokines can activate this pathway, such as the binding of epidermal growth factor (EGF) to its receptor, EGFR. This promotes the activation of KRAS with a GEF, such as SOS1, and results in recruitment of RAF, which is a family of serine/threonine-specific kinases, to the plasma membrane. This initiates the activation of a three-tier kinase signaling cascade that involves phosphorylation of MAPK by MAPKK, which is phosphorylated by MAPKKK (Pierce KL, 2001).

RAF (MAPKKK) is activated by a combination of factors, including protein-protein interaction with KRAS, dephosphorylation of the inhibitory site S259-RAF, and phosphorylation of various activating sites at the N-terminal region, such as S338-RAF (Dhillon AS, 2007). Once activated, RAF phosphorylates MEK1/2 (MAPKK), which then phosphorylates ERK1/2 (MAPK). Activating ERK1/2 activity results in phosphorylation of a large set of transcription factors that
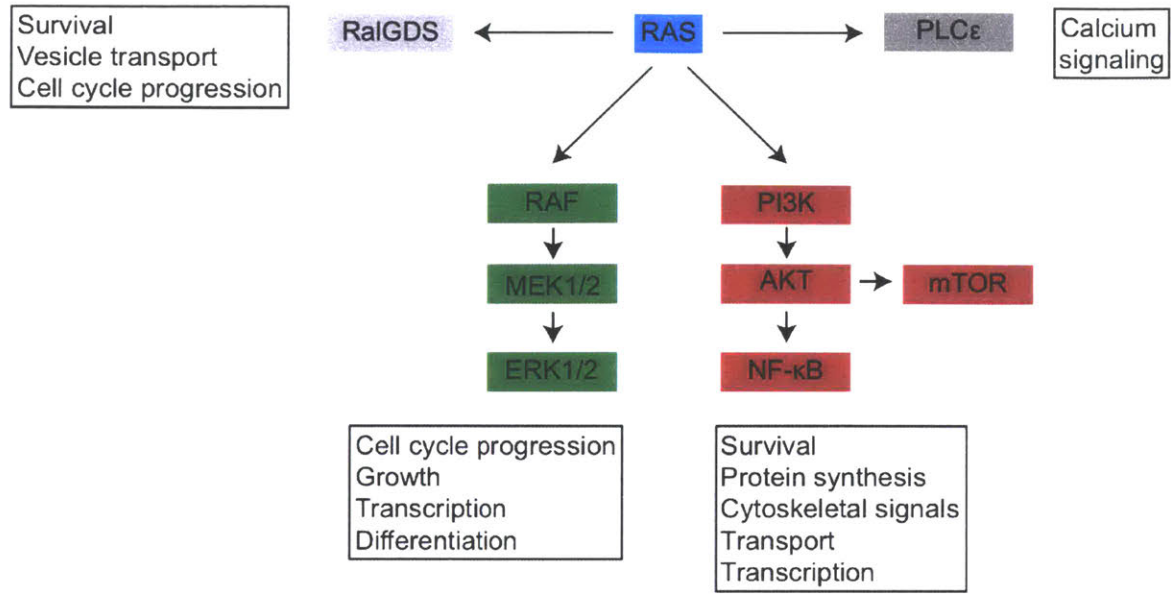
Figure 3: Basic schematic of RAS-associated signaling pathways, activated by various effector molecules including RAF, PI3K, RAL-GDS, and PLCε.

positively promote cell proliferation and growth, such as c-FOS, STAT and ELK1. Phosphorylation of these transcription factors results in the elevated transcriptional response of genes involved in cell cycle entry and progression (Downward J, 2003; Monje P, 2005).

Another main effector pathway of KRAS involves the activity of phosphoinositide 3-kinase (PI3K), a heterodimeric lipid kinase that acts as a signal transducer for a complex network of intracellular signaling pathways, including the mTOR and NF-κB pathways, that regulate cell growth, survival and metabolism (Sulciner et al., 1996; Pylayeva-Gupta Y, 2011). The family of PI3K proteins is divided into 3 distinct classes based off of structure and substrate specificity, where many of the different isoforms have divergent roles in cellular signaling (Mellor P, 2011). Activation of Class I PI3Ks through RTKs or G protein-coupled receptors (GPCR) involves the heterodimerization of the catalytic subunit, p110, and regulatory subunit, p85. The p110 catalytic subunit has three isoforms, each expressed from distinct genes, where the p110α and p110β subunits are expressed in all tissue but p110δ is predominantly restricted to leukocytes. The regulatory subunit, p85, has five isoforms, where three are splice isoforms and the other two are expressed from their own genes (Thorpe LM, 2015). As opposed to the role of Class I PI3Ks in cell signaling, Class II and III PI3Ks have distinct roles in membrane-trafficking but are not as well characterized (Jean and Kiger, 2014).

In normal physiological settings, PI3K activity is repressed through p85 inhibitory signals. When activated by extracellular signals via RTKs or GPCRs,

PI3K is recruited to the plasma membrane where p85-mediated repression is relieved and the p110 subunit interacts with the switch II domain of KRAS. Subsequently, an active kinase signaling cascade is initiated, where PI3K phosphorylates phosphatidylinositol (4,5) (PIP_2), converting it into phosphatidylinositol (3,4,5) (PIP_3). As a result, PIP_3 acts as a secondary messenger by interacting with the pleckstrin homology (PH) domain of AKT, a serine/threonine-specific kinase, to initiate its phosphorylation. Once activated, the kinase activity of AKT phosphorylates downstream effectors involved in the NF-κB and/or mTOR pathways. Furthermore, an additional mode of regulation through PTEN, a tumor suppressive phosphatase, is involved in the PI3K-AKT signaling pathway, where it represses PI3K activity by dephosphorylating the 3' phosphate of PIP_3 in the absence of extracellular signals (Vasudevan et al. 2004).

Nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) activity is stimulated by a diverse set of stimuli, many of which are involved in B cell differentiation and function. There are five mammalian NF-κB subunits: NF-κB1 (p50), NF-κB2 (p52), RelA (p65), RelB, c-Rel. Generally, NF-κB activity is repressed by cytoplasmic inhibitor complexes (IκB) that masks the NLS sequence of mature p50 and p52 subunits, which are processed from p105 and p100, respectively, sequestering these inactive subunits in the cytoplasm. Induction of this pathway involves IκB kinase (IKK) mediated phosphorylation of the IκB complex, resulting in its ubiquitin-dependent degradation. Furthermore, relief of this inhibitory complex promotes p50 heterodimerization with p65/c-Rel

and/or p52 heterodimerization with RelB, which are parts of the canonical and non-canonical NF-κB signaling pathways, respectively. The heterodimerized complexes are able to translocate to the nucleus, where they bind to specific DNA motifs to promote transcriptional activity of genes involved in inflammatory responses, cell survival and proliferation. Additionally, all NF-κB subunits, with the exception of RelB, are able to form homodimers but they are generally associated with repressive functions (Dobrzanski P, 1993; Siggers et al., 2011).

Depending on the source of stimulus, the pathway responsible for NF-κB activation is quite variable, such that numerous ligand:receptor interactions have been identified that are able to activate the NF-κB transcription factors. NF-κB activity induced by way of Ras signaling occurs primarily through the PI3K-AKT signal transduction pathway, where AKT-mediated phosphorylation of downstream effector molecules derepresses NF-κB function (Hussain et al., 2012; Zhang Q, 2017). Thus, it is not surprising that NF-κB misregulation is associated with many pathological conditions, such as cancer.

As demonstrated in a study of KRAS-driven pancreatic cancer using mouse models, oncogenic Ras activity promoted cellular transformation by activating NF-κB (Bang D, 2013). Intriguingly, both the canonical and non-canonical pathways were activated. In this study, oncogenic Kras was demonstrated to upregulate GSK-3α, a serine/threonine specific kinase, which binds and stabilizes the TAK1/TAB complex, an uprstream regulator of IKK. In this instance, IKK activation results in IκB ubiquitination and subsequent proteasome-dependent degradation, promoting p50/p65 activity. Concurrently,

GSK-3α upregulation increased p100 processing into the active p52 subunit in these pancreatic cancer cells, resulting in significant nuclear localization of activated NF-κB subunits, promoting transcriptional activation of anti-apoptotic genes.

The intracellular signaling pathways are stimulated and regulated through intricate mechanisms, involving extensive crosstalk and feedback mechanisms. Therefore, dysregulation of various components can result in activation of the same downstream signaling pathways. For example, in another study of Kras activated pancreatic tumors, induction of mitochondrial reactive oxidative species (mROS) triggers the activation of the ROS-receptive kinase Protein Kinase D1 (PKD1) (Liou et al., 2016). This results in NF-κB activation via PKD1-mediated phosphorylation of the IKK complex, which promotes transcriptional activation of genes involved in cell proliferation and survival, including EGFR.

Mutations in KRAS and TP53 in cancer

p53, often regarded as the guardian of the genome, is a potent tumor suppressor that is activated by various cellular stresses, such as DNA damage. Under normal physiological conditions, p53 protein expression is suppressed by MDM2, which functions as ubiquitin ligase to promote p53 degradation. Stress signals induced in healthy cells repress MDM2 through many mechanisms, including but not limited to DNA damaged induced MDM2 phosphorylation and ARF-mediated sequestration of MDM2 to the nucleolus (Haupt Y, 1997; Weber JD, 1999). Upon activation, p53 functions as a transcription factor, where it

regulates key genes involved in cellular processes such as cell cycle arrest or apoptosis, by promoting transcription of p21 and pro-apoptotic BCL-2, respectively.

Not surprisingly, *TP53* is the most frequently mutated gene in human cancers, where numerous genetic aberrations have been identified that contributes to p53 loss-of-function (LOF) (Kastenhuber and Lowe, 2017). These genetic alterations can range from missense mutations with dominant negative effects to segmental loss of the gene. Generally, p53 deficiency results in genomic instability due to inhibition of key DNA repair pathways and the inability to initiate apoptosis (Murphy and Rosen, 2000). These are key features in the initiation and progression of tumor development.

As *TP53* is frequently mutated in cancer, it is not surprising that these mutations co-occur with oncogenic *KRAS* mutations (Ciriello et al., 2013). The most frequent *KRAS* mutations detected in cancer occur at codons 12, 13 and 61 (Scheffzek et al., 2007; Simanshu DK, 2017). Missense mutations at these codons, with the exception of proline substitutions, result in constitutive activation of KRAS, where hydrolysis of GTP is inhibited. It is thought that mutations at these codons block the interaction between GAPs and Ras, preventing GAP-mediated GTP hydrolysis. Additionally, codon 61 has been demonstrated through biochemical reactions to play a critical role in the intrinsic GTPase activity of KRAS, which supports the increased KRAS activity detected with mutations at this residue. In terms of therapeutics, small molecule inhibitors targeting oncogenic KRAS have been largely ineffective due to the lack of binding

pockets/cavities, with the exception of KRAS^{G12C} mutations (Athuluri-Divakar et al., 2016). Consequently, a larger effort to inhibit oncogenic KRAS signaling has been focused on the development of small molecules targeting RAS effector molecules.

Non-small cell lung cancer (NSCLC), the most common subtype of lung cancer, is an example of a malignancy where KRAS gain-of-function (GOF) mutations often co-occur with TP53 LOF mutations (Skoulidis et al., 2015). Although lung cancer is currently the second most frequently diagnosed cancer, it is leading cause of cancer-related mortalities in the United States (Siegel RL, 2017). Even within the subtype of lung cancer, different combinations of driver mutations have been identified, which speaks to the heterogeneity of cancer in general. As there has been a general shift toward targeted therapeutic approaches to combat cancer, it is imperative to understand the numerous mechanisms that contribute to cellular mechanism for each subtype of cancer.

A feasible approach is the usage of mouse models to engineer genetic alterations detected in human cancers and study their mechanistic and functional consequences in the context of tumor initiation and progression. For example, a widely used "KP" mouse model to study NSCLC involves the deletion of *Tp53* and activation of oncogenic KRAS in a Cre recombinase-inducible system (DuPage M, 2009). This model has been popular due to the striking resemblance to the human form of the malignancy at the genetic, molecular and histological levels. Notably, tumorigenesis can be initiated in this model with oncogenic KRAS activation alone, but ablating p53 function simultaneously significantly

41

accelerates tumor progression, where tumors are more rapidly initiated and display features of advanced stage disease.

Additionally, novel genetic alterations can be introduced to investigate their oncogenic or tumor suppressive effects in the KP background. Because *AGO2* amplifications are often detected in human NSCLC, this KP mouse model can be manipulated to overexpress AGO2 via intratracheal administration of viral particles to constitutively express *Ago2* using an ectopic promoter, in addition to the endogenously expressed *Ago2* gene, to investigate whether AGO2 overexpression promotes cellular transformation (Figure 4) (Cerami et al., 2012).
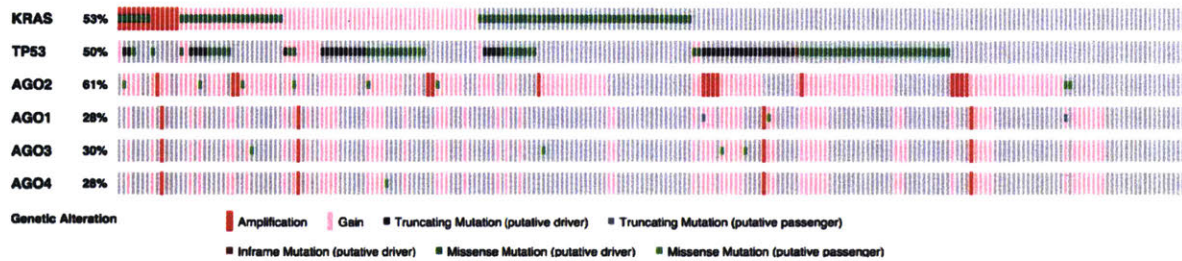

Figure 4: Summary chart displaying mutation status of individual NSLCL patients (Cerami et al., 2012).

TP53 mutations in B cell lymphomas

Mutations in *TP53* are not only prevalent in solid tumors but they also contribute to the malignant transformation of a wide range of liquid tumors, including acute lymphoblastic leukemia (ALL) and diffuse large B cell lymphoma (DLBCL) (Figure 5) (Cerami et al., 2012). It is estimated that nearly one-third of all Burkitt lymphoma patients harbor *TP53* mutations, promoting the hyperproliferative nature of transformed B cells through pro-survival cues (Schmitz R, 2014). Furthermore, inactivating mutations associated with chromatin modifiers, such as EP300 and CREBBP, have been demonstrated to result in defective acetylation-mediated activation of p53 and derepression of *Bcl-6*, a gene critical in B cell maturation. In normal physiological settings, regulated BCL-6 activity is vital in germinal center (GC) reactions by desensitizing GC B cells to somatic hypermutations and class switching during B cell maturation. This provides a permissive environment for DNA remodeling without activating cell cycle arrest (Basso et al., 2009).

To understand the potency of p53-mediated tumor suppression, models of p53 reactivation in established tumors have been used to study the p53 anti-tumor response. The response is highly context dependent but in most cases results in cell cycle arrest, senescence and apoptosis. In Eμ-Myc mouse models of B cell lymphomas, restoration of p53 function in established tumors resulted in massive apoptosis; however, p53-resistant clones are always selected for resulting in relapse (Martins CP, 2006). Understanding the synergistic effects between loss of p53 function and other driver mutations promoting tumor
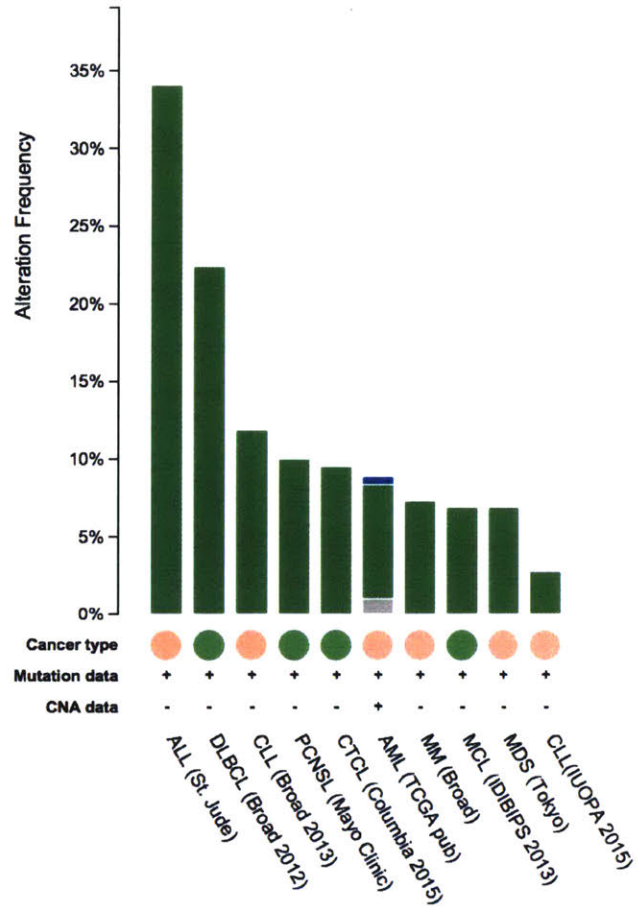
Figure 5: TP53 genetic alteration frequencies from a diverse spectrum of patients with blood-associated cancers (Cerami et al. 2012).

progression and p53-resistance will help optimize p53-based cancer therapies.

B cell development

To understand the complexity of lymphomagenesis, it is crucial to understand the coordinated processes that occur during normal B cell maturation, where a series of checkpoints are in place to prevent the initiation of autoimmune disease or cellular transformation arising from defected B cells. In a healthy adult, hematopoietic stem cells (HSCs) reside in the bone marrow where they engage in self-renewal and differentiation into a variety of immune cells, including B-lymphocytes. B-lymphocytes primarily function in the adaptive immune response, where they can be stimulated by foreign pathogens to differentiate into antibody-secreting plasma cells or facilitate antigen presentation to activate T-cell-mediated clearance of foreign molecules (Melchers F, 2015).

Once committed to the B cell lineage, HSCs initially undergo differentiation into progenitor and precursor B cells, where they gain expression of key B cell surface markers, such as B220 and CD19. Addition, these developing B cells gain expression of IgM, the first class of immunoglobulin to appear on the surface of B cells, displayed as a B cell receptor (BCR) (Nagasawa T, 2006). Immature B cells are assessed for autoreactivity through early differentiation checkpoints. Non-autoreactive B cells then escape the bone marrow into peripheral blood, where they migrate to secondary lymphoid structures, such as the spleen. In the spleen, B-lymphocytes primarily reside in the white pulp space, where they form germinal centers (GC) upon antigen stimulation (Mebius and Kraal, 2005).

45

Here, a very coordinated GC reaction takes place to gain specificity toward the presented antigen. Although MYC, a frequently mutated transcription factor involved in promoting cell cycle progression, is required for the formation of germinal centers, BCL-6 suppresses *MYC* expression during the early onset of GC reactions (Basso and Dalla-Favera, 2015). These proliferative B cells initially reside in the dark zone of the GC, where activation-induced cytidine deaminase (AID) mediates somatic hypermutations (SHM) in the variable regions of the heavy and light chain genes (Muramatsu et al., 2000). To enhance their ability to target presented antigens, B cells migrate from the dark zone to the light zone of the germinal center, where they lose their proliferative capacity (Figure 10). Here, AID mediates class switch recombination (CSR) of the heavy chain gene, from IgM to either the IgG, IgA, or IgE isoforms. Class switching involves rearrangement of the constant region of the heavy chain locus, which does not affect the antibody specificity to antigen, but alters the mechanism of action.

With the guidance of helper T-cells, a checkpoint during the light zone reaction assesses the overall antigen recognition capacity of B cells that have undergone SHM and CSR. If further SHM are necessary to enhance recognition to antigen, MYC expression is induced in a subset of light zone B cells to gain re-entry into the dark zone. B cells with weak affinity toward antigen are signaled to undergo cell death processes, such as apoptosis (Mayer et al., 2017). Once antigen recognition is considered ideal, mature B cells are signaled to exit the germinal center through NF-κB activation. Mature B cells destined to further differentiate into either memory B cells or plasma cells migrate into the red pulp

46

region of the spleen, primarily occupied by monocytes and macrophages, to filter out antigens and cellular debris (Bronte and Pittet, 2014).

As mature B cells exit GC reactions to engage in terminal differentiation into plasma cells, a drastic switch in the transcriptional program is induced, where many general B cell markers are downregulated and key plasma cell markers are induced. Initially, NF-κB activates MUM1, a splice isoform of IRF4, in immature plasma cells, termed plasmablasts (Shapiro-Shelef and Calame, 2005). Plasmablasts are highly proliferative, retain the ability to present antigens to T-cells and capable of producing and secreting large volumes of antibody albeit less than mature plasma cells. The transition from plasmablast to mature plasma cell requires expression of BLIMP1, the master regulator of plasma cell differentiation. Normally, BCL-6 transcriptionally represses *BLIMP1*. Many mechanisms have been described to relieve BCL-6-mediated repression in this context, such as MUM1 suppression of BCL6 or MAPK-mediated BCL-6 phosphorylation, marking BCL-6 as a target for ubiquitin/proteasome-mediated degradation (Niu H, 1998). With BLIMP1 activation in plasma cells, a distinct transcriptional profile separates these cells apart from their precursors, where signatures associated with cell proliferation and germinal center functions are depleted (Shaffer et al., 2002).

It is evident that B cell development is a tightly orchestrated progression involving multi-tiered regulators and stimuli. Maturation of B cells is far from being a unidirectional process and involves many checkpoints to prevent progression of abnormal or low-affinity B cells. Dysregulation of signaling pathways involved in

B cell development can result in developmental arrest, thereby promoting

uncontrolled cell proliferation in a specific B cell state. Thus, the onset of

lymphomagenesis can give rise to a variety of lymphomas, dependent on the cell

of origin and mutation status of the transformed cell (Figure 6).

Role of AGO2 in hematopoiesis

As previously alluded, miRNAs play a major role during development by

regulating heterochronic genes. Impairments in miRNA biogenesis induced in

hematopoietic systems have demonstrated to trigger differentiation defects using

bone marrow cells derived from mice (O'Carroll et al., 2007). From studies of

reconstituting bone marrow cells into lethally irradiated mice, depletion of AGO2

in transferred bone marrow resulted in differentiation of cells into many

hematopoietic lineages *in vivo*, but drastic defects were noticeable from both

erythroid and B lymphoid lineages. FACs analysis of cells derived from various

lymphoid tissues revealed that B cells were unable to differentiate into precursor

B cells. As a consequence, this impaired their ability to migrate into peripheral

blood and migrate to secondary lymphoid tissue. This is not surprising as many

miRNAs have been identified to be essential in hematopoiesis, such as miR-150

and miR-155, and depleting AGO2 would presumably diminish the function of

those miRNAs (Chen CZ, 2004; Rodriguez et al., 2007). Retroviral restoration of

AGO2 expression in the AGO2-depleted bone marrow cells rescued the

differentiation defects observed from the $Ago2^{-/-}$ transplantation, confirming

AGO2-dependency for proper B cell maturation.
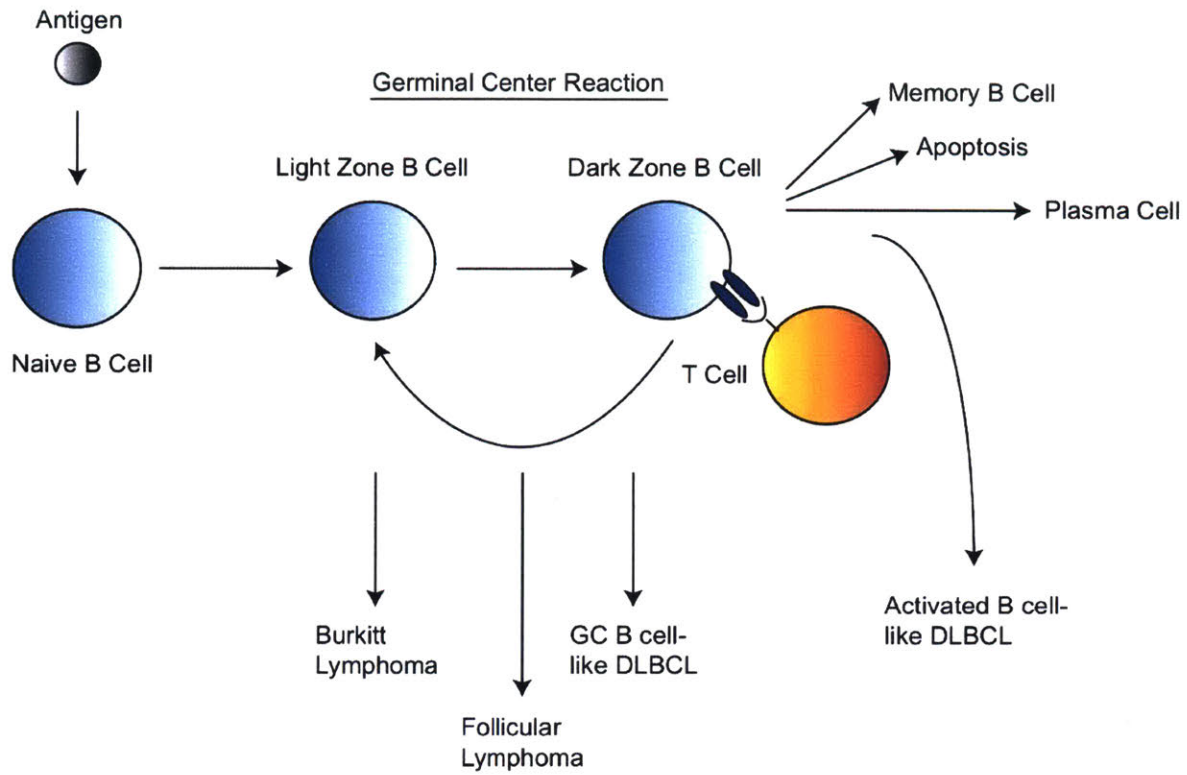
Figure 6: Schematic of B cell development. Also shown is the origin of each subtype of B cell lymphoma.

Interestingly, the transplantation of bone marrow cells engineered with slicer-deficient *Ago2* mutations, which results in the loss of endonucleolytic activity, recapitulated the results from transfer of wild-type *Ago2* bone marrow cells. In both conditions, defects in erythroid and B cell development were not observed. Additionally, transplantation of bone marrow cells with either *Ago1* or *Ago4* ablation resulted in mice surviving to adulthood without any noticeable developmental defects (Martinez and Busslinger, 2007). This argues that silencing of target transcripts through AGO2-mediated translational repression and/or mRNA destabilization is essential to promote proper B cell maturation but slicing activity is dispensable.

Genetically engineered mouse models

Cancer is the second leading cause of deaths worldwide. Although we have major strides in cancer therapeutics, ranging from non-specific combination chemotherapy regimens to targeted small molecule therapies, the fact that all cancers are incredibly heterogeneous in terms of their genetic aberrations and mechanisms of cellular transformation reflect the large variability in patient responses to each form of therapy. One of the most effective approaches in studying the initiation and progression of specific tumor types is the usage of genetically engineered mouse models (GEMMs) (Talmadge JE, 2006).

GEMMs offer many advantages in studying human cancers due to their extensive genomic similarities to the human genome and anatomical organization that resembles the human body plan. Additionally, the life cycle of

mice makes them conducive to mammalian studies as their gestation period is roughly 3 weeks and each female gives birth to an average litter size of 8, depending on genetic and environmental factors. The small size of the laboratory mouse, mus *musculus*, makes it manageable to house large quantities of cohorts, which is advantageous in studies requiring large numbers of replicates and conditions (Frese and Tuveson, 2007).

Over the years, there has been increased sophistication with generating mouse models to improve many aspects of cancer studies. Features that are constantly improved on include initiating tumors that recapitulate the human form of the disease on the genetic, molecular and histological levels and developing tools to accelerate the process of engineering genetic mutations, such as optimization of CRISPR/Cas9 technology for *in vivo* studies (Cong et al., 2013; Xue et al., 2014). Recently, large efforts to optimize the CRISPR/Cas9 genome editing system have resulted in much more rapidly produced mouse models that are highly specific in terms of targeted genomic editing. This system utilizes two specific elements: guide RNAs (gRNAs) that contain complimentary sequences to specific genomic regions and an endonuclease, Cas9, that co-localizes with the gRNA to introduce a cut in the targeted DNA sequence. This activates the intrinsic DNA machinery to repair the nicked DNA, resulting in genetic alterations such as indels via error-prone non-homologous end joining (NHEJ) (Sánchez-Rivera et al., 2014).

There are two main classes of GEMMs, the endogenous and transgenic models. In the endogenous model, mutations introduced in mouse embryonic

stem cells (ESCs) can be engineered to express dominant-negative forms of tumor suppressor genes, LOF variants, or oncogenes through the endogenous promoter (Figure 7A). These genetically altered ESCs can then be transferred into blastocysts, which are subsequently transplanted into pseudopregnant female mice. This process gives rise to chimeric mice, allowing for further crosses to generate mice with desired genotypes (Eakin and Hadjantonakis, 2006; Longenecker and Kulkarni, 2009). However, the disadvantage with this system is that the germline alterations will likely be constitutively expressed systemically, which can result in developmental defects or embryonic lethality.

The advent of the conditional system relieves many of these concerns by allowing for spatial and temporal control over the expression of the genetically modified genes. One of the most widely used conditional models utilizes the Cre-Lox system, which takes advantage of inverted DNA repeat elements (LoxP sites) that are recognized by Cre recombinase (Figure 7B) (Orban PC, 1992). LoxP sites are engineered between DNA sequences of interest, allowing for Cre recombinase-mediated catalysis, resulting in either deletion or inversion of the intervening sequences depending on the orientation of the LoxP sites. For instance, to prevent systemic activation of the oncogenic $Kras^{G12D}$ allele, a STOP cassette engineered between the promoter and first exon of $Kras^{G12D}$ inhibits transcription of this oncogenic allele (DuPage M, 2009). Furthermore, LoxP sites oriented in the same direction can be introduced to flank the STOP cassette (LSL), allowing for Cre-mediated deletion, resulting in constitutive activation of the mutant $Kras$ allele. Notably, mice can only be heterozygous for the
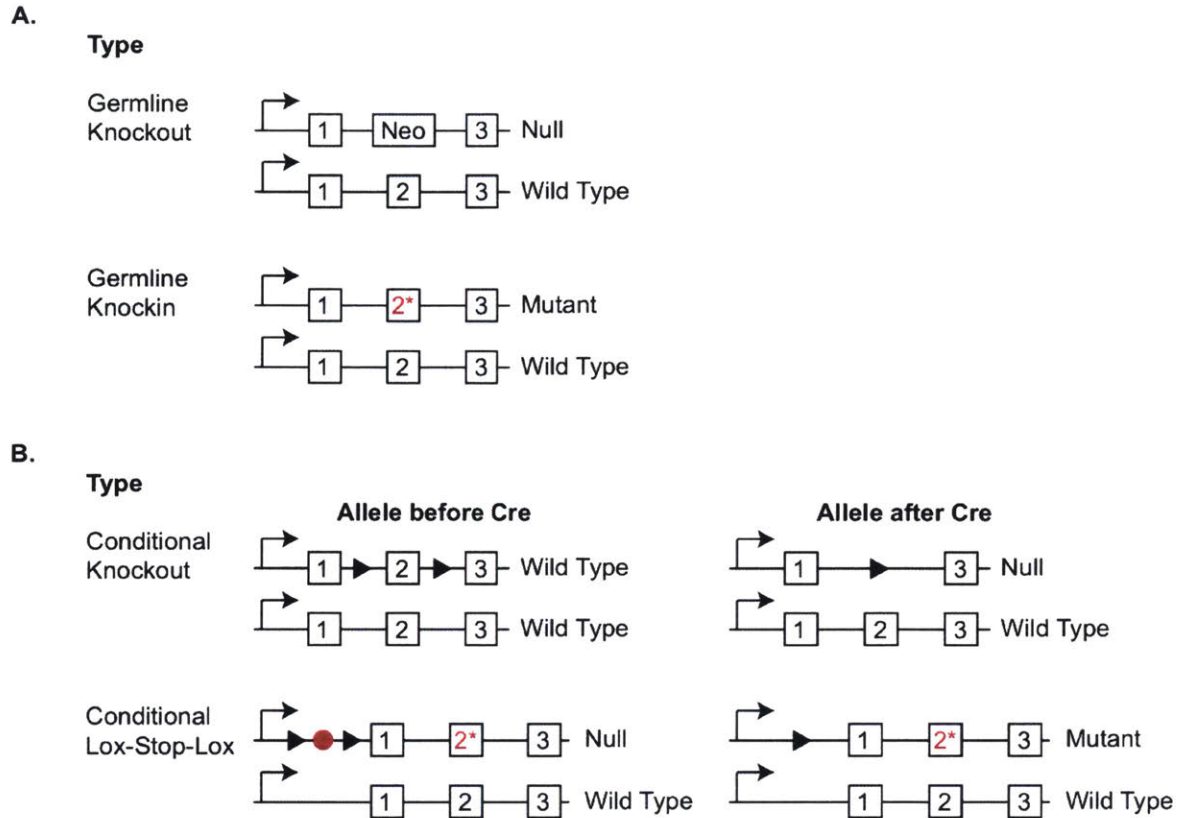
Figure 7: Endogenous mouse models. (A) Germline endogenous models where selectable markers replace exons to generate knockouts or point mutations are introduced into exons using knockin alleles. (B) Conditional alleles are used to express mutant endogenous genes in a tissue specific and temporally regulated manner. This is usually accomplished by integrating DNA motifs recognized by specific recombinases to recombine DNA sequences. Arrowheads represent LoxP sites, circle represents STOP cassette and asterisks denote mutations.

53

oncogenic $Kras^{LSL-G12D}$ allele since $Kras$ is an essential gene during embryonic development. This recombineering technique is also applicable to ablating tumor suppressor gene function. In many cases, this is achieved by introducing homozygous loxP sites flanking exons encoding regions critical to protein function or resulting in nonsense-mediated mRNA decay (NMD) (Möröy and Heyd, 2007).

Cre recombinase can be induced to mediate recombination of loxP sites using multiple methods, but it is commonly expressed using an ectopic promoter introduced by way of viral infection, a key feature in many transgenic mouse models. Alternatively, Cre recombinase can be introduced into the mouse germline, where its expression is regulated by a tissue specific promoter or drug-inducible promoter (Lewandoski M, 2001).

In the second class of GEMMs, ectopic promoters are used to express genes of interest, including site-specific recombinases, oncogenes or dominant-negative tumor suppressor genes (Figure 8A). These transgenes are often introduced into mouse models by way of viral infections using either lentiviruses or adenoviruses (Merentie et al., 2016). Transgenes can be expressed using a variety of constitutive promoters, including but not limited to pGK and CMV, tissue-specific promoters, including lung specific SPC, and inducible promoters, such as tetracycline-inducible systems.

There are various forms of the tetracycline-inducible promoter, such as the Tet-on and Tet-off systems (Figure 8B) (Gossen et al., 1995). In the Tet-off system, a tetracycline transactivator (tTA) protein is expressed, where it binds

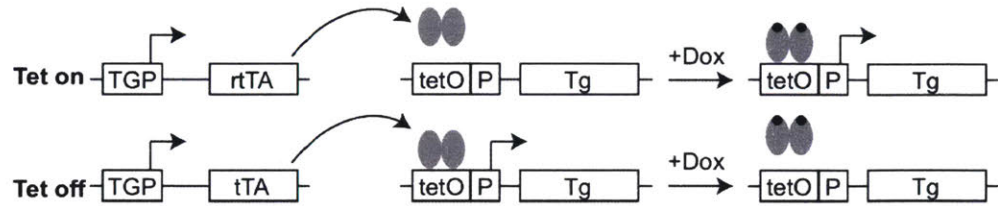**Figure 8:** Transgenic GEMMs. (A) Transgenes that are driven by a constitutive promoter (TGP) are introduced via germline or random viral integration. (B) Expression of transgene (Tg) is temporally regulated using an inducible promoter (P), such as the Tet-on or Tet-off systems.

the tetracycline responsive element (TRE), composed of Tet operator (TetO) sequences followed by a minimal promoter, that is placed upstream of the gene of interest. Here, tTA binds to the TetO sequence resulting in transcriptional activation of the gene. Transcription of this gene is suppressed in this system by addition of doxycycline, a tetracycline derivative that binds to tTA, thereby inhibiting tTA binding to the TRE sequence.

The Tet-on system is more commonly used, where it takes advantage of a reverse tetracycline-controlled transactivator (rtTA), which does not bind to TRE sequences in the absence of doxycycline. This inducible system is a powerful tool in cancer studies, as it allows for temporal activation or inactivation of specific genes to investigate its role in many aspects of tumor development, such as tumor initiation and progression.

An additional type of mouse model involves the transplantation of genetically engineered cancer cells into immunocompromised mice through either subcutaneous or intravenous administration (McClatchey AI, 1999). Subcutaneous injection of tumor cells into the flanks of mice assays for perturbations in tumor progression by comparing various conditions that are altered either genetically or by drug administration. Here, comparisons of tumor size, histological features and presence of metastases are used to assess the degree of tumorigenicity in each condition.

The mouse model involving intravenous administration of tumor cells is frequently utilized in studying various core features of metastasis in addition to lymphomagenesis. The features of metastasis commonly analyzed with this

assay include the ability for tumor cells derived from solid tumors to circulate, extravasate out of blood vessels and colonize distant organs. Furthermore, tail vein injection of genetically altered lymphocytes is a method of assessing variations in B- or T cell development, as well as their ability to initiate lymphomagenesis (Reddy P, 2008). In many instances, this method requires the usage of immunocompromised mice to prevent graft rejection. Commonly used immunocompromised mice include "nude" mice that carry germline genetic alterations that ablate thymic function or irradiated mice. The hematopoietic system is always mitotically active, resulting in hypersensitivity to irradiation treatment.

References:

Adams BD, Claffey KP, and White BA. 2009. Argonaute-2 Expression Is regulated by Epidermal Growth Factor Receptor and Mitogen-Activated Protein Kinase Signaling and Correlates with a Transformed Phenotype in Breast Cancer Cells. *Endocrinology* **150:** 14-23.

Athuluri-Divakar SK, Carpio RV, Dutta K, Baker SJ, Cosenza SC, Basu I et al. 2016. A Small Molecule RAS-Mimetic Disrupts RAS Association with Effector Proteins to Block Signaling. *Cell* **165:** 643-655.

Axtell MJ, Westholm JO, and Lai EC. 2011. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. *Genome Biol* **12:** 221.

Bang D, Wilson W, Ryan M, Yeh JJ, and Baldwin AS. 2013. GSK-3α Promotes Oncogenic KRAS Function in Pancreatic Cancer via TAK1-TAB Stabilization and Regulation of Non-Canonical NF-κB. *Cancer Discov* **3:** 690-703.

Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. *Cell* **136:** 215-233.

Basso K, Saito M, Sumazin P, Margolin AA, Wang K, Lim WK et al. 2009. Integrated biochemical and computational approach identifies BCL6 direct target genes controlling multiple pathways in normal germinal center B cells. *Blood* **115:** 975-984.

Basso K and Dalla-Favera R. 2015. Germinal Centres and B cell Lymphomagenesis. *Nat Rev Immunol* **15:** 172-184.

Bronte V and PIttet MJ. 2014. The spleen in local and system regulation of immunity. *Immunity* **39:** 806-818.

Carmell MA, Xuan Z, Zhang MQ, and Hannon GJ. 2002. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes & Dev* **16:** 2733-2742.

Castellano E and Santos E. 2011. Functional Specificity of Ras Isoforms. *Genes Cancer* **2:** 216-231.

Cech TR and Steitz JA. 2014. The Noncoding RNA Revolution- Trashing Old Rules to Forge New Ones. *Cell* **157:** 77-94.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E et al. 2012. The cBio cancer genomics

portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**: 401-404.

Chang SS, Smith I, Glazer C, Hennessey P, and Califano JA. 2010. *EIF2C* Is Overexpressed and Amplified in Head and Neck Squamous Cell Carcinoma. *ORL* **72**: 337-343.

Cheloufi S, Dos Santo CO, Chong MM, and Hannon GJ. 2010. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**: 584-589.

Chen CZ, Li L, Lodish HF, and Bartel DP. 2004. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**: 83-86.

Chen C, Er T, Liu Y, Hwang J, Barrio MJ, Rodrigo M et al. 2013. Computational Analysis of KRAS Mutations: Implications for Different Effects on the KRAS p.G12D and p.G13D Mutations. *PLoS One* **8**: e55793.

Cheng N, Li Y, and Han Z. 2013. Argonaute2 Promotes Metastasis by Way of Up-regulating Focal Adhesion Kinase Expression in Hepatocellular Carcinoma. *Hepatology* **57**: 1906-1918.

Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, and Sander C. 2013. Emerging landscape of oncogenic signatures across human cancers. *Nat Genet* **45**: 1127-1133.

Concepcion CP, Bonetti C, and Ventura A. 2012. The miR-17-92 family of microRNA clusters in development and disease. *Cancer J* **18**: 262-267.

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819-823.

Denli AM, Tops BB, Plasterk RH, Ketting RF, and Hannon GJ. 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**: 231-235.

Dhillon AS, Hagan S, Rath O, and Kolch W. 2007. MAP kinase signaling pathways in cancer. *Oncogene* **26**: 3279-3290.

Dobrzanski P, Ryseck R, and Bravo R. 1993. Both N- and C-Terminal Domains of RelB Are Required for Full Transactivation: Role of the N-Terminal Leucine Zipper-Like Motif. *Mol Cell Biol* **13**: 1572-1582.

Downward J. 2003. Targeting RAS signaling pathways in cancer therapy. *Nat Rev Cancer* **3**: 11-22.

DuPage M, Dooley AL, and Jacks T. 2009. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nat Protoc* **4:** 1064-1072.

Eakin GS and Hadjantonakis A. 2006. Production of chimeras by aggregation of embryonic stem cells with diploid or tetraploid mouse embryos. *Nat Protoc* **1:** 1145-1163.

Faehnle CR, Elkayam E, Haase AD, Hannon GJ, and Joshua-Tor L. 2013. The Making of a Slicer: Activation of Human Argonuate-1. *Cell Reports* **3:** 1901-1909.

Franovic A, Gunaratnam L, Smith K, Robert I, Patten D, and Lee S. 2007. Translational up-regulation of the EGFR by tumor hypoxia provides a nonmutational explanation for its overexpression in human cancer. *PNAS* **104:** 13092-13097.

Frese KK and Tuveson DA. 2007. Maximizing mouse cancer models. *Nat Rev Cancer* **7:** 645-658.

Friedman RC, Farh KK, Burge CB, and Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19:** 92–105.

Frohn A, Eberl HC, Stöhr J, Glasmacher E, Rüdel S, Heissmeyer V et al. 2012. Dicer-dependent and –independent Argonaute2 Protein Interaction Networks in Mammalian Cells. *Mol Cell Proteomics* **11:11:** 1442-1456.

Gagnon KT, Li L, Chu Y, Janowski BA, and Corey DR. 2014. RNAi Factors Are Present and Active in Human Cell Nuclei. *Cell Reports* **6:** 211-221.

Ghildiyal M and Zamore PD. 2009. Small Silencing RNAs: an expanding universe. *Nat Rev Genet* **10:** 94-108.

Gossen M, Freundlieb S, Bender G, Müller G, Hillen W, and Bujard H. 1995. Transcriptional Activation by Tetracyclines in Mammalian Cells. *Science* **268:** 1766-1769.

Griffiths-Jones S, Saini HK, Dongen SV, and Enright AJ. 2008. miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36:** D154–8.

Gu S, Jin L, Zhang F, Huang Y, Grimm D, Rossi JJ et al. 2011. Thermodynamic stability of small hairpin RNAs highly influences the loading process of different mammalian Argonuates. *PNAS* **108:** 9208-9213.

Ha M and Kim NV. 2014. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* **15:** 509-524.

Han J, Lee Y, Yeom K, Nam J, Heo I, Rhee J et al. 2006. Molecular Basis for the Recognition of Primary microRNAs by the DROSHA-DGCR8 Complex. *Cell* **125:** 887-901.

Hancock JF and Parton RG. 2005. Ras plasma membrane signaling platforms. *Biochem J* **389:** 1-11.

Haupt Y, Maya R, Kazaz A, and Oren M. 1997. Mdm2 promotes the rapid degradation of p53. *Nature* **387:** 296-299.

Hauptmann J, Kater L, Löffler P, Merkl R, and Meister G. 2014. Generation of catalytic human Ago4 identifies structural elements important for RNA cleavage. *RNA* **20:** 1532-1538.

Hayes J, Peruzzi PP, and Lawler S. 2014. MicroRNAs in cancer: biomarkers, functions and therapy. *Trends Mol Med* **20:** 460-469.

Höck J and Meister G. 2008. The Argonaute protein family. *Genome Biology* **9:** 210.

Hunter C, Sun H, and Poethig RS. 2003. The *Arabidopsis* Heterochronic Gene ZIPPY Is an *ARGONAUTE* Family Member. *Curr Biol* **13:** 1734-1739.

Hussain AR, Ahmed SO, Ahmed M, Khan OS, AbdulMohsen SA, Platanias LC et al. 2012. Cross-Talk between NFkB and te PI3-Kinase/AKT Pathway Can Be Targeted in Primary Effusion Lymphoma (PEL) Cell Lines for Efficient Apoptosis. *PLoS One* **7:** e39945.

Hutvágner G, McLachlan J, Pasquinelli AE, Bálint E, Tuschl T, Zamore PD. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293:** 834-838.

Hutvágner G and Zamore PD. 2002. A microRNA in a Multiple-Turnover RNAi Enzyme Complex. *Science* **297:** 2056-2060.

Hutvágner G, Simard MJ. 2008. Argonaute proteins: key players in RNA silencing. *Nature Reviews Molecular Cell Biology* **9:** 22-32.

Jean S and Kiger AA. 2014. Classes of phosphoinositide 3-kinases at a glance. *J Cell Sci* **127:** 923-928.

Jinek M and Doudna JA. 2009. A three-dimensional view of the molecular machinery of RNA interference. *Nature* **457:** 405-412.

Jonas S and Izaurralde E. 2015. Towards a molecular understand of microRNA-

mediated gene silencing. *Nat Rev Genet* **16:** 421-433.

Kastenhuber ER and Lowe SW. 2017. Putting p53 in context. *Cell* **170:** 1062-1078.

Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK et al. 2014. Defining functional elements in the human genome. *PNAS* **111:** 6131-6138.

Kim VN, Han J, and Siomi MC. 2009. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10:** 126-139.

Kiriakidou M, Tan GS, Lamprinaki S, Planell-Saguer MD, Nelson PT, and Mourelatos Z. 2007. An mRNA m(7)G cap binding-like motif within human Ago2 represses translation. *Cell* **129:** 1141-1151.

Kirsten WH and Mayer LA. 1967. Morphological responses to a murine erythroblastosis virus. *J Natl Cancer Inst* **39:** 311-35.

Kobayashi T. 2011. Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast. *Cell Mol Life Sci* **68:** 1395-1403.

Kwak PB and Tomari Y. 2012. The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat Struct Mol Biol* **19:** 145-152.

Lander E. 2011. Initial impact of the sequencing of the human genome. *Nature* **470:** 187-197.

Lee RC, Feinbaum RL, and Ambros V. 1993. The C. elegans Heterochronic Gene lin-4 Encodes Small RNAs with Antisense Complementarity to lin-14. *Cell* **75:** 843-854.

Lewandoski M. 2001. Conditional control of gene expression in the mouse. *Nat Rev Genet* **2:** 743-755.

Li L, Yu C, Gao H, and Li Y. 2010. Argonaute proteins: potential biomarkers for human colon cancer. *BMC cancer* **10:** 38.

Liou G, Doppler H, DelGiorno KE, Zhang L, Leitges M, Crawford HC et al. 2016. Mutant Kras-induced mitochondrial oxidative stress in acinar cells upregulates EGFR signaling to drive formation of pancreatic precancerous lesions. *Cell Rep* **14:** 2325-2336.

Liu Y, Ye X, Jiang F, Liang C, Chen D, Peng J et al. 2009. C3PO, an

Endoribonuclease That Promotes RNAi By Facilitating RISC Activation. *Science* **325**: 750-753.

Longenecker G and Kulkarni AB. 2009. Generation of Gene Knockout Mice by ES Cell Microinjection. *Curr Protoc Cell Biol* **19**: 1436.

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D et al. 2005. MicroRNA expression profiles classify human cancers. *Nature* **435**: 834-838.

Margolis B and Skolnick EY. 1994. Activation of Ras by Receptor Tyrosine Kinases. *J Am Soc Nephrol* **5**: 1288-1299.

Martinez J and Busslinger M. 2007. Life beyond cleavage: the case of Ago2 and hematopoiesis. *Genes & Dev* **21**: 1983-1988.

Martins CP, Brown-Swigart L, and Evans GI. 2006. Modelling the Therapeutic Efficacy of p53 Restoration in Tumors. *Cell* **127**: 1323-1334.

Mayer CT, Gazumayan A, Kara EE, Gitlin AD, Golijanin J, Viant C et al. 2017. The microanatomic segregation of selection by apoptosis in the germinal center. *Science* **358**: eaao2602.

McClatchey AI. 1999. Modeling metastasis in the mouse. *Oncogene* **18**: 5334-5339.

Mebius RE and Kraal G. 2005. Structure and function of the spleen. *Nat Rev Immunol* 5: 606-616.

Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, and Tuschl T. 2004. Human Argonaute2 Mediates RNA Cleavage Targeted by miRNAs and siRNAs. *Mol Cell* **15**: 185-197.

Meister G. 2013. Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet* **14**: 447-459.

Melchers F. 2015. Checkpoints that control B cell development. *J Clin Invest* **125**: 2203-2210.

Melo SA, Moutinho C, Ropero S, Calin GA, Rossi S, Spizzo R et al. 2010. A Genetic Defect in Exportin-5 Traps Precursor MicroRNAs in the Nucleus of Cancer Cells. *Cancer Cell* **18**: 303-315.

Mellor P, Furber LA, Nyarko JNK, and Anderson DH. 2011. Multiple roles for the p85α isoform in the regulation and function of PI3K signalling and receptor trafficking. *Biochem J* **441**: 23-37.

Merentie M, Lottonen-Raikasehto L, Parviainen V, Huusko J, Pikkarainen S, Mendel M et al. 2016. Efficacy and safety of myocardial gene transfer of adenovirus, adeno-associated virus and lentivirus vectors in the mouse heart. *Gene Ther* **23:** 296-305.

Milburn MV, Tong L, deVos AM, Brünger A, Yamaizumi Z, Nishimura S et al. 1990. Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science* **247:** 939-945.

Miyoshi T, Ito K, Murakami R, and Uchiumi T. 2016. Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat Commun* **7:** 1-12.

Modzelewski AJ, Holmes RJ, Hilz S, Grimson A, and Cohen PE. 2012. AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline. *Dev Cell* **23:** 251-264.

Monje P, Hernández-Losa J, Lyons RJ, Castellone MD, and Gutkind JS. 2005. Regulation of the Transcriptional Activity of c-Fos by ERK: A Novel Role For The Prolyl Isomerase PIN1. *J Biol Chem* **280:** 35081-35084.

Möröy T and Heyd F. 2007. The impact of alternative splicing in vivo: Mouse models show the way. *RNA* **13:** 1155-1171.

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, and Honjo T. 2000. Class Switch Recombination and Hypermutation Required Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. *Cell* **102:** 553-563.

Murphy KL and Rosen JM. 2000. Mutant p53 and genomic instability in a transgenic mouse model of breast cancer. *Oncogene* **19:** 1045-1051.

Nagasawa T. 2006. Microenvironmental niches in the bone marrow required for B cell development. *Nat Rev Immunol* **6:** 107-116.

Nishi K, Nishi A, Nagasawa T, and Ul-Tei K. 2013. Human TNRC6A is an Argonaute-navigator protein for microRNA-mediated gene silencing in the nucleus. *RNA* **19:** 17-35.

O'Carroll D, Mecklenbrauker I, Das PP, Santana A, Koenig U, Enright AJ et al. 2007. A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. *Genes & Dev* **21:** 1999-2004.

Orban PC, Chui D, and Marth JC. 1992. Tissue- and site-specific DNA

recombination in transgenic mice. *Proc Natl Acad Sci USA* **89:** 6861-6965.

Parker R and Sheth U. 2007. P Bodies and the Control of mRNA Translation and Degradation. *Mol Cell* **25:** 635-646.

