Spring 2021

Identification and Characterization of the P53-Induced Long Noncoding RNA Isoform Pvt1b and Its Role in Stress-Specific Growth Inhibition via Myc Repression

Christiane Olivero
Yale University Graduate School of Arts and Sciences, christianeolivero@gmail.com

Abstract

Identification and Characterization of the P53-Induced Long Noncoding RNA Isoform *Pvt1b* and Its Role in Stress-Specific Growth Inhibition via *Myc* Repression

Christiane Elizabeth Olivero
2021

The tumor suppressor p53 and proto-oncogenic Myc transcription factors are frequently deregulated in cancer, with common loss-of-function and gain-of-function mutations observed in the p53 and Myc networks, respectively. Referred to as the 'guardian of the genome,' p53 regulates genes important for curtailing cellular proliferation and tumorigenesis under conditions of stress, while the proto-oncogene Myc induces genes that, in contrast, promote cellular growth and can, in overcoming growth inhibitory signals, support cancer development. While previous literature has documented decreased *Myc* expression in response to cellular stress, researchers have long puzzled over identifying the specific regulatory lever responsible. The work presented here identifies a novel regulatory axis positioned at the intersection of the p53 and Myc pathways, which represses Myc and restricts cellular proliferation downstream of p53 activation.

Long noncoding RNAs (lncRNAs) are a diverse class of transcripts lacking protein-coding potential and implicated in gene expression regulation. Here I present my work on the identification of an isoform of the lncRNA Plasmacytoma variant translocation 1 (*Pvt1*) and the characterization of its role in the p53-mediated response to stress. I found that the stress-specific *Pvt1b*, expressed 50 Kb downstream of the *Myc* locus, is induced by p53 in response to oncogenic and

genotoxic stress and accumulates at its site of transcription. I demonstrated that production of the *Pvt1b* RNA is necessary and sufficient to repress *Myc* transcription *in cis* without altering the chromatin organization of the locus. I investigated the functional outputs of *Pvt1b*-mediated *Myc* downregulation and found that inhibition of *Pvt1b* increased both *Myc* levels and transcriptional activity and promoted cellular proliferation. Notably, *Pvt1b* loss accelerated tumor growth, but not tumor progression, in an autochthonous mouse model of lung cancer. Further examination of the *Pvt1b* mechanism of action failed to identify *Pvt1b*-specific sequences required for its function, but uncovered a potential role for histone deacetylation in *Pvt1b* regulation of *Myc*. Finally, I initiated development of a suite of genetically engineered *Pvt1* mouse models, the characterization of which will shed light on *Pvt1* function *in vivo* and benefit future mechanistic studies.

Taken together, this work conceptually advances our understanding of stress-induced growth inhibition orchestrated by p53. Specifically, I identify *Pvt1b* as the primary mediator of stress-specific *Myc* repression, providing insight into the long-standing question of how p53 activation triggers *Myc* downregulation. As such, this work has far-reaching implications not only for our understanding of *cis*-acting lncRNAs, which can fine-tune local gene expression downstream of broadly active transcription programs, but also for the exciting therapeutic possibility of restricting *Myc* levels in cancer via *Pvt1b* modulation.

Identification and Characterization of the P53-Induced Long Noncoding RNA Isoform *Pvt1b* and Its Role in Stress-Specific Growth Inhibition via *Myc* Repression

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by

Christiane Elizabeth Olivero

Dissertation Director: Nadya Dimitrova, PhD

June 2021

Acknowledgements

The work presented here would not have been possible without the enduring support of my colleagues, friends, and family.

My development as a scientist has been shaped, guided, and immeasurably enriched by exceptional female mentors. I am immensely grateful not only for their expertise, critiques, encouragement, and advice, but also for the opportunity to learn from the example they set as successful women in science. First and foremost, I would like to thank my thesis advisor, Nadya Dimitrova, for her passion, drive, and unfaltering enthusiasm for our work. She has challenged and inspired me in equal measure, and I am deeply appreciative of her dedicated mentorship and commitment to my intellectual growth. My dissertation research has also benefitted from the guidance of three outstanding scientists that comprise my thesis committee: Valerie Horsley, Josien van Wolfswinkel, and Joan Steitz. I am thankful for their discerning questions, insightful suggestions, and support of my professional aspirations. I would also like to thank my undergraduate research advisors, Priscilla Van Wynsberghe and Natasha Caplen, for fostering my budding scientific curiosity and encouraging me to pursue post-graduate education.

I feel infinitely lucky to have spent my time at the bench surrounded by people I now count as treasured friends. My colleagues in the Dimitrova lab built a community filled with camaraderie, humor, and mutual respect and admiration. I am deeply grateful to the current and former members of this uplifting environment: Ephrath Tesfaye, Lauren Winkler, Elena Martínez-Terroba, Peppe

Militello, Clara Liao, Rahul Nagvekar, Sonja Herter, Nina Grexa, Yee Nip, Ben Shao, and Meg Tian. It is rare to stumble upon a group of people that can function as intellectual sounding board, emotional cheerleaders, and trusted confidantes in turn, and I credit our frequent coffee runs with maintaining my spirit and sanity through the hills and valleys of doctoral research.

I would also like to thank the inclusive scientific community nurtured by the Molecular, Cellular, and Developmental Biology (MCDB) department at Yale. I am grateful for every MCDB seminar, retreat, and happy hour, and the many gifted colleagues I had the pleasure of interacting with during my time here.

My work was supported, in part, by the National Institutes of Health (NIH)-funded Training Program in Cellular, Molecular, and Quantitative Biology; I appreciate all the scientific and career development opportunities it afforded.

Finally, I am grateful for the lovely and kindhearted friends I've had the good fortune to meet at Yale, of which there are too many to name here. Above all, I would like to thank my family for the unconditional love and support they've given so freely over the past five years, and indeed, my entire life. To my grandparents, Jim and Arlene, thank you for your unwavering pride in your favorite granddaughter. To my brother, AJ, thank you for reminding me not to take myself too seriously. To my dad, Bob, thank you for endeavoring to solve any problem I faced, no matter how abstract or intractable. And to my mom, Kerin, thank you for your calming presence during my worst moments of self-doubt. To my parents especially, I would be lost without you both as my guideposts.

Table of Contents

Nonstandard Abbreviations

AAH	atypical adenomatous hyperplasia
AFAP1	actin filament-associated protein 1
AFAP1-AS1	AFAP1 antisense RNA 1
Airn	antisense of Igf2r ncRNA
ALV	avian leukosis virus
ANRIL	antisense noncoding RNA in the INK4 locus
APC	adenomatous polyposis coli protein
AR	androgen receptor
Arf	alternate reading frame of Cdkn2a
ARLNC1	AR-regulated long noncoding RNA 1
ASO	antisense oligonucleotide
BAC	bacterial artificial chromosome
BANCR	BRAF-activated non-protein coding RNA
BC	breast cancer
BCAR4	breast cancer anti-estrogen resistance 4
CARLo-5	cancer-associated region long noncoding RNA 5
CASC14	cancer susceptibility candidate-14
CCAT1/2	colon cancer-associated transcript 1/2
CDKN2A	cyclin-dependent kinase inhibitor 2A/B
ChIP	chromatin immunoprecipitation
CHX	cycloheximide
circRNA	circular RNA
CLIP	crosslinking and immunoprecipitation
CNV	copy number variation
CON	control
CRISPR	clustered regulatory interspaced short palindromic repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR inhibition

CRISPR-SAM	CRISPR synergistic activation mediator
CRNDE	colorectal neoplasia differentially expressed
CTL	cytotoxic T cell
DANCR	differentiation antagonizing non-protein coding RNA
DAPI	4',6-diamidino-2-phenylindole
DINO	damage-induced noncoding
DMBA	7,12-Dimethylbenz[a]anthracene
DNA-FISH	DNA fluorescence *in situ* hybridization
DOXO	doxorubicin
dRNA	'dead' RNA
ecDNA	extrachromosomal DNA
EPIC1	epigenetically induced lncRNA 1
ER	estrogen receptor; ERAR, ER agitation-related
eRNA	enhancer RNA
ETOP	etoposide
EV	empty vector
FAL1	focally amplified lncRNA 1
FALEC	focally amplified long noncoding RNA in epithelial cancer
Firre	functional intergenic repeating RNA element
F1	first filial
Gadd7	growth-arrested DNA damage-inducible gene 7
GAPLINC	gastric adenocarcinoma predictive long intergenic noncoding RNA
GAS5	growth arrest-specific 5
GATA6	GATA-binding protein 6
GEMM	genetically engineered mouse model
GFP	green fluorescent protein
GOF	gain-of-function
gRNA	guide RNA
GWAS	genome-wide associated study
Gy	gray
HDAC	histone deacetylase

HDR	homology directed repair
H&E	hematoxylin and eosin
hnRNP	heterogeneous ribonucleoprotein
HOTAIR	HOX transcript antisense RNA
HOTTIP	HOXA distal transcript antisense RNA
HPV	human papillomavirus
HULC	highly upregulated in liver cancer
IDR	intrinsically disordered region
Igf2r	insulin-like growth factor receptor 2
IκB	inhibitor of kappa B
JAX	Jackson Laboratory for Genomic Medicine
KC	*K-ras$^{LSL\text{-}G12D/+}$; Rosa26-Cas9$^{LSL/LSL}$*
KI	knock-in
KO	knock-out
KPC	*K-ras$^{LSL\text{-}G12D/+}$; p53$^{FL/FL}$; Rosa26-Cas9$^{LSL/LSL}$*
KRAS	kirsten rat sarcoma viral oncogene homolog
KPR	*Kras$^{LA2\text{-}G12D/+}$; p53$^{LSL/LSL}$; Rosa26-CreER^{T2}*
LAST	lncRNA-assisted stabilization of transcripts
LED	lncRNA activator of enhancer domains
LINC-PINT	long intergenic non-protein coding RNA, p53-induced transcript
LLPS	liquid-liquid phase separation
LNA	locked nucleic acid
lncGATA6	lncRNA GATA6
lncPRESS1	lncRNA p53-regulated and ESC-associated 1
lncRNA	long noncoding RNA
LOF	loss-of-function
LOH	loss of heterozygosity
LSAMP	limbic system-associated membrane protein
LSL	*loxP*-STOP-*loxP*
LUAD	lung adenocarcinoma
MALAT1	metastasis-associated lung adenocarcinoma transcript 1

MAPK	mitogen-activated protein kinase
MaTAR	mammary tumor-associated RNA
MAX	Myc associated factor X
MBP	Ms2 binding protein
MCL1	myeloid cell leukemia sequence 1
Mdm2	mouse double minute 2
MEF	mouse embryonic fibroblast
MEG3	maternally expressed 3
miRNA	microRNA
MITF	microphthalmia-associated transcription factor
MMTV-PyMT	mouse mammary tumor virus-polyomavirus middle T antigen
Morrbid	myeloid RNA regulator of Bim-induced death
mRNA	messenger RNA
MYC	myelocytomasis oncogene
MYCLo	MYC-regulated lncRNA
NANCI	Nkx2.1-associated noncoding intergenic RNA
NBAT-1	neuroblastoma-associated transcript-1
ncRNA	noncoding RNA
NEAT1	nuclear enriched abundant transcript 1
NF-κB	nuclear factor-kappa B
NKILA	NF-κB interacting long noncoding RNA
NKX2.1	NK2 homeobox 1
NMD	nonsense-mediated decay
NORAD	noncoding RNA activated by DNA damage
OIS1	oncogene-induced senescence 1
ORF	open reading frame
Orilnc1	oncogenic RAS-induced lncRNA 1
PAM	protospacer adjacent motif
PANDA	p21-associated ncRNA DNA damage activated
PANDAR	promoter of CDKN1A antisense DNA damage activated RNA
PanINs	pancreatic intraepithelial neoplasias
PAS	polyadenylation sequence

PCa	prostate cancer
PCA3	prostate cancer antigen 3
PCATs	prostate cancer-associated ncRNA transcripts
PCAT1/19	prostate cancer-associated transcript 1/19
PcG	polycomb group
PCGEM1	prostate cancer gene expression marker 1
PDAC	pancreatic ductal adenocarcinoma
PDX	patient-derived xenograft
PHH3	phosphorylated histone H3
PINCR	p53-induced noncoding RNA
PRAL	p53 regulation-associated lncRNA
PRC1/2	polycomb repressive complex 1/2
PR-lncRNA	p53-regulated lncRNA
PR MEF	*p53$^{LSL/LSL}$; Rosa-CreER^{T2}* mouse embryonic fibroblast
PRNCR1	prostate cancer- associated noncoding RNA 1
PSA	prostate-specific antigen
PTENP1	phosphatase and tensin homolog pseudogene 1
PURPL	p53 upregulated regulator of p53 levels
PVT1	plasmacytoma variant translocation 1
p53BER	p53-bound enhancer regions
p53RE	p53 response element
qRT-PCR	quantitative reverse transcription polymerase chain reaction
RAP-MS	RNA antisense purification mass spectrometry
RBP	RNA binding protein
RIP	RNA immunoprecipitation
RNAi	RNA interference
RNA Pol II	RNA polymerase II
RNP	ribonucleoprotein
SAHA	suberoylanilide hydroxamic acid
SALNR	senescence-associated lncRNA
SAMMSON	survival-associated mitochondrial melanoma-specific oncogenic noncoding RNA

SCC	squamous cell carcinoma
SCNV	somatic copy number variation
SHARP	SMART/HDAC1-associated repressor protein
siRNA	small interfering RNA
smRNA-FISH	single molecule RNA fluorescence *in situ* hybridization
SMRT	silencing mediator of retinoic acid and thyroid hormone receptor
SNHG15	small nucleolar RNA host gene 15
snoRNA	small nucleolar RNA
SNP	single nucleotide polymorphism
SpCas9	*Streptococcus pyogenes* Cas9
SPRY4	sprouty RTK signaling antagonist 4
SPRY4-IT1	SPRY4 intronic transcript 1
ssODN	single-stranded oligo donors
TAD	topologically associated domain
TAM	tamoxifen
TCGA	The Cancer Genome Atlas
THOR	testis-associated highly conserved oncogenic long noncoding RNA
TIDE	tracking of indels by decomposition
TPA	12-O-tetradecanoylphorbol-13-acetate
TSA	trichostatin A
TSS	transcription start site
TWI	twister ribozyme
UCA1	urothelial cancer associated 1
Wnt	wingless/integrated
WT	wild-type
XCI	X-chromosome inactivation
XIST	X-inactive specific transcript
XRN2	5'-3' Exoribonuclease 2

List of Figures and Tables

Chapter 1: Background

Long noncoding RNAs in gene regulation

Recent sequencing advances have revealed pervasive transcription of mammalian genomes, far exceeding the level of RNA production required for protein synthesis alone. Specifically, despite protein-coding sequences comprising just 2% of the human genome, as much as 75% of the genome is transcribed (Djebali et al., 2012). This discrepancy has largely upended one of the foundational tenets of molecular biology: that RNA (with a few notable exemptions) provides the cellular instructions, copied from DNA, to produce proteins (Rinn and Chang, 2012). Why cells expend the energy to transcribe a majority of the genome is still not well understood, and some have speculated that these noncoding RNAs (ncRNAs), not being destined for translation, may amount to nothing more than extensive transcriptional noise. However, there is increasing evidence for noncoding transcripts with critical roles in cellular homeostasis, implicating these RNAs as having nuanced and previously unappreciated functions that go far beyond mere messenger (*reviewed in* (Statello et al., 2020)).

Once simply referred to as "junk DNA," recent years have seen a collective reframing of noncoding DNA sequences as mysterious genomic "dark matter" with unexplored functional depths. Breaking from a protein-centric view of cellular operations, there has been growing interest in how the RNAs produced from these enigmatic parts of the genome might impose an additional regulatory layer on cellular activities. Constituting perhaps the most nebulous of these ncRNA classes,

long noncoding RNAs (lncRNAs) encompass a heterogenous collection of RNA molecules broadly defined as transcripts exceeding 200 nucleotides in length and, like other ncRNAs, lacking protein-coding potential (Rinn and Chang, 2012). In contrast to other ncRNAs classes, such as miRNAs (microRNAs), with their comparatively well-defined role in post-transcriptional gene regulation, lncRNAs have far more varied and versatile roles in biology, stemming in part from their imprecise operational definition (Cech and Steitz, 2014).

LncRNA transcription and processing share many similarities with those of protein-coding genes. For one thing, lncRNA genes possess chromatin marks consistent with other actively transcribed genes: H3K4me3, H3K9ac, H3K27ac (Guttman et al., 2009). Like mRNAs (messenger RNAs), many lncRNAs are transcribed by RNA Pol II, and are often subject to 5'-capping, splicing, and polyadenylation (Quinn and Chang, 2016). Following transcription, lncRNAs have comparable stability to mRNAs, with slightly shorter half-lives on average (Clark et al., 2012), and are also subject to nonsense-mediated decay (NMD), although perhaps with increased susceptibility over mRNAs (Mendell et al., 2004). However, there are several key features that distinguish lncRNAs from coding RNA transcripts, apart from their lack of open reading frame (ORF). First, the structures of lncRNA loci can vary widely, with lncRNA genes found overlapping, antisense to, or divergent from protein-coding genes, or located in intronic or intergenic regions (Rinn and Chang, 2012). Notably, lncRNAs can also undergo unique processing events, including RNase P 3' end cleavage (as in MALAT1 processing) (Wilusz et al., 2008), RNA back-splicing to form circRNAs (circular RNAs) (Salzman et al., 2012), and the trimming of snoRNAs (small nucleolar RNAs) from

the introns of protein-coding genes (Yin et al., 2012) (*reviewed in* (Quinn and Chang, 2016)). Importantly, it is likely that our understanding of the full breadth of diversity in lncRNA form, structure, and processing is incomplete.

LncRNA genes engulf protein-coding genes in abundance. While one study identified roughly 60,000 lncRNA genes in the human genome (Iyer et al., 2015), estimates of that number now approach 100,000, far outstripping the 21,000 known protein-coding genes (Fang et al., 2018). Interestingly, some have noted a correlation between higher numbers of lncRNA genes and increasing organismal complexity, suggesting lncRNAs may have played an outsized role in recent evolutionary history (Jandura and Krause, 2017). Indeed, the ratio of noncoding to protein-coding DNA sequences increases as a factor of developmental complexity and is especially high in vertebrates (Mattick, 2004). This observation highlights the importance of the expansion of the noncoding genome, although it is still unknown whether this relationship between noncoding DNA and organismal intricacy is causal or simply incidental.

Analyses of the human transcriptional landscape have determined that lncRNAs are expressed with exquisite cell-type and disease-state specificity (Iyer et al., 2015). While this may point to critical lncRNA functions in normal and disease states, it is thus far from clear whether these specific lncRNA expression patterns are a cause or consequence of underlying biology. To date, examination of individual lncRNAs has revealed functions in cell cycle regulation (Dimitrova et al., 2014; Marin-Bejar et al., 2013), nuclear organization (Hacisuleyman et al., 2014; Sunwoo et al., 2009), and differentiation (Jain et al., 2016; Kretz et al., 2013), among myriad other ubiquitous cellular processes (*reviewed in* (Statello et

al., 2020)). As a result of lncRNA-specific expression signatures, the putative prognostic value of lncRNAs in disease is high. For example, several lncRNAs are expressed at higher levels in accordance with increased tumor stage or metastasis risk, and can provide valuable information about cancer severity (Lu et al., 2017; Shi et al., 2015). UCA1 (*Urothelial Cancer Associated-1*) and HULC (*Highly Upregulated in Liver Cancer*) have been proposed as biomarkers for bladder and liver cancer, respectively, in keeping with the tissues in which they were originally identified (Milowich et al., 2015; Xie et al., 2013). Surprisingly, the only lncRNA to be used in an FDA-approved diagnostic test to date is PCA3 (*Prostate Cancer Antigen 3*); its presence in patient urine samples enables prostate cancer diagnosis with a high degree of accuracy and sensitivity, surpassing the previous diagnostic standard, prostate-specific antigen (PSA) (Fradet et al., 2004; Hessels et al., 2003). Continuing advances in our knowledge of the specificity of lncRNA expression patterns in disease will provide ample occasion for lncRNA-based diagnostic and prognostic test development moving forward.

Of the lncRNAs which have been functionally characterized, many regulate gene expression, having described roles in modulating virtually every step of RNA production (Dimitrova et al., 2014; Marin-Bejar et al., 2013), processing (Tripathi et al., 2010; Yap et al., 2018), stability (Cao et al., 2017; Kretz et al., 2013), and translation (Carrieri et al., 2012; Zhang et al., 2013) (*reviewed in* (Statello et al., 2020)). Gene regulation by lncRNAs is an extensive area of research; there has been significant interest in the role of lncRNAs in epigenetic modifications in particular, with several studies describing lncRNAs interacting with polycomb proteins to elicit gene repression (Khalil et al., 2009; Rinn et al., 2007; Tsai et al.,

2010). While recent work demonstrating the promiscuous binding of PRC2 (Polycomb Repressive Complex 2) to RNAs has cast doubt on the prevalence of this mechanism of action (Davidovich et al., 2013), lncRNAs have been shown to frequently associate with chromatin-modifying complexes to control transcription of target genes (*reviewed in* (Statello et al., 2020)). It is important to note that lncRNA functions are by no means restricted to gene regulation, and an increasing number of studies highlight regulatory roles for lncRNAs in organizing subcellular compartments and governing organelle activities (Clemson et al., 2009; Leucci et al., 2016; Yap et al., 2018).

Despite immense growth in the field of lncRNA biology over the past three decades, efforts to distinguish *bona fide* functional lncRNAs from transcriptional noise have been bogged down by the sheer volume of uncharacterized lncRNA transcripts, raising questions about how to best prioritize them for additional study. Bioinformatic analyses have attempted to tackle this problem by examining the molecular features and regulation of lncRNAs to provide clues about their potential functions. One popular method involves assigning putative functions to lncRNAs based on their co-expression with protein-coding genes (Guttman et al., 2009; Hung et al., 2011). These so-called "guilt by association" studies use what is known about different cellular pathways as a proxy for lncRNA function based on whether a lncRNA is co-regulated with protein-coding genes in the same network. The database decodeRNA catalogues lncRNAs based on this strategy (Lefever et al., 2017), although these functional projections should be approached with caution due to the correlative nature of these analyses. Others have attempted to predict function by evaluating lncRNA composition. For example, one recent study

grouped lncRNAs based on short sequence motifs called kmers, finding functional similarities between lncRNAs with related kmer profiles, despite an absence in linear sequence homology (Kirk et al., 2018). Notwithstanding these advances in computational methods, experimental validation of lncRNA candidates is the only way to confirm their functional importance. As such, recent genome-wide screens for functional lncRNAs that evaluate a specific cellular output following perturbation, such as those that have been performed using CRISPR/Cas9 genome editing, or its variants CRISPRa (CRISPR activation) or CRISPRi (CRISPR inactivation), can provide useful insights into function (Bester et al., 2018; Joung et al., 2017; Liu et al., 2017; Zhu et al., 2016).

Long noncoding RNA mechanisms of action

The study of lncRNAs poses unique challenges, but perhaps the most irksome of these is the inability to define lncRNAs based on an unequivocal set of characteristics. Unlike mRNAs, or indeed some other ncRNA classes, there is no distinct (or even hazy) lncRNA model that can provide clear instructions for functional characterization, because any two lncRNAs can vary greatly in form, function, and mechanism of action. As a class, lncRNAs are extremely modular, able to bind DNA, proteins, and other RNAs, and multifaceted in their regulatory output, capable of eliciting either positive or negative feedback within a variety of cellular pathways (*reviewed in* (Statello et al., 2020)). Therefore, while the existing body of lncRNA literature can provide clues as to what a lncRNA might do, there is no standard path to follow when it comes to experimental design. Many frameworks have been developed which attempt to classify lncRNAs by their

modes of action (Rinn and Chang, 2012; Wang and Chang, 2011). For example, some have proposed dividing lncRNAs by "mechanistic themes" into those which act as either "decoys," which interfere with DNA-protein interactions, "scaffolds," which bring together two or more proteins in a complex, or "guides," which localize proteins within a specific genomic area (Rinn and Chang, 2012). While these groupings are certainly useful, there are many gray areas and potential for overlap that preclude precise categorization.

One framework that is particularly helpful for initial lncRNA characterization without extensive *a priori* knowledge of mechanism involves broadly categorizing lncRNAs by their localization, or more specifically, based on the cellular compartment in which they reside. Subcellular fractionation and single molecule RNA fluorescence *in situ* hybridization (smRNA-FISH) are useful experimental tools for determining lncRNA location within the cell (Cabili et al., 2015; Conrad and Orom, 2017). Notably, the majority of lncRNAs are nuclear-enriched, perhaps reflecting the large number of lncRNAs involved in some aspect of gene expression regulation (Cabili et al., 2015; Derrien et al., 2012). LncRNA localization patterns have significant bearing on the spectrum of potential functions a lncRNA can execute, and can therefore guide further mechanistic studies. As such, lncRNA subcellular position informs an extremely related framework: categorization based on whether a lncRNA regulates *in cis* or *in trans* (Kopp and Mendell, 2018; Quinn and Chang, 2016).

***Cis*-acting lncRNAs** generally reside close to their site of transcription and regulate genes located *in cis*, or which are expressed from the same chromosome (Figure 1) (Gil and Ulitsky, 2020). These lncRNAs may act on their

nearest neighbor, or cross large linear distances to enact regulation, with spatially disparate genes brought in close proximity by long-range chromatin interactions (Cai et al., 2016). These distant contacts are often, but not always, facilitated between genes residing within the same topologically associated domain (TAD) (Groff et al., 2018). While many *cis*-regulatory lncRNAs exert control over one or more target genes, their effects can also be far more widespread.

One lncRNA with extensive regulatory influence is *XIST* (*X-Inactive-Specific Transcript*), a lncRNA which performs the critical task of dosage compensation in females by coordinating epigenetic repression of the entire X-chromosome from which it is expressed in a process known as X-chromosome inactivation (XCI) (*reviewed in* (Loda and Heard, 2019)). *XIST* becomes upregulated in early development, eventually coating one of two X-chromosomes to induce widespread chromosomal reorganization and a heterochromatic state (Loda and Heard, 2019). Indispensable for gene silencing, *XIST* is responsible for coordinating the repression of >1000 genes through functions encoded in discrete and genetically separable elements of the *XIST* transcript (Loda and Heard, 2019). For example, the repeat A region of *XIST* is essential for gene silencing (Wutz et al., 2002), while other regions have been identified as necessary for recruitment of PRC1 (Polycomb Repressive Complex 1) or its localization on the inactive X-chromosome, with some occasional redundancy (Colognori et al., 2019; Wutz et al., 2002). The *XIST* sequences and RNA binding proteins (RBPs) required for the establishment and maintenance of gene silencing, and the precise order of events in XCI, are incompletely understood. Attempts to identify *XIST* interactors have revealed numerous binding partners including various PcG (polycomb group)

proteins, hnRNPs (heterogeneous ribonucleoproteins), and others with unknown functional significance (Chu et al., 2015; McHugh et al., 2015). Different approaches have defined different, albeit overlapping, sets of *XIST* interacting proteins (Chu et al., 2015; McHugh et al., 2015). Notably, multiple studies have identified SHARP (SMART/HDAC1-Associated Repressor Protein; also known as Spen) as an *XIST* binding partner with a critical role in recruiting the nuclear corepressor SMRT and the histone deacetylase HDAC3 to remove activating histone acetylation marks from the X-chromosome (Chu et al., 2015; McHugh et al., 2015). However, a recent study demonstrated that HDAC3 is not essential to the establishment of XCI (Zylicz et al., 2019), raising questions about the role of the *XIST*-SHARP interaction in gene repression. While study of *XIST* dates back to the beginning of the lncRNA field itself (Brown et al., 1991), our understanding of *XIST*'s mechanism of action is still expanding and evolving, 30 years later. This highlights the complexity of lncRNA mechanisms and the importance of employing multiple orthogonal approaches in lncRNA functional characterization.

The lncRNA *Morrbid* operates by a similar general mechanism of action, although with a far more restricted regulatory output than *XIST*. Expressed exclusively in a subset of myeloid cells, *Morrbid* influences cellular lifespan by repressing the pro-apoptotic gene *Bim in cis* (Kotzin et al., 2016). This negative *cis*-regulation is mediated by *Morrbid* recruitment of PRC2 to the *Bim* promoter facilitated by chromatin contacts between the neighboring *Morrbid* and *Bim* loci. *XIST* and *Morrbid* exemplify an archetype of *cis*-acting lncRNAs, namely those that interact with chromatin-modifying complexes (either activating or repressive) to engage in epigenetic regulation of target gene(s). Many lncRNAs, *cis*-regulatory

or otherwise, execute their functions through their association with various RBPs (*reviewed in* (Statello et al., 2020)). However, lncRNA-mediated *cis*-regulation does not always require RNA-protein interactions, or indeed the RNA molecule itself. A notable example of this is *Airn* (*Antisense Igf2r ncRNA*), a lncRNA oriented overlapping and antisense to *Igf2r* (*Insulin-like growth factor receptor 2*); both genes are encoded in a paternally imprinted gene cluster that also includes *Slc22a2* and *Slc22a3* (Statello et al., 2020). *Airn* expression from the paternal allele, specifically transcription through the *Igr2r* promoter region, is required for *Igf2r* silencing via transcriptional interference, a function which does not require the *Airn* RNA molecule, only its production (Latos et al., 2012; Sleutels et al., 2002). In this way, the placement and architecture of the *Airn* locus confers its entire *Igf2r*-repressing function. However, the *Airn* RNA transcript has been shown to play a role in epigenetically repressing *Slc22a3* through the recruitment of a histone methyltransferase G9a, revealing distinct transcript- and transcription- based mechanisms (Nagano et al., 2008).

These difficulties in elucidating the mechanism by which *Airn* represses genes in the *Igf2r* locus emphasize a key and pervasive challenge in the lncRNA field; specifically, how can we accurately discern the element(s) of a lncRNA locus required for its function(s)? Gene regulation by *cis*-acting lncRNAs can be mediated by **(1)** DNA elements in the locus, **(2)** the act of transcription or RNA processing, **(3)** the RNA molecule itself, or some combination thereof (Figure 1) (Gil and Ulitsky, 2020). Employing either deletion of a lncRNA locus or insertion of a premature polyadenylation signal downstream of its transcriptional start site (TSS) can aid in dissociating the contributions of DNA elements in the locus from

both the act of transcription and the activity of the RNA molecule (Engreitz et al., 2016; Paralkar et al., 2016). However, differentiating between the interconnected roles of transcription and the RNA transcript has proven a more thorny problem. Some experimental tools may be too intractable or imprecise to effectively isolate one lncRNA feature from the other and can produce muddled results. For example, recent work has demonstrated that antisense oligonucleotides (ASOs), often a preferred choice for RNA depletion in lncRNA studies (especially for nuclear-enriched transcripts), can cause premature transcription termination and therefore do not constitute a loss-of-function (LOF) model specific to the RNA molecule (Lee and Mendell, 2020). Given such shortcomings in current technologies, necessity dictates that we continue to develop diverse and innovative experimental tools and approaches with lncRNA biology in mind.

Cis-activating lncRNAs can be especially difficult to functionally dissect, as their transcription alone may be responsible for their ascribed regulatory outputs. Indeed, the transcription of protein-coding genes, like lncRNAs, has been shown to activate the expression of neighboring loci (Engreitz et al., 2016), perhaps because active transcription increases interactions mediated by *cis*-regulatory elements, thus supporting promoter-enhancer contacts (Gu et al., 2018). These observations may point to a widespread mechanism of transcription-facilitated *cis*-activation that is not specific to lncRNAs. Interestingly, many *cis*-activating lncRNAs are transcribed from enhancers and regulate neighboring genes by modulating enhancer availability through the act of their transcription or recruitment of the Mediator complex (Isoda et al., 2017; Lai et al., 2013). Yet, the role of other enhancer RNAs (eRNAs), such as the RNAs produced from p53-

bound enhancer regions (p53BERs), is less clear, and may depend only on DNA elements to elicit gene expression changes (Melo et al., 2013). Other lncRNAs are not transcribed from enhancers per se, but may have enhancer elements associated with their loci. *LincRNA-p21*, for example, has been proposed to activate the transcription of its neighboring gene p21 through *cis*-regulatory DNA elements (Groff et al., 2016), despite other studies demonstrating a role for the RNA molecule (Dimitrova et al., 2014; Huarte et al., 2010). Controversies such as these highlight a need for robust characterization of *cis*-acting lncRNAs and clear identification of the elements of lncRNA loci required for their function(s).