Pasquinelli AE. 2013. MicroRNAs and their targets: recognition, regulation, and an emerging reciprocal relationship. *Nature Reviews Genetics* **13:** 271-282.

Peters L and Meister G. 2007. Argonaute Proteins: Mediators of RNA Silencing. *Mol Cell* **26:** 603-609.

Pierce KL, Luttrell LM, and Lefkowitz RJ. 2001. New mechanisms in heptahelical receptor signaling to mitogen activated protein kinase cascades. *Oncogene* **20:** 1532-1539.

Pylayeva-Gupta Y, Grabocka E, and Bar-Sagi D. 2011. RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer* **11:** 761-774.

Qi HH, Ongusaha PP, Myllyharju J, Cheng D, Pakkanen O, Shi Y et al. 2008. Prolyl 4-hydroxylation regulates Argonaute 2 stability. *Nature* **455:** 421-424.

Reddy P, Negrin R, and Hill GR. 2008. Mouse Models of Bone Marrow Transplantation. *Biol Blood Marrow Transplant* **14:** 129-135.

Rehwinkel J, Behm-Ansmant I, Gatfield D, and Izaurralde E. 2005. A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA* **11:** 1640-1647.

Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE et al. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403:** 901-906.

Rodriguez A, Vigorito E, Clare S, Warren MV, Couttet P, Soond DR et al. 2007. Requirement of bic/microRNA-155 for normal immune function. *Science* **316:** 608-611.

Rüdel S, Wang Y, Lenobel R, Körner R, Hsiao H, Urlaub H et al. 2010. Phosphorylation of human Argonaute proteins affects small RNA binding. *Nucleic Acids Res* **39:** 2330-2343.

Rybak A, Fuchs H, Hadian K, Smirnova L, Wulczyn EA, Michel G et al. 2009. The let-7 target gene mouse *lin*-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. *Nat Cell Biol* **11:** 1411-1422.

Saini HK, Griffiths-Jones S, and Enright AJ. 2007. Genomic analysis of human microRNA transcripts. *PNAS* **104:** 17719-17724.

Salomon DS, Brandt R, Ciardiello F, and Normanno N. 1995. Epidermal growth factor-related peptides and their receptors in human malignancies. *Crit Rev Oncol Hematol* **19:** 183-282.

Sánchez-Rivera FJ, Papagiannakopoulos T, Romero R, Tammela T, Bauer MR, Bhutkar A et al. 2014. Rapid modeling of cooperating genetic events in cancer through somatic genome editing. *Nature* **516:** 428-431.

Sasaki T, Shiohama A, Minoshima S, and Shimizu N. 2003. Identification of eight members of the Argonaute family in the human genome. *Genomics* **82:** 323-330.

Schmick M, Vartak N, Papke B, Kovacevic M, Truxius DC, Rossmanneck L et al. 2014. KRas Localizes to the Plasma Membrane by Spatial Cycles of Solubilization, Trapping and Vesicular Transport. *Cell* **157:** 459-471.

Schraivogel D and Meister G. 2014. Import routes and nuclear functions of Argonaute and other small RNA-silencing proteins. *Trends Biochem Sci* **39:** 420-431.

Scheffzek K, Ahmadian MR, Kabsch W, Weismuller L, Lautwein A, Schmitz F et al. 2007. The Ras-RasGAP Complex: Structural Basis for GTPase Activation and Its Loss in Oncogenic Ras Mutants. *Science* **277:** 333-339.

Schmitz R, Ceribelli M, Pittaluga S, Wright G, Staudt LM. 2014. Oncogenic Mechanisms in Burkitt Lymphoma. *Cold Spring Harb Perspect Med* **4:** a014282.

Schwarz DS, Tomari Y, and Zamore PD. 2004. The RNA-Induced Silencing Complex Is a Mg^{2+}-Dependent Endonuclease. *Curr Biol* **14:** 787-791.

Shaffer AL, Lin K, Kuo TC, Yu X, Hurt EM, Rosenwald A et al. 2002. Blimp-1 Orchestrates Plasma Cell Differentiation by Extinguishing the Mature B Cell Gene Expression Program. *Immunity* **17:** 51-62.

Shankar S, Pitchiaya S, Malik R, Kothari V, Hosono Y, Yocum AK et al. 2016. KRAS Engages AGO2 to Enhance Cellular Transformation. *Cell Reports* **14:** 1-14.

Shapiro-Shelef M and Calame K. 2005. Regulation of Plasma Cell Development. *Nat Rev Immunol* **5:** 230-242.

Sharma SV, Bell DW, Settleman J, Haber DA. 2007. Epidermal growth factor

receptors mutations in lunch cancer. *Nat Rev Cancer* **7:** 169-181.

Shen J, Xia W, Khotskaya YB, Huo L, Nakanishi K, Lim S et al. 2013. EGFR modulates microRNA maturation in response to hypoxia through phosphorylation of AGO2. *Nature* **497:** 383-387.

Shieh A, Ward Af, Donlan KL, Hardin-Theobald ER, Xu J, Mullighan CG et al. 2013. Defective K-Ras oncoproteins overcome impaired effector activation to initiate leukemia in vivo. *Blood* **121:** 4884-4893.

Shih TY, Weeks MO, Young HA, and Scholnick EM. 1979. Identification of a sarcoma virus-coded phosphoprotein in nonproducer cells by Kirsten or Harvey murine sarcoma virus. *Virology* **96:** 64-79.

Siegel RL, Miller KD, and Jemal A. 2017. Cancer Statistics, 2017. *CA Cancer J Clin* **67:** 7-30.

Siggers T, Chang AB, Teixeira A, Wong D, Williams KJ, Ahmed B et al. 2011. Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-κB family DNA binding. *Nat Immunol* **13:** 95-103.

Simanshi DK, Nissley DV, and McCormick F. 2017. RAS Proteins and Their Regulators in Human Disease. *Cell* **170:** 17-33.

Skoulidis F, Byers LA, Diao L, Papadimitrakopoulou VA, Tong P, Izzo J et al. 2015. Co-occuring genomic alterations define major subsets of *KRAS*-mutant lung adenocarcinoma with distinct biology, immune profiles, and therapeutic vulnerabilities. *Cancer Discov* **5:** 860-877.

Smibert P, Yang J, Azzam G, Liu J, and Lai EC. 2013. Homeostatic control of Argonaute stability by microRNA availability. *Nat Struct Mol Biol* **20:** 789-797.

Song J, Smith SK, Hannon GJ, and Joshua-Tor L. 2004. Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science* **305:** 1434-1437.

Stambolic V, Suzuki A, Pompa JL, Brothers GM, Mirtsos C, Sasaki T et al. 1998. Negative Regulation of PKB/Akt-dependent Cell Survival by the Tumor Suppressor PTEN. *Cell* **95:** 29-39.

Sulciner DJ, Irani K, Yu Z, Ferrans VJ, Goldschmidt-Clermont P, and Finkel T. 1996. rac1 Regulates a Cytokine-Stimulated, Redox-Dependent Pathway Necessary for NF-κB Activation. *Mol Cell Biol* **16:** 7115-7121.

Suwaki N, Vanhecke E, Atkin KM, Graf M, Swabey K, Huang P et al. 2011. A

HIF-Regulated VHL-PTP1B-Src Signaling Axis Identifies a Therapeutic Target in Renal Cell Carcinoma. *Sci Transl Med* **3:** 85ra47.

Swarts DC, Makarova K, Wang Y, Nakanishi K, Ketting RF, Koonin EV et al. 2014. The evolutionary journey of Argonaute proteins. *Nat Struct Mol Biol* **21:** 743-753.

Talmadge JE, Singh RK, Fidler IJ, Raz Avraham. 2006. Murine Models to Evaluate Novel and Conventional Therapeutic Strategies for Cancer. *AJP* **170:** 793-804.

Thorpe LM, Yuzugullu H, and Zhao JJ. 2015. PI3K in cancer: divergent roles of isoforms, modes of activation and therapeutic targeting. *Nat Rev Cancer* **15:** 7-24.

Valdmanis PN, Gu S, Schüermann N, Sethupathy P, Grimm D, and Kay MA. 2012. Expression determinants of mammalian argonaute proteins in mediating gene silencing. *Nucleic Acids Res* **40:** 3704-3713.

Vasudevan KM, Gurumurthy S, and Rangnekar VM. 2004. Supression of PTEN Expression by NF-κB Prevents Apoptosis. *Mol Cell Biol* **24:** 1007-1021.

Wang B, Li S, Qi HH, Chowdhury D, Shi Y, and Novina CD. 2009. Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins. *Nat Struct Mol Biol* **16:** 1259-1267.

Weber JD, Taylor LJ, Roussel MF, Sherr CJ, and Bar-Saji D. 1999. Nucleolar Arf sequesters Mdm2 and activates p53. *Nat Cell Biol* **1:** 20-26.

Winter J and Diederichs S. 2011. Argonaute proteins regulate microRNA stability: Increased microRNA abundance by Argonaute proteins is due to microRNA stabilization. *RNA Biol* **8:6:** 1149-1157.

Worringer KA, Rand TA, Hayashi Y, Sami S, Takahashi K, Tanabe K et al. 2014. The let-7/LIN-41 Pathway Regulates Reprogramming to Human Induced Pluripotent Stem Cells by Controlling Expression of Prodifferentiation Genes. *Cell Stem Cell* **14:** 40-52.

Wu S, Yu W, Qu X, Wang R, Xu J, Zhang Q et al. 2014. Argonaute 2 promotes myeloma angiogenesis via microRNA dysregulation. *J Hemat Oncol* **7:** 40.

Xue W, Chen S, Yin H, Tammela T, Papagiannakopoulos T, Joshi NS et al. 2014. CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature* **514:** 380-384.

Yang M, Haase AD, Huang F, Coulis G, Rivera KD, Dickinson BC et al. 2014.

Dephosphorylation of Tyrosine 393 in Argonaute 2 by Protein Tyrosine Phosphatase 1B Regulates Gene Silencing in Oncogenic RAS-Induced Senescence. *Mol Cell* **55:** 782-790.

Ye R, Wang W, Iki T, Liu C, Wu Y, Ishikawa M et al. 2012. Cytoplasmic Assembly and Selective Nuclear Import of *Arabidopsis* ARGONAUTE4/siRNA Complexes. *Mol Cell* **46:** 859-870.

Yigit E, Batista PJ, Bei Y, Pang KM, Chen CG, Tolia NH et al. 2006. Analysis of the *C. elegans* Argonaute Family Reveals that Distinct Argonautes Act Sequentially during RNAi. *Cell* **127:** 747-757.

Zamudio JR, Kelly TJ, and Sharp PA. 2014. Agonaute-bound small RNAs from promoter-proximal RNA polymerase II. *Cell* **156,** 920-934.

Zeng Y, Sankala H, Zhang X, and Graves PR. 2008. Phosphorylation of Argonaute 2 at serine-387 facilitates its localization to processing bodies. *Biochem J* **413:** 429-436.

Zhang W and Liu HT. 2002. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Res* **12:** 9-18.

Zhang X, Zuo X, Yang B, Li Z, Xue Y, Zhou Y et al. 2014. MicroRNA Directly Enhances Mitochondrial Translation during Muscle Differentiation. *Cell* **158:** 607-619.

Zhang Q, Lenardo MJ, and Baltimore D. 2017. 30 Years of NF-κB: A Blossoming of Relevance to Human Pathobiology. *Cell* **168:** 37-57.

Zhou Y, Chen L, Barlogie B, Stephens O, Wu X, Williams DR et al. 2010. High-risk myeloma is associated with global elevation of miRNAs and overexpression of *EIF2C2/Ago2*. *PNAS* **107:** 7904-7909.

Chapter 2

AGO2 overexpression in KP lung adenocarcinomas promotes metastasis

Abstract

AGO2 is an essential component of the RISC complex that facilitates microRNA-mediated silencing of target transcripts. While misregulation of microRNAs is a common feature across the vast majority of cancer types, recent cancer genome studies suggest uncharacterized AGO2 functions in the progression of select human cancers, such as NSCLC. High throughput genomic studies reveal frequent alterations in *AGO2* copy number from NSCLC, which is associated with AGO2 upregulation. Here, we use a previously established KP mouse model to initiate NSCLC to study the effects of AGO2 overexpression in the context of KRAS activation and p53 deletion. In the absence of exogenous AGO2, high-grade tumors express ~25% higher levels of AGO2 protein compared to adjacent normal lung tissue and benign tumors. Furthermore, our results indicate that inducing AGO2 overexpression promotes the development of advanced stage disease by accelerating tumor progression and increasing cell motility and invasiveness, thereby providing insight into uncharacterized oncogenic functions of AGO2.

Introduction

The family of Argonaute proteins are evolutionarily well-conserved between species (Drinnenberg et al., 2009). These proteins are integral components of the cytosolic RNA-induced silencing complex (RISC) where they promote microRNA (miRNA) processing and activity. Of the four mammalian Argonaute proteins, Argonaute-2 (AGO2) is unique due to its ubiquitous

71

expression and catalytic RNase H domain (Song JJ, 2004; Jinek and Doudna, 2009). Consequently, AGO2-bound miRNAs repress their targets through mRNA destabilization and/or translation repression (Huntzinger and Izaurralde, 2011).

Well over 2,000 miRNAs have been identified in the human genome that may contribute to coordinating gene expression programs to regulate many biological processes, such as cell proliferation, differentiation, and apoptosis (Griffiths-Jones S, 2008). Over 60% of the protein-coding genes of the human genome contain at least one predicted 3' UTR miRNA binding site (Friedman RC, 2009). Thus abnormal perturbations in miRNA activity can have profound deleterious effects on normal cellular functions, some of which contribute to tumorigenesis (Calin et al., 2002; Eis et al., 2005).

Misregulation of miRNAs is a common feature across the vast majority of cancer types. In the context of cancer, miRNAs that are aberrantly downregulated typically target genes that are associated with cell proliferation and growth while those that are upregulated are associated with genes that have tumor suppressive properties. The family of let-7 miRNAs is an example of tumor suppressive miRNAs that are often downregulated in a variety of cancers, including non-small cell lung cancer (NSCLC). They are known to target many oncogenes, including *HMGA2* and *KRAS*, contributing to the stem-like properties of tumor cells (Lu et al., 2005; Kumar MS, 2007; Büssing I, 2008).

Alterations in miRNA activity can be attributed to numerous factors, including dysregulation of transcription or mutations in components of the RNA interference (RNAi) pathway that affect miRNA processing (O'Donnell KA, 2005;

72

Lin and Gregory, 2015). For example, previous work from mouse model systems has demonstrated that loss of a single copy of *DICER1*, an RNase III enzyme involved in processing miRNA hairpin structures, in transformed cells provide a selective advantage for growth of lung adenocarcinomas (LUAD) but complete loss of DICER1 was not observed (Kumar et al., 2009; Ravi et al., 2012). This suggests that a select population of microRNAs is crucial for general cellular functions and that DICER1 has tumor suppressor properties, which favors the haploinsufficient state in tumor cells.

Though misregulation of many components of the RNAi pathway have been well studied in various cancers, only recent studies have emerged implicating an oncogenic role for AGO2. Studies of breast cancer models have identified MAPK signaling as an important regulator of increased AGO2 expression (Adams BD, 2009). Interestingly, AGO2 has been recently associated with advanced stage breast cancer. More specifically, the poorly vascularized region of tumors was shown to modulate miRNA processing via increased association between EGFR and AGO2. This results in the EGFR-mediated phosphorylation of Y393-AGO2, impairing the interaction between DICER1 and AGO2. Precursor miRNAs with short stem-loop structures were preferentially selected for in this context, many of which targeted tumor suppressor genes (Shen et al., 2013).

In normal physiological conditions, the stability of AGO2 is highly dependent on the expression of miRNAs (Martinez and Gregory, 2013). Interestingly, miRNA profiling in a diverse array of human cancers has suggested

a general downregulation of miRNAs (Lu et al., 2005). This is perplexing as AGO2 expression is generally upregulated in many cancers and has been implicated in tumor progression (Li L, 2010; Cheng N, 2012; Vaksman O, 2012; Zhang et al., 2013). Furthermore, datasets from The Cancer Genome Atlas (TCGA) show that *AGO2* gene amplification events are common in NSCLC patients, which is also associated with elevated AGO2 expression (Figure 1A) (Cerami et al., 2012). This dichotomy between microRNA and AGO2 expression suggests that there may be microRNA-independent AGO2 functions in transformed cells.

These discoveries demonstrate that AGO2 is a dynamic protein and further characterization is required to determine how misregulation of AGO2 in transformed cells affects its function and localization. To this end, we sought to determine how AGO2 overexpression in NSCLC impacts tumor progression using a genetically engineered mouse model of human lung adenocarcinomas (LUAD). In this model, lung tumors are initiated by inducing KRAS activation and *Tp53* deletion in lung epithelial cells through lentiviral infection administered intratracheally (DuPage M, 2009). To characterize the biological effects of AGO2 overexpression, lentiviral constructs were modified to introduce a doxycycline-inducible *Ago2* cDNA.

Results

AGO2 is upregulated in malignant lung tumors

We initially sought to profile the protein expression pattern of AGO2 in various grades of lung adenocarcinomas derived from mice. Lung tumors were initiated in $Kras^{LSL-G12D/+}$; $Tp53^{fl/fl}$ (KP) mice by infecting them with lentivirus to constitutively express Cre recombinase via intratracheal inhalation. This resulted in the activation of oncogenic *KRAS* and homozygous deletion of tumor suppressor *Tp53* primarily in lung epithelial cells. Mice were sacrificed 24 wk post-lentiviral administration.

Hematoxylin and eosin (H&E) staining of cross sections from lung tissue revealed various grades of lung tumors developed, ranging from grade 1 (benign) to grade 4 (malignant). Similar to human lung adenocarcinomas, AGO2 protein expression appeared to be upregulated in high-grade tumors compared to both adjacent normal tissue and low-grade lung tumors (Figures 1B). Western blot analysis confirmed that there was approximately 25% increase in AGO2 protein in grade 3 lung tumors compared to adjacent normal tissue (Figure 1C).

To determine whether AGO2 expression fluctuates between metastases and their primary tumors, we utilized previously generated LUAD cell lines derived from metastases (Met) and their primary tumor of origin (Tmet) (Winslow et al., 2011). Metastases were traced back to their tumor of origin based off of lentiviral integration sites. *Ago2* expression was evaluated from two independent pair of tumors via qPCR. Ago2 levels were consistently upregulated in Met cell lines compared to Tmets (Figure 1D). Together, these results suggest that *AGO2*
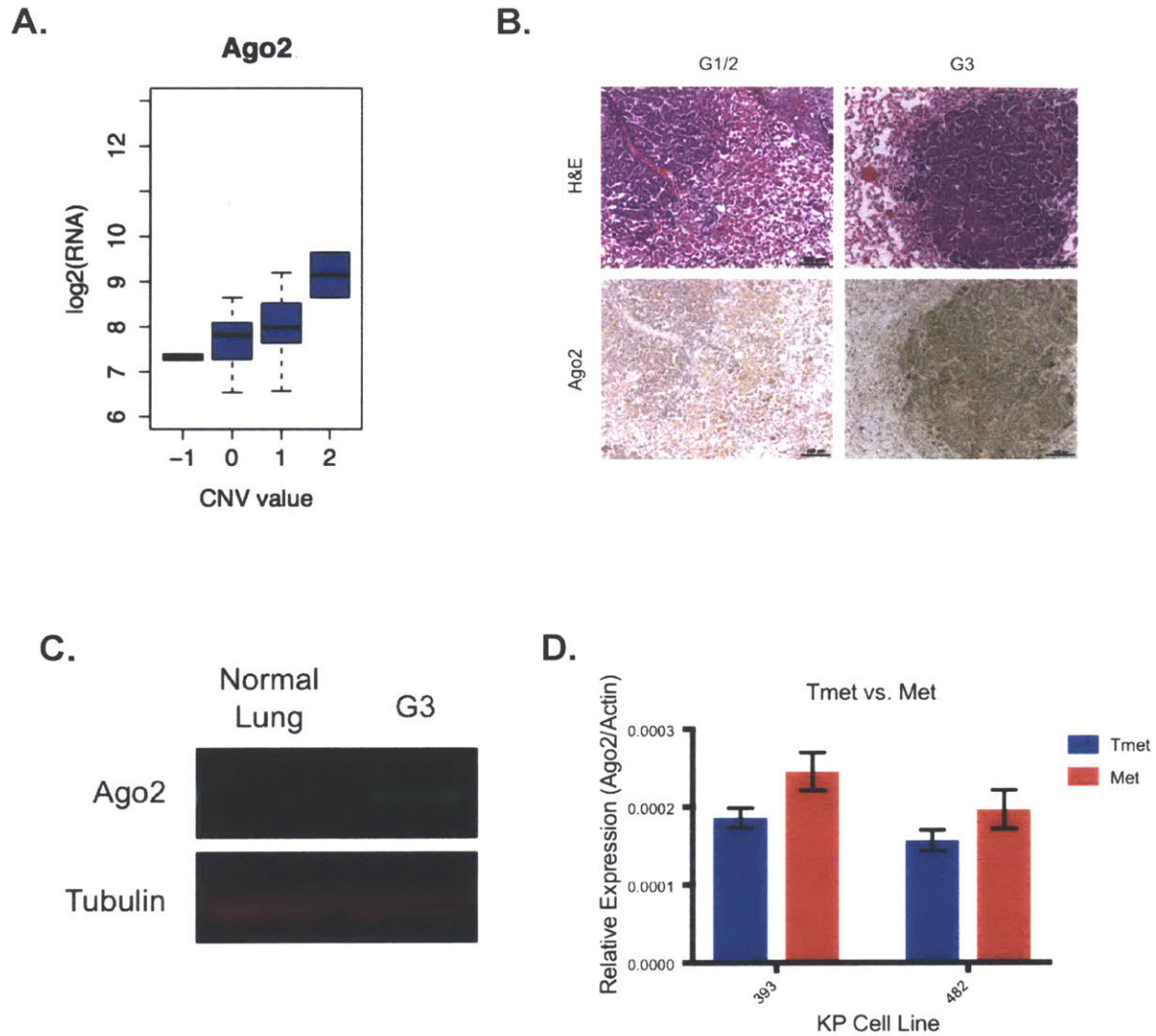
Figure 1: AGO2 expression is elevated in high-grade lung tumors. (A) *AGO2* copy number variations (CNV) was plotted against *AGO2* mRNA expression for tumors from lung adenocarcinoma patients (Cerami et al., 2013). (B) Hematoxylin & eosin (H&E) staining was used to confirm lung tumor grades (top panel) and AGO2 protein expression was evaluated using immunohistochemistry (IHC) (bottom panel). (C) AGO2 protein expression from normal lung tissue was compared to a grade 3 lung tumor with western blot analysis. Tubulin was used as a loading control. (D) qPCR was used to analyze *Ago2* expression from two pairs of KP cell lines, comparing cell lines generated from a metastatic tumor (Met) to its primary lung tumor (Tmet).

overexpression facilitates the transition between benign to malignant states, contributing to tumor dedifferentiation and metastasis.
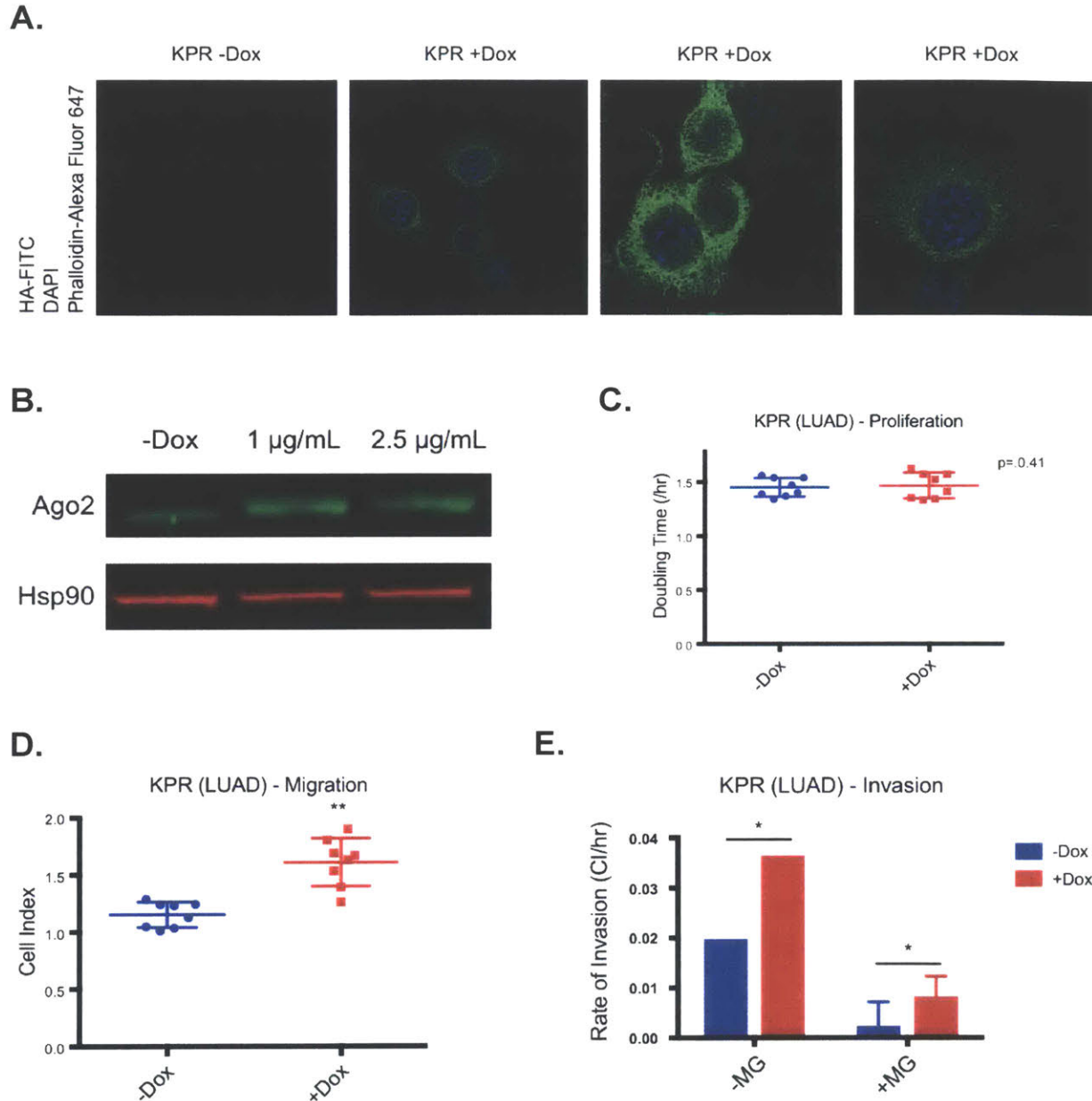
Overexpressing AGO2 increases cell motility in KP lung tumor cells

Given that AGO2 is upregulated in advanced tumors, we sought to explore the functional consequences of AGO2 overexpression in KP lung adenocarcinoma tumor cells. These KP cell lines were derived from primary lung tumors and transduced *in vitro*, where lentiviral integration of a Tet-inducible FLAG-HA-tagged *Ago2* construct allowed for AGO2 overexpression in the presence of doxycycline, now referred to as KPR cells.

Transduced KP cells were single-cell sorted in hygromycin-containing media to select for KPR tumor cells with successful lentiviral integration. Immunofluorescence of HA-tagged AGO2 confirmed doxycycline-inducible expression of exogenous AGO2 in multiple clones (Figure 2A). To quantify the level of AGO2 overexpression in the presence of doxycycline, we used a fluorescence-based western blot approach. This demonstrated a two-fold increase in total AGO2 expression in the presence of 1µg/mL doxycycline (Figure 2B).

To explore the effects of AGO2 overexpression on cell proliferation rates, we quantified the doubling time of KPR cells in the presence and absence of doxycycline. Surprisingly, there was no difference detected in the doubling time of AGO2 overexpressed KPR cells compared to controls (Figure 2C).

Next, we investigated whether AGO2 overexpression affects cell invasion or migration by using a transwell-based system in the presence and absence of Matrigel, respectively. In the absence of Matrigel, there was a significant increase in cell migration of AGO2 overexpressed KPR cells compared to cells cultured in the absence of doxycycline (Figure 2D). Furthermore, the ability of AGO2 overexpressed KPR cells to migrate through the reconstituted basement membrane was significantly enhanced compared to their controls, showing increased invasiveness (Figure 2E). These results suggest that upregulation of AGO2 promote metastasis in the KP LUAD mouse model.

A.

| KPR -Dox | KPR +Dox | KPR +Dox | KPR +Dox |

HA-FITC
DAPI
Phalloidin-Alexa Fluor 647

B.

-Dox 1 µg/mL 2.5 µg/mL

Ago2

Hsp90

C.

KPR (LUAD) - Proliferation

p=.0.41

Doubling Time (/hr)

-Dox   +Dox

D.

KPR (LUAD) - Migration

**

Cell Index

-Dox   +Dox

E.

KPR (LUAD) - Invasion

Rate of Invasion (CI/hr)

*

*

-MG   +MG

■ -Dox
■ +Dox

Figure 2: Overexpressing Ago2 *in vitro* promotes cell migration and invasion. (A) Immunofluorescence was used to analyze Ago2 localization and detect HA-tagged Ago2 from doxycycline-treated KPR cells. (B) Western blot was used to confirm Ago2 overexpression in KPR cells. Hsp90 was used as a loading control. (C) The proliferation rate was measured using Xcelligence on KPR cells in the presence or absence of doxycycline. (D) Cell migration was measured using an Xcelligence transwell assay, using KPR cells in the presence and absence of doxycycline. (E) The rate of invasion was measured using an Xcelligence transwell assay, using KPR cells in the presence or absence of doxycycline. Wells coated with matrigel (+MG) was used to assess cell invasion.

KPR lung tumors are significantly more advanced than KP tumors

Because results from *in vitro* models do not always recapitulate

phenotypes produced from *in vivo* environments, we chose to overexpress AGO2

in the KPR autochthonous mouse model to investigate its effect on tumor

progression and metastasis (Figure 3A). In this model, KPR mice were generated

by crossing KP mice to mice harboring a transgenic CAG-rtTA allele. KP mice

with germline-introduced CAG-rtTA (KPR) allowed for doxycycline-induced

expression of TRE-regulated genes. For this study, KPR and KP mice were

infected with lentivirus, administered through intratracheal inhalation, to co-

express Cre recombinase and doxycycline-inducible AGO2. This promotes *in*

vivo transduction of epithelial cells lining the lungs. Additionally, both cohorts of

mice were fed doxycycline-containing food where only KPR infected cells were

expected to overexpress AGO2.

H&E stains from lung cross-sections collected at 20 wk post-lentiviral

delivery revealed no significant difference in tumor burden between KPR and KP

mice. Furthermore, the number of tumors detected from KPR mice was similar to

those from KP mice, suggesting that AGO2 overexpression neither affects the

initiation of lung tumors nor the proliferation rate of these tumor cells (Figure 3B).

Interestingly, there was a significant shift in the grades of tumors that developed.

Tumors derived from KPR mice appeared more aggressive as determined by

increased dedifferentiated and pleomorphic features compared to KP tumors,

(Figure 3C).

To confirm these findings, KPR cells were transplanted subcutaneously in the flanks of nude mice. These mice were divided into two conditions, where one group was fed doxycycline-containing food while the second group was fed normal chow. The size of the tumors was measured during a 6 wk time-course. Consistent with our previous *in vitro* and *in vivo* findings, we did not detect a proliferative advantage in lung tumor cells overexpressing AGO2 (Figure 3D).
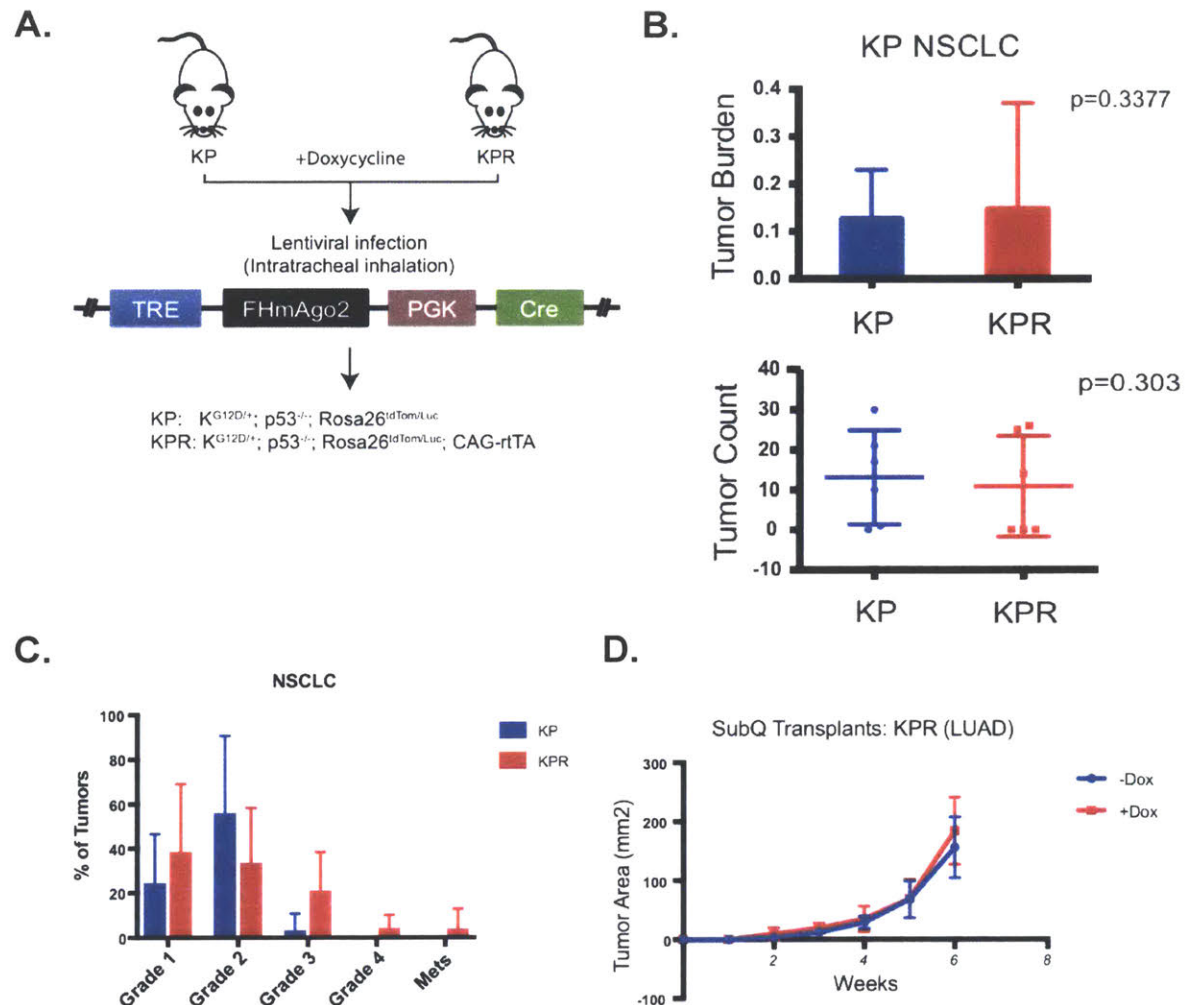
Figure 3: Development of advanced lung tumors is favored with AGO2 overexpression. (A) Schematic of initiating lung tumors from autochthonous KPR mouse model. Lentivirus co-expressing Cre recombinase and doxycycline-inducible AGO2 was administered via intratracheal inhalation. KP mice were used as a control for doxycycline-induced effects (n=6 per group). (B) Tumor burden was measured from H&E stained lung cross-sections collected 20 wk after virus delivery. (C) Lung tumors were graded from a 1 (benign) to 4 (malignant) scale. The number of tumors for each grade was quantified from KPR and KP mice. A single metastasis from kidney tissue was detected from a KPR mouse. (D) KPR cells were transplanted into nude mice via subcutaneous injection. Mice were maintained in the presence and absence of doxycycline-containing food. Tumor size was measured weekly using a caliper.

Tumor cells are more invasive with AGO2 overexpression

Our results thus far suggest that AGO2 overexpression in KP lung tumor cells promote the transition between benign to malignant tumor states. To this end, we sought to determine whether AGO2 overexpression increases the rate of metastasis via intravenous delivery of KPR cells in immunocompromised mice. This mode of cell transplantation serves as a readout for metastasis as it assays the ability of cells to extravasate out of the blood vessels and colonize distant organs. Here, we evaluated the number of lung tumor nodules detected as the lung capillary bed is initially encountered through tail-vein injection. These KPR cells harbored a transgene to express a tdTomato fluorescent reporter, allowing for *ex vivo* quantification of colonized tumor cells.

Fluorescent imaging of the lungs revealed colonization of tdTomato-expressing tumor cells from mice in both the doxycycline-containing food and normal diet conditions (Figure 4A). However, there was a significant difference in the number of surface tumors detected. In the condition where mice were fed doxycycline-containing food, AGO2 overexpression was induced in transplanted KPR cells. Here, there was an increase in the number of tumor nodules detected compared to control (Figure 4B). This result is consistent with our previous findings that show AGO2 overexpressed tumor cells promote cell invasion in the context of KRAS activation and *Tp53* deletion.
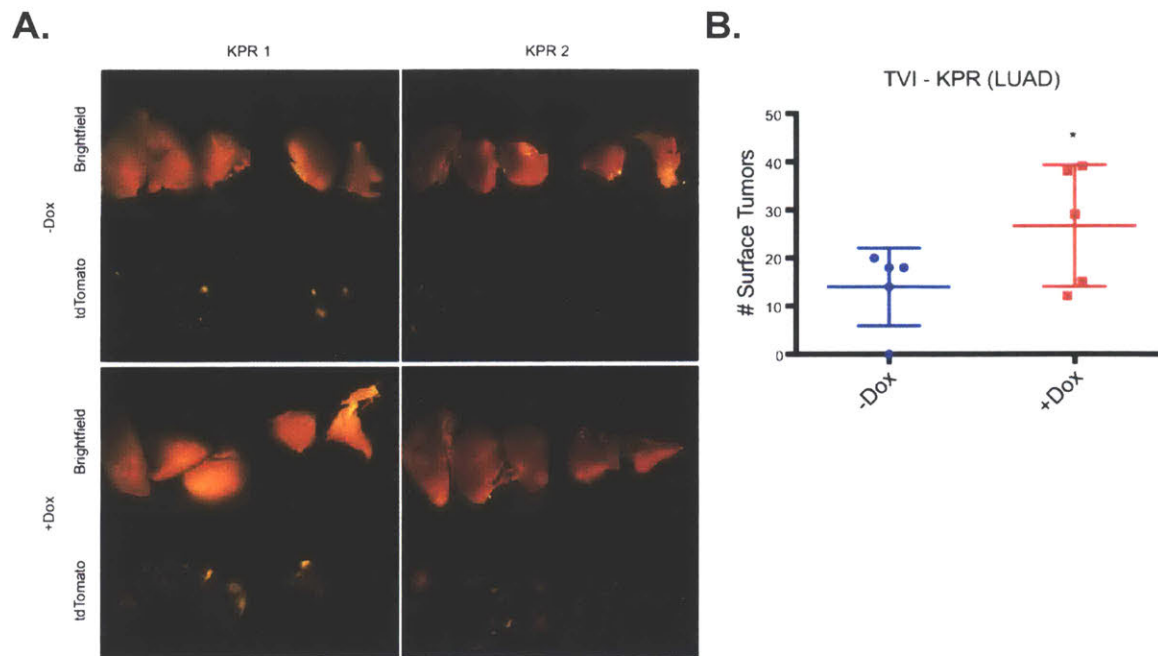
Figure 4: Increased metastases formed with AGO2 overexpression *in vivo*. (A) Fluorescent imaging of tdTomato-expressing KPR tumor cells from lung tissue harvested from nude mice. KPR cells were transplanted via tail-vein injection and mice were either fed normal food (top panels) or doxycycline-containing food (bottom panels). (B) The number of surface tumors detected from the lungs was quantified from *ex vivo* imaging (-Dox: n=5; +Dox: n=5).

MAPK signaling enhanced in AGO2 overexpressed KP tumors

Previous studies have identified the MAPK signaling pathway as a critical component of malignant transformation in the KP NSCLC mouse model (Feldser DM et al., 2010). As our results indicate that AGO2 upregulation promotes the progression toward advanced grade tumors, we sought to profile ERK1/2 (MAPK) activity in different stages of KP lung tumor development using IHC. This revealed a significant increase in the phosphorylation of ERK1/2 during the transition between grade 2 and grade 3 tumors, which represents the malignant transformation event (Figure 5A). This is consistent with the notion of MAPK signaling being an essential component of lung tumor progression in this system and together with our previous results, suggests that constitutive AGO2 overexpression enhances MAPK activity at earlier stages, thereby accelerating tumor progression.

Similarly, we sought to profile the phosphorylation of AKT during tumor progression since PI3K activity is also stimulated through KRAS activation. Intriguingly, the activity of this pathway is active during all stages of tumor development but is dampened during malignant transformation (Figure 5B). Therefore, the anti-correlation between MAPK and PI3K activity during lung tumor progression shows that MAPK signaling is the dominant effector pathway promoting advanced stage tumors while PI3K signaling is more critical for tumor initiation, which has been previously described in the KP model (Gupta et al., 2007).
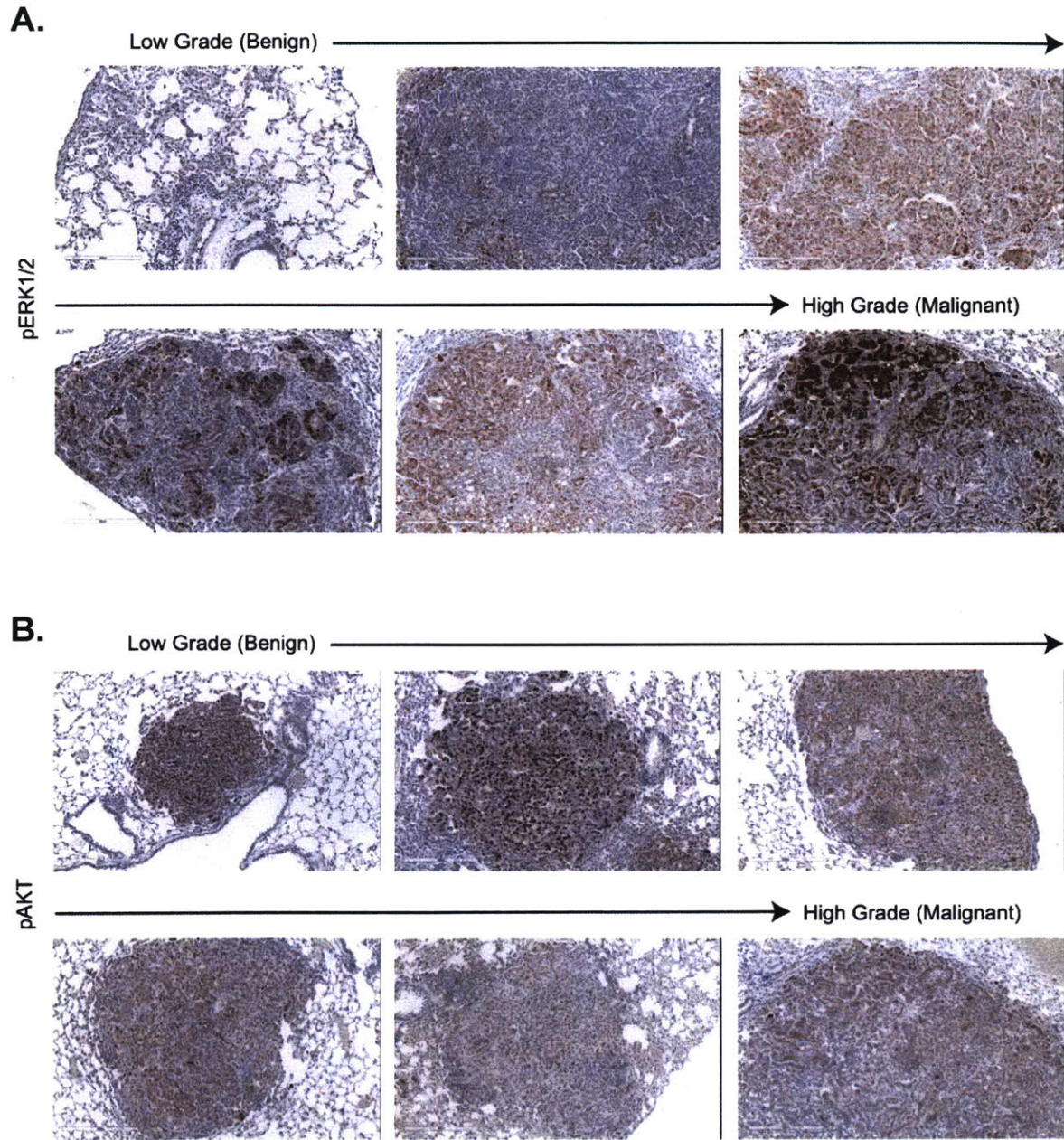
Figure 5: Anti-correlation between MAPK and PI3K signaling during lung tumor progression. (A) pERK1/2 expression was assessed through IHC to measure relative MAPK activity from low grade to high grade lung tumors in Ago2 overexpressed conditions. (B) Relative pAKT expression was assessed through IHC as a readout for relative PI3K activity from lung tumors at different stages during tumor progression. All tissue was harvested from KPR mice infected with lentivirus delivered through intratracheal inhalation.

Discussion

We investigated the biological effects of AGO2 overexpression in lung adenocarcinomas in the context of KRAS activation and *Tp53* deletion. Overexpressing AGO2 in established KP LUAD cell lines resulted in an increase in cell migration and invasion in both *in vitro* and *in vivo* settings using transwell and transplantation assays, respectively. Strikingly, variations in cell doubling rates were not observed with perturbations in AGO2 levels. Hence, AGO2 upregulation in this system predominantly enhances metastatic characteristics of lung tumor cells without noticeable changes in proliferation rates.

Consistent with these findings, ectopic expression of AGO2 in infected cells from the KP autochthonous model accelerated tumor progression with no observable differences in tumor burden. In KPR mice, the development of high-grade lung tumors was enhanced in comparison to tumors initiated from KP control mice. This suggests that elevated AGO2 expression promotes the transition from a benign lesion to a malignant tumor. This was further supported by our observation that increased metastases developed from transplantation of KPR lung tumor cells into nude mice in the presence of doxycycline. Consistent with these findings, AGO2 was previously reported to upregulate focal adhesion kinase (FAK) expression in a model of hepatocellular carcinoma, resulting in enhanced metastasis in this model (Cheng N, 2013). This suggests that AGO2 perturbs the transcriptional program of tumor cells to disrupt cell-cell adhesion.

The canonical functions of AGO2 have been well established in the context of miRNA-mediated gene regulation. It is plausible that AGO2

overexpression in the described KP LUAD model promotes activity of oncogenic

miRNAs to facilitate the progression from benign to aggressive tumor state.

However, as many miRNAs have been found to be generally downregulated in

human cancers, including NSCLC, the notion that there are miRNA-independent

functions of AGO2 is both provocative and plausible (Lu et al., 2005).

It was recently reported that an engagement between KRAS and AGO2 in

a panel of lung and pancreatic cell lines promotes cellular transformation

(Shankar et al., 2016). This interaction was found to upregulate signaling

pathways involved in cell growth, including the PI3K signaling cascade. If this is a

general phenomenon in KRAS-driven cancers, this provides novel insight into the

mechanism of AGO2 overexpression in promoting advanced stage cancers.

In a related study, a recent genome-wide loss-of-function screen using

shRNA libraries had been conducted to identify dependencies for cancer cell

survival. This screen revealed heightened sensitivity to AGO2 knockdown from a

variety of cancer cell lines, specifcally in the context of *KRAS* and *TP53*

mutations. However, sensitivity to AGO2 depletion was not observed in cancer

cell lines with mutually exclusive mutations in either *KRAS* or *TP53* (Tsherniak et

al., 2017). Moreover, AGO2 dependent cancers were generally found to be co-

dependent on TWIST1, which is a transcription factor involved in promoting

metastasis (Figure 6A). This further supports the synergistic effects of activating

Kras mutations, *Tp53* ablation and AGO2 overexpression involved in cellular

transformation in our KP mouse model.
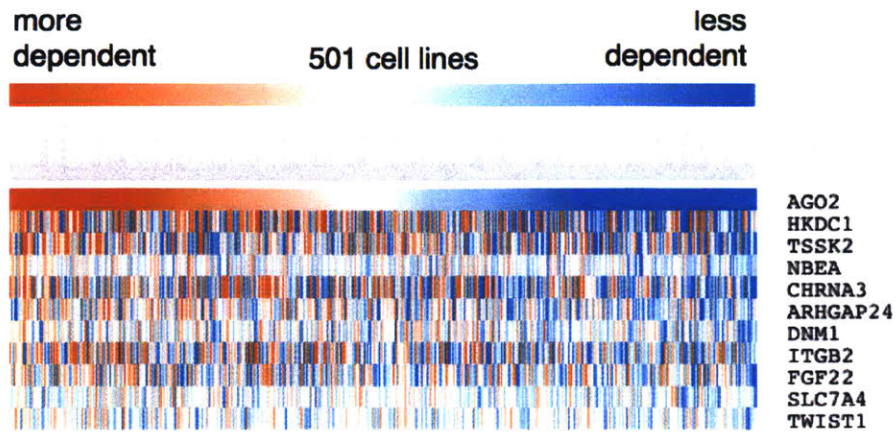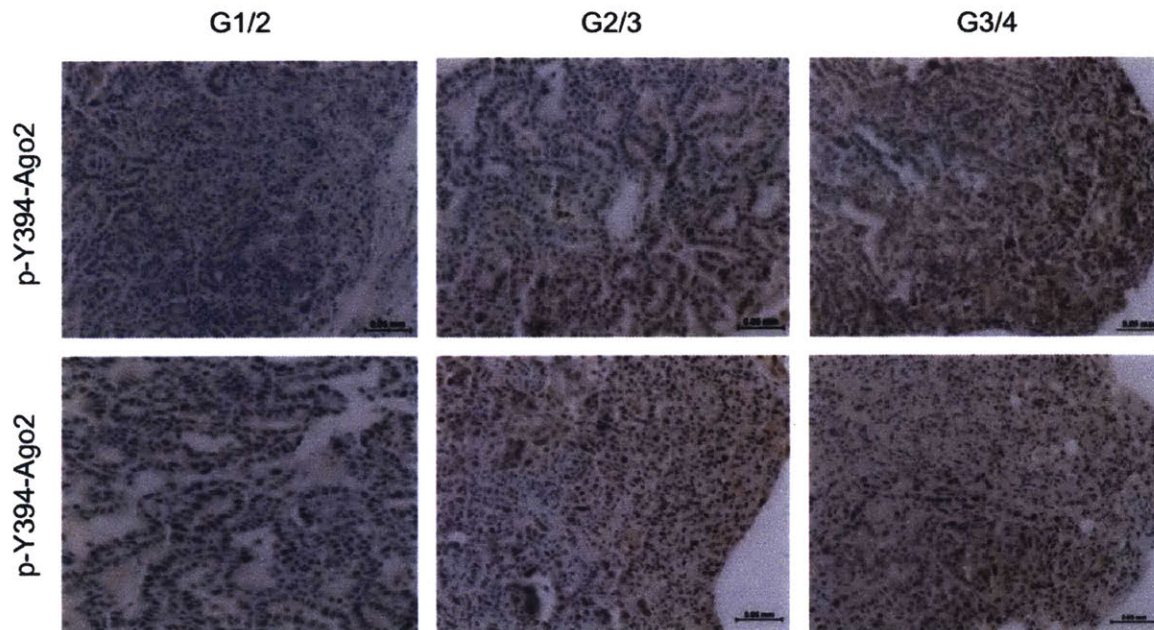
Figure 6: Association between AGO2 dysregulation and malignant transformation. (A) Top 10 AGO2 co-dependent genes were identified from a LOF screen in 501 cancer cell lines using gonome-wide shRNA library (Tsherniak et al., 2017). (B) IHC demonstrates that p-Y394-AGO2 is predominantly expressed in high-grade lung tumors. Cross sections of lung tissue derived from KP mice were pathologically graded and were subsequently stained for p-Y394-AGO2.