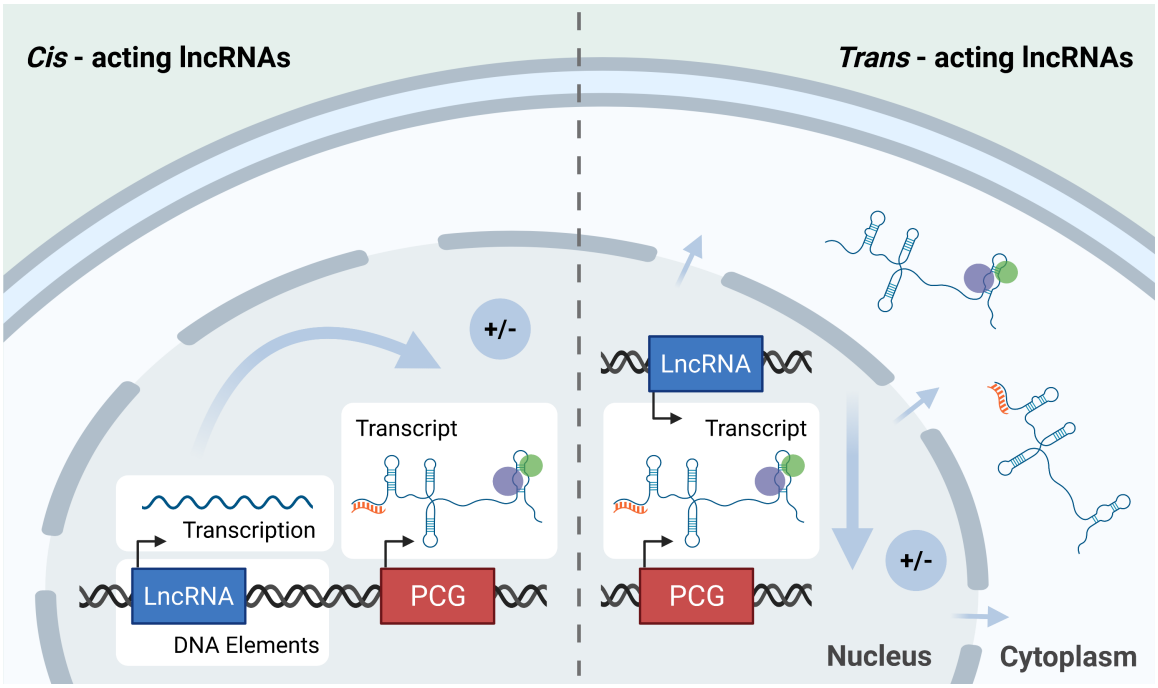

Figure 1. *Cis*- and *trans*- acting mechanisms of action by lncRNAs. Long noncoding RNAs (lncRNAs) can function in either *cis* or *trans*. *Cis*-acting lncRNAs are nuclear-enriched and can regulate neighboring protein-coding genes (PCGs) through DNA elements in the lncRNA locus, the act of transcription, or the RNA transcript. Functions of the lncRNA transcript (blue) are mediated through interactions with target genes (red), RNA binding proteins (green and purple), or other RNAs (orange). *Trans*-acting lncRNAs can reside in the nucleus and regulate target genes expressed from different chromosomes in a similar manner, or can reside in the cytoplasm and engage in RNA-RNA and RNA-protein interactions to influence various cellular processes.

In contrast to *cis*-acting lncRNAs, ***trans*-acting lncRNAs** can travel far from their site of transcription, and either regulate genes elsewhere in the nucleus or are exported to the cytoplasm, enabling a range of other activities (Figure 1) (*reviewed in* (Statello et al., 2020)). A notable example of a nuclear-enriched *trans*-acting lncRNA is *HOTAIR* (*HOX Transcript Antisense RNA*), expressed antisense to the *HOXC* gene cluster and proposed to epigenetically repress the distally located *HOXD* cluster via PRC2 recruitment (Rinn et al., 2007). However, this *trans*-acting mechanism has been called into question by a study noting nonspecific interactions between PRC2 and RNA (Davidovich et al., 2013). Additionally, the function of *Hotair* in the mouse has come under scrutiny after conflicting results were obtained using the same locus deletion model. One group documented homeotic transformations in response to *Hotair* loss (Li et al., 2013), while another group found no developmental defects whatsoever, and therefore concluded *Hotair* was dispensable for mouse development (Amandio et al., 2016). Despite these inconsistencies, there is functional evidence for *HOTAIR* overexpression in supporting breast cancer metastasis, highlighting a need for improved mechanistic elucidation (Gupta et al., 2010).

Firre (*Functional intergenic RNA repeat element*), in contrast to other *trans*-acting nuclear-enriched lncRNAs, resides close to its site of transcription on the X-chromosome, but engages in various *trans*-chromosomal contacts in order to spatially concentrate distal genomic elements (Hacisuleyman et al., 2014). While the function of *Firre* was not initially clear, recent studies have proposed roles in hematopoiesis and XCI, potentially involving both *trans*- and *cis*-regulatory activities (Fang et al., 2020; Lewandowski et al., 2019). This unique

example illustrates that lncRNA location is not always a proxy for its function and that localization patterns *in cis* and *trans* may leave space for nuanced and unexpected mechanisms of action.

Examples of *trans*-acting lncRNAs that function in the cytoplasm include *NORAD* (*Noncoding RNA Activated by DNA Damage*), which binds PUMILIO proteins to support genomic stability (Lee et al., 2016); *LAST* (*LncRNA-Assisted Stabilization of Transcripts*), which stabilizes the mRNA encoding Cyclin D1 (*CCND1*) to promote cellular proliferation (Cao et al., 2017); and *SAMMSON* (*Survival Associated Mitochondrial Melanoma Specific Oncogenic Noncoding RNA*), which associates with the mitochondrial regulator p32 to increase its localization and function in mitochondrial homeostasis (Leucci et al., 2016).

On the whole, *cis*-acting lncRNA mechanisms appear to be more prevalent than *trans*-acting. This idea is supported by enrichment of the majority of lncRNAs in the chromatin fraction, the low copy number of some lncRNAs, and the conservation of lncRNA genomic organization (or synteny) despite poor sequence conservation, all of which favor local *cis*-regulatory lncRNA modes of action (*reviewed in* (Gil and Ulitsky, 2020)). These observations in no way preclude abundant *trans*-acting lncRNA mechanisms, and indeed some lncRNAs have been suggested to regulate in both *cis* and *trans*. For example, in this work I describe a *cis*-acting role for an isoform of the lncRNA *Pvt1* in negatively regulating the expression of a neighboring protein-coding gene (see **Chapters 2-5)**. However, various *trans*-regulatory mechanisms for *Pvt1* have been previously described, including a role in protein stabilization (Tseng et al., 2014), and one in which a circular form of *Pvt1* (circPVT1) acts as a miRNA decoy (Panda et al., 2017).

Moreover, there is disagreement even with respect to the *cis*-activity of *Pvt1*, and whether it depends on the RNA or DNA elements in the locus (Cho et al., 2018). Such distinct regulatory functions described for *Pvt1* and other lncRNAs may be cell-type specific and in keeping with observed differences in localization patterns. Another exciting possibility is that they reflect *bona fide* isoform-specific functions that become important in different cellular contexts, underscoring a need for further exploration.

Long noncoding RNAs in cancer

The work in this section has been published as part of the following invited review: Olivero, C., and Dimitrova, N. Identification and characterization of functional long noncoding RNAs in cancer. (2020). The FASEB Journal 34, 15630-15646.

Introduction

Cancer is a disease of aberrant cell growth arising from a complex genetic landscape of inherited and sporadic mutations and environmental factors. Historically, cancer research has prioritized examining alterations to protein-coding genes in molecular pathways influencing the hallmarks of cancer (Hanahan and Weinberg, 2000, 2011). While these analyses have provided extensive insights into key players in tumorigenesis, protein-coding sequences account for only 2% of the genome (International Human Genome Sequencing, 2004). Both the pervasive transcription of the human genome (Djebali et al., 2012) and the

presence of cancer- associated mutations in noncoding regions (Freedman et al., 2011) have suggested a potential wealth of unexplored cancer targets. Notably, the heterogeneous class of long noncoding RNAs (lncRNAs) occupies a significant space within the noncoding transcriptome, with recent estimates suggesting the existence of over 100,000 human lncRNA transcripts (Bertone et al., 2004; Carninci et al., 2005; Consortium, 2012; Kapranov et al., 2007).

LncRNAs are operationally defined as RNA molecules exceeding 200 nucleotides in length that lack protein-coding potential (Mercer et al., 2009; Rinn and Chang, 2012). Able to dynamically fold into intricate secondary structures (Qian et al., 2019) to interact with DNA, proteins and other RNAs, lncRNAs are diverse in their structure, localization, and pattern of expression, enabling them to regulate the flow of cellular information at many levels (Wang and Chang, 2011). Frequently the targets of transcriptional programs, lncRNAs influence many fundamental cellular processes including cell division, genome maintenance, and pluripotency (Lee et al., 2016; Loewer et al., 2010; Yap et al., 2010).

As lncRNAs are expressed with exquisite cell-type and disease-state specificity, they are ideally positioned to act as biomarkers for a number of pathologies, including different cancers (Derrien et al., 2012; Iyer et al., 2015; Yan et al., 2015). Identifying lncRNA expression changes, or their association with recurrent copy number variations (CNVs) or cancer susceptibility single nucleotide polymorphisms (SNPs) have the potential to become useful tools in cancer diagnosis and treatment planning. Beyond their diagnostic and prognostic utility, over the past decade, individual lncRNAs have been mechanistically and functionally dissected, revealing critical roles in cancer-related pathways at the

cellular and organismal level. These studies have pointed to lncRNAs as operators within proto-oncogenic and tumor suppressive networks, suggesting that lncRNAs themselves may play active roles in promoting or limiting tumor development (Huarte, 2015; Prensner and Chinnaiyan, 2011; Wapinski and Chang, 2011).

Despite growing data supporting the involvement of lncRNAs in tumorigenesis, it is often difficult to surmise whether changes in individual lncRNAs are *bona fide* drivers of human cancer development and whether targeting altered lncRNAs in patients would be expected to produce therapeutic benefit. Here, we present an overview of how functional lncRNAs in cancer are identified. We highlight promising therapeutic targets based on patient data and on experimental evidence from *in vitro* and *in vivo* cancer models. We also discuss important discrepancies to suggest a best-practice roadmap for further characterization of the roles of lncRNAs in cancer.

Identification of cancer-associated lncRNAs

Mining global human cancer genomic and transcriptomic data

Integrating genomic and transcriptomic data from diverse human cancers has provided a starting point for the identification of lncRNAs with functional roles in cancer. In particular, recurrent genetic alterations have implicated many genes involved in oncogenesis, and the capacity to identify such genes has expanded in the last several years due to rapid advances in sequencing technologies. These studies have uncovered that many recurrent somatic copy number variations (SCNVs) map to noncoding regions (Beroukhim et al., 2010). Notably, analysis of 5000 human tumor samples across 13 cancer types from The Cancer Genome Atlas

(TCGA) revealed that, on average, as many as one quarter of all lncRNAs manifest frequent cancer-related copy number gains or losses (Yan et al., 2015). A more recent study probed the copy number of over 10,000 lncRNAs in 80 cancer cell lines across 11 cancer types, identifying 136 lncRNAs involved in focal SCNVs (Volders et al., 2018). Importantly, 76 of these lncRNAs lacked copy number changes in flanking protein-coding genes, suggesting potential lncRNA-driven genomic alterations in cancer. Cancer risk SNPs in noncoding loci can also point to a potential role for specific lncRNAs in tumorigenesis. One study identified nearly 4000 lncRNAs overlapping disease-associated SNPs, while another estimated that roughly 12% of all cancer-associated SNPs mapped within 5 Kb of lncRNA loci (compared to 55% mapping near protein-coding genes) (Iyer et al., 2015; Yan et al., 2015).

Apart from harboring genomic alterations, lncRNAs have also been found to exhibit differential expression patterns in tumor samples compared to normal tissues. A comprehensive meta-analysis of over 7000 gene expression datasets, including a range of normal and cancer samples, identified as many as 60,000 lncRNAs with altered expression (Iyer et al., 2015). Notably, many previously unannotated lncRNAs were found in disease-associated regions and the expression of roughly 8000 lncRNAs clustered with specific cancer or cell lineages, suggesting the potential for lncRNAs to execute cancer-specific functions (Iyer et al., 2015). Along similar lines, an analysis of seven cancer types revealed that, on average, 26% of expressed lncRNAs were significantly deregulated in at least one cancer type (15% upregulated and 11% downregulated) with 60% of these altered lncRNAs demonstrating cancer specificity (Yan et al., 2015). In addition, a recent study of

lncRNA-associated epigenetic alterations across 20 different cancers identified over 2000 lncRNAs either epigenetically activated or silenced in at least one cancer type (Wang et al., 2018). Altogether, these studies led to the consensus that, as a class, lncRNAs are subject to frequent genetic and epigenetic alterations in cancer.

LncRNA loci with recurrent SCNVs in cancer

In addition to global patterns of lncRNA deregulation in cancer, several individual lncRNAs have been identified based on frequent large-scale genomic alterations. One of the first cancer-associated lncRNAs was identified in murine lymphomas due to the frequent translocations and viral insertions involving the as-yet uncharacterized *Pvt1* (Plasmacytoma Variant Translocation 1) lncRNA (Cory et al., 1985; Graham et al., 1985), located approximately 72 Kb downstream of the *MYC* (Myelocytomatosis) proto-oncogene. Later studies extended these results to human cancer and demonstrated a correlation between *PVT1* genomic amplification and poor prognosis in acute myeloid leukemia and in breast and ovarian cancers, among others (*reviewed in* (Colombo et al., 2015)). Significantly, *PVT1* amplification is observed frequently in a range of cancer types including in 33% of ovarian cancers, 20% of esophageal cancers, 13% of invasive breast carcinomas and 7% of lung adenocarcinomas based on TCGA data (Hoadley et al., 2018). Moreover, *PVT1* alterations are associated with a significant reduction in overall and disease-free survival (Cerami et al., 2012; Gao et al., 2013; Hoadley et al., 2018).

Another prominent example of a lncRNA initially characterized by genomic alterations is *FAL1* (Focally Amplified LncRNA 1, also known as *FALEC*) located

on chromosome 1q21 (Hu et al., 2014a). *FAL1* copy number gains have been observed across many cancer types, including in approximately 10% of liver cancers, invasive breast carcinomas and lung adenocarcinomas according to TCGA data (Hoadley et al., 2018). *FAL1* amplification and overexpression are associated with late stage tumors and with decreased survival of patients with ovarian cancer (Hoadley et al., 2018; Hu et al., 2014a). Similarly, the lncRNA *SAMMSON* (Survival Associated Mitochondrial Melanoma Specific Oncogenic Noncoding RNA) was identified in a region of focal amplification on chromosome 3p13-14 in 10% of melanomas (Leucci et al., 2016). High *SAMMSON* copy number and expression levels are correlated with a reduction in disease-free survival of melanoma patients and associated with resistance to MAPK (Mitogen Activated Protein Kinase) inhibitors (Cerami et al., 2012; Gao et al., 2013; Hoadley et al., 2018; Leucci et al., 2016).

The locus of the lncRNA *LOC285194* on chromosome 3q, on the other hand, is subject to recurrent monoallelic deletions in as many as 80% of osteosarcomas, often followed by loss of heterozygosity (LOH) (Pasic et al., 2010). Loss of *LOC285194* is associated with decreased survival in osteosarcoma patients (Pasic et al., 2010). The focal deletion of *PRAL* (p53 Regulation-Associated LncRNA) on chromosome 17p in hepatocellular carcinoma has also been associated with reduced survival (Zhou et al., 2016). Similarly, recurrent loss of the 9p21 locus, where the lncRNA *ANRIL* (Antisense Noncoding RNA in the INK4 Locus) resides, is observed in over 50% of glioblastomas, more than 40% of mesotheliomas, and roughly 30% of bladder cancers (Hoadley et al., 2018). Interestingly, a 403 Kb

germline deletion encompassing the *ANRIL* locus is associated with a strong hereditary predisposition to melanoma development (Pasmant et al., 2007).

Many regions of recurrent SCNVs, however, harbor both lncRNAs and protein-coding genes. Therefore, determining the specific contribution of the lncRNA has been challenging. For example, the presence of multiple overlapping transcripts in the *ANRIL* locus, including the p15INK4B (*CDKN2B*), p16INK4A (*CDKN2A*), and p19ARF tumor suppressors, has confounded the role of *ANRIL* (Yap et al., 2010). Analogously, *PVT1* is co-amplified with *MYC* and the *PVT1* gene body contains DNA regulatory elements, which promote *MYC* expression (Fulco et al., 2016). Likewise, *SAMMSON* is expressed near *MITF* (Microphthalmia Associated Transcription Factor), a key factor in melanocyte differentiation, whereas the commonly amplified genomic region in which *FAL1* resides contains the proto-oncogene *MCL1* (Myeloid Cell Leukemia Sequence 1). Finally, the *LOC285194*-associated region of deletion also harbors the tumor suppressor *LSAMP* (Limbic System-Associated Membrane Protein). Given the complex chromatin architecture and transcriptional profiles in these loci, further studies are needed to deconvolve the specific roles of the lncRNAs and to determine whether lncRNAs act in cooperation with or independently of their neighboring protein-coding genes.

LncRNA loci with cancer-associated SNPs

The link between inherited germline variants in lncRNA loci and cancer predisposition or prognosis has been probed extensively in large-scale genome-

wide associated studies (GWAS). These studies have identified a plethora of lncRNA-linked SNPs associated with altered cancer risk or patient prognosis.

As an example, the 2 Mb region mapping to 8q24 has emerged as a major hotspot for over a hundred SNPs strongly associated with multiple diseases, including cancers of the breast, colon, ovaries, prostate, and bladder (Easton and Eeles, 2008; Ghoussaini et al., 2008; Grisanzio and Freedman, 2010; Huppi et al., 2012). Many of these SNPs are significantly correlated with cancer development and highly predictive of poor patient outcome (Bertucci et al., 2012; Garcia-Closas et al., 2008; Haiman et al., 2007; Yeager et al., 2007; Zhang et al., 2012b). While *MYC* is the dominant oncogene in the locus, many of the cancer risk SNPs are linked to the expression of lncRNAs in the surrounding region, including *PVT1* (Meyer et al., 2011), *CCAT1* (Colon Cancer Associated Transcript 1, also known as *CARLo-5*) (Zhao et al., 2016), *CCAT2* (Colon Cancer Associated Transcript 2) (Ling et al., 2013), *PCAT1* (Prostate Cancer Associated Transcript 1) (Guo et al., 2016), *PCAT19* (Prostate Associated Transcript 19) (Hua et al., 2018), and *PRNCR1* (Prostate Cancer Associated Noncoding RNA 1) (Huang et al., 2018b). The *ANRIL* locus is another example of a hotspot harboring more than 10 cancer risk SNPs, some of which are correlated with *ANRIL* expression (Cunnington et al., 2010; Khorshidi et al., 2017). Other lncRNAs linked to cancer SNPs include *HOTAIR* (HOX Transcript Antisense RNA) (Botti et al., 2018), *HOTTIP* (HOXA Distal Transcript Antisense RNA) (Huang et al., 2018b), *MALAT1* (Metastasis-Associated Lung Adenocarcinoma Transcript 1) (Huang et al., 2018b), *HULC* (Highly Upregulated in Liver Cancer) (Huang et al., 2018b), *MEG3* (Maternally Expressed 3) (Dong et al., 2020), *H19* (Hashemi et al., 2019), *GAS5* (Growth Arrest Specific

5) (Dong et al., 2020), and *PTENP1* (Phosphatase And Tensin Homolog Pseudogene 1) (Ge et al., 2017).

Mechanistic investigations of SNPs associated with lncRNAs have suggested that the risk variants may, in some cases, affect regulatory DNA sequences, thereby resulting in altered lncRNA expression. For example, the *PCAT1*-linked risk variant rs7463708 was found to increase the activity of a distal enhancer, resulting in increased *PCAT1* expression (Guo et al., 2016), whereas the *PCAT19*-linked SNP rs11672691 was proposed to perturb transcription factor binding sites, resulting in the increased expression of a pro-metastatic *PCAT19* isoform (Gao et al., 2018; Hua et al., 2018). Finally, a high-risk neuroblastoma associated SNP rs693940 on chromosome 6p22 was found to contribute to differential CpG methylation and decreased expression of *NBAT-1* (Neuroblastoma Associated Transcript-1, also known as *CASC14*), a lncRNA with tumor suppressor properties (Pandey et al., 2014). Apart from these intriguing examples, however, the majority of lncRNA-associated SNPs lack experimental support that would robustly link the cancer-susceptibility variants with deregulation of lncRNA levels or function, and have thus had limited impact on the identification and characterization of functional lncRNAs in cancer.

LncRNAs differentially expressed in cancer

Global gene expression analyses of normal and cancer samples have also led to the identification of numerous differentially expressed lncRNAs hypothesized to contribute to disease development. Some of the initial analyses revealed frequent upregulation of lncRNAs, such as the imprinted lncRNA *H19* in Wilms'

tumors and lung cancer (Hibi et al., 1996; Kondo et al., 1995; Rainier et al., 1993), the prostate cancer-specific lncRNA *PCGEM1* (Prostate Cancer Gene Expression Marker 1) (Srikantan et al., 2000), the lung metastasis-promoting lncRNA *MALAT1* (Ji et al., 2003) and the hepatocellular carcinoma overexpressed lncRNA *HULC* (Panzitt et al., 2007).

The differential expression of some of these lncRNAs has been associated with clinical outcomes. For example, altered *H19* expression correlates with poor clinical outcomes across various cancer types including breast cancer, non-small cell lung cancer and acute myeloid leukemia (Shima et al., 2018; Zhang et al., 2018b; Zhou et al., 2017). On the other hand, increased expression of *PCGEM1* in normal prostate tissue is a prostate cancer risk factor (Petrovics et al., 2004; Srikantan et al., 2000). At the same time, a large body of literature has cemented the strong correlation between high *MALAT1* expression levels and poor patient prognosis across over 20 cancer types (Amodio et al., 2018; Zhang et al., 2015). Finally, high expression of *HULC* is associated with poor overall survival and distant metastases (Chen et al., 2017).

Notably, integrated analysis of gene expression and methylation datasets has also led to the identification of differentially expressed lncRNAs arising from cancer-associated epigenetic changes, including *AFAP1-AS1* (AFAP1 Antisense RNA 1) and *EPIC1* (Epigenetically Induced LncRNA1), both identified as hypomethylated and overexpressed in Barrett's esophagus and esophageal adenocarcinoma, and breast cancer, respectively (Wang et al., 2018; Wu et al., 2013).

Gene expression profiling in cohorts of cancer patients have further fueled the discovery of lncRNAs associated with specific cancer types. Transcriptome sequencing across a cohort of prostate cancer patients identified *PCAT-1* amongst 121 unannotated prostate cancer-associated ncRNA (noncoding RNA) transcripts (*PCATs*) (Prensner et al., 2011). Similarly, comprehensive lncRNA profiling in colorectal carcinoma led to the identification of *CCAT1* (Kim et al., 2014b; Nissan et al., 2012), *CCAT2* (Ling et al., 2013) and other *CCAT* family members (Kim et al., 2015b), whereas the lncRNA *GAPLINC* (Gastric Adenocarcinoma Predictive Long Intergenic Noncoding RNA) stood out as aberrantly overexpressed in gastric tumors (Hu et al., 2014b). A different set of analyses led to the identification of stage-specific lncRNAs, such as the lncRNA *CRNDE* (Colorectal Neoplasia Differentially Expressed) (Graham et al., 2011), a marker of early stages of colorectal cancer development, although the protein-coding capacity of *CRDNE* remains an open question (Szafron et al., 2015). Transcriptome profiling of breast cancer subtypes, on the other hand, highlighted sets of lncRNAs which are either differentially expressed in tumor samples compared to normal tissues or uniquely enriched in specific stages or subtypes of breast cancer. Examples include *MALAT1* (Arun et al., 2016; Jadaliha et al., 2016), *HOTAIR* (Gupta et al., 2010), and *BCAR4* (Breast Cancer Anti- Estrogen Resistance 4) (Meijer et al., 2006; Xing et al., 2014). In parallel, mouse models of cancer were recently employed for the identification of 30 murine *MaTARs* (Mammary Tumor Associated RNAs), many of which were found to have human counterparts (*hMaTARs*) with potential clinical significance determined based on differential expression and correlation with cancer subtype and/or hormone receptor status (Diermeier et al., 2016). Interestingly, many of

these examples of cancer-specific lncRNAs were later found to show differential expression across multiple cancer types, hinting at universal roles in cancer pathogenesis.

LncRNAs in cancer pathways

In addition to profiling tumor samples, many researchers have undertaken diverse functional approaches to identify novel lncRNAs, including dissecting tumor suppressive and pro-oncogenic transcriptional networks, analyzing various cancer-related cellular states and processes, and performing genome-wide functional screens.

Analysis of the p53 (also known as Trp53) transcriptional network, in particular, has revealed a wealth of lncRNAs with potential tumor suppressor functions. By comparing gene expression profiles and p53 binding patterns in the absence and in the presence of genotoxic or oncogenic stress, known to activate the p53 pathway, as well as in p53-proficient and -deficient cells, researchers have identified multiple direct lncRNA targets of p53. These included *lincRNA-p21* (Huarte et al., 2010); *PANDAR* (Promoter Of CDKN1A Antisense DNA Damage Activated RNA, also known as *PANDA*) (Hung et al., 2011); p53BERs (p53-Bound Enhancer Regions) (Melo et al., 2013); *Pint* (P53 Induced Transcript) (Marin-Bejar et al., 2013); *LED* (LncRNA Activator of Enhancer Domains) (Leveille et al., 2015); *PR-lncRNAs* (p53-Regulated lncRNAs) (Sanchez et al., 2014; Younger et al., 2015); *DINO* (Damage Induced Noncoding) (Schmitt et al., 2016); *lncPRESS1* (LncRNA P53 Regulated And ESC Associated 1) (Jain et al., 2016); *NEAT1* (Nuclear Enriched Abundant Transcript 1) (Adriaens et al., 2016; Blume et al., 2015; Mello et al.,

2017); *PURPL* (P53 Upregulated Regulator Of P53 Levels) (Li et al., 2017); *PINCR* (P53-Induced Noncoding RNA) (Chaudhary et al., 2017); *GUARDIN* (Hu et al., 2018); and an isoform of *Pvt1*, *Pvt1b* (Olivero et al., 2020). Functional characterizations have suggested that many of these lncRNAs contribute to p53 tumor suppressor activities.

Other lncRNAs have been identified downstream of oncogenic signaling networks, giving insight into their potential functions. For example, *Orilnc1* (Oncogenic RAS-Induced lncRNA 1) was identified as a target of oncogenic RAS signaling with a proposed role in promoting cell growth (Zhang et al., 2017). *LncRNA-OIS1* (Oncogene- Induced Senescence 1) was found to modulate senescence induced by activation of oncogenic RAS (Li et al., 2018), whereas *BANCR* (BRAF-Activated Non-Protein Coding RNA) was identified as a transcript induced upon expression of oncogenic BRAFV600E (Flockhart et al., 2012). Analogously, investigation of estrogen receptor (ER) signaling targets identified 33 ER agitation-related (ERAR) lncRNAs and suggested potential roles in ER-positive breast cancer (Wu et al., 2016). A similar study was performed to examine lncRNAs regulated by androgen receptor (AR) signaling, which identified *ARLNC1* (AR-Regulated Long Noncoding RNA 1) as both a downstream target and upstream effector of AR signaling during prostate cancer progression (Zhang et al., 2018c). MYC-regulated lncRNAs have also been identified, including a set of *MYCLos* (MYC-regulated lncRNAs) (Kim et al., 2015b); *LAST* (LncRNA-Assisted Stabilization of Transcripts) (Cao et al., 2017); *DANCR* (Differentiation Antagonizing Non-Protein Coding RNA) (Lu et al., 2018), and *SNHG15* (Small Nucleolar RNA Host Gene 15) (Jiang et al., 2018).

Alterations of cancer hallmarks that enable tumorigenesis have also been linked to the functions of specific lncRNAs (*reviewed in* (Gutschner and Diederichs, 2012)). Examples include lncRNA *gadd7* (growth-arrested DNA damage-inducible gene 7) with a proposed role in suppressing cell cycle progression (Liu et al., 2012), *SPRY4-IT1* (SPRY4 Intronic Transcript 1) with a proposed role in inhibiting apoptosis in melanoma (Khaitan et al., 2011), and *SALNR* (Senescence-Associated lncRNA), proposed to regulate senescence (Wu et al., 2015).

Finally, genome-wide functional screens for lncRNAs involved in promoting or inhibiting specific cellular outcomes important in cancer have aimed to identify candidates for further study. A CRISPR/Cas9-based genome editing approach used a paired guide RNA (gRNA) strategy to target for deletion a set of 700 human lncRNAs, identifying 51 lncRNAs able to regulate cancer cell growth (Zhu et al., 2016). Alternatively, CRISPRi (CRISPR inactivation) and CRISPRa (CRISPR activation) screens, involving a nuclease-dead Cas9 to tether transcriptional repressors or activators to lncRNA loci have provided effective epigenetic loss-of-function and gain-of-function approaches to query on a genome-wide level the role of lncRNAs in processes such as cellular proliferation or therapeutic resistance (Bester et al., 2018; Joung et al., 2017; Liu et al., 2017; Liu et al., 2020).

Functional characterization of lncRNAs in cancer

Common approaches and limitations

For the hundred or so lncRNAs identified in the approaches described above, the pressing question has become how to accurately distinguish functional lncRNAs from lncRNAs that are subject to passenger genetic and epigenetic alterations in cancer. RNA interference (RNAi)-mediated downregulation of lncRNAs has been a common approach for functional characterization. In parallel, antisense oligonucleotides (ASOs) have provided a convenient and efficient loss-of-function alternative. While RNAi is most effective for lncRNAs exported to the cytoplasm, ASOs lend broader efficacy by triggering RNase H-mediated co-transcriptional RNA cleavage and degradation, in some cases accompanied by transcriptional repression (Lai et al., 2020; Lee and Mendell, 2020). Frequently, RNAi and ASO approaches have been performed in parallel with exogenous lncRNA overexpression. Regrettably, few studies have complemented RNAi or ASO loss-of-function experiments with knockdown-resistant lncRNA rescue mutants, missing an important opportunity to both demonstrate specificity and establish a system to investigate the sequence basis for lncRNA function. CRISPR-based epigenetic inhibition (CRISPRi) and activation (CRISPRa) have also been employed as successful loss-of-function and gain-of-function approaches, respectively.

Genetically engineered mouse models (GEMMs) of lncRNAs and CRISPR-based editing of lncRNA loci in cell lines have also brought important insights. In contrast to protein-coding genes, where genetic approaches aim to perturb the open reading frame (ORF) and therefore, the functional output of the transcript, methods to target lncRNAs have been, by necessity, more diverse and creative (*reviewed in* (Bassett et al., 2014)). Some loss-of-function studies have undertaken

deletion of the entire gene body, the promoter region, or narrower functional regions, while others have employed introduction of a premature polyadenylation signal (PAS) or polyadenylation cassette (STOP) to terminate transcription. Conversely, gain-of-function studies in animal models have involved the introduction of a transgenic lncRNA sequence or amplification of an entire lncRNA locus.

Strikingly, for many lncRNAs, observed phenotypes have varied with the use of alternative approaches. For example, initial RNAi knockdown of the p53-regulated lncRNA, *lincRNA-p21*, suggested that it acts globally to modulate the expression of multiple p53 target genes, whereas subsequent genetic deletion of its promoter in the mouse revealed a more restricted role in promoting the expression of the neighboring *p21/CDKN1a* gene (Dimitrova et al., 2014; Huarte et al., 2010). Further investigation involving a locus deletion genetic approach, however, raised doubts about whether the lncRNA plays any functional role at all (Groff et al., 2016). The metastasis-promoting lncRNA *HOTAIR* has provided additional examples of the complexity in developing lncRNA models. While ectopic expression of *HOTAIR* in breast cancer cells induced global gene expression changes and increased metastases in a xenograft mouse model, supporting an oncogenic function (Gupta et al., 2010), loss-of-function models, including RNAi-mediated knockdown, a 4 Kb gene body deletion, and a 140 Kb locus deletion have led to significant discrepancies (Amandio et al., 2016; Li et al., 2013; Rinn et al., 2007; Schorderet and Duboule, 2011). The differences between alternative models have highlighted the need to use multiple independent and complementary approaches to investigate the functional roles of lncRNAs in cancer biology.

Multi-pronged approaches to lncRNA characterization

In this section, we focus on a small set of lncRNAs for which work from multiple groups or involving an array of *in vitro* and *in vivo* approaches has revealed exciting functional insights and provided starting points for further exploration of their contributions to tumor development.

MALAT1

MALAT1 remains one of the most studied cancer-associated lncRNAs, with proposed roles in influencing nuclear speckles (Hutchinson et al., 2007), pre-mRNA splicing (Tripathi et al., 2010), and epigenetically regulating gene transcription (West et al., 2014). While initial studies pointed to a pro-metastatic function (Ji et al., 2003), further characterization resulted in discrepancies (Figure 2). Three different loss-of-function GEMMs, including an insertion of a LacZ reporter and polyadenylation cassette 69 nucleotides downstream of the *Malat1* transcription start site, a 3 Kb deletion of the 5' end and promoter region of *Malat1*, and a conditional deletion of 7 Kb encompassing the entire *Malat1* gene body, revealed that *Malat1* is dispensable for organismal development and viability (Eissmann et al., 2012; Nakagawa et al., 2012; Zhang et al., 2012a). Strikingly, none of the mouse models showed effects on global gene expression, nuclear speckle formation, or alternative pre-mRNA splicing. This opposed previous findings using RNAi to downregulate *MALAT1* levels in cancer cell lines *in vitro* (Tripathi et al., 2010; West et al., 2014), perhaps suggesting a cancer-specific function. Furthermore, different *in vivo* models have yielded conflicting results about the function of *MALAT1* in cancer. On the one hand, crossing the promoter deletion

model (Zhang et al., 2012a) to the MMTV-PyMT (mouse mammary tumor virus-polyomavirus middle T antigen) mouse model of breast cancer resulted in reduced metastases to the lung, without affecting primary tumor burden, an effect largely recapitulated by ASO-depletion of *Malat1 in vivo* (Arun et al., 2016). This pro-metastatic function was also observed in a mouse xenograft model of lung cancer where *MALAT1* knockout human lung tumor cells formed fewer tumor nodules (Gutschner et al., 2013). In this model, targeting *MALAT1* with ASOs after tumor implantation prevented metastasis formation, pointing to *MALAT1* as a viable therapeutic target (Gutschner et al., 2013). On the other hand, crossing the *Malat1* premature termination model (Nakagawa et al., 2012) to the MMTV-PyMT breast cancer model led to a significant increase in the number and area of metastatic nodules in the lungs (Kim et al., 2018). This surprising tumor suppressive effect could be rescued with a *Malat1* transgene expressed from the Rosa26 locus (Kim et al., 2018). A similar effect was observed *in vitro* in human breast cancer cells, with expression of *MALAT1* from an exogenous construct rescuing the increased metastatic ability conferred by *MALAT1* knockout in clonal cell populations (Kim et al., 2018). The debate surrounding the precise contribution of *MALAT1* to cancer development is ongoing. It is unclear whether the phenotypic differences arising from *MALAT1* loss might be due to differences in experimental setup, such as mouse strain or knockout approach, or reflect the complex biology of *MALAT1*. Altogether, investigations of *MALAT1* using *in vitro* and *in vivo* approaches have highlighted the biological and technical complexities associated with studying the functional roles of lncRNAs in cancer (Arun and Spector, 2019; Sun and Ma, 2019).

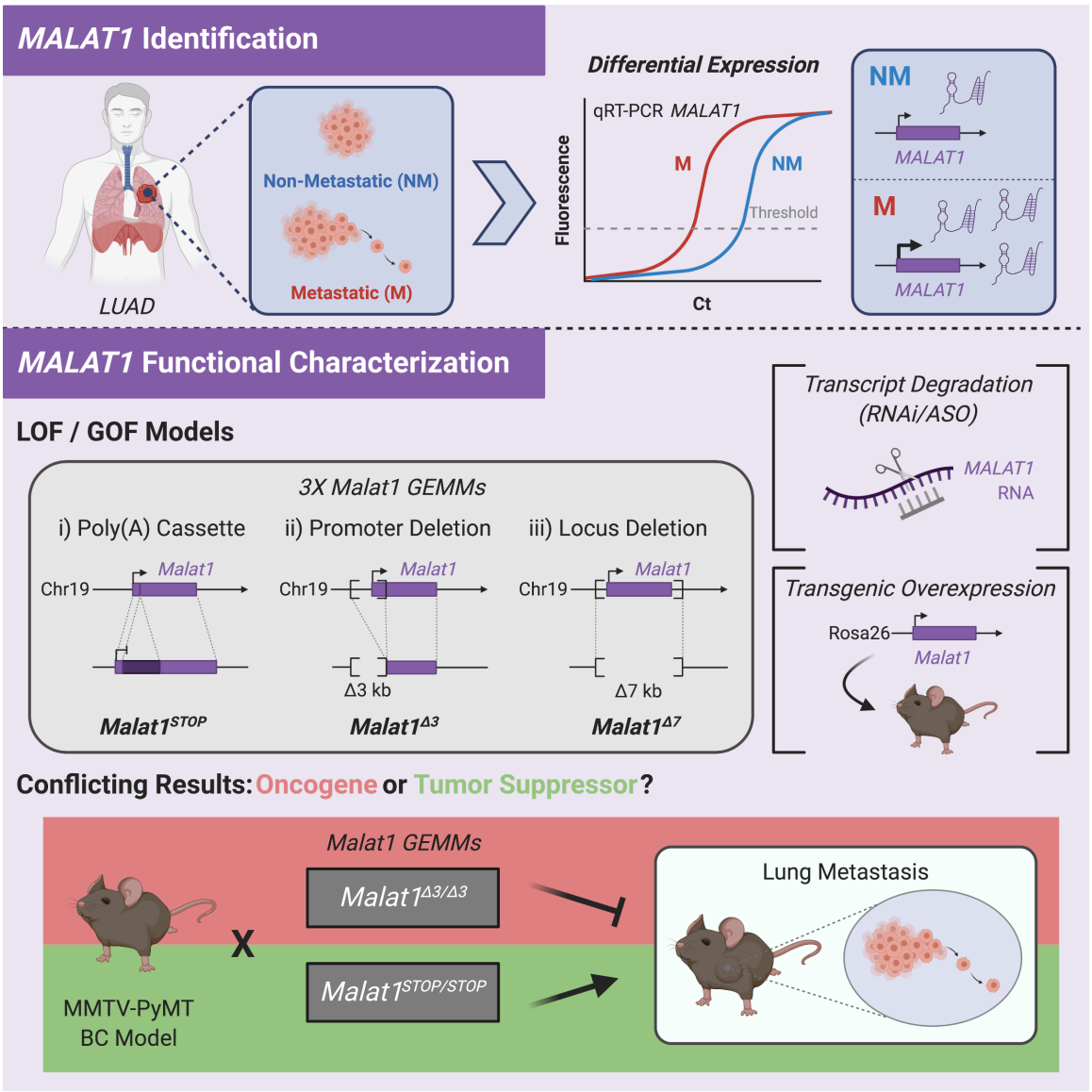

Figure 2. Identification and functional characterization of *MALAT1*. *MALAT1* was identified as upregulated in metastatic (M) LUAD (lung adenocarcinoma) compared to non-metastatic (NM) tissue. Functional characterization of *MALAT1* has utilized various loss-of-function (LOF) and gain-of-function (GOF) models including polyadenylation cassette insertion (*Malat1STOP*, (Nakagawa et al., 2012)), promoter deletion (*Malat1* Δ*3*, (Zhang et al., 2012a)), and locus deletion (*Malat1* Δ*7*, (Eissmann et al., 2012)) genetically engineered mouse models (GEMMs), as well as transcript degradation with RNAi and ASO, and transgenic overexpression. Crossing *Malat1* Δ*3* or *Malat1STOP* GEMMs to the MMTV-PyMT BC (breast cancer) mouse model has resulted in either oncogenic (red box, (Arun et al., 2016)) or tumor suppressor (green box, (Kim et al., 2018)) models for *Malat1* function, due to observed decreases and increases in lung metastases, respectively.

NEAT1

Similarly to *MALAT1*, several studies have examined the role of *NEAT1* during cancer development, leading to opposing views (Figure 3). Initial studies suggested that *NEAT1* levels were elevated in a variety of human cancers relative to normal tissues and correlated with worse prognosis, suggesting a pro-oncogenic role for *NEAT1* ((Chakravarty et al., 2014) and *reviewed in* (Yang et al., 2017)). This conclusion was supported by a study of *Neat1* knockout mice subjected to chemical induction of skin squamous cell carcinoma with the carcinogen DMBA and the pro- inflammatory agent TPA (Adriaens et al., 2016; Nassar et al., 2015). While *Neat1*-deficient animals displayed no obvious phenotypes in the absence of stress (Nakagawa et al., 2011), loss of *Neat1* conferred resistance to chemically-induced squamous cell carcinoma (Adriaens et al., 2016). Interestingly, studies have also suggested that *NEAT1* may be a target of the p53 pathway and, therefore, may have tumor suppressive activities in some contexts (Blume et al., 2015; Idogawa et al., 2017). Indeed, tumor suppressive functions of *Neat1* were unveiled in primary mouse embryonic fibroblasts (MEFs), where *Neat1* knockout led to increased colony formation in an E1A; HrasG12V transformation experiment, as well as in an autochthonous mouse model of pancreatic cancer, where *Neat1* deficiency increased the occurrence of premalignant lesions, known as pancreatic intraepithelial neoplasias (PanINs) (Mello et al., 2017). Interestingly, *Malat1* and *Neat1* are neighboring genes and studies have suggested that genomic deletion of either lncRNA may impact the epigenetic organization and transcriptional profiles of the entire locus, raising questions about the specificity of each approach (Nakagawa et al., 2012).

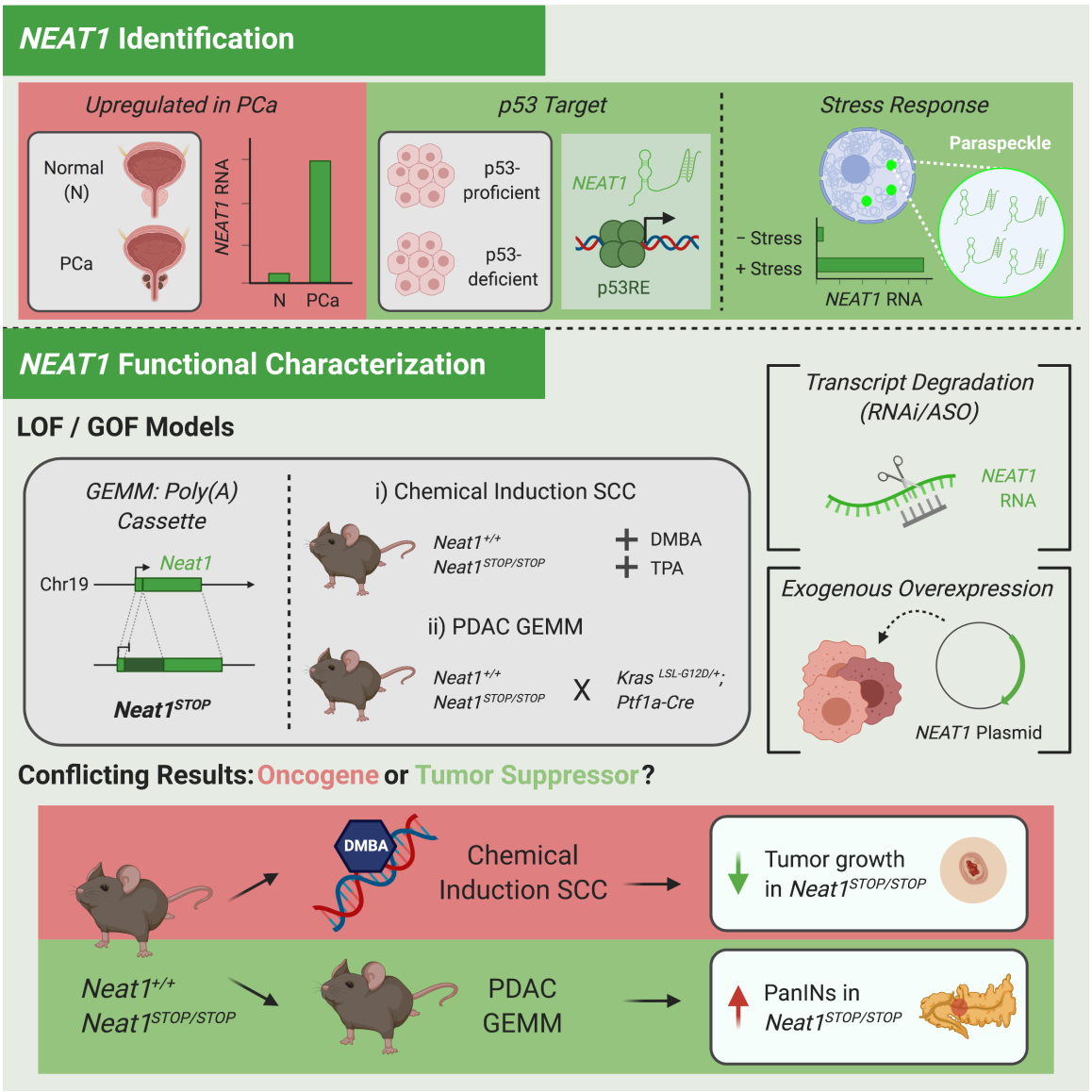

Figure 3. Identification and functional characterization of *NEAT1*. *NEAT1* was initially identified as being upregulated in prostate cancer (PCa) compared to normal (N) tissue, suggesting a potential oncogenic function (top, red box). Later, it was also identified as a p53 target with p53 binding to a conserved p53 Response Element (p53RE) in the *NEAT1* promoter, as well as a paraspeckle component induced by cellular stress, suggesting a potential tumor suppressor function (top, green box). Functional characterization of *NEAT1* has utilized various loss-of-function (LOF) and gain-of-function (GOF) models including a polyadenylation cassette insertion genetically engineered mouse model (GEMM) (Nakagawa et al., 2011), transcript degradation with RNAi or ASO, and exogenous overexpression. The *Neat1STOP* GEMM has been shown to either decrease (Adriaens et al., 2016) or increase (Mello et al., 2017) tumor growth following chemical induction of SCC (squamous cell carcinoma) or when crossed to a PDAC (pancreatic ductal adenocarcinoma) GEMM, respectively, suggesting either oncogenic (bottom, red box) or tumor suppressor (bottom, green box) models for *Neat1* function in cancer.

PVT1

As one of the lncRNAs strongly associated with advanced disease and poor patient prognosis, *PVT1* has been the subject of extensive investigation (Figure 4). In keeping with the finding that *PVT1* is frequently co- amplified with the *MYC* proto-oncogene, *Myc-Pvt1* co-amplification in a mouse model of breast cancer was found to be more tumorigenic than *Myc* amplification alone (Tseng et al., 2014). This study suggested that *PVT1* acts *in trans* to promote MYC protein stability, based on evidence that a 300 Kb genomic deletion of the *PVT1* locus in a human colorectal carcinoma cell line resulted in decreased MYC protein levels (Tseng et al., 2014). However, later studies found evidence for *MYC* enhancers within the region of deletion, raising questions about the role of the *PVT1* locus and its associated RNA in *MYC* regulation (Fulco et al., 2016). Subsequent studies confirmed the presence of DNA regulatory elements in the locus but challenged the understanding of *PVT1* as a strictly pro-oncogenic lncRNA (Cho et al., 2018; Porter et al., 2017). On the one hand, deletion of a ~600 bp region containing a p53 binding site and mapping to the 5' end of *PVT1* led to defects in p53-mediated *MYC* repression, although the contribution of *PVT1* to the p53 response was unclear (Porter et al., 2017). On the other hand, CRISPRi-based inhibition of *PVT1* in breast cancer cell lines revealed a role for the *PVT1* promoter as a DNA tumor suppressor boundary element that limits *MYC* promoter accessibility to enhancers within the *PVT1* gene body, resulting in restricted *MYC* expression (Cho et al., 2018). In this setting, the *PVT1* RNA appeared to be dispensable (Cho et al., 2018). In contrast, our group identified a stress-induced, p53- dependent isoform of *Pvt1*, *Pvt1b*, which is both necessary and sufficient to repress *Myc* transcription (Olivero

et al., 2020). These findings were recapitulated *in vitro* using a genetic loss-of-function approach to mutate the p53 binding site required for *Pvt1b* expression (Olivero et al., 2020). Importantly, mutagenesis of the *Pvt1*-associated p53 binding site at the time of tumor initiation in an autochthonous mouse model of lung cancer led to larger tumors and indicated a key role for *Pvt1b* in restraining tumor growth downstream of p53 (Olivero et al., 2020). In the future, it would be interesting to deconvolve the oncogenic and tumor suppressive elements in the *PVT1* locus and to differentiate between DNA elements and RNA isoforms with potentially distinct functions.

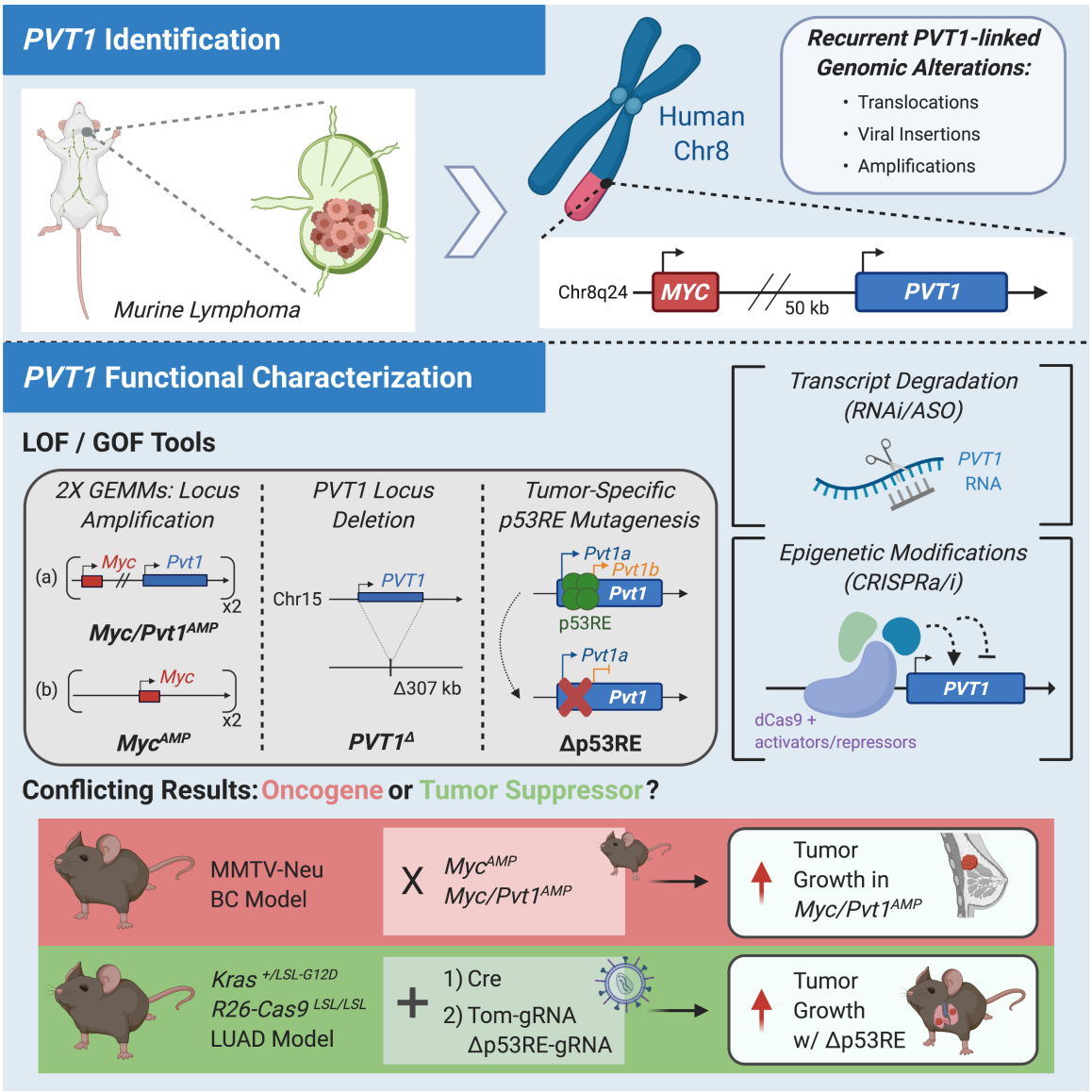

Figure 4. Identification and functional characterization of *PVT1*. *PVT1* was identified in murine lymphomas following the observation of translocations, viral insertions, and amplifications involving the *Pvt1* locus. Functional characterization of *PVT1* has utilized various loss-of- function (LOF) and gain-of-function (GOF) models including amplification genetically engineered mouse models (GEMMs) (*Myc/Pvt1AMP*, *MycAMP*, (Tseng et al., 2014)), locus deletion (*PVT1*), tumor-specific mutagenesis of the *Pvt1*-associated p53 Response Element (p53RE) (*p53RE*, (Olivero et al., 2020)), transcript degradation with RNAi and ASO, and CRISPR-mediated epigenetic activation and inhibition (CRISPRa/i). The increased tumor growth observed in a *Myc/Pvt1* co-amplification GEMM (*Myc/Pvt1AMP*) compared to Myc amplification alone (*MycAMP*) when crossed to the MMTV-Neu BC (breast cancer) GEMM suggests an oncogenic function for *Pvt1* (red box, (Tseng et al., 2014)). However, the increased tumor growth in *Pvt1*-associated p53RE mutagenized lung tumors following Cre-mediated tumor initiation in a *Kras*-driven lung adenocarcinoma (LUAD) GEMM suggests a tumor suppressor function (green box, (Olivero et al., 2020)).

XIST

With a critical role in X chromosome inactivation and dosage compensation that has been investigated for decades (*reviewed in* (Brockdorff et al., 2020; Sahakyan et al., 2018)), the potential role of *XIST* (X Inactive Specific Transcript) in tumorigenesis has intrigued researchers. Historically, it has been observed that altered chromosome copy numbers and inappropriate dosage compensation are frequently associated with human cancer. Notably, men with Klinefelter syndrome characterized by an extra X chromosome have an increased risk of many malignancies including breast cancer and non-Hodgkin lymphoma (Swerdlow et al., 2005), and loss of X chromosome inactivation has been observed in breast cancer cell lines (Sirchia et al., 2005) and testicular germ cell tumors (Kawakami et al., 2003). In support of these correlative observations, a conditional *Xist* deletion model in mouse blood cell lineages led to aggressive myeloproliferative neoplasm and myelodysplastic syndrome with complete penetrance, likely as the result of widespread gene expression changes (Yildirim et al., 2013). The tumor suppressive role of *XIST* was recapitulated in RNAi and overexpression studies in breast cancer cell lines as well as by crossing the *Xist* knockout to the MMTV-PyMT mouse model of breast cancer (Xing et al., 2018). Further studies should determine the prevalence of *XIST* and X inactivation perturbations in human cancer and investigate the possibility of targeting this pathway as a therapeutic strategy.

ANRIL

High ANRIL expression in tumor tissues has been linked to aggressive pathological features and poor overall survival (*reviewed in* (Kong et al., 2018)).

In initial studies, targeted deletion of a 70 Kb region in the *Anril* locus, which harbors multiple cancer and coronary artery disease associated SNPs, led to viable progeny but showed increased mortality during development and as adults (Visel et al., 2010). Primary cultures of smooth muscle cells, isolated from mutant mice, exhibited excessive proliferation and diminished senescence, cellular phenotypes consistent both with accelerated coronary disease pathogenesis and increased cancer risk. Mechanistic investigation revealed that the effects were mediated *in cis* through the reduced expression of *Cdkn2a* and *Cdkn2b* and led to the conclusion that the risk region contained key regulatory elements. Subsequent investigation using exogenous overexpression of *ANRIL* in primary human fibroblasts suggested that the lncRNA may be responsible for *CDKN2A/2B* repression through the locus-specific recruitment of the repressive PRC1 complex (Yap et al., 2010). Unfortunately, little progress has been made over the past decade in determining whether *ANRIL* transcription or transcript accumulation is required for its *cis*-regulatory function, in part due to the limited conservation of *ANRIL* sequence and exonic structure between human and mouse.

Promising lncRNA candidates warranting further investigation

In this section we examine exciting, albeit limited, initial studies of lncRNAs with putative cancer functions, the validation of which could benefit from the development of alternative approaches and further characterization.

SAMMSON

To investigate the role of *SAMMSON* as a lineage addiction oncogene in melanoma, researchers employed ASO-mediated knockdown and exogenous overexpression as loss-of-function and gain-of-function tools (Leucci et al., 2016). They observed that *SAMMSON* amplification and increased expression led to altered mitochondrial metabolism and homeostasis. In turn, this caused increased melanoma cell viability and clonogenic potential and resulted in sensitization of melanoma cells to MAPK targeting therapeutics *in vitro* and in patient-derived xenograft (PDX) models *in vivo*. Further mechanistic studies clarified the role of *SAMMSON* in balancing mitochondrial translation rates (Vendramin et al., 2018). The generation of genetic models of *SAMMSON* may reveal further insights into its role in melanoma development.

NKILA

NKILA (NF-κB interacting long noncoding RNA) was identified as both a target and negative modulator of the NF-κB signaling pathway, with low *NKILA* levels observed in metastatic breast cancer cell lines and correlated with decreased disease-free survival in a cohort of breast cancer patients (Liu et al., 2015). Mechanistically, a series of deletion mutants demonstrated that *NKILA* interacts directly and stably with the NF-κB:IκB complex in the cytoplasm to prevent IκB phosphorylation and suppress activation of the NF-κB pathway, suggesting a tumor suppressive role for *NKILA* in limiting inflammatory processes in cancer (Liu et al., 2015). A different study from the same group showed that RNAi downregulation of *NKILA* in cytotoxic T cells (CTLs) led to increased tumor

infiltration and reduced tumor volume in a breast cancer PDX mouse model, implicating *NKILA* as a potential target in the field of cancer immunotherapy (Huang et al., 2018a).

LncGata6

LncGata6 (LncRNA GATA6) was identified as a divergent transcript expressed from the promoter of *Gata6*, which is specifically enriched in a subset of intestinal stem cells (Zhu et al., 2018). Deletion of exons 2-4 of *lncGata6* in the mouse did not affect *Gata6* levels but resulted in decreased intestinal regeneration due to decreased proliferative capacity of intestinal stem cells (Zhu et al., 2018). Consistent with the key role of stem cells in intestinal tumorigenesis, genetic and ASO-mediated depletion of *lncGata6*/*lncGATA6* were found to impair tumor growth in the APCmin mouse model of intestinal adenoma and in a PDX model (Zhu et al., 2018). Future studies should focus on elucidating the mechanism by which *lncGATA6* is upregulated in colorectal cancer and on determining the extent to which it contributes to aberrant Wnt signaling, a known colorectal cancer driver.

DINO

The p53 target lncRNA *DINO* binds to and stabilizes p53 in a positive feedback loop, enhancing the activation of p53 target genes (Schmitt et al., 2016). Importantly, RNAi knockdown of *DINO* in human fibroblasts and a deletion of the *Dino* promoter in MEFs led to impaired cell cycle arrest following genotoxic stress (Schmitt et al., 2016). Interestingly, ectopic expression of *DINO* in HPV-positive cervical cancer cells, which suppress p53 stabilization and express *DINO* at low

levels, led to reactivation of dormant p53, resulting in sensitization of the cancer cells to chemotherapeutic agents and vulnerability to metabolic stress (Sharma and Munger, 2020). To date, however, evidence that *DINO* acts as a tumor suppressor in human cancer is limited.

LINC-PINT

Like *DINO*, *linc-Pint* was also identified as a p53 target (Marin-Bejar et al., 2013). A knockout mouse generated by replacing the *linc-Pint* locus with a LacZ reporter cassette yielded smaller pups, suggesting a role for *linc-Pint* in early development (Sauvageau et al., 2013). Characterization of *LINC-PINT* function in cancer suggested a role in limiting cell invasion, with *LINC-PINT* overexpression leading to decreased liver metastases in a mouse model (Marin-Bejar et al., 2017). In a transwell migration and invasion assay, invasiveness increased following treatment with *LINC-PINT* -targeting ASOs or following CRISPR-mediated deletion of a highly conserved *LINC-PINT* sequence element (Marin-Bejar et al., 2017). Analysis of the previously generated *linc-Pint* knockout mouse (Sauvageau et al., 2013) in a cancer background could help support these results. However, the potential role of the *LINC-PINT* RNA may be confounded by the identification of a peptide with a function in suppressing cell proliferation encoded by a circular form of *LINC-PINT* (Zhang et al., 2018a).

THOR

While examples of alternative organismal models for lncRNA function in cancer are limited, in part due to low evolutionary conservation of lncRNAs,

investigation of the highly conserved lncRNA *THOR* (Testis-associated Highly conserved Oncogenic long noncoding RNA) in human and zebrafish cancer models has implicated this lncRNA in promoting melanoma development (Figure 5) (Hosono et al., 2017). *THOR* expression is normally restricted to the testis, but has been found aberrantly overexpressed in multiple cancer types, including lung adenocarcinoma, lung squamous carcinoma, and melanoma (Hosono et al., 2017). Knockdown of *THOR* via RNAi and ASOs in lung adenocarcinoma and melanoma cell lines led to decreased proliferation and reduced colony formation (Hosono et al., 2017). These findings were corroborated in two independently derived lung adenocarcinoma cell lines harboring approximately 3 Kb CRISPR-mediated deletions within the *THOR* gene body. Conversely, *THOR* overexpression gave the opposite phenotype, leading to increased proliferative capacity and anchorage-independent growth. Importantly, ectopic expression of human *THOR* in zebrafish cooperated with oncogenic NRAS and p53 loss to promote melanoma development, whereas knockout of *THOR* in zebrafish embryos delayed mutant NRAS-induced melanoma formation (Hosono et al., 2017). Further studies may reveal the potential of using *THOR* expression as a biomarker or targeting *THOR* as a therapeutic strategy.

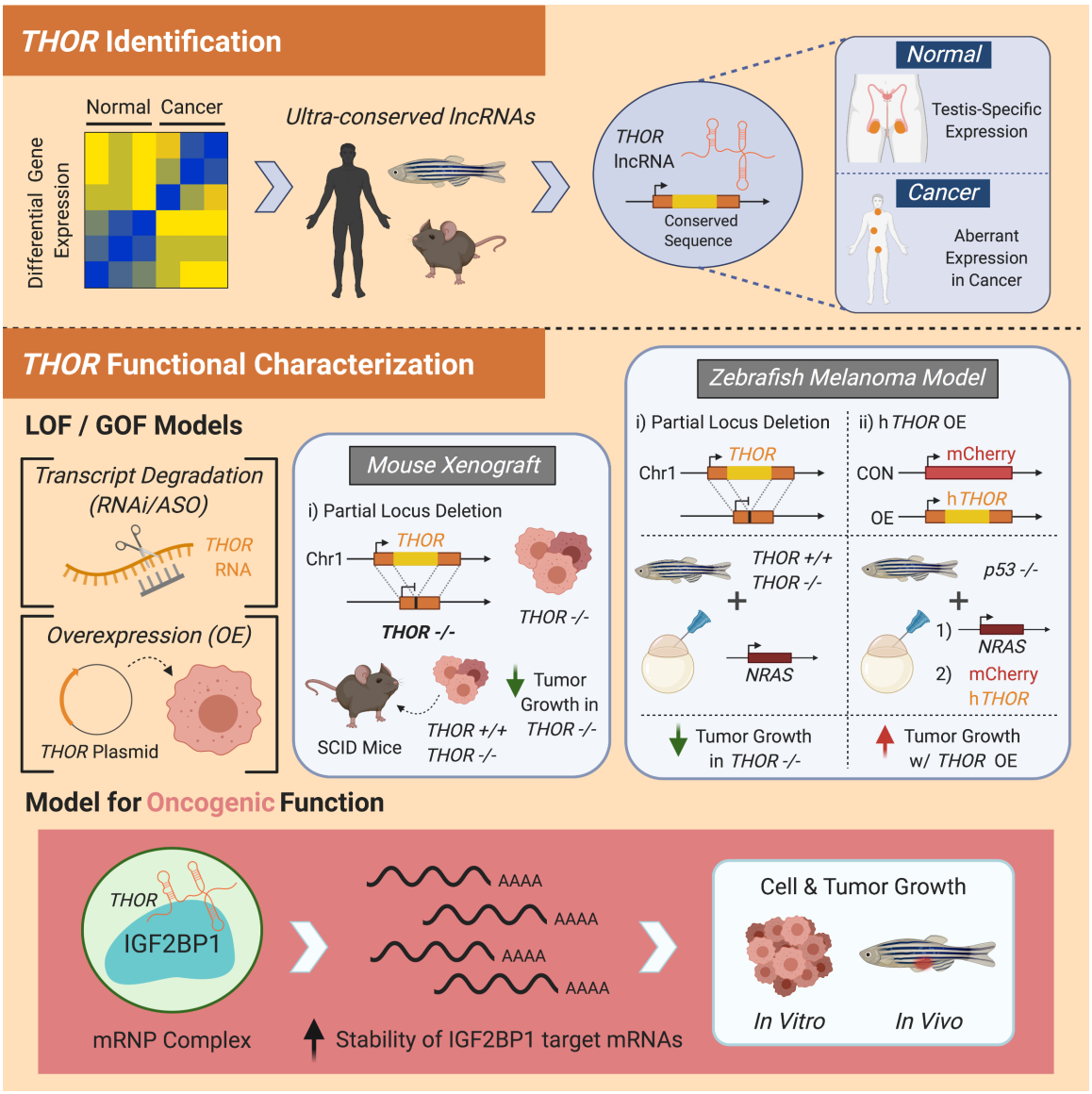

Figure 5. Identification and functional characterization of *THOR*. *THOR* was identified as a testis-specific ultra-conserved lncRNA aberrantly expressed in cancer tissues (Hosono et al., 2017). Hosono and colleagues generated several *in vitro* and *in vivo* loss-of-function (LOF) and gain-of-function (GOF) models to functionally characterize *THOR*. LOF models included transcript degradation with RNAi and ASO, and *THOR* partial locus deletion (*THOR-/-)* in both human cells injected in severe combined immunodeficiency disease (SCID) mice and in a genetically engineered zebrafish model (*THOR-/-*) embryonically injected with *NRAS* to induce melanoma. GOF models included *THOR* overexpression (OE) *in vitro* and OE of human *THOR* (h*THOR*) in p53-deficient zebrafish (*p53-/-*) embryonically injected with *NRAS* to induce melanoma. Overexpression of *THOR* plays an oncogenic role (red box) in cancer by binding to IGF2BP1 and increasing the stability of its mRNA targets to promote cancer progression.

Future Perspectives

Identification of lncRNAs that are genetically or epigenetically perturbed in cancer has risen sharply over the past decade. The precipitous increase in the number of cancer-associated lncRNAs has been accompanied by a growing excitement that many lncRNAs may act as novel drivers of cancer development. Yet, lagging understanding of how lncRNAs function in physiologic and pathologic contexts has limited our insights into the roles of lncRNAs in tumorigenesis. The current literature points to many lncRNAs acting as both oncogenes and tumor suppressors. While these seemingly contradictory findings may stem from differences in experimental models, they may also be reflective of complex and context-dependent lncRNA biology, analogous to the dual oncogenic and tumor suppressor roles played by cancer-associated protein-coding genes (Shen et al., 2018). Future studies should prioritize the identification and validation of true dual functions from technical inconsistencies.

LncRNAs make attractive drug targets, particularly in diseases where protein candidates are not amenable to pharmacological inhibition (Dang et al., 2017). Both siRNA- and ASO-mediated lncRNA degradation as well as locked nucleic acid (LNA)-mediated interference with lncRNA function have emerged as clinic-ready approaches (Arun et al., 2018; Lieberman, 2018). The successful deployment of these approaches in cancer, however, is predicated upon robust functional characterization. In the future, it would be essential to develop *in vitro* and *in vivo* models that closely recapitulate the recurrent genetic or epigenetic changes of lncRNAs observed in human cancer. In parallel, experiments that uncover the functional elements of perturbed lncRNA loci will inform whether

motives or structural features of the lncRNA molecules, the act of their transcription, or underlying DNA elements mediate their roles in disease development. These questions will be best answered through the integration of diverse and complementary approaches and by corroboration from multiple independent studies.

The p53 and Myc duet in cancer

Transcription networks coordinate the expression of a variety of genes in response to cellular inputs. As such, transcription factors and their downstream gene expression programs are frequently deregulated in cancer, with many the target of alterations that promote tumor growth (*reviewed in* (Bradner et al., 2017)). Two significant transcription networks in cancer are regulated by the p53 tumor suppressor and the Myc proto-oncogene, respectively. As a consequence of their central roles in controlling cellular growth and survival, albeit activated by very different cellular impulses, both networks are often subject to pro-tumorigenic genetic and epigenetic modifications (Dang, 2012; Kastenhuber and Lowe, 2017).