Although the association between AGO2 and key features of metastasis is substantial, the mechanism that links AGO2 upregulation to promoting cell migration and invasion is still unclear. We have demonstrated an increase in MAPK signaling associated with malignant transformation in AGO2 overexpressed lung tumors, consistent with previous reports of MAPK dependency for lung tumor progression (Feldser et al., 2010). It is also peculiar that there is a strong correlation between MAPK activity and p-Y394-AGO2 expression during tumor progression (Figure 6B). Accumulation of p-Y394-AGO2 also appeared to be predominantly nuclear. Coincidentally, p38 MAPK activity has been demonstrated to promote phosphorylation of serine 387 of AGO2, resulting in AGO2 localization to processing bodies (p-bodies), which is involved in facilitating miRNA-mediated mRNA destabilization (Zeng Y, 2008). As MAPK signaling is enhanced during tumor progression in our system, it is conceivable that activated MAPK mediates phosphorylation of Y394-AGO2, resulting in its nuclear localization where it promotes a transcriptional response favoring tumor dedifferentiation and/or metastasis. Notably, it is currently unclear whether phosphorylation of this tyrosine residue affects AGO2 binding to KRAS.

The post-translational modification of Y393-AGO2, the human equivalent of Y394-AGO2 in mice, has previously been attributed to the progression of pathological conditions. For example, accumulation of p-Y393-AGO2 has been associated with poor prognosis of EGFR-driven breast cancer patients, where phosphorylation of tyrosine 393 reduces DICER1-AGO2 interactions, thereby negatively affecting miRNA processing (Shen et al., 2013). Here, miRNAs with

long stem-loop hairpin structures were disfavored for AGO2 loading. These specific miRNAs were generally associated with tumor suppressive functions, revealing a novel mechanism of post-transcriptional regulation favored by cancer cells through AGO2 dysregulation. Though we did not assay the functional consequences of AGO2 phosphorylation in KP lung tumor cells, these studies implicate a broad advantage of AGO2 overexpression to promote tumor progression, by promoting enhanced KRAS signaling and/or miRNA dysregulation.

In conclusion, we have identified an oncogenic function for AGO2 in our KP mouse model of human NSCLC, where AGO2 overexpression promotes metastasis. Because targeting oncogenic KRAS has been a therapeutic challenge due to structural limitations, identifying novel KRAS effector molecules to impede oncogenic signaling has been appealing. This study demonstrates a critical role for AGO2 function in a KRAS-driven malignancy, supporting further efforts to characterize the AGO2-KRAS association and identify novel approaches to suppress Kras activation.

Experimental Procedures

Cell Culture

Mouse lung adenocarcinoma cell lines were derived from plucked tumors. These cells were homogenized and then plated on 10cm cell culture-treated plates in DMEM media supplemented with 10% fetal bovine serum (FBS), L-Glutamine and pen/strep. All cells were grown at 37°C with 5% CO_2. All experiments using doxycycline to induce Ago2 overexpression was used at a concentration of 2.5 μg/mL unless noted.

Lentiviral Constructs

GMAP-compatible pLL3 lentiviral backbone was previously engineered to generate viral constructs via Gibson Assembly® with the following construction: backbone-pA-gA-pB-gB. TRE promoter was amplified from a DNA template with "promoter A" (pA) Gibson homology arms (Clontech, #R040A), purified from 1% agarose gels and gel-extracted (Macherey-Nagel, #740609.50). Tandem FLAG and HA tags were engineered onto the N-terminus of mAGO2 via PCR from cDNA templates derived from mouse embryonic stem cells. PCR primers were designed with "gene A" (gA) Gibson homology arms. PCR products were purified from 0.7% agarose gels and gel-extracted. pGK was amplified from DNA templates with "promoter B" (pB) Gibson homology arms, purified from 1% agarose gel and gel-extracted. Cre recombinase was amplified with "gene B" (gB) Gibson homology arms, purified from a 0.8% agarose gel and gel-extracted. Gibson Assembly was performed by first preparing 5X isothermal assembly reaction buffer (0.5M Tris-HCl pH7.5, 50mM $MgCl_2$, 1mM dNTPs, 50mM DTT, 1.5g PEG-8000, 20mg NAD, and bring up to 6mL with nuclease-free water). Isothermal master mix was then prepared by combining 320μL 5X isothermal assembly reaction buffer, 1.2μL T5 exonuclease (NEB®, #M0363L), 20μL Phusion polymerase (NEB®, #M0530L), 160μL Taq ligase (NEB®, #M0208L), and 700μL nuclease-free water. 4μL inserts (5.7 x 10^{-2} pmol each of pA,gA,pB,gB) and 1μL of 50ng digested-backbone (BsrGI/PmeI-digested lentiviral backbone) were added to 15μL isothermal master mix and incubated at 50°C for 1 hour. These reactions were then transformed into competent bacteria (Thermo Scientific™, #C737303), grown at 30°C for 20 hours, and screened with Sanger sequencing. pSLIK-hygro (Addgene, #25737) lentiviral backbone was used to clone in mAgo2, which was PCR amplified with 5' EcoRI and 3' NotI restriction sites. PCR products were purified and extracted with the above protocol. All primers used are listed in primers table in supplemental materials.

Virus Production and Titering

To produce lentiviruses, we passaged and split 293T cells onto 15cm tissue culture-treated plates until they reached 70% confluency. We co-transfected these 293T with 10μg lentiviral constructs, 7.5μg delta8.2 and 2.5μg VSV-G packaging plasmids, and 75μL TransIT-LT1 (Mirus Bio, #MIR2300). Media was changed on the following day with 15mL fresh media (DMEM, 10% FBS, 2mM L-Glutamine). Supernatant was collected 48 and 72 hours post-transfection and

run through 0.45µm syringe filter (Pall Corporation, #4184). Virus was concentrated at 25,000 r.p.m. for 90 minutes at 4°C and resuspended in 200µL Gibco™ Opti-MEM™ (Thermo Scientific™, #31985070). Virus was titered by infecting Green-Go cells with serial dilutions of virus. Green-go cells were generated by transducing 3TZ cells with a bicistronic retrovirus containing an LTR promoter-driven inverted GFP sequence, flanked by two sets of incompatible loxP sites. Viral titers were calculated based off of Cre-mediated GFP-expression.

Immunofluorescence

A sterile coverslip (VWR, #48383-042) was placed into each well of a 12-well plate. 25,000 cells were seeded into each well and allow for cells to adhere overnight at 37°C with 5% CO_2. Media was aspirated and cells washed 2x with PBS. Ice cold methanol:acetone (1:1) was added for 10min at -20°C. Wells were washed with PBS and blocked with PNB Blocking Reagent (PerkinElmer, #FP1020) at room temperature for 1 h. Blocking reagent was removed and fixed cells were incubated with primary antibody (diluted in PNB blocking reagent) overnight at 4°C. Cells were washed 5x with PBS and incubated with secondary antibody at room temperature for 1 h. Wells were washed 5x with PBS and coverslips were mounted onto slides with mounting medium containing DAPI (Vector Laboratories, #H-1200). Slides were stored in the dark at 4°C until ready to image. All antibodies used are listed in antibodies table in supplemental materials.

Proliferation, migration and invasion assays

To assay for cell proliferation rates, KPR cells were diluted in media, either containing or lacking doxycycline, at a concentration of 100 cells/µL. 100µL of media was added to each well of the proliferation plates (ACEA Biosciences Inc, #05469830001) and incubated at room temperature for 30 m to equilibrate. The background reading was measured on the xCELLigence RTCA DP apparatus (ACEA Biosciences Inc, #00380601050). 100µL of diluted cells were added to each well and incubated at room temperature for 30 minutes. Plates were mounted on xCELLigence RTCP DP at 37°C with 5% CO_2 and the proliferation analysis was initiated, where impedance measurements were taken every 15 m for 48 h. Wells plated without cells were used as negative controls. Cell doubling time was calculated.

To assay for cell migration and invasion, KPR cells were diluted in serum-free media (SFM), either containing or lacking doxycycline, at a concentration of 250 cells/µL. Matrigel® (Corning, #354230) was diluted to a concentration of 800 µg/mL and 50µL was added to wells in the upper chamber intended for cell invasion measurements (ACEA Biosciences Inc, #05665817001). 30µL of Matrigel® was removed from each well and the upper chamber was incubated at 37°C for 4 h. 160µL of media was added to each well of the bottom chamber of the transwell plate. SFM was added to wells in the bottom chamber as negative controls for cell migration/invasion. The upper chamber was attached to the bottom chamber. 25µL of SFM was added to each well of the upper chamber and

incubated at 37°C for 1 h to equilibrate plate. The background reading was measured on the xCELLigence RTCA DP (ACEA Biosciences Inc, #00380601050). 100µL of diluted cells were added to each well and incubated at room temperature for 30 minutes. Plates were mounted on xCELLigence RTCP DP at 37°C with 5% CO_2 and migration/invasion analysis was started, where measurements were taken every 15 m for 24 h.

Histology and IHC

Tissues were fixed in zinc formalin (Polysciences, Inc., #21516) overnight at room temperature. They were then transferred to 70% ethanol prior to paraffin embedding. Sections were stained with Hematoxylin and Eosin (H&E) on the Varistain Gemini automated slide stainer (Thermo Shandon). IHC was performed with 5µm-thick sections. Tissues were blocked with Dako Dual Endogenous Enzyme-Blocking Reagent (Agilent Technologies, #S200389-2) and stained using the ImmPRESS™ HRP Polymer Reagents and DAB staining kit (Vector Laboratories, #SK-4100). All antibodies used are listed in antibodies table in supplemental materials.

Western Blotting

Cells were lysed with RIPA supplemented with protease inhibitor cocktail (Roche, #04693159001), Benzonase® nuclease (Sigma, #E8263) and Halt phosphatase inhibitor cocktail (Thermo Scientific™, #78420). Protein lysates were quantified using Pierce™ BCA™ Protein Assay (Thermo Scientific™, #23225). Samples and Kaleidoscope ladder (Bio-Rad, #161-0375) were prepared in NuPAGE® loading buffer (Thermo Scientific™, #NP0007) and reducing agent (Thermo Scientific™, #NP0004) and boiled for 5 minutes at 95°C followed by 2-minute incubation on ice. Western blotting was performed using the Xcell II® Blot Module (Thermo Scientific™). Proteins were run on NuPAGE™ 4-12% Bis-Tris protein gels (Thermo Scientific™, #NP0321BOX) and then transferred to nitrocellulose membrane for LI-COR fluorescent imaging (LI-COR® Biosciences, #926-31092). Membranes were blocked in Odyssey® Blocking Buffer (LI-COR®, #927-40000) incubated with primary antibodies overnight at 4°C with rocking. Membranes were washed 3x with PBST, 5 minutes each, and incubated in secondary antibody for 1 hour at room temperature with rocking. Membranes were washed 4x with PBST and exposed/developed. All antibodies used are listed in antibodies table in supplemental materials.

qRT-PCR

Total RNA from cells were TRIzol-extracted (Thermo Scientific™, #15596026). For mRNA qRT-PCR, total RNA was converted to cDNA using M-MLV (Thermo Scientific™, #28025013) and Oligo d(T) (Gene Link™, #26-4000-05). In each qRT-PCR reaction, we added 10ng cDNA, 1X Taqman probe mix (Thermo Scientific™, #4331182), and 1X TaqMan™ Fast Advanced Master Mix (Thermo Scientific™, #4444556). These samples were amplified and analyzed on the Roche LightCycler® 480 instrument. Gene expression was normalized to β-Actin (Thermo Scientific™, #4453320).

Tumor models

All animal studies described in this study were approved by the MIT institutional Animal Care and Use Committee. All animals were maintained on a mixed C57BL/6J x 129SvJ genetic background. KP mice were previously developed by engineering transcriptional and translational stop elements, flanked by loxP sites, into the first intron of the endogenous K-ras gene. Furthermore, these mice contain an oncogenic mutation in K-ras, where a glycine to aspartic acid mutation in present at codon 12, and conditional "floxed" p53 alleles between exons two through ten to constitutively activate K-ras and mimic loss of p53 function upon Cre-mediate recombination, respectively. These KP mice were then crossed to the Ai9 Cre reporter mice harboring loxP-flanked STOP cassette preventing transcription of a CAG-promoter driven tdTomato fluorophore at the ROSA26 locus. These KP tdTomato reporter mice were then crossed to Cre reporter mice generated from Inder Verma's laboratory harboring loxP-flanked STOP cassette preventing transcription of luciferase at the ROSA26 locus. These KP(tdTom/Luc) reporter mice were then crossed to CAGs-rtTA3 mice generated from Scott Lowe's laboratory to enable a conditional expression system in the presence of doxycycline.

The generated KP reporter mice used for our studies either harbored or lacked the rtTA3 transgene, allowing us to control for doxycycline-specific effects. In the autochthonous setting, we delivered 8.0×10^3 infectious particles of our lentiviral construct via intratracheal inhalation to each mouse that was briefly anesthetized with isoflorane. These mice were placed on doxycycline food to induce Ago2 overexpression in transduced cells harboring rtTA allele.

For subcutaneous injections, KP lung adenocarcinoma cells were diluted using cold PBS to a concentration of 1,000 cells/μL. 100μL of cells were injected into the flanks of nude mice (The Jackson Laboratory, #007850) and monitored for tumor growth.

For intravenous injections, KP lung adenocarcinoma cells were diluted using cold PBS to a concentration of 1,000 cells/μL. 100μL of cells were injected into the tail vein of nude mice and sacrificed 4 wk post-transplantation.

Statistical Analysis

Student's *t*-test was used for all measurements of tumor burden and IHC quantifications. All error bars denote s.e.m.

Supplemental Materials

Antibody	Vendor/Catalog	Assay(s)	Dilution
Ago2	Cell Signaling Technology/2897S	W.B./IHC	1:1000/1:200
HA	Cell Signaling Technology/3724S	IHC	1:500
Hsp90	BD Biosciences/610418	W.B.	1:10000
phospho-Y394-Ago2	Cell Signaling Technology/N.A.	IHC	1µg/mL
phospho-AKT	Cell Signaling Technology/4060S	IHC	1:200
phospho-ERK1/2	Cell Signaling Technology/4370S	IHC	1:400
Tubulin	Abcam/ab6046	W.B.	1:500

Table 1: List of antibodies used.

Primer	Orientation	Sequence
FHmAgo2	F	CGCGAATTCATGTACTCGGGAGC
FHmAgo2	R	GCGGCGGCCGCTCAAGCAAAG
FHmAgo2 gA	F	CTAACTCGAACGCTAGCTGTGCGATCGTTTGCG ACTAGTCCACCATGGACTACAAGGACGACGATG ACAAGTACCCTTATGACGTGCCCGATTACGCTT ACTCGGGAGCCGGCCCCG
FHmAgo2 gA	R	AGGCCTCGGGATTCCTAGGAACAGCGGTTTTCA AGCAAAGTACATGGTGCGCAGTGTGTC
TRE pA	F	GATCAGTGTGAGGGAGTGTAAAGCTGGTTTCGA CGGTtTCGATATGTCGAG
TRE pA	R	AAACGATCGCACAGCTAGCGTTCGAGTTAGTGA CTTACGtATTCTCCAGGC
pGK pB	F	AAACCGCTGTTCCTAGGAATCCCGAGGCCTATT CTACCGGGTAGGGGAGG
pGK pB	R	CCGCTTAAGCTGTAGCGCTAATGTCGGGTCTCC CGGAGATGAGGAAGAGG
Cre gB	F	GACCCGACATTAGCGCTACAGCTTAAGCGGACC ATGCCCAAGAAGAAGAGG
Cre gB	R	AGAGTAATTCAACCCCAAACAACAACGTTTCTAA TCGCCATCTTCCAGCAG

Table 2: List of primers used.

References

Adam BD, Claffey KP, and White BA. 2009. Argonaute-2 Expression is
    Regulated by Epidermal Growth Factor Receptor and Mitogen-Activated
    Protein Kinase Signaling and Correlates with a Transformed Phenotype in
    Breast Cancer Cells. *Endocrinology* **150:** 14-23.

Büssing I, Slack FJ, and Großhans H. 2008. let-7 microRNAs in development,
    stem cells and cancer. *Cell Press* **14:** 400-409.

Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch et al. 2002. Frequent
    deletions and down-regulation of micro-RNA genes *miR15* and *miR16* at
    13q14 in chronic lymphocytic leukemia. *Proc. Natl Acad Sci. USA* **99:**
    15524-15529.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BU et al. 2013. The
    cBio Cancer Genomics Portal: An Open Platform for Exploring
    Multidimentional Cancer Genomics Data. *Cancer Discov* **2:** 401.

Cheng N, Li Y, and Han Z. 2013. Argonaute2 Promotes Metastasis by Way of
    Up-regulating Focal Adhesion Kinase Expression in Hepatocellular
    Carcinoma. *Hepatology* **57:** 1906-1918.

Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR et al.
    2009. RNAi in budding yeast. *Science* **326:** 544-550.

DuPage M, Dooley AL, and Jacks T. 2009. Conditional mouse lung cancer
    models using adenoviral or lentiviral delivery of Cre recombinase. *Nat
    Protoc* **4:** 1063-1072.

Eis PS, Tam W, Sun L, Chadburn A, Li Z, Gomez MF et al. 2005. Accumulation
    of miR-155 and *BIC* RNA in human B cell lymphomas. *Proc. Natl. Acad.
    Sci. USA* **102:** 3627-3632.

Feldser DM, Kostova KK, Winslow MM, Taylor SE, Cashman C, Whittaker CA et
    al. 2010. Stage-specific sensitivity to p53 restoration during lung cancer
    progression. *Nature* **468:** 572-575.

Friedman RC, Farh KK, Burge CB, and Bartel DP. 2008. Most mammalian
    mRNAs are conserved targets of microRNAs. *Genome Res* **19:** 92-
    105.

Griffiths-Jones S, Saini HK, Dongen SV, and Enright AJ. 2008. miRBase: tools
    for microRNA genomics. *Nucleic Acids Res* **36:**154-158.

Gupta S, Ramjaun AR, Haiko P, Wang Y, Warne PH, Nicke B et al. 2007.

Binding of ras to phosphoinositide 3-kinase p110 alpha is required for ras-driven tumorigenesis in mice. *Cell* **129**: 957-968.

Huntzinger E and Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* **12**: 99-110.

Jinek M, Doudna JA. 2009. A three-dimensional view of the molecular machinery of RNA interference. *Nature* **457**: 405-412.

Kumar MS, Lu J, Mercer KL, Golub TR, and Jacks T. 2007. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet* **39**: 673-677.

Kumar MS, Pester RE, Chen CY, Lane K, Chin C, Lu J et al. 2009. *Dicer1* functions as a haploinsufficient tumor suppressor. *Genes & Dev* **23**: 2700-2704.

Li L, Yu C, Gao H, and Li Y. 2010. Argonaute proteins: potential biomarkers for human colon cancer. *BMC cancer* **10**: 38.

Lin S and Gregory RI. 2015. MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer* **15**: 321-333.

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D et al. 2005. MicroRNA expression profiles classify human cancer. *Nature* **435**: 834-838.

Martinez NJ and Gregory RI. 2013. Argonaute2 expression is post-transcriptionally coupled to microRNA abundance. *RNA* **19**: 605-612.

O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, and Mendell JT. 2005. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**: 839-843.

Ravi A, Gurtan AM, Kumar MS, Bhutkar A, Chin C, Lu V et al. 2012. Proliferation and Tumorigenesis of a Murine Sarcoma Cell Line in the Absence of DICER1. *Cancer Cell* **21**: 848-855.

Shen J, Xia W, Khotskaya YB, Huo L, Nakanishi K, Lim S et al. 2013. EGFR modulates microRNA maturation in response to hypoxia through phosphorylation of AGO2. *Nature* **497**: 383 -387.

Song JJ, Smith SK, Hannon GJ, and Joshua-Tor L. 2004. Crystal structure of Argonaute and its implications. *Science* **305**: 1434-1437.

Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kyrukov G, Cowley GS et al.

2017. Defining a Cancer Dependency Map. *Cell* **170:** 564-576.

Vaksman O, Hetland TE, Trope CG, Reich R, and Davidson B. 2012. Argonaute, Dicer and Drosha are up-regulated along tumor progression in serious ovarian carcinoma. *Hum Pathol* **43:** 2062-2069.

Zeng Y, Sankala H, Zhang X, and Graves PR. 2008. Phosphorylation of Argonaute 2 at serine-387 facilitates its localization to processing bodies. *Biochem J* **413:** 429-436.

Zhang J, Fan XS, Wang CX, Liu B, Li Q, and Zhou XJ. 2013. Up-regulation of Ago2 expression in gastric carcinoma. *Med Oncol* **30:** 628.

Chapter 3

A novel mouse model of ABC-like DLBCL initiated by oncogenic KRAS and overexpression of AGO2

This chapter is adapted from the following manuscript in preparation:

Thai KK, Louissaint Jr A, Suzuki H, Bhutkar A, Sharp PA, and Jacks T. A novel mouse model of ABC DLBCL initiated by oncogenic KRAS and overexpression of AGO2. *In preparation*.

The work in this chapter had contributions from Vasilena Gocheva and Yadira Soto-Feliciano, who provided assistance with mouse procedures, Abner Louissaint Jr. and Roderick T. Bronson, who provided pathology consultations on mouse tissue, and Hiroshi Sazuki and Arjun Bhutkar, who provided assistance with computational analysis of sequencing data. KT designed and performed all experiments.

Authors: Kevin K. Thai[1,2], Abner Louissaint Jr[3], Hiroshi Suzuki[1], Arjun Bhutkar[1], Phillip A. Sharp[1,2*] & Tyler Jacks[1,2,4*]

Affiliations:
[1]David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
[3]Massachusetts General Hospital, Department of Pathology, Boston, Massachusetts 02114, USA.
[4]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

Abstract

Activated B cell (ABC) lymphoma is an aggressive form of diffuse large B cell lymphoma (DLBCL). Here, we describe a mouse model resembling ABC DLBCL made by overexpressing AGO2 in the context of oncogenic *Kras* mutation and p53 inactivation. These animals develop B cell lymphomas that display abnormal plasmablastic features. We demonstrate that cellular transformation in this system is facilitated by the physical interaction between AGO2 and KRAS. Disruption of this interaction via expression of KRAS-binding AGO2 mutant proteins resulted in proliferation defects and loss of cell viability. Furthermore, ectopic AGO2 expression enhanced activity of KRAS-mediated signaling pathways, including NF-κB, which is often stimulated in ABC DLBCL. Downregulation of AGO2 resulted in loss of survival signals and induction of cell death. These results demonstrate a novel mechanism to induce lymphomagenesis through a functional interaction between KRAS and AGO2 and provide a novel model for the study of an ABC-like DLBCL.

Introduction

Several genes encoding components of the RNAi machinery have been found to be mutated in various human cancers, including *DROSHA* and *DICER1* (Lin et al., 2010; Lin S and Gregory RI, 2015). Mutations in these genes have been documented to alter gene expression programs via improper microRNA (miRNA) processing and are associated with uncontrolled cell proliferation and tumor growth (Kumar et al., 2007; Ravi et al., 2012; Wegert et al., 2015). By contrast, the involvement of members of the Argonaute family of proteins in cancer has been less well studied.

In mammals, four Argonaute (AGO) proteins are highly conserved and serve as critical components of the RNAi pathway (Drinnenberg et al., 2009). Their primary function under normal physiological conditions is to regulate gene expression programs that affect various biological processes, such as cell proliferation and differentiation, by guiding small RNAs to the 3'UTRs of their mRNA targets. As members of the RNA-induced silencing complex (RISC), AGO proteins promote target suppression through either mRNA destabilization and/or translational repression (Ambros 2004; Baek 2008; Bartel 2009).

Of these four AGO orthologues, AGO2 is the only member that is ubiquitously expressed in mammalian tissues and possesses catalytic activity for transcript destabilization, termed "slicing" (Liu et al., 2004). *Ago2* is an essential gene in the mouse and plays an important role in development and maintenance of cell identity by, for instance, regulating hematopoiesis and B cell differentiation, respectively, in a microRNA-dependent manner (O'Carroll et al.,

2007; Su et al., 2008). Despite the well-characterized role of AGO2 in the canonical RNAi pathway, little is known about how misregulation of AGO2 impacts pathological conditions.

AGO2 was previously identified to play a key role in tumorigenesis by facilitating activity of oncogenic microRNAs (Calin et al., 2002; Eis et al., 2005). However, AGO2 has also been shown to have distinct functions, some of which are microRNA-independent, in various human cancers. For example, in breast cancer, hypoxia in the core of high-grade tumors was found to enhance the association between EGFR and AGO2. This resulted in EGFR-mediated phosphorylation of AGO2-Y393, which inhibited loading of microRNAs with tumor suppressive properties due to AGO2 structural differences. The elevated level of p-AGO2-Y393 correlated with overall poor survival of breast cancer patients (Shen at al., 2013). In a separate study, the phosphorylation state of AGO2-Y393 was found to be regulated by protein tyrosine phosphate 1B (PTP1B) in oncogenic H-Ras-expressing fibroblasts (Yang et al., 2014).

Recently, a direct interaction between KRAS and AGO2 has been reported in oncogenic *KRAS*-driven cancers, such as numerous lung and pancreatic cancer cell lines. This interaction was found to occur on the endoplasmic reticulum, between the switch II domain of KRAS and the N-terminal domain of AGO2. Engagement of KRAS with AGO2 reduced RISC activity and promoted stabilization of oncogenic KRAS. Furthermore, the association between KRAS and AGO2 resulted in enhanced PI3K signaling and promoted cellular transformation (Shankar at al., 2016).

Activation of oncogenic KRAS typically results in enhanced activity of

many signaling cascades, including MAPK signaling (Drosten et al., 2010).

Interestingly, the MAPK signaling pathway has been implicated in the regulation

of AGO2 expression in transformed cells, including breast cancer models.

Elevated AGO2 expression in this model is associated with enhanced cell

proliferation and migratory ability (Adams et al., 2009).

AGO2 overexpression has been reported in many other human cancer

types, such as colon cancer and hepatocellular carcinoma. This upregulation of

AGO2 has been implicated in affecting tumor development and metastasis *in*

vivo (Li et al., 2010; Vaksman et al., 2012). Additionally, many cancer types have

been documented to have *AGO2*-associated copy number alterations, including

multiple myeloma (MM), hepatocellular carcinoma and non-small cell lung cancer

(NSCLC). For example, over 50% of lung adenocarcinomas (LUAD) have copy

number alterations affecting the *AGO2* locus, primarily as a result of gene

amplification events (Figure S1A) (Cerami et al., 2013). These copy number

alterations are significantly associated with poor disease outcome (Zhou et al.,

2010, Cheng et al., 2013).

As AGO2 is commonly overexpressed in many human cancers and recent

reports suggests that it has novel oncogenic properties, we sought to investigate

the role of AGO2 overexpression in tumor progression. We initially utilized a

genetically engineered "KP" autochthonous mouse model of human LUAD. The

KP mice harbor a heterozygous $Kras^{LSL-G12D}$ allele as well as homozygous

"floxed" *Tp53* alleles. Expression of viral-mediated Cre recombinase in these

lung epithelial cells results in activation of oncogenic KRAS and elimination of p53 function (DuPage et al., 2009). This KP model was modified to incorporate conditional overexpression of AGO2.

AGO2 overexpression in this system had the unexpected effect of initiating a B cell lymphoma with abnormal plasmablastic features, resembling ABC DLBCL, in these animals. We demonstrate that constitutive KRAS activation and AGO2 overexpression together stimulate downstream signaling cascades to promote cell survival in a lymphoid malignancy. Additionally, AGO2 downregulation resulted in a dampening of these signals, changes in cell identity and initiation of cell death.
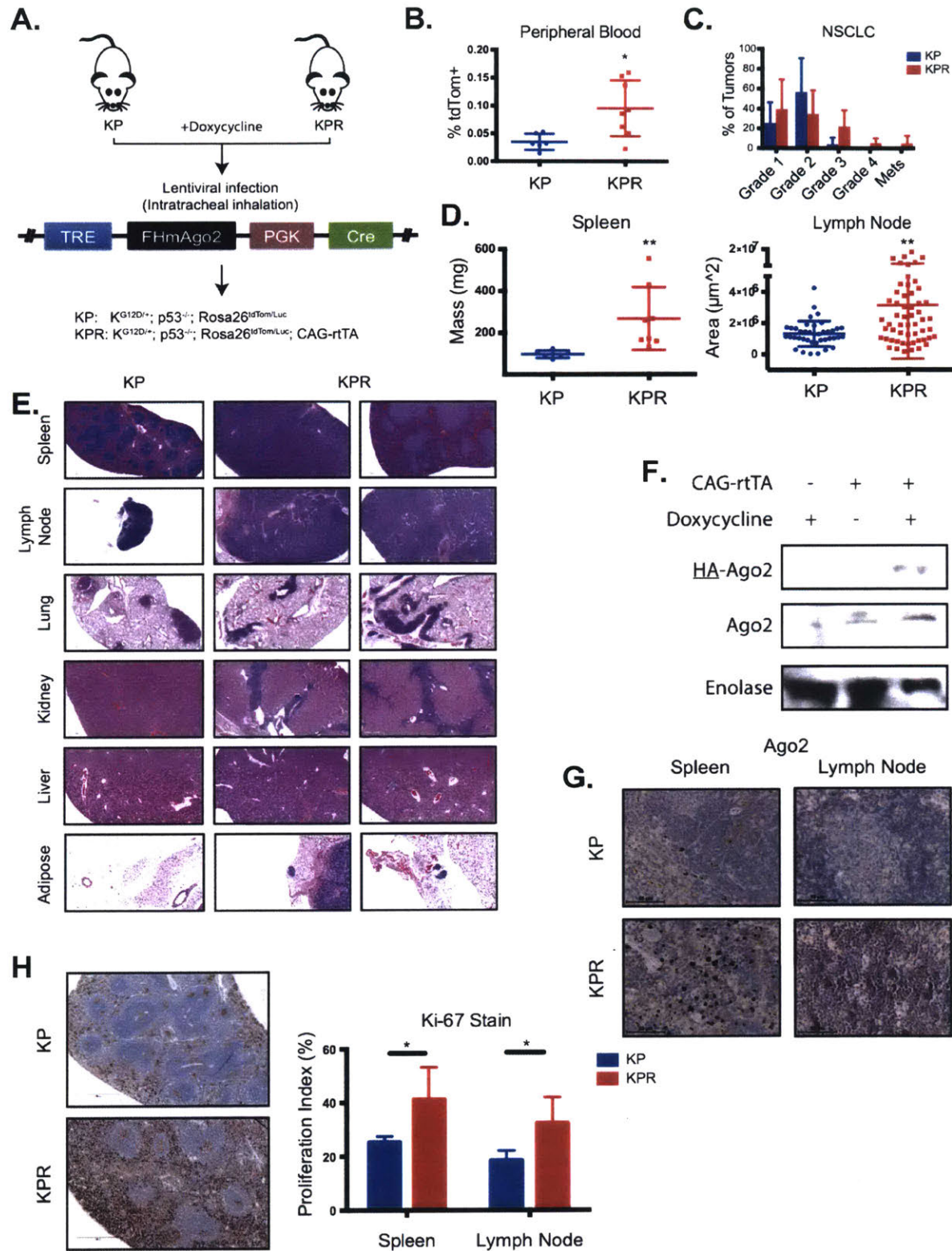

Results

Overexpression of AGO2 induces lymphomagenesis in KP autochthonous mouse model

We first set out to test the effects of AGO2 overexpression in LUAD. To this end, we generated a doxycycline-inducible, epitope tagged *Ago2* (FLAG-HA-AGO2) lentivirus co-expressing Cre recombinase. In this study, KP mice carried a germline transmitted CAG-rtTA transgene (KPR mice) for induction of doxycycline-regulated transcription from the lentiviral-integrated *Ago2* cDNA. Intratracheal administration of this virus in KPR mice resulted in deletion of *Tp53*, constitutive activation of oncogenic KRAS, and doxycycline-regulated AGO2 overexpression. These mice also carried reporter alleles, *Rosa26$^{LSL-tdTomato;LSL-Luciferase}$*, allowing for *in vivo* monitoring of Cre-expressing cells. In parallel, KP

mice lacking the rtTA allele were used to control for doxycycline-induced effects (Figure 1A).

Initially, we hypothesized that AGO2 overexpression in this system would accelerate lung tumor progression by taking into account our previous findings. From *in vitro* transwell experiments, an increase in cell migration and invasion were detected in AGO2 overexpressed KP LUAD cell lines over control (Figure S1B). Moreover, intravenous administration of these cell lines into immunocompromised mice resulted in an increase in tumor cell colonization of lung tissue in conditions with ectopic AGO2 expression (Figure S1C). These findings imply a functional role for AGO2 upregulation in promoting tumor progression by enhancing features associated with metastasis.

At week 4 post-infection in the presence of doxycycline, an increase in tdTomato-positive cells in circulation was detected from peripheral blood of KPR mice as compared to KP control (Figure 1B). *In vivo* imaging of these mice revealed a significant increase in luciferase signal at distant organ sites of KPR mice, suggesting AGO2 overexpression affects cell motility and promotes metastasis (Figure S1D). Furthermore, KPR mice displayed phenotypes associated with advanced stage cancer, such as weight loss and hunched posture, which was not seen in the control group. This suggested an acceleration of tumor progression with AGO2 overexpression, but upon sacrifice we detected no difference in lung tumor burden between KP and KPR mice (Figure S1E). However, a noticeable shift in tumor grades was detected, where malignant transformation was favored in KPR mice, consistent with our initial hypothesis of

106

A.

+Doxycycline

Lentiviral infection
(Intratracheal inhalation)

| TRE | FHmAgo2 | PGK | Cre |

KP: K^G12D/+; p53^-/-; Rosa26^tdTom/Luc
KPR: K^G12D/+; p53^-/-; Rosa26^tdTom/Luc; CAG-rtTA

B. Peripheral Blood

C. NSCLC

D. Spleen Lymph Node

E.

F.

| CAG-rtTA | − | + | + |
| Doxycycline | + | − | + |

HA-Ago2

Ago2

Enolase

G. Ago2
Spleen   Lymph Node

H. Ki-67 Stain

107

Figure 1. AGO2 overexpression in KP autochthonous model induces lymphomagenesis. (A) An autochthonous mouse model to overexpress AGO2 in a doxycycline-inducible system. Delivery of lentivirus via intratracheal inhalation resulted in Cre-mediated recombination of $KRAS^{LSL\text{-}G12D/+}$ and $Tp53^{fl/fl}$ primarily in lung epithelial cells. Cre-expressing cells overexpress AGO2 in mice harboring CAG-rtTA transgene (KPR mice) in the presence of doxycycline. KP mice were used as controls for doxycycline-induced effects. (B) Peripheral blood was collected via retro-orbital bleeding and analyzed for tdTomato-expressing cells from KPR (n=8) and KP mice (n=6) 16 wk after lentiviral delivery. (C) Lung tumors were scored for histological grade (1/2= benign; 3/4=malignant). (D) Mass and size of spleen and lymph nodes (LN), respectively, were measured in KPR mice (spleen: n=8; LN: n=56) and compared to KP mice (spleen: n=6; LN: n=40) at 20 wk post-lentiviral delivery. (E) Haematoxylin and eosin (H&E) staining showed an effacement of lymphoid tissue with prominent lymphoid infiltrates. (F) AGO2 expression was analyzed from cross sections of KPR and KP mouse tissue using IHC. (G) Splenocytes were harvested from KPR and KP mice cultured *in vitro*. Western blot analysis was used to determine protein expression of total AGO2 and HA-tagged AGO2 from these splenocytes. Enolase was used as a loading control. (H) Ki-67 staining using IHC was quantified to determine the proliferation index of lymphocytes from KPR mice (n=6) compared to KP mice (n=4). $*p < 0.05$, $**p < 0.01$.

AGO2 accelerating lung tumor progression (Figure 1C).

Surprisingly, full necropsy of all KPR mice revealed significantly enlarged secondary lymphoid organs, such as spleen and various lymph nodes, that were not observed in the KP control mice (Figures 1D and S1F). Histology displayed effacement of these lymphoid structures as well as infiltration of other distant organ sites, including kidneys, due to massive expansion of lymphocytes (Figure 1E). Moreover, we detected extensive infiltration of adipose tissue by lymphoid aggregates in KPR mice, consistent with lymphoid malignancy. Conversely, all KP mice displayed normal lymphoid tissue and lymphoid infiltrates were not observed in distant organs.

To determine whether these expanding lymphocyte populations displayed AGO2 overexpression, we used immunohistochemistry (IHC) and western blot analysis. We detected an increase in AGO2-expressing lymphocytes from cross-sections of lymphoid tissue of KPR mice compared to KP control group (Figures 1F and S1G). At the time of sacrifice, splenocytes were also harvested from both KP and KPR mice. The virally introduced HA-tagged AGO2 protein was only detected in KPR splenocytes in the presence of doxycycline, suggesting lymphocytes were transduced *in vivo* via intratracheal administration of the virus. Doxycycline-induced AGO2 expression resulted in a two-fold increase in total AGO2 expression in KPR cells. Predictably, the epitope tagged AGO2 protein was not detectable in KP splenocytes (Figure 1G).

As secondary lymphoid structures were enlarged and many non-lymphoid tissue displayed notable lymphocyte infiltration, we sought to compare the

proliferation index of splenocytes between KP and KPR mice. Unlike KP control, the majority of KPR lymphocytes stained positively for Ki-67, with a proliferation fraction of over 40%. Additionally, many of these cells were actively undergoing mitosis as detected by phospho-H3 staining (Figures 1H and S1H), confirming the hyperplastic phenotype. Together, the extensive infiltration of nodal and extranodal tissues by lymphoid infiltrates expressing HA-tagged AGO2 and high proliferation fraction are consistent with lymphomagenesis resulting from AGO2 overexpression in the context of activated KRAS and *Tp53* deletion.
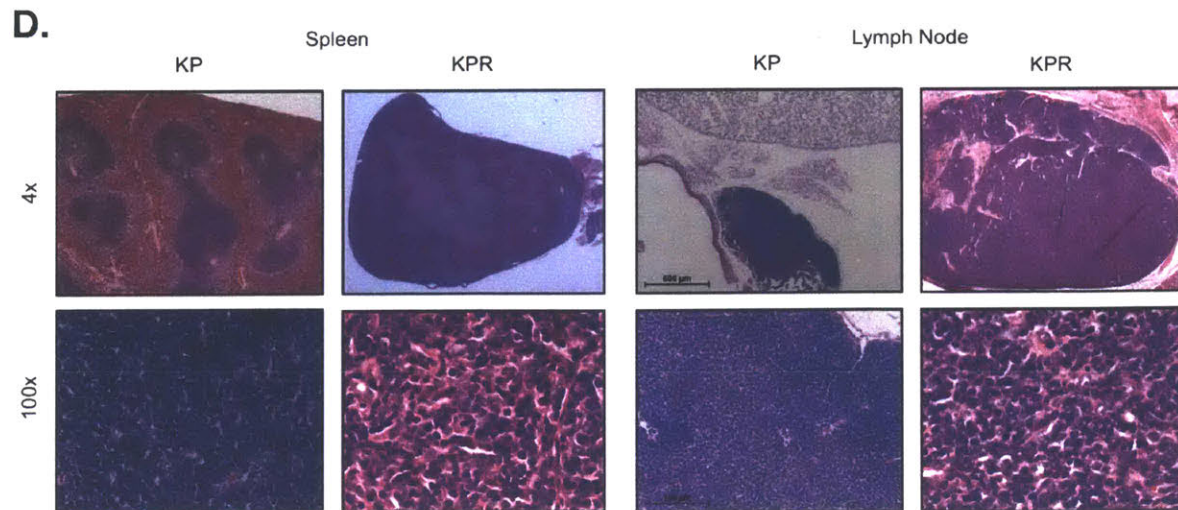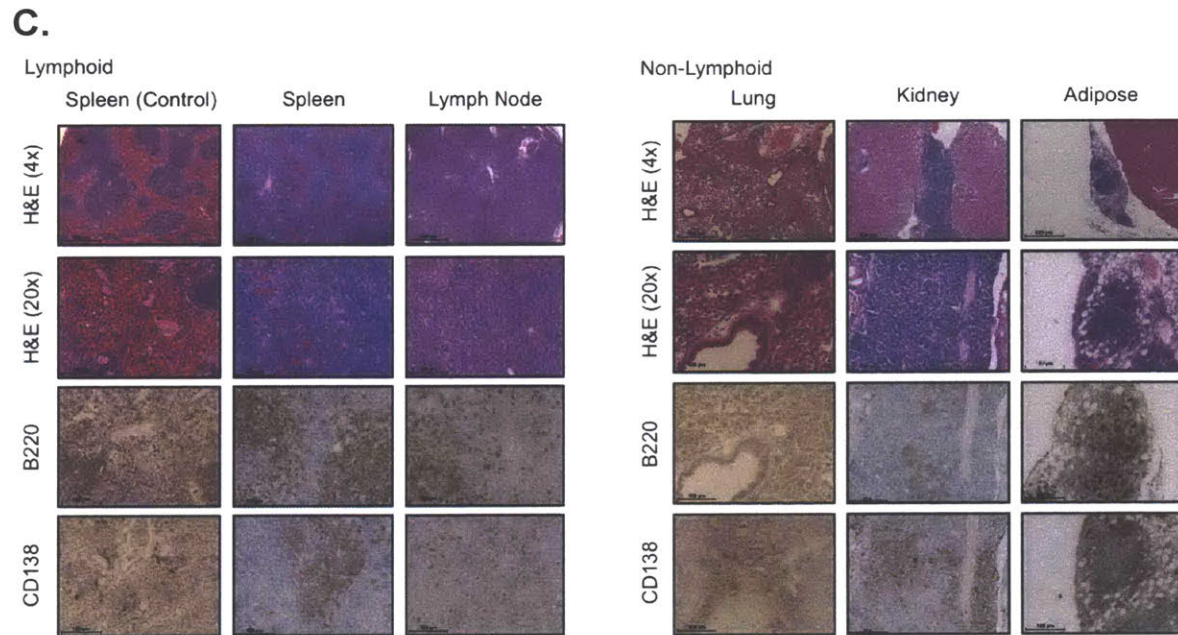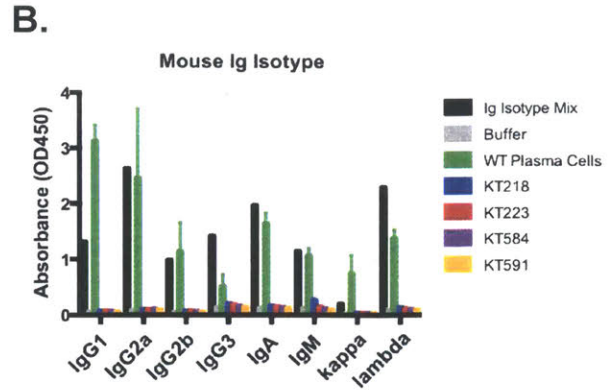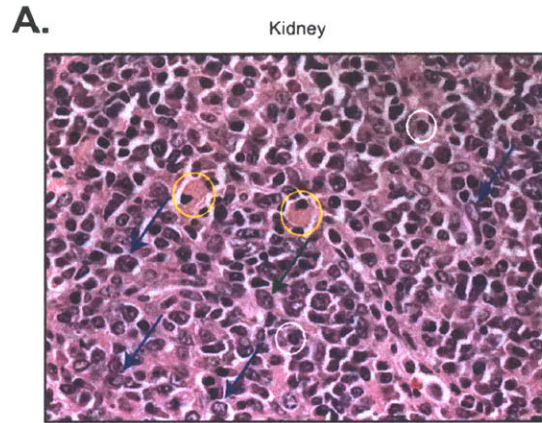
Overexpression of AGO2 induces ABC-like DLBCL with plasmablastic features

Closer examination of KPR lymphoid aggregates revealed the infiltrates consisted of a heterogeneous proliferative lymphoplasmacytic cell population. These infiltrates were predominated by medium-sized to large abnormal lymphocytes with scant to moderate cytoplasm and vesicular nuclei. These abnormal cells frequently had eccentric nuclei with coarse to vesicular chromatin and moderate to abundant cytoplasm with perinuclear clearing, consistent with plasmablastic cells. The infiltrates also included numerous maturing plasma cells, with frequent intracytoplasmic and intranuclear inclusions, such as Russell bodies and Dutcher bodies, respectively, indicative of defective antibody secretion (Figures 2A and S2A). Immunoglobulin isotyping of these abnormal plasmablastic cells in culture confirm that they were defective in antibody secretion (Figure 2B).

110

IHC and flow cytometry (FACS) showed that a large population of the KPR lymphocytic infiltrates were B220-expressing B cells and included abundant CD138-positive cells, confirming plasmacytic differentiation (Figures 2C and S2B). The majority of these abnormal lymphocytes were also positive for IRF4-encoded MUM1 protein, an early regulator of plasma cell differentiation that is often expressed in multiple myeloma (MM) and plasmablastic lymphoma (PBL) patients (Figure S2C) (Vega et al., 2005). Moreover, the vast majority of these proliferative infiltrates were negative for CD3, consistent with a B cell specific malignancy (Figure S2D). KP splenocytes derived from control mice displayed a normal fraction of B220- and CD138-expressing cells expected of healthy mice, confirming that the plasmacytic differentiation phenotype occurred exclusively in the doxycycline-treated KPR mice (Colovai et al., 2004).

Given the high percentage of abnormal plasma cells, we sought to assess the lymphoplasmacytic infiltrates for evidence of clonality using stains for kappa/lambda immunoglobulin light chains. The majority of the KPR mice (10/14) showed expansion of lambda-positive cells, while the remaining KPR mice harbored infiltrates with kappa light chain restriction. This preferential selection for light chain isoforms is consistent with a clonal expansion originating from a small subset of neoplastic cells (Figure S2E).

The combination of the morphology and immunophenotype of these hyperproliferative cells were consistent with a plasmablastic neoplasm phenotypically resembling human ABC DLBCL with plasmablastic features. To confirm that the hyperplastic KPR lymphocytes were tumorigenic, we

111

A.

Kidney

B.

Mouse Ig Isotype

- Ig Isotype Mix
- Buffer
- WT Plasma Cells
- KT218
- KT223
- KT584
- KT591

C.

Lymphoid

	Spleen (Control)	Spleen	Lymph Node
H&E (4x)			
H&E (20x)			
B220			
CD138			

Non-Lymphoid

	Lung	Kidney	Adipose
H&E (4x)			
H&E (20x)			
B220			
CD138			

D.

Spleen

	KP	KPR
4x		
100x		

Lymph Node

	KP	KPR
4x		
100x		

112

Figure 2. Mice develop an ABC-like DLBCL phenotype from AGO2 overexpression in autochthonous model. (A) H&E staining of a KPR kidney cross-section (yellow circles= Mott cells; white circles= plasma cells with Russell bodies; blue arrows= large neoplastic cells with vesicular nuclei and open chromatin; green arrows= large neoplastic cells with scant cytoplasm). (B) An ELISA-based approach for mouse immunoglobulin isotyping was used to determine levels of antibody secretion (positive control= premixed Ig control mixture; negative control= wash buffer; WT plasma cells= cell culture medium from WT CD138-expressing plasma cells grown overnight). Cell culture medium was collected from KPR cell lines grown in the presence of doxycycline. (C) Serial cross-sections of tissue from KP and KPR mice at 20 wk after lentiviral administration were analyzed for expression of B220 (B cell specific marker) and CD138 (plasma-cell specific marker) with IHC (Bottom panels) in addition to H&E staining (top panel). (D) Splenocytes from KP and KPR mice were transplanted into recipient sublethally irradiated mice. Histology and H&E staining at 18 wk post-transplantation was used to assess the extent of tissue effacement and lymphocytic infiltration from KPR and KP splenocytes in the presence of doxycycline.

transplanted these cells into sublethally-irradiated recipient mice. Transplantation in the presence of doxycycline resulted in the development of a similar ABC-like DLBCL neoplasm in recipient mice, confirming that AGO2 overexpression in this setting leads to tumorigenesis and gives rise to an aggressive lymphoma (Figure 2D). Moreover, these lymphoid neoplasms were predominantly CD10-negative, consistent with an activated B cell malignancy (Figure S2F) (Hans et al., 2004).

Transplanting AGO2 overexpressed B lymphocytes into irradiated mice induced an ABC-like DLBCL phenotype

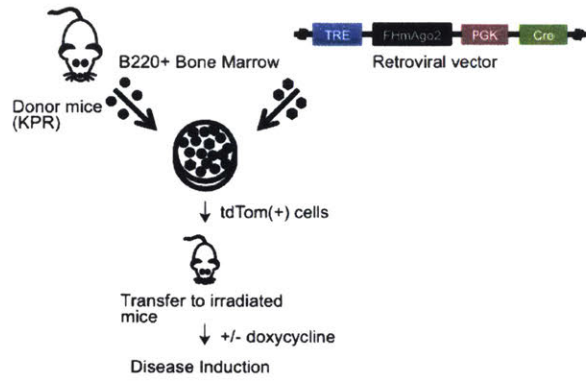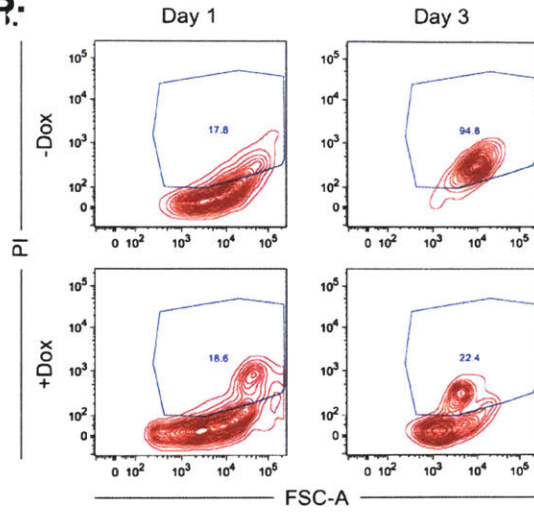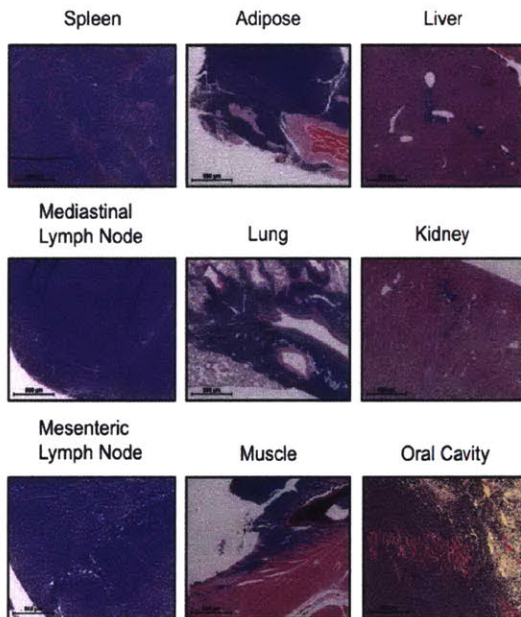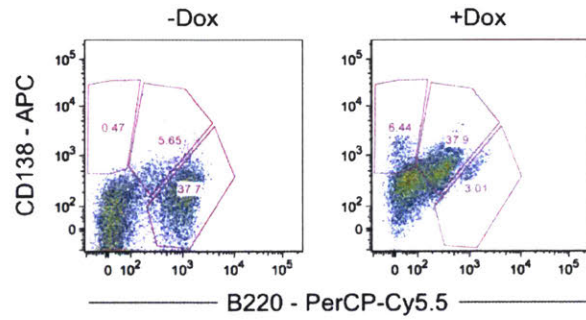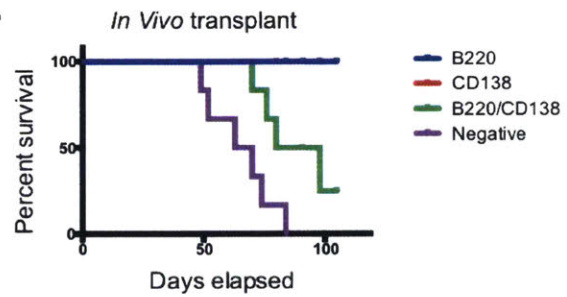To develop a more robust model of ABC DLBCL involving overexpression of AGO2, we investigated the effects of direct infection of early stage B cells from KPR mice. B220-positive B cells were isolated from bone marrow of healthy KPR mice harboring the $Rosa26^{LSL-tdTomato/LSL-Luciferase}$ reporter allele. These bone marrow-derived, B220-positive lymphocytes were retrovirally transduced *in vitro* with a vector that co-expressed Cre recombinase and a doxycycline-inducible *Ago2* construct (Figure 3A). Following infection, tdTomato-positive populations were isolated through FACS for subsequent studies.

TdTomato-positive cells were subsequently passaged in tissue culture in the presence and absence of doxycycline. As expected, continual passaging of primary B-lymphocytes in medium lacking doxycycline led to a gradual decline in cell viability in the absence of IL-7, stem cell factor (SCF) and stromal feeder cells. However, a population of the KPR lymphocytes expanded under these conditions *in vitro* in the presence of doxycycline, where the percentage of

propidium iodide (PI)-negative cells exceeded that of PI-positive cells after two weeks in culture (Figure S3A). This suggested the combination of KRAS activation, *Tp53* deletion and AGO2 overexpression might be sufficient to maintain cell viability or transform B-lymphocytes *in vitro*.

The dependency on AGO2 overexpression for cell survival was tested further by passaging KPR cells that were derived in the presence of doxycycline in medium lacking the drug. At day 3 post-doxycycline withdrawal, KPR lymphocytes failed to survive as determined by PI staining. However, these cells continued to proliferate with minimal cell death detected in the presence of doxycycline (Figure 3B). Thus the transduced KPR lymphocytes were dependent on AGO2 overexpression for survival.

In parallel, tdTomato-positive KPR lymphocytes were transferred into sublethally irradiated mice via intravenous injection to test for tumorigenicity. Transplantation of these KPR cells into recipient mice fed doxycycline-containing food recapitulated the ABC-like DLBCL phenotype observed from the KPR autochthonous model. Luciferase-expressing lymphocytes expanded well *in vivo* and exhibited both hyperproliferation and extensive infiltration of nodal and extranodal tissues (Figures 3C and S3B). Moreover, prominent populations of td-Tomato-expressing splenocytes were detected from doxycycline-treated mice, demonstrating engraftment of transplanted lymphocytes (Figure S3C). In the absence of doxycycline, transplanted mice appeared to be healthy and showed no signs of lymphocytic hyperproliferation.

A.

Donor mice (KPR) → B220+ Bone Marrow

Retroviral vector: TRE — FHmAgo2 — PGK — Cre

↓ tdTom(+) cells

Transfer to irradiated mice

↓ +/- doxycycline

Disease Induction

B.

Day 1 | Day 3

PI vs FSC-A

-Dox: 17.8 (Day 1), 94.8 (Day 3)

+Dox: 18.6 (Day 1), 22.4 (Day 3)

C.

Spleen | Adipose | Liver

Mediastinal Lymph Node | Lung | Kidney

Mesenteric Lymph Node | Muscle | Oral Cavity

D.

CD138 - APC vs B220 - PerCP-Cy5.5

-Dox: 0.47, 5.65, 37.7

+Dox: 6.44, 37.9, 3.01

E.

In Vivo transplant

Percent survival vs Days elapsed

- B220
- CD138
- B220/CD138
- Negative

116

Figure 3. Transfer of AGO2 overexpressed B cells into mice induces ABC-like DLBCL phenotype. (A) A transplant mouse model to overexpress AGO2 in a doxycycline-inducible system to initiate lymphomagenesis. *In vitro* retroviral-mediated transduction of B220-expressing primary KPR bone marrow resulted in Cre-mediated recombination of $Kras^{LSL-G12D/+}$, $p53^{fl/fl}$, and $Rosa26^{LSL-tdTom/LSL-Luc}$. TdTomato-expressing cells were sorted and subsequently transplanted into recipient sublethally irradiated mice. These mice were divided into two cohorts: doxycycline-containing food and regular food. (B) KPR lymphocytes dependent on AGO2 overexpression were passaged into +/- doxycycline conditions. PI staining was used to monitor for cell death at 1 and 3 d after cells were passaged. (C) H&E staining from sections of tissue from mice transplanted with KPR lymphocytes and treated with doxycycline at 20 wk after transplantation. (D) Splenocytes derived from mice (either treated or untreated with doxycycline) were compared with FACS analysis based on B220 and CD138 expression. (E) AGO2 overexpression-dependent KPR cells were sorted into distinct populations based on B220 and CD138 expression and subsequently transferred into recipient sublethally irradiated mice that were fed doxycycline-containing food. Survival curve from this transplantation revealed two tumorigenic populations: $B220^{low}/CD138^+$ and DN.

To confirm that the hyperproliferative population of B cells were derived from the transplanted KPR cells, we isolated splenocytes from recipient mice and performed FACS analysis to detect tdTomato-expressing cells. As expected, the majority of splenocytes harvested from doxycycline-treated mice were tdTomato-positive but were not detected from the untreated condition (Figure S3C). This indicated that the transplanted KPR cells were successfully seeded in recipient mice and expanded well *in vivo* only in the AGO2-overexpressed condition.

Unlike the control mice on regular chow, we observed that mice on doxycycline treatment to induce AGO2 overexpression led to B cells with plasmacytic differentiation. FACS analysis for CD138 and B220 demonstrated that a prominent population of splenic B cells co-expressed CD138 and low levels of B220 (B220^{low}), characteristic of plasmablastic cells. As expected, many of these CD138-expressing cells also expressed IRF4 (Figure S3D). Interestingly, the population of terminally differentiated plasma cells, as characterized by CD138-positive and B220^{neg} expression, was significantly lower than the immature plasma cells, suggesting that the tumorigenic cells were arrested in the plasmablastic stage of B cell development, consistent with ABC DLBCL (Figures 3D and S3D).

Subsequent transplantation of lymphoma cells induced with AGO2 overexpression into recipient animals caused rapid death of the recipient mice compared to both the parental tumor-bearing mice and control mice not treated with doxycycline (Figure S3E), suggesting the presence of an aggressive clonal population. To address the clonality of these cells, we used a J_H4 probe to

analyze VDJ recombination events via Southern blotting (Zhu et al., 1996). This analysis revealed that the KPR lymphomas were clonal, as a unique VDJ recombination event was detected in each sample. Predominant VDJ recombination events were not detected in the control cell lines generated from a heterogeneous population of wild type (WT) splenocytes (Figure S3F).

We next sought to determine the tumorigenic potential of distinct subpopulations of lymphoma cells based off of immunophenotyping. KPR lymphoma cells were sorted into four gates based on B220 and CD138 expression (Figure S3G). Both $CD138^+/B220^{low}$ and "double negative" (DN) populations were able to proliferate in culture in the presence of doxycycline, but neither the $B220^{high}$ and $CD138^+/B220^{neg}$ populations expanded *in vitro* under the conditions tested (Figure S3H).

To confirm the tumorigenicity of these sorted populations, we transferred them into recipient mice and monitored for expansion *in vivo*. As expected, both the $CD138(+)/B220^{low}$ and DN populations were tumorigenic as transferring them into recipient mice resulted in reduced survival compared to both the $B220^{high}$ and mature plasma cell populations (Figure 3E). Interestingly, although all recipient animals that developed lymphoma presented with an ABC-like DLBCL phenotype, the DN population of cells appeared to produce a more aggressive form of the disease than the $CD138^+/B220^{low}$ population. This further confirmed that a diverse spectrum of abnormal plasmablasts was present with distinct subpopulations displaying greater tumorigenicity than others.

Withdrawal of AGO2 overexpression perturbs arrest in plasmablastic stage and promotes apoptosis

We confirmed that outgrowth of KPR tumor cell lines were depleted of p53 protein as a consequence of Cre-mediated recombination prior to *in vivo* transplantation (Figure S4A). As withdrawal of doxycycline from the AGO2 overexpression-dependent tumor cells resulted in significant cell death, we sought to determine the mechanism promoting their loss of viability. FACS analysis with Annexin V and PI revealed that reduced AGO2 expression resulted in a dramatic increase in cell death 48 h. post-doxycycline withdrawal, where a prominent population lost viability via apoptosis (Figure 4A). Caspase-3 activity and PARP1 cleavage, indicative of apoptotic cells, were detectable in the doxycycline-withdrawn KPR cells, corroborating loss of cellular viability (Figures 4B and S4B).

Interestingly, downregulation of exogenous AGO2 expression also affected the differentiation state of the plasmablastic tumor cells. The majority of cells overexpressing AGO2 were CD19-positive and expressed low levels of B220, but two distinct populations of cells emerged upon withdrawal of doxycycline. One population was CD19$^-$/B220^{low}, resembling the expression profile of mature plasma cells, and the second population was CD19$^+$/B220^{high}, resembling mature B-lymphocytes (Figure 4C). Furthermore, there was a dramatic reduction in the abundance of IRF4-expressing cells upon doxycycline withdrawal (Figure 4D). These data suggests that Ago2 overexpression causes arrest in the highly proliferative plasmablastic stage, and withdrawal of
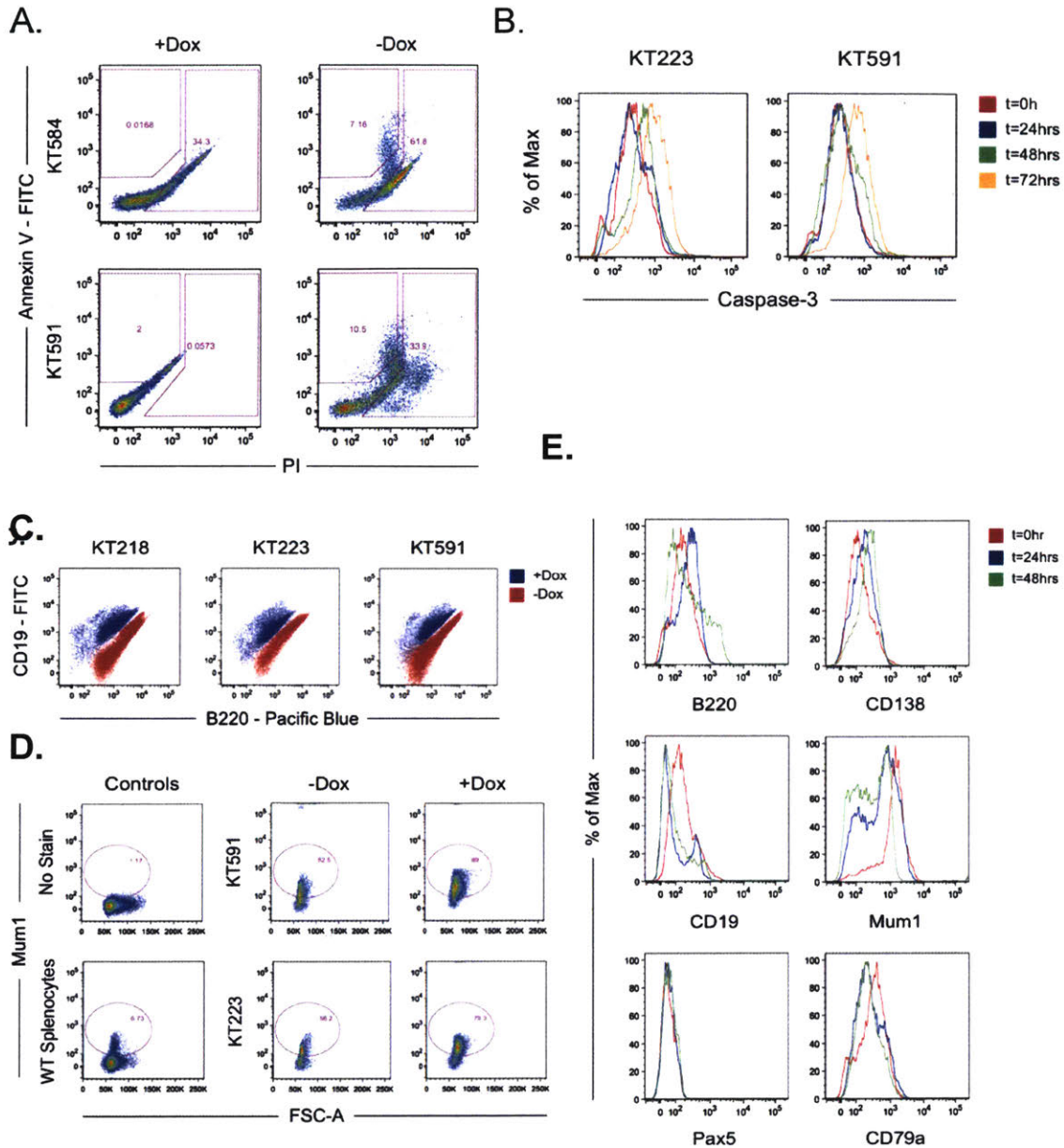
Figure 4. Loss of Ago2 overexpression perturbs block in plasmablastic stage and promotes cell death. (A) Apoptosis analysis using Annexin V and PI staining was used on two Ago2 overexpression-dependent KPR lymphoma cell lines (KT584 and KT591) in the presence and absence of doxycycline. (B) Caspase-3 staining on KPR cells at 0, 24, 48, and 72 h after doxycycline withdrawal was used to confirm the initiation of apoptosis. (C) KPR cell lines were analyzed for CD19 expression using FACs analysis. (D) Irf4 expression was examined from KPR cells to determine population of cells exhibiting plasmacytic differentiation. (E) The kinetics of gene expression changes of specific B cell (B220, CD19, Pax5, CD79a) and plasma cell markers (CD138, Irf4) were analyzed with FACS analysis. With the exception of Irf4, changes in expression of these markers do not occur until after 24 h of doxycycline withdrawal.

121

doxycycline relieved this differentiation block, resulting in cell death.

To explore the kinetics of the resumption of differentiation following loss of AGO2 overexpression, expression of various B cell markers were characterized in a time course during doxycycline withdrawal. There appeared to be minimal changes after 24 h of doxycycline withdrawal regarding B220 and CD138 expression, but a large shift toward CD19-negative cells was noted, in agreement with terminally differentiated plasma cells. Interestingly, KPR lymphoma cells appeared to be predominantly PAX5-negative, an early B cell marker often misregulated in DLBCL, and no changes were detected upon doxycycline withdrawal. Furthermore, there was a slight shift toward a cell population with lower CD79a expression, suggesting a decrease in BCR signaling with loss of ectopic AGO2 expression (Figure 4E). These findings are consistent with the immunophenotype of human ABC DLBCL (Castillo et al., 2015).

Inactivation of BLIMP1, the master regulator of plasma cell differentiation, is a common feature in ABC DLBCL. Homozygous deletion of *BLIMP1* is frequently detected from these patients (Mandelbaum et al., 2010). However, as we overexpressed AGO2 in our system, we reasoned that inactivation of BLIMP1 could be caused by miRNA-mediated repression. We first analyzed expression of let-7a and miR-9-5p, both of which are predicted to target *Blimp1* based off of seed binding sites, in KPR cells in the presence and absence of doxycycline. KPR cells treated with doxycycline expressed high levels of both of these miRNAs, but they were downregulated upon withdrawal of doxycycline (Figure S4C). This suggests that KPR cells were arrested in the plasmablastic stage
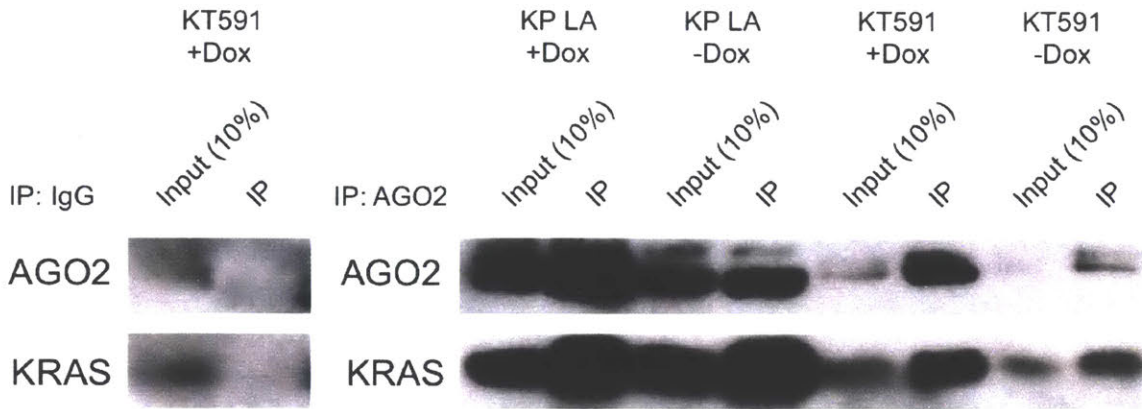
primarily through miRNA-mediated repression of *Blimp1*, and this differentiation

block was relieved upon loss of exogenous AGO2. Furthermore, miR-148a, a

positive regulator of *Blimp1*, was upregulated upon withdrawal of exogenous

AGO2 (Figure S4C). These findings suggests that overexpression of AGO2 in

KPR cells enhance activity of a subset of miRNAs responsible for blocking

terminal differentiation into plasma cells.

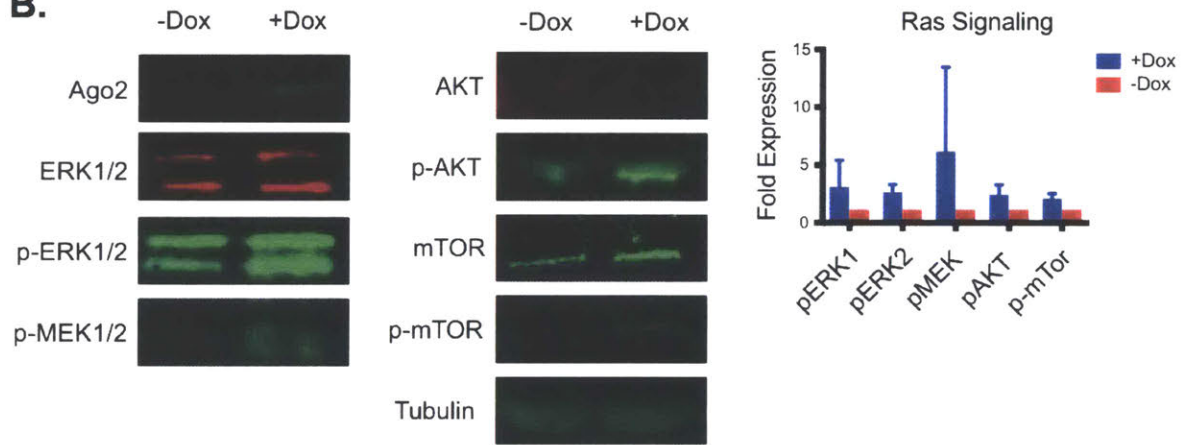## Lymphoma cells are dependent on interaction between AGO2 and oncogenic KRAS for cell survival

Based on the results of Shankar *et al.* (2016) on the physical and functional

interaction between AGO2 and oncogenic KRAS, we investigated whether this

mechanism might underlie the effects of AGO2 overexpression in this system.

Co-immunoprecipitation (CoIP) showed that AGO2-KRAS complexes could be

detected in the AGO2 overexpressing lymphoma cells (Figures 5A and S5A). As

this physical interaction facilitated cellular transformation in various solid tumor

systems, we sought to evaluate whether upregulation of AGO2 affects signaling

pathways downstream of KRAS. With AGO2 overexpression, we detected active

signaling from both the MAPK and PI3K pathways (Figure 5B). Upon doxycycline

withdrawal, there was a significant decrease in activity of both pathways,

consistent with these pathways being crucial for cell survival in the AGO2

overexpressing cells.

To determine whether this interaction was necessary to promote cell

proliferation and survival in this system, we generated KRAs binding deficient
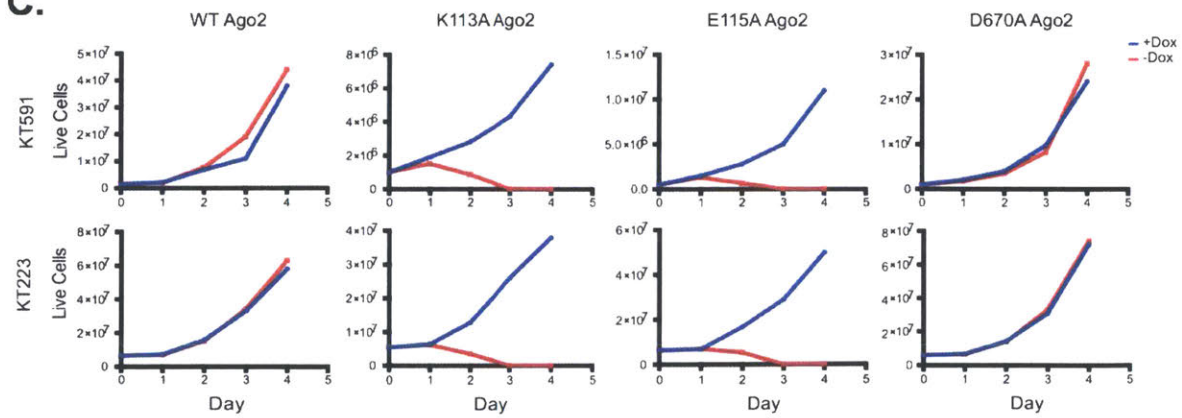
A.

KT591
+Dox

KP LA
+Dox

KP LA
−Dox

KT591
+Dox

KT591
−Dox

IP: IgG   Input (10%)   IP   IP: AGO2   Input (10%)   IP   Input (10%)   IP   Input (10%)   IP   Input (10%)   IP

AGO2   AGO2

KRAS   KRAS

B.

−Dox   +Dox   −Dox   +Dox   Ras Signaling

Ago2   AKT

ERK1/2   p-AKT

p-ERK1/2   mTOR

p-MEK1/2   p-mTOR

                               Tubulin

Fold Expression: pERK1, pERK2, pMEK, pAKT, p-mTor

+Dox (blue), −Dox (red)

C.

WT Ago2   K113A Ago2   E115A Ago2   D670A Ago2

KT591 — Live Cells vs Day
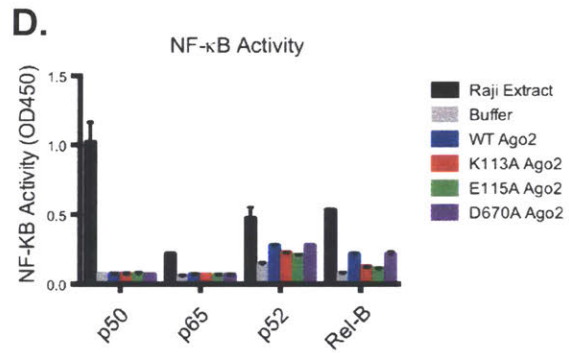
KT223 — Live Cells vs Day
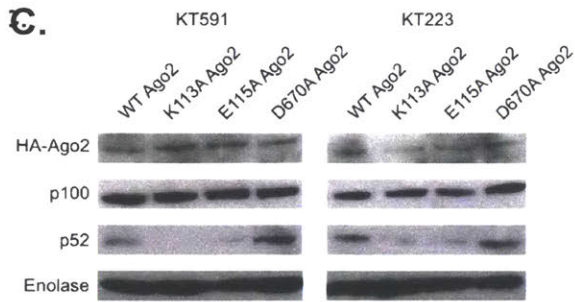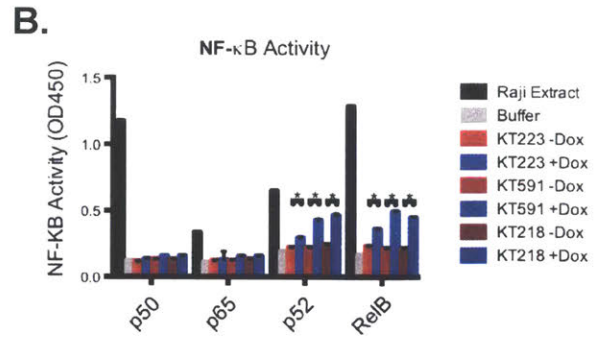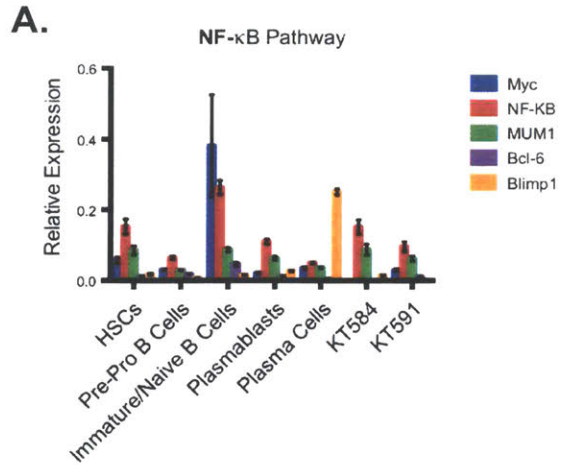
+Dox (blue), −Dox (red)

Figure 5. KPR lymphocytes are dependent on the AGO2-KRAS complex for cell survival. (A) The interaction between KRAs and AGO2 was assessed in KPR cells (KT591) by AGO2 immunoprecipitation (IP) followed by immunoblotting (AGO2, KRAS). Protein extract from a KP lung adenocarcinoma was used as a positive control for KRAS. An IgG IP control was used to test for antibody specificity. (B) MAPK, PI3K, and mTOR signaling was analyzed with western blotting using KPR cells in the presence and absence of doxycycline. Quantification of effector proteins (n=6) showed a significant increase in activity from all 3 signaling pathways when AGO2 is overexpressed (n=6). (C) The dependency on the AGO2-KRAS complex to promote cell survival was explored by generating KRAS-binding AGO2 mutants (K113A and E115A). The slicer-deficient AGO2 mutant (D670A) was used here as a WT AGO2 mimic. These cMyc-tagged mutants, in addition to the WT isoform of AGO2, were constitutively expressed with a pGK promoter in KPR lymphoma cells in both the presence and absence of doxycycline.

AGO2 mutants. Shankar *et al.* (2016) previously reported critical amino acid residues for AGO2-KRAS binding in human cancer cell lines. We confirmed that mutations of these residues, K113A and E115A, in our mouse system perturb the AGO2-KRAS interaction (Figure 5SB). Additionally, the AGO2 slicer mutant, D670A, which lacks catalytic RNase activity, was generated as a control to mimic the wild type AGO2 interaction with KRAS. Doxycycline-inducible AGO2 tumor-derived cell lines were infected *in vitro* with a lentiviral construct that overexpressed wild type or mutant AGO2 protein (PGK-cMyc-AGO2-EFS-Venus); the virus encodes Venus as a marker of transduced cells, allowing for FACS sorting of Venus-positive populations for subsequent functional studies.

The dependency on the AGO2-KRAS interaction to promote cellular survival was tested by sorting Venus-expressing cells into medium containing or lacking doxycycline, in which populations maintained on doxycycline continually expressed the doxycycline-inducible wild type isoform of AGO2 in addition to the various constitutively expressed AGO2 variants (WT, K113A, E115A, D670A). Intriguingly, in the absence of doxycycline, tumor cells overexpressing KRAS-binding deficient AGO2 mutants exhibited proliferation defects by day 1 of doxycycline withdrawal and ultimately failed to survive (Figure S5D). By contrast, overexpression of either WT-AGO2 or D670A-AGO2 rescued cell survival in the absence of doxycycline (Figures 5C). These data suggest that the AGO2-KRAS interaction is critical to promote cellular transformation and survival in our model system. Of note, we confirmed that the various AGO2 mutants could be detected in the absence of doxycycline in the proliferating tumor cells (Figure S5E).

Oncogenic KRAS and AGO2 interaction potentiate NF-κB signaling to promote plasmacytic differentiation

NF-κB activation plays an integral role in promoting B cell differentiation and cellular transformation in the context of ABC DLBCL (Davis et al., 2001). Because the AGO2-overexpressing tumor cells had a plasmablastic phenotype, we investigated whether regulation of NF-κB signaling was involved in the increased proliferation and defect in differentiation described above. To this end, we investigated specific components of the NF-κB signaling pathway critical for plasma cell differentiation. Transcriptionally, these lymphoma cells expressed relatively high levels of *Nfκb2* and *Irf4*, but failed to express *Blimp1*, the master regulator of plasma cell differentiation (Figure 6A) (Pasqualucci et al., 2006). To determine the specific subunits of the NF-κB complex that are active in the lymphoma cells, we used an ELISA-based technique to quantitate their levels in cellular extracts derived from KPR cells in the presence and absence of doxycycline. There was an increase in activity of NF-κB in the AGO2-overexpressing lymphoma cells, associated with activation of p52 and Rel-B specifically (Figures 6B and S6A). These two subunits are known to Heterodimerize and translocate to the nucleus when activated through the non-canonical NF-κB signaling pathway. Interestingly, miR-155, a miRNA highly expressed in many DLBCL and involved in stimulating NF-κB activity, was upregulated upon AGO2 overexpression (Figure S6B) (Kluiver et al., 2005). This suggests that ectopic expression of AGO2 stabilizes miR-155 to promote pro-survival signals.

127

A. NF-κB Pathway

Relative Expression

Myc
NF-KB
MUM1
Bcl-6
Blimp1

HSCs, Pre-Pro B Cells, Immature/Naive B Cells, Plasmablasts, Plasma Cells, KT584, KT591

B. NF-κB Activity

NF-KB Activity (OD450)

Raji Extract
Buffer
KT223 -Dox
KT223 +Dox
KT591 -Dox
KT591 +Dox
KT218 -Dox
KT218 +Dox

p50, p65, p52, RelB

C.

KT591   KT223

WT Ago2, K113A Ago2, E115A Ago2, D670A Ago2   WT Ago2, K113A Ago2, E115A Ago2, D670A Ago2

HA-Ago2
p100
p52
Enolase

D. NF-κB Activity

NF-KB Activity (OD450)

Raji Extract
Buffer
WT Ago2
K113A Ago2
E115A Ago2
D670A Ago2

p50, p65, p52, Rel-B

Figure 6. Interaction between oncogenic KRAS and AGO2 promotes NF-κB activation. (A) Relative expression of various components of NF-κB pathway involved in plasmacytic differentiation was analyzed with qPCR. RNA from control cells (sorted populations of WT lymphocytes) were harvested from healthy mice and compared to RNA from AGO2 overexpressed lymphoma cell lines. (B) Activity of NF-κB subunits was measured by an ELISA assay. Whole cell extracts were harvested from KPR cell lines in the presence and absence of doxycycline. Raji extract and buffer were used as positive and negative controls, respectively. (C) Western blot analysis was used to compare levels of p100 processing in doxycycline withdrawn (24 h) KPR cells. These cells constitutively overexpress various isoforms of AGO2 (WT, K113A, E115A, D670A). (D) Whole cell extracts from KT591 cell lines overexpressing various AGO2 isoforms (WT, K113A, E115A, and D670A) were used to determine the activity of NF-κB subunits using an ELISA assay. These cells were grown in the absence of doxycycline for 24 h.
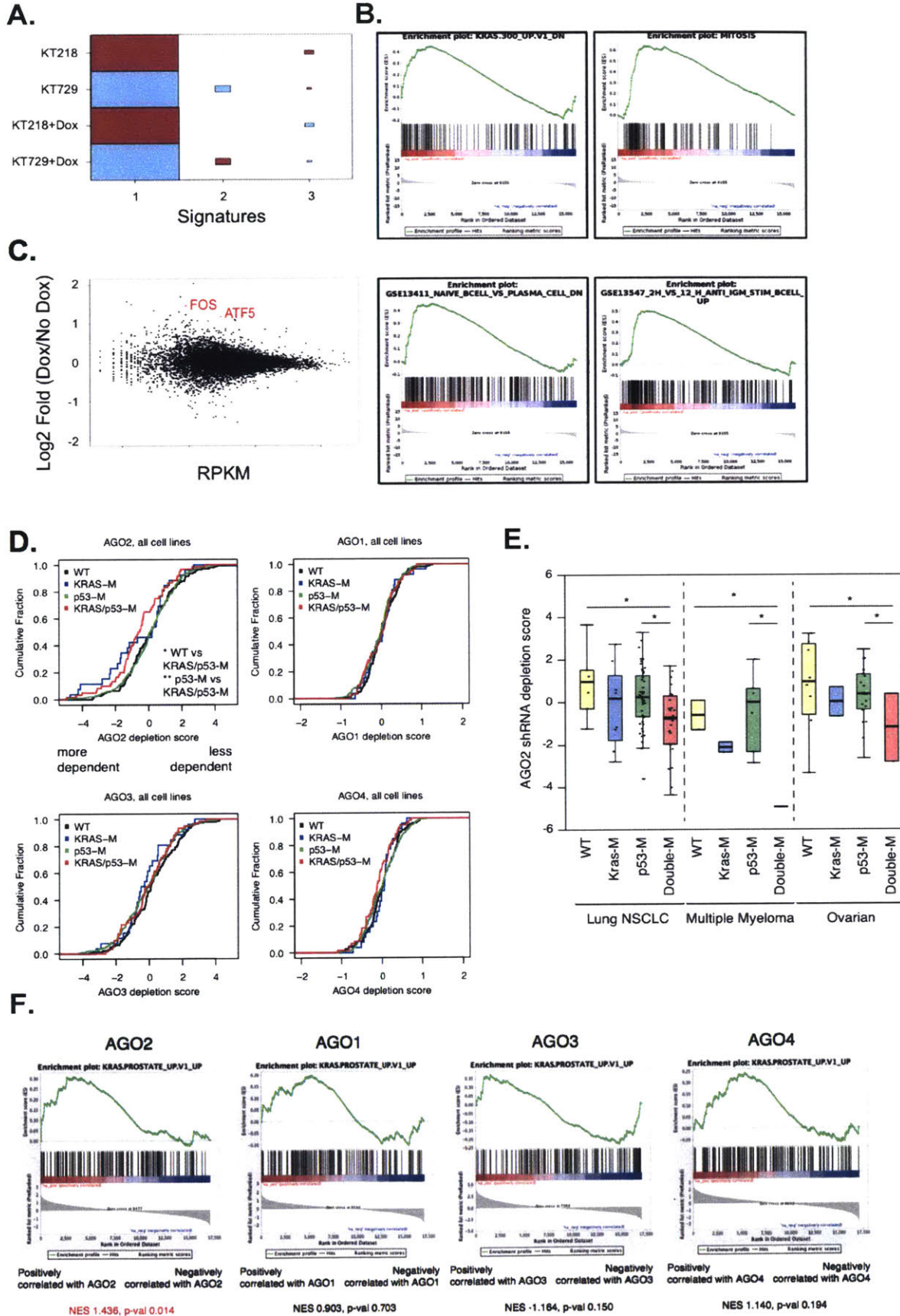
As the NF-κB signaling pathway has been found to be stimulated through activated KRAS signaling, we suspected that overexpressing AGO2 might be amplifying signal transduction pathways downstream of KRAS to activate specific NF-κB subunits (Bang et al., 2013). Expression of the KRAS-binding deficient AGO2 mutants, K113A and E115A, in the absence of doxycycline resulted in reduction in p100 processing into the active p52 NF-κB subunit (Figure 6C). In agreement with this, activity of both the p52 and Rel-B subunits were decreased with overexpression of the K113A and E115A AGO2 mutant proteins. By contrast, constitutive expression of both the wild type AGO2 and D670A AGO2 mutant proteins in the absence of doxycycline maintained relatively high levels of activated p52 and Rel-B subunits (Figure 6D). This suggests activation of NF-κB in this system is predominantly regulated by the KRAS/AGO2 complex.

AGO2 upregulation in KRAS-driven cancer is generally advantageous to maintain pro-survival signals

To determine the impact of AGO2 overexpression on the transcriptome of KPR tumor cells, we performed RNA-seq on AGO2 overexpression-dependent cell lines in the presence and absence of doxycycline. RNA-seq analysis revealed a few distinct gene expression profiles determined by independent component analysis (ICA), including a strong transcriptional signature that separates biological replicates independent of doxycycline-induced AGO2 overexpression, referred to as "signature 1" (Figure 7A). This suggests a high degree of heterogeneity between KPR cell lines. Nevertheless, we focused on "signature

3," which represented a set of genes that were upregulated in doxycycline-treated KPR cells compared to controls. This signature contained gene sets associated with mitosis, oncogenic KRAS-driven cancers, and B cell differentiation (Figure 7B). Furthermore, we detected upregulation of genes in KPR cells ectopically expressing AGO2 that are often misregulated in human cancers, such as FOS and ATF5. These transcription factors are typically activated through the PI3K and MAPK signaling pathways to promote cell proliferation and survival (Figure 7C), suggesting that their enhanced activity is facilitated by the KRAS-AGO2 interaction (Shaulian and Karin, 2001; Greene et al., 2009).

The dependency on AGO2 to promote KRAS-mediated cellular transformation was further supported from a LOF screen performed on 501 human cancer cell lines using an shRNA library (Tsherniak et al., 2017). This study demonstrated that cancer cells harboring both *KRAS* and *TP53* mutations were generally hypersensitive to depletion of AGO2, but not AGO1, AGO3, or AGO4 (Figure 7D). Moreover, this dependency was evident by grouping cell lines used in the study by cancer type, where *KRAS* mutant cancer cell lines showed slight sensitivity to AGO2 depletion but cell viability was significantly impaired in the *KRAS/TP53* double mutant cancer cells (Figure 7E). Consistent with our findings from the KPR mouse model, genes that were co-dependent on AGO2 expression for survival were significantly associated with gene sets upregulated in *KRAS*-mutant cancer cells (Figure 7F). These activated KRAS gene sets were not statistically significant when compared to AGO1, AGO3, or AGO4

A.

Signatures

B.

Enrichment plot: KRAS.300_UP.V1_DN

Enrichment plot: MITOSIS

C.

Log2 Fold (Dox/No Dox)

RPKM

FOS ATF5

Enrichment plot:
GSE13411_NAIVE_BCELL_VS_PLASMA_CELL_DN

Enrichment plot:
GSE13547_2H_VS_12_H_ANTI_IGM_STIM_BCELL_UP

D.

AGO2, all cell lines

AGO1, all cell lines

* WT vs
KRAS/p53-M
** p53-M vs
KRAS/p53-M

more   less
dependent   dependent

AGO3, all cell lines

AGO4, all cell lines

E.

Lung NSCLC   Multiple Myeloma   Ovarian

F.

AGO2

Enrichment plot: KRAS.PROSTATE_UP.V1_UP

AGO1

Enrichment plot: KRAS.PROSTATE_UP.V1_UP

AGO3

Enrichment plot: KRAS.PROSTATE_UP.V1_UP

AGO4

Enrichment plot: KRAS.PROSTATE_UP.V1_UP

Positively   Negatively
correlated with AGO2   correlated with AGO2

Positively   Negatively
correlated with AGO1   correlated with AGO1

Positively   Negatively
correlated with AGO3   correlated with AGO3

Positively   Negatively
correlated with AGO4   correlated with AGO4

NES 1.436, p-val 0.014

NES 0.903, p-val 0.703

NES -1.164, p-val 0.150

NES 1.140, p-val 0.194

132

Figure 7. AGO2 dependency is associated with *KRAS/TP53* mutant cancer cells. (A) RNA-seq analysis of KPR cell lines (KT223 and K591) in the presence and absence of doxycycline revealed distinct gene expression profiles using independent principle component analysis. (B) Signature 3 was associated with gene sets involved in oncogenic Kras signaling, mitosis and B cell differentiation via GSEA analysis. (C) MA plot revealed genes that were differentially expressed based off of log2-fold changes. ATF5 and FOS were significantly upregulated in AGO2-overexpressed conditions. (D) AGO2 depletion score was calculated from a LOF screen on 501 cell lines, categorized by *KRAS* and *TP53* mutation status. (E) GSEA analysis comparing genes found to be dependent on individual AGO proteins from LOF screen to a gene set associated with upregulated genes in mutant *KRAS* pancreatic cancer cells. (F) Cancer cell lines of the NSCLC, multiple myeloma and ovarian cancer types were separated by their *KRAS* and *TP53* mutation status. The AGO2 depletion score was calculated for each.

dependency.

Together, these results suggest that overexpressing AGO2 in B-lymphocytes, in the context of oncogenic *Kras* mutation and *Tp53* deletion, causes arrest in the plasmablastic stage of B cell development and enhances cellular transformation. The interaction between AGO2 and oncogenic KRAS appear to facilitate activation of downstream pathways, including NF-κB signaling, that is instrumental in promoting cell proliferation, differentiation, and survival (Figure 8). However, the plasticity of these cellular processes was demonstrated upon withdrawal of ectopic AGO2, whereby cell death signals were induced in KPR tumor cells.

Discussion

AGO2 is frequently amplified and overexpressed in various human cancers, but its role in tumor progression has been incompletely characterized. Our study demonstrates the oncogenic effect of AGO2 overexpression in the initiation and progression of a B cell malignancy in mice that phenotypically resembles human ABC DLBCL. We confirmed that AGO2 overexpression in B lymphocytes resulted in cellular transformation by developing a transplantation model to induce oncogenic KRAS activity, *Tp53* deletion and ectopic AGO2 expression in bone marrow-derived B cells. Histology and FACS analysis of lymphoid aggregates derived from recipient mice revealed extensive plasmacytic differentiation, consisting largely of abnormal plasmablasts, associated with
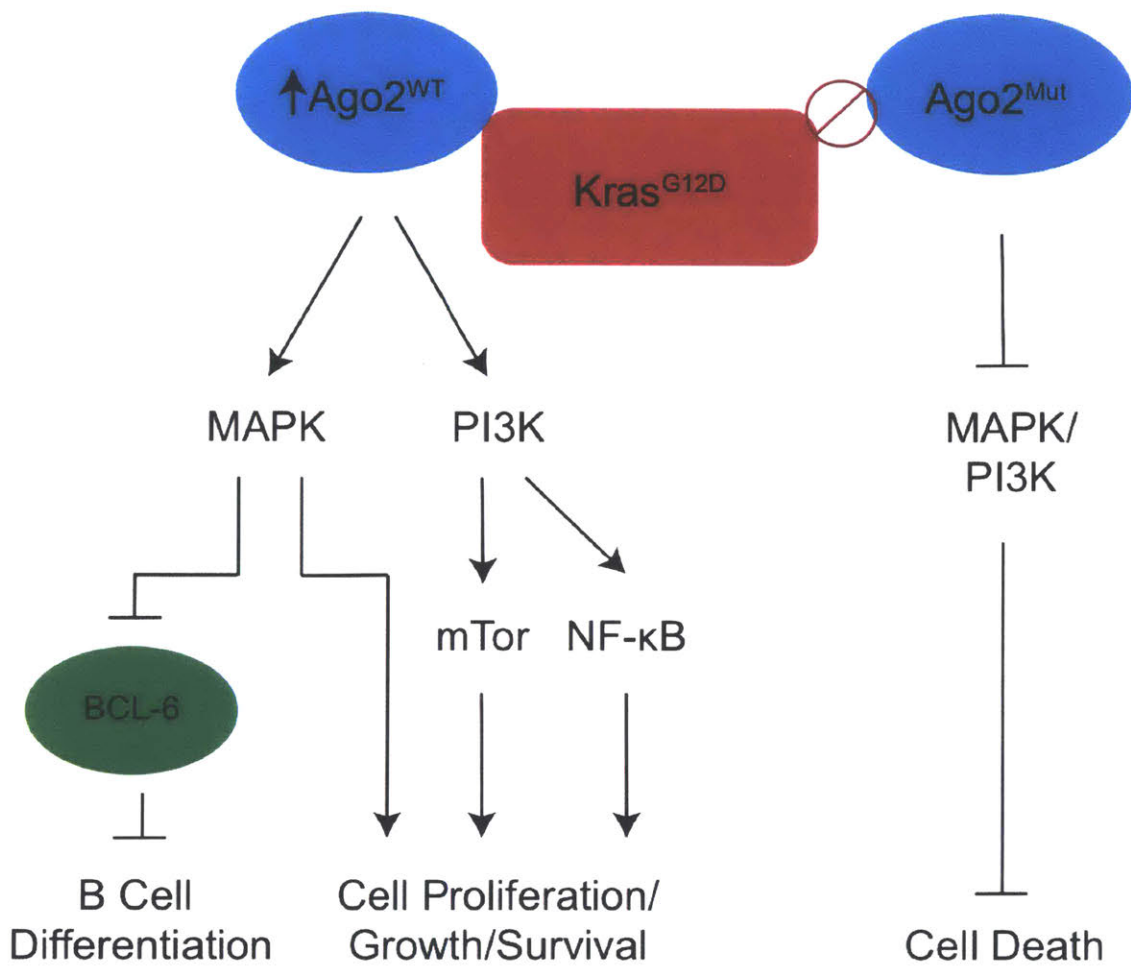
Figure 8. Model of mechanism of cellular transformation in p53-deficient B-lymphocytes with Ago2 overexpression and oncogenic Kras activation.

AGO2 overexpression. The ABC-like DLBCL phenotype was dependent on the physical interaction between KRAS and AGO2.

This pathological condition was initially detected using the KP autochthonous mouse model, where virus was administered via intratracheal inhalation to infect lung epithelial cell. Thus, the initiation of lymphomagenesis was unexpected. We suspect that escape of lentiviral particles into local draining lymph nodes resulted in transformed B lymphocytes *in vivo.*

We have confirmed that cellular transformation in our model is dependent on the physical interaction between AGO2 and KRAS, suggesting a mechanism exploited by cancer cells to efficiently enhance oncogenic signaling pathways and dysregulate miRNA activity. This notion is further supported by the defect in miRNA processing resulting from the KRAS-AGO2 complex, which was characterized in prior studies (Shen et al., 2013). Previous reports have also demonstrated that the N-terminal domain of AGO2, critical for KRAS binding, plays an important role in RNA duplex unwinding (Kwak and Tomari, 2012). Impediment of this domain through KRAS binding has been implicated in perturbation of miRNA processing and is associated with enhanced cell proliferation (Shankar et al., 2016).

Moreover, the KRAS-AGO2 interaction stimulated activity of signaling pathways frequently perturbed in human ABC DLBCL, including the PI3K and NF-κB signaling cascades. Loss of AGO2 overexpression resulted in a decrease in activity of these pathways and, consequently, initiation of cell death primarily through apoptotic cues. Consistent with this observation, amino acid residue Y64

of KRAS, the direct binding site of AGO2, was previously identified to bind to other effector molecules, such as PI3K, to promote hematopoietic malignancies (Shieh et al., 2013; Shankar et al., 2016). Missense mutations at this tyrosine residue resulted in a significant decrease in cellular transformation. Together, this implicates AGO2 in being an important KRAS regulator that broadly enhances oncogenic signaling pathways.

Although human ABC DLBCL is known to be heterogeneous in terms of mode of NF-κB activation, most are largely reliant on constitutive B cell receptor (BCR) signaling (Wilson et al., 2015). To this end, it is not surprising that Ibrutinib, a small molecule that targets Bruton's tyrosine kinase, is clinically effective in treating ABC DLBCL. Here, we describe a distinct mechanism where activation of the non-canonical NF-κB signaling pathway is mediated by interaction between KRAS and AGO2. Though the precise mechanism of NF-κB activation is unclear in this model, oncogenic KRAS has been documented to activate both canonical and non-canonical NF-κB via GSK-3α stabilization in other cancer models (Bang et al., 2013). Though activating *KRAS* mutations are only detected in approximately 5% of all human lymphoid malignancies, it is possible that oncogenic KRAS activity is advantageous in a specific subset of lymphomas, including ABC DLBCL (Pylayeva-Gupta et al., 2011).

KRAS was recently shown to be critical for B cell lymphopoiesis (Chen et al., 2016). Hematopoietic-specific deletion of *Kras* impaired early B cell development and late B cell maturation. This specific KRAS requirement for development of B- but not T-cells may partly explain the occurrence of an

aggressive B cell lymphoma in this model system. Furthermore, KRAS is essential for activation of the Raf-1/MEK/ERK pathway induced by BCR in mature B cells. Thus, while *KRAS* mutation is infrequent in newly diagnosed DLBCL patients, the AGO2-KRAS axis may also mimic chronic activation of BCR signaling by promoting constitutive NF-κB activation, thereby leading to a phenotype resembling ABC DLBCL (Pylayeva-Gupta et al., 2011; Reddy et al., 2017). Additionally, given the frequent AGO2 amplification and KRAS mutation detected in multiple myeloma, the AGO2-KRAS pathway may have broad pathogenic impacts in B cell malignancies with post-germinal center phenotypes.

We also report an increase in MAPK signaling as a function of AGO2 overexpression. Despite a low frequency of perturbation of this specific pathway in many human lymphomas, MAPK signaling has been reported to promote B cell differentiation via BCL6-targeted degradation (Niu et al., 1998). As *Bcl6* is downregulated in the KPR lymphomas described here, AGO2-enhanced MAPK signaling could partially be responsible for the preference for an ABC-like lymphoma.

BLIMP1, the master regulator of plasma cell differentiation, is frequently inactivated through homozygous deletion and LOF mutations in human ABC DLBCL patients (Mandelbaum et al., 2010). It is mechanistically unclear how expression of BLIMP1 is inhibited in the KPR lymphomas. As AGO2 is overexpressed in the KPR model, the possibility of miRNA-mediated repression of BLIMP1 is intriguing.

Finally, our data demonstrates that oncogenic KRAS and AGO2 overexpression promotes an ABC-like lymphoma in the context of p53 deficiency. p53 LOF mutations occur in a fraction of human DLBCL cases, many of which are found from relapsed patients (Ichikawa et al., 1997). To this end, it would not be surprising that p53 inactivation is a necessary prerequisite in B cell transformation involving KRAS activation and AGO2 overexpression. In addition, the sensitivity of human cancer cell lines carrying *KRAS* and *TP53* mutations to shRNA-mediated inhibition of AGO2 also supports a functional role of p53 mutation in promoting dependency on AGO2 expression.

ABC lymphoma is an aggressive subtype of human DLBCL, where neoplastic plasmablasts are prevented from terminally differentiating. There are well-characterized differences in the gene expression profiles between germinal center B cell like DLBCL and ABC DLBCL. Consequently, conventional therapeutic strategies for GCB DLBCL have remained largely ineffective to treat ABC DLBCL patients. The mouse model of an aggressive lymphoma with plasmablastic features described here could be used to gain insight into oncogenic mechanisms in this disease state, which could be useful in designing new treatments for human ABC DLBCL.

Experimental Procedures

Cell Culture

Mouse cell lines were generated from lymphocytes derived from both splenocytes and bone marrow. For primary cell cultures, we initially harvested and expanded stromal cells to plate primary lymphocytes on top of. We harvested bone marrow from healthy mice into Roswell Park Memorial Institute (RPMI) media supplemented with penicillin/streptomycin (pen/strep). The cells were filtered through a 70μm cell strainer (Fisher Scientific®, #08-771-2) and pelleted at 1500 R.P.M for 5 m. After media removal, we lysed red blood cells (RBCs) in 5mL ACK lysis buffer (Thermo Scientific™, #A1049201) for 1.5 m and then inactivated it with 45mL sterile PBS. These cells were then plated on 10cm cell culture-treated plates in RPMI media supplemented with 10% fetal bovine serum (FBS), 2mM L-Glutamine, and pen/strep. Media was replaced every day to remove non-adherent cells and monitored for expansion of adherent stromal cells. Once stromal cells reached confluency, they were split once onto 6-well tissue culture treated plates. Primary B220-expressing lymphocytes of interest, derived from bone marrow, were sorted for and plated on top of stromal feeder cells in B cell media (Dulbecco's Modified Eagle Medium (DMEM), Iscove's Modified Dulbecco's Medium (IMDM), 50μM β-mercaptoethanol (BME), 10% FBS, 2mM L-Glutamine, pen/strep, 10ng/mL recombinant murine SCF, 10ng/mL recombinant murine IL-7). These lymphocytes were transduced *in vitro* via spinfection in the presence of 8μg/mL polybrene (EMD Millipore, #TR-1003-G) and 1mL retrovirus at 1500g for 1 hour at room temperature. Cells were transferred to a new plate the following day, as polybrene is toxic to lymphocytes, and allowed to expand in culture in the presence of SCF and IL-7. For tumor cells from transplant model, splenocytes were harvested into RPMI media supplemented with pen/strep. The cells were homogenized and filtered through a 70μm cell strainer and pelleted at 1500 R.P.M for 5 m. After media removal, we lysed RBCs and cells were resuspended in B cell media and subsequently plated onto tissue culture treated plates without stromal cells. All cells were grown at 37°C with 5% CO_2 in the presence of doxycycline (Sigma, #D9891). All experiments using cells in doxycycline-withdrawal conditions were collected at 48 h unless noted.