The human *TP53* gene, which encodes the p53 tumor suppressor protein, is mutated in as many as 50% of all human cancers, with p53 loss often predictive of advanced tumor grade and poor overall survival (Kandoth et al., 2013; Olivier et al., 2010). Germline *TP53* mutations, the defining characteristic of Li Fraumeni Syndrome, dramatically increase the risk of developing a range of tumor types (Olivier et al., 2010). Similarly, mice with germline loss of one or both copies of

p53 (*Trp53-/-* or *Trp53+/-*) are prone to spontaneous tumorigenesis, with complete p53 loss often resulting in neoplasm development within the first year of life (Donehower et al., 1992; Jacks et al., 1994). Apart from mutations which inhibit p53 directly, mutations that disturb critical p53 effector genes are also common, and can be equally as disruptive to p53 network function. For example, transgenic mice that overexpress *Mdm2* (*Mouse double minute 2*), a negative regulator of p53, and mice lacking *Arf* (*Alternate reading frame of Cdkn2a*), a positive regulator of p53, are both similarly prone to spontaneous tumorigenesis (Jones et al., 1998; Kamijo et al., 1999).

On the other hand, cancer-driving mutations affecting the Myc network are often characterized not by alterations to *Myc* coding sequences, but by alterations that increase *Myc* expression or activity. Upregulation of *MYC* in avian leukosis virus (ALV)-induced lymphomas via retroviral insertion upstream of the *MYC* promoter is well-documented (Hayward et al., 1981; Payne et al., 1982). Furthermore, an analysis of somatic copy number alterations (SCNAs) in human cancer revealed *MYC* alterations as one of the most prominent (Beroukhim et al., 2010). The *MYC* locus often participates in chromosomal rearrangements, and is a frequent translocation partner of the immunoglobulin heavy chain locus (Taub et al., 1982), a fusion that increases *MYC* expression and is modeled in the *Eμ–Myc* mouse model of B-cell lymphoma (Adams et al., 1985). Focal genomic amplifications of the *MYC* locus in the form of homogeneously staining regions and double minutes are also common (Alitalo et al., 1983; Collins and Groudine, 1982). Notably, several studies have illustrated MYC oncogene addiction, with

suppression of MYC sufficient to cause tumor regression (*reviewed in* (Dang, 2012)), suggesting its importance for viability in MYC-driven cancers.

The observation that loss-of-function and gain-of-function mutations in *TP53* and *MYC*, respectively, have such pervasive effects on cancer pathology points to critical roles for these transcription networks in maintaining normal cellular growth and homeostasis.

The p53 tumor suppressor pathway

Under normal conditions, p53 is constitutively produced and degraded, enabling rapid action to be taken against cellular assaults (Kastenhuber and Lowe, 2017). In response to a variety of cellular stressors, such as doxorubicin-induced DNA damage or the expression of an activated oncogene, p53 is stabilized via repression of MDM2 activity and p53 post-translational modifications, thus enabling its tetrameric binding to canonical p53 response elements (p53REs) and the induction of the p53 transcriptional program (Figure 6) (Beckerman and Prives, 2010; Kastenhuber and Lowe, 2017). The ability of p53 to bind specific DNA sequences is critical to its function. Genes directly induced by p53 share the presence of a conserved p53RE, usually in their promoter or first intron (Beckerman and Prives, 2010), composed of two 10 bp motifs separated by a stretch of nucleotides between 10-13 bp in length (el-Deiry et al., 1992). There is some evidence of gene repression downstream of p53 (Allen et al., 2014), but its contribution to the p53 response remains unclear. Significantly, p53 stabilization strongly correlates with reduced MYC levels, suggesting that p53 may actively suppress positive regulators of the cell cycle rather than simply promoting the

expression of negative regulators of cell growth (Ho et al., 2005; Porter et al., 2017).

Through the activation of its transcription program, p53 drives a variety of cellular outcomes, the most prominent of which are transient cell cycle arrest (Brugarolas et al., 1995), senescence (Ferbeyre et al., 2002), and apoptosis (Lowe et al., 1994). Moreover, several studies have suggested that canonical p53 responses and full p53 transcriptional activation may be dispensable for tumor suppression, suggesting that some non-canonical p53-driven responses may play equally important roles in cellular homeostasis (Jiang et al., 2011; Li et al., 2012). The targets of p53 are diverse and include coding and noncoding genes with functions in a range of cellular processes such as cell cycle regulation, pluripotency, metabolism, and inflammation (*reviewed in* (Kastenhuber and Lowe, 2017)). Notably, the characterization of numerous lncRNA targets of p53 in recent years suggests the growing importance of this class of noncoding transcripts in the p53 response to stress (as discussed previously in this chapter). What dictates selection of a particular cellular outcome in response to p53 activation is unclear, although evidence suggests it may depend on cell type and stressor (Attardi et al., 2004; Paris et al., 2008). Despite these gaps in knowledge, it is clear that p53 is critical in preventing the outgrowth of aberrantly proliferating cell populations that might otherwise initiate cancer development.

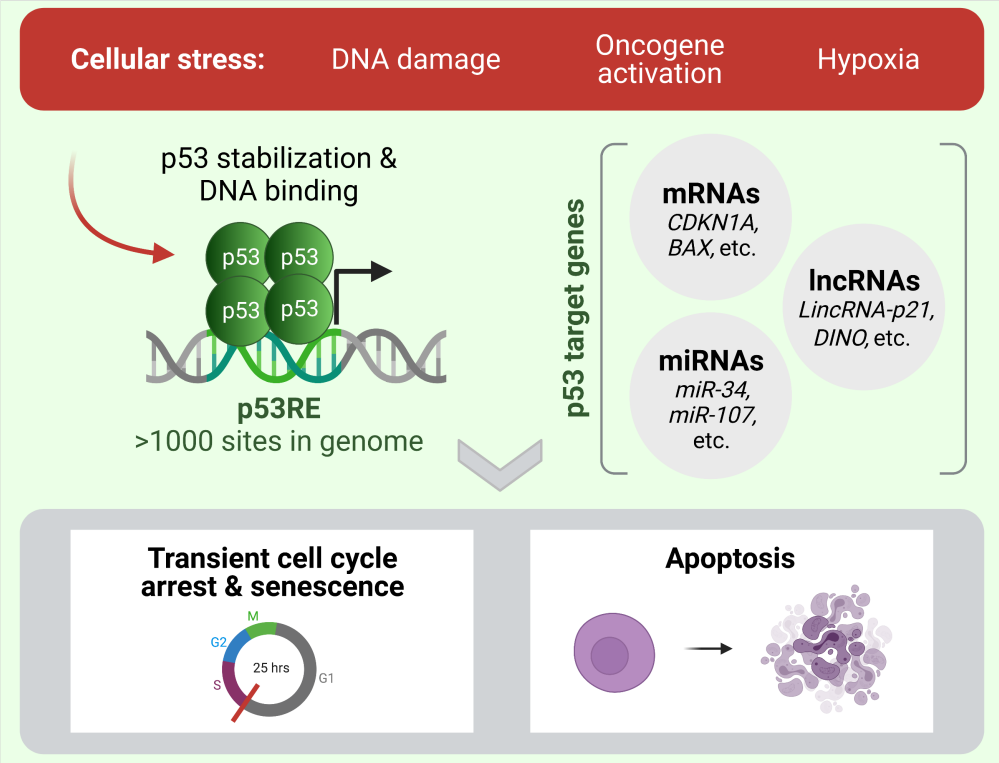

Figure 6. The p53 tumor suppressor pathway. The tumor suppressor protein p53 is a transcription factor that is stabilized by cellular stress (i.e. DNA damage, oncogene activation, hypoxia, etc) and binds p53 Response Elements (p53REs) to upregulate various coding and noncoding target genes. The induction of the p53 transcription program drives cellular outcomes (i.e. cell cycle arrest, senescence, and apoptosis) that limit cell growth.

The proto-oncogenic Myc network

Myc responds to pro-growth stimuli by orchestrating the induction of cell cycle genes to promote cell division. It does so most commonly through heterodimerization with its partner Max (Myc-associated factor X) and cooperative binding to E-box (CACGTG) sequences in the promoters of target genes (Amati et al., 1993; Blackwell et al., 1990). While Myc half-life is short, Max is more stable, pointing to both *Myc* expression and protein synthesis as rate limiting steps in cellular proliferation (Amati et al., 1993).

Researchers have long sought to define a universal set of Myc target genes, but precise characterization of the Myc transcription network has remained elusive. Analysis of serum-stimulated fibroblasts revealed nearly 300 genes induced in a Myc-dependent manner (Perna et al., 2012), consistent with an understanding of Myc as driving a selective gene expression program (Sabo et al., 2014; Walz et al., 2014). However, several studies have advocated a less discriminatory role for Myc function, suggesting that Myc acts not necessarily as a sequence-specific transcription factor, but as a global amplifier of genes expressed from open chromatin (Lin et al., 2012; Rahl et al., 2010). These differing viewpoints stem, in part, from normalization tactics used in RNA-sequencing experiments to identify Myc target genes. Specifically, there has been disagreement over how to contend with the proposed ability of Myc to increase global transcription as a potential confounding factor in standard normalization practices. Some have argued that normalization based on cell equivalents overcomes biases introduced by Myc-induced surges in total RNA content, but in practice accuracy may require multiple analyses (Kress et al., 2015). Of note, Myc

roles in gene repression have also been suggested, such as its proposed function in influencing the cellular response to DNA damage by downregulating the CDK inhibitor p21 to promote apoptosis, perhaps indicating feedback between the p53 and Myc transcription programs (Seoane et al., 2002).

Ambiguities concerning Myc regulation and function in cancer also persist. Expressed from the 8q24 locus, *MYC* is impacted by a range of *cis*-acting DNA elements (Fulco et al., 2016). An abundance of lncRNAs in the *MYC* locus have also been identified, with some proposed to have cancer-specific functions in *MYC* regulation (Ling et al., 2013; Xiang et al., 2014). How these lncRNAs and other enhancer and repressor elements adjacent to the *MYC* locus drive *MYC* expression in different cancer types is still an open question. While elevated Myc levels are known to promote aberrant cell growth under circumstances favorable to tumorigenesis, its dysregulation has also been associated with an increased propensity for cells to undergo programmed cell death (Zindy et al., 1998). Interestingly, some have speculated about a *Myc* expression threshold cells cannot surpass without triggering apoptosis (Murphy et al., 2008). This has been proposed as especially critical in early cancer development when cells retain functional tumor suppressor pathways, suggesting that lower levels of Myc dysregulation may be optimal for initial cancer cell survival.

Project framework: p53, Myc, and the missing lncRNA

This chapter links together disparate elements of lncRNA biology and cancer biology in broad strokes. A key question raised at the intersection of these

fields is whether, and to what extent, lncRNA aberrations in cancer can constitute true drivers of tumorigenesis. There is abundant evidence for genetic and epigenetic alterations of lncRNAs in cancer and their presence in pathways essential to cancer growth and development is well-documented. However, ultimately only a small fraction of cancer-related lncRNAs have been shown to have significant, and functionally consistent, physiological impacts when they are manipulated *in vivo* (discussed in detail previously in this chapter). As such, the functional impacts of individual lncRNAs, and indeed lncRNAs as a class, on cancer pathogenesis are incompletely understood.

An additional, and ultimately related, question highlighted here concerns the elusive threads connecting the p53 and Myc transcription networks. Given the significant, albeit often contradictory, influences of the tumor suppressor p53 and the proto-oncogene Myc on cell growth, and by extension, cancer development, one might expect a need for cells to coordinate their activities in order to avoid the transmission of opposing cellular impulses (Figure 7). As alluded to previously, p53 activation results in repression of Myc as a mechanism for temporarily curtailing cellular proliferation in stressed cells (discussed more in **Chapter 2**). This phenomenon has long been of interest to cancer biologists, but remained poorly understood despite decades spent studying Myc regulation. We find that this cellular fail-safe is primarily facilitated by the lncRNA isoform *Pvt1b*, the subject of this work. Considering the heterogeneity of lncRNAs in form and function, they make apt candidates for enacting specific local regulation downstream of broadly active transcription factors. However, the lncRNA field has been plagued by challenges in ascertaining lncRNA functional roles, particularly

under tumorigenic conditions; many controversies center on lncRNAs with both ascribed pro- and anti- oncogenic capabilities rooted in evidence from conflicting model systems.

Here I describe *Pvt1b* as an emissary connecting two extensive gene expression programs, a potential archetype for lncRNA function, which may exist in other cellular processes and pathways. Incorporating lessons in experimental design derived from lncRNA literature and guided by an understanding of the roles of p53 and Myc in cancer, I employ a range of orthogonal approaches to assess the function and biological significance of *Pvt1b* under physiologic and tumorigenic conditions. I first identify and characterize *Pvt1b*, presenting evidence for its role in repressing *Myc* transcription downstream of p53 activation (see **Chapter 2**). I further investigate the functional elements of the *Pvt1b* transcript (see **Chapter 3**) and pursue a potential mechanism of action involving *Pvt1b*-guided histone deacetylation at the *Myc* promoter (see **Chapter 4**). Finally, I provide a foundation for future dissection of *Pvt1* function *in vivo* through my role in generating three genetically engineered *Pvt1* mouse models, the incisive combination of which should help illuminate features of the *Pvt1* locus required for its activities (see **Chapter 5**). Rationales for specific experiments are discussed in detail in **Chapters 2-5**.

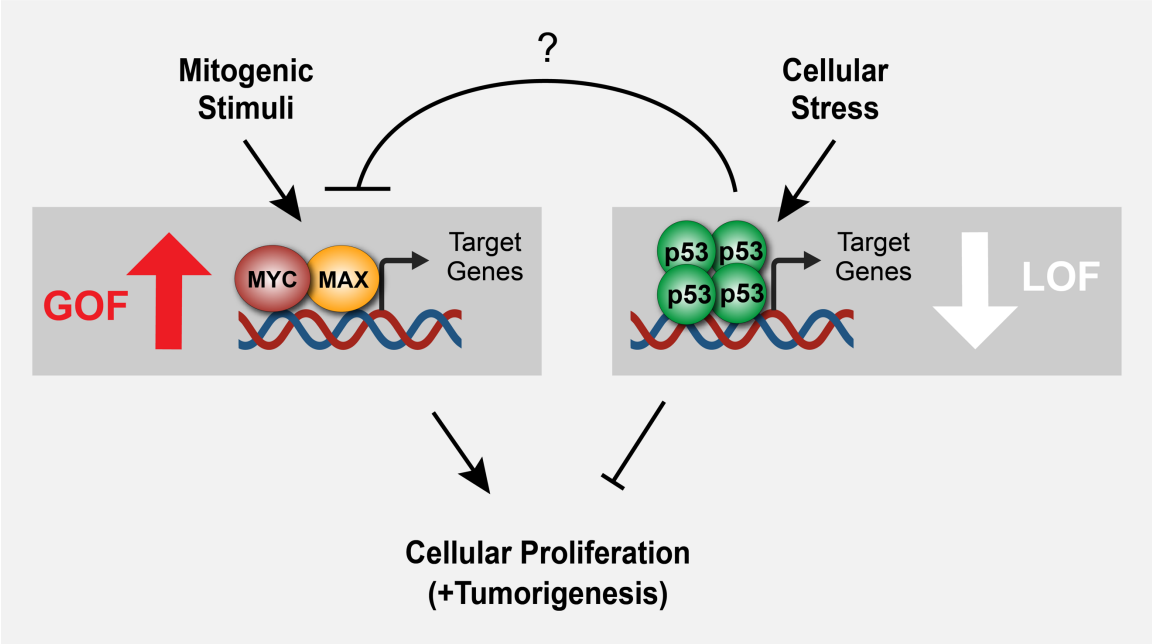

Figure 7. The p53 and Myc transcription networks in cancer. In response to mitogenic stimuli, MYC (red) dimerizes with its partner MAX (yellow) to upregulate target genes that promote cellular proliferation. In contrast, p53 (green) responds to cellular stress, binding as a tetramer to induce genes that suppress cellular growth. In cancer, MYC is frequently the target of gain-of-function (GOF) mutations, while p53 is frequently the target of loss-of-function (LOF) mutations. MYC levels have been observed to decrease following p53 activation, and various mechanisms, both direct and indirect, have been proposed to explain this phenomenon.

Chapter 2:

p53 activates the long noncoding RNA Pvt1b to inhibit Myc and suppress tumorigenesis

The work described in this chapter has been published as follows: Olivero, C.E., Martinez-Terroba, E., Zimmer, J., Liao, C., Tesfaye, E., Hooshdaran, N., Schofield, J.A., Bendor, J., Fang, D., Simon, M.D., et al. (2020). p53 Activates the Long Noncoding RNA Pvt1b to Inhibit Myc and Suppress Tumorigenesis. Mol Cell 77, 761-774 e768. Co-author contributions that have also been included in this thesis are specified in figure legends and/or text.

Introduction

The p53 (also known as TP53) network is a central tumor suppressive mechanism in mammalian cells that is inactivated in the vast majority of human cancers (Vousden and Prives, 2009). In response to cellular stress induced by DNA damage or oncogenic signaling, p53 transcriptionally activates target genes to limit cellular proliferation or to permanently eliminate damaged cells (Vousden and Prives, 2009). Transcriptional activation by p53 relies on its binding to conserved p53 response elements (p53REs) in the promoters of target genes (Levine and Oren, 2009). p53 has also been implicated in the repression of cell cycle regulators (Engeland, 2018). One of the prominent targets of p53 repression is the Myelocytomasis (Myc) oncogene (Ho et al., 2005; Levy et al., 1993; Sachdeva et al.,

2009), a global transcriptional amplifier that responds to mitogenic signals to promote cellular proliferation (Lin et al., 2012). Multiple models for how p53 negatively affects *Myc* levels have been proposed, including p53 binding to the *Myc* promoter to suppress histone acetylation, binding to a distal regulatory element to alter nucleosome positioning in the *Myc* promoter, or activating repressive *Myc*-targeting microRNAs (Ho et al., 2005; Porter et al., 2017; Sachdeva et al., 2009). However, the mechanism of p53-mediated *Myc* downregulation and its contribution to tumor suppression *in vivo* have remained unclear.

Long noncoding RNAs (lncRNAs) can modulate gene expression locally by accumulating near their sites of transcription (Kopp and Mendell, 2018). In dosage compensation, *Xist* and other lncRNAs expressed from the X-chromosome specifically repress genes across the entire X-chromosome through the recruitment of epigenetic regulators (Lee, 2012). Other *cis*-regulatory lncRNAs act in a more limited, locus-specific manner, such as the p53 target *lincRNA-p21* proposed to promote the levels of its neighbor *p21* (also known as *Cdkn1a*) by recruiting activating factors (Dimitrova et al., 2014). While studies of locus-specific *cis*-regulatory lncRNAs have revealed important roles in diverse biological processes (Dimitrova et al., 2014; Elling et al., 2018; Kotzin et al., 2016), characterization of the RNA molecule is often confounded by potential functional roles of DNA regulatory sequences in the lncRNA locus (Bassett et al., 2014; Engreitz et al., 2016; Groff et al., 2016). Defining the RNA-mediated regulation provides important opportunities for RNA-based therapeutics that can alter hardwired molecular interactions to change cellular responses.

Plasmacytoma variant 1 (Pvt1), a lncRNA expressed 50 Kb downstream of *Myc,* is altered in a large fraction of human cancers. Frequent translocations and viral integrations in the *Pvt1* locus in lymphomas suggest important roles for *Pvt1* in cancer progression (Cory et al., 1985; Graham and Adams, 1986; Graham et al., 1985). In addition, co-amplification of *Myc* and *Pvt1* across multiple cancer types correlates with poor cancer patient prognosis, suggesting cooperation between the two genes during tumorigenesis (Cui et al., 2016; Tseng and Bagchi, 2015; Zeng et al., 2017). This pro-oncogenic cooperation between *Myc* and *Pvt1* was recently confounded by the identification of a p53-binding site in the *Pvt1* locus and by the description of the *Pvt1* promoter as a transcriptional repressor of *Myc* (Cho et al., 2018; Porter et al., 2017). These studies suggested undefined roles for *Pvt1* in cancer progression and a potential crosstalk between the tumor suppressor p53 pathway and the oncogenic Myc network.

In this study, I characterized *Pvt1b,* a p53-induced isoform of the lncRNA *Pvt1*, and determined its contribution to Myc regulation and the p53 response to stress. I show that production of the *Pvt1b* RNA downstream of p53 represses *Myc* transcription and suppresses cellular proliferation during stress and in the early stages of tumorigenesis. The model presented here illuminates a role for the lncRNA isoform *Pvt1b* as a locus-specific transcriptional regulator that serves to enact selective gene repression downstream of the broad p53 transcriptional activation network.

Results

p53 suppresses *Myc* under conditions of genotoxic and oncogenic stress

To gain insight into the mechanism by which p53 causes suppression of *Myc*, I used multiple independent approaches to model the p53-dependent response to stress. To model the cellular response to genotoxic stress, I utilized wild-type (WT) mouse embryonic fibroblasts (MEFs) treated with the genotoxic agent Doxorubicin (Doxo) (Figure 8A). I observed that activation of the p53 transcriptional program following Doxo treatment for 24 hours resulted in 3-fold induction of the p53 target *p21* and a concomitant reduction in *Myc* RNA and protein levels by 34±6% (p=0.008, Figure 8B) and 44±15% (p=0.0051, Figure 8C), respectively, consistent with previous findings (Ho et al., 2005; Porter et al., 2017). I also found that p53 activation by oncogenic stress, modeled by Tamoxifen (Tam)-CreER-dependent restoration of endogenous p53 expression in a murine lung adenocarcinoma cell line (*K-ras$^{LA2\text{-}G12D/+}$; p53$^{LSL/LSL}$; Rosa26-CreER^{T2}+, KPR*) (Figure 8D) (Feldser et al., 2010), similarly led to a 70-fold activation of *p21*, a 34±7% repression of *Myc* RNA (p=0.0020, Figure 8E) and a 37±10% decrease in Myc protein (p=0.0028, Figure 8F). *Myc* repression by 39±5% was also observed in intestinal epithelium cells isolated from mice exposed to 6 Grays (Gy) of whole-body irradiation, which leads to a well-characterized p53-mediated response to genotoxic stress *in vivo* (p=0.0007, Figures 8G and 8H) (Clarke et al., 1994). Altogether, these results suggested that *Myc* repression is a general event downstream of p53 transcriptional activation.

In an effort to elucidate the mechanism by which p53 activation results in

Myc repression, Nadya Dimitrova examined whether p53 associates with the *Myc* locus. She observed that both in Doxo-treated MEFs and Tam-treated *KPR* cells, stress-dependent *Myc* repression was accompanied by binding of p53 to a distal p53RE, located 50 Kb downstream of *Myc*, which has previously been implicated in limiting *Myc* expression (Figure 8I)(Porter et al., 2017).

Consistent with p53 dependency, the changes in *Myc* RNA and protein levels were present in p53-proficient, but not p53-deficient MEFs (Figures 8J and 8K). Additionally, the decrease in *Myc* RNA levels was detectable as early as 4 hours following p53 activation and was coincident with the decrease in Myc protein levels, suggesting direct transcriptional modulation by p53 (Figures 8L and 8M). Inhibition of protein translation with Cycloheximide (Chx) revealed that Myc protein stability was not significantly affected by the presence of stress, suggesting that the decrease in Myc levels was not primarily due to post-translational regulation (Figure 8N).

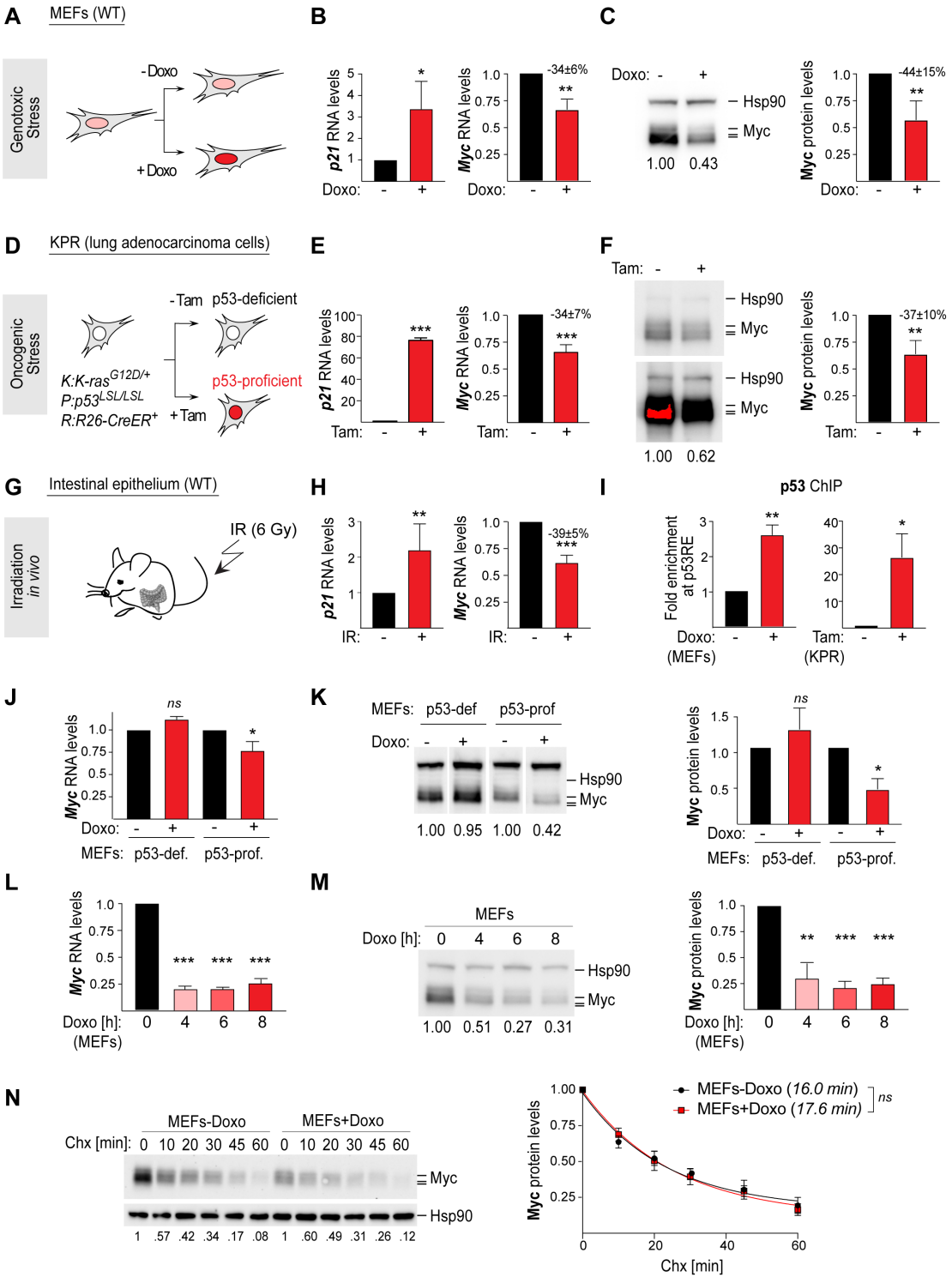

Figure 8. p53 suppresses *Myc* in response to genotoxic and oncogenic stress. (A) Schematic of the model system for studying p53-mediated response to genotoxic stress in WT MEFs untreated or treated with Doxo for 24 h. Activation of p53 by passaging or by genotoxic stress is represented by light and dark red nuclei, respectively. **(B)** *p21* and *Myc* RNA levels in cells from (A). Data show mean ± SEM (n=4, biological replicates), *p<0.05, **p<0.01, paired t test. **(C)** *Left* Representative image and quantification of Myc protein levels from cells in (A). Hsp90 as a loading

control. *Right* Bargraph of Myc protein levels showing mean±SEM (n=5, biological replicates), **p<0.01, paired t test. **(D)** Schematic of the model system for studying p53-mediated response to oncogenic stress in *KPR* cells untreated or treated with Tam for 24 h. Activation of p53 by oncogenic stress is represented by red nucleus. **(E)** *p21* and *Myc* RNA levels in cells from (D). Data show mean ± SEM (n=6, biological replicates), ***p<0.001, paired t test. **(F)** *Left* Representative image and quantification of Myc protein levels from cells in (D). Hsp90 as a loading control. *Right* Bargraph of Myc protein levels showing mean±SEM (n=5, biological replicates), **p<0.01, paired t test. **(G)** Schematic of the model system for studying p53-mediated response *in vivo* in intestinal epithelial cells isolated from WT mice at 6 h post 6 Gy whole-body irradiation. Samples provided by Nadya Dimitrova. **(H)** *p21* and *Myc* RNA levels from mice in (G). Data show mean ± SEM (n=3, biological replicates) **p<0.01, ***p<0.001, unpaired t test. **(I)** Enrichment of p53 binding at the *Pvt1*-associated p53RE by ChIP-qPCR in *Left* Doxo-treated MEFs and *Right* Tam-treated *KPR* cells. Data show mean ± SEM (MEFs: n=4; KPR: n=3, biological replicates) *p<0.05, **p<0.01, paired t test. P53 ChIP performed by Nadya Dimitrova. **(J)** *Myc* RNA levels in p53-deficient or p53-proficient MEFs, untreated or treated with Doxo for 24 h. Data show mean ± SEM (n=3, biological replicates), *ns* = not significant, *p<0.05, paired t test. **(K)** *Left* Representative image and quantification of Myc protein levels from cells in (J). Hsp90 as a loading control. *Right* Bargraph of Myc protein levels showing mean ± SEM (n=3, biological replicates), ns = not significant, *p<0.05, paired t test. **(L)** *Myc* RNA levels in WT MEFs, untreated or treated with Doxo for the indicated times. Data show mean ± SEM (n=4, biological replicates), ***p<0.001, paired t test. **(M)** *Left* Representative image and quantification of Myc protein levels from cells in (L). Hsp90 as a loading control. *Right* Bargraph of Myc protein levels showing mean ± SEM (n=4, biological replicates), **p<0.01, ***p<0.001, paired t test. **(N)** *Left* Representative image and quantification of Myc protein levels following treatment with cycloheximide (Chx) for indicated times in WT MEFs, untreated or treated with Doxo for 8 h. *Right* Myc protein half-life (n=3, biological replicates), ns = not significant, paired t test.

Contributions from N. Dimitrova in (G) and (I) are described above.

Myc repression correlates with activation of a p53-dependent *Pvt1* isoform, *Pvt1b*

I was intrigued that the distal p53RE was located within the gene body of the lncRNA *Pvt1* (Figure 9A), which has previously been implicated as a p53 target (Barsotti et al., 2012). Considering lncRNAs can act *in cis* to regulate the transcription of neighboring genes, I examined whether *Pvt1* played a role in restricting *Myc* expression during stress. I noted significant stress-dependent induction of an isoform of *Pvt1*, termed *Pvt1b,* initiated at a transcription start site located immediately downstream of the p53RE. I observed a 3.1±0.2-fold induction of *Pvt1b* in Doxo-treated MEFs (Figure 9B) and a 38±6-fold induction of *Pvt1b* in Tam-treated *KPR* cells (Figure 9C). *Pvt1a*, an isoform of *Pvt1* initiated

at exon 1a, was induced to a lesser extent in Doxo-treated MEFs (Figure 9B) and was not significantly induced by Tam in *KPR* cells (Figure 9C). Copy number calculations suggested that *Pvt1b* was induced from 20 to 210 copes per cell, while *Pvt1a* was expressed at 300-400 copies per cell (Figure 10A). Notably, activation of *Pvt1b* was coincident with *Myc* repression and occurred as early as 4 hours following Doxo treatment in MEFs (Figure 10B) or 6 hours following Tam treatment in *KPR* cells (Figure 10C), consistent with direct transcriptional regulation by p53. Similarly, Doxo-treated human fibroblasts exhibited a 2-fold decrease in *MYC* levels and an 8-fold increase of human *PVT1B* (Figure 10D). These findings indicated that the downregulation of *Myc* and the activation of a p53-dependent, stress-specific *Pvt1* variant are conserved between mouse and human.

To further characterize the transcripts produced from the *Pvt1* locus, I performed RT-PCR with forward primers located in either exon 1a or 1b and a reverse primer in exon 5. I found evidence for extensive alternative splicing and confirmed that variants containing exon 1b were induced by p53, while exon 1a-containing variants were constitutively expressed (Figures 9D and 9E). Despite the splicing heterogeneity, sequencing of nascent RNA revealed that stress-induced *Pvt1b* differed from constitutively expressed *Pvt1a* solely by the use of exon 1b versus exon 1a, and exhibited comparable splicing patterns to downstream exons (Figure 9F). I concluded that p53 activation during genotoxic and oncogenic stress initiated transcription in the *Pvt1* locus from exon 1b, leading to the production of the p53-dependent isoform, *Pvt1b*, while *Pvt1a* represented a largely constitutively expressed isoform.