Lentiviral and Retroviral Constructs

GMAP-compatible pLL3 lentiviral and MSCV retroviral backbones were previously engineered to generate viral constructs via Gibson Assembly® with the following construction: backbone-pA-gA-pB-gB. TRE and pGK promoters were amplified from DNA templates with "promoter A" (pA) Gibson homology arms (Clontech, #R040A), purified from 1% agarose gels and gel-extracted (Macherey-Nagel, #740609.50). FLAG-HA and cMyc tags were independently engineered onto the N-terminus of mAGO2 via PCR from cDNA templates derived from mouse embryonic stem cells. PCR primers were designed with "gene A" (gA) Gibson homology arms. PCR products were purified from 0.7% agarose gels and gel-extracted. pGK and EFS were amplified from DNA

templates with "promoter B" (pB) Gibson homology arms, purified from 1% agarose gels and gel-extracted. Cre recombinase and Venus were amplified with "gene B" (gB) Gibson homology arms, purified from 0.8% and 1% agarose gels, respectively, and gel-extracted. Gibson Assembly was performed by first preparing 5X isothermal assembly reaction buffer (0.5M Tris-HCl pH7.5, 50mM $MgCl_2$, 1mM dNTPs, 50mM DTT, 1.5g PEG-8000, 20mg NAD, and bring up to 6mL with nuclease-free water). Isothermal master mix was then prepared by combining 320µL 5X isothermal assembly reaction buffer, 1.2µL T5 exonuclease (NEB®, #M0363L), 20µL Phusion polymerase (NEB®, #M0530L), 160µL Taq ligase (NEB®, #M0208L), and 700µL nuclease-free water. 4µL inserts (5.7 x 10^{-2} pmol each of pA,gA,pB,gB) and 1µL of 50ng digested-backbone (BsrGI/PmeI-digested lentiviral backbone or PmeI-digested retroviral backbone) were added to 15µL isothermal master mix and incubated at 50°C for 1 hour. These reactions were then transformed into competent bacteria (Thermo Scientific^{TM}, #C737303), grown at 30°C for 20 hours, and screened with Sanger sequencing. All primers used are listed in primer table.

pSLIK-hygro (Addgene, #25737) lentiviral backbone was used to clone in FLAG-HA-mAgo2, which was PCR amplified with 5' EcoRI and 3' NotI restriction sites. PCR products were purified and extracted with the above protocol. All primers used are listed in primers table in supplemental materials.

Migration and invasion assay

To assay for cell migration and invasion, KP LUAD cells were diluted in serum-free media (SFM), either containing or lacking doxycycline, at a concentration of 250 cells/µL. Matrigel® (Corning, #354230) was diluted to a concentration of 800 µg/mL and 50µL was added to wells in the upper chamber intended for cell invasion measurements (ACEA Biosciences Inc, #05665817001). 30µL of Matrigel® was removed from each well and the upper chamber was incubated at 37°C for 4 h. 160µL of media was added to each well of the bottom chamber of the transwell plate. SFM was added to wells in the bottom chamber as negative controls for cell migration/invasion. The upper chamber was attached to the bottom chamber. 25µL of SFM was added to each well of the upper chamber and incubated at 37°C for 1 h to equilibrate plate. The background reading was measured on the xCELLigence RTCA DP (ACEA Biosciences Inc, #00380601050). 100µL of diluted cells were added to each well and incubated at room temperature for 30 minutes. Plates were mounted on xCELLigence RTCP DP at 37°C with 5% CO_2 and migration/invasion analysis was started, where measurements were taken every 15 m for 24 h.

Site-directed Mutagenesis

To generate Ago2 mutants, we designed primers that contained point mutations that would introduce missense mutations of interest. We used PCR to amplify the mutant construct from our lentiviral plasmid containing the WT AGO2 sequence by incorporating 1ng WT AGO2 template, 200nM mutagenic primers (F/R), 5µL 10X Pfu Buffer, 1µL Turbo PfuTurbo DNA Polymerase (Agilent Technologies,

141

#600410), 200nM dNTPs, and nuclease-free water to bring reaction up to 50µL. PCR products were digested with 1µL DpnI (NEB®, #R0176S) for 1 hour to remove WT AGO2 templates and digested templates were transformed into competent bacterial cells. Transformed cells were selected for ampicillin resistance and mutants were screened with Sanger sequencing.

Virus Production and Titering

To produce lentiviruses, we passaged and split 293T cells onto 15cm tissue culture-treated plates until they reached 70% confluency. We co-transfected these 293T with 10µg lentiviral constructs, 7.5µg delta8.2 and 2.5µg VSV-G packaging plasmids, and 75µL TransIT-LT1 (Mirus Bio, #MIR2300). Media was changed on the following day with 15mL fresh media (DMEM, 10% FBS, 2mM L-Glutamine). Supernatant was collected 48 and 72 hours post-transfection and run through a 0.45µm syringe filter (Pall Corporation, #4184). Virus was concentrated at 25,000 r.p.m. for 1.5 h at 4°C and resuspended in 200µL Gibco™ Opti-MEM™ (Thermo Scientific™, #31985070). Virus was titered by infecting Green-Go cells with serial dilutions of virus. Green-go cells were generated by transducing 3TZ cells with a bicistronic retrovirus containing an LTR promoter-driven inverted GFP sequence, flanked by two sets of incompatible loxP sites. Viral titers were calculated based off of Cre-mediated GFP-expression. To produce retroviruses, we transfected helper-free Phoenix cells plated on 10cm tissue culture-treated plates with 10µg retroviral constructs and 20µL TransIT-LT1 (Mirus Bio, #MIR2300). Media was changed on the following day with 8mL fresh media (DMEM, 10% FBS, 2mM L-Glutamine). Supernatant was collected 48 and 72 hours post-transfection and run through 0.45µm syringe filter.

Histology and IHC

Tissues were fixed in zinc formalin (Polysciences, Inc., #21516) overnight at room temperature. They were then transferred to 70% ethanol prior to paraffin embedding. Sections were stained with Hematoxylin and Eosin (H&E) on the Varistain Gemini automated slide stainer (Thermo Shandon). IHC was performed with 5µm-thick sections. Tissues were blocked with Dako Dual Endogenous Enzyme-Blocking Reagent (Agilent Technologies, #S200389-2) and stained using the ImmPRESS™ HRP Polymer Reagents and DAB staining kit (Vector Laboratories, #SK-4100). All antibodies with used are listed in antibodies table in supplemental materials.

Mouse Ig Isotyping

All cell lines were cultured overnight before collecting supernatant for immunoglobulin isotyping. 10µL of media was collected for each isotyping reaction. Analysis of relative abundance of each Ig isotype was determined using the Ig Isotyping Mouse Instant ELISA™ Kit (Thermo Scientific™, #88-50660-22).

Western Blotting

Cells were lysed with RIPA supplemented with protease inhibitor cocktail (Roche, #04693159001), Benzonase® nuclease (Sigma, #E8263) and Halt phosphatase inhibitor cocktail (Thermo Scientific™, #78420). Protein lysates were quantified using Pierce™ BCA™ Protein Assay (Thermo Scientific™, #23225). Samples and Kaleidoscope ladder (Bio-Rad, #161-0375) were prepared in NuPAGE® loading buffer (Thermo Scientific™, #NP0007) and reducing agent (Thermo Scientific™, #NP0004) and boiled for 5 minutes at 95°C followed by 2-minute incubation on ice. Western blotting was performed using the Xcell II® Blot Module (Thermo Scientific™). Proteins were run on NuPAGE™ 4-12% Bis-Tris protein gels (Thermo Scientific™, #NP0321BOX) and then transferred to either PVDF membrane for film exposure (Merck Millipore Ltd., IPVH00010) or nitrocellulose membrane for LI-COR fluorescent imaging (LI-COR® Biosciences, #926-31092). Membranes were blocked in 5% milk for PVDF or Odyssey® Blocking Buffer (LI-COR®, #927-40000) for nitrocellulose and incubated with primary antibodies overnight at 4°C with rocking. Membranes were washed 3x with PBST, 5 minutes each, and incubated in secondary antibody for 1 hour at room temperature with rocking. Membranes were washed 4x with PBST and exposed/developed.

Southern Blotting
gDNA was extracted from cell pellets that were resuspended in lysis buffer (100mM Tris pH8.5, 5mM EDTA, 0.2% SDS, 200mM NaCl, 100μg/mL Proteinase K). Pellets were incubated overnight at 37°C with gentle agitation. The following day, we added 1 volume of isopropanol and incubated the tubes on a shaker at room temperature for 4 hours. The gDNA webs were transferred to a new tube with 100uL TE (10mM Tris pH7.5, 0.1mM EDTA) and gently agitated overnight at 55°C. gDNA was quantified with the Thermo Scientific™ NanoDrop the following day and 10μg of gDNA was digested with EcoRI for 2 hours (NEB, #3101S). Digested samples, and ladder (NEB®, #N3014S), were run on a 0.8% agarose gel (Thermo Scientific™, #16500500). The gel was run at 25 volts overnight at 4°C. Gel was imaged on Kodak Gel Logic Imaging System to ensure efficient digestion and separation of gDNA. The gel was then incubated in 2 volumes of 0.25N HCL for 15 minutes with gentle shaking, rinsed 2x with distilled water, and transferred to 2 volumes of 0.4N NaOH for 15 minutes with gentle shaking. The gel was then transferred to an assembled transfer apparatus consisting on pre-wet Whatman filter paper and Hybond XL membrane (GE Life Sciences, #RPN2020S) in 0.4N NaOH. gDNA was transferred overnight at room temperature and membranes were prehybridized in ULTRAhyb®-Oligo Buffer (Thermo Scientific™, #AM8663) at 42°C with rotation for 30 minutes. J$_H$4 probe was labeled with... Radiolabeled probes were added to prehybridized membrane overnight at 42°C. The membrane was washed 3x with wash buffer (2X SSC, 0.5% SDS) at 42°C with rotation. Membrane was exposed overnight on a phosphor screen (GE Healthcare/Molecular Dynamics, #2670) and scanned on the GE Typhoon TRIO imager.

Co-Immunoprecipitation

Cells were lysed with modified RIPA (10mM Tris pH 7.4, 150mM NaCl, 1% Triton-X, 0.1% SDS, 1mM EDTA) supplemented with protease inhibitor cocktail, Benzonase® nuclease and phosphatase inhibitor cocktail (Thermo Scientific™, #78420). Protein lysates were quantified using Pierce™ BCA™ Protein Assay. 20µL Protein G Dynabeads™ (Thermo Scientific™, #10003D) was prepared per sample at 4°C. First, beads were washed 3x with 0.5% BSA and primary antibodies were conjugated to beads overnight in 0.5% BSA at 4°C. We washed antibody-conjugated beads 3x with 0.5% BSA and 1x with modified RIPA. Beads were resuspended in 50µL modified RIPA for every 10µL beads used. We added 500µg protein lysate to antibody-conjugated beads and it was incubated overnight at 4°C with rotation. 10% of the lysate was aliquotted and frozen for input. We removed 10% supernatant for flowthrough the following day and washed beads 4x with modified RIPA. We resuspended beads in NuPAGE® loading buffer and reducing agent. Input and flowthrough samples were also prepared in NuPAGE® loading buffer and reducing agent. Samples and ladder were boiled for 5 minutes at 95°C followed by 2-minute incubation on ice. Western blotting procedure is described above.

qRT-PCR
Total RNA from cells were TRIzol-extracted (Thermo Scientific™, #15596026). For mRNA qRT-PCR, total RNA was converted to cDNA using M-MLV (Thermo Scientific™, #28025013) and Oligo d(T) (Gene Link™, #26-4000-05). In each qRT-PCR reaction, we added 10ng cDNA, 200nM forward and reverse primers, and KAPA SYBR FAST (KAPA Biosystems, #KK4611). These samples were amplified and analyzed on the Roche LightCycler® 480 instrument. Gene expression was normalized to β-Actin. For miRNA qRT-PCR, we prepared miRNA templates from 400ng total RNA using miR-X™ miRNA First Strand Synthesis Kit (TaKaRa, #638313). In each 10µL qRT-PCR reaction, we added 4µL template, 200nM miRNA-specific primer, 0.2µL mRQ 3' primer, and 5µL SYBR Advantage Premix (Clontech, #638314). These samples were amplified and analyzed on the Roche LightCycler® 480 instrument. miRNA expression was normalized to U6. U6 primer set is supplied in Clontech SYBR kit.

RNA-Seq
Total RNA from cells was extracted with Qiagen RNeasy Plus Micro kit (Qiagen, #74034). RNA concentration and quality was assessed with Agilent BioAnalyzer. 50ng of total RNA was used in each library preparation reaction, using KAPA RNA Hyperprep Kit with RiboErase (KAPA Biosystems, #KK8560). Each sample was prepared with a unique barcode (KAPA, #KK8712). Library concentration and quality was assessed with Agilent BioAnalyzer, pooled and loaded onto Illumina HiSeq 2000 for single-end sequencing.

Flow Cytometry
To stain for cell surface markers, cells were isolated and washed with 1X sterile PBS. Cells were resuspended in FACS Buffer (10% FBS in PBS) and fluorophore-conjugated primary antibodies were added and incubated on ice for

30 minutes. The cells were spun down at 350g for 5 minutes and resuspended in appropriate volume of FACS buffer to analyze on BD LSR II Flow Cytometer.
To stain for intracellular proteins, we used the BD Transcription Factor Buffer Set (BD Biosciences, #562725). We initially stained the cells with 1μL fixable viability dye (BD Biosciences, #562247) in 1mL PBS for 10 minutes at room temperature. We then washed the cells with FACS buffer and resuspended the cells in 1mL 1X Fix/Perm buffer. The samples were vortexed for 3 s and incubated at 4°C in the dark for 40 m. 1mL 1X perm/wash buffer was directly added to the samples and centrifuged for 6 m at 350g. The supernatant was discarded and cells were washed with 2mL perm/wash buffer. Primary antibody was added to each sample in 100μL wash/perm buffer and incubated at 4°C for 40 m. The samples were vortexed briefly and washed with 2mL wash/perm buffer. The cells were resuspended in 100μL wash/perm buffer with fluorophore-conjugated secondary antibody. Samples were incubated at 4°C for 40 m in the dark. The cells were washed with 2mL wash/perm buffer and the cells were resuspended in 350μL FACS buffer. Samples were analyzed on the BD LSR II Flow Cytometer.

Cell viability
Dead cells were detected with propidium iodide (Thermo Scientific^{TM}, #P1304MP) and trypan blue (Thermo Scientific^{TM}, #15250061) staining.

Apoptosis Assay
Cells were washed 2x with cold PBS and resuspended in 1X binding buffer (BD Biosciences, #556547). 5μL of Annexin V (FITC) antibody and 5μL of PI were added to each sample and incubated for 15 minutes at RT in the dark. 400μL of 1X binding buffer was added to each tube and cells were analyzed by flow cytometry.

NF-κB Activity Assay
Whole cell extract from cell lines were extracted using Complete Lysis Buffer from TransAM® NF-κB family kit (Active Motif, #43296). Concentration of whole cell extracts were determined using Pierce^{TM} BCA^{TM} Protein Assay. Analysis of active p50, p52, p65 and RelB subunits were assessed with the TransAM® NF-κB family kit (Active Motif, #43296).

In Vivo **luciferase imaging**
Mice were given 100μL D-Luciferin (PerkinElmer, #122799) via intraperitoneal injection. They were then subjected to mild anesthesia for 10 m with isoflurane inhalation (Southern Anesthesia Surgical Inc., #001725CS) before bioluminescence imaging on the PerkinElmer IVIS imaging System. Luciferase signals were analyzed with Living Image Software.

Retro-orbital bleeding
Mice were subjected to mild anesthesia via isoflurane inhalation (Southern Anesthesia Surgical Inc., #001725CS). Retro-orbital bleeding was performed using heparinized blood collecting tubes (Fisher Scientific®, #02-668-10).

Collected blood (~150µL per mouse) was placed into tubes containing 100µL 3.2% sodium citrate to prevent coagulation. We then added 500µL RPMI media and underlayed the samples with 500µL Histopaque (Sigma Aldrich, #10771-100mL). Samples were centrifuged for 20 minutes at 670g to pellet red blood cells. The mononuclear cells were harvested at the interface and tdTomato+ cells were detected with EMD Millipore Guava® easyCyte Flow Cytometer.

Tumor models
All animal studies described in this study were approved by the MIT institutional Animal Care and Use Committee. All animals were maintained on a mixed C57BL/6J x 129SvJ genetic background. KP mice were previously developed by engineering transcriptional and translational stop elements, flanked by loxP sites, into the first intron of the endogenous K-ras gene. Furthermore, these mice contained an oncogenic mutation in K-ras, where a glycine to aspartic acid mutation in present at codon 12, and conditional "floxed" p53 alleles between exons two through ten to constitutively activate K-ras and mimic loss of p53 function upon Cre-mediate recombination, respectively. These KP mice were then crossed to the Ai9 Cre reporter mice harboring loxP-flanked STOP cassette preventing transcription of a CAG-promoter driven tdTomato fluorophore at the ROSA26 locus. These KP tdTomato reporter mice were then crossed to Cre reporter mice generated from Inder Verma's laboratory harboring loxP-flanked STOP cassette preventing transcription of luciferase at the ROSA26 locus. These KP(tdTom/Luc) reporter mice were then crossed to CAGs-rtTA3 mice generated from Scott Lowe's laboratory to enable a conditional expression system in the presence of doxycycline.

The generated KP reporter mice used for our studies either harbored or lacked the rtTA3 transgene, allowing us to control for doxycycline-induced effects. In the autochthonous setting, we delivered 8.0×10^3 infectious particles of our lentiviral construct via intratracheal inhalation to each mouse that was briefly anesthetized with isoflorane. These mice were placed on doxycycline food to induce Ago2 overexpression from *in vivo* transduced cells. In the transplant model, we isolated B220-positive lymphocytes from bone marrow of the KP reporter mice harboring the rtTA transgene and transduced them *in vitro* with our retroviral construct as described above. The tdTomato-positive cells were selected for using the BD FACSAria™ III sorter. We sublethally irradiated F1 mice (C57BL/6J x 129SvJ) and transferred 400,000 tdTomato-positive cells to each mouse 16 h post-irradiation via tail-vein injection. Mice were split into +/- doxycycline conditions. For intravenous injection of KP LUAD into immunocompromised mice (The Jackson Laboratory, #007850), cells were diluted using cold PBS to a concentration of 1,000 cells/µL. 100µL of cells were injected into the tail vein of nude mice and sacrificed 4 wk post-transplantation.
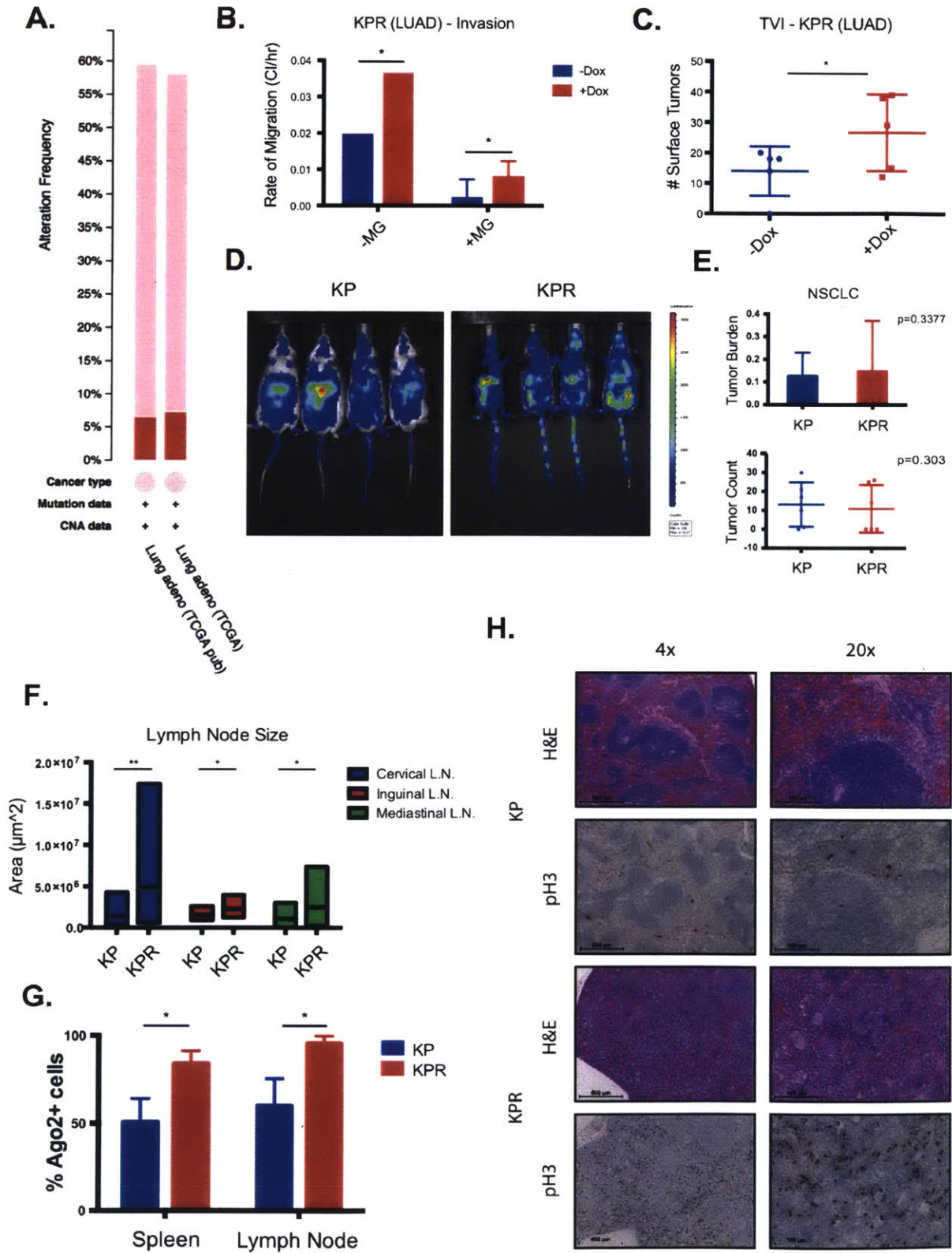
Statistical Analysis
Student's *t*-test was used for all measurements of tumor burden and IHC quantifications. All error bars denote s.e.m.

146
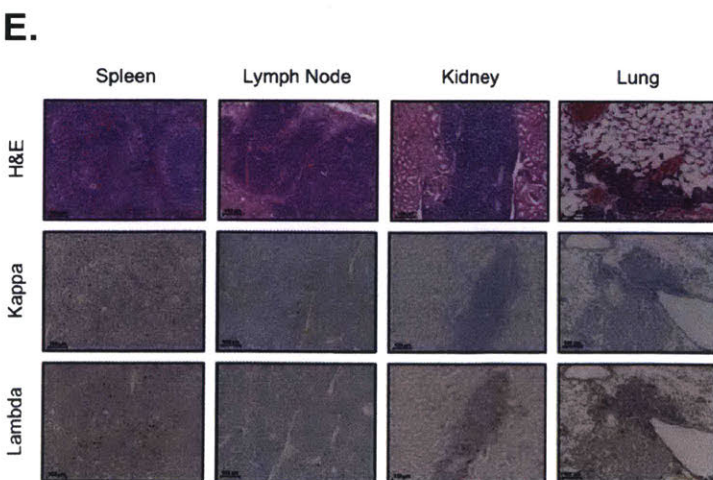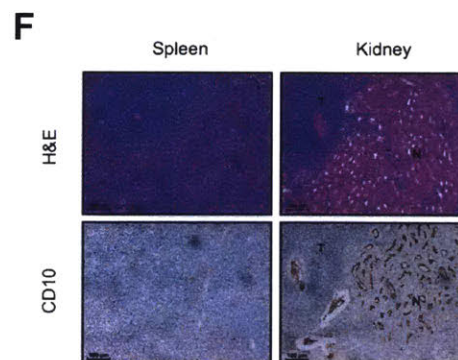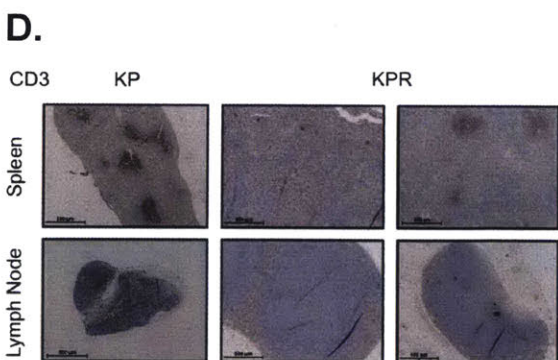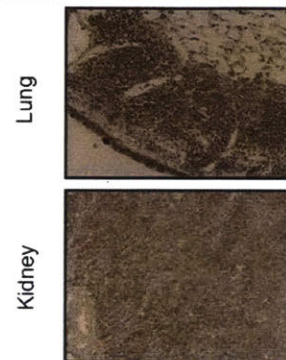
Acknowledgements

Supplemental Figures and Tables

A.

B.

KPR (LUAD) - Invasion

C.

TVI - KPR (LUAD)

D.

E.

NSCLC

F.

Lymph Node Size

G.

H.
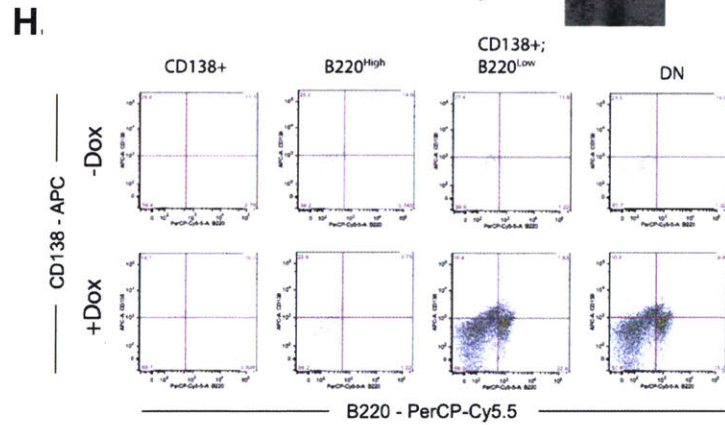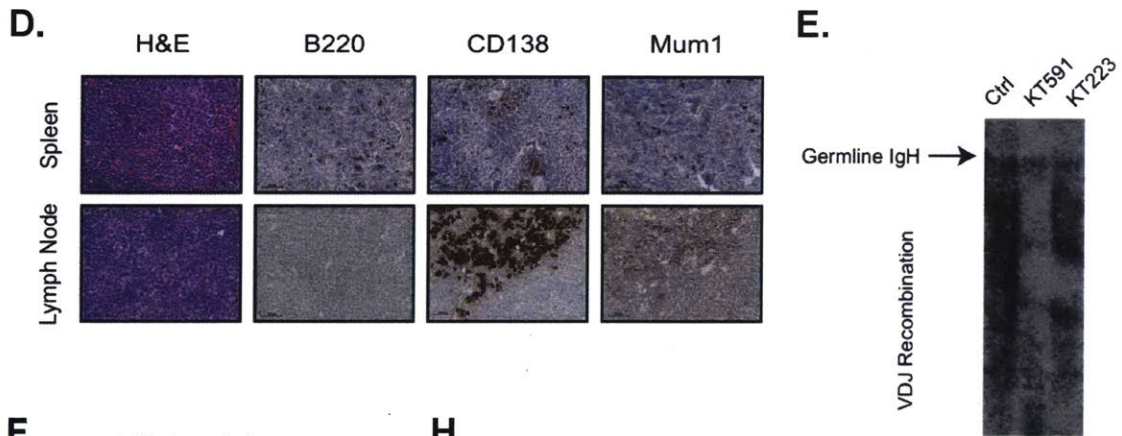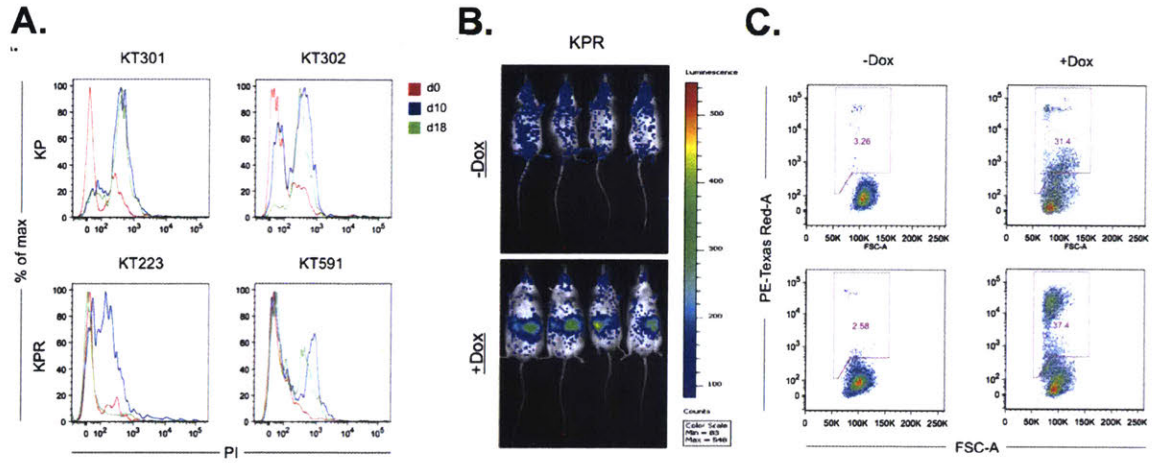
Figure S1. AGO2 overexpression in KP autochthonous model induces lymphomagenesis. (A) TCGA datasets were analyzed via cBioPortal to quantify AGO2 CNV from NSCLC patients (Cerami et al., 2013). (B) A KP lung adenocarcinoma (LUAD) cell line was infected with lentivirus (pSLIK vector) to overexpress AGO2 in a doxycycline-inducible system. Rate of cell migration was measured in the presence and absence of both Matrigel and doxycycline. (C) The number of surface tumors from lung tissue was quantified 4 wk after intravenous administration of the previously described LUAD cell line into immunocompromised mice in the presence and absence of doxycycline. (D) Luciferase imaging of KP and KPR mice at 16 wk after lentivirus delivery. (E) Tumor burden and number of initiated lung tumors (n=6) were quantified from cross sections of lung tissue from KP and KPR mice. (F) Quantification of lymph node sizes from KPR and KP mice were calculated using cross sections of tissue. (G) AGO2 protein expression was assessed with IHC from spleens and lymph nodes of KP (Spleen: n=3; LN: n=9) and KPR (Spleen: n=3; LN: n=16) mice. The percentage of AGO2-expressing cells from each tissue was quantified. (H) The number of KPR and KP lymphocytes actively in mitosis was assessed by phospho-H3 staining using IHC from cross sections of spleens 22 wk after infection. *p<0.05, **p<0.01.

A.

Spleen

B.

Mum1

KP   KPR

Spleen

Lymph Node

Lung

Kidney

C.

KP

KPR

CD138 - APC

B220 - PerCP-Cy5.5

D.

CD3   KP   KPR

Spleen

Lymph Node

F.

Spleen   Kidney

H&E

CD10

E.

Spleen   Lymph Node   Kidney   Lung
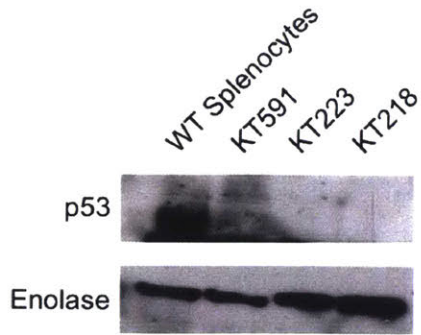
H&E

Kappa

Lambda
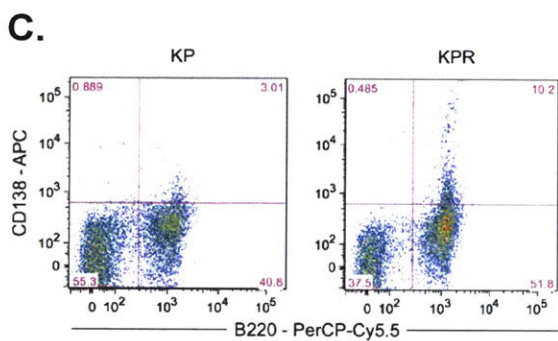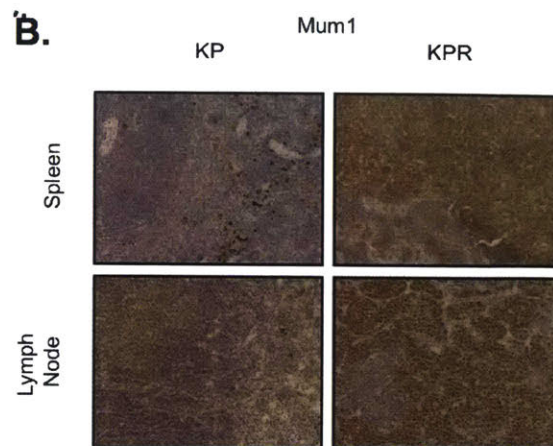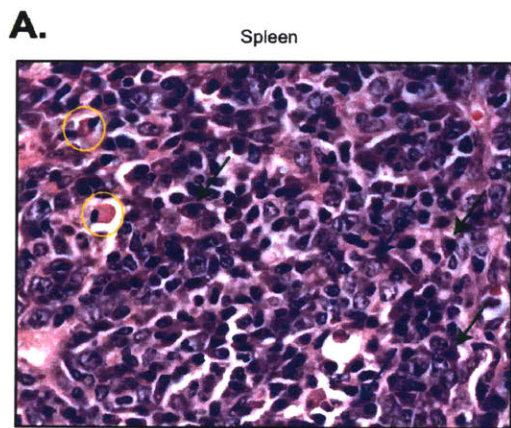
Figure S2. Mice develop an ABC-like DLBCL phenotype from AGO2 overexpression in autochthonous model. (A) Morphological features of hyperplastic lymphocytes were examined with high magnification of H&E stains from a spleen cross-section (yellow circles= Mott cells; blue arrows= large neoplastic cells with vesicular nuclei and open chromatin; green arrows= large neoplastic cells with scant cytoplasm). (B) The population of B220- and CD138-expressing splenocytes derived from KP and KPR mice were analyzed with FACs analysis. Cells were harvested 22 wk after infection. (C) IRF4 protein expression was analyzed with IHC from cross-sections of various tissues from KPR and KP mice. (D) CD3 expression was assessed with IHC to identify T-cells from tissue of KPR and KP mice. (E) Clonal expansion of B-lymphocytes was examined with kappa and lambda staining on serial cross sections of various tissues (spleen, lymph node, kidney and lung) using IHC. (F) Expression of CD10 was analyzed with IHC. Normal kidney tissue was used as a positive control for CD10 (CD10-negative tumor (T) is adjacent to normal kidney (N) tissue).

A.

KT301   KT302

KT223   KT591

d0
d10
d18

% of max

KP

KPR

PI

B.

KPR

-Dox

+Dox

Luminescence

Counts
Color Scale
Min = 83
Max = 546

C.

-Dox   +Dox

PE-Texas Red-A

3.26   31.4

2.58   37.4

FSC-A

D.

	H&E	B220	CD138	Mum1
Spleen				
Lymph Node				

E.

Ctrl   KT591   KT223

Germline IgH →

VDJ Recombination

F. *In Vivo* transplant

Percent survival

Days elapsed

-Dox
+Dox

H.

CD138+   B220^High   CD138+;B220^Low   DN

-Dox

+Dox

CD138 - APC

B220 - PerCP-Cy5.5

G.

CD138 - APC

0.347   32.4

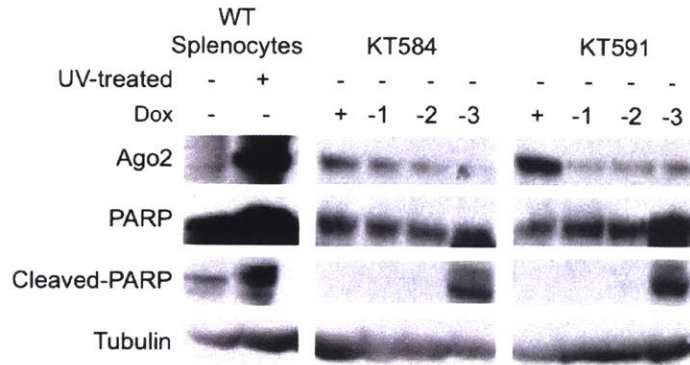61.6   0.458

B220 - PerCP Cy5.5

152

Figure S3. Transplantation of AGO2 overexpressed B-lymphocytes into mice induces ABC-like DLBCL phenotype. (A) Cell death was detected by PI staining and FACS analysis. KPR (KT223 and KT591) and KP (KT301 and KT302) cells were collected at 0, 10 and 18 d after doxycycline withdrawal. (B) FACS analysis of tdTomato-expressing splenocytes harvested from recipient mice. (C) Luciferase imaging was used to monitor transplanted KPR cells *in vivo* in the presence and absence of doxycycline-containing food at 10 wk post-transplantation. (D) Expression pattern of B220, CD138 and IRF4 were analyzed from abnormal lymphoid tissue of KPR recipient mice fed doxycycline. (E) Splenocytes derived from KPR recipient mice that were fed doxycycline-containing food were passaged *in vivo* to determine their tumorigenicity in the presence (n=3) or absence (n=3) of doxycycline. A survival curve was generated. (F) Southern blot analysis was used to determine the clonality of KPR lymphoma cell lines using a J_H4 probe. Genomic DNA isolated from wild type splenocytes was used as a control. (G) KPR cells were FACS sorted into four distinct populations: $B220^{High}$, CD138+, CD138+/$B220^{Low}$, and double negative (DN). (H) KPR lymphoma cells were initially sorted into 4 distinct populations based on B220 and CD138 expression (CD138+, $B220^{High}$, CD138+/$B220^{Low}$ and double negative (DN)). These sorted cells were passaged either in the presence or absence of doxycycline. Here, B220 and CD138 expression was reassessed for each population after 1 wk using FACS analysis.

153

A.

WT Splenocytes | KT591 | KT223 | KT218

p53

Enolase

B.

WT Splenocytes | KT584 | KT591

UV-treated: - +

Dox: - - + -1 -2 -3 + -1 -2 -3

Ago2

PARP

Cleaved-PARP

Tubulin

C.

let-7a

miR-9-5p

miR-148a

Fold Enrichment (Log10)

Immature B Cells
Plasma Cells
KT591
KT591D
KT223
KT223D
KT1220
KT1220D
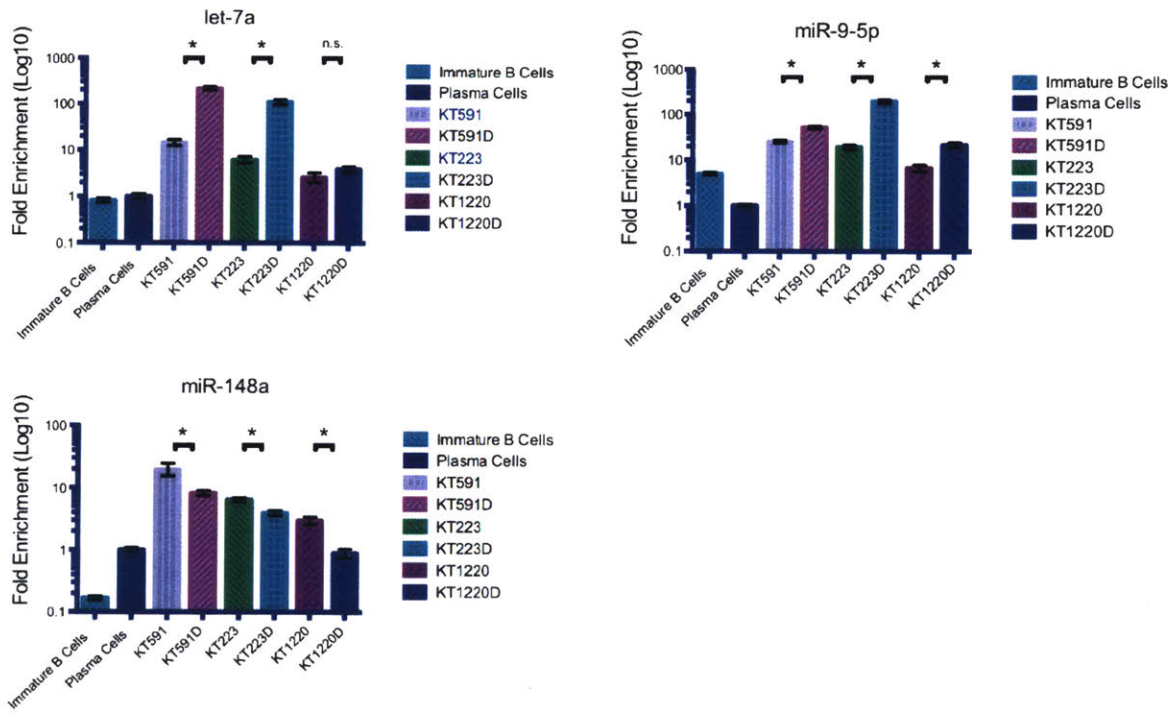
Figure S4. Loss of AGO2 overexpression perturbs block in plasmablastic stage and promotes cell death. (A) Loss of p53 from KPR cell lines was confirmed with western blotting. WT splenocytes was used as a positive c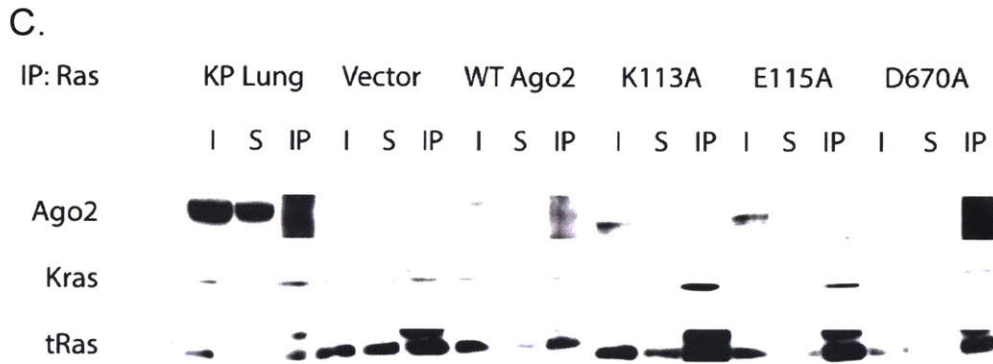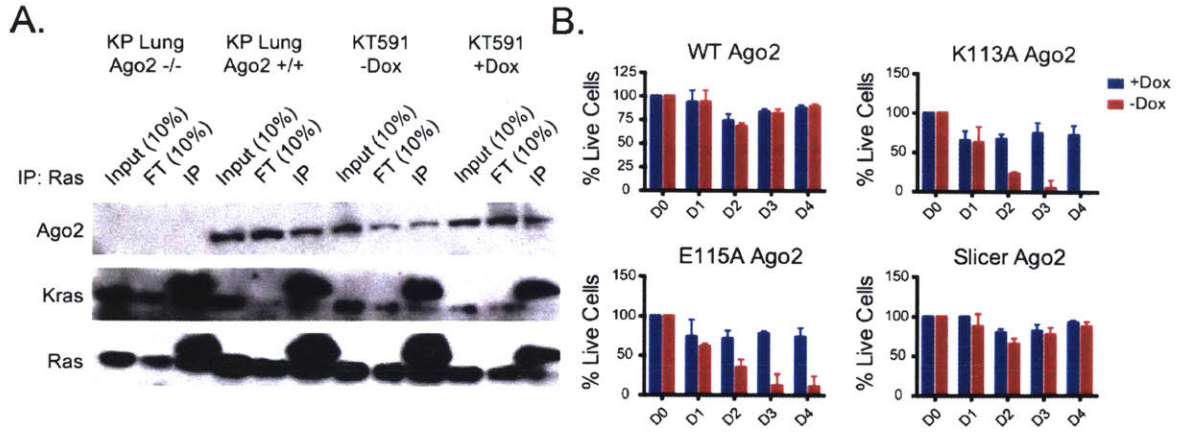ontrol. Enolase was used as a loading control. (B) PARP1 cleavage in KPR cells was detected using western blot. These cells were either grown in the presence of doxycycline or subjected to doxycycline withdrawal for 24, 48 or 72 h before harvesting protein lysate. WT splenocytes were either UV-treated or untreated to induce DNA damage as positive and negative controls, respectively. Tubulin was used as a loading control and total AGO2 expression was assessed. (C) Expression of let-7a, miR-9-5p and miR-148a were analyzed with qPCR and normalized to U6. Immature B cells and Plasma cells were FACS sorted and used as controls. Fold enrichment was normalized to miRNA expression from plasma cells. KT591D, KT223D and KT1220D were maintained on doxycycline.

A.

IP: Ras	KP Lung Ago2 -/-	KP Lung Ago2 +/+	KT591 -Dox	KT591 +Dox
	Input (10%) FT (10%) IP	Input (10%) FT (10%) IP	Input (10%) FT (10%) IP	Input (10%) FT (10%) IP

Ago2

Kras

Ras

B.

WT Ago2

K113A Ago2

E115A Ago2

Slicer Ago2

+Dox
-Dox

C.

IP: Ras   KP Lung   Vector   WT Ago2   K113A   E115A   D670A

           I S IP   I S IP I S IP   I S IP I S IP I S IP

Ago2

Kras

tRas

D.

Ago2:   WT   K113A   E115A   D670A

cMyc-Ago2

Ago2

Enolase

156
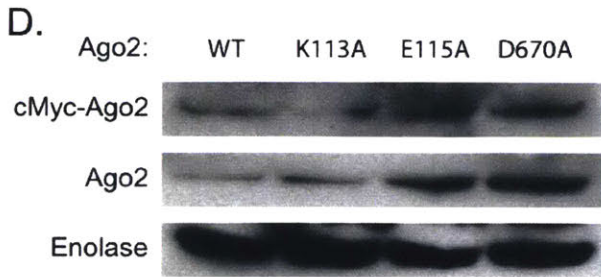
Figure S5. KPR lymphocytes are dependent on the AGO2-KRAS complex for cell survival. (A) Immunoprecipitation (IP) of RAS followed by immunoblotting (AGO2, KRAS, Ras) using protein extract from KT591 cells in the presence and absence of doxycycline. Protein extract from a KP lung adenocarcinoma (KP L.A.) cell line with the genotypes $Ago2^{+/+}$ or $Ago2^{-/-}$ were used as positive and negative controls, respectively, for KRAS and AGO2. (B) IP of cMyc-epitope tag followed by immunoblotting (AGO2, KRAS) using protein extract from KT591 constitutively expressing various cMyc-AGO2 isoforms (WT, K113A, E115A, D670A). (C) The percentage of viable KT591 cells was measured using Trypan blue staining at 0, 1, 2, 3 and 4 d in the presence and absence of doxycycline. (D) Various cMyc-tagged AGO2 isoforms were constitutively expressed in KT591 cells in the presence of doxycycline. Western blotting was used to confirm expression of these cMyc-tagged AGO2 isoforms. Enolase was used as a loading control.

A.

Active NF-KB Subunits

B.

miR-155
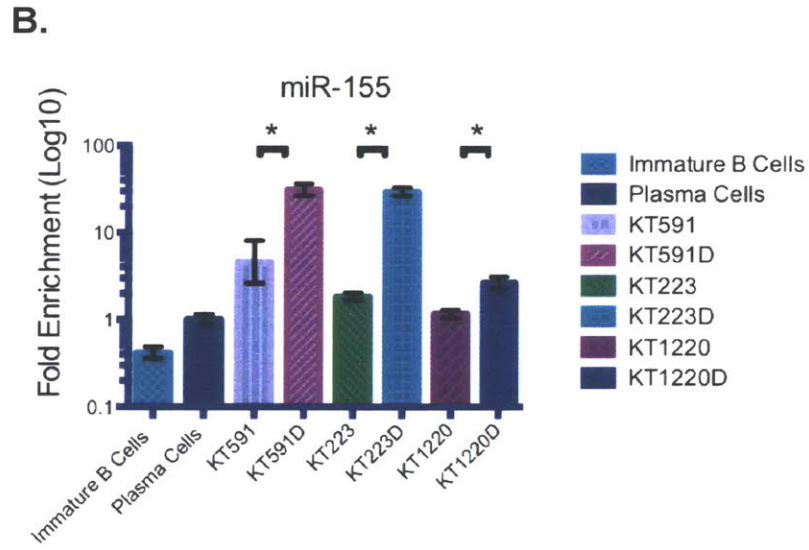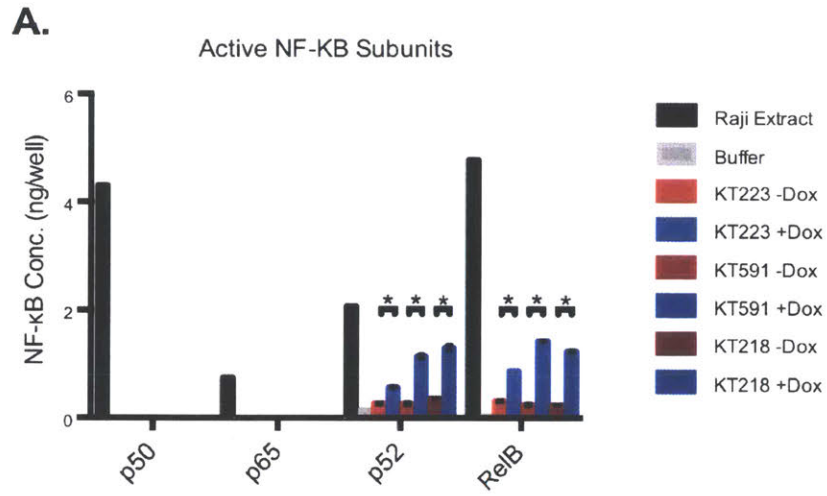
Figure S6. Interaction between oncogenic KRAS and AGO2 promote NF-κB activation. (A) Concentration of activated NF-κB subunits was measured by an ELISA assay. Whole cell extracts from KPR cell lines in the presence and absence of doxycycline were compared. Raji extract and buffer were used as positive and negative controls, respectively. A standard curve was generated using known amounts of recombinant p50 (10, 5, 2.5, 1.25, 0.624, 0.312, 0.156 and 0 ng/well). (B) Expression of miR-155 was analyzed with qPCR and normalized to U6. Immature B cells and Plasma cells were FACS sorted and used as controls. Fold enrichment was normalized to miRNA expression from plasma cells. KT591D, KT223D and KT1220D were maintained on doxycycline.

Primer Name	Orientation	Sequence
GAPDH*	F	GGTCCTCAGTGTAGCCCAAG
GAPDH*	R	AATGTGTCCGTCGTGGATCT
Bcl-6*	F	GACGTTGTCATCGTGGTGAG
Bcl-6*	R	GGTTGCATTTCAACTGGTCA
Blimp1	F	TCCAGAGAGGAGTTCCACCA
Blimp1	R	ACACAAGAAGTTCCTGGTTGGC
cMyc-mAgo2 gA	F	CTAACTCGAACGCTAGCTGTGCGATCGTTTGC GACTAGTGAACAAAAACTCATCTCAGAAGAGG ATCTGGAACAAAAACTCATCTCAGAAGAGGAT CTGTACTCGGGAGCCGGCCCCGTTCTTGC
Cre gB	F	GACCCGACATTAGCGCTACAGCTTAAGCGGAC CATGCCCAAGAAGAAGAGG
Cre gB	R	AGAGTAATTCAACCCCAAACAACAACGTTTCTA ATCGCCATCTTCCAGCAG
EFS pB	F	AAACCGCTGTTCCTAGGAATCCCGAGGCCTTC GAGTGGCTCCGGTGCCCG
EFS pB	R	CCGCTTAAGCTGTAGCGCTAATGTCGGGTCCC GCGTCACGACACCTGTG
FHmAgo2	F	CGCGAATTCATGTACTCGGGAGC
FHmAgo2	R	GCGGCGGCCGCTCAAGCAAAG
FH-mAgo2 gA	F	CTAACTCGAACGCTAGCTGTGCGATCGTTTGC GACTAGTCCACCATGGACTACAAGGACGACGA TGACAAGTACCCTTATGACGTGCCCGATTACG CTTACTCGGGAGCCGGCCCCG
FH-mAgo2 gA	R	AGGCCTCGGGATTCCTAGGAACAGCGGTTTTC AAGCAAAGTACATGGTGCGCAGTGTGTC
let-7a*	F	ACT ATA CAA CCT ACT ACC TCA
miR-9-5p*	F	TCA TAC AGC TAG ATA ACC AAA GA
miR-148a*	F	ACAAAGTTCTGTAGTGCACTGA
miR-155*	F	CCC CTA TCA CAA TTA GCA TTA A
Mum1	F	GAGAGAAGCAGAGGATGCCC
Mum1	R	TGTGGACTTGAACCCTGTGG
Myc*	F	AGCTGTTTGAAGGCTGGATT
Myc*	R	AATAGGGCTGTACGGAGTCG
NF-κB2	F	GTTGAGAAGCCTGGTGGACA
NF-κB2	R	CGGGTCTCAGGACCTTTCAG
pGK pA	F	GATCAGTGTGAGGGAGTGTAAAGCTGGTTTAT TCTACCGGGTAGGGGAGG
pGK pA	R	AAACGATCGCACAgctagcGTTCGAGTTAGTCCC GGAGATGAGGAAGAGG
pGK pB	F	AAACCGCTGTTCCTAGGAATCCCGAGGCCTAT TCTACCGGGTAGGGGAGG
pGK pB	R	CCGCTTAAGCTGTAGCGCTAATGTCGGGTCTC

		CCGGAGATGAGGAAGAGG
TRE pA	F	GATCAGTGTGAGGGAGTGTAAAGCTGGTTTCG ACGGTtTCGATATGTCGAG
TRE pA	R	AAACGATCGCACAGCTAGCGTTCGAGTTAGTG ACTTACGtATTCTCCAGGC

Table 1: List of primers used.