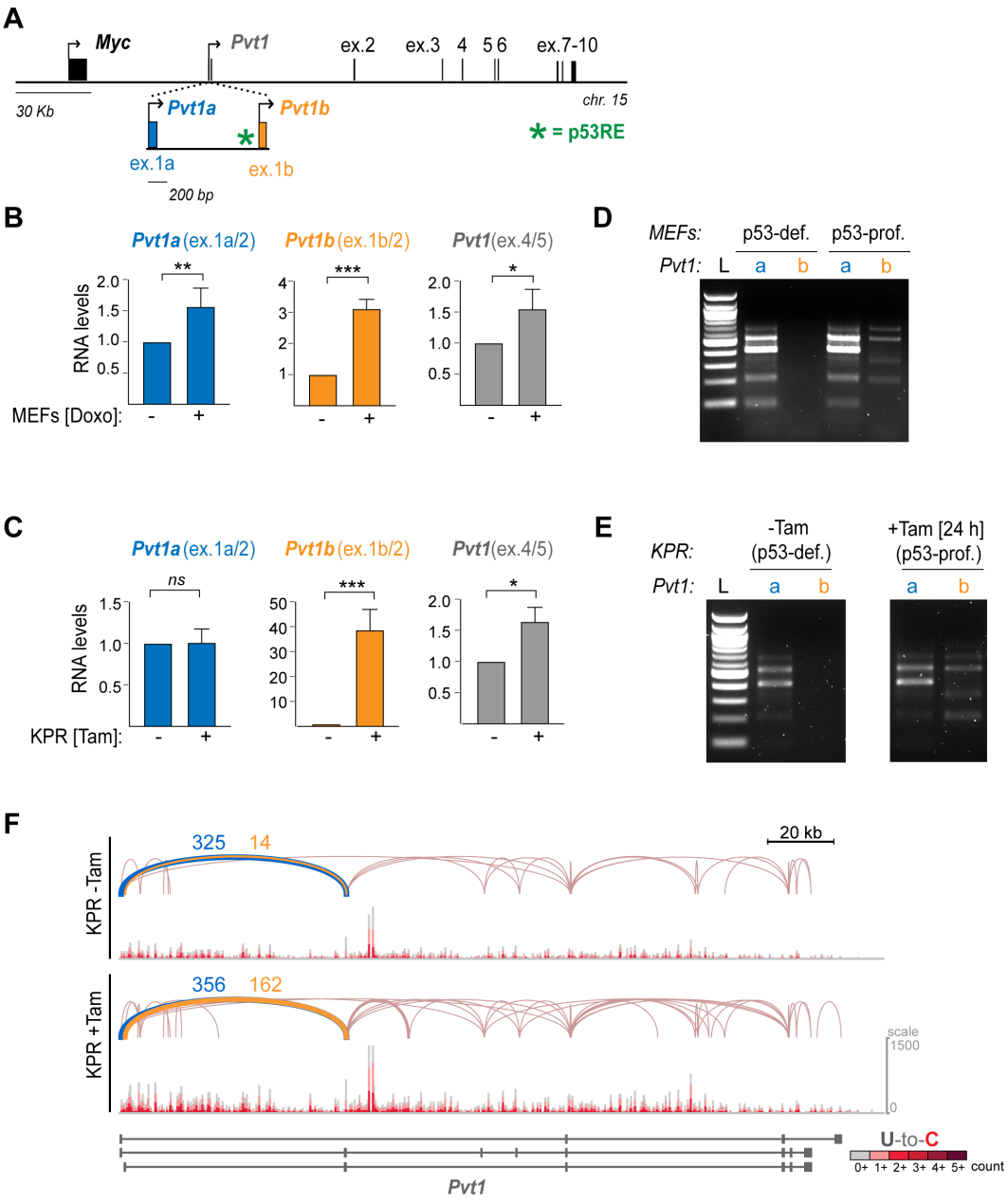

Figure 9. p53-dependent induction of the *Pvt1* isoform, *Pvt1b*. **(A)** Schematic of the mouse *Myc-Pvt1* locus, highlighting exons 1a and 1b of *Pvt1* and the location of the p53RE (green *). **(B, C)** Isoform-specific and total *Pvt1* RNA levels detected with primers located in indicated exons in (B) WT MEFs and (C) *KPR* cells, treated as indicated. Data show mean ± SEM (n=3, biological replicates), *p<0.05, **p<0.01, ***p<0.001, paired t test. **(D, E)** RT-PCR detection of *Pvt1a* isoforms (a, blue), amplified with primers from exon 1a to exon 5, and *Pvt1b* isoforms (b, orange), amplified with primers from exon 1b to exon 5, in RNA isolated from (D) MEFs and (E) *KPR* cells, ladder (L). **(F)** Genome browser tracks and Sashimi plots from TimeLapse-seq data in *KPR* cells, treated as indicated. Average number of splice junctions from 2 biological replicates from exon 1a to exon 2 (blue) and from exon 1b to exon 2 (orange) are indicated. Processing of samples for TimeLapse-seq and data analysis performed by Jeremy Schofield and Josh Zimmer.

Contributions from J. Schofield and J. Zimmer in (F) are described above.

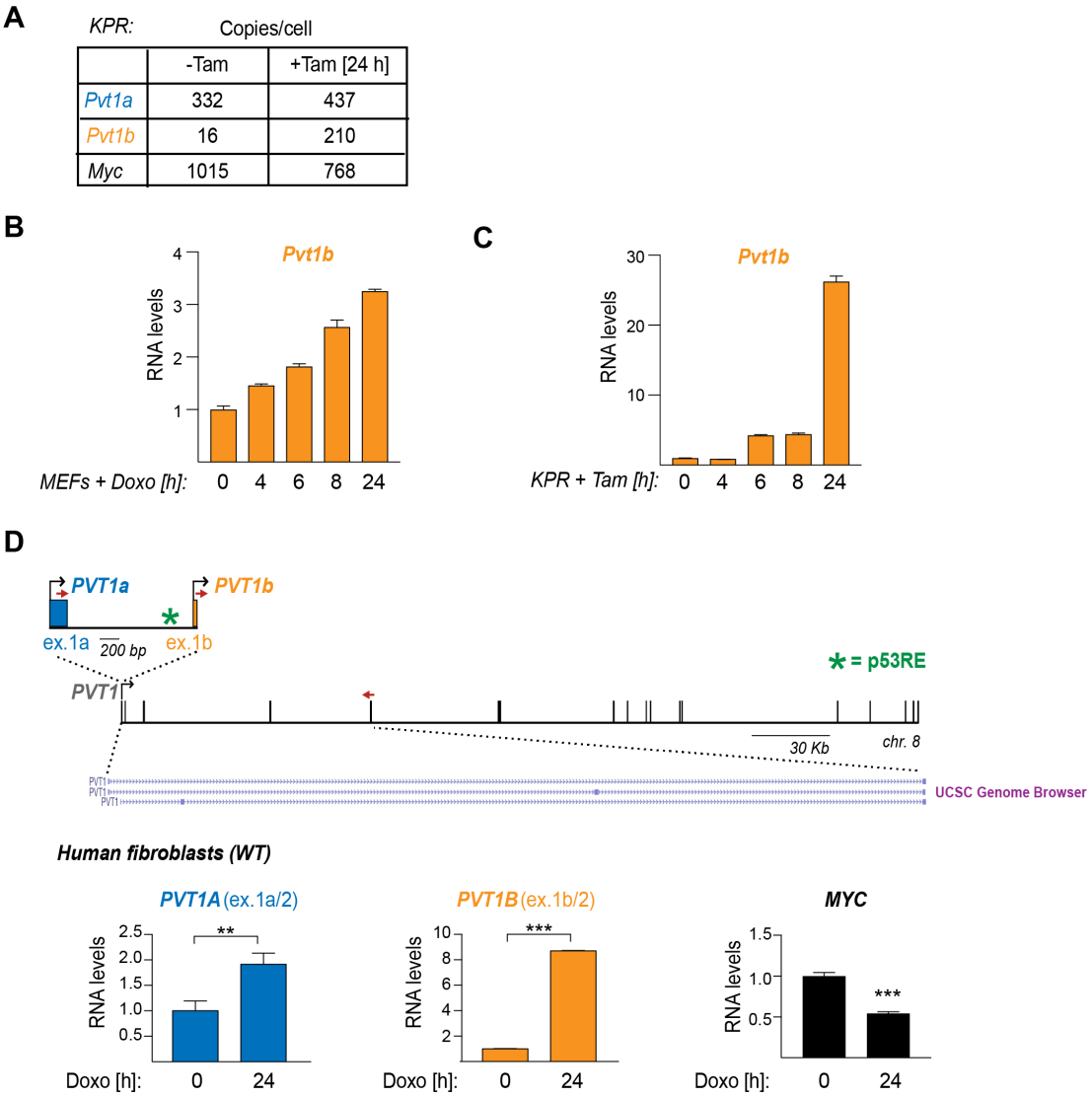

Figure 10. p53 activates the lncRNA isoform *Pvt1b*. (A) Copy number calculations of *Pvt1a*, *Pvt1b*, and *Myc* by qRT-PCR analysis of *KPR* cells untreated or 24 h post-treatment with Tam. **(B, C)** qRT-PCR analysis of *Pvt1b* levels in (B) MEFs treated with Doxo for the indicated times and (C) *KPR* cells treated with Tam for the indicated times. The observed induction of *Pvt1b* as early as 4-6 hours post stress suggests direct transcriptional activation by p53. **(D)** *Top* Schematic of the human *PVT1* locus, highlighting exon 1a (blue), exon 1b (orange), the conserved p53RE (green star) and showing the location of qPCR primers (red arrows), *Bottom* qRT-PCR analysis of relative *PVT1A*, *PVT1B,* and *MYC* RNA levels in normal human fibroblasts untreated or treated with Doxo for 24 h. Data show the mean ± SEM of 3 technical replicates from a representative example of two biological replicates. Human fibroblast samples provided by Nadya Dimitrova.

Contributions from N. Dimitrova in (D) are described above.

Stress-induced *Myc* repression occurs in the absence of promoter-enhancer contact reorganization

Previous work had shown that CRISPR-mediated transcriptional regulation of the *Pvt1* promoter in p53-deficient cancer cells causes reorganization of the chromatin architecture in the locus and impacts the access of *Myc* to downstream enhancers (Cho et al., 2018). To test whether the stress-responsive, p53-dependent induction of *Pvt1b* was associated with changes of these chromatin contacts, I performed Chromosome Conformation Capture (3C) in MEFs and *KPR* cells. Using an anchor in the *Myc* promoter, I confirmed that the *Myc* promoter accessed multiple upstream and downstream enhancers, including previously described *Pvt1* intragenic enhancers (Figures 11A and 11B) (Cho et al., 2018). However, I did not detect significant changes in the chromatin looping between the *Myc* promoter and *Myc*-associated enhancers during the p53-mediated stress response (Figures 11A and 11B). These results argue against a model where p53-dependent activation of *Pvt1b* leads to reorganization of the three-dimensional architecture of the locus.

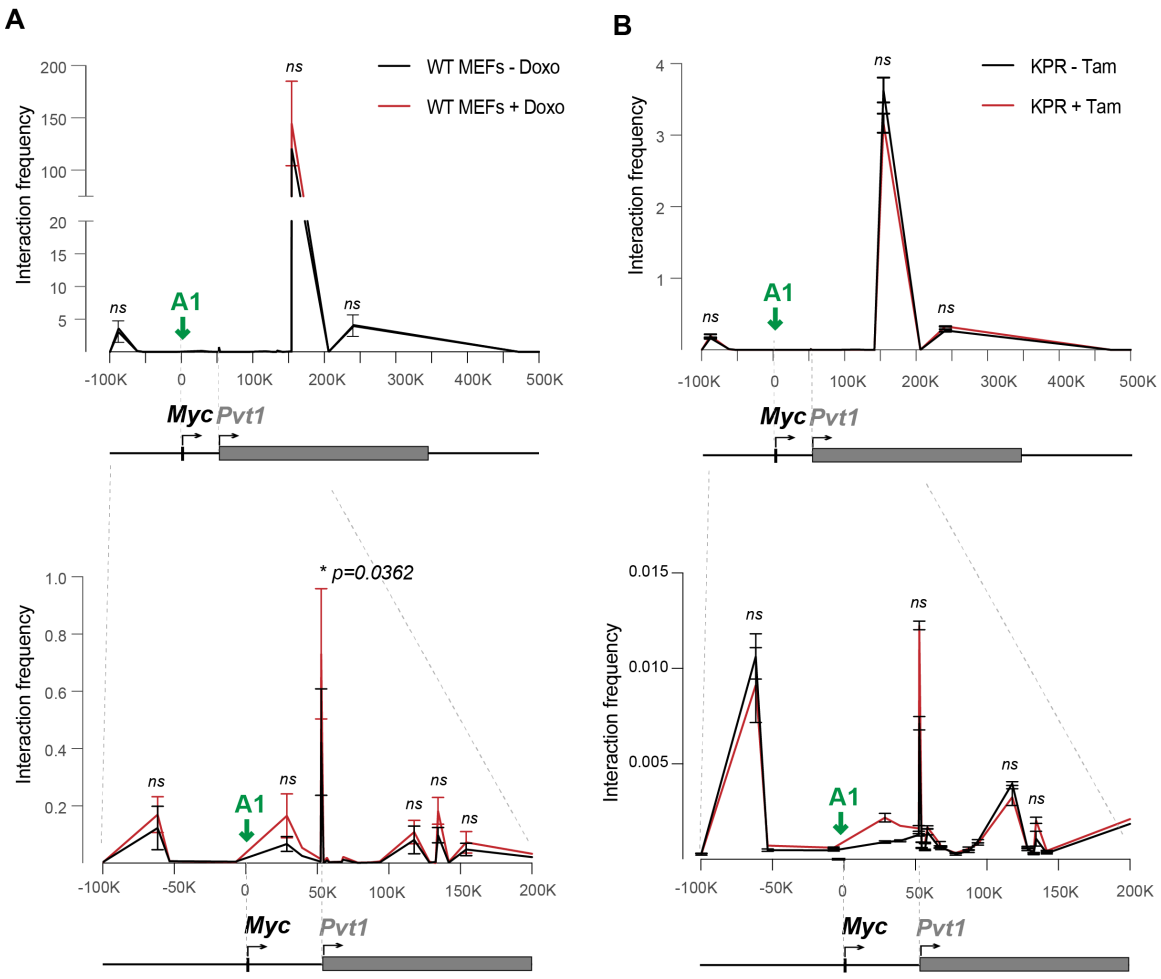

Figure 11. p53 activation does not substantially alter chromatin architecture of the *Myc-Pvt1* locus. (A-B) 3C analysis of BamHI-digested DNA from (A) WT MEFs, untreated or treated for 24 h with Doxo and (B) *KPR* cells, untreated or treated for 24 h with Tam. Interaction frequency with an anchor in the M*yc* promoter (A1, green arrow) is plotted relative to genomic location. A unidirectional forward primer strategy was used to probe chromatin interactions within the *Myc- Pvt1* locus and primers were designed to query previously described enhancer regions interacting with the *Myc* promoter (Cho et al., Cell 2018), published H3K4me1 peaks (Meyer et al., JBC 2016), and control regions. *Top* Interaction frequency plots depicting the strongest interactions between the *Myc* promoter and upstream and downstream enhancers. Note that the chromatin architecture is largely unchanged by the presence of stress. *Bottom* Interaction frequency plot depicting weaker interactions between the *Myc* promoter and upstream and downstream enhancers. Note the change of scale and that the three strongest peaks, depicted in the *Top* panels, are excluded in the *Bottom* panels to allow visualization of smaller peaks. Note that there is a pre-existing chromatin looping interaction between the *Myc* and *Pvt1* promoters, which exhibits a mild increase following Doxo treatment in (A) WT MEFs but is not significantly changed in (B) Tam-treated compared to untreated *KPR* cells. Data represent (A) n=3 biological replicates or (B) representative plot from n=2 biological replicates, ns = not significant, *p<0.05, paired t-test.

Accumulation of *Pvt1b* in the chromatin surrounding the *Pvt1-Myc* locus

To gain insight into the potential regulatory function of *Pvt1b*, I performed single-molecule RNA Fluorescence *in situ* Hybridization (smRNA-FISH), which allows visualization of individual RNA molecules by utilizing multiple fluorescently-labeled probes per transcript. I designed four independent probesets to detect *Pvt1* transcripts. *Pvt1a*- and *Pvt1b*-specific probesets (named *Pvt1a (ex.1a)* and *Pvt1b (ex.1b)*) were designed against the first exon of each isoform. While isoform-specific, the two probesets were not expected to detect single RNA molecules due to the low number of probes per transcript. The probeset *Pvt1 (ex.1a-10)* was designed to detect both full-length *Pvt1a* and full-length *Pvt1b* at single-molecule resolution, while the *Pvt1 (introns)* probeset was specific to unspliced Pvt1 molecules. Finally, I designed a probeset to detect *Myc* intronic regions (*Myc (intron))* and mark the site of *Myc* transcription. I observed that *Pvt1a* and *Pvt1b* exhibited a primarily 2- or 4-dot nuclear pattern in Etoposide (Etop)-treated MEFs, reflective of G1 or S/G2 stages of the cell cycle, respectively (Figures 12A and 13A). *Pvt1a* and *Pvt1b* formed larger clouds in Tam-treated *KPR* cells (Figures 12B and 13B), which have amplified the locus, as shown by DNA Fluorescence *in situ* Hybridization (DNA-FISH) (Figure 13C). By co-staining either *Pvt1a* or *Pvt1b* with total *Pvt1*, I concluded that both isoforms exhibited an identical localization pattern (Figures 12A, 12B, 13A and 13B). Notably, *Pvt1a*- and *Pvt1b*-containing foci co-localized with signals specific to the introns of nascent *Myc* (Figures 12C and 12D) as well as with nascent *Pvt1* transcripts (Figures 12E and 12F). These results led me to conclude that, following transcription, *Pvt1a* and

Pvt1b are retained on the chromatin surrounding the *Pvt1-Myc* locus. Subcellular fractionation analysis confirmed enrichment of both *Pvt1* variants in the chromatin fraction (Figure 12G).

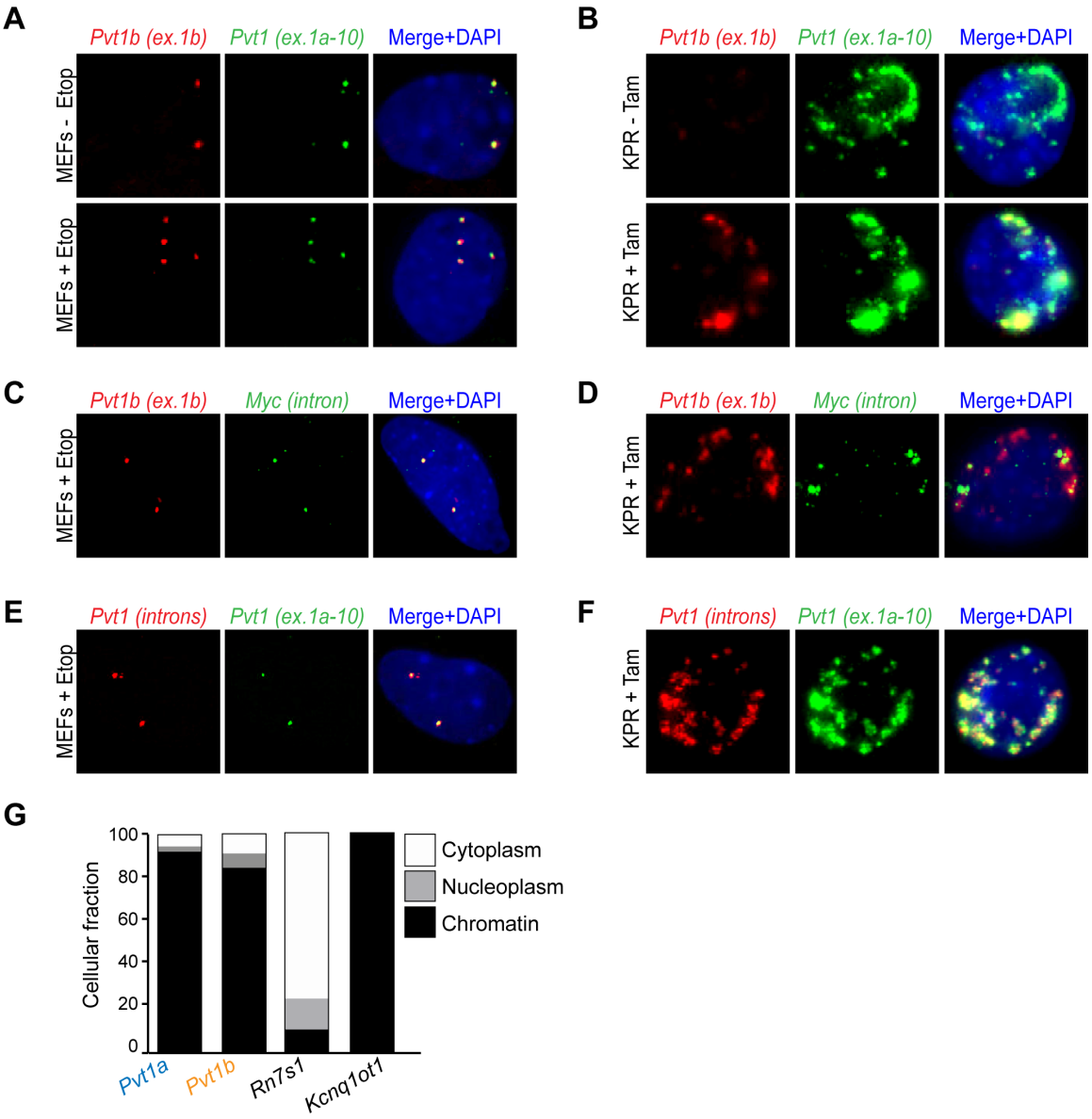

Figure 12. Accumulation of *Pvt1* isoforms in the chromatin surrounding the *Pvt1-Myc* locus. (A-F) smRNA-FISH with indicated probes in (A, C, E) WT MEFs, untreated or treated with Etop for 24 h and in (B, D, F) *KPR* cells untreated or treated with Tam for 24 h. DNA, DAPI. Note: *Pvt1b* is detectable in untreated, p53-proficient MEFs likely due to activation of the p53 pathway by passaging in primary cells but is undetectable in untreated, p53-deficient *KPR* cells. **(G)** *Pvt1a* and *Pvt1b* RNA levels in Doxo-treated WT MEFs following subcellular fractionation (representative from n=2 biological replicates). *Rn7s1* and *Kcnq1ot1* used as controls for the cytoplasmic and chromatin fractions, respectively. Subcellular fractionation and analysis performed by Ephrath Tesfaye.

Contributions from E. Tesfaye in (G) are described above.

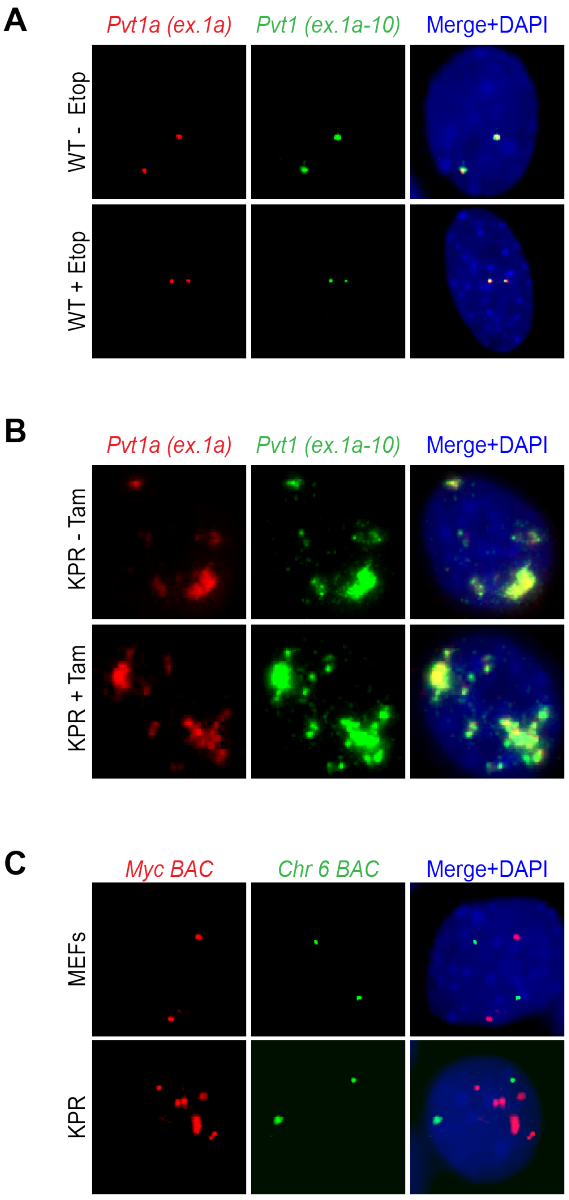

Figure 13. Co-localization of *Pvt1*a and total *Pvt1* by smRNA-FISH. (A-B) smRNA-FISH with probes designed against indicated regions in (A) WT MEFs, untreated or treated with Etop for 24 h and in (B) *KPR* cells untreated or treated with Tam for 24 h. The following probesets are shown: *Pvt1a (ex.1a, red)* detecting *Pvt1a* isoform with 11 probes spanning exon 1a and *Pvt1 (ex.1a-10, green)* detecting total *Pvt1* with 48 probes spanning exons 1a-10. Note: The *Pvt1a* probeset does not detect at the single molecule level. **(C)** DNA-FISH with probes generated using a Bacterial Artificial Chromosome (BAC) of the Myc locus (*Myc BAC, red*) or a control region in chromosome 6 (*Chr 6 BAC, green*) in untreated WT MEFs and *KPR* cells, highlighting increased copy number of the *Myc* locus in *KPR* cells. DNA-FISH performed by Dorthy Fang.

Contributions from D. Fang in (C) are described above.

Pvt1b RNA represses *Myc* levels *in cis*

Based on the stress-dependent expression of *Pvt1b* and its local chromatin accumulation, I hypothesized that *Pvt1b* could be involved in *Myc* repression through an RNA-dependent mechanism. To directly test this hypothesis, I designed three independent antisense oligonucleotides (ASOs) specific to exon 1b (Figure 14A). I used a non-targeting ASO (CON) as a negative control. As ASOs lead to co-transcriptional RNA cleavage and degradation, ASO1, 2, and 3 significantly downregulated both *Pvt1a* and *Pvt1b* (Figure 14B).

Next, I examined how *Pvt1*-targeting ASOs affected *Myc* expression levels. In untreated MEFs, *Myc* RNA and protein levels were not significantly altered in ASO compared to CON samples, indicating that knockdown of *Pvt1* isoforms did not affect *Myc* regulation in the absence of stress, consistent with previous findings (Figures 14C, 14D and 15A) (Cho et al., 2018). As expected, upon treatment with Doxo, CON MEFs experienced a significant decrease in *Myc* RNA (Figure 14C) and protein levels (Figures 14D and 15A). On the other hand, I found that *Pvt1*-targeting ASOs completely rescued stress-induced downregulation of *Myc* RNA and protein (Figures 14C, 14D and 15A). These findings revealed that transcriptional activation of *Pvt1b* by p53 is required for *Myc* repression during stress. As a control, the absence of *Myc* downregulation was not due to altered association of p53 with the *Pvt1b*-associated p53RE (Figure 15B).

To test the sufficiency of *Pvt1b* in suppressing *Myc*, Elena Martínez-Terroba employed the CRISPR-SAM (Synergistic Activation Mediator) system to activate the expression of endogenous *Pvt1b* in p53-deficient cells (Dahlman et al., 2015). CRISPR-SAM combines nuclease-proficient Cas9 with 15-nucleotide 'dead RNAs'

(dRNAs), which are competent for Cas9 recruitment but do not support Cas9 nuclease activity. In CRISPR-SAM, the dRNA scaffold is extended by two MS2 binding loops (dRNA-MS2), which serve to recruit the MS2-binding protein (MBP) fused to the transcriptional activator domains of p65 and HSF1, allowing CRISPR activation (CRISPRa) of target genes (Dahlman et al., 2015). Martínez-Terroba designed A1 and A2 dRNA-MS2 targeting the promoters of *Pvt1a* and *Pvt1b*, respectively (Figure 14E). Compared to a non-targeting control (C), CRISPRa using A1 led to 1.6-fold induction of *Pvt1a*, without altering *Pvt1b* levels, while A2 resulted in a 20-fold activation of *Pvt1b* with no significant induction of *Pvt1a* (Figure 14F). Next, she examined the effect of activation of endogenous *Pvt1a* and *Pvt1b* on *Myc* levels. In support of the model, she found that CRISPRa of *Pvt1b*, but not *Pvt1a,* was sufficient to significantly repress *Myc* RNA in p53-deficient cells compared to control dRNA-expressing cells (p=0.023, Figure 14G). Activation of *Pvt1b* did not further downregulate *Myc* levels following p53 restoration, indicating that *Pvt1b* acted downstream of p53 (Figure 14G). On the other hand, activation of *Pvt1b* was not sufficient to suppress Myc protein levels, opening the possibility for *Pvt1b*-independent input at the post-transcriptional level (Figures 15C and 14H).

To distinguish between activity *in cis* versus *in trans*, I tested whether exogenous overexpression of *Pvt1a* and *Pvt1b* by transfection of cDNA constructs containing exons 1a-10 (1a) or 1b-10 (1b) affected Myc expression (Figure 14I). I observed a 6.5-fold overexpression of *Pvt1a* as well as a 23-fold overexpression of *Pvt1b,* which were comparable to CRISPRa-induced overexpression (Figure 14J). However, I found that exogenously delivered *Pvt1a* or *Pvt1b* did not significantly

affect *Myc* RNA or protein levels, arguing against an effect *in trans* (Figures 14K, 14L and 15D). Altogether, these data supported a previously unappreciated role for *Pvt1b,* but not *Pvt1a,* in the repression of *Myc in cis*.

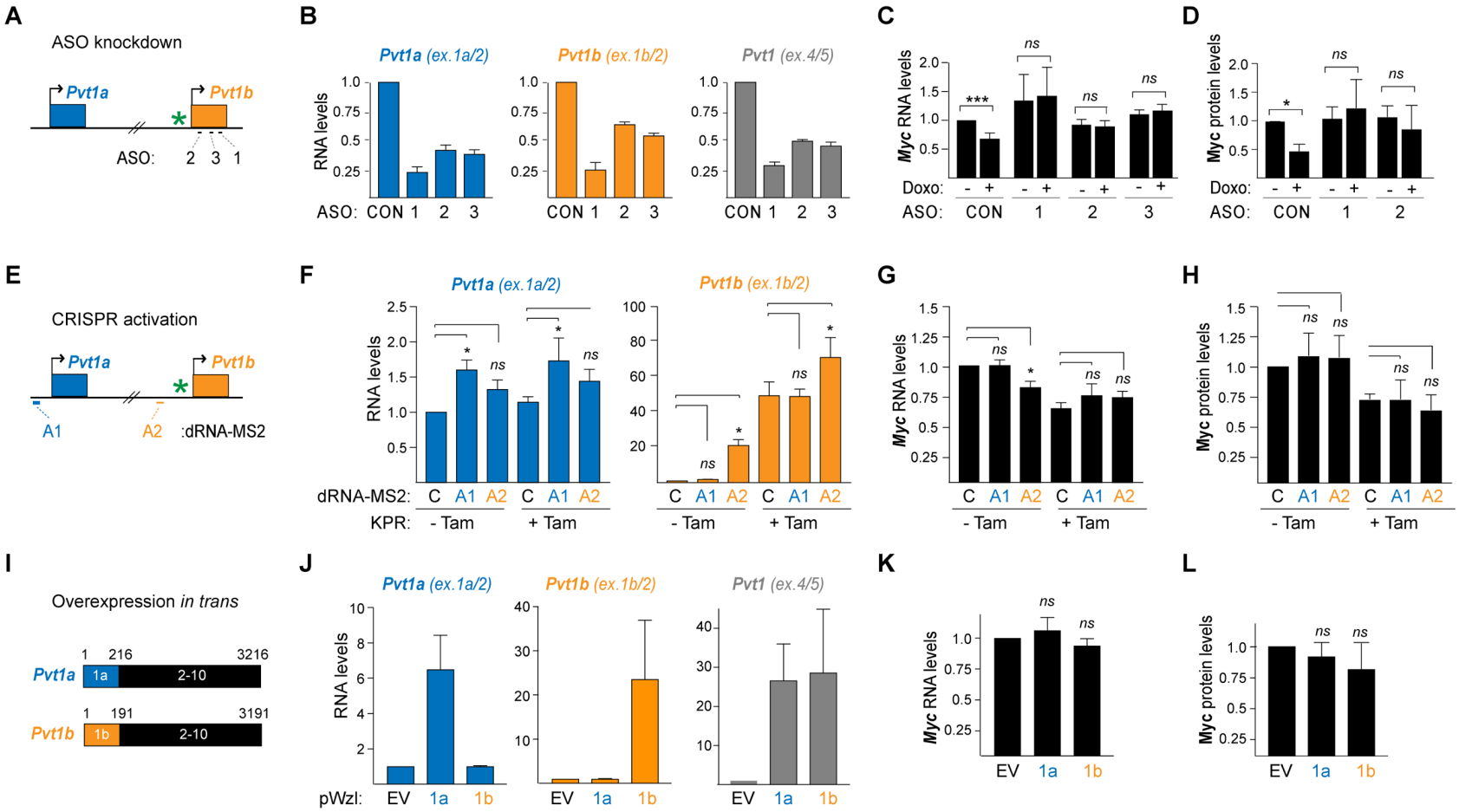
A
ASO knockdown
Pvt1a
Pvt1b
ASO: 2 3 1
B
Pvt1a (ex.1a/2)
Pvt1b (ex.1b/2)
Pvt1 (ex.4/5)
RNA levels
ASO: CON 1 2 3
C
Myc RNA levels
Doxo: - + - + - + - +
ASO: CON 1 2 3
D
Myc protein levels
Doxo: - + - + - +
ASO: CON 1 2
E
CRISPR activation
Pvt1a
Pvt1b
A1
A2
:dRNA-MS2
F
Pvt1a (ex.1a/2)
Pvt1b (ex.1b/2)
RNA levels
dRNA-MS2: C A1 A2 C A1 A2
KPR: - Tam + Tam
G
Myc RNA levels
C A1 A2 C A1 A2
- Tam + Tam
H
Myc protein levels
C A1 A2 C A1 A2
- Tam + Tam
I
Overexpression in trans
Pvt1a
1 216 3216
1a 2-10
Pvt1b
1 191 3191
1b 2-10
J
Pvt1a (ex.1a/2)
Pvt1b (ex.1b/2)
Pvt1 (ex.4/5)
RNA levels
pWzl: EV 1a 1b
K
Myc RNA levels
EV 1a 1b
L
Myc protein levels
EV 1a 1b

Figure 14. Production of *Pvt1b* RNA suppresses *Myc* expression *in cis*. **(A)** Schematic of ASO design. * denotes p53RE. **(B)** Isoform-specific and total *Pvt1* RNA levels in WT MEFs transfected with indicated control (CON) or *Pvt1*-targeting ASOs and harvested 24 h post Doxo treatment. Data are normalized to CON and show mean ± SEM (n=3, biological replicates). **(C)** *Myc* RNA levels in cells from (B), untreated or treated with Doxo for 24 h. Data are normalized to CON-Doxo and show mean ± SEM (n=3, biological replicates), ***p<0.001, ns = not significant, paired t test. **(D)** Quantification of Myc protein levels in cells from (B). Data are normalized to CON-Doxo and show mean ± SEM (n=3, biological replicates), *p<0.05, ns = not significant, paired t test. **(E)** Schematic of CRISPRa dRNA design. * denotes p53RE. CRISPRa cell lines generated by Elena Martínez-Terroba. **(F)** *Pvt1a* and *Pvt1b* RNA levels following *Pvt1a* (A1) or *Pvt1b* (A2) transcriptional activation in *KPR* cells, untreated or treated with Tam for 24 hours. Data are normalized to control dRNA (C) and show mean ± SEM (n=5, biological replicates), ns = not significant, *p<0.05, paired t test. Data collection and analysis performed by Elena Martínez-Terroba. **(G)** *Myc* RNA levels from experiment in (F). Data collection and analysis performed by Elena Martínez-Terroba. **(H)** Quantification of Myc protein levels in cells from (F). Data show mean ± SEM (n=3, biological replicates), ns = not significant, paired t test. Protein samples provided by Elena Martínez-Terroba. **(I)** Schematic of *Pvt1a* and *Pvt1b* overexpression constructs. **(J)** Isoform-specific and total *Pvt1* RNA levels in WT MEFs transiently overexpressing full length *Pvt1a* (1a) or *Pvt1b* (1b). Data are normalized to empty vector (EV) and show mean ± SEM (n=3, biological replicates), ns = not significant, paired t-test. **(K)** *Myc* RNA levels from experiment in (J). **(L)** Quantification of Myc protein levels in cells from (J). Data show mean ± SEM (n=3, biological replicates), ns = not significant, paired t test.

Contributions from E. Martínez-Terroba in (E), (F), (G) and (H) are described above.

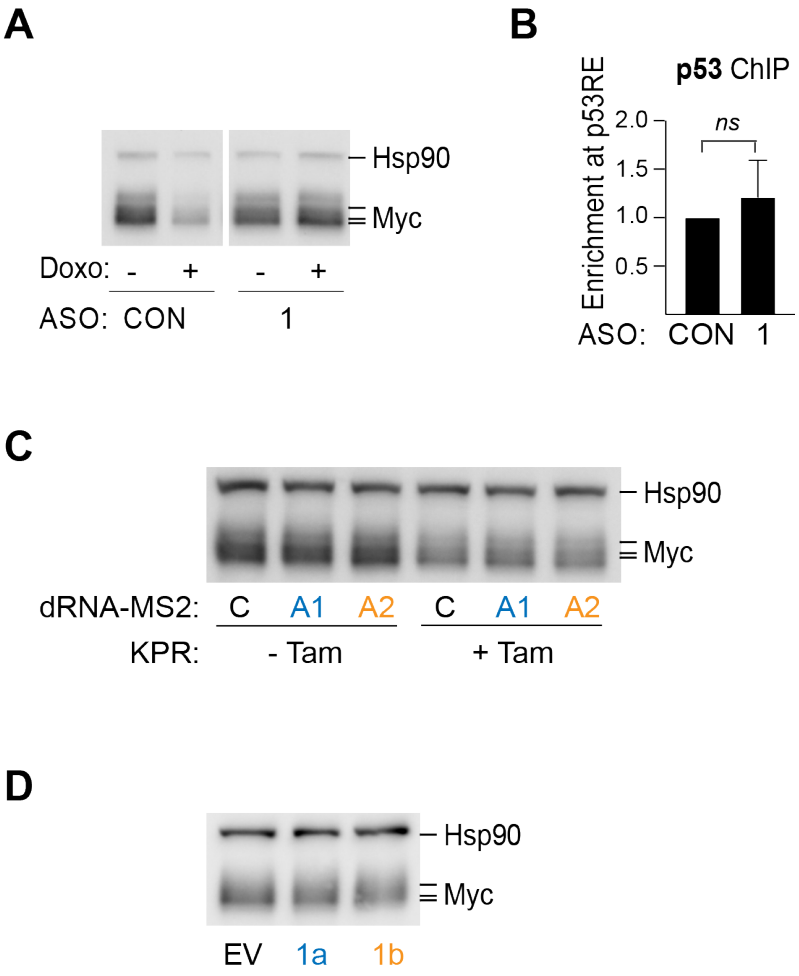

Figure 15. Effects of *Pvt1b* manipulation *in cis* and *in trans* on Myc protein levels. (A) Representative immunoblot analysis of Myc protein levels in whole-cell extracts isolated from ASO knockdown experiments in Fig. 4D. Hsp90 as a loading control. **(B)** ChIP-qPCR analysis showing the enrichment of p53 binding at p53RE relative to input in CON- and ASO1- treated WT MEFs following 8 h Doxo treatment. Data represented mean ± - SEM of biological replicates, ns = not significant, paired t-test. P53 ChIP performed by Nadya Dimitrova. **(C)** Representative immunoblot analysis of Myc protein levels in whole-cell extracts isolated from CRISPRa experiments in Fig. 4H. Hsp90 as a loading control. Protein samples provided by Elena Martínez-Terroba. **(D)** Representative immunoblot analysis of Myc protein levels in whole-cell extracts isolated from exogenous overexpression experiments in Fig. 4L. Hsp90 as a loading control.

Contributions from N. Dimitrova in (B) and E. Martínez-Terroba in (C) are described above.

Genetic inhibition of *Pvt1b* reverses stress-induced *Myc* downregulation

To investigate the functional contribution of *Pvt1b* to the p53 tumor suppressor pathway, I developed a genetic approach to specifically inhibit *Pvt1b* expression by mutating the p53RE required for its expression. I targeted Cas9 to the *Pvt1b* p53RE by designing a guide RNA (ΔRE) adjacent to the GGG protospacer adjacent motif (PAM) site located in the central region of the p53 consensus binding motif (Figure 16A). A non-targeting gRNA (Con) was used as a negative control. I generated control (Con) and mutant (ΔRE) *KPR* population, MEF population, and *KPR* clonal cell lines, which contain numerous or clone-specific CRISPR/Cas9-induced mutations of the *Pvt1b*-associated p53RE. I confirmed mutagenesis of the p53RE by Sanger sequencing (Figures 16A, 17A and 17B) and ChIP showed that ΔRE mutagenesis reduced p53 binding by 15-fold (Figure 16B). Importantly, by qRT-PCR, *Pvt1b* levels were significantly suppressed in ΔRE cells compared to controls (Figures 16C, 17C, 17D and 17G), and, by smRNA-FISH, I observed loss of *Pvt1b*-specific signal in Tam-treated ΔRE *KPR* cells compared to Tam-treated controls (Figures 16D and 16E). These observations led to the conclusion that mutagenesis of the *Pvt1b*-associated p53RE leads to efficient abrogation of stress-dependent *Pvt1b* activation.

Next, I queried whether ΔRE mutagenesis led to isoform-specific inhibition. By qRT-PCR and smRNA-FISH, I found that *Pvt1a* RNA levels and localization pattern were not significantly altered in ΔRE *KPR* population and clonal cell lines compared to controls, indicating that mutation of the p53RE led to specific

inhibition of *Pvt1b* in *KPR* cells (Figures 16C, 16D, 17C and 17D). On the other hand, mutagenesis of the p53RE in MEFs led to a significant reduction of *Pvt1a* (Figure 17G), consistent with our findings that *Pvt1a* expression has a p53-dependent component in this cell type (Figure 9B).

Finally, I examined by qRT-PCR and immunoblotting the effects of the ΔRE mutation and the resulting loss of *Pvt1b* expression on *Myc* levels during the cellular response to stress. In Con *KPR* population, *KPR* clonal, and MEF lines, exposure to oncogenic or genotoxic stress led to the expected significant decrease in *Myc* RNA (Figures 16F, 17E, 17F and 17H) and protein levels (Figures 16G, 16H, 17I and 17J). In contrast, exposure to stress in ΔRE *KPR* population, K*PR* clonal, and MEF lines did not lead to a significant decrease in *Myc* RNA levels compared to unstressed cells, consistent with the ASO data (Figures 16F, 17E, 17F and 17H). These results provided an independent, genetic confirmation that *Pvt1b* regulates *Myc* RNA levels downstream of p53.

Interestingly, while Myc protein levels were significantly elevated in ΔRE *KPR*+Tam and ΔRE MEF+Doxo lines compared to Con *KPR*+Tam and Con MEF+Doxo lines, respectively, the rescue was not complete (Figures 16G, 16H, 17I and 17J), consistent with the possibility of *Pvt1b*-independent regulatory input at the post-transcriptional level (Figure 14H).

Of note, mutagenesis of the *Pvt1b*-associated p53RE did not impact the long-range chromatin interactions in the locus, consistent with chromatin architecture not playing a significant role in p53-mediated Myc repression (Figure 18A).

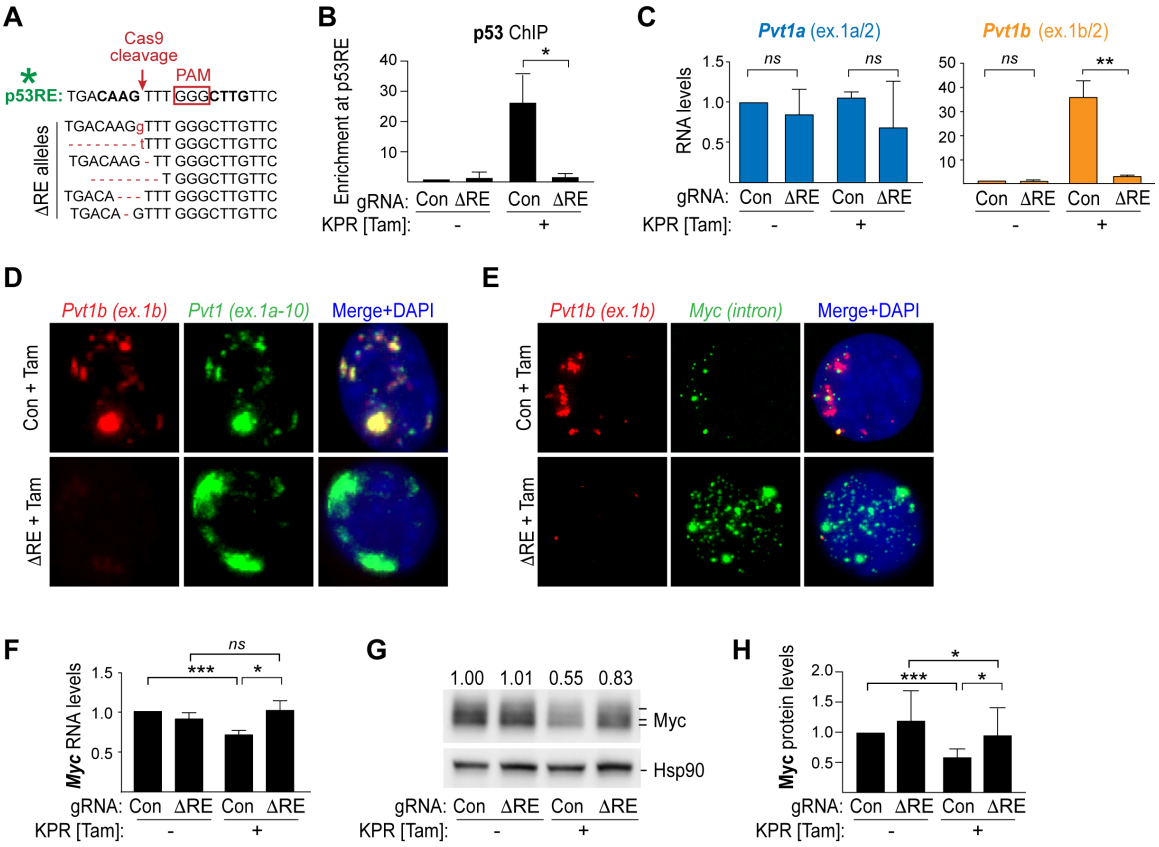

Figure 16. Genetic inhibition of *Pvt1b* leads to increased *Myc* levels. (A) *Top* Schematic of p53RE mutagenesis, indicating the PAM site (red box) and Cas9 cleavage site (red arrow). *Bottom* Mutant alleles, determined by Sanger sequencing. **(B)** ChIP-qPCR analysis of p53 enrichment at *Pvt1b*-associated p53RE in indicated cells and treatments. Data show mean ± SEM (n=3, biological replicates) *p<0.05, paired t test. P53 ChIP performed by Nadya Dimitrova. **(C)** *Pvt1a* and *Pvt1b* RNA levels in indicated cells and treatments. Data show mean ± SEM (n=3, biological replicates), **p<0.01, ns = not significant, paired t test. **(D, E)** smRNA-FISH of *Pvt1b* (ex.1b, red) co-localized with (D) total *Pvt1* (ex1a-10, green) or (E) nascent *Myc* (intron, green) in indicated cells and treatments. DNA, DAPI. **(F)** *Myc* RNA levels in indicated cells and treatments. Data show mean ± SEM (n=3, biological replicates), *p<0.05, ***p<0.001, ns = not significant, paired t test. **(G)** Representative image and quantification of Myc protein levels in indicated cells and treatments. Hsp90 as a loading control. **(H)** Quantification of Myc protein levels from experiments in (G). Data show mean ± SEM (n=6, biological replicates), *p<0.05, ***p<0.001, paired t test.

Contributions from N. Dimitrova in (B) are described above.

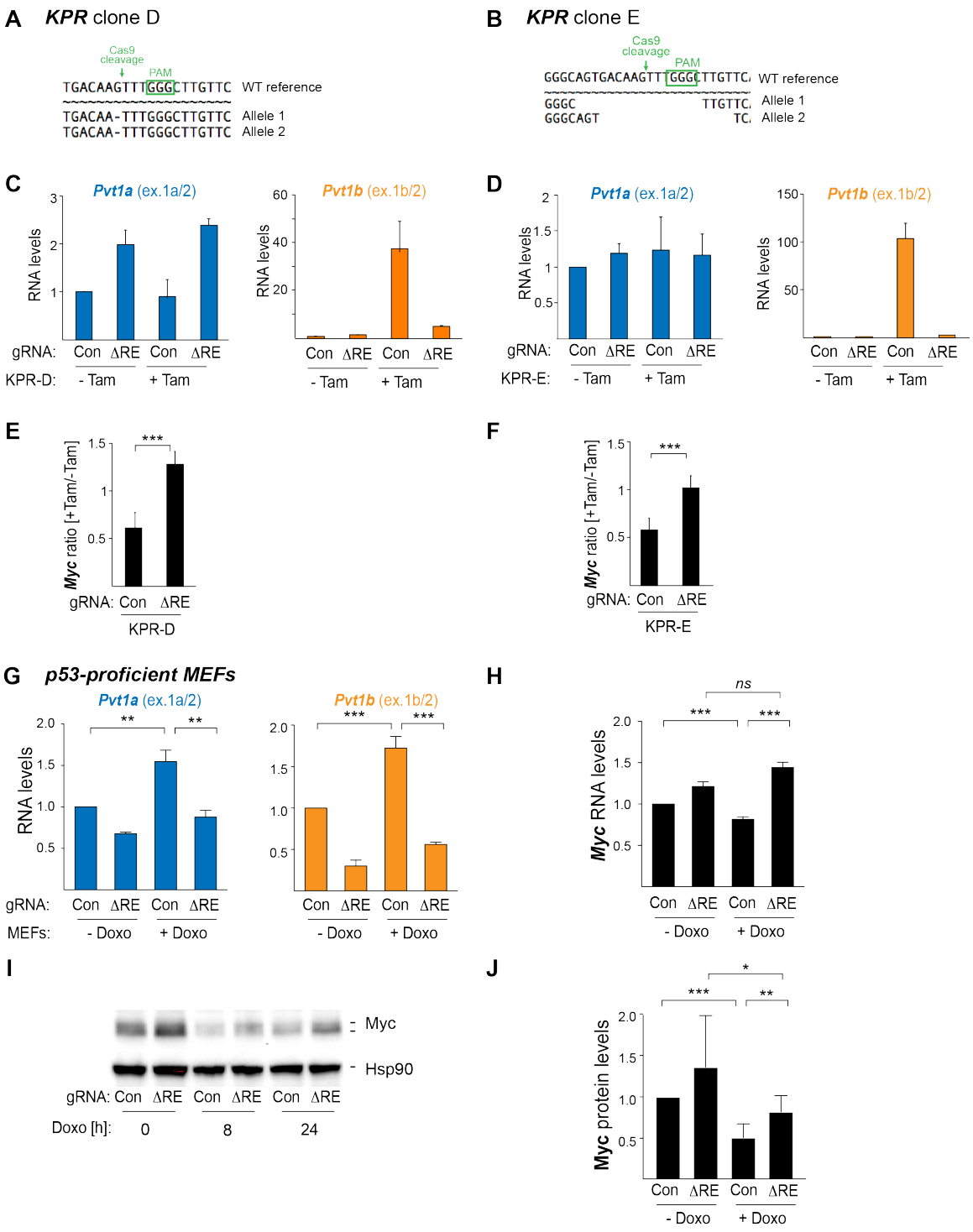

Figure 17. Genetic inhibition of *Pvt1b* rescues stress-dependent *Myc* repression in *KPR* clonal and MEF cell lines. (A-F) Analysis of (A, C, E) clone D and (B, D, F) clone E, isolated from *KPR* cells infected with a gRNA targeting the *Pvt1b* p53RE (ΔRE). As a control, *KPR* clones expressing Con gRNA were analyzed. **(A, B)** Sanger sequencing of the region containing the *Pvt1b*-associated p53RE. **(C, D)** qRT-PCR analysis of relative *Pvt1a* and *Pvt1b* RNA levels in indicated samples. **(E, F)** qRT-PCR of the ratio (+Tam/-Tam) of *Myc* expression in indicated clones. Bars show the mean ± SEM of n=3 biological replicates, ***p<0.001, paired t-test. **(G-J)**

Analysis of MEF population cell lines, infected with a gRNA targeting the *Pvt1b* p53RE (ΔRE) or a non-targeting control (Con) and harvested untreated or 8-24 h post Doxo treatment. **(G)** qRT-PCR analysis of relative *Pvt1a* and *Pvt1b* RNA levels in indicated samples. Data represent mean ± SEM of n=3 biological replicates, **p<0.01, ***p<0.001, paired t- test. **(H)** qRT-PCR analysis of relative *Myc* RNA levels in indicated samples. Data represented as mean±SEM of n=3 biological replicates, ns = not significant, ***p<0.001, paired t- test. **(I)** Immunoblot analysis of Myc protein levels in whole-cell extracts from indicated cells. Hsp90 as a loading control. **(J)** Quantification of Myc protein levels from cells in (I). Data represent mean ± SEM of n=7 biological replicates, *p<0.05, **p<0.01, ***p<0.001, paired t- test.

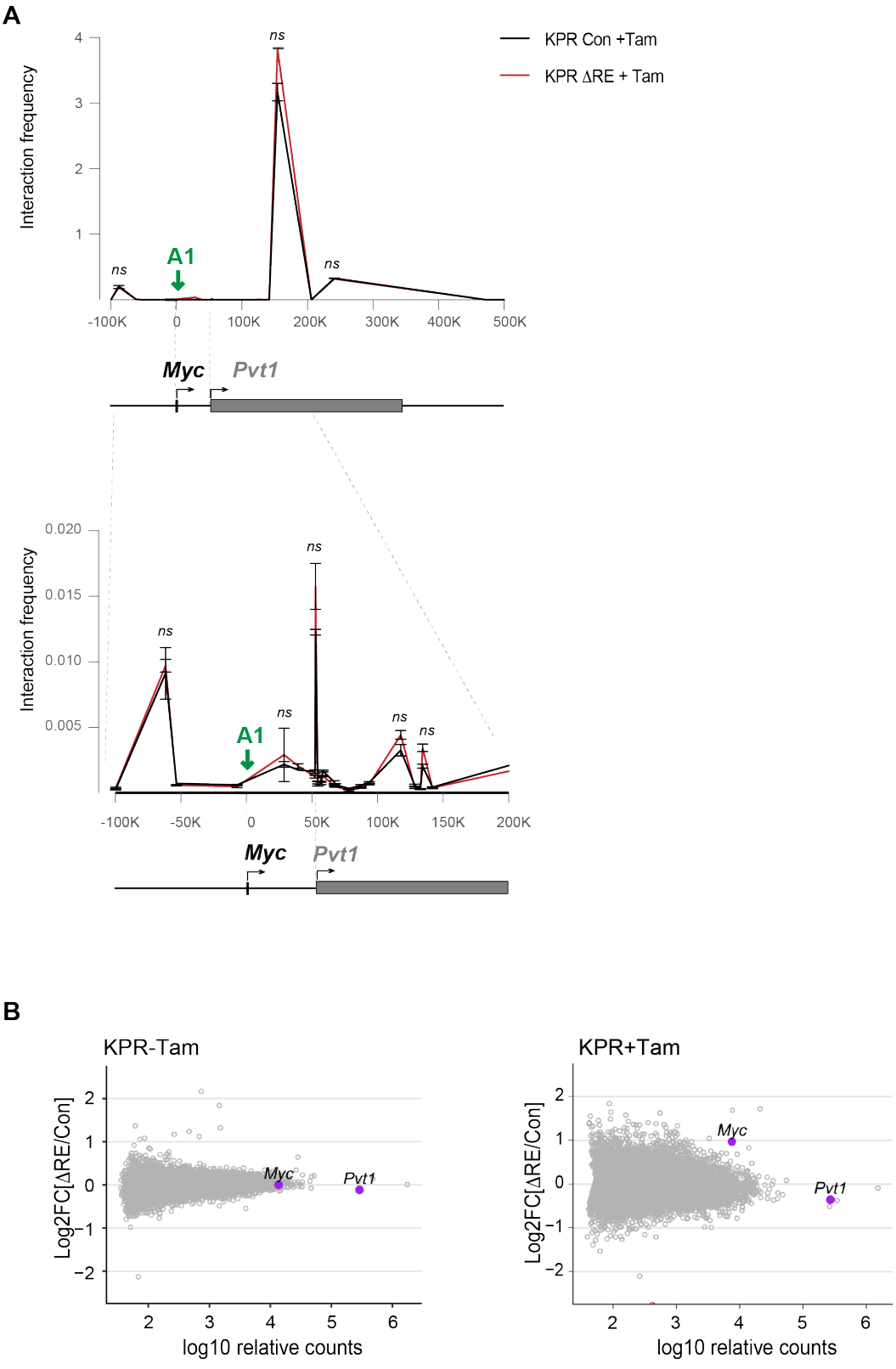

Figure 18. Effects of genetic inhibition of *Pvt1b* on the chromatin architecture of the *Myc-Pvt1* locus and *Myc* transcription. (A) 3C analysis in *KPR* cells, infected with *Pvt1b*

p53RE-targeting (ΔRE) or Control (Con) gRNAs and treated for 24 h with Tam. Interaction frequency relative to an anchor in the *Myc* promoter (A1, green arrow) is plotted from three technical replicates from a representative experiment of two biological replicates, as described in Figure S2. **(B)** MA plots from TT-TimeLapse-seq data depicting log2 fold change in nascent RNA (n >11,000 for each condition) in (A) untreated or (B) Tam-treated KPR cells. *Myc* and total *Pvt1* are indicated (purple dots). Processing of samples for TT-TimeLapse-seq and data analysis performed by Jeremy Schofield, Josh Zimmer and Matt Simon.

Contributions from J. Schofield, J. Zimmer, and M. Simon in (B) are described above.

Pvt1b suppresses Myc transcriptional activity and cellular proliferation *in vitro*

By analyzing the effects of the ΔRE mutation on gene expression in total RNA from untreated and Doxo-treated ΔRE and Con MEFs in collaboration with Nima Hooshdaran and Jesse Zamudio, I confirmed that *Myc* is a target of *Pvt1b* regulation in response to stress (Figure 19A). Next, to test whether *Pvt1b* acted at the transcriptional or post-transcriptional level, my collaborators Josh Zimmer, Jeremy Schofield and Matt Simon sequenced nascent RNA from untreated and Tam-treated ΔRE and Con *KPR* cells (Schofield et al., 2018). They found that nascent *Myc* transcripts were significantly upregulated in ΔRE+Tam compared to Con+Tam *KPR* cells, indicative of transcriptional regulation (Figures 19B and 19C). These data revealed that *Pvt1b* production promotes transcriptional suppression of *Myc*.

Next, Hooshdaran and Zamudio queried how the changes in *Myc* RNA levels affected the Myc transcriptional program by examining the consequence of *Pvt1b* loss on a curated set of 196 Myc target genes (Gene Set Enrichment Analysis, HALLMARK_MYC_TARGETS_V1 (Liberzon et al., 2015)). They plotted the cumulative frequency distribution of the fold change of Myc target genes in ΔRE

cells relative to Con cells in the presence of stress (logFC [ΔRE/Con+stress]). Compared to a randomly generated set of control genes expressed at comparable levels, they found a significant increase in the levels of Myc targets in MEFs and *KPR* cells (Figures 19D and 19E). They concluded that *Myc* derepression by ΔRE mutagenesis leads to a small but significant increase in the transcriptional activity of *Myc*.

Considering *Myc* target genes include factors that promote cellular growth, I compared the proliferation of mutant cells compared to controls. It has previously been shown that Tam-mediated p53 restoration in *KPR* cells leads to a permanent cell cycle arrest, called senescence (Feldser et al., 2010). While loss of *Pvt1b* expression did not overcome senescence, it led to a significant increase in cellular proliferation and colony formation compared to control cells (Figures 19F and 19G). As a control, the ΔRE mutation did not impact *Myc* levels and proliferation in p53-deficient cells, ruling out off target effects (Figures 18B and 19F). These data suggested that *Pvt1b* mediates specific aspects of p53 function to suppress the proliferative potential of cells *in vitro*.

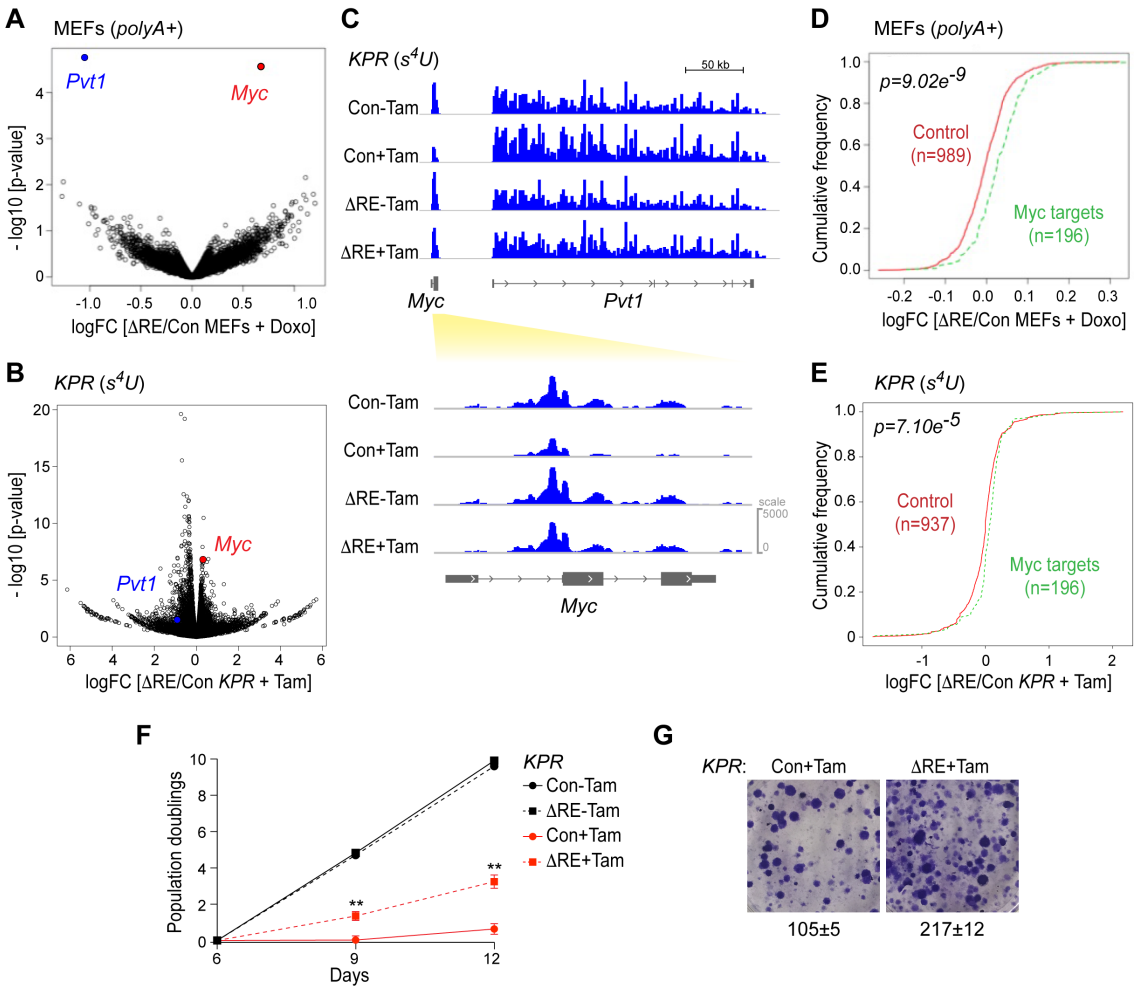

Figure 19. *Pvt1b* suppresses *Myc* transcription and proliferative function. (A, B) Butterfly plot depicting the fold change (logFC) in gene expression of indicated samples relative to statistical significance (-log10(p-value), MEF: n=3; *KPR*: n=2, biological replicates). Gene expression profiling was performed by (A) RNAseq of polyA-selected RNA isolated from Con or ΔRE gRNA-expressing MEFs, untreated or treated with Doxo for 24 hours or (B) TimeLapse-seq of ribosomal cDNA-depleted s4U-labeled RNA isolated from Con or ΔRE gRNA-expressing *KPR* cells, untreated or treated with Tam for 16 hours. Total *Pvt1* (blue) and *Myc* (red) are labeled. Library preparation performed by Nadya Dimitrova. Data analysis performed by Nima Hooshdaran and Jesse Zamudio. **(C)** *Top* Genome browser tracks depicting the *Myc-Pvt1* locus and *Bottom* Detail of the *Myc* locus from TT-TimeLapse-seq. Processing of samples for TT-TimeLapse-seq and data analysis performed by Jeremy Schofield, Josh Zimmer and Matt Simon. **(D, E)** Cumulative frequency distribution plot of differential expression for a set of curated Myc target genes and a matched set of control genes from analyses in (A, B). Library preparation performed by Nadya Dimitrova. Data analysis performed by Nima Hooshdaran and Jesse Zamudio. **(F)** Population doublings in Con or ΔRE gRNA-expressing *KPR* cells, untreated or treated with Tam over indicated timecourse. Data show mean ± SEM (n=3, biological replicates), **p<0.01, unpaired t test. **(G)** Representative images of colony formation assay of Tam-treated *KPR* cells, infected with Con or ΔRE gRNAs. Numbers show mean ± SEM (n=3, biological replicates), **p<0.01, unpaired t test.

Contributions from N. Hooshdaran and J. Zamudio in (A), (B), (D) and (E), and from J. Schofield, J. Zimmer, and M. Simon in (C) are described above.

Tumor-specific inhibition of *Pvt1b* promotes tumor growth *in vivo*

Inactivation of p53 in the *K-ras*$^{LSL\text{-}G12D/+}$*(K)* autochthonous mouse model of lung cancer has been shown to increase tumor burden and promote tumor progression from benign to aggressive disease (DuPage et al., 2009; Jackson et al., 2005; Jackson et al., 2001). To elucidate whether *Pvt1b* mediated some aspects of p53 function, Nadya Dimitrova and Clara Liao performed tumor-specific mutagenesis of the *Pvt1b*-associated p53RE (Figure 20A). Dimitrova built a bifunctional lentiviral construct (*U6-gRNA PGK-Cre, UGPC*) for co-expression of the ΔRE gRNA (UGPC-ΔRE) and Cre recombinase, required for Cas9 targeting and tumor initiation, respectively (DuPage et al., 2009). Expression of Cas9 in a tumor-specific manner was achieved by crossing the *K* model to *Rosa26-Cas9*^{LSL} (*C*) mice to generate *KC* animals (Platt et al., 2014). As a negative control, they used a non-targeting control (UGPC-Con). As a positive control, they used a previously described gRNA that targets the open reading frame of p53 (UGPC-p53KO) (Xue et al., 2014). Sanger sequencing confirmed successful mutagenesis of the *Pvt1b*-associated p53RE in UGPC-ΔRE-infected animals (Figure 20B).

They next examined hematoxylin and eosin (H&E) sections of lungs from mice infected with UGPC-Con, -p53KO and -ΔRE virus and sacrificed at 16 weeks post tumor initiation. In the K model, progression of atypical adenomatous hyperplasia (AAH, grade 1) and lung adenoma (grade 2) to adenocarcinoma (grade 3) and invasive adenocarcinoma (grade 4) is promoted by loss of p53 function (Jackson et al., 2005). Indeed, histopathological analysis revealed that all of the tumors (53/53 tumors) in UGPC-Con-infected animals manifested grade 1 features (Figures 20C and 20D). In contrast, 70% of UGPC-p53KO-expressing tumors

(39/56 tumors) were marked by atypical nuclei, desmoplasia, and transition to a poorly differentiated phenotype and were classified as grade 2 or 3 (Figures 20C and 20D) (DuPage et al., 2009). Based on these data, they estimated that a large portion of the tumors underwent successful CRISPR/Cas9 editing *in vivo*. Editing of the *Pvt1b*-associated p53RE resulted in tumors with histopathological features comparable to controls and only 3% of tumors (2/67 tumors) in UGPC-ΔRE-infected animals were classified grade 2 or 3, suggesting that tumor progression was not accelerated by *Pvt1b* inhibition (Figures 20C and 20D). They concluded that *Pvt1b* does not likely mediate the ability of p53 to restrain tumor progression from benign hyperplasia to advanced disease.