Antibody	Vendor/Catalog	Assay(s)	Dilution
Ago2	Cell Signaling Technology/2897S	W.B./IHC/ IP	1:1000/1:200/ 1:50
Akt	Cell Signaling Technology/2920S	W.B.	1:2000
B220	eBioscience/14-0452-82	IHC	1:100
B220-Pacific Blue	BD Bioscience/103227	FACS	1:400
B220-PerCP Cy5.5	BD Bioscience/552771	FACS	1:400
CD79a-BV421	BD Biosceience/562852	FACS	1:200
CD138	Abcam/ab104803	IHC	6µg/mL
CD138-APC	BD Bioscience/558626	FACS	1:200
CD19-FITC	BD Bioscience/115506	FACS	1:400
CD3	Abcam/ab5690	IHC	1:150
Cleaved-PARP	Cell Signaling Technology/9544S	W.B.	1:2000
cMyc-epitope tag	Cell Signaling Technology/2278S	W.B./IP	1:1000/1:200
Enolase	Cell Signaling Technology/3810S	W.B.	1:1000
ERK1/2	Cell Signaling Technology/9107S	W.B.	1:1000
IgG	Abcam/ab171870	IP	1:100
HA	Roche/11867423001	W.B.	200ng/mL
Kappa Light Chain-HRP	Novus/NB7549	IHC	1:200
Ki-67	Abcam/ab15580	IHC	1:500
Kras	Santa Cruz Biotechnology/sc-30	W.B.	1:200
Lambda Light Chain	Thermo Scientific/PA5-16651	IHC	1:100
MEK1/2	Cell Signaling Technology/4694S	W.B.	1:1000
mTOR	Cell Signaling Technology/2983S	W.B.	1:1000
Mum1/Irf4	Abcam/ab104803	IHC/FAC S	1:125/1:100

p53	Cell Signaling Technology/2524S	W.B.	1:1000
p100	Cell Signaling Technology/4882S	W.B.	1:1000
Parp	Cell Signaling Technology/9542S	W.B.	1:1000
Pax5	BD Bioscience/562688	FACS	1:200
phospho-AKT	Cell Signaling Technology/4060S	W.B.	1:1000
phospho-ERK1/2	Cell Signaling Technology/4370S	W.B.	1:1000
phospho-H3	Cell Signaling Technology/9701	IHC	1:200
phospho-MEK1/2	Cell Signaling Technology/9154S	W.B.	1:1000
phospho-mTor	Cell Signaling Technology/5536S	W.B.	1:1000
phospho-γH2A.X	Millipore/05-636	W.B.	1:1000
Sca1-PE	eBioscience/15-5981-81	FACS	1:400
Total H2A.X	Cell Signaling Technology/7631	W.B.	1:4000
Total Ras	Cell Signaling Technology/3965S	W.B./IP	1:1000/1:50
Tubulin	Abcam/ab6046	W.B.	1:500

Table 2: List of primers used.

References

Adam BD, Claffey KP, White BA. 2009. Argonaute-2 Expression is Regulated by Epidermal Growth Factor Receptor and Mitogen-Activated Protein Kinase Signaling and Correlates with a Transformed Phenotype in Breast Cancer Cells. *Endocrinology* **150:** 14-23.

Ambros V. 2004. The functions of animal microRNAs. *Nature* **431:** 350-355.

Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP. 2008. The impact of microRNAs on protein output. *Nature* **455:** 64-71.

Bang D, Wilson W, Ryan M, Yen JJ, Baldwin AS. 2013. GSK-3α Promotes Oncogenic KRAS function in Pancreatic Cancer via TAK1-TAB Stabilization and Regulation of Non-Canonical NF-κB. *Cancer Discov* **3:** 690-703.

Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. *Cell* **136:** 215-233.

Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch et al. 2002. Frequent deletions and down-regulation of micro-RNA genes *miR15* and *miR16* at 13q14 in chronic lymphocytic leukemia. *Proc. Natl Acad Sci. USA* **99:** 15524-15529.

Castillo JJ, Bibas Michele, Miranda RN. 2015. The biology and treatment of plasmablastic lymphoma. *Blood* **125:** 2323-2330.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BU et al. 2013. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimentional Cancer Genomics Data. *Cancer Discov* **2:** 401.

Chen Y, Zheng Y, You X, Yu M, Fu G, Su X et al. 2016. Kras is Critical for B Cell Lymphopoiesis. *J Immunol* **196:** 1678-1685.

Cheng N, Li Y, Han Z. 2013. Argonaute2 Promotes Metastasis by Way of Up-regulating Focal Adhesion Kinase Expression in Hepatocellular Carcinoma. *Hepatology* **57:** 1906-1918.

Colovai AI, Giatzikis C, Ho EK, Farooqi M, Suciu-Foca N, Cattoretti G et al. 2004. Flow cytometry analysis of normal and reactive spleen. *Mod Pathol* **17:** 918-927.

Davis RE, Brown KD, Siebenlist U, Staudt LM. 2001. Constitutive nuclear factor kappa B activity is required for survival of activated B Cell-like diffuse large B cell lymphoma cells. *J Exp Med* **194:** 1861-1874.

Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR et al. 2009. RNAi in budding yeast. *Science* **326**: 544-550.

Drosten M, Dhawahir A, Sum EY, Urosevic J, Lechuga CG, Esteban LM et al. 2010. Genetic analysis of Ras signaling signaling pathways in cell proliferation, migration and survival. *EMBO J* **29**: 1091-1104.

DuPage M, Dooley AL, Jacks T. 2009. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nat Protoc* **4**: 1063-1072.

Eis PS, Tam W, Sun L, Chadburn A, Li Z, Gomez MF et al. 2005. Accumulation of miR-155 and *BIC* RNA in human B cell lymphomas. *Proc. Natl. Acad. Sci. USA* **102**: 3627-3632.

Greene LA, Lee HY, Angelastro JM. 2009. The transcription factor ATF5: role in neuraldevelopment and neural tumors. *J Neurochem* **108**: 11-22.

Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Delabie J, Ott G et al. 2004. Confirmation of the molecular classification of diffuse large B cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* **103**: 275-282.

Ichikawa A, Kinoshita T, Watanabe T, Kato H, Nagai H, Tsushita K et al. 1997. Mutations of the p53 gene as a prognostic factor in aggressive B cell lymphoma. *N Engl J Med* **337**: 529-534.

Kluiver J, Poppema S, de Jong D, Blokzijl T, Harms G, Jacobs S et al. 2005. BIC and miR-155 are highly expressed in Hodgkin, primary mediastinal and diffuse large B cell lymphomas. *J Pathol* **207**: 243-249.

Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. 2007. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet* **39**: 673-677.

Kwak PB and Tomari Y. 2012. The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat Struct Mol Biol* **19**: 145-152.

Li L, Yu C, Gao H, Li Y. 2010. Argonaute proteins: potential biomarkers for human colon cancer. *BMC cancer* **10**: 38.

Lin R, Lin Y, Chen J, Kuo H, Chen Y, Diccianni MB et al. 2010. microRNA Signature and Expression of *Dicer* and *Drosha* Can Predict Prognosis and Delineate Risk groups in Neuroblastoma. *Cancer Res* **70**: 7841-7850.

Lin S, Gregory RI. 2015. 2015. MicroRNA biogenesis pathways in cancer. *Nat*

Rev Cancer **15:** 321-333.

Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song J et al. 2004. Argonaute2 is the Catalytic Engine of Mammalian RNAi. *Science* **305:** 1437-1441).

Niu H, Ye BH, and Dalla-Favera R. 1998. Antigen receptor signaling induces MAP kinase-mediated phosphorylation and degradation of the BCL-6 transcription factor. *Genes & Dev* **12:** 1953-1961.

O'Carroll D, Mecklenbrauker I, Das PP, Santana A, Koenig U, Enright AJ et al. 2007. A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. *Genes & Dev* **21:** 1999-2004.

Pasqualucci L, Compagno M, Houldsworth J, Monti S, Grunn A, Nandula SV et al. 2006. Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *JEM* **203:** 311-317.

Pylayeva-Gupta Y, Grabocka E, Bar-Sagi D. 2011. RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer* **11:** 761-774.

Ravi A, Gurtan AM, Kumar MS, Bhutkar A, Chin C, Lu V et al. 2012. Proliferation and Tumorigenesis of a Murine Sarcoma Cell Line in the Absence of DICER1. *Cancer Cell* ***21***: 848–855.

Reddy A, Zhang J, Davis NS, Moffitt AB, Love CL, Waldrop A et al. 2017. Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell* **171:** 481-494.

Shankar S, Pitchiaya S, Malik R, Kothari V, Hosono Y, Yocum AK et al. 2016. KRAS engages AGO2 to Enhance Cellular Transformation. *Cell Rep* **14:** 1-14.

Shaulian E and Karin M. 2001. AP-1 in cell proliferation and survival. *Oncogene* **20:** 2390-2400.

Shen J, Xia W, Khotskaya YB, Huo L, Nakanishi K, Lim S et al. 2013. EGFR modulates microRNA maturation in response to hypoxia through phosphorylation of AGO2. *Nature* **497:** 383 -387.

Shieh A, Ward Af, Donlan KL, Hardin-Theobald ER, Xu J, Mullighan CG et al. 2013. Defective K-Ras oncoproteins overcome impaired effector activation to initiate leukemia in vivo. *Blood* **121:** 4884-4893.

Su H, Trombly MI, Chen J, Wang X. 2008. Essential and overlapping functions

for mammalian Argonautes in microRNA silencing. *Genes & Dev* **23**: 304-317.

Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kyrukov G, Cowley GS et al. 2017. Defining a Cancer Dependency Map. *Cell* **170**: 564-576.

Vaksman O, Hetland TE, Trope CG, Reich R, Davidson B. 2012. Argonaute, Dicer and Drosha are up-regulated along tumor progression in serious ovarian carcinoma. *Hum Pathol* **43**: 2062-2069.

Vega F, Chang C, Medeiros LJ, Udden MM, Cho-Vega JH, Lau C et al. 2005. Plasmablastic lymphomas and plasmablastic plasma cell myelomas have nearly identical immunophenotypic profiles. *Mod Pathol* **18**: 806-815.

Wegert J, Ishaque N, Vardapour R, Geörg C, Gu Z, Bieg M et al. 2015. Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie blastemal type Wilms tumors. *Cancer Cell* **27**: 298-311.

Wilson WH, Young RM, Schmitz R, Yang Y, Pittaluga S, Wright G et al. 2015. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med* **21**: 922-926.

Yang M, Haase AD, Huang F, Coulis G, Rivera KD, Dickinson BC et al. 2014. Dephosphorylation of Tyrosine 393 in Argonaute 2 by Protein Tyrosine Phosphatase 1B Regulates Gene Silencing in Oncogenic RAS-induced Senescence. *Mols Cell* **55**: 782-790.

Zhou Y, Chen L, Barlogie B, Stephens O, Wu X, Williams DR. et al. 2010. High-risk myeloma is associated with global elevation of miRNAs and overexpression of EIF2C2/AGO2. *PNAS* **107**: 7904-7909.

Zhu C, Bogue MA, Lim DS, Hasty P, Roth DB. 1996. Ku86-deficient mice exhibit severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. *Cell* **86**: 379-389.

Chapter 4

Discussion and Future Directions

In this thesis, we investigated the role of AGO2 in promoting tumor progression in two mouse models of KRAS-activated and *Tp53*-deleted cancers. In the case of non-small cell lung cancer (NSCLC), we determined that AGO2 overexpression plays a prominent role in the progression, but not initiation, of NSCLC. From *in vitro* assays, AGO2 overexpression promoted cell motility and invasiveness, but had no effect on cell doubling rates. Consistently, *in vivo* studies ectopically expressing AGO2 did not enhance the lung tumor burden, but promoted malignancy transformation of initiated tumors and increased the capacity of tumor cells to invade and colonize distant tissues. Additionally, we generated a novel mouse model to initiate an ABC-DLBCL-like malignancy using a transplantation approach. We thoroughly characterized this model, demonstrating that AGO2 overexpression-induced cellular transformation results in the arrest of transformed cells in the hyperproliferative plasmablastic stage of B cell development. In this system, we have identified that the KRAS-AGO2 interaction, enhanced through AGO2 overexpression, is essential in inducing pro-survival signals through multiple pathways. Moreover, we have provided evidence that the KRAS and AGO2 dependency is possibly a general phenomenon that occurs in many cancer types.

Uncovering novel oncogenic AGO2 functions in promoting tumor progression

Given that miRNA dysregulation frequently contributes to tumorigenesis and cancer progression and elevated AGO2 expression is detected in many types of human malignancies, we reasoned that overexpressing AGO2 with

common driver mutations could accelerate tumor progression in our mouse models (Iorio and Croce, 2012; Cerami et al., 2013). In the KP mouse model of NSCLC, we determined that overexpressing AGO2 enhances the ability of established tumor cell lines to migrate and colonize distant tissues. Similarly, initiated tumors from the KP autochthonous model frequently developed into malignant lesions in AGO2 overexpressed conditions compared to control mice. Therefore, it would be interesting to explore the dependency on AGO2 overexpression in promoting metastasis. Here, I propose to initiate lung tumors in the KP autochthonous model with doxycycline-inducible AGO2 overexpression using a low viral titer. At week 8 post lentiviral delivery, half of the animals are switched to regular food to withdraw ectopic AGO2. Based on previous results, I hypothesize that animals maintained on doxycycline food will develop metastatic tumors at a higher frequency than animals with doxycycline withdrawal. Moreover, since lung tumors can be inititated and progress in conditions where only endogenous AGO2 is expressed, I reason that withdrawal of ectopic AGO2 from established tumors will have negligible effects on tumor regression.

Interestingly, human cancer cells, such as NSCLC cell lines, that were dependent on AGO2 for survival were commonly co-dependent on TWIST1, a positive regulator of metastasis (Yang et al., 2004). Together, these results argue that elevated AGO2 expression in lung tumor cells promotes metastasis but the mechanism that facilitates this is still unclear. There are a couple of plausible mechanisms that could contribute to this AGO2 overexpressed phenotype. First, since it has been well established that high *Nkx2-1*, *Foxa2*, and *Cdx2* expression

in KP lung tumors work synergistically to inhibit lung adenocarcinoma metastasis, it is possible that elevated AGO2 expression preferentially stabilizes miRNAs that target these specific transcription factors (Li et al., 2015). In doing so, the transcriptional profile of these tumor cells would be altered to favor a metastasis program. Alternatively, upregulation of AGO2 in this system could enhance its physical interaction with KRAS, thereby promoting activation of signal transduction pathways required to rewire transcriptional circuits to favor metastasis.

In our novel B cell transplantable mouse model to induce lymphomagenesis via AGO2 overexpression, we have identified that the interaction between oncogenic KRAS and AGO2 is important in promoting cellular transformation. We demonstrated that overexpressing KRAS-binding AGO2 mutant proteins does not rescue cell survival in KPR tumor cells that are depleted of ectopically expressed wild type AGO2. Although this does not prove that miRNA-independent AGO2 functions are solely responsible for facilitating malignant transformation in this system, it does suggest that the physical interaction between AGO2 and KRAS is crucial in promoting cellular transformation by altering the protein interactome of both molecules.

Interestingly, Shankar et al. (2016) has demonstrated that KRAS and AGO2 primarily co-localize to various organelle membranes of various human cancer cell lines, with a significant fraction of this complex detected at the endoplasmic reticulum. This opens the possibility for novel interactions between the KRAS-AGO2 complex with other signaling molecules to enhance cell

proliferation and survival. To explore this further in our system, we could perform

co-localization experiments using immunofluorescence microscopy and

fractionation techniques to determine the intracellular localization pattern of this

complex. Additionally, it would be informative to identify changes in the KRAS

protein interactome of KPR tumor cells via IP-mass spectrometry in the presence

and absence of doxycycline. This could reveal additional KRAS effector

molecules to help elucidate novel mechanisms contributing to the tumorigenicity

of the KPR cell lines induced by AGO2 upregulation.

Upregulation of AGO2 cooperatively activates signaling pathways involved

in cell proliferation and survival with oncogenic KRAS

As a consequence of ectopic AGO2 expression in KPR tumor cells in our

activated B cell-like lymphoma mouse model, we detected a significant increase

in activity of multiple signaling pathways, including NF-κB. Using KRAS-binding

AGO2 mutants, we have demonstrated that the KRAS-AGO2 complex is at least

partially responsible for elevated NF-κB signaling. Here, a reduction in specific

activated NF-κB subunits, p52 and Rel-B, were detected with loss of ectopically

expressed wild type AGO2. Given that NF-κB activity often promotes cell

proliferation and anti-apoptotic signals, these results suggest that the physical

interaction between KRAS and AGO2 enhances cellular transformation through

NF-κB activity in this system (Guttridge et al., 1999; Hoesel and Schmid, 2013).

However, we have also detected increased activity of additional KRAS-regulated

pathways, including the PI3K-mTOR and MAPK signaling cascades. This opens

the possibility of crosstalk between multiple signaling pathways that synergistically contribute to malignant transformation in the context of AGO2 overexpression.

One approach to identify which pathways are essential for cell survival in this system would be to use a genome-scale RNAi screening strategy (Possemato et al. 2011). With this strategy, a lentivirally delivered genome-wide shRNA library consisting of multiples shRNAs targeting each gene could be used to infect AGO2 overexpressed KPR cells. Here, we can determine which genes exhibit the highest sensitivity to their suppression by analyzing the depletion of each shRNA in all screened KPR cell lines. Subsequently, drug-based inhibitors targeting these identified genes can be used to confirm KPR cell sensitivity to their suppression and, hence, elucidate essential pathways that AGO2 overexpressed KPR cells are dependent on for cell proliferation and survival.

Exploring the synergistic effects of AGO2 overexpression, KRAS activation and *Tp53* deletion

In both the NSCLC and ABC-like DLCBC mouse models used, we investigated the role of AGO2 overexpression in KRAS activated and *Tp53* deleted tumor cells. We have identified that the physical interaction between AGO2 and KRAS is essential in promoting cell proliferation and survival in AGO2 overexpressed lymphoma cell lines. However, the dependency on p53 ablation has yet to be explored in this system. We have provided evidence of a general dependency on AGO2 expression in *KRAS* and *TP53* mutant human cancer cell

lines, including those of the NSCLC subtype (Tsherniak et al., 2017). Thus, I hypothesize that the phenotypes observed with AGO2 overexpression in both KP-associated mouse models is dependent on *Tp53* deletion. To ascertain this notion, we can explore the effects of AGO2 overexpression on lung tumor development and metastasis in a $Kras^{LSL-G12D/+}$;$p53^{+/+}$ autochthonous mouse model. Similarly, the consequences of ectopic expression of AGO2 on B cell lymphoma initiation and maintenance can be explored using B cells derived from bone marrow of $Kras^{LSL-G12D/+}$;$p53^{+/+}$ mice.

Establishing clinical relevance of AGO2 overexpressed ABC-like DLBCL mouse model

DLBCL is the most common subtype of non-Hodgkin lymphoma. Based on previous molecular characterization, there are two major subtypes of DLBCL: germinal center B cell (GCB) and activated B cell (ABC) DLBCL (Lenz et al., 2008). Both standard and aggressive forms of chemotherapeutic regimens have been fairly successful in treating a wide range of GCB DLBCL patients but have remained largely ineffective in treating those diagnosed with ABC DLBCL (Rosenwald et al., 2002). Hence, ABC DLBCL is associated with substantially worse prognosis and outcome. Currently, there are no well-established mouse models of ABC DLBCL to study the molecular mechanisms that contribute to the initiation and progression of this malignancy. As expected, lack of understanding of the genetic and molecular components that promote ABC DLBCL makes it difficult to develop effective therapies against this aggressive lymphoma subtype.

In this thesis, we have used ectopic expression of AGO2 in KRAS activated and *Tp53* deleted B cells to initiate lymphomagenesis, resulting in a malignancy that resembles human ABC DLBCL based off of histological and molecular features. Nevertheless, to use this novel mouse model to gain insights into the mechanisms that contribute to the human form of the disease, we must compare genetic and transcriptional aberrations between our KPR lymphoma cell lines to tumor cells originating from ABC DLBCL patients. In the event that both oncogenic KRAS mutations and AGO2 amplifications are detected from patient samples, our ABC-like DLBCL mouse model could be used to identify novel targets for targeted cancer therapies that could be effective in treating these select patients.

Differentiating ABC DLBCL from plasmablastic lymphoma

As previously alluded, hematopoietic malignancies are highly complex due to the extensive subtypes classified by their cell of origin and histological and molecular features. Both ABC DLBCL and plasmablastic lymphoma (PBL) are post-germinal center lymphomas associated with tumorigenic plasmablastic cells. ABC DLBCL is charactized by the arrest of tumorigenic cells in the plasmablastic stage of B cell development; however, PBL is associated with "late" plasmablasts that are capable of terminally differentiating into plasma cells (Lenz G and Staudt LM, 2010; Said JW, 2013). Because there is a lack of understanding of the molecular and genetic features of human PBL due to this malignancy's rarity, it is occasionally difficult to distinguish PBL from ABC DLBCL.

To date, immunophenotyping a small subset of PBL patients has revealed that PBL is positive for CD138 and BLIMP1. Though not always the case, PBL is also generally positive for CD79a, MUM1, and CD10 and negative for PAX5, BCL-6 and CD20. Similarly, ABC DLBCL is negative for BCL-6 and generally positive for CD138, but not always the case. Because ABC DLBCL is associated with "early" plasmablasts, which requires NF-κB activation to exit GC reactions, incapable of terminally differentiating, it is positive for MUM1/IRF4 and negative for BLIMP1, the master regulator of plasma cell differentiation (Vega F et al., 2005; Lenz G and Staudt LM, 2010; Castillo JJ et al., 2015; Said JW, 2013; Harmon CM and Smith LB, 2016).

In our B cell lymphoma mouse model inititated by AGO2 overexpression, histology revealed a very heterogenous population of abnormal plasmablasts and plasma cells with features that commonly found in both ABC DLBCL and PBL. We also provided evidence that these lymphoid neoplasms were predominantly negative for BCL-6, CD10, and BLIMP1, which is characteristic of ABC DLBCL. However, FACS analysis revealed a large subset of tumorigenic cells that were positive for IRF4 but only approximately 50% of these cells were positive for CD138. Together, these data suggest that AGO2 overexpression in KP B cells initiated an ABC-like DLBCL with plasmablastic features found in PBL.

References

Castillo JJ, Bibas Michele, Miranda RN. 2015. The biology and treatment of plasmablastic lymphoma. *Blood* **125:** 2323-2330.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BU et al. 2013. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimentional Cancer Genomics Data. *Cancer Discov* **2:** 401.

Guttridge DC, Albanese C, Reuther JY, Pestell RG, Baldwin AS Jr. 1999. NF-kappaB controls cell growth and differentiation through transcriptional regulation of cyclin D1. *Mol Cell Biol* **19:** 5785-5799.

Harmon CM and Smith LB. 2016. Plasmablastic Lymphoma: A Review of Clinicopathologic Features and Differential Diagnosis. *Arch Pathol Lab Med* **140:** 1074-1078.

Hoesel B and Schmid JA. 2013. The complexity of NF-κB signaling in inflammation and cancer. *Mol Cancer* **12:** 86.

Iorio MV and Croce CM. 2012. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO Mol Med* **4:** 143-159.

Lenz G, Wright GW, Emre NC, Kohlhammer H, Dave SS Davis RE et al. 2008. Molecule subtypes of diffuse large B cell lymphoma arise by distinct genetic pathways. *Proc Natl Acad Sci U.S.A.* **105:** 13520-13525.

Lenz G and Staudt LM. 2010. Aggressive Lymphomas. *N Engl J Med* **362:** 1417-1429.

Possemato R, Marks KM, Shaul YD, Pacold ME, Kim D, Birsoy K et al. 2011. Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476:** 346-350.

Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI et al. 2002. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B cell lymphoma. *New Engl J Med* **346:** 1937-1947.

Said JW. 2012. Aggressive B-cell lymphomas: how many categories do we need?. *Mod Pathol* **1:** S42-56.

Shankar S, Pitchiaya S, Malik R, Kothari V, Hosono Y, Yocum AK et al. 2016. KRAS Engages AGO2 to Enhance Cellular Transformation. *Cell Reports* **14:** 1-14.

Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kyrukov G, Cowley GS et al. 2017. Defining a Cancer Dependency Map. *Cell* **170:** 564-576.

Vega F, Chang C, Medeiros LJ, Udden MM, Cho-Vega JH, Lau C et al. 2005. Plasmablastic lymphomas and plasmablastic plasma cell myelomas have nearly identical immunophenotypic profiles. *Mod Pathol* **18:** 806-815.

Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C et al. 2004. Twist, a Master Regulator of Morphogenesis, Plays an Essential Role in Tumor Metastasis. *Cell* **117:** 927-939.

Chapter 5

Appendix

The work in this chapter is adapted with permission from the following manuscript:

Title: Dicer loss and recovery induces an oncogenic switch driven by transcriptional activation of the oncofetal Imp1-3 family

Authors: Courtney K. JnBaptiste[1,2], Allan M. Gurtan[2], Kevin K. Thai[1,2], Victoria Lu[1,2], Arjun Bhutkar[2], Mei-Ju Su[3], Asaf Rotem[3], Tyler Jacks[1,2], Phillip A. Sharp[1,2‡]

Affiliations:

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]David H. Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA 02139, USA

[3]Department of Medical Oncology and Center for Cancer Precision Medicine, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA.

[‡]To whom correspondence should be addressed: sharppa@mit.edu

Abstract

miRNAs are post-transcriptional regulators of gene expression critical for organismal viability. Changes in miRNA activity are common in cancer, but how these changes relate to subsequent alterations in transcription and the process of tumorigenesis is not well understood. Here, we report a deep transcriptional, oncogenic network regulated by miRNAs. We present analysis of the gene expression and phenotypic changes associated with global miRNA restoration in miRNA-deficient fibroblasts. This analysis uncovers a miRNA-repressed network containing oncofetal genes Imp1-3 that is upregulated primarily transcriptionally greater than 100-fold upon Dicer loss and is resistant to re-silencing by complete restoration of miRNA activity. This Dicer-resistant epigenetic switch confers tumorigenicity to these cells. Let-7 targets Imp1-3 are required for this

179

tumorigenicity and feed back to reinforce and sustain expression of the oncogenic network. Together, these Dicer-resistant genes constitute an mRNA expression signature that is present in numerous human cancers and is associated with poor survival.

Introduction

MicroRNAs (miRNAs) are an abundant class of noncoding RNAs that repress gene expression typically ~2-fold through target degradation and/or translational inhibition (Baek et al. 2008; Bartel 2009; Ghildiyal and Zamore 2009). miRNAs regulate many processes, including development (Bernstein et al. 2003; Wienholds et al. 2003), differentiation (Kanellopoulou et al. 2005; Xu et al. 2009), apoptosis (Cimmino et al. 2005; Harfe BD 2005), and proliferation (Wang et al. 2008).

Consistent with such broad influence, miRNA deregulation is commonly observed in cancer, with both oncogenic and tumor suppressive roles having been reported (Ventura and Jacks 2009; Lin and Gregory 2015). Let-7 miRNAs comprise a tumor suppressor family that is downregulated in multiple cancers and correlates inversely with patient survival (Calin 2004; Takamizawa et al. 2004; Boyerinas et al. 2010; Nguyen et al. 2014; Kugel et al. 2016; Manier et al. 2016; Powers et al. 2016). Underlying the potent tumor suppressor activity of let-7 is repression of well-established oncogenic mRNAs such as Lin-28, Hmga2,

Myc, Ras and the Igf2 mRNA-binding protein family Imp1-3 (Johnson et al. 2005; Lee and Dutta 2007; Sampson et al. 2007; Boyerinas et al. 2008).

Imp1, Imp2, Imp3 (Imp1-3) are an RNA-binding protein (RBP) family that control RNA localization, stability and translation (Bell et al. 2013). Along with Hmga2 and Lin-28, Imp1-3 are categorized as oncofetal genes, defined based on their high expression during embryogenesis and re-expression in diverse cancer types, where they collectively promote stemness, proliferation, invasion and metastasis (Boyerinas et al. 2008; Janiszewska et al. 2012; Gurtan et al. 2013; Nishino et al. 2013; Degrauwe et al. 2016). Let-7 and its oncofetal targets have been studied intensely in cancer and several mechanisms for neutralizing let-7 activity have emerged. Besides genetic deletion (Calin 2004), biogenesis of let-7 is suppressed by Lin-28, another developmentally regulated RNA-binding protein (RBP) family (Viswanathan et al. 2008; Molenaar et al. 2012; Madison et al. 2013). Furthermore, alternative 3'UTR usage by Imp1 and Hmga2 promotes transformation (Mayr C 2007.; Mayr and Bartel 2009). A more recent study described a novel mechanism in neuroblastoma, where excessive amplification of Myc mRNA might sponge let-7 activity away from other targets (Powers et al. 2016). Additionally, Imp1-3 may sequester Hmga2, Lin-28 and other mRNA targets by directly binding let-7 target sites to impede miRNA activity or modulate the association of transcripts with RISC (Jonson et al. 2014; Busch et al. 2016; Degrauwe et al. 2016; Ennajdaoui et al. 2016). Thus, in cancer, inhibition of let-7 activity and extensive co-regulation among its oncofetal targets are thought to synergize to promote oncogenesis.

Beyond perturbation of individual miRNAs in cancer, global miRNA deregulation through impairment of the miRNA processing machinery drives tumor development (Melo et al. 2009; Melo et al. 2010; Walz et al. 2015; Wegert et al. 2015). Mouse models have demonstrated that global miRNA loss through reduced Dicer expression promotes tumorigenesis (Kumar et al. 2007; Kumar et al. 2009). In humans, germline or somatic Dicer mutations predispose patients to cancer and are collectively defined as Dicer1 syndrome (Foulkes et al. 2014). Germline loss of function mutations in Dicer have been identified in pediatric pleuropulmonary and pituitary blastomas, familial cystic nephroma, and Wilms' tumor (Hill DA 2009.; Bahubeshi et al. 2010; Foulkes et al. 2011; de Kock et al. 2014). Hypomorphic somatic mutations in Dicer, primarily altering its catalytic residues are associated with multiple cancer types (Heravi-Moussavi A 2012.; Rakheja et al. 2014; Chen et al. 2015).

It is unclear why a global reduction in miRNAs promotes cancer (Heravi-Moussavi A 2012.; Chen et al. 2015). Here, we use a cell-based model of miRNA loss and subsequent reconstitution to identify miRNA-regulated gene networks and characterize their effects on tumorigenesis. In immortalized Dicer null mesenchymal stem cells, a fibroblast cell type, we restored global miRNA expression and subsequent post-transcriptional activity by stably re-expressing Dicer. Although post-transcriptional repression by let-7 and other miRNAs is restored, there persists a gene signature of transcriptionally upregulated oncogenes, including the let-7-regulated Imp1-3 family, whose de-repression is largely resistant to Dicer and let-7 restoration. These Dicer-rescued cells are

182

highly tumorigenic, consistent with expression of this oncogenic signature. Expression of this oncogenic network is reinforced in a feedback loop by three of its members, namely Imp1-3. Demonstrating clinical relevance, this set of irreversibly activated miRNA targets is associated with poor survival in cancer patients.

Results

### hsDicer restores global miRNA expression in Dicer null fibroblasts

Previously, we reported the generation of SV40-Large T immortalized, Dicer wild-type and Dicer null mesenchymal stem cells (KO-MSCs) as a model system for delineating gene networks regulated by miRNAs (Ravi et al. 2012). miRNA loss through Dicer deletion results in dramatic upregulation of known oncofetal let-7 targets, such as Hmga2 and Imp1-3, that are highly expressed in both embryos and tumors. Beyond loss of miRNA mediated post-transcriptional repression, these genes exhibit evidence of transcriptional upregulation (Gurtan et al. 2013; Gosline et al. 2016).

To investigate the reversibility of previously observed post-transcriptional and transcriptional changes resulting from global miRNA loss, we expressed human Dicer (hsDicer) in monoclonal KO-MSCs wherein the endogenous murine Dicer had previously been ablated (Dicer$^{-/-}$ in Fig. 1A). Specifically, we HA-Flag-epitope tagged and cloned wild-type hsDicer into a retroviral construct, infected KO-MSCs and selected with puromycin for a heterogeneous population of cells. Introduction of Dicer was well tolerated with no evidence of cell death due to re-

expression of miRNAs. Protein analysis on these stable lines indicated robust expression of both the tag and hsDicer (Fig. 1B). Northern blots on two monoclonal cell lines (independently isolated from the stable heterogeneous populations) revealed abundant expression of highly expressed miRNAs including let-7c and miR-22 in the parental, wild-type MSCs expressing an empty vector (WT Vec). These miRNAs are undetectable in the corresponding KO-MSCs (KO Vec in Fig. 1B) but stable expression of wild-type hsDicer resulted in levels of mature miRNAs similar to the parental wild-type (Fig. 1B). Thus, hsDicer was stably expressed and functional in processing murine pre-miRNAs to mature miRNAs.

To globally profile the miRNAs expressed across these cell lines, we performed sequencing of the small RNA population (small RNA-seq) on each independent clone. We observed strong reproducibility between the clones (Fig. 1A, 1B). On average, mature miRNAs comprised 20% of all reads in wild-type cells but only 1% in the Dicer nulls, demonstrating depletion of mature miRNAs globally (Supplemental Fig. 1C). Consistent with the northern blot results, expression of wild-type hsDicer restored mature miRNAs to levels comparable to wild-type cells (18% of all reads; Supplemental Fig. 1B). Greater than 250 miRNAs were detectable in both wild-type clones at a threshold of at least 5 reads per million. The 10 most abundant miRNAs included miR-143, miR-100, miR-22 and multiple members of the let-7 family. Upon collapsing by seeds, the let-7 family was most abundant, constituting on average 37% of reads mapping to miRNAs.
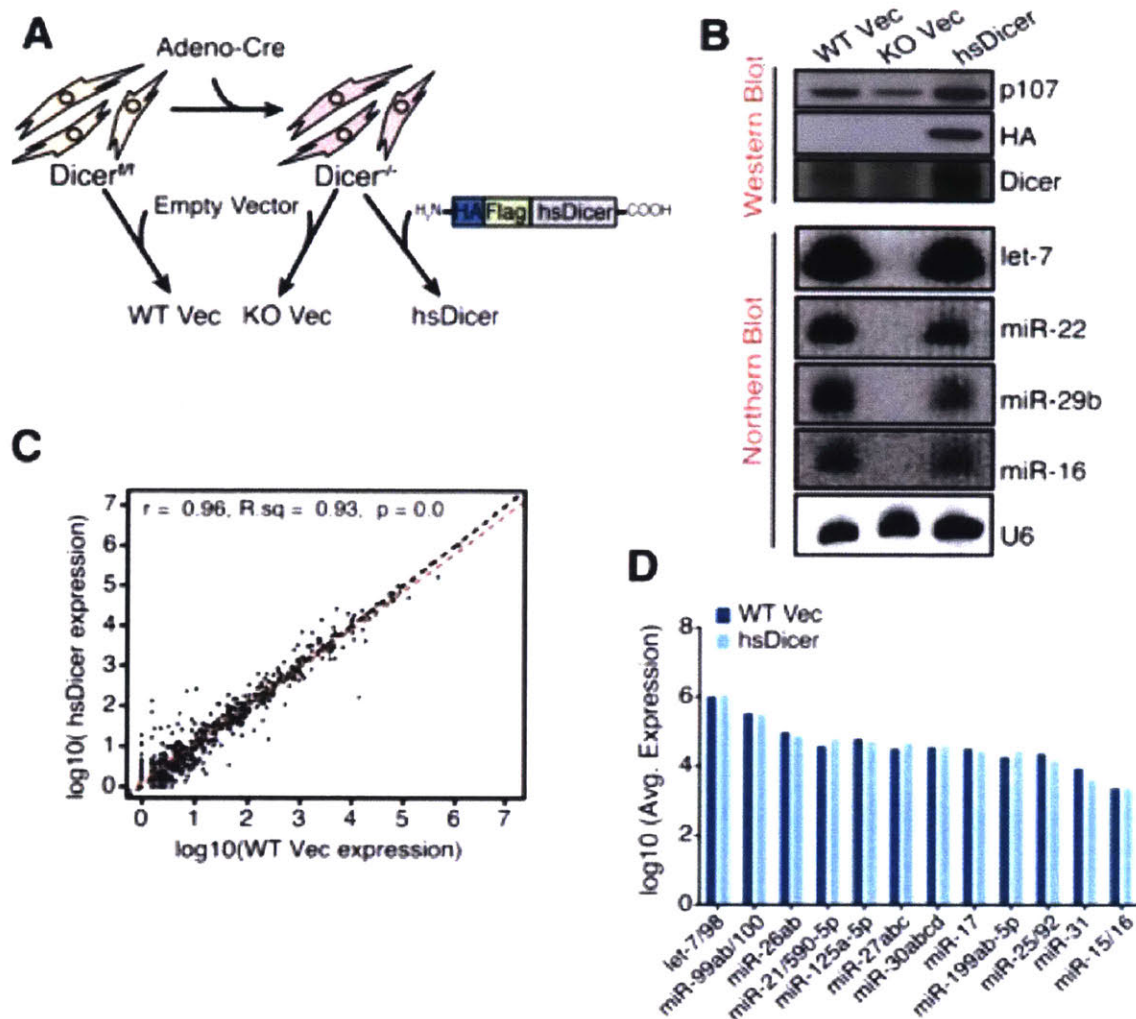
Fig. 1. hsDicer rescues murine miRNA expression. (A) Schematized experimental design. Wild-type human Dicer (hsDicer) was HA-Flag epitope tagged at the N-terminus and cloned into pMMP-Puro, a puromycin resistant vector. SV40-Large T immortalized Dicer$^{f/f}$ or Dicer$^{-/-}$ cells were infected with virus encoding the empty vector or HA-Flag-hsDicer, drug selected and passaged prior to experiments. **(B)** Upper panel: western blot analysis across the heterogeneous population of resistant cells. P107 is shown as a loading control. Lower panel: representative northern blot analysis across a set of monoclonal lines isolated by low density seeding from the heterogeneous population of resistant cells. U6 is shown as a loading control. **(C)** Global miRNA profile comparison between WT Vec and hsDicer. miRNAs collapsed by TargetScan family. The indicated values are normalized average expression counts across both replicates within each condition (2 wild-type clones, 2 hsDicer clones). Black line: y=x; red line: line of best fit. **(D)** Normalized expression counts for the 12 TargetScan miRNA families that show significant Ago2 crosslinking at their target sites over background (Bosson et al. 2014). Adjusted p-value =1 for all comparisons.

We then directly compared the wild-type and hsDicer reconstituted miRNA libraries using DE-Seq (Anders and Huber 2010). The expression of miRNA families was largely restored to wild-type levels (Fig. 1C). The variation between wild type and hsDicer observed at lowly/ non-expressed miRNAs was not statistically significant, and globally, only 3 small RNAs could be assessed as significantly different. These 3 are all low abundance and together constitute less than 0.5% of the miRNA population (Supplemental Fig. 1D). Notably, let-7, the most abundant seed family was almost identically expressed between wild type and hsDicer (Fig. 1B, D). To consider other dominant seeds, we evaluated the 12 most active miRNA seeds in MSCs, as defined by statistically significant, genome wide Ago2 dependent crosslinking (iCLIP) at their target sites (Bosson et al. 2014). Importantly, these 12 families were all completely restored at the mature miRNA level with hsDicer reconstitution (Fig. 1D; p=n.s.). Together, these results show that stable wild-type hsDicer expression in a $Dicer^{-/-}$ background returns the mature miRNA expression profile to the parental wild-type ($Dicer^{f/f}$) state.

hsDicer expression completely reconstitutes endogenous miRNA activity

To determine if miRNAs generated by hsDicer mediate repression of mRNA targets, we used a previously reported two-color fluorescent reporter (Mukherji et al. 2011; Wilusz et al. 2012) to compare post-transcriptional miRNA activity between parental wild-type and hsDicer. A bi-directional promoter drives the expression of mCherry containing miRNA binding sites in its 3'UTR, and non-targeted eYFP that serves as an internal control (Fig. 2A). We measured the

repression mediated by miRNAs representing a range of expressions. In WT Vec cells, compared to a non-targeted mCherry control ('0x'), the addition of miRNA binding sites resulted in an average of 1.3 fold repression for the lowly expressed miR-24 and 5.3 fold repression for the more abundant let-7 (Fig. 2A). Remarkably, we observed 1.2 and 5.5 fold repression for these miRNAs respectively in wild-type hsDicer expressing cells (Fig. 2A). These results are consistent across the second population of clones (data not shown). Hence, the levels of repression across all tested miRNAs are not statistically different between WT Vec and hsDicer expressing cells (p-value=0.71; Fig. 2A). Furthermore, there was negligible repression of the reporters in KO Vec cells confirming that these cells are indeed miRNA deficient (Fig. 2A).

We next examined whether the above results could be extended to endogenous targets. To do so, we carried out mRNA sequencing (mRNA-seq) on polyA-selected total RNA isolated from both clones of each cell type. We detected just under 13,000 genes expressed across all conditions. A pairwise comparison of the transcriptomes of WT Vec and KO Vec identified a population of 1,524 differentially expressed genes that exhibited a median upregulation of ~3-fold and downregulation of ~3.6-fold (Supplemental Fig. 2A). The degree of change occurring with Dicer loss showed a strong and significant correlation to previously reported expression changes in MSCs (Supplemental Fig. 2B) (Gurtan et al. 2013). Gene set enrichment analysis (Subramanian et al. 2005) confirmed that the most upregulated genes were enriched for targets of 'active' miRNAs (Supplemental Fig. 2C).
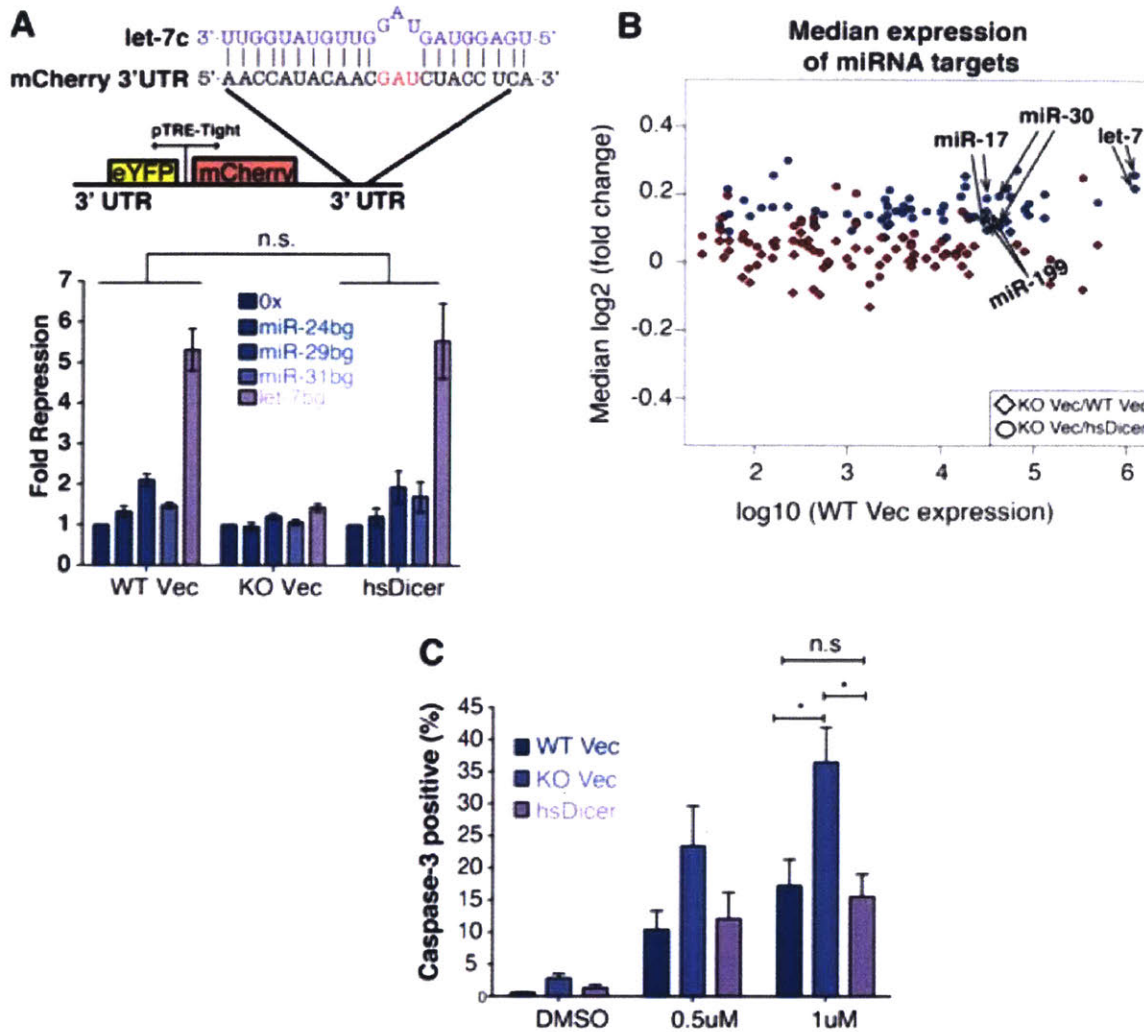
Fig. 2. hsDicer expression recovers miRNA activity. (A) Upper panel: schematic of dual color reporter used to assay miRNA repression. The 3'UTR of mCherry contains imperfect complementary sites to the miRNA of interest. The let-7c sequence is shown in purple. Lower panel: After transfecting with the reporter, flow cytometry was used to measure mCherry and eYFP levels. For each cell type, fold repression is relative to the non-targeted 0x reporter. Data shown is the mean and standard deviation of three independent experiments. P-values calculated by paired Students t-test. (B) Scatter plot of the median change in expression of miRNA targets relative to control genes matched for 3' UTR length, GC content, and expression. Each point represents conserved targets of a single TargetScan miRNA family. miRNA expression is based on the average WT Vec expression reported in this study. Only expressed miRNAs are shown. Blue indicates significant change (Wilcoxon rank sum p≤0.05), while red indicates not significant. (C) Doxorubicin induced apoptosis measured by caspase-3 cleavage. Mean ±S.E.M. of three independent experiments indicated. P-values calculated by unpaired Students t-test. (*) P< 0.05. (n.s) not significant.

188

To perform a more miRNA-focused comparison of cell based activity on endogenous targets, we derived empirical cumulative distributive functions for TargetScan miRNA families expressed in WT Vec cells. Upon loss of Dicer, conserved targets of most miRNAs were upregulated, with let-7 showing the most significant change (Fig. 2B; miRNAs whose targets significantly change statistically are shown in blue). The significant upregulation indicates strong silencing of targets in the wild-type state. HsDicer expression in the Dicer knockout background resulted in significant repression of conserved targets (Fig. 2B) indicating that the restored miRNAs were highly active. Notably, targets of more abundant and 'active' miRNA seeds (such as let-7, miR-17, miR-199 and miR-30) that showed significant upregulation upon Dicer loss, exhibited comparable median downregulation with hsDicer. Hence, hsDicer comprehensively restores miRNA activity on endogenous targets to a degree comparable to the parental wild-type state.

Finally, as miRNAs are known to buffer changes due to environmental stress (Leung and Sharp 2010), we tested for Dicer dependent changes in stress response by treating the cells with the DNA damaging agent doxorubicin. KO Vec cells were highly sensitive to this drug compared to WT Vec ($p < 0.05$; Fig. 2C). Expression of hsDicer in KO cells completely rescued the doxorubicin-induced cell death at all concentrations tested, such that the response between WT Vec and hsDicer was indistinguishable ($p = 0.77$; Fig. 2C). Together, these results demonstrate that on a global scale, hsDicer expression completely restores

miRNA post-transcriptional activity, which in turn is sufficient to rescue miRNA loss-induced stress defects.

let-7 oncofetal targets remain activated after restoration of Dicer

Next, we examined the expression of oncofetal targets reported previously to contribute to the let-7 signature in Dicer KO MSCs (Gurtan et al. 2013). Since our data indicates full reconstitution of miRNA expression and activity, we hypothesized that the expression of these targets would revert to levels observed in WT Vec. Quantitative-PCR analysis showed a dramatic upregulation of Imp family proteins upon Dicer loss, from 45-fold for Imp2 to upwards of 200-fold for Imp1 and Imp3 (Fig. 3A). Hmga2, an additional let-7 oncofetal target that is often correlated with Imp2-3 in expression, was upregulated to a lesser degree (~17-fold, Fig. 3A). These fold changes are consistent with the RNA-seq results (Supplemental Fig. 3A).

Surprisingly, restoration of miRNA activity with introduction of hsDicer only marginally downregulated Imp mRNA expression relative to KO Vec levels (Fig. 3A). Imp mRNAs that were upregulated 45 to 200-fold, were only suppressed 2 to 4-fold with restoration of miRNA activity. Less dramatically, Hmga2 mRNA which increased 17-fold with loss of miRNAs was suppressed by 4-fold upon restoration with hsDicer. At the protein level, western blot analysis mirrored the mRNA results. While Hmga2 was very lowly expressed and the Imp family was
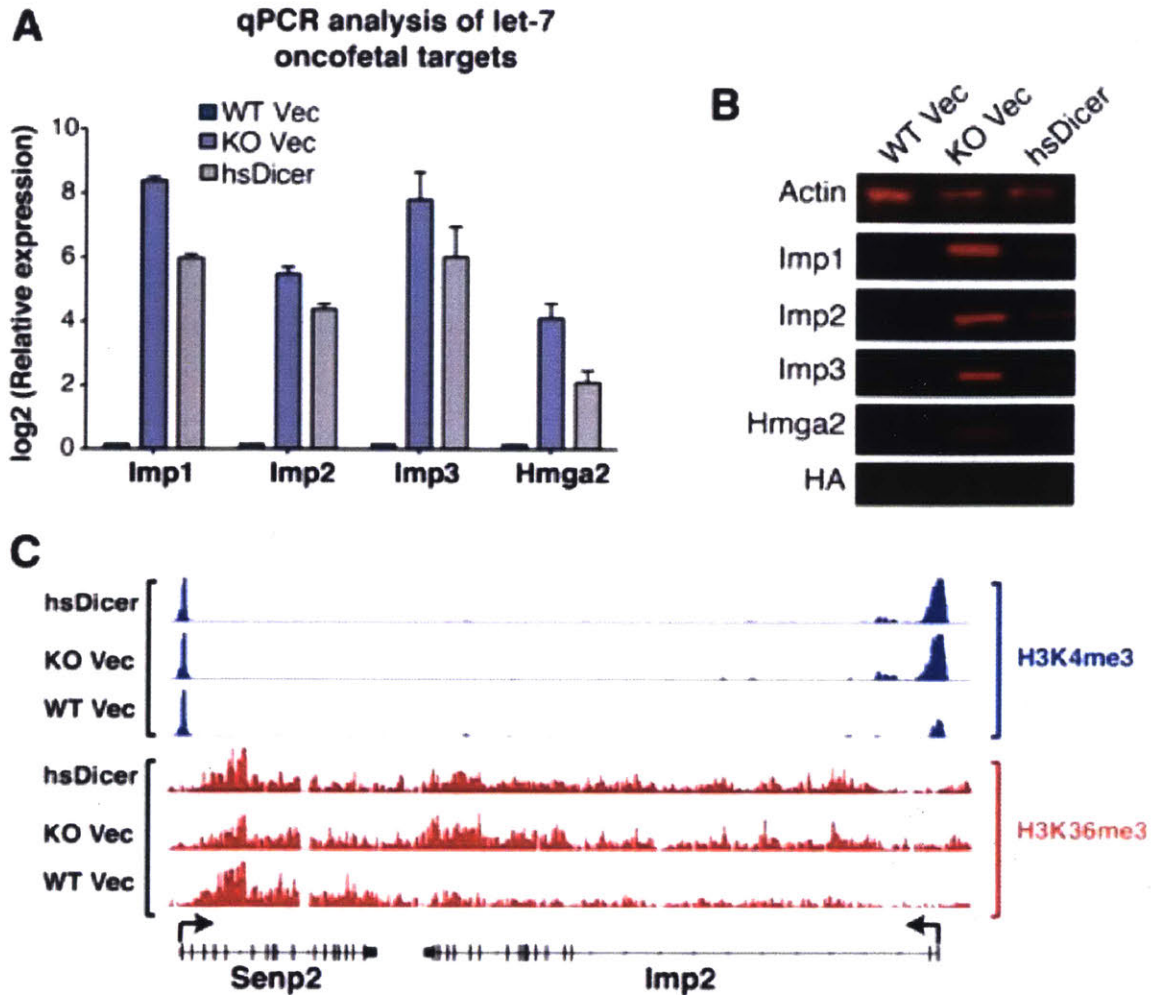
Fig. 3. Despite miRNA rescue, let-7 oncofetal targets remain activated. (A) qPCR analysis of four let-7 regulated oncofetal genes. Note the logarithmic scale. For each gene, expression is relative to WT Vec levels. Data are plotted as the mean ± s.e.m of four or more independent experiments. **(B)** Western blot analysis of the oncofetal gene set. HA-tag is used to indicate hsDicer expression. Actin is used as a loading control. **(C)** Normalized read counts for H3K4me3 and H3K36me3 marks at the Imp2 locus across the three cell lines. Within each chromatin mark, all conditions are set to the same scale. Flanking genes are shown as controls. Arrowheads indicate transcription start sites (TSS).

undetectable in WT Vec, all four genes showed robust protein expression in KO Vec cells (Fig. 3B). Subsequent to miRNA rescue with hsDicer, Imp1 Imp2 and Imp3 all showed a partial reduction consistent with the typical post-transcriptional repression by miRNAs. Importantly, even after restoration of miRNA activity, all three genes were highly expressed compared to original wild-type cells (Fig. 3B). In contrast, Hmga2 protein expression was more significantly repressed than the Imp family. This is consistent with the mRNA results, where in hsDicer cells, Hmga2 was only ~4-fold more highly expressed than in WT Vec (Fig. 3A; Supplemental Fig. 3A).

The inability of complete miRNA reconstitution to revert Imp gene expression suggests that somatic Dicer loss activates a self-sustaining, 'oncofetal dependent' state. This possibility is consistent with our previous studies showing major transcriptional changes occurring at these loci upon Dicer loss (Gurtan et al. 2013; Gosline et al. 2016). Hence, we evaluated the transcriptional state at these loci through chromatin immunoprecipitation followed by massively parallel DNA sequencing (ChIP-Seq) on H3K4me3 and H3K36me3 modifications. Upon inspection of the Imp2 locus, we observed increased H3K4me3 near the transcription start site (TSS) and H3K36me3 towards the 3' end of the gene body in KO Vec compared to WT Vec cells (Fig. 3C). Intriguingly, add-back of hsDicer effected negligible change from the KO condition, indicating that there was 'miRNA resistant' transcriptional activity across this locus in hsDicer cells compared to the WT Vec. As a control comparison, the flanking gene Senp2 was indistinguishable across all three conditions for both

modifications (Fig. 3C). These results indicate a targeted, miRNA-resistant transcriptional activation of Imp2 upon Dicer loss.

This trend was observed for the other Imp family members but not for Hmga2, which showed no evidence of Dicer-dependent transcriptional changes (Supplemental Fig. 3B-D). As mentioned above, the degree of upregulation for Hmga2 at the mRNA level was substantially less than that of the other oncofetal targets. Given that the Hmga2 3'UTR has seven let-7 seed matches (of which six are conserved) we speculate that the 17-fold upregulation is primarily due to multiplicative or cooperative post-transcriptional miRNA effects (Doench et al. 2003; Grimson et al. 2007; Saetrom et al. 2007). For the other oncofetal targets however, we propose that transcriptional changes upon Dicer loss account for most of their upregulation (beyond direct miRNA effects on mRNA stability) and this change is not reversed upon restoration of miRNA activity.

Transcriptomic profiling reveals a larger, miRNA-resistant gene expression signature

We tested whether the irreversible behavior of the oncofetal let-7 targets Imp1-3 was part of a more global phenomenon. To do this, we used an unsupervised blind source separation approach employing Independent Component Analysis (ICA) (Hyvarinen and Oja 2000) to identify distinct gene expression signatures. We have previously used this approach to elucidate cancer-associated gene signatures in lung adenocarcinoma (Li et al. 2015; Dimitrova et al. 2016; Papagiannakopoulos et al. 2016). ICA provides the ability to detect statistically independent gene expression signatures within the context

of the entire dataset (i.e. across all samples), as opposed to performing multiple, pair-wise differential expression analyses. When applied to the mRNA-Seq data across all cell lines, ICA identified two significant and biologically relevant signatures (p= 0.01; Fig. 4A): signature 1 reflects a pattern of gene upregulation in both KO Vec and hsDicer conditions relative to parental wild-type, while signature 2 indicates a pattern of higher gene expression in miRNA deficient conditions relative to miRNA competent ones (Fig. 4A).

qPCR analysis of a sampling of signature 1 correlated genes (z-score >3) confirmed that hsDicer reconstitution did not revert their expression to wild-type levels (Fig. 4B). In contrast, signature 2 related genes showed near complete rescue to wild-type expression levels (Fig. 4C). As such, hereafter signature 1 and signature 2 will be referred to as miRNA-resistant and miRNA-sensitive respectively. Consistent with the observations for the Imp1-3 family, as a group, the miRNA-resistant genes were dramatically upregulated upon Dicer loss, significantly more so than miRNA-sensitive genes (Fig. 4D). Upon quantification of the ChIP-Seq signal at promoters, we observed that the miRNA-resistant genes exhibited significantly higher gains in H3K4me3 than reversible genes (Fig. 4E; compare 2nd and 3rd box plots) when Dicer was originally lost in these cells. With hsDicer addback, there was no significant reduction in H3K4me3 at the miRNA-resistant genes (compare 2nd to 5th box plots), while H3K4me3 at the miRNA-sensitive genes were significantly reduced (compare 3rd to 6th box plots) (Fig. 4E). A similar trend was observed for H3K36me3 in the body of genes (Fig. 4F).
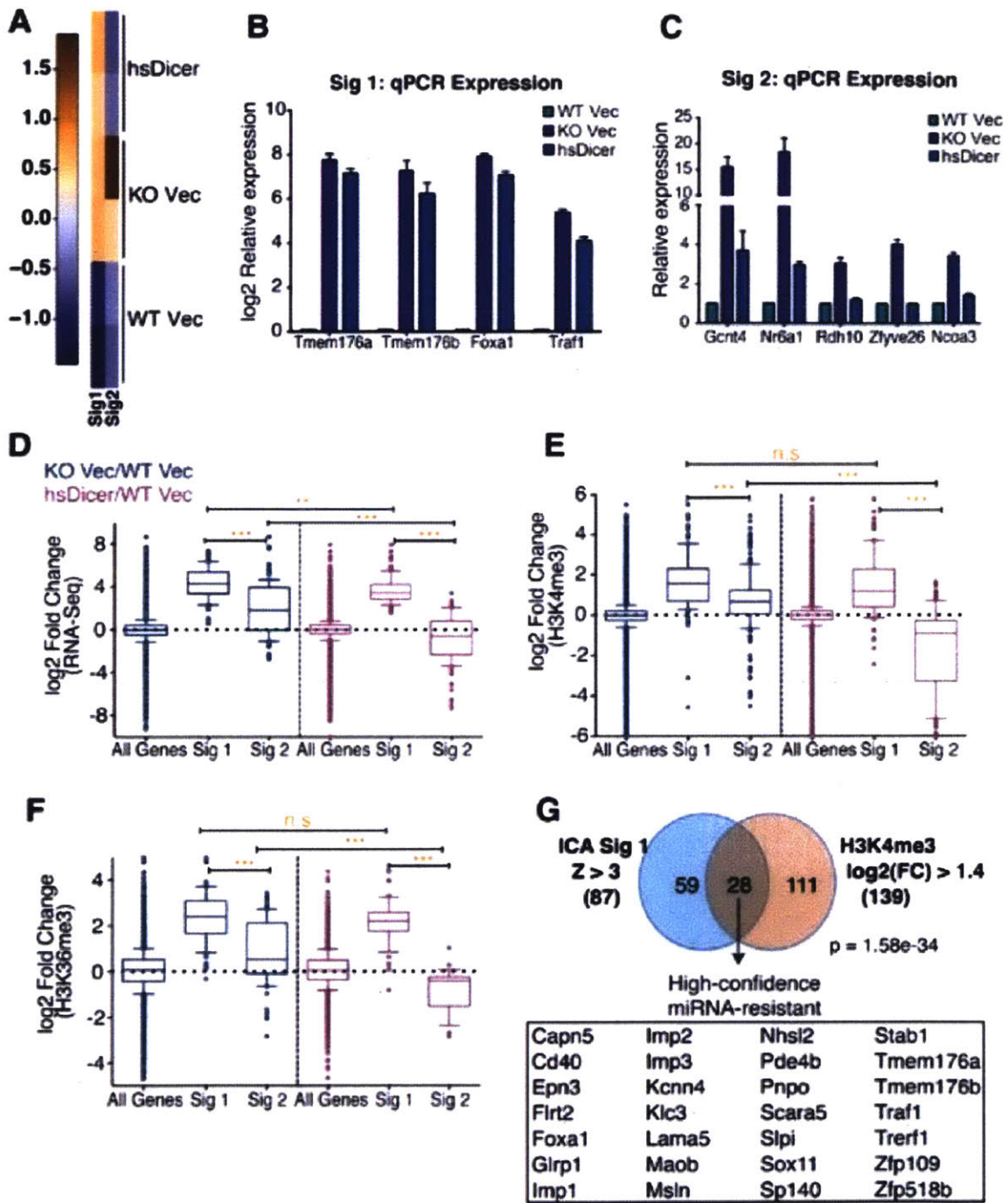
Fig. 4. Identification of a high-confidence miRNA-resistant gene signature.
(A) Heatmap depicting two independent, statistically significant gene signatures detected in the RNA-Seq expression dataset using ICA. Sig. 1 represents an expression pattern of genes up-regulated in KO Vec and hsDicer conditions relative to WT Vec. Sig. 2 represents genes that are down-regulated to WT Vec levels upon introduction of hsDicer. **(B)** qPCR validation of 'irreversible' or 'miRNA-resistant' genes identified from signature 1 (Sig 1) and **(C)** 'reversible' or 'miRNA-sensitive' genes identified from signature 2 (Sig 2). Results are plotted relative to WT Vec levels. Bars represent mean ± S.E.M. of at least 3 independent experiments. **(D)** Box and whisker plots of normalized RNA-Seq expression fold changes for all expressed genes (n = 12834), Sig 1 correlated genes (n = 87) and Sig 2 correlated genes (n = 112). **(E-F)** Same as in (D) but for fold changes in normalized counts for gene associated H3K4me3 peaks **(E)** and H3K36me peaks **(F)**. Whiskers represent 10th-90th percentile and all other points are shown as individual dots. The p-values were calculated by Mann Whitney-U test. (**) $P<0.01$; (***) $P<0.0001$; (n.s) not significant. **(G)** Overlap of Signature 1 correlated genes and genes with a log2 fold change of at least 1.4 in H3K4me3 peak counts near their transcription start sites. The H3K4me3 criteria is satisfied by both KO Vec /WT Vec and hsDicer/WT Vec comparisons. High-confidence miRNA-resitant genes are listed. Hypergeometric test's p-value indicated.

Thus, ICA uncovered a broader miRNA-resistant gene set beyond Imp1-3, which exhibits dramatic upregulation upon Dicer loss followed by incomplete repression to wild-type expression. Our results indicate that increased transcription (measured through ChIP-Seq) strongly correlates to this resistance (Fig 4D-F). Therefore, we defined a 'high-confidence' set of miRNA-resistant genes due to changes in transcription using our combined datasets. We overlapped the set of genes defined by ICA (Signature 1; z-score >3; n=87) with genes that show a log2 change of at least 1.4-fold in their associated H3K4me3 peaks in both KO Vec and hsDicer compared to WT Vec (KO Vec/WT Vec ≥ 1.4 and hsDicer/WT Vec ≥ 1.4; n = 139). This analysis identified a set of 28 genes that we designate as 'high-confidence miRNA-resistant' genes (Fig. 4G). In contrast, an overlap using miRNA-sensitive genes (Signature 2; z-score >3; n=112) with the same H3K4me3 gene set was insignificant and only contained 3 genes (p=0.12; data not shown). These results indicate that transcriptional activation is a feature of miRNA-resistant genes.

miRNA restoration confers a stably inherited transformed phenotype

The high confidence miRNA-resistant signature includes let-7 regulated oncogenes that are elevated dramatically in various cancers and function to promote tumor growth, drug-resistance and metastasis (Tessier et al. 2004; Boyerinas et al. 2008; Samanta et al. 2013). Thus, we examined whether re-expression of hsDicer and hence miRNAs in Dicer KO cells resulted in a tumorigenic phenotype. In a soft agar assay, hsDicer cells exhibited significant growth potential not observed for WT Vec and KO Vec cells (Fig. 5A). These

results suggest that hsDicer cells acquired properties of anchorage-independent growth and anoikis resistance, which could possibly contribute to tumorigenicity *in vivo*. To directly test this possibility, we performed subcutaneous injections of WT Vec, KO Vec and hsDicer cells into the flank of immune-compromised mice. Strikingly, we observed 100% incidence of tumor formation with hsDicer cells, but none for WT Vec or KO Vec (p= 3e-07; Fig. 5B). Histopathological analysis of the tumors indicated undifferentiated sarcomas, which is consistent with the mesenchymal origin of these cells (Fig. 5C).
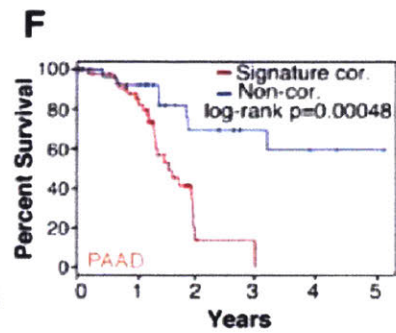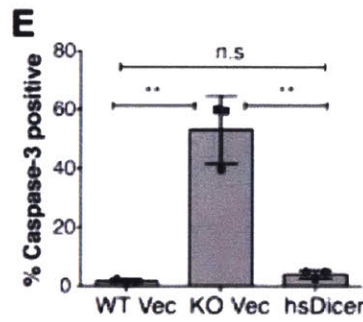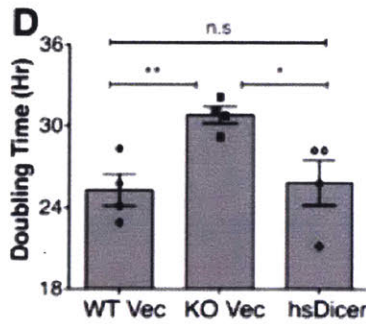
To confirm that the tumors were of hsDicer cell origin, we generated cell lines from five tumors and confirmed absence of endogenous murine Dicer via PCR genotyping (Supplemental Fig. 4A). Furthermore, the tumor cells were all puromycin resistant due to the vector containing hsDicer (data not shown) and strongly expressed let-7 and other mature miRNAs, thereby confirming the retention of hsDicer *in vivo* (Supplemental Fig. 4B). Importantly, a sampling of a subset of miRNA-resistant genes indicates that they were highly expressed in all the tumor cell lines (Supplemental Fig. 4C-D). Thus, we conclude that the observed tumors developed from hsDicer cells.

Our results indicate that only hsDicer cells are tumorigenic despite the fact that KO Vec cells also have high expression of Imp1-3 and other high confidence miRNA-resistant genes. To evaluate whether miRNA deficiency adversely impacted growth of KO Vec cells, we measured doubling times and apoptosis for each genotype. KO Vec cells exhibited a significantly slower doubling, while

A — # of colonies / field; WT Vec, KO Vec, hsDicer; *** , **
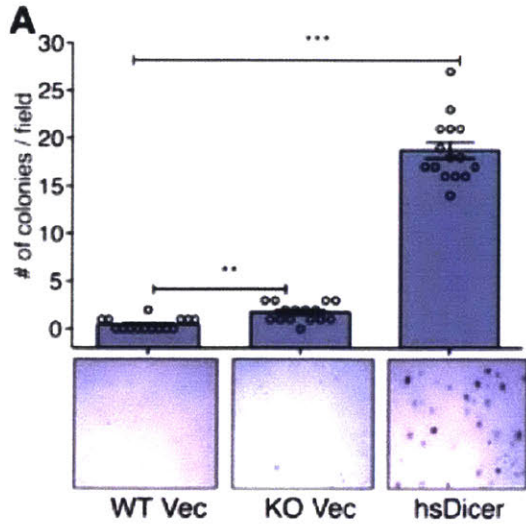
B

Cell line	WT Vec	KO Vec	hsDicer
Tumor Incidence # with tumors/ total	0/15	0/10	15/15

C

D — Doubling Time (Hr); WT Vec, KO Vec, hsDicer; n.s, **, *

E — % Caspase-3 positive; WT Vec, KO Vec, hsDicer; n.s, **, **

F — Percent Survival vs Years; Signature cor.; Non-cor.; log-rank p=0.00048; PAAD

G

Characteristic	Univariate		Multivariable		
	HR (95% CI)	p	HR (95% CI)	p	p_interaction
High-confidence miRNA-resistant Sig.	1.57(1.19-2.08)	0.0013	1.51(1.09-2.11)	0.01379	-
Gender (Male vs Female)	0.88(0.57-1.35)	0.56209	0.96(0.58-1.57)	0.86122	-
Age (Years)	1.03(1.01-1.05)	0.00722	1.04(1.01-1.06)	0.00296	0.05533
T score (T3 vs T1/T2)	2.52(1.25-5.08)	0.01005	1.22(0.57-2.61)	0.60888	-
N score (N1 vs N0)	2.16(1.24-3.75)	0.00647	0.93(0.45-1.91)	0.84298	-
Completeness of resection (R1 vs R0)	1.95(1.23-3.09)	0.00461	1.63(0.95-2.80)	0.07796	-
Number of lymph nodes	1.06(1.01-1.12)	0.02096	1.10(1.01-1.20)	0.03794	0.1359

HR = Hazard ratio; CI = Confidence Interval; T score = Primary tumor size/invasiveness; N score = Lymph node metastasis;

$p_{interaction}$ = p-value of interaction between significant covariates (model comparison; likelihood ratio test)

Fig. 5. miRNA restoration through hsDicer expression transforms MSCs.
(A) Colony formation of MSCs after ~15 days of growth in agarose containing media. Representative 4x brightfield images for each genotype are shown. Colonies were counted by eye in 5 random fields for 3 independent experiments. Data are plotted as the mean ± S.E.M. Student's t-test p-value indicated. **(B)** Frequency of tumor formation by the three parental MSC cell lines 8 weeks after injection of 10^5 cells into the flank of immune-compromised mice. **(C)** Representative hematoxylin and eosin staining of an hsDicer tumor section at 20x magnification. Scale bar represents 200um. **(D)** Proliferation assay indicating the mean ± S.E.M. of 2 experiments with 2 replicates each. **(E)** Apoptosis assayed by capase-3 cleavage. Mean and standard deviation are plotted. P-values calculated by Student's t-test. (*) $P<0.05$; (**) $P<0.01$; (***) $P<0.0001$; (n.s) not significant. **(F)** Kaplan-Meier survival analysis of TCGA pancreatic adenocarcinoma (PAAD) patients stratified by their correlation score with the high-confidence miRNA-resistant signature ($|z| > 0.5$ extremes of score distribution, n=46 most correlated, n=28 least correlated). Patients with most correlated gene expression scores (red) exhibit significantly reduced survival times compared to least-correlated patients (blue). See supplementary methods for details. **(G)** Results of univariate and multivariable Cox proportional hazards model on overall survival in the TCGA pancreatic adenocarcinoma cohort (all patients). The high-confidence miRNA-resistant signature is found to be independently prognostic within the cohort of TCGA PAAD patients.

hsDicer completely rescued this phenotype (Fig. 5D). In addition, KO Vec cells exhibited an elevated basal apoptosis at high culture density compared to WT Vec and hsDicer (Fig. 5E). *In vivo*, these combined defects are probably deleterious to KO Vec cells and dominate over any oncogenic potential.

The high-confidence miRNA-resistant gene set is a pan-cancer signature

To assess the relevance of high expression levels of this high-confidence miRNA-resistant network that is enriched in oncofetal genes, we first queried the cBioPortal database (Cerami et al. 2012). A cross-cancer analysis revealed a high degree of alterations in this gene set in breast, lung, bladder, and pancreatic cancers (Supplemental Fig. 4E). Notably, these alterations were predominantly amplifications, suggesting oncogenic activity that is important for human malignancy. To address this more directly, we examined whether the high-confidence miRNA-resistant gene signature could provide prognostic information relevant to human cancer. Using gene expression profiles and clinical information from the Cancer Genome Atlas' (TCGA; cancergenome.nih.gov) pancreatic adenocarcinoma (PAAD) cohort, we found the signature to be strongly correlated with clinical outcome: high expression was significantly associated with worse survival in Kaplan-Meier analysis comparing patients at the extremes of the signature correlation score distribution ($|z| > 0.5$; log-rank $P < 0.0005$; Fig. 5F). Furthermore, increasing signature correlation score was also associated with poor prognosis across all patients in the cohort (Hazard ratio = 1.51) in a Cox proportional hazards model while controlling for various other clinical covariates (Fig. 5G), suggesting that this signature is independently prognostic in PAAD. To

assess the pan-cancer relevance of this signature, we investigated additional cancer types within the TCGA dataset. Strikingly, this signature is also significantly correlated with reduced survival across other solid cancers (lung, kidney and glioblastoma/low-grade glioma; Supplemental Fig. 4F), indicating that upregulation of this high-confidence miRNA-resistant gene set is a more general feature of human cancers.

Co-targeted disruption of Imp1-3 adversely impacts hsDicer tumorigenicity

The miRNA-resistant signature identified in this study is almost certainly oncogenic in human cancer. However, it is unclear whether a core subset of these genes is primarily responsible for the tumorigenicity of the hsDicer cells. To investigate this, we focused on the Imp family because Imp1-3 are not expressed in most adult tissues, and are often *de novo* activated in cancers. To test whether loss of any Imp member would negatively affect tumorigenesis, we used the CRISPR-Cas9 system to generate individual knockouts of Imp1-3 in hsDicer cells and propagated two independent clones of each knockout cell line. Western blot analysis indicated complete loss of the individual Imp proteins in the targeted clones and the results of genomic DNA sequencing were consistent with a loss of the reading frame (Fig. 6A; data not shown). However, each Imp knockout retained robust expression of the remaining two family members suggesting a lack of co-regulation. We then evaluated the potential of the Imp knockouts to grow *in vitro* via the soft agar assay. Individual loss of Imp1-3 in hsDicer cells did not confer any meaningful reduction in colony formation (Fig. 6B; Supplemental Fig. 5A). As an alternative measure of tumorigenic potential, we used the Growth

in Low Attachment assay (GILA) which has been shown to capture aspects of transformation that could be missed by soft agar growth (Rotem et al. 2015). We validated that GILA recapitulated the soft agar results; only hsDicer cells were able to grow over a 7-day time course in GILA (Supplemental Fig. 5B). In contrast, WT Vec and KO Vec cells underwent anoikis despite exhibiting strong growth under normal adherent conditions (Supplemental Fig. 5B and data not shown). Consistent with the soft agar results, individual knockouts of Imp1-3 did not dramatically reduce cellular growth via GILA (Fig. 6C). Hence, our data indicate that loss of any one Imp family member on the background of the endogenous expression of the other two Imp proteins is not sufficient to abrogate the tumorigenic properties of hsDicer cells.

Given this result, we next explored the effects of simultaneous loss of all three Imp proteins in the hsDicer condition. We isolated two independent clones of these triple knockout cells, designated $Imp^{\Delta 3}$, and confirmed loss of function by western blot and genomic DNA sequencing (Fig. 6D and Supplemental Fig. 5C). These clones exhibited a significant reduction in the number of colonies formed in soft agar as compared to the hsDicer cells from which they were derived (Fig. 6E). Similarly, the growth of both $Imp^{\Delta 3}$ clones was severely compromised in the low adherent conditions of the GILA assay (Fig. 6F). To extend these results *in vivo*, we performed subcutaneous tumor formation assays in immune-compromised mice, injecting the parental hsDicer cells and $Imp^{\Delta 3}$ cells on the left and right flanks respectively of each mouse. hsDicer derived tumors first appeared after 28 days compared to 35 days for $Imp^{\Delta 3}$ (Fig. 6G). $Imp^{\Delta 3}$ cells
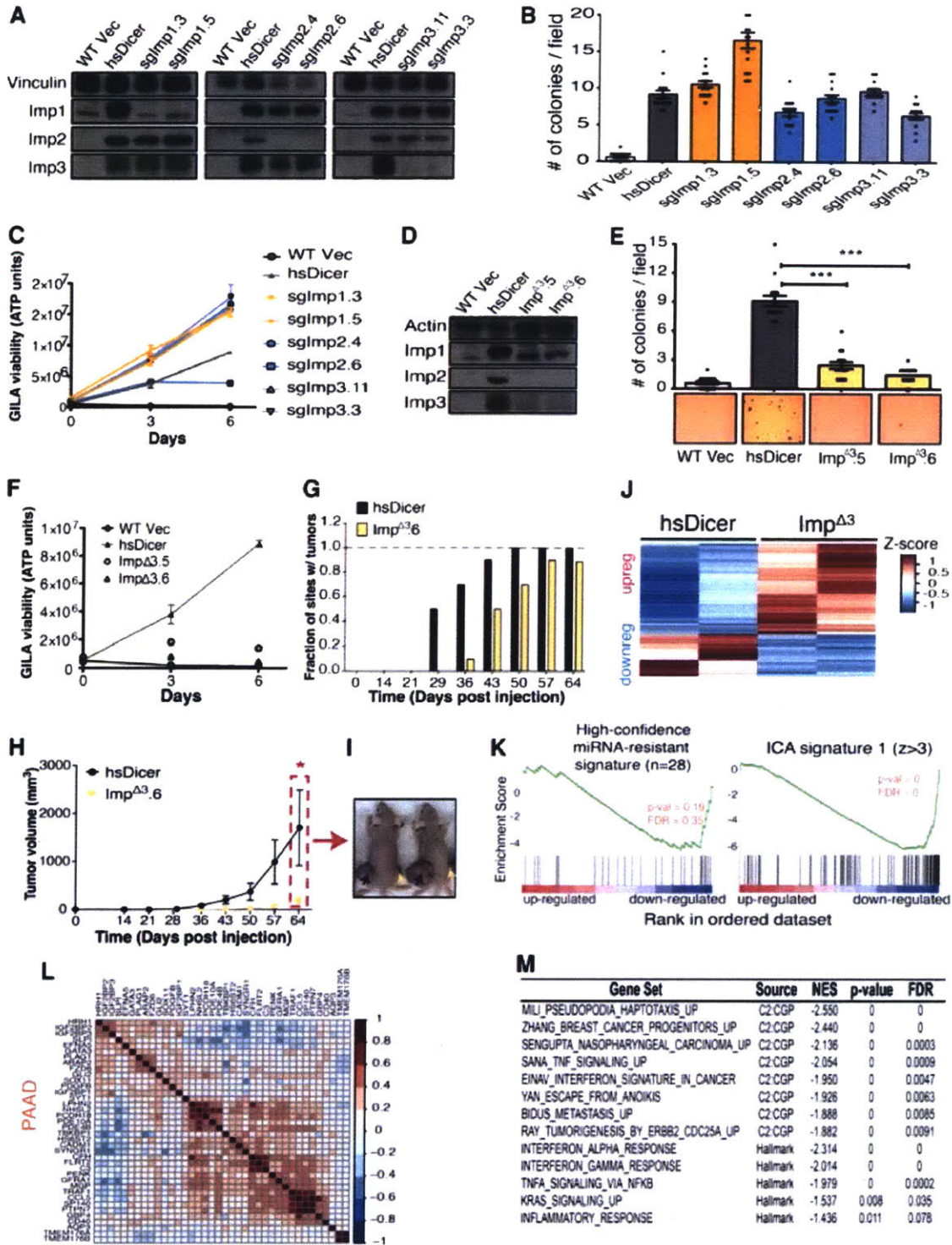
Fig. 6. Combined loss of Imp1-3 greatly impairs tumorigenicity *in vivo.* **(A)** Western blot of lysates derived from CRISPR-Cas9 generated Imp knockout cells. Two independent clones of each knockout cell type are indicated. Left panel: Imp1 knockouts (the upper band is the specific band), middle panel: Imp2 knockouts, right panel: Imp3 knockouts. Vinculin is used as a loading control. **(B)** Quantification of soft agar colony counts from the cell lines depicted in A. Data are plotted as the mean ± s.e.m. **(C)** Growth time course of cells in GILA assay. The ATP yield at each time point is a measure of cell viability. Data are plotted as the mean ± standard deviation (n=3). **(D)** Western blot of whole cell lysates from two independent clones of CRISPR-Cas9 generated Imp1-3 triple knockout ($Imp^{\Delta 3}$) cells. Actin is used as a loading control. **(E)** Quantification of soft agar colony counts from the cell lines depicted in D. Representative brightfield soft agar images for each genotype is shown at 4x magnification. **(F)** Same as in C but for $Imp^{\Delta 3}$ cells. **(G)** Cumulative fraction of injected sites where tumors formed as a function of time. Each mouse from this cohort was subcutaneously injected with two cell types: hsDicer cells on the left flank and $Imp^{\Delta 3}$ cells on the right flank (n =10). **(H)** Kinetics of subcutaneous tumor growth for hsDicer and $Imp^{\Delta 3}$ derived tumors. (*) p < 0.05 from paired Wilcoxon signed rank test. **(I)** Representative image depicting gross subcutaneous tumor sizes at the experimental end point (64 days post-injection). Black and yellow arrowheads indicate site of hsDicer or $Imp^{\Delta 3}$ cell injection respectively **(J)** Heatmap of the 302 genes differentially expressed between hsDicer and $Imp^{\Delta 3}$ cells. For each gene (each row), normalized gene expression is z-score standardized. **(K)** GSEA enrichment plots for the high-confidence miRNA-resistant (left panel) and ICA-derived signature 1 (right panel) signatures. Genes are ranked according to descending ($Imp^{\Delta 3}$/ hsDicer) log2-fold change such that most downregulated genes are skewed to the right of the plot. Nominal p-values and FDR q-values are indicated. **(L)** Heatmap depicting gene-by-gene spearman correlation coefficients in tumor samples of pancreatic adenocarcinoma (PAAD) patients from TCGA. Depicted genes are the leading-edge core genes driving the enrichment of ICA signature 1 in K above. **(M)** Summary table for GSEA results showing enriched gene sets from the MSigDB curated or hallmark gene set collections. Genes downregulated with Imp1-3 loss are significantly enriched for gene sets associated with pseudopodia, cancer, anoikis and inflammation associated signaling.

205

required ~42 days to form tumors at 50% of the sites of injection, while 90% of sites injected with hsDicer cells had formed tumors at that time point (Fig. 6G). To better assess this difference in tumorigenicity, we measured the growth of tumors with time. The Imp$^{\Delta 3}$ tumors that formed grew at a significantly reduced rate, ultimately resulting in dramatically stunted tumors when compared to the parental hsDicer tumors (p<0.05; Fig. 6H, 6I). Histologically, the Imp$^{\Delta 3}$ tumors were undifferentiated sarcomas with pleomorphic nuclei, similar to hsDicer derived tumors (data not shown). Altogether, these data demonstrate that the combined activation of the Imp family is important for the oncogenic switch acquired by Dicer rescued cells and that additional miRNA-resistant genes may contribute to tumorigenesis.

Imp1-3 reinforce expression of miRNA-resistant genes

To understand the mechanism by which combined loss of Imp1-3 effects this reduced tumorigenicity, we used poly(A) selected RNA-Seq to analyze the steady state mRNA levels in Imp$^{\Delta 3}$ and the parental hsDicer cells. Upon loss of Imp1-3 proteins, 302 genes were significantly differentially expressed (fold change ≥ 2, FDR ≤ 0.05). 207 genes were upregulated compared to hsDicer with a median fold change of 2.9, while 95 genes were downregulated ~6-fold on average (Fig. 6J, Supplemental Fig. 5D). First, we queried whether loss of Imp1-3 predominantly affected other members of the high-confidence miRNA-resistant signature. Only 5 genes from this signature were significantly differentially expressed (Supplemental Fig. 5E), and globally, these were not biased to either the highly up- or down- regulated genes (Fig. 6K). Given this result, we extended

the analysis to include the larger set of miRNA-resistant genes from which the high-confidence set had been derived (ICA signature 1, z>3, n=87; Fig. 4G). GSEA revealed that ICA signature 1 associated genes were significantly enriched among genes downregulated by Imp1-3 loss, 12 of which met the significance threshold for differential expression and of these, 11 decreased in expression (exception being Msln) (Fig. 6K; Supplemental Fig. 5F). This indicates that majority of the genes that were strongly upregulated in KO Vec and hsDicer cells, but yet lacked transcriptional activation, were stabilized by expression of Imp1-3.

We then examined the GSEA leading-edge subset, which constitutes the core group of genes accounting for this enrichment and reflects the genes most biologically important for the phenotype of interest (in this case Imp1-3 loss), and noticed a strong contribution from high-confidence miRNA-resistant genes. Consistent with expression changes being specific to Imp1-3 activity, in all, 10 of the 44 genes comprising the leading-edge have been implicated as direct Imp1-3 targets (statistically significant signal in either CLIP or RIP data sets) across studies in pluripotent stem and pancreatic cancer cells (Taniuchi et al. 2014; Conway et al. 2016; Ennajdaoui et al. 2016). This included, but was not limited to components of the high-confidence miRNA-resistant signature that we found to be prognostic in human cancers: Sox11, Traf1, Flrt2. However, the leading-edge gene set was also populated by oncogenes that have not been formally linked with Imp1-3, such as Plag1 and Fzd6. To explore potentially novel relationships, we assessed expression patterns of leading-edge genes in cancer. Slpi, Plag1

and Fzd6 showed strong correlation with Imp1-3 family members in pancreatic tumors (Fig. 6L) and this trend was consistent in kidney and lung tumors as well (Supplemental Fig. 5G-H). This observation implicates Slpi, Plag1 and Fzd6 as genuine downstream effectors of Imp1-3 oncogenic activity in our cell system, and in human cancers. In summary, in hsDicer cells, Imp1-3 primarily sustain a post-transcriptional upregulation of miRNA-resistant genes populated by likely direct Imp targets and other oncogenes.

Finally, we broadened our analysis to investigate externally annotated gene sets from the Molecular Signature Database (MSigDB) (Liberzon et al. 2011). This unbiased approach implicated multiple biological pathways as being perturbed upon Imp1-3 loss. For genes upregulated in Imp$^{\Delta 3}$, few gene sets were enriched, among which a coherent biological theme was not apparent. However, among downregulated genes, there was significant enrichment for transcripts localized to pseudopodia, consistent with previously described roles for Imp1 and Imp3 (Stohr et al. 2012; Taniuchi et al. 2014; Ennajdaoui et al. 2016), as well as genes important for avoiding anoikis, and various signaling and cancer pathways (Fig. 6M). In all, this analysis confirmed that the expression changes downstream of Imp1-3 loss are concordant with genes and biological processes with established importance in cancer.

Discussion

We have demonstrated that loss of Dicer results in upregulation of a subset of miRNA targets that are transcriptionally activated and are not re-

silenced upon reconstitution of miRNA expression. The expression of this miRNA-resistant gene set is reinforced by at least three of its members, Imp1-3, and is self-sustaining. The miRNA-resistant signature was observed in independently derived clones of Dicer restored cells. More important and surprising, these clones were tumorigenic in contrast to their progenitors. Finally, the miRNA-resistant gene set is associated with poor patient prognosis across multiple cancers.

Irreversible upregulation of oncofetal targets of let-7

Previous work has demonstrated that the most dramatic changes in gene expression upon loss of the post-transcriptional activity of miRNAs in MSCs surprisingly occur at the transcriptional level (Gosline et al. 2016). This dramatic transcription activation was documented as an increase in nascent RNAs, as assessed through intronic reads, and as well as a corresponding increase in the H3K4me3 and H3K36me3 chromatin modifications (Gosline et al. 2016). We have now extended these findings by showing that rescue of miRNA expression and activity restored the vast majority of changes that occurred with Dicer loss but failed to suppress the transcriptional activation of a core oncogenic gene set, including Imp1-3. The inability of restored miRNAs to completely suppress this oncogenic signature is surprising and reveals important principles regarding the functional organization of miRNAs in gene-expression networks. For example, the let-7-targeted genes Imp1-3, are highly expressed during early embryonic development, but their expression dramatically declines in a window that coincides with induction of mature let-7. It is unknown whether their repression is

driven solely by let-7-mediated post-transcriptional activity, or whether loss/gain of transcriptional regulators acts on them redundantly to facilitate complete silencing in differentiated tissues (Zhu et al. 2011; Nguyen et al. 2014). Our study demonstrates that despite the fact that sole loss of miRNA activity is sufficient for Imp1-3 transcriptional activation, rescue of this post-transcriptional layer of regulation cannot significantly re-silence these genes. This state is distinct from induction of a pluripotent state characterized by expression of Oct4, Sox2 and Lin-28, which are not expressed in any of these MSC related cells. Thus, loss of miRNA activity in MSCs activates components that are epigenetically resistant to complete restoration of miRNA expression.

Transcriptional and post-transcriptional regulation in an epigenetic switch

The observation that the cells expressing the miRNA-resistant set of genes are highly malignant, whereas their progenitor wild-type cells before loss and restoration of Dicer are not, suggests epigenetic activation of an oncogenic program. A stable inheritance of an acquired transformed phenotype has been described elsewhere (Iliopoulos et al. 2009). That study elucidated a positive feedback loop between transcriptional (NF-κB regulation of IL-6) and post-transcriptional (Lin-28b, let-7) regulators as key components underlying maintenance of transformation. However, neither Lin-28 paralogue are expressed in the MSCs. Furthermore, extensive perturbation of candidate transcriptional factors identified by computational approaches did not identify the key mediators of the miRNA-resistant signature's induction (Gosline et al. 2016). While the complete mechanism underlying the switch described in this system is

unknown, the positive feedback loop involving Imp1-3 that is activated by temporary loss of Dicer partially underlies this phenomenon.

Reinforcement of sustained oncogene expression by Imp1-3 RNA-binding proteins

Disruption of let-7 regulation of oncofetal targets (Imp1-3, Hmga2, Lin-28) is an important aspect of oncogenesis and is achieved through a variety of mechanisms. Let-7 may be genetically deleted or suppressed by Lin-28 (Calin 2004; Viswanathan et al. 2008; Viswanathan et al. 2009), let-7 activity may be sponged (Powers et al. 2016), and alternative 3'UTR usage may mediate escape of let-7 mediated repression (Mayr et al. 2007; Mayr and Bartel 2009). Based on the above observations on the activation of Imp1-3 in our model system, simultaneous elevated expression of multiple Imp proteins through transcriptional activation may be an additional mechanism overriding let-7 activity. This mechanism is probably independent of observations that Imp1-3 activity converges on post-transcriptionally opposing let-7 and other miRNAs (Jonson et al. 2014; Sheen et al. 2015; Busch et al. 2016). Post-transcriptional silencing by let-7 is fully active in these tumorigenic cells.

There has been a concerted effort towards understanding the mechanisms through which Imp1-3 exert oncogenic effects by use of pull-down experiments to characterize their 'RNA-interactomes' (Hafner et al. 2013; Taniuchi et al. 2014; Conway et al. 2016; Ennajdaoui et al. 2016). Our work complements such studies by investigating the gene expression changes and functional outcomes downstream of simultaneous inactivation of all Imp

paralogues, and places Imp1-3 at the core of an oncogenic switch. The pro-tumorigenic phenotype of hsDicer cells was only impaired with combined Imp1-3 loss and not by loss of the individual members, suggesting strong functional redundancy among this family. This observation is consistent with reports of overlapping mRNA binding as well as auto-regulation by Imp1-3 (Conway et al. 2016; Ennajdaoui et al. 2016). Additionally, we identified Plag1, Fzd6 and Slpi as novel possible downstream effectors of Imp1-3. These oncogenes belonged to the miRNA-resistant signature, were strongly downregulated by Imp1-3 loss, and exhibited correlated expression amongst themselves as well as with Imp1-3 in various human cancers.

In summary, we described an oncogenic switch involving the conversion of non-transformed fibroblasts to stable, transformed cells, by toggling the activity of miRNAs off and on. Future studies will enhance our understanding of the underlying mechanism and role of Imp1-3 in oncogenic networks.

Materials and Methods

All sequencing data are available under Gene Expression Omnibus accession number GSE.

Cell culture

Mesenchymal stem cells were cultured at 37°C with 5% CO_2 in Alpha-MEM supplemented with penicillin/streptomycin and 10% fetal bovine serum.

Protein analysis

Cells were harvested in RIPA supplemented with protease inhibitor cocktail (Roche) and samples were diluted in Laemilli buffer with DTT. Samples were electrophoresed using the Nu-PAGE Bis-Tris Electrophoresis System (Life Technologies) as per the manufacturer's instructions and transferred to a polyvinylidene fluoride (PVDF) or nitrocellulose membrane in a Mini Trans-Blot wet transfer apparatus (Bio-Rad). Membranes were blocked in 5% milk in TBST (for ECL based detection) or Odyssey blocking buffer (for Licor based detection). Primary antibody incubation was performed overnight and samples were detected on film using ECL (Perkin Elmer) or visualized using Odyssey (Licor).

The Imp1-3 polyclonal antibodies were obtained from MBL International (catalogue #: RN007P, RN008P, RN009P), anti-Hmga2 from Cell Signaling (catalogue #: 5269), anti-p107 and anti-Actin from SantaCruz Biotechnology (catalogue #: sc-318, sc-130656), anti-HA and anti-Vinculin from Sigma-Aldrich (catalogue #: 11867423001, V9131), and anti-Dicer from Bethyl laboratories (catalogue #: A301-936A).

qPCR

Total RNA was isolated using Trizol or the RNeasy mini kit (Qiagen) following the manufacturer's instructions, treated with the TURBO DNA-free kit (Life Technologies), and reverse transcribed with Oligo(dT) primers using Superscript III (Life Technologies). QPCR reactions on resulting cDNAs were performed using Power SYBR Green (Life Technologies) run on an Applied Biosystems 7500 Real-Time PCR instrument. QPCR primers are listed in Supplemental Table.

Northern blot

Total RNA was separated on a 12% polyacrylamide/Urea/TBE gel (Sequagel, National Diagnostics) and transferred to a HyBond N+ membrane (GE Healthcare) using the TransBlot SD Semi-dry Transfer System (Biorad). RNA was then UV crosslinked to the membrane. DNA oligo probes perfectly complementary to the miRNAs of interest were γ-^{32}P end labeled using T4 polynucleotide kinase (New England Biolabs) and purified using Illustra G-25 MicroSpin columns (GE Healthcare). Membranes were pre-hybridized in ULTRAHyb Oligo (Ambion) at 42°C for at least 30 min and then hybridized with a radiolabeled DNA probe overnight at 42°C. Blots were washed twice for 30 min with 2xSSC and 0.5% SDS prewarmed to 42°C. RNA was visualized by

exposure to phosphor screens and then imaged on a Storm scanner (Molecular Dynamics).

Animal work

All animal work was performed under the guidelines of the MIT Division of Comparative Medicine, with protocols approved by the MIT Committee for Animal Care, and were consistent with the Guide for the Care and Use of Laboratory Animals, National Research Council, 1996 (Institutional Animal Welfare Assurance no. A-3125-01).

Soft agar assay

Cells were suspended in 0.4% SeaPlaque agarose (Lonza) in complete DMEM, and seeded over a first layer of 0.8% agarose in complete DMEM. $2x10^4$-$4x10^4$ cells were seeded per well of a 6 well plate and cells were seeded in triplicate for each cell type. Cells were grown at 37°C and 5% CO_2 and colonies were counted 2-3 weeks after seeding. For each replicate, 5 random fields were captured by light microscopy at 4x magnification, and the number of colonies in each field was counted manually by eye.
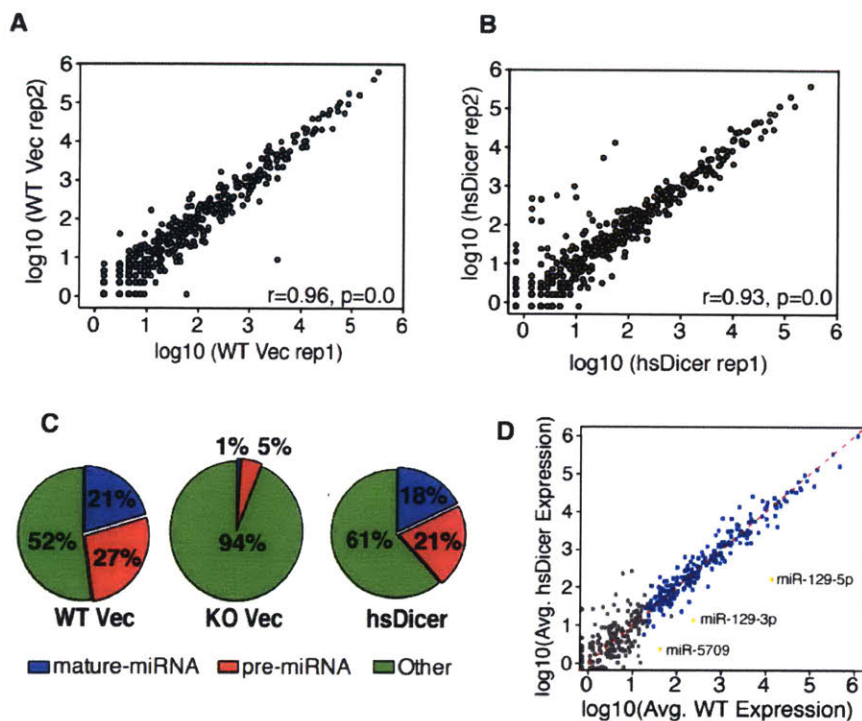
GILA

Subconfluent cells were trypsinized, counted and seeded at 1,000 cells per well of a 96-well plate in 100ul of cell culture media on ultra low-attachment, round-bottom plates (catalogue #: 7007, Corning). Cells were cultured at 37°C with 5% CO_2, and at the respective time points, plates were transferred to -80° in order to freeze the cell suspension. At the completion of the time course, cells were assayed for ATP content as a proxy of cell viability, using the CellTiter-Glo luminescent cell viability assay (Promega) according to the previously described protocol (Rotem et al. 2015).

Acknowledgements

Supplemental Figures and Tables

Supplemental Fig. 1. Restoration of miRNA expression through hsDicer. (A, B) Comparison of normalized miRNA counts between the 2 independently isolated clones each for WT vec and hsDicer. **(C)** Small RNA library composition for each cell type. Numbers indicated are averaged across the 2 libraries for each genotype. **(D)** Same as in Fig. 1C. Gray indicates non-expressed miRNAs (<5 RPM in WT Vec), yellow indicates the 3 differentially expressed miRNAs (Adj. p-value < 0.05), and blue indicates all other miRNAs. Red line: y=x.
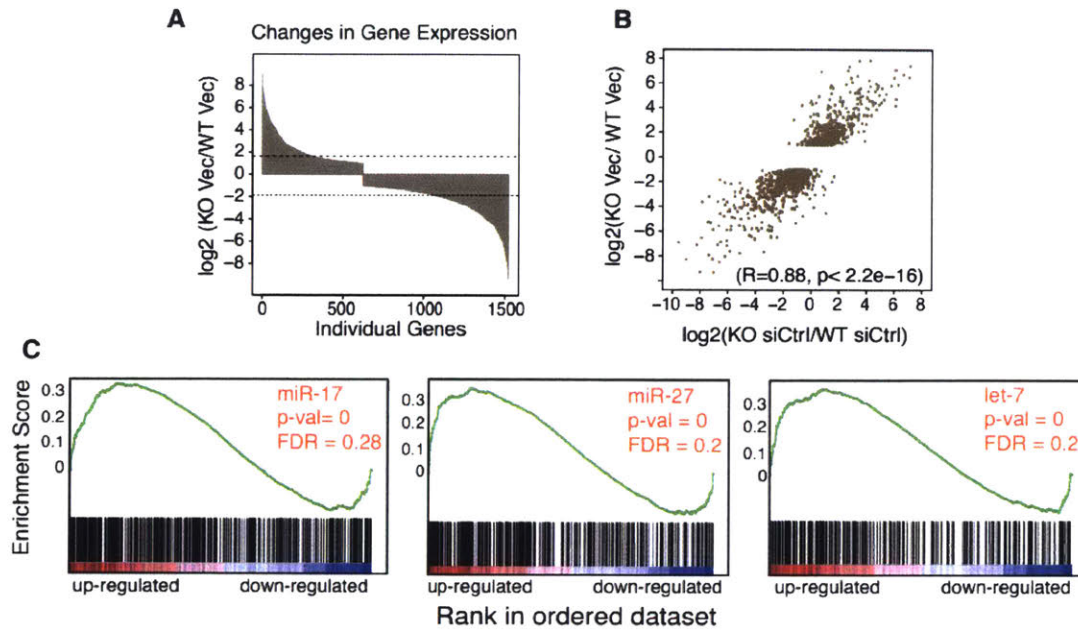
Supplemental Fig. 2. Effects of miRNA loss on gene expression. (A) Fold change waterfall plot for genes differentially expressed between WT Vec and KO Vec cells (FDR < 0.05; fold change ≥ 2 in either direction). Dotted horizontal lines represent median up and downregulation. (B) Scatter plot comparing fold changes reported in this study (y-axis) for differentially expressed genes from (A) to a previous report (Gurtan et al., 2013) (x-axis) with wild-type and knockout cells transfected with a non-targeting siRNA (siCtrl). Pearson correlation and p-value are indicated. (C) GSEA enrichment plots indicate that the most upregulated genes in KO Vec are enriched for targets of miR-17, miR-27 and let-7. Genes are ordered from left to right with decreasing fold change.