On the other hand, quantification of the tumor area relative to the total lung area revealed that the tumor burden in UGPC-ΔRE-infected animals (21±4%) was significantly increased compared to the burden of control mice (12±2%) (p=0.0040, Figure 20E). Notably, the tumor burden in p53RE-edited mice was comparable to the tumor burden in UGPC-p53KO-infected mice (26±3%) (Figure 20E). These findings suggested that *Pvt1b* mediated in large part the growth-restrictive functions downstream of p53, particularly during the pre-malignant stages of the disease. As a control for potential off-target effects of Cas9 expression and CRISPR editing, Ephrath Tesfaye used two independent sgRNAs (sg1 and sg2) to target the p53RE in intron 1 of an unrelated lncRNA, *Gm26542*, for which we had evidence for direct p53 regulation (Figures 21A, 21B and 21C). In contrast to *Pvt1b*, inhibition of *Gm26542* did not affect proliferation in Tam-treated *KPR* cells *in vitro* (Figure 21D) and did not significantly alter the tumor burden in *KC* mice *in vivo* (Figure 21E).

The increase in tumor burden in UGPC-ΔRE-infected animals compared to UGPC-Con mice was not due to decreased apoptosis as there was no evidence for Cleaved Caspase 3 (CC3) immunohistochemistry (IHC) staining in lung sections. Instead, the increase in tumor burden could be attributed to enhanced proliferation, as manifested by the significantly greater number of phosphorylated histone H3 (pHH3)-positive mitotic cells in *Pvt1b*-deficient tumors from UGPC-ΔRE-infected animals compared to tumors from UGPC-Con-infected mice (p=0.0026, Figures 20F and 20G).

Finally, to investigate whether *Pvt1b* acted downstream or independent of p53, Dimitrova and Liao performed an epistasis experiment. They generated cohorts of either *KC* or *K-ras*$^{LSL\text{-}G12D/+}$*; p53*$^{FL/FL}$*; Rosa26-Cas9*$^{LSL/LSL}$ *(KPC)* animals, which have genetically engineered Cre-inducible loss-of-function alleles of p53. They analyzed tumor burden at 12 weeks post tumor initiation with UGPC-Con or -ΔRE virus. Consistent with their findings above, they observed a significant increase in the tumor burden of UGPC-ΔRE-infected mice compared to UGPC-Con-infected *KC* animals (p=0.0035, Figure 20H). In contrast, they found that the tumor burden was not significantly different between UGPC-ΔRE and UGPC-Con-infected *KPC* animals (Figure 20H). Moreover, there was no statistically significant difference between the tumor burden of *KC* mice infected with UGPC-ΔRE and *KPC* mice infected with UGPC-Con (Figure 20H). Altogether, these results revealed that *Pvt1b* and p53 enhance the expansion of pre-malignant tumors through a common pathway.

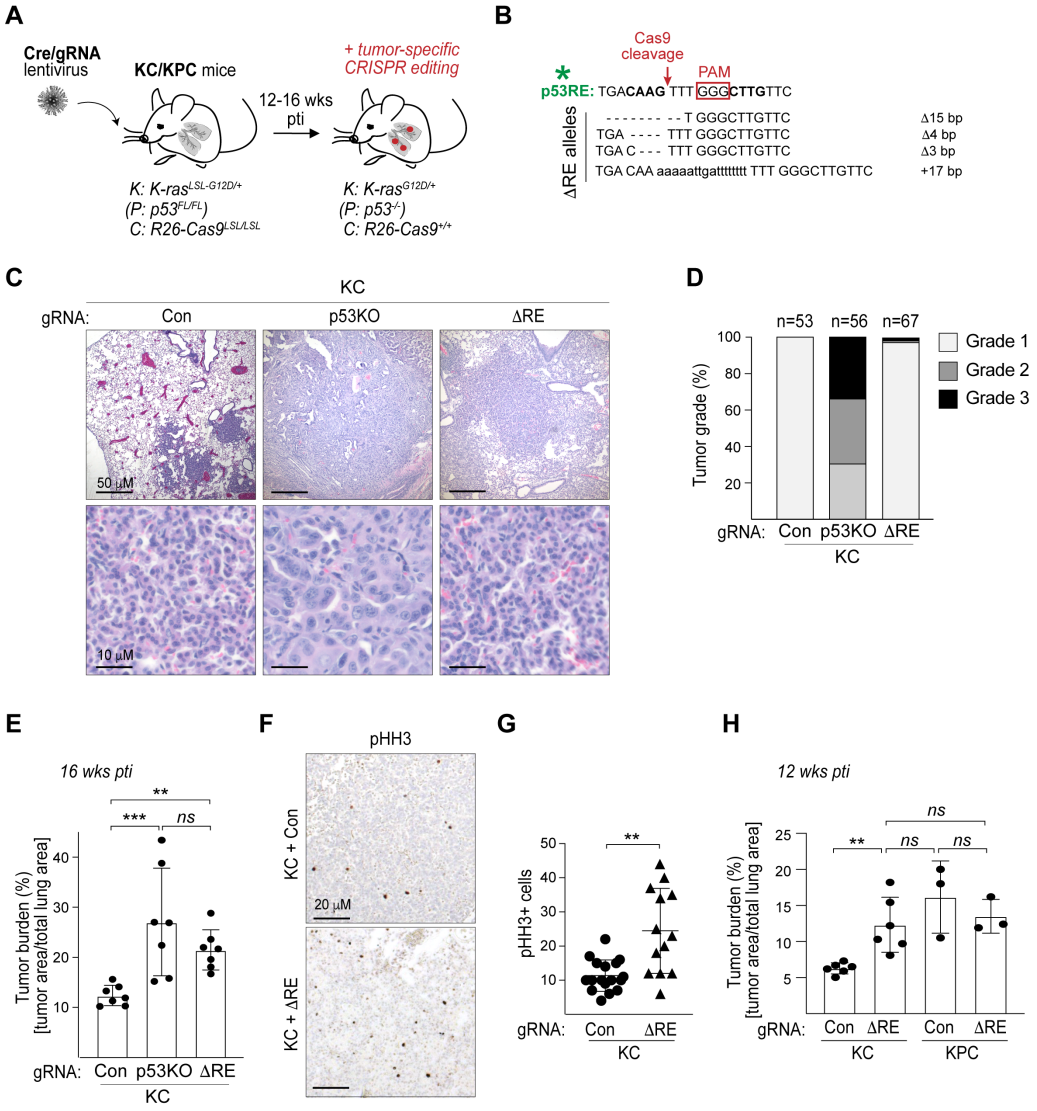

Figure 20. Tumor-specific editing in a lung cancer model reveals a role for *Pvt1b* in suppressing tumor growth, but not progression. (A) Schematic of tumor-specific gene editing in *KC* and *KPC* lung cancer mouse models. **(B)** Mutant ΔRE alleles, determined by Sanger sequencing of bulk DNA isolated from tumor-bearing lungs. **(C)** H&E staining of lung sections of *KC* mice infected with indicated gRNAs and analyzed at 16 weeks post tumor initiation (pti). Scale bars as indicated. **(D)** Quantification of tumor grade in mice described in (C). The number of tumors analyzed from n=5 mice is indicated for each group. **(E)** Quantification of tumor burden in mice described in (C). Dots represent individual animals and bargraph shows mean ± SEM (n=7 mice), ***p<0.001, **p<0.01, ns = not significant, unpaired t test. **(F)** Representative images of immunohistochemistry for the mitotic marker pHH3 in lung sections from (C). Scale bars as indicated. **(G)** Quantification of images in (F). Data show mean ± SEM of n=13-15 tumors from n=5 mice, **p<0.01, Mann-Whitney test. **(H)** Quantification of tumor burden in *KC* and *KPC* mice infected with indicated gRNAs and analyzed at 12 weeks pti. Dots represent individual animals and bargraph shows mean ± SEM (KC: n=6 mice, KPC: n=3 mice), *p<0.05, ns=not significant, unpaired t-test. All *in vivo* experiments and data analysis performed by Nadya Dimitrova and Clara Liao.

Contributions from N. Dimitrova and C. Liao in (A-H) are described above.

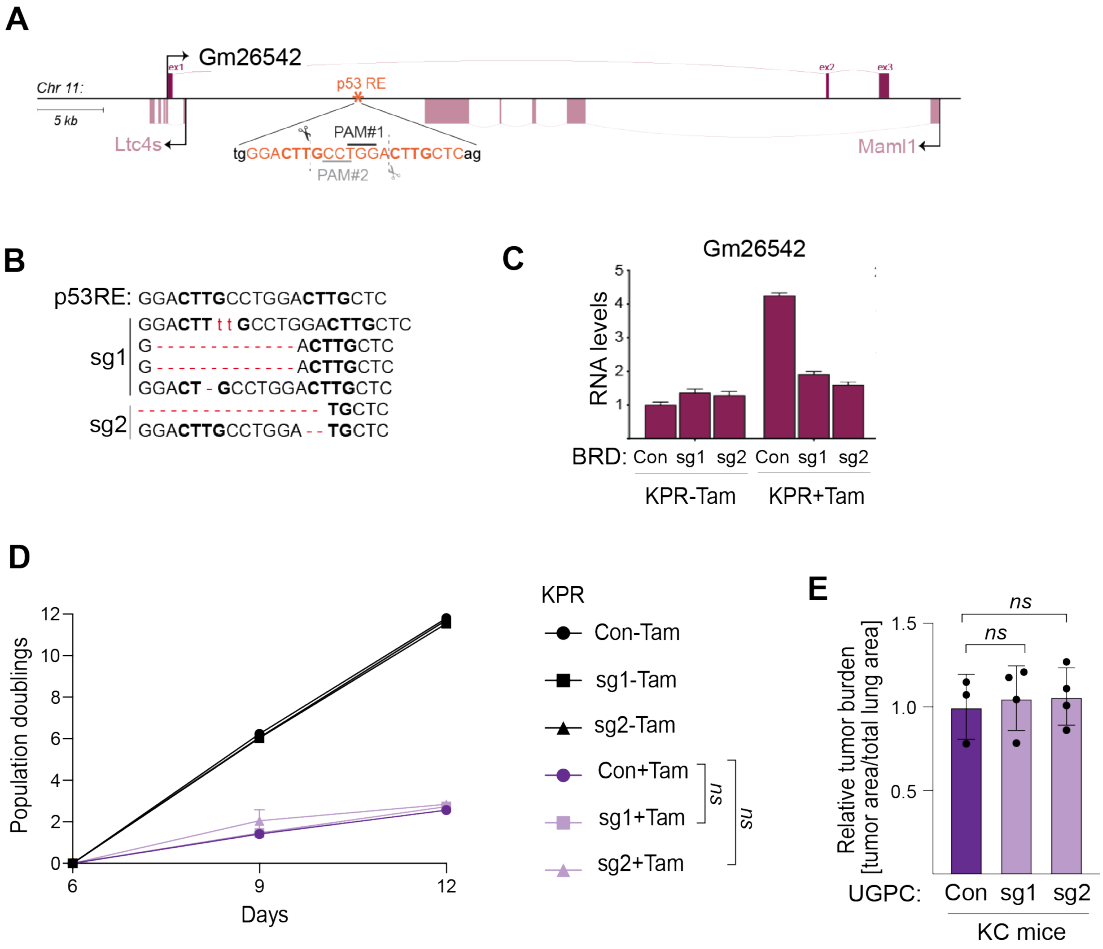

Figure 21. Mutagenesis of Gm26542-associated p53RE does not affect proliferation *in vitro*, tumor growth *in vivo*. (A) Schematic of the mouse Gm26542 lncRNA locus depicting the sequence of the p53RE located in intron 1, and the PAM sites (PAM #1 and PAM #2) utilized by guide RNAs sg1 and sg2, respectively. Schematic provided by Ephrath Tesfaye. **(B)** Sanger sequencing of the region containing the Gm26542-associated p53RE in *KPR* cells infected with sg1 or sg2. Data collection and analysis performed by Ephrath Tesfaye. **(C)** qRT-PCR analysis of relative RNA levels of the Gm26542 lncRNA in Con-, sg1- and sg2- infected *KPR* cell populations showing the p53-dependent induction of Gm26542 at 24 hours post Tam treatment and the abrogation of this induction by p53RE mutagenesis. Data collection and analysis performed by Ephrath Tesfaye. **(D)** Growth analysis showing population doublings in Con, sg1, or sg2 gRNA-expressing *KPR* cells, untreated or treated with Tam. Data show mean ± SEM of n=3 biological replicates, ns= not significant, unpaired t-test. Cell lines provided by Ephrath Tesfaye. **(E)** Quantification of tumor burden as tumor area relative to total lung area in KC mice infected with indicated gRNAs and analyzed at 16 weeks post tumor initiation (pti) as described in Fig. 20. Data show tumor burden of individual mice and mean±SD, ns = not significant, unpaired t-test. *In vivo* experiments and data analysis performed by Nadya Dimitrova.

Contributions from E. Tesfaye in (A-D) and N. Dimitrova in (E) are described above.

Discussion

This study provides new mechanistic insights into the function of the lncRNA *Pvt1* in the context of the p53 tumor suppressor pathway. I identify a conserved isoform of *Pvt1, Pvt1b*, which is directly activated by p53 in response to genotoxic and oncogenic stress. My data reveal that production of *Pvt1b* functions as a p53-dependent mechanism that is wired into the *Myc-Pvt1* locus to directly and swiftly down-regulate *Myc* transcription during stress (Figure 22). This appears to be the primary mechanism underlying stress-induced Myc reduction at the transcriptional level, although these data are also consistent with *Pvt1b*-independent regulation at the post-transcriptional level.

Functionally, I observed that *Pvt1b* activation leads to restricted Myc levels and transcriptional activity and suppressed cellular proliferation. Furthermore, use of an autochthonous mouse model of lung cancer demonstrated that *Pvt1b* acts downstream of p53 during the early stages of cancer development to limit tumor growth. Strikingly, in this respect, epistasis analysis suggested that *Pvt1b* acts as the primary mediator of p53. On the other hand, I found that *Pvt1b* is not involved in other aspects of p53 function, such as promoting senescence or limiting tumor progression to advanced disease. Altogether, these analyses define the specific contributions of *Pvt1b* downstream of p53, pointing to growth limiting and tumor suppressive functions of *Pvt1b* in the context of cancer. These conclusions contrast the common classification of *Pvt1* as an oncogene, which is based on extensive correlative evidence linking *Pvt1* aberrations with increased invasive capacities of cancer cells and poor patient survival (Guan et al., 2007; Guo et al., 2018; Kong et al., 2015; Riquelme et al., 2014; Tseng et al., 2014; Zeng et al., 2017; Zhang et al.,

2019; Zhao et al., 2018; Zheng et al., 2016; Zhu et al., 2017). On the other hand, these data are consistent with recent reports of tumor suppressive elements in the *Pvt1* locus (Barsotti et al., 2012; Cho et al., 2018; Porter et al., 2017).

My findings shed light on a subset of genomic aberrations reported across a variety of malignancies, which represent translocations between the first exon of *Pvt1a* fused to various 3' gene partners (Iwakawa et al., 2013; Kim et al., 2014a; Nagoshi et al., 2012; Northcott et al., 2012). Such rearrangements would be expected to separate the *Myc* locus from *Pvt1b*, providing cells with a proliferative advantage due to the inability of p53 to suppress *Myc* levels during early stages of tumor development. On the other hand, the proposed tumor suppressive role of *Pvt1b* is at odds with the common amplification of the *Pvt1* locus in cancer (Guan et al., 2007; Riquelme et al., 2014). I propose that amplification of other elements, such as the *Pvt1a* transcript or *Pvt1*-associated *Myc* enhancers may be the drivers of oncogenic activities in this setting, as proposed by others (Cho et al., 2018; Tseng et al., 2014). Alternatively, these alterations might be occurring following p53 inactivation, which would preclude *Pvt1b* expression.

Mechanistically, I provide direct evidence for a role of *Pvt1b* RNA production in *Myc* regulation. Antisense-mediated depletion experiments reveal that *Pvt1b* is required for stress-induced *Myc* inhibition, whereas epigenetic activation from the endogenous locus shows that *Pvt1b* is sufficient to repress *Myc* in the absence of stress or a functional p53 pathway. While ASO-based knockdown and CRISPR-guided epigenetic experiments cannot formally differentiate between the mature *Pvt1b* molecules or the production of nascent *Pvt1b* transcripts as the mediator of *Myc* repression, these data support an RNA-based mechanism.

This conclusion differs from the recent finding that the *Pvt1a* promoter suppresses *Myc* levels in an RNA-independent manner (Cho et al., 2018). The discrepancy can potentially be explained by the previous focus on the constitutive *Pvt1* isoform, by the use of p53-deficient cell lines, or by the use of ineffective ASOs (Cho et al., 2018). Alternatively, I propose that the two tumor suppressive activities in the *Pvt1* locus, one p53- and RNA-dependent and the other p53- and RNA-independent, may co-exist and operate in distinct cellular contexts. My findings also do not contradict studies that have implicated *Pvt1a* or circular *Pvt1* isoforms as oncogenes via diverse mechanisms, such as oncoprotein stabilization or competition for miRNA binding (Tseng et al., 2014; Xu et al., 2017; Zhao et al., 2018). Indeed, the complexity of the *Pvt1* locus highlights the need for further rigorous dissection of the various alternative start site- and splice-variants.

It is important to note that *Pvt1b* mediates a repressive event downstream of p53, which is a well-characterized transcriptional activator. Considered in the context of the previously characterized p53-dependent *cis*-regulatory *lincRNA-p21* (Dimitrova et al., 2014), it appears that transcription factors use lncRNAs to either enhance their inherent activity or to allow reverse regulation within local circuits. LncRNAs which accumulate at their sites of transcription, such as *Pvt1b*, are poised to act as modulators of gene expression in a locus-specific manner. Indeed, *Pvt1b* activation leads to *Myc* repression within four hours of exposure to genotoxic stress, which is comparable to the kinetics of activation of p53 target genes. I propose that production and/or chromatin accumulation of p53-induced *Pvt1b* transcripts act *in cis* during the cellular response to stress to rapidly influence the transcriptional environment at the *Myc* promoter. Thus, locus-specific

transcriptional regulation by lncRNAs may provide additional tools within a transcriptional program that allow dynamic and swift responses to cellular challenges. As the mechanisms of more p53-dependent lncRNAs are revealed, we can gain new insight into how regulatory RNAs contribute to the cellular responses to stress mediated by p53. Although future work will determine the functional elements of *Pvt1b* transcripts, the widespread importance of this regulatory circuit in normal and transformed cells *in vitro* and *in vivo* suggests the possibility of controlling *Myc* levels in cancer by modulating *Pvt1b* activity.

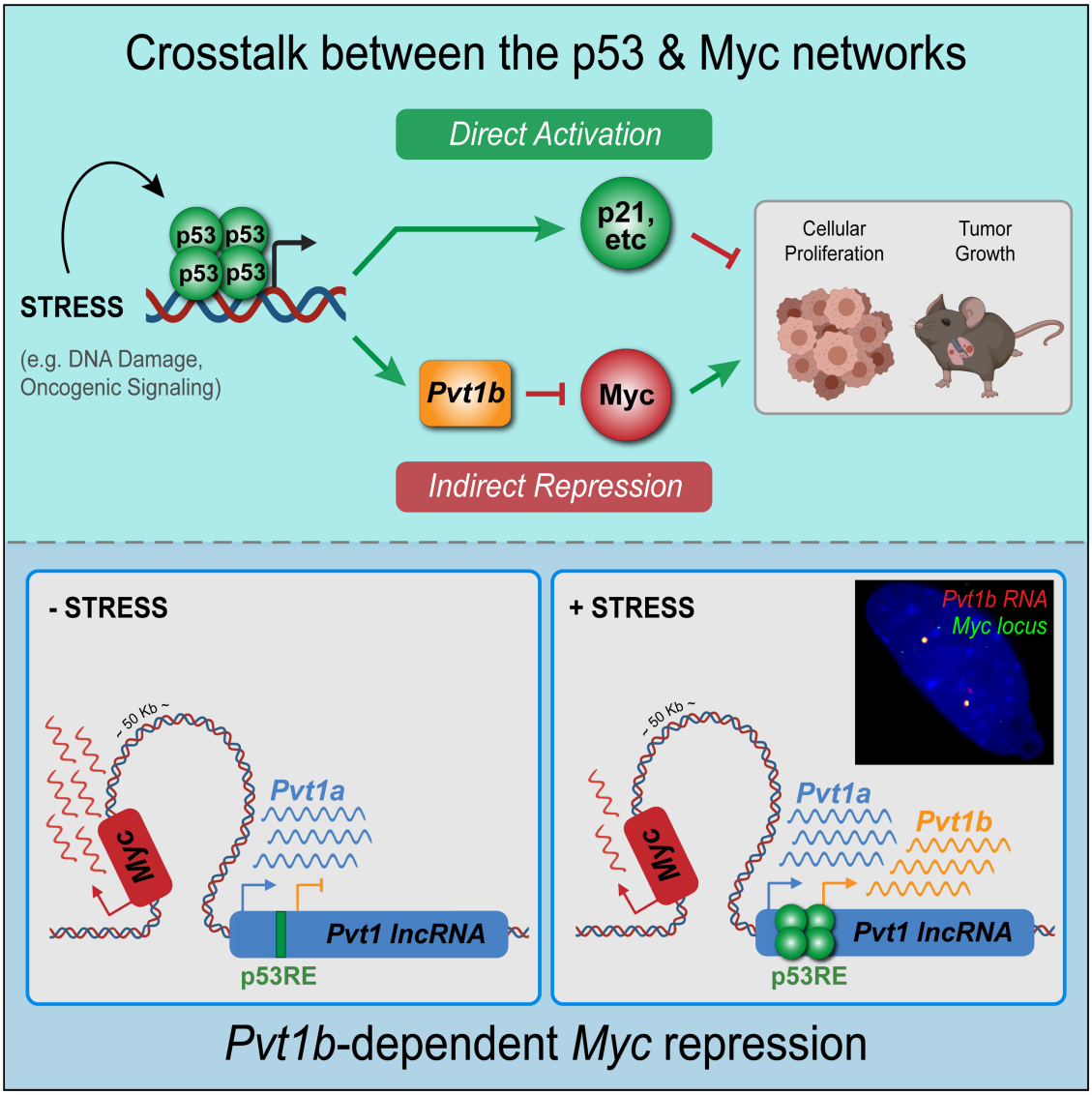

Figure 22. LncRNA *Pvt1b* mediates crosstalk between the Myc and p53 transcriptional networks. *Top* Under conditions of cellular stress, p53 (green) directly activates lncRNA *Pvt1b* (orange) to indirectly repress Myc (red), inhibiting cellular proliferation and tumor growth. *Bottom* p53-mediated induction of *Pvt1b* following cellular stress leads to local repression of *Myc* transcription, with smRNA-FISH inset showing *Pvt1b* accumulation at the *Myc-Pvt1* locus.

Chapter 3:

Investigating the functional elements of the Pvt1b transcript

Introduction

The elements of lncRNAs required to perform specific functions can comprise a range of characteristics encoded in the transcripts themselves, including sequence and/or structural motifs (see **Chapter 1**). The sequence of a lncRNA may enable its association with specific RNA binding proteins (RBPs) or may fold into unique hairpins or more complex structures that otherwise confer function (*reviewed in* (Zampetaki et al., 2018)). Given that *Pvt1b* differs from *Pvt1a* primarily based on whether transcript initiation is from exon 1b or exon 1a, respectively, I reasoned that the function of *Pvt1b* may depend explicitly on the exon 1b sequence or any structures therein. In this chapter, I investigate the importance of *Pvt1* exon 1b for p53-dependent *Myc* repression by attempting to either alter *Pvt1b*-specific sequences or disrupt their production. In doing so, I provide insight into the potential mechanism by which *Pvt1b* downregulates *Myc* under conditions of cellular stress.

Results

Investigating the function of the *Pvt1* exon 1b sequence in p53-dependent *Myc* repression

To determine if *Pvt1* exon 1b harbors any sequence and/or structural motifs required for *Pvt1b* function, I developed a genetic approach to test the functionality of regions within exon 1b whose disruption could rescue p53-induced *Myc* repression. To accomplish this, I employed a strategy to generate numerous CRISPR/Cas9-induced mutations throughout the 191 bp exon 1b sequence (Figure 23A). After analyzing available PAM sites, I selected 11 high specificity guide RNAs, excluding those with low specificity scores (Concordet and Haeussler, 2018) and those targeting similar areas of the exon 1b sequence (Figure 23B). I generated 11 mutant (gALT1-11) *KPR* population and *PR MEF* population cell lines (refer to **Chapter 2**, Figure 8 for descriptions of *KPR* and *PR MEF* cell lines), each of which contain heterogeneous mutations in the guide RNA-specific region targeted by Cas9 (Figure 23B). A guide RNA targeting dTomato (gTOM) was used as a negative control, while the guide RNA targeting the *Pvt1b*-associated p53RE (gRE, see **Chapter 2**, Figure 16) was used as a positive control due to its previously documented ability to rescue stress-induced *Myc* downregulation (see **Chapter 2**, Figure 16). I confirmed mutagenesis by each guide RNA (gALT1-11) by Sanger sequencing and Tracking of Indels by DEcomposition (TIDE) analysis (Brinkman et al., 2014), observing estimated mutagenesis efficiencies ranging from 37.2 – 94.4% (Figure 23C). Importantly, by qRT-PCR, *Pvt1b* levels were significantly suppressed in gRE cells compared to controls (Figures 23C and 23E), consistent with previous results (see **Chapter 2**, Figure 16). These observations led us to conclude that I had successfully mutagenized sequences in *Pvt1* exon 1b.

Infection with gALT1-11 resulted in fluctuations in *Pvt1b* expression levels in both *KPR* and *PR MEF* cell lines, variations which appeared guide RNA-

independent, with some guide RNAs yielding differing *Pvt1b* levels depending on cell line (Figures 23C and 23E). Considering that the qRT-PCR *Pvt1b* forward primer overlaps the exon 1b regions targeted by gALT8 and gALT9, I anticipated *Pvt1b* levels in these cell lines might be unusually low, although any underestimation of *Pvt1b* expression appeared to be minimal (Figures 23C and 23E). Importantly, any reductions in *Pvt1b* expression resulting from infection with gALT1-11 did not approach the extent of *Pvt1b* inhibition observed following infection with gRE (Figures 23C and 23E).

Finally, I used qRT-PCR to assess whether mutations within the *Pvt1* exon 1b sequence had any effect on *Myc* levels during the cellular response to stress. In negative control *KPR* population and *PR MEF* population cell lines infected with gTOM, exposure to oncogenic or genotoxic stress led to an expected decrease in *Myc* RNA (Figures 23D and 23F). Importantly, positive control *KPR* population and *PR MEF* population cell lines infected with gRE experienced a rescue in stress-dependent *Myc* downregulation (Figures 23D and 23F), consistent with previous findings (see **Chapter 2**, Figure 16). Next, I found that exposure to stress in gALT1-11 *KPR* population and *PR MEF* population cell lines did not result in any notable increases in *Myc* levels approaching the rescue observed in gRE cell lines. Taken together, these results suggested mutagenesis of sequences throughout *Pvt1* exon 1b was insufficient to rescue p53-dependent *Myc* repression.

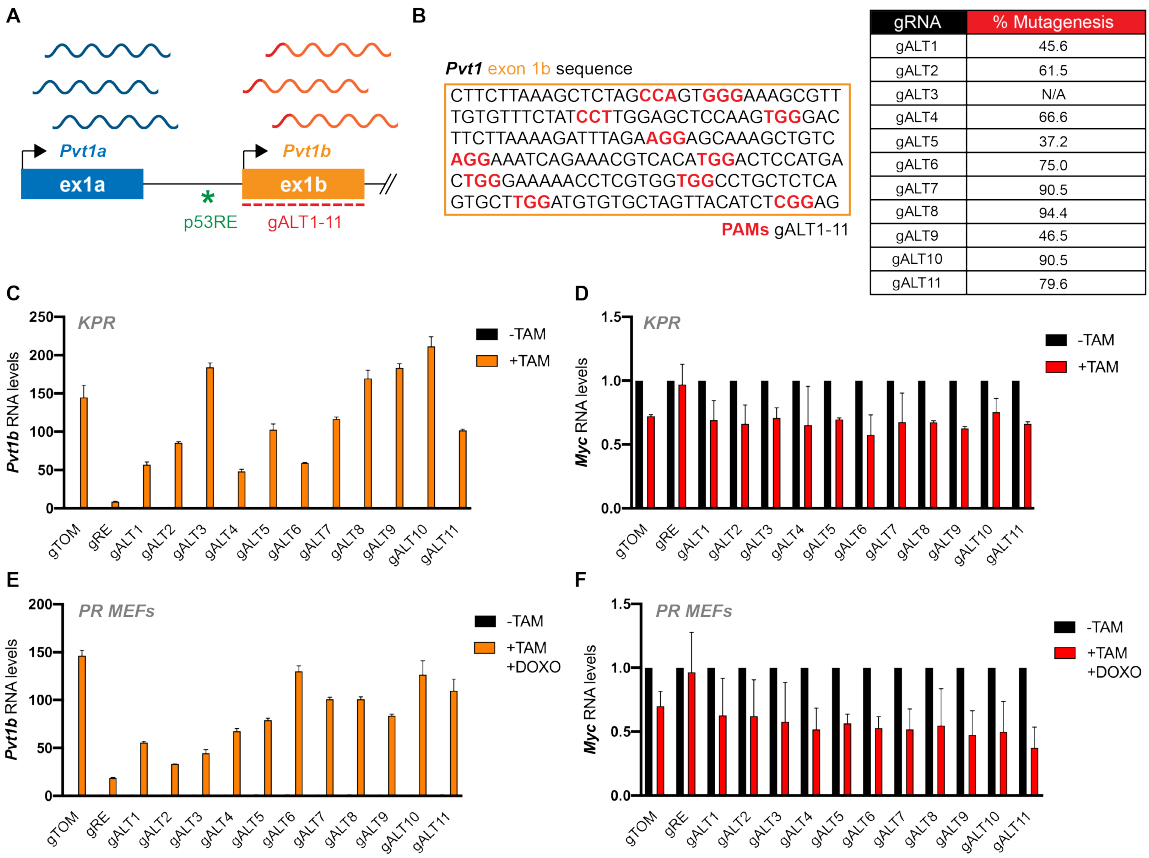

Figure 23. Probing the role of the *Pvt1* ex1b sequence in p53-dependent *Myc* repression. (A) Schematic of the 5' end of the *Pvt1* locus showing transcript initiation sites for *Pvt1a* (blue) and *Pvt1b* (orange) at exon 1a and exon 1b, respectively. Guide RNAs to mutate the exon 1b sequence (gALT1-11) shown in red, with red 5' ends of *Pvt1b* transcripts indicating mutagenesis. Location of *Pvt1b*-associated p53RE indicated by green asterisk. **(B)** *Left, Pvt1* exon 1b sequence, with PAM sites (red) utilized by guide RNAs gALT1-11 indicated sequentially. *Right,* percent mutagenesis efficiency yielded by each guide RNA (gRNA) as estimated by Tracking of Indels by DEcomposition (TIDE) analysis in *KPR* cells as described in (Brinkman et al., 2014). N/A = data not available. **(C)** *Pvt1b* RNA levels in *KPR* cells infected with indicated guide RNAs, untreated or treated with tamoxifen (TAM) for 24 h. Data show mean ± SEM (n=3, technical replicates, confirmed in independent biological replicates). **(D)** *Myc* RNA levels in same cells as (C). Data are normalized to untreated within each cell line and show mean ± SEM (n=2, biological replicates). **(E)** *Pvt1b* RNA levels in *PR MEFs* infected with indicated guide RNAs, untreated or treated with tamoxifen (TAM) for 48 h and doxorubicin (DOXO) for 24 h. Data show mean ± SEM (n=3, technical replicates, confirmed in independent biological replicates). **(F)** *Myc* RNA levels in the same cells as (E). Data are normalized to untreated within each cell line and show mean ± SEM (n=2, biological replicates).

Investigating the function of spliced *Pvt1b* in p53-dependent *Myc* repression

To determine if the spliced *Pvt1b* transcript is required for p53-dependent *Myc* repression, I developed a genetic approach to abrogate splicing between *Pvt1* exon 1b and its downstream exon 2 and thereby decrease processing of nascent *Pvt1b* transcripts. I designed a guide RNA (gdeltaSS) to target Cas9 close to the 3' end of *Pvt1* exon 1b (Figure 24A) with the goal of mutagenizing the AG|GU sequence spanning the exon|intron junction, which comprises a key sequence element of the splice donor site (Mount, 1982). I generated *KPR* population and *PR MEF* population cell lines (refer to **Chapter 2**, Figure 8 for descriptions of *KPR* and *PR MEF* cell lines) containing numerous mutations at the 3' end of *Pvt1* exon 1b. As above, a guide RNA targeting dTomato (gTOM) was used as a negative control, while the guide RNA targeting the *Pvt1b*-associated p53RE (gRE, see **Chapter 2**, Figure 16) was used as a positive control due to its previously documented ability to rescue stress-induced *Myc* downregulation. I confirmed mutagenesis of the exon|intron junction by Sanger sequencing and Tracking of Indels by DEcomposition (TIDE) analysis (Brinkman et al., 2014) in *KPR* cells and *PR MEFs*, observing 92.8% and 90.5% estimated efficiencies, respectively (Figure 24B and Figure 24C). However, I observed low frequencies of deletions larger than 5 nucleotides (Figure 24B and Figure 24C), and given the location of the Cas9 cut site 4-5 nucleotides away from the exon|intron junction, it is possible that a low percentage of mutations directly affected the AG|GU splice site.

To determine the efficacy of this approach in inhibiting the production of spliced *Pvt1b*, I analyzed the expression of *Pvt1* transcripts in gdeltaSS-infected

population cell lines. While exon 1b 3' end mutagenesis did not affect *Pvt1a* expression, it resulted in a significant 34% decrease in spliced *Pvt1b* in tamoxifen-treated *KPR* cells relative to control (Figure 24D) and a similar, albeit nonsignificant, decrease in spliced *Pvt1b* in *PR MEFs* treated with tamoxifen and doxorubicin (Figure 24F). Importantly, infection with gRE inhibited *Pvt1b* expression in both *KPR* cells (Figure 24D) and *PR MEFs* (Figure 24F), consistent with previous results (see **Chapter 2**, Figure 16). To determine if the observed reduction in spliced *Pvt1b* was due to abrogated splicing and not simply decreased expression, I analyzed nascent *Pvt1b* RNA levels using a qRT-PCR primer set spanning the *Pvt1* exon 1b|intron junction. Concurrent with a decrease in spliced *Pvt1b*, gdeltaSS-infected *KPR* and *PR MEF* cells exhibited increases in nascent *Pvt1b* compared to controls, suggesting a defect in pre-mRNA processing (Figures 24D and 24F). Collectively, these data suggested partial inhibition of the splicing event linking *Pvt1* exon 1b and exon 2.