A — RNA-Seq analysis of let-7 oncofetal targets

B

C

D
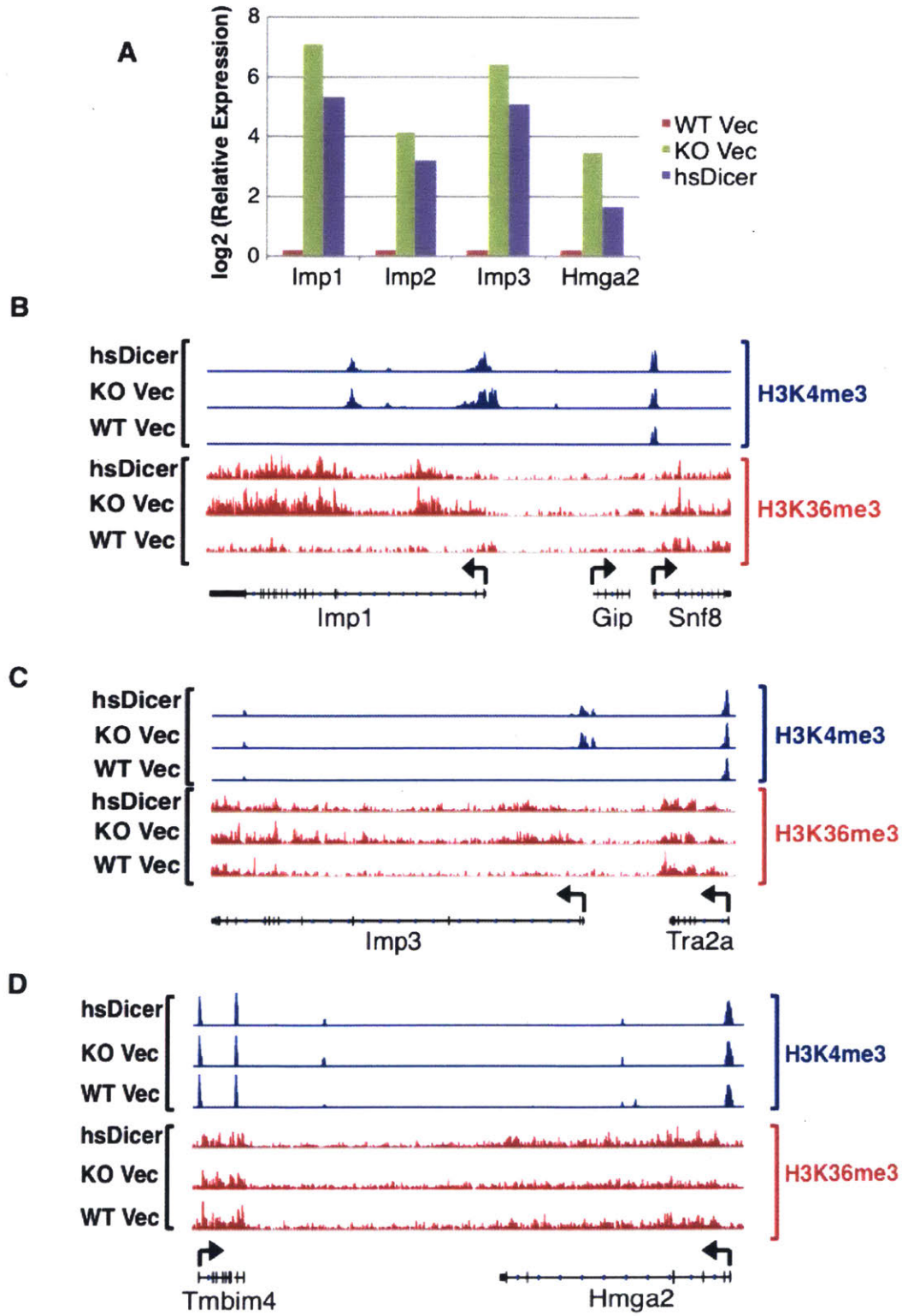
Supplemental Fig. 3. Transcriptional profiling of let-7 oncofetal targets. (A) Log2 RNA-Seq expression of four let-7 regulated oncofetal genes. For each gene, expression is relative to WT Vec levels. **(B-D)** Normalized read counts for H3K4me3 and H3K36me3 marks at the Imp1, Imp3 and Hmga2 loci. Within each chromatin mark, WT Vec, KO Vec and hsDicer are all set to the same scale. Flanking genes are shown as controls. Arrowheads indicate transcription start sites (TSS).
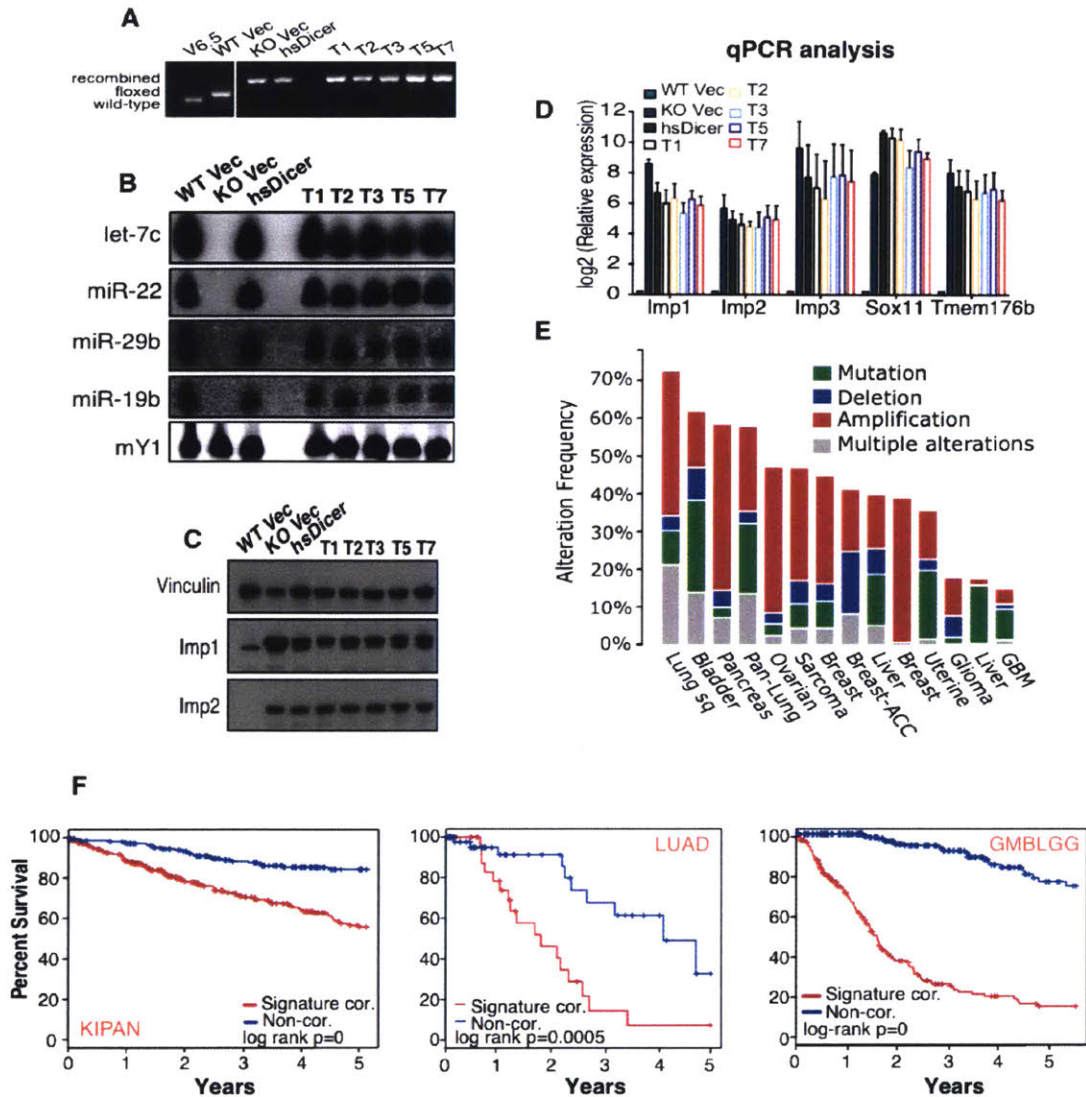
Supplemental Fig. 4. miRNA-resistant genes remain expressed in subcutaneous tumors – related to Figure 5. (A) PCR genotyping of hsDicer subcutaneous tumor derived cell lines shows recombination of the floxed alleles. V6.5: Embryonic stem cells not containing floxed alleles used as a control. **(B)** Northern blot analysis of tumor derived cell lines indicates robust expression of mature miRNAs. mY1 is used as a loading control. **(C)** western blot and **(D)** qPCR analyses also confirm high expression for a panel of irreversible genes. Bars represent mean and standard deviation of 3

219

independent experiments. In all cases, WT Vec, KO Vec and hsDicer are the parental cell lines, while T1-T7 are cell lines derived from hsDicer tumors. **(E)** Pan-cancer alteration profile of the high-confidence miRNA-resistant gene signature using cBioPortal. The cancer types are indicated. Lung sq: Lung Squamous Cell Carcinoma; Pan-Lung: pan non small cell lung cancer; Breast-ACC: Adenoid Cystic Carcinoma of the Breast; GBM: Glioblastoma Multiforme. **(F)** Kaplan-Meier survival analysis for patients with higher correlation (red) of the high-confidence miRNA-resistant signature in KIPAN (pan kidney, $|z| > 0.5$ extremes of score distribution), LUAD (lung adenocarcinoma, $|z| > 1.25$) and GBMLGG (pan glioma, , $|z| > 0.5$) cohorts compared to patients with lower signature correlation scores (blue).
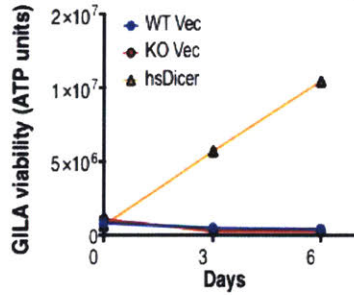
Name	Sequence

Dicer PCR Genotyping

Name	Sequence
MC93	CATGACTCTTCAACTCAAACT
MC94	CCTGACAGTGACGGTCCAAAG
MC113	AGCATGGGGGCACCCTGGTCCTGG

Northern Blot probes

Name	Sequence
let-7c	AACCATACAACCTACTACCTCA
miR-22	ACAGTTCTTCAACTGGCAGCTT
miR-29b	AACACTGATTTCAAATGGTGCTA
miR-19b	TCAGTTTTGCATGGATTTGCACA
miR-16	CGCCAATATTTACGTGCTGCTA
mY1	AAGGGGGGAAAGTGTAGAACAGGA
U6	GCTAATCTTCTCTGTATCGTTCCAATTTTAGTATATGTGCTGCCG

CRISPR/Cas9 sgRNA guides

Gene	Sequence
Imp1 SG1	TCATCGCCCAGTGCTCGTCG
Imp2 SG1	CTGGTCGGGGTAGTCCACGA
Imp2 SG3	CCGTCACCGCCGACGACCTC
Imp3 SG1	ACGCGTAGCCCGTCTTCACC
Imp3 SG2	ACTTTCCAGGTCCGCGGGGC
Imp3 SG3	AGACACTTTCCAGGTCCGCG

qPCR	Column1	Column2
Gene	Forward Sequence	Reverse Sequence
Tmem176a	GAGCAGCAGCACCCAGATAC	TCTCGGAGATGATGTCTGTCA
Tmem176b	GTGCAGAGAGTACCTGAACATG	GGAAGCCACAGACACAACAA
Foxa1	GTAGGACATGTTGAAGGAAGCC	CATGAGAGCAACGACTGGAAC
Sox11	TCCTCTTTATCCTGACCGCC	TTCGAGTTCCCCGACTACTG
Traf1	ATCACGATGAAGAGGGACAGG	AGATCACCAATGTCACCAAGC
Imp1	ACCACCAGAAACACCTGACTCCAA	ATTCTTCCCTGGGCCTTGAACTGA
Imp2	AGCCTGTGCCAATGCTGAGATAGA	TTAGCCCTGGGATCAGATTGGCTT
Imp3	TCTGACTCAGGTAAAGCAGCACCA	TGAGCCTTTACTTCCGCCTTGACT
Gcnt4	CATGTCCTGCAGTTCCATTG	GGGAACCTGTTCCTCACAAA
Rdh10	ACGAGTCTCTGAGCCATGAG	GTCACGATGTACATGAGGCG
Zfyve26	GAGCTGAGCACAAGTACTTCAG	AGGTTGAACATGAGTACCACCT
Ncoa3	GCGCACTCACAATAGTTCCA	GGCCTGGCTTTGAAGACATA
Hmga2	TCACTTGGGTGGGTTCATTGGGTA	TGTCAGCCTTGAAGCATCGGAGAT

Table 1: List of primers used.

A

WT Vec | hsDicer | sgImp1.3 | sgImp1.5 | sgImp2.4 | sgImp2.6 | sgImp3.11 | sgImp3.3

B

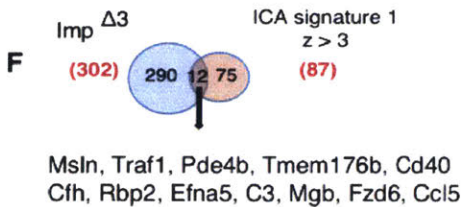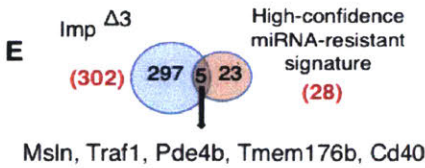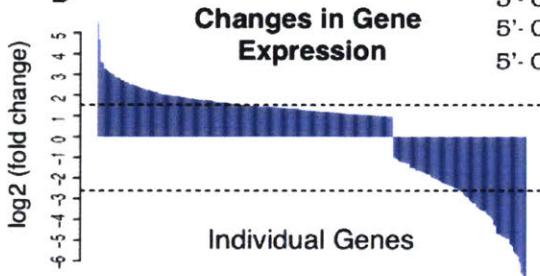GILA viability (ATP units)

WT Vec
KO Vec
hsDicer

Days
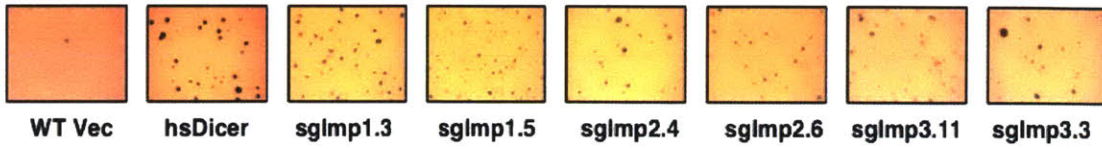
C

Imp1
5'- TGCCCCGA CGAGCACTGGGCGATGAAGGC **WT**
5'- TGCCCC - - CGAGCACTGGGCGATGAAGGC **(both)**
5'- TGCCCCGAAG CGAGCACTGGGCGATGAAGGC **(clone 5)**
5'- TGCCCCGAC CGAGCACTGGGCGATGAAGGC **(clone 6)**

Imp2
5'- CGCCTTCGTGGACTACCCCGACCAGAACT **WT**
5'- CGCCTT - - TGGACTACCCCGACCAGAACT **(both)**
5'- CGCCTTCGT - - - - TACCCCGACCAGAACT **(both)**

Imp3
5'- CCACGCCGGCCCCGCGG ACCTGGAAAGT **WT**
5'- CC - - - - - - - - - - - -CGG ACCTGGAAAGT **(clone 5)**
5'- CCACGCCGGCCCCGCGGG ACCTGGAAAGT **(clone 5)**
5'- CCACGCCG - - - - - -CGG ACCTGGAAAGT **(clone 6)**

D

log2 (fold change)

Changes in Gene Expression

Individual Genes

E

Imp Δ3
(302) 297 5 23

High-confidence miRNA-resistant signature
(28)

Msln, Traf1, Pde4b, Tmem176b, Cd40

F

Imp Δ3
(302) 290 12 75

ICA signature 1
z > 3
(87)

Msln, Traf1, Pde4b, Tmem176b, Cd40
Cfh, Rbp2, Efna5, C3, Mgb, Fzd6, Ccl5

G

KIRP

H

LUAD
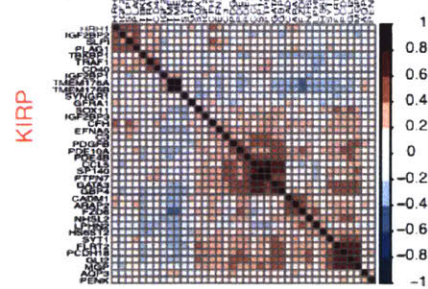
Supplemental Fig. 5. Simultaneous expression of Imp-3 drives the oncogenic switch – related to Figure 6. (A) Representative brightfield soft agar images at 4x magnification for each clone of Imp1, Imp2 and Imp3 single knockout cells. **(B)** GILA assay on the WT Vec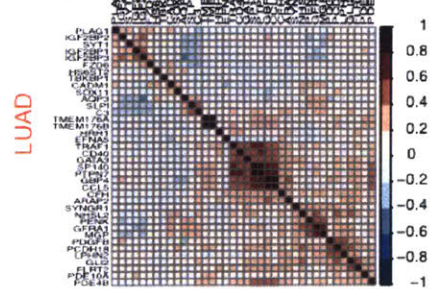, KO Vec and hsDicer cell lines. Only hsDicer cells are able to grow over a 6-day time course in GILA, consistent with their ability to grow and form colonies in soft agar. Data are plotted as the mean ± standard deviation (n=3). **(C)** Genomic PCR product sequencing at the Imp1-3 loci in $Imp^{\Delta 3}$ cells. WT designates the wild type sequence. The 20-nucleotide target sequence is indicated in green; and the PAM is shown in blue. Red dashes denote deletions, red nucleotides denote insertions, and parentheses indicate whether that particular allele is recovered from both or either $Imp^{\Delta 3}$ clone. **(D)** Waterfall plot of log2-fold changes ($Imp^{\Delta 3}$/ hsDicer) for differentially expressed genes (FDR ≤ 0.05 and fold change ≥ 2 in either direction). Dashed horizontal lines represent median up and downregulation. **(E)** Five genes belonging to the high-confidence miRNA-resistant signature (discovered in Fig. 4G) are significantly differentially expressed between $Imp^{\Delta 3}$ and hsDicer cells. **(F)** Overlap of $Imp^{\Delta 3}$ and hsDicer differentially expressed genes with ICA signature 1 associated genes. **(G-H)** Heatmap depicting gene-by-gene spearman correlation coefficients (of leading-edge genes) in tumor samples of **(G)** cervical kidney renal papillary cell carcinoma (KIRP) and **(H)** lung adenocarcinoma (LUAD) patients from TCGA.

References:

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* **11**: R106.

Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP. 2008. The impact of microRNAs on protein output. *Nature* **455**: 64-71.

Bahubeshi A, Bal N, Rio Frio T, Hamel N, Pouchet C, Yilmaz A, Bouron-Dal Soglio D, Williams GM, Tischkowitz M, Priest JR et al. 2010. Germline DICER1 mutations and familial cystic nephroma. *J Med Genet* **47**: 863-866.

Bartel DP. 2009. MicroRNAs: target recognition and regulatory functions. *Cell* **136**: 215-233.

Bell JL, Wachter K, Muhleck B, Pazaitis N, Kohn M, Lederer M, Huttelmaier S. 2013. Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? *Cell Mol Life Sci* **70**: 2657-2675.

Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, Mills AA, Elledge SJ, Anderson KV, Hannon GJ. 2003. Dicer is essential for mouse development. *Nature Genetics* **35**: 215-217.

Bosson AD, Zamudio JR, Sharp PA. 2014. Endogenous miRNA and target concentrations determine susceptibility to potential ceRNA competition. *Mol Cell* **56**: 347-359.

Boyerinas B, Park SM, Hau A, Murmann AE, Peter ME. 2010. The role of let-7 in cell differentiation and cancer. *Endocr Relat Cancer* **17**: F19-36.

Boyerinas B, Park SM, Shomron N, Hedegaard MM, Vinther J, Andersen JS, Feig C, Xu J, Burge CB, Peter ME. 2008. Identification of Let-7-Regulated Oncofetal Genes. *Cancer Research* **68**: 2587-2591.

Busch B, Bley N, Muller S, Glass M, Misiak D, Lederer M, Vetter M, Strauss HG, Thomssen C, Huttelmaier S. 2016. The oncogenic triangle of HMGA2, LIN28B and IGF2BP1 antagonizes tumor-suppressive actions of the let-7 family. *Nucleic Acids Res* **44**: 3845-3864.

Calin GA. 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences* **101**: 2999-3004.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E et al. 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**: 401-404.

Chen J, Wang Y, McMonechy MK, Anglesio MS, Yang W, Senz J, Maines-Bandiera S, Rosner J, Trigo-Gonzalez G, Grace Cheng SW et al. 2015. Recurrent DICER1 hotspot mutations in endometrial tumours and their impact on microRNA biogenesis. *J Pathol*.

Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE, Aqeilan RI, Zupo S, Dono M et al. 2005. miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 13944-13949.

Conway AE, Van Nostrand EL, Pratt GA, Aigner S, Wilbert ML, Sundararaman B, Freese P, Lambert NJ, Sathe S, Liang TY et al. 2016. Enhanced CLIP Uncovers IMP Protein-RNA Targets in Human Pluripotent Stem Cells Important for Cell Adhesion and Survival. *Cell Rep* **15**: 666-679.

de Kock L, Sabbaghian N, Plourde F, Srivastava A, Weber E, Bouron-Dal Soglio D, Hamel N, Choi JH, Park SH, Deal CL et al. 2014. Pituitary blastoma: a pathognomonic feature of germ-line DICER1 mutations. *Acta Neuropathol* **128**: 111-122.

Degrauwe N, Schlumpf TB, Janiszewska M, Martin P, Cauderay A, Provero P, Riggi N, Suva ML, Paro R, Stamenkovic I. 2016. The RNA Binding Protein IMP2 Preserves Glioblastoma Stem Cells by Preventing let-7 Target Gene Silencing. *Cell Rep* **15**: 1634-1647.

Dimitrova N, Gocheva V, Bhutkar A, Resnick R, Jong RM, Miller KM, Bendor J, Jacks T. 2016. Stromal Expression of miR-143/145 Promotes Neoangiogenesis in Lung Cancer Development. *Cancer Discov* **6**: 188-201.

Doench JG, Petersen CP, Sharp PA. 2003. siRNAs can function as miRNAs. *Genes & Development* **17**: 438-442.

Ennajdaoui H, Howard JM, Sterne-Weiler T, Jahanbani F, Coyne DJ, Uren PJ, Dargyte M, Katzman S, Draper JM, Wallace A et al. 2016. IGF2BP3 Modulates the Interaction of Invasion-Associated Transcripts with RISC. *Cell Rep* **15**: 1876-1883.

Foulkes WD, Bahubeshi A, Hamel N, Pasini B, Asioli S, Baynam G, Choong CS, Charles A, Frieder RP, Dishop MK et al. 2011. Extending the phenotypes associated with DICER1 mutations. *Hum Mutat* **32**: 1381-1384.

Foulkes WD, Priest JR, Duchaine TE. 2014. DICER1: mutations, microRNAs and mechanisms. *Nature Reviews Cancer* **14**: 662-672.

Ghildiyal M, Zamore PD. 2009. Small silencing RNAs: an expanding universe. *Nature Reviews Genetics* **10**: 94-108.

Gosline SJ, Gurtan AM, JnBaptiste CK, Bosson A, Milani P, Dalin S, Matthews BJ, Yap YS, Sharp PA, Fraenkel E. 2016. Elucidating MicroRNA Regulatory Networks Using Transcriptional, Post-transcriptional, and Histone Modification Measurements. *Cell Rep* **14**: 310-319.

Grimson A, Farh KKH, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. 2007. MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. *Molecular Cell* **27**: 91-105.

Gurtan AM, Ravi A, Rahl PB, Bosson AD, JnBaptiste CK, Bhutkar A, Whittaker CA, Young RA, Sharp PA. 2013. Let-7 represses Nr6a1 and a mid-gestation developmental program in adult fibroblasts. *Genes Dev* **27**: 941-954.

Hafner M, Max KE, Bandaru P, Morozov P, Gerstberger S, Brown M, Molina H, Tuschl T. 2013. Identification of mRNAs bound and regulated by human LIN28 proteins and molecular requirements for RNA recognition. *RNA* **19**: 613-626.

Harfe BD MM, Mansfield JH, Hornstein E, Tabin CJ. 2005. The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate

limb. *Proceedings of the National Academy of Sciences* **102**: 10898-10903.

Heravi-Moussavi A AM, Cheng SW, Senz J, Yang W, Prentice L, Fejes AP, Chow C, Tone A, Kalloger SE, Hamel N, Roth A, Ha G, Wan AN, Maines-Bandiera S, Salamanca C, Pasini B, Clarke BA, Lee AF, Lee CH, Zhao C, Young RH, Aparicio SA, Sorensen PH, Woo MM, Boyd N, Jones SJ, Hirst M, Marra MA, Gilks B, Shah SP, Foulkes WD, Morin GB, Huntsman DG. 2012. Recurrent somatic DICER1 mutations in nonepithelial ovarian cancers. *N Engl J Med* **366**: 234-242.

Hill DA IJ, Priest JR, Gurnett CA, Dehner LP, Desruisseau D, Jarzembowski JA, Wikenheiser-Brokamp KA, Suarez BK, Whelan AJ, Williams G, Bracamontes D, Messinger Y, Goodfellow PJ. 2009. DICER1 mutations in familial pleuropulmonary blastoma. *Science* **325**: 965.

Hyvarinen A, Oja E. 2000. Independent component analysis: algorithms and applications. *Neural Netw* **13**: 411-430.

Iliopoulos D, Hirsch HA, Struhl K. 2009. An epigenetic switch involving NF-kappaB, Lin28, Let-7 MicroRNA, and IL6 links inflammation to cell transformation. *Cell* **139**: 693-706.

Janiszewska M, Suva ML, Riggi N, Houtkooper RH, Auwerx J, Clement-Schatlo V, Radovanovic I, Rheinbay E, Provero P, Stamenkovic I. 2012. Imp2 controls oxidative phosphorylation and is crucial for preserving glioblastoma cancer stem cells. *Genes Dev* **26**: 1926-1944.

Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, Cheng A, Labourier E, Reinert KL, Brown D, Slack FJ. 2005. RAS Is Regulated by the let-7 MicroRNA Family. *Cell* **120**: 635-647.

Jonson L, Christiansen J, Hansen TV, Vikesa J, Yamamoto Y, Nielsen FC. 2014. IMP3 RNP safe houses prevent miRNA-directed HMGA2 mRNA decay in cancer and development. *Cell Rep* **7**: 539-551.

Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* **19**: 489-501.

Kugel S, Sebastian C, Fitamant J, Ross KN, Saha SK, Jain E, Gladden A, Arora KS, Kato Y, Rivera MN et al. 2016. SIRT6 Suppresses Pancreatic Cancer through Control of Lin28b. *Cell* **165**: 1401-1415.

Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. 2007. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet* **39**: 673-677.

Kumar MS, Pester RE, Chen CY, Lane K, Chin C, Lu J, Kirsch DG, Golub TR, Jacks T. 2009. Dicer1 functions as a haploinsufficient tumor suppressor. *Genes Dev* **23**: 2700-2704.

Lee YS, Dutta A. 2007. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes & Development* **21**: 1025-1030.

Leung AKL, Sharp PA. 2010. MicroRNA Functions in Stress Responses. *Molecular Cell* **40**: 205-215.

Li CM, Gocheva V, Oudin MJ, Bhutkar A, Wang SY, Date SR, Ng SR, Whittaker CA, Bronson RT, Snyder EL et al. 2015. Foxa2 and Cdx2 cooperate with Nkx2-1 to inhibit lung adenocarcinoma metastasis. *Genes Dev* **29**: 1850-1862.

Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**: 1739-1740.

Lin S, Gregory RI. 2015. MicroRNA biogenesis pathways in cancer. *Nature reviews Cancer* **15**: 321-333.

Madison BB, Liu Q, Zhong X, Hahn CM, Lin N, Emmett MJ, Stanger BZ, Lee JS, Rustgi AK. 2013. LIN28B promotes growth and tumorigenesis of the intestinal epithelium via Let-7. *Genes Dev* **27**: 2233-2245.

Manier S, Powers JT, Sacco A, Glavey SV, Huynh D, Reagan MR, Salem KZ, Moschetta M, Shi J, Mishima Y et al. 2016. The LIN28B/let-7 axis is a novel therapeutic pathway in multiple myeloma. *Leukemia*.

Mayr C, Bartel DP. 2009. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**: 673-684.

Mayr C, Hemann MT, Bartel DP. 2007. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* **315**: 1576-1579.

Mayr C HM, Bartel DP. 2007. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* **315**: 1576-1579.

Melo SA, Moutinho C, Ropero S, Calin GA, Rossi S, Spizzo R, Fernandez AF, Davalos V, Villanueva A, Montoya G et al. 2010. A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. *Cancer Cell* **18**: 303-315.

Melo SA, Ropero S, Moutinho C, Aaltonen LA, Yamamoto H, Calin GA, Rossi S, Fernandez AF, Carneiro F, Oliveira C et al. 2009. A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. *Nat Genet* **41**: 365-370.

Molenaar JJ, Domingo-Fernandez R, Ebus ME, Lindner S, Koster J, Drabek K, Mestdagh P, van Sluis P, Valentijn LJ, van Nes J et al. 2012. LIN28B induces neuroblastoma and enhances MYCN levels via let-7 suppression. *Nat Genet* **44**: 1199-1206.

Mukherji S, Ebert MS, Zheng GXY, Tsang JS, Sharp PA, van Oudenaarden A. 2011. MicroRNAs can generate thresholds in target gene expression. *Nature Genetics* **43**: 854-859.

Nguyen LH, Robinton DA, Seligson MT, Wu L, Li L, Rakheja D, Comerford SA, Ramezani S, Sun X, Parikh MS et al. 2014. Lin28b is sufficient to drive liver cancer and necessary for its maintenance in murine models. *Cancer Cell* **26**: 248-261.

Nishino J, Kim S, Zhu Y, Zhu H, Morrison SJ. 2013. A network of heterochronic genes including Imp1 regulates temporal changes in stem cell properties. *Elife* **2**: e00924.

Papagiannakopoulos T, Bauer MR, Davidson SM, Heimann M, Subbaraj L, Bhutkar A, Bartlebaugh J, Vander Heiden MG, Jacks T. 2016. Circadian

Rhythm Disruption Promotes Lung Tumorigenesis. *Cell Metab* **24**: 324-331.

Powers JT, Tsanov KM, Pearson DS, Roels F, Spina CS, Ebright R, Seligson M, de Soysa Y, Cahan P, Theissen J et al. 2016. Multiple mechanisms disrupt the let-7 microRNA family in neuroblastoma. *Nature* **535**: 246-251.

Rakheja D, Chen KS, Liu Y, Shukla AA, Schmid V, Chang TC, Khokhar S, Wickiser JE, Karandikar NJ, Malter JS et al. 2014. Somatic mutations in DROSHA and DICER1 impair microRNA biogenesis through distinct mechanisms in Wilms tumours. *Nat Commun* **2**: 4802.

Ravi A, Gurtan Allan M, Kumar Madhu S, Bhutkar A, Chin C, Lu V, Lees Jacqueline A, Jacks T, Sharp Phillip A. 2012. Proliferation and Tumorigenesis of a Murine Sarcoma Cell Line in the Absence of DICER1. *Cancer Cell* **21**: 848-855.

Rotem A, Janzer A, Izar B, Ji Z, Doench JG, Garraway LA, Struhl K. 2015. Alternative to the soft-agar assay that permits high-throughput drug and genetic screens for cellular transformation. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 5708-5713.

Saetrom P, Heale BSE, Snove O, Aagaard L, Alluin J, Rossi JJ. 2007. Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic Acids Research* **35**: 2333-2342.

Samanta S, Pursell B, Mercurio AM. 2013. IMP3 protein promotes chemoresistance in breast cancer cells by regulating breast cancer resistance protein (ABCG2) expression. *The Journal of biological chemistry* **288**: 12569-12573.

Sampson VB, Rong NH, Han J, Yang Q, Aris V, Soteropoulos P, Petrelli NJ, Dunn SP, Krueger LJ. 2007. MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. *Cancer Res* **67**: 9762-9770.

Sheen YS, Liao YH, Lin MH, Chu CY, Ho BY, Hsieh MC, Chen PC, Cha ST, Jeng YM, Chang CC et al. 2015. IMP-3 promotes migration and invasion of melanoma cells by modulating the expression of HMGA2 and predicts poor prognosis in melanoma. *J Invest Dermatol* **135**: 1065-1073.

Stohr N, Kohn M, Lederer M, Glass M, Reinke C, Singer RH, Huttelmaier S. 2012. IGF2BP1 promotes cell migration by regulating MK5 and PTEN signaling. *Genes Dev* **26**: 176-189.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 15545-15550.

Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y et al. 2004. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res* **64**: 3753-3756.

Taniuchi K, Furihata M, Hanazaki K, Saito M, Saibara T. 2014. IGF2BP3-mediated translation in cell protrusions promotes cell invasiveness and metastasis of pancreatic cancer. *Oncotarget* **5**: 6832-6845.

Tessier CR, Doyle GA, Clark BA, Pitot HC, Ross J. 2004. Mammary tumor induction in transgenic mice expressing an RNA-binding protein. *Cancer Res* **64**: 209-214.

Ventura A, Jacks T. 2009. MicroRNAs and cancer: short RNAs go a long way. *Cell* **136**: 586-591.

Viswanathan SR, Daley GQ, Gregory RI. 2008. Selective blockade of microRNA processing by Lin28. *Science* **320**: 97-100.

Viswanathan SR, Powers JT, Einhorn W, Hoshida Y, Ng TL, Toffanin S, O'Sullivan M, Lu J, Phillips LA, Lockhart VL et al. 2009. Lin28 promotes transformation and is associated with advanced human malignancies. *Nature Genetics* **41**: 843-848.

Walz AL, Ooms A, Gadd S, Gerhard DS, Smith MA, Guidry Auvil JM, Meerzaman D, Chen QR, Hsu CH, Yan C et al. 2015. Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. *Cancer Cell* **27**: 286-297.

Wang Y, Baskerville S, Shenoy A, Babiarz JE, Baehner L, Blelloch R. 2008. Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat Genet* **40**: 1478-1483.

Wegert J, Ishaque N, Vardapour R, Georg C, Gu Z, Bieg M, Ziegler B, Bausenwein S, Nourkami N, Ludwig N et al. 2015. Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors. *Cancer Cell* **27**: 298-311.

Wienholds E, Koudijs MJ, van Eeden FJ, Cuppen E, Plasterk RH. 2003. The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nat Genet* **35**: 217-218.

Wilusz JE, Jnbaptiste CK, Lu LY, Kuhn CD, Joshua-Tor L, Sharp PA. 2012. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.*

Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS. 2009. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell* **137**: 647-658.

Zhu H, Shyh-Chang N, Segre AV, Shinoda G, Shah SP, Einhorn WS, Takeuchi A, Engreitz JM, Hagan JP, Kharas MG et al. 2011. The Lin28/let-7 axis regulates glucose metabolism. *Cell* **147**: 81-94.

Supplementary Material and Methods

Expression plasmid construction and cell line generation

Wild-type human Dicer (hsDicer) was PCR-amplified from pCAGGs-Flag-hsDicer (Addgene plasmid # 41584) using Forward: 5'-AAAAACATGTTATACCCATACGACGTACCAGATTACGCTGACTACAAAGACG ATGACGAC-3' and reverse: 5'-AAAAAAAAACATGTTCAGCTATTGGGAACC TGAGGTTG -3' primers. An HA-Flag tandem epitope was encoded within the 5' forward primer used for amplification. A single N-terminal leucine residue precedes the HA-Flag tag. Similarly, hsDicer is preceded by a single leucine residue after the flag epitope. HA-Flag-hsDicer was subcloned into the NcoI sites of the retroviral vector pMMP-Puro. Phoenix packaging cells were transfected with either an empty pMMP-Puro retroviral vector plasmid or the equivalent vector with the hsDicer insert using Lipofectamine 2000 (Invitrogen). Viral supernatants were cleared with a 0.45 um syringe filter and concentrated using Ultracel 100K centrifugal filter units (Millipore). MSCs at subconfluency were infected with viral supernantant supplemented with 8µg of polybrene (Sigma). Stably transduced lines were selected and maintained in alpha-MEM containing 2.5µg/ml puromycin (Life Technologies). After a heterogenous population of puromycin resistant cells grew out, cells were seeded at low density and two monoclonal lines were isolated per genotype.

Dual color miRNA reporter assay

The bidirectional pTRE-Tight-BI (Clontech) eYFP and mCherry reporter construct has been previously described (Mukherji et al. 2011). Oligonucleotides encoding target sites for miR-24, miR-29b, miR-31 and let-7c were synthesized by IDT and inserted into the 3' UTR of mCherry using HindIII and SalI cloning sites. The sequences are provided in Supplemental Table S7. Each miRNA target is imperfectly complementary across nucleotides 9-11 and consists of 2 tandem sites separated by a 6-nucleotide spacer. MSCs were seeded at 10,000 cells per well of a 96-well plate, and the next day were transfected with equivalent amounts (100ng each) of the reporter plasmid and rtTA using Lipofectamine 2000. At the time of transfection, the media was changed to complete Alpha-MEM supplemented with 2mg/ml doxycycline (Sigma). FACS measurements were taken on a LSR II HTS instrument (BD Biosciences) 24 hrs post-transfection and population averages of mCherry and eYFP intensities were used for determining the fold repression.

Doxorubicin and apoptosis assays

For basal apoptosis analysis, MSCs were seeded at 80,000 cells per well of a 12 well plate. Cells were harvested for analysis 4 days later. The cell culture media was harvested and after washing, was pooled with the PBS supernatant to collect floating cells. Wells were trypsinized, quenched with cell culture media and then cells were harvested by centrifugation and added to the pool of floating cells. The combined cell suspension was washed in cold PBS and harvested via centrifugation. Apoptosis was assayed using the Active Caspase3-FITC kit (BD Biosciences) following the manufacturer's protocol. The percentage of FITC stained cells was measured by an LSR

II machine (BD Biosciences).

For Doxorubicin induced stress analysis, MSCs were seeded at 10000 cells per well of a 96 well plate. The next day, cells were treated with Alpha-MEM supplemented with Doxorubicin at the indicated concentrations. Cells were harvested approximately 24 hours post drug treatment. Analysis of cell death was carried out as described above for the basal apoptosis assay.

PCR genotyping

Cells were genotyped for Dicer status as previously described (Harfe BD 2005). Briefly, genomic DNA was isolated using QuickExtract (catalogue #: QE0905T, Epicenter) following the manufacturer's protocol. Genomic DNA was then PCR amplified using the following primer mix in a 2:1:1 ratio respectively- MC93, MC94, MC113 with a mouse genotyping kit from KAPA Biosystems (catalogue #: KK5621). See Supplemental Table S7 for primer sequences. PCR products were analyzed on a 2% agarose gel. The endogenous wild-type allele yields a product of ~350bp, the floxed allele yields a product of ~420bp, and the deletion allele yields a product of ~470bp.

Doubling time

For calculating cell-doubling times, we used the xCELLigence system (Acea Biosciences) for the real-time monitoring of cell proliferation and viability. Cell index (CI- an arbitrary unit reflecting the electronic cell-sensor impedance) measurements were performed according to the manufacturer's instructions. Briefly, background impedance of 100ul of the cell culture media alone in the E-plate was first determined. Then, sub-confluent cells were trypsinized, counted by hematocytometer and resuspended at 25,000 cells/ml. 100ul of the cell suspension was added to each well. The plate was shaken gently to disperse the cells, then left to stand for 30 minutes in the tissue culture hood at room temperature. The E-plate was then transferred to the xCELLigence instrument and the CI (as a measure of cell proliferation) was monitored at 15-minute intervals over 4 days. Normalized cell index were plotted against incubation time and doubling times were determined based on a linear part of the cell index/ proliferation curve using the RTCA-integrated software of the xCELLigence system.

Subcutaneous injections & Histology

Subconfluent Dicer wild-type, Dicer knockout and hsDicer rescued cells were harvested with trypsin and washed in cold PBS. After centrifugation, cells were resuspended at 1×10^5 cells/ 100ul PBS. 100ul of the cell suspension was subcutaneously injected into the flanks of six week old Nu/J nude female mice.

Eight weeks post injection, mice were sacrificed via CO_2 asphyxiation. Excised tumors were fixed in 4% paraformaldehyde, transferred to 70% ethanol, and then embedded in paraffin. Tumors were then sectioned and stained with hematoxylin and eosin, and analysis for tumor type was performed by a pathologist (R.T.B.). Some tumors were trypsinized and replated in Alpha-MEM to generate secondary cell lines.

For the tumorigenicity comparison of hsDicer and $Imp^{\Delta 3}$ cells, 1×10^5 cells were injected into the left flank or the right flank respectively of the same nude mouse. A total of 10 mice were used in this cohort. Mice were monitored weekly for appearance of tumors. From initial appearance, tumors were measured manually using calipers by the

231

same investigator throughout the entire time course. Tumor volume was calculated as follows: volume = ½(length x width^2).

CRISPR/Cas9 knockouts

Imp1-3 targeting guide RNAs were designed using the MIT Optimized CRISPR design tool (http://crispr.mit.edu). Using protocols from http://crispr.genome-engineering.org, these guides were subsequently cloned into a Cas9/gRNA co-expressing construct (PX458, Addgene plasmid # 48138). hsDicer cells were transfected with the plasmid and ~72 hours later were single cell sorted into a 96 well plate based on GFP signal on an AriaII sorter. Single cell clones were expanded, screened via western blot and genomic PCR for successful knockouts, and two validated clones were subsequently used for experiments. For the Imp2 and Imp3 single knockout cells, the two clones used were derived from different sgRNAs. For the triple knockout cells, plasmids targeting all three Imp proteins were co-transfected into cells. In all knockouts, gRNAs were targeted to the first exons of the respective gene. Oligonucleotides used for gRNA cloning are listed in Supplemental Table S7.

For genomic characterization of the knockouts, we PCR amplified the locus of interest using primers listed in Supplemental Table S7. Blunt-end PCR products were either directly sequenced, or were first topocloned (ThermoFisher), propagated in bacteria, then mini-prepped and Sanger sequenced.

Survival Analyses

The high-confidence miRNA-resistant gene signature derived from RNA-seq analysis was used to score expression profiles of individual TCGA tumors (http://cancergenome.nih.gov/) in the indicated datasets using a single sample enrichment approach (Barbie et al. 2009). Tumors were stratified based on the standardized score and extremes of the score distribution (at various z-thresholds) were compared for differences in survival times using the Kaplan-Meier methodology. Additionally, the Cox proportional hazards regression model was used to analyze the prognostic value of the high-confidence miRNA-resistant signature across all patients in the TCGA pancreatic adenocarcinoma cohort, in the context of additional clinical covariates. Hazard ratio proportionality assumptions for the Cox regression model were validated by testing for all interactions simultaneously (p = 0.611). Interactions between the signature score and other significant covariates (age, number of lymph nodes) were tested using a likelihood ratio test (LRT) to contrast a model consisting of both covariates against another model consisting of both covariates plus an interaction term. All survival analyses were performed using the survival package in R.

Imp1-3 downstream gene expression correlations

RSEM RNASeqV2 Level3 normalized mRNA-Seq expression datasets were downloaded from the Broad Institute TCGA Genome Data Analysis Center (2016): Firehose 2016_01_28 run, Broad Institute of MIT and Harvard. doi:10.7908/C11G0KM9 via the URL: https://gdac.broadinstitute.org. The datasets were filtered to restrict it to tumor-derived expression values for the GSEA leading-edge genes from Figure 6K. Genes with an upper quartile value across all tumor samples of less than 10 were filtered

232

out, normalized expression was log transformed, and then gene-by-gene Spearman correlations were calculated in the R programming language. Heatmaps were generated using the corrplot package in R.

miRNA cloning, sequencing and quantitation

From MSCs, total RNA was isolated using Trizol (Life Technologies) and then 20μg was electrophoresed on a 10% denaturing polyacrylamide gel. Small RNAs in the range of 15-65nt were excised (deliberately avoiding the dominant tRNA band). DNA libraries from this size selected RNA population was then generated using the NEBNext Multiplex Small RNA Library Prep Set for Illumina, following the manufacturer's instructions (New England Biolabs).

Multiplexed sequencing reads from an Illumina HiSeq 2000 instrument were bucketed by sample barcode and processed using the FASTX-Toolkit (from the laboratory of Gregory Hannon: http://hannonlab.cshl.edu/fastx_toolkit/index.html) to strip adapters. Clipped reads less than 10bp were dropped form downstream analyses. Within each individual library, the remaining reads were aggregated by sequence identity. Processed reads were first mapped to mouse mature miRNA sequences annotated in miRBase release 19 (http://www.mirbase.org/), using the Bowtie short read alignment tool (Langmead et al. 2009) allowing for unique and repeat alignments with up to a single base pair mismatch per alignment. Reads that did not align to mature micro-RNA sequences were similarly mapped to micro-RNA hairpin sequences where possible, allowing for up to two mismatches per alignment. Individual miRNA mapped reads were summarized by TargetScan microRNA family (Lewis et al. 2005). Differential analysis for miRNA family counts was performed using DESeq (Anders and Huber 2010).

mRNA-Seq analysis

For the first mRNA sequencing dataset (across WT Vec, KO Vec and hsDicer cells), total RNA was isolated with an RNeasy Kit (Qiagen). Samples were DNase treated, and submitted to the BioMicroCenter at M.I.T. for library preparation (Illumina Tru-Seq protocol) and sequencing (Illumina HiSeq 2000; 51nt single end reads). For the second mRNA-Seq dataset (on triple knockout cells as well as the parental hsDicer cells), total RNA was isolated using Trizol and DNase treated with the TURBO DNA-free kit. RNA was then phenol chloroform extracted and analyzed for quality on an Agilent 2100 Bioanalyzer. 1 μg of total RNA with a RIN above 9, was then used as input for library preparation with poly(A) selection, using Illumina's TruSeq Stranded mRNA Library Preparation Kit and protocol. The final libraries were sequenced on a NextSeq 500 machine (75/75 paired end reads).

Illumina HiSeq 2000 or NextSeq 500 reads were mapped to the UCSC mm9 mouse genome build (http://genome.ucsc.edu/) using RSEM (Li and Dewey 2011). For consistency, reads from the second mRNA-Seq run were trimmed down to 51nts using the FASTX-Trimmer tool from the laboratory of Gregory Hannon. Raw estimated expression counts were upper-quartile normalized to a count of 1000 (Bullard et al., 2010) and log2 transformed. Independent Component Analysis (ICA) was used to identify statistically significant (Mann-Whitney-Wilcoxon test) and biologically relevant signatures that characterize the global gene expression profiles of these samples, as

233

described previously (Li et al. 2015; Dimitrova et al. 2016; Papagiannakopoulos et al. 2016). The R implementation of the core Joint Approximate Diagonalization of Eigenmatrices algorithm (JADE) (Nordhausen et al., 2012; Rutledge and Jouan-Rimbaud Bouveresse, 2013; Biton et al., 2013) was used along with custom R utilities.

For pairwise comparisons of RNA-seq samples, reads were mapped using RSEM as before, and differential expression analyses were performed using using EBSeq v1.12 (Leng et al. 2013). For EBSeq, we used the default settings of the EBTest function allowing the algorithm to filter out genes with 0 reads across all conditions. Genes were called differentially expressed if they changed atleast 2-fold in either direction and also had a PPDE of ≥0.95. Gene set enrichment analysis (GSEA) (Subramanian et al. 2005) was carried out using the pre-ranked mode filtering out gene sets larger than 2000 from the Hallmark or C2 Curated gene sets of the Molecular Signature Database Collection (http://software.broadinstitute.org/gsea/msigdb). For testing enrichment of our custom signatures, we added them to the larger set of C2 Curated or Hallmark gene sets and performed GSEA as before. All analyses were conducted in the R Statistical Programming language (http://www.r-project.org/).

ChIP-Seq library preparation

Chromatin immunoprecipitation (ChIP) was performed as described previously (Gurtan et al. 2013). Briefly, ~ 50 million MSCs were cross-linked for 10 minutes by the addition of 1/10 of the cell culture volume of 11% formaldehyde solution. The cross-linking was quenched by addition of 1/20 the volume of 2.5M glycine. Cross-linked cells were washed 2x with PBS, pelleted and stored at -80°C.

Cross-linked cells were lysed with LB1, and washed with LB2 then resuspended and sonicated in LB3. Samples were sonicated for 14-16 minutes total with a Qsonica Q700 sonicator using a 1/8" microtip and ½" coupler, in 2 minute cycles of 5 seconds on 5 seconds off. The initial 2 minutes were done at 30% amplitude and the remaining time was done in sets of 2 minutes at 65% amplitude, with 1 minute rests in between. Sonicated lysates were cleared and incubated overnight at 4°C with Dynal Protein G magnetic beads that had been previously conjugated with 5µg of the respective antibody (H3K4me3- Millipore 07-473; H3K36me3 -Abcam ab9050-100). Beads were then washed twice in LB3, once in LB3 supplemented with 500 mM NaCl, once in LiCl buffer, and once with TE + 50 mM NaCl. Bound complexes were eluted in elution buffer at 65°C for 15 minutes. Cross-links were reversed overnight at 65°C. RNA and protein were digested using RNAse A and Proteinase K respectively, and DNA was purified with phenol chloroform extraction and ethanol precipitation. Barcoded libraries were prepared and sequenced on Illumina HiSeq 2000 (single end, 45nt reads). Relevant buffer recipes are listed below and a detailed library preparation protocol is available upon request

LB1 (50mM HEPES-KOH pH 7.5, 140mM NaCl, 1mM EDTA, 10% Glycerol, 0.5% NP-40, 0.25% Triton X-100)
LB2 (10mM Tris-HCl pH 8.0, 200mM NaCl, 1mM EDTA, 0.5mM EGTA)
LB3 (50mM Tris pH 8, 140mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate)
LiCl Buffer (10mM Tris-HCl pH 8.0, 250mM LiCl, 1mM EDTA, 1% NP-40)
Elution Buffer (50mM Tris-HCl, pH 8.0, 10mM EDTA, 1% SDS)

ChIP-Seq alignments, peak calling and quantification

Reads were de-multiplexed based on library barcodes and mapped to mouse genome mm9 using bowtie (Langmead et al. 2009) requiring unique mapping with at most two mismatches (-n 2 -m 1 –best–strata). Mapped reads were collapsed, and for each histone ChIP the same number of reads was randomly sampled from the corresponding whole cell extract (WCE) input control libraries to match sequencing depth. For example: WT Vec H3K4me3 ChIP and the respective WT Vec WCE were reduced to the same read depth. For each ChIP in a particular cell type, peaks were called relative to the matching genotype's WCE using MACS (Zhang et al. 2008) with default settings and a P-value threshold of 1e-8. Significant H3K4me3 peaks were assigned to genes if they occurred within +/- 2Kb of the transcription start site and significant H3K36me3 peaks were assigned if they overlapped with the gene body (transcription start- transcription end).

For analysis of histone modifications across different cell types, we used MANorm (Shao et al. 2012), a method for quantitative comparison of ChIP-Seq datasets. MANorm normalizes the raw read counts across the two samples in a comparison, and the read density in common peaks are presumed to reflect the scaling relationship of ChIP-Seq signals between two samples. MANorm outputs an M-value, which represents the log2 fold change of normalized reads across the two conditions, as well as a p-value. The log2 fold change in H3K4me3 ChIP signal (for significant peaks called by MACS as described above) was calculated for each pair-wise comparison (KO Vec vs WT Vec and hsDicer vs WT Vec) using the default settings of the MANorm script. Since all peaks input into MANorm were significantly called by MACS (over the respective WCE), we disregarded MANorm's output p-values and instead filtered peaks and their associated genes solely based on M-value >=1.4. In case of any one gene having more than 1 associated H3K4me3 peak, we assigned the peak with the largest M-value to that gene.

Supplemental References:

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* 11: R106.

Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C et al. 2009. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462: 108-112.

Biton, A., Zinovyev, A., Barillot, E., and Radvanyi, F. (2013). MineICA: Independent component analysis of transcriptomic data.

Dimitrova N, Gocheva V, Bhutkar A, Resnick R, Jong RM, Miller KM, Bendor J, Jacks T. 2016. Stromal Expression of miR-143/145 Promotes Neoangiogenesis in Lung Cancer Development. *Cancer Discov* 6: 188-201.

Gurtan AM, Ravi A, Rahl PB, Bosson AD, JnBaptiste CK, Bhutkar A, Whittaker CA, Young RA, Sharp PA. 2013. Let-7 represses Nr6a1 and a mid-gestation developmental program in adult fibroblasts. *Genes Dev* 27: 941-954.

Harfe BD MM, Mansfield JH, Hornstein E, Tabin CJ. 2005. The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. *Proceedings of the National Academy of Sciences* **102**: 10898-10903.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BM, Haag JD, Gould MN, Stewart RM, Kendziorski C. 2013. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **29**: 1035-1043.

Lewis BP, Burge CB, Bartel DP. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**: 15-20.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.

Li CM, Gocheva V, Oudin MJ, Bhutkar A, Wang SY, Date SR, Ng SR, Whittaker CA, Bronson RT, Snyder EL et al. 2015. Foxa2 and Cdx2 cooperate with Nkx2-1 to inhibit lung adenocarcinoma metastasis. *Genes Dev* **29**: 1850-1862.

Mukherji S, Ebert MS, Zheng GXY, Tsang JS, Sharp PA, van Oudenaarden A. 2011. MicroRNAs can generate thresholds in target gene expression. *Nature Genetics* **43**: 854-859.

Nordhausen, K., Cardoso, J.F., Miettinen, J., Oja, H., and Ollila, E. (2012). JADE: JADE and other BSS methods as well as some BSS performance criteria (R package version).

Papagiannakopoulos T, Bauer MR, Davidson SM, Heimann M, Subbaraj L, Bhutkar A, Bartlebaugh J, Vander Heiden MG, Jacks T. 2016. Circadian Rhythm Disruption Promotes Lung Tumorigenesis. *Cell Metab* **24**: 324-331.

Rutledge, D.N., and Jouan-Rimbaud Bouveresse, D. (2013). Independent Components Analysis with the JADE algorithm. TrAC Trends in Analytical Chemistry *50*, 22–32.

Shao Z, Zhang Y, Yuan GC, Orkin SH, Waxman DJ. 2012. MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome Biol* **13**: R16.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 15545-15550.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.