To assess whether the processing of nascent *Pvt1b* into spliced *Pvt1b*, and the resulting *Pvt1* ex 1b – exon 2 sequence is required for stress-dependent *Myc* repression, I analyzed *Myc* RNA levels in response to oncogenic or genotoxic stress. In negative control *KPR* population and *PR MEF* population cell lines infected with gTOM, exposure to oncogenic or genotoxic stress led to an expected decrease in *Myc* RNA (Figures 24E and 24G). Importantly, positive control *KPR* population and *PR MEF* population cell lines infected with gRE experienced a rescue in stress-dependent *Myc* downregulation (Figures 24E and 24G), consistent with previous findings (see **Chapter 2**, Figure 16). However, mutagenesis of the 3' end of exon 1b did not increase *Myc* levels following cellular stress to the extent

observed in gRE-infected cells (Figures 24E and 24G). Given that *Pvt1b*-associated p53RE mutagenesis resulted in near total loss of *Pvt1b* expression and yielded a ~30% rescue of *Myc* levels, I would expect to observe a fraction of that rescue in *Myc* expression with a loss of spliced *Pvt1b* reaching only as high as 34% in gdeltaSS cells (Figures 24D and 24F). As such, these data and the importance of spliced *Pvt1b* in p53-dependent *Myc* repression are largely inconclusive, but could be further resolved with the isolation of gdeltaSS clones with biallelic modifications.

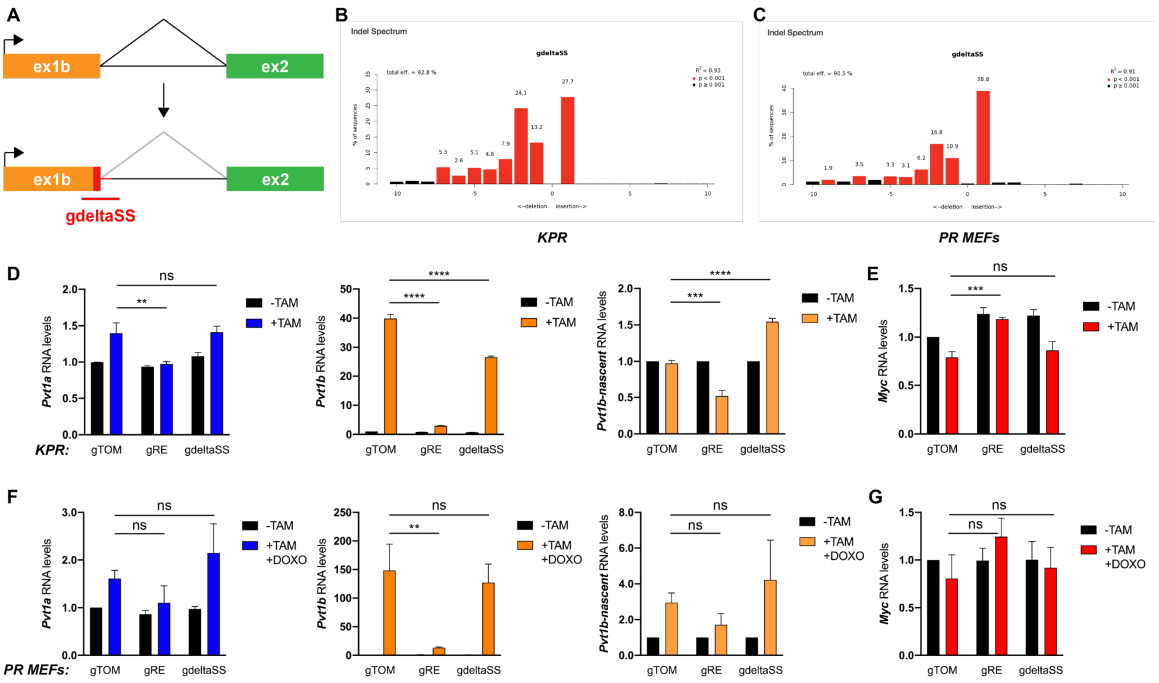

Figure 24. Probing the role of spliced *Pvt1b* in p53-dependent *Myc* repression. (A) Schematic of the 5' end of the *Pvt1b* transcript, indicating splicing between exon 1b and exon 2 and its disruption due to CRISPR-induced mutagenesis of the splice site at the 3' end of exon 1b. Region targeted by guide RNA (gdeltaSS) shown in red. **(B, C)** Mutagenesis efficiency of gdeltaSS estimated by Tracking of Indels by DEcomposition (TIDE) analysis in *KPR* cells (B) and *PR MEFs* (C) as described in (Brinkman et al., 2014). **(D)** *Pvt1a,* spliced *Pvt1b*, and nascent *Pvt1b* RNA levels in *KPR* cells infected with indicated guide RNAs, untreated or treated with tamoxifen (TAM) for 24 h. Data show mean ± SEM (n=3, biological replicates); ns = not significant; **p < 0.01, ***p < 0.001 and ****p < 0.0001, unpaired t test. **(E)** *Myc* RNA levels from cells in (D). Data show mean ± SEM (n=3, biological replicates); ns = not significant; ***p < 0.001, unpaired t test. **(F)** *Pvt1a,* spliced *Pvt1b*, and nascent *Pvt1b* RNA levels in *PR MEFs* infected with indicated guide RNAs, untreated or treated with tamoxifen (TAM) for 48 h and doxorubicin (DOXO) for 24 h. Data show mean ± SEM (n=3, biological replicates); ns = not significant; **p < 0.01, unpaired t test. **(G)** *Myc* RNA levels from cells in (F). Data show mean ± SEM (n=3, biological replicates); ns = not significant; unpaired t test.

Discussion

The research presented in this chapter attempts to address whether *Pvt1b* sequence and/or structural specificity is necessary for its stress-dependent function in repressing *Myc* expression. Neither mutagenesis of sequences throughout *Pvt1* exon 1b, nor abrogation of *Pvt1b* splicing was sufficient to rescue p53-induced *Myc* downregulation to the extent observed following *Pvt1b* inhibition via mutagenesis of the p53 binding site required for its expression. While the experiments outlined in this chapter were ultimately unsuccessful in identifying discrete elements of the *Pvt1b* transcript required for its function, these data do not conclusively disprove the existence of such sequence and/or structural motifs.

The analysis of cell populations with mutation heterogeneity, while useful in rapidly assessing the broad importance of numerous sequence elements at once, may not be sensitive enough to pinpoint essential sequences. There are several possible outcomes of such cell population-based mutagenesis experiments that may muddle our interpretation: (1) mutagenesis efficiency is not high enough or, depending on the types of indels produced by a particular guide RNA, yields mutations that preserve rather than disrupt critical motifs, (2) the guide RNAs chosen do not target close enough to key sequences to effectively alter them, and (3) any useful mutations that might inhibit RNA function in isolation are easily obscured by other, less impactful, mutations. In short, it is possible that these approaches are not precise enough to yield interpretable results. For example, the decrease in spliced *Pvt1b* observed in gdeltaSS *KPR* and *PR* cell lines is significantly less than the decrease in *Pvt1b* levels following p53RE mutagenesis

and may not have been sufficient to rescue *Myc* levels. Therefore, it is possible a more robust abrogation of splicing would nullify *Myc* repression and recapitulate the results observed with *Pvt1b* transcriptional inhibition. One method for overcoming some of these pitfalls involves the isolation of clonal cell lines and comprehensive analysis of individual mutations. This approach can be fruitful but may result in clone-specific behaviors unrelated to the mutation in question.

Choosing the appropriate cell model in which to perform such cell population-based mutagenesis experiments poses an additional challenge. *KPR* cells, while a robust model for p53-dependent *Myc* repression under conditions of oncogenic stress, harbor multiple copies of the *Myc-Pvt1* locus in the form of extrachromosomal DNA circles (see **Chapter 2**, Figure 13). This leads to a landscape of numerous distinct Cas9-induced *Pvt1* mutations existing in a single cell, the effects of which on the expression of individual *Myc* alleles may vary and produce an average that does not reflect complex heterogeneity between loci. However, my experiments in *PR MEFs*, which do not exhibit the same extensive *Myc-Pvt1* amplifications, support and provide an independent confirmation of our data in *KPR* cells. In summary, while this set of genetic queries did not successfully discover a sequence-based and/or structural mechanism for *Pvt1b* function, they do not preclude the existence of such a mechanism and alternative approaches are needed to better understand the functional elements of the *Pvt1b* transcript.

Chapter 4:

Investigating the mechanism of Pvt1b-mediated Myc repression

Introduction

Previous studies have investigated the role of histone deacetylases (HDACs) in p53-dependent *Myc* repression (Harms and Chen, 2007; Ho et al., 2005). Notably, Ho et al. observed decreased histone H4 acetylation marks at the *Myc* promoter following p53 activation and further posited a mechanism for p53-induced *Myc* downregulation involving mSin3a, a corepressor that associates with HDAC1 (Ho et al., 2005). However, the function of HDACs in regulating *Myc* expression is unclear, with different studies investigating the effect of HDAC inhibitors on *Myc* levels reporting conflicting results (Majumdar et al., 2012; Sasakawa et al., 2003; Xu et al., 2005; Yu et al., 2020).

The function of lncRNAs in gene regulation via epigenetic modification is well-documented (*reviewed in* (Statello et al., 2020)). A significant example of an HDAC-dependent mechanism of lncRNA-mediated gene regulation is the function of *XIST* in *cis*-repression of the X-chromosome via SHARP/HDAC3 (McHugh et al., 2015). One of a family of transcriptional repressors, SHARP (also known as Spen) interacts with SMRT (Ariyoshi and Schwabe, 2003), a component of the nuclear corepressor complex with a known role in HDAC3-mediated chromatin deacetylation (You et al., 2013). A direct interaction between *XIST* and SHARP is required to recruit SMRT and HDAC3, enabling transcriptional silencing of the X-

chromosome via histone deacetylation and RNA Pol II exclusion (Chu et al., 2015; McHugh et al., 2015; Moindrot et al., 2015; Monfort et al., 2015). Recent evidence from the Guttman (unpublished) and Chang (Carter et al., 2020) labs notes an association between SHARP/Spen and exon 2 of human and mouse *Pvt1*. In preliminary RAP-MS experiments performed by Giuseppe Militello in our lab, we also detected evidence for SHARP/Spen binding the *Pvt1* RNA in *KPR* cells (data not shown). This may suggest a mechanism of *cis*-repression by *Pvt1b* similar to that employed by *XIST* during X-chromosome inactivation (XCI), involving histone deacetylation via HDAC3 recruited by SHARP/Spen (McHugh et al., 2015). In this chapter, I present preliminary evidence supporting a role for *Pvt1b* in stress-dependent *Myc* repression via histone deacetylation at the *Myc* promoter.

Results

Histone deacetylation may be required for stress-dependent *Myc* repression

To confirm previous findings (Ho et al., 2005) and determine whether histone acetylation marks at the *Myc* promoter change in response to cellular stress, I performed H3K27ac ChIP in *KPR* cells following treatment with or without tamoxifen. Concurrent with a decrease in *Myc* expression, I observed commensurate decreases in H3K27ac marks upstream of the *Myc* transcriptional start site following exposure to oncogenic stress (Figure 25A). To gain insight into whether *Pvt1b* is required for histone deacetylation at the *Myc* promoter following p53 activation, I performed H3K27ac ChIP in *KPR* cells infected with gRE to mutagenize the *Pvt1b*-associated p53RE and inhibit *Pvt1b* expression (see

Chapter 2). In contrast to the results observed in wildtype cells (Figure 25A), H3K27ac marks did not decrease in response to oncogenic stress in cells in which *Pvt1b* expression was inhibited (Figure 25B). These results, while preliminary and in need of repetition, suggest a potential HDAC-based mechanism underpinning *Pvt1b* function.

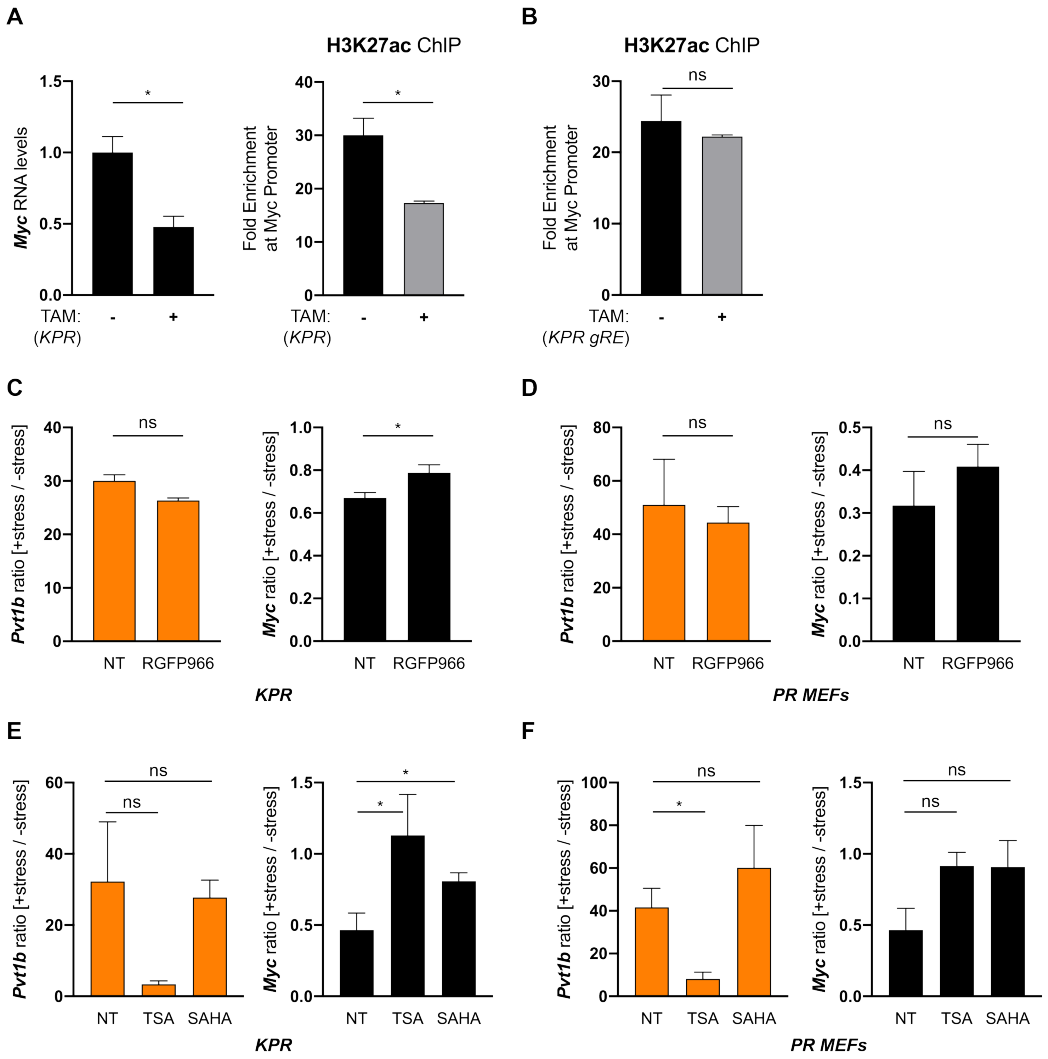

Figure 25. Histone deacetylation may be required for stress-dependent *Myc* repression. (A) *Left, Myc* RNA levels in *KPR* cells, untreated or treated with tamoxifen (TAM) for 24 h. *Right,* ChIP-qPCR analysis of H3K27ac enrichment at the *Myc* promoter in the same cells. Data show mean ± SEM (n=3, technical replicates of one biological replicate); ns = not significant; *p < 0.05, paired t test. **(B)** ChIP-qPCR analysis of H3K27ac enrichment at the *Myc* promoter in *KPR* cells infected with a guide RNA targeting the *Pvt1b*-associated p53RE (gRE), untreated or treated with tamoxifen (TAM) for 24 h. Data show mean ± SEM (n=3, technical replicates of one biological replicate); ns = not significant; paired t test. **(C,D)** qRT-PCR of the ratio (+stress/-stress) of *Pvt1b* and *Myc* expression levels in *KPR* cells (C) and *PR MEFs* (D) treated with or without the HDAC3 inhibitor RGFP966 for 6 h (C) or 4 h (D). *KPR* cells were treated with 24 h tamoxifen (C) and *PR MEFs* were treated with 48 h tamoxifen and 24 h doxorubicin (D) to induce cellular stress. Data in (C) show mean ± SEM (n=3, technical replicates of one biological replicate); *p < 0.05, paired t test. Data in (D) show mean ± SEM (n=3, biological replicates); ns = not significant; paired t test. NT = no treatment. **(E,F)** qRT-PCR of the ratio (+stress/-stress) of *Pvt1b* and *Myc* expression levels in *KPR* cells (E) and *PR MEFs* (F) treated with or without the pan-HDAC inhibitors TSA or SAHA for 4 h. Treatment with cellular stress and fold induction of *Pvt1b* performed as in (C,D). Data show mean ± SEM (n=3, biological replicates); ns = not significant; *p < 0.05, paired t test. NT = no treatment. Data collection and analysis performed by Giuseppe Militello.

Contributions from G. Militello in (E) and (F) are described above.

Histone deacetylase inhibition rescues stress-dependent *Myc* repression

To query whether HDACs play a role in p53-dependent *Myc* repression more directly, I sought to determine whether the inhibition of one or more HDACs could rescue stress-induced *Myc* downregulation. Given the importance of HDAC3 in *XIST*-mediated XCI, we treated *KPR* cells and *PR MEFs* with the HDAC3 inhibitor RGFP966 in the presence and absence of stress. HDACs, including HDAC3, have a well-documented role in repressing the p53 pathway (Ito et al., 2002; Monte et al., 2006; Narita et al., 2010), and treatment with RGFP966 caused a minor decrease in *Pvt1b* compared to untreated controls (Figures 25C and 25D). While *Myc* levels decreased in response to both oncogenic and genotoxic stress in control samples, we observed a partial, albeit nonsignificant, rescue of *Myc* levels with HDAC3 inhibition (Figures 25C and 25D). These results indicated that HDAC3 may play a role in p53-induced *Myc* repression, but do not completely explain the observed decreases in *Myc* expression.

To determine whether additional HDACs might be important for *Myc* downregulation under conditions of cellular stress, Giuseppe Militello performed experiments utilizing pan-HDAC inhibitors trichostatin A (TSA) and suberoylanilide hydroxamic acid (SAHA), which inhibit class I and class II HDACs, including HDAC3 (Xu et al., 2007). While SAHA did not have significant effects on *Pvt1b* expression relative to controls, TSA treatment caused robust downregulation of *Pvt1b*, perhaps indicating systemic effects on the p53 network (Figures 25E and 25F). While *Myc* RNA levels displayed an expected decrease in the presence of stress, he observed a significant increase in *Myc* RNA in *KPR* cells treated with

either TSA or SAHA, amounting to a near-full rescue of *Myc* expression (Figure 25E). Similarly, he observed robust, albeit nonsignificant, increases in *Myc* expression in stressed *PR MEFs* following treatment with pan-HDAC inhibitors (Figure 25F). Taken together, these data point to an HDAC-dependent mechanism for *Pvt1b*-mediated *Myc* repression.

Discussion

The results presented in this chapter comprise a broad investigation of the hypothesis that *Pvt1b* represses *Myc* transcription by facilitating histone deacetylation at the *Myc* promoter in response to cellular stress. These preliminary data thus far point to a mechanism of p53-dependent *Myc* repression via histone deacetylation, which may require *Pvt1b* expression. That H3K27ac marks did not decrease at the *Myc* promoter in cells in which *Pvt1b* expression was inhibited, in contrast to wildtype cells, suggests the specific importance of *Pvt1b* in histone deacetylation. Significantly, preliminary findings from an epistasis experiment involving HDAC inhibition in cells lacking *Pvt1b* implicate HDACs as functioning in the same pathway as *Pvt1b* (data not shown), providing further support for *Pvt1b* modulating *Myc* expression via HDACs. However, additional mechanistic studies are needed to establish the link between *Pvt1b*, HDACs, and *Myc* repression more definitively. While HDAC3 inhibition prompted a partial rescue of *Myc* downregulation under conditions of oncogenic or genotoxic stress, the more prominent *Myc* rescue observed following pan-HDAC inhibition suggests the potential involvement of additional HDACs. In the future, it will be necessary to

determine which class I/II HDACs inhibited by TSA and SAHA are responsible for the stress-dependent decrease in *Myc* expression.

While my results certainly point to a role for histone deacetylation in stress-specific *Myc* regulation, any potential involvement of SHARP/Spen in keeping with the mechanism of *XIST*-mediated XCI still needs to be investigated (Chu et al., 2015; McHugh et al., 2015; Moindrot et al., 2015; Monfort et al., 2015). I have designed a strategy for generating a CRISPR-mediated Spen knockout model, which will be useful for assessing any requirement for SHARP/Spen in p53-dependent *Myc* regulation. Thoroughly validating any putative interaction between *Pvt1b* and Spen (Carter et al., 2020) via RNA immunoprecipitation (RIP) or crosslinking and immunoprecipitation (CLIP) studies will also be important for elucidating the mechanism of *Pvt1b*-mediated *Myc* repression. LncRNAs often engage epigenetic regulators to elicit repression (*reviewed in* (Statello et al., 2020)), and the data presented here currently support this paradigm. Future work in our lab will focus on both further defining the functional elements of the *Pvt1b* isoform and elucidating the role of HDACs in the p53-*Pvt1b*-*Myc* regulatory axis.

Chapter 5:

Generation of Pvt1 genetically engineered mouse models

Introduction

Historically, the study of lncRNAs in cancer has benefitted from the use of GEMMs to probe lncRNA function *in vivo* (see **Chapter 1**). Perturbation of lncRNA function via genetic or epigenetic modifications to the endogenous locus can help overcome many issues with the study of lncRNAs in *in vitro* model systems (see **Chapter 1**) and can solidify our understanding of lncRNA function at the organismal level. Considering the power of *in vivo* models in resolving lncRNA function(s), I sought to deploy a suite of molecular tools that would enable elucidation of the function(s) of *Pvt1* isoforms under both physiologic and tumorigenic conditions.

First, I took advantage of a well-characterized synthetic polyadenylation signal (PAS) to elicit premature transcription termination (Levitt et al., 1989), a robust genetic tool that has been used previously to clarify mechanisms of local gene regulation enacted by lncRNAs (Engreitz et al., 2016). The short 49 bp sequence, when transcribed as part of the *Pvt1* locus, should cause efficient 3' cleavage and polyadenylation of nascent *Pvt1* transcripts (Figure 26A) (Levitt et al., 1989). In addition to stimulating 3' end-processing machinery via transcription of a A(A/U)UAAA hexamer followed by a GU-rich tract 30 bp downstream (Millevoi and Vagner, 2010), PAS insertion may also suppress *Pvt1* transcription

initiation due to the tight connection between the splicing and transcriptional processes (Engreitz et al., 2016).

I chose to insert the synthetic PAS within *Pvt1* exon 1b, just downstream of the *Pvt1b* transcription start site (TSS), in order to abrogate production of the full length *Pvt1b* transcript (*Pvt1b-PAS*, Figures 26A and 26C). Due to the efficiency with which polyadenylation occurs and the location of *Pvt1* exon 1b downstream from exon 1a, I expected that a PAS insertion within exon 1b would also abrogate the production of nascent *Pvt1a* transcripts, effectively acting as a LOF model for both *Pvt1a* and *Pvt1b*. Therefore, to distinguish between any potentially divergent functions ascribed to *Pvt1a* and *Pvt1b* isoforms, I designed a separate LOF model specific to *Pvt1a* (*Pvt1a-PAS*) by inserting a PAS into *Pvt1* exon 1a, just downstream of the *Pvt1a* TSS (Figures 26A and 26C). This model should result in specific abrogation of *Pvt1a* transcription, without negatively affecting *Pvt1b*, thus providing both a control for any experiments using the *Pvt1b-PAS* model and an important tool for elucidating *Pvt1a* function independent of *Pvt1b*.

To complement the *Pvt1a-PAS* and *Pvt1b-PAS* GEMMs, I designed an additional *Pvt1* GEMM using a novel ribozyme-based tool developed in our lab to elicit co-transcriptional cleavage and subsequent transcript degradation (Winkler et al., *in preparation*). This strategy relies on the 74 bp-length self-cleaving ribozyme *Twister*, initially identified and characterized by the Breaker lab at Yale (Figure 26B) (Roth et al., 2014). In *in vitro* studies, *Twister* undergoes efficient self-cleavage under simulated physiologic conditions, approaching rates as high as ~1000 min^{-1} (Roth et al., 2014). *Twister*'s small size and ability to rapidly self-cleave via site-specific phosphodiester scission (Jimenez et al., 2015) make it an

ideal candidate for a lncRNA LOF tool. Previous studies have demonstrated success using ribozymes to inhibit ncRNA accumulation in yeast and mouse cells (Camblong et al., 2009; Tuck and Buhler, 2021; Tuck et al., 2018) and unpublished results from our lab indicate that *Twister* robustly destabilizes transcript aggregation *in vivo* when inserted into the *LincRNA-p21* locus (Winkler et al., *in preparation*).

I chose to insert the *Twister* sequence into *Pvt1* exon 1b in order to abrogate accumulation of both *Pvt1a* and *Pvt1b* transcripts, similar to the *Pvt1b-PAS* model (*Pvt1b-TWI*, Figure 26C). Inserting *Twister* into the same location as the PAS in the *Pvt1b-PAS* model enables direct comparison between the effects of premature polyadenylation and *Twister*-mediated transcript cleavage on neighboring *Myc* expression. While PAS-mediated transcript cleavage and polyadenylation leaves a downstream nascent cleavage product with a terminal 5' phosphate group, a ready substrate for degradation by the 5' end surveillance factor XRN2 (West et al., 2004), the downstream fragment resulting from *Twister* self-cleavage instead possesses a 5' hydroxyl group (Roth et al., 2014), which is unlikely to engage XRN2-mediated transcriptional termination (Doamekpor et al., 2020; Jinek et al., 2011; Mathy et al., 2007; West et al., 2004). Indeed, unpublished results from our lab suggest that *Twister*-mediated RNA cleavage enables some level of transcription to proceed through the locus (Winkler et al., *in preparation*), in contrast to the rapid drop-off in transcription resulting from PAS insertion, usually within 2 Kb (Core et al., 2008). Considering that local gene regulation by lncRNAs may depend on the RNA transcript or the process of transcription through the locus alone (Engreitz et al., 2016), the *Pvt1b-TWI* model will provide further

insight into the elements of the *Pvt1* transcript required for its function(s). Whether the *Pvt1b-PAS* and *Pvt1b-TWI* models successfully abrogate both *Pvt1a* and *Pvt1b* levels will need to be empirically determined, as the increased distance between either the polyadenylation sequence or the *Twister* ribozyme and the *Pvt1a* TSS may result in reduced efficiency of these LOF tools (Engreitz et al., 2016; Tuck and Buhler, 2021; Tuck et al., 2018).

In this chapter, I describe the successful generation of the *Pvt1a-PAS*, *Pvt1b-PAS* and *Pvt1b-TWI* alleles in the mouse and provide evidence for their germline transmission. These GEMMs set the stage for further extensive characterization of the contributions of *Pvt1a* and *Pvt1b* isoforms to organismal development and normal homeostasis, the cellular response to stress, and tumorigenesis.

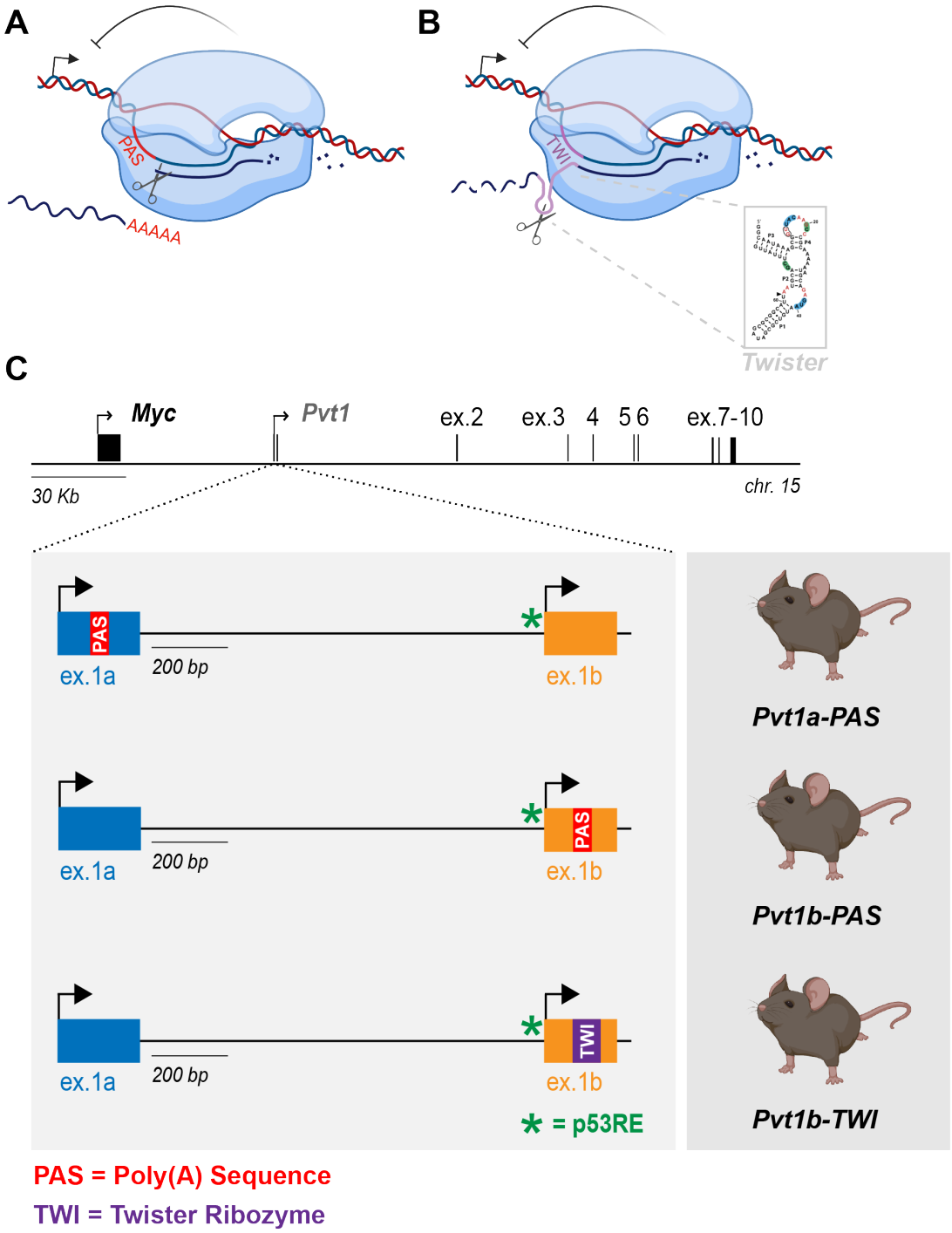

Figure 26. Schematic of *Pvt1* GEMMs. (A) Illustration of premature RNA cleavage and polyadenylation induced by insertion of a 49 bp synthetic polyadenylation signal (PAS) into an endogenous gene. Inhibitory line indicates transcriptional suppression. **(B)** Illustration of co-transcriptional RNA degradation induced by insertion of the 74 bp self-cleaving *Twister* ribozyme (TWI) into an endogenous gene. Inhibitory line indicates transcriptional suppression. The structure of the *Twister* ribozyme is shown (adapted from Roth et al. 2014). **(C)** *Top*, schematic of the *Myc-Pvt1* locus. *Bottom*, schematics of the *Pvt1a-PAS*, *Pvt1b-PAS* and *Pvt1b-TWI* alleles, as present in their associated genetically engineered mouse models (GEMMs).

Results

Design and generation of *in vivo Pvt1a-PAS*, *Pvt1b-PAS* and *Pvt1b-TWI* alleles

Pvt1a-PAS, *Pvt1b-PAS* and *Pvt1-TWI* mice were generated in collaboration with Adam Williams and Rick Maser at the Jackson Laboratory for Genomic Medicine (JAX). We chose candidate guide RNAs based on predicted specificity scores (Concordet and Haeussler, 2018) and the criteria that Cas9 cleavage and subsequent insertion of either PAS or TWI occur <100 bp from *Pvt1a* and *Pvt1b* TSSs to increase the likelihood of efficient premature transcript termination or co-transcriptional transcript degradation, respectively. The final two guide RNAs, one targeting *Pvt1* exon 1a (gEx1a) and one targeting *Pvt1* exon 1b (gEx1b), were selected based on cleavage efficiency estimates from *in vitro* analyses performed at JAX (Figure 27A; data not shown). Homology directed repair (HDR) templates were subsequently designed with either the 49 bp PAS or the 74 bp TWI sequence inserted at the expected Cas9 cleavage site with appropriate length homology arms on either side of the altered DNA. Single-stranded oligo donors (ssODNs) were constructed complementary to the non-target strand to increase repair efficiency (Richardson et al., 2016).

Briefly, mouse embryos were electroporated with guide RNA:Cas9 ribonucleoprotein (RNP) complexes and the appropriate homology directed repair (HDR) templates to generate *Pvt1a-PAS*, *Pvt1b-PAS* and *Pvt1b-TWI* alleles. These embryos were surgically implanted into pseudo-pregnant mice and we received the resulting litters from JAX including at least five potential founder mice for each

allele with successful knock-ins based on initial genotyping performed by JAX (data not shown).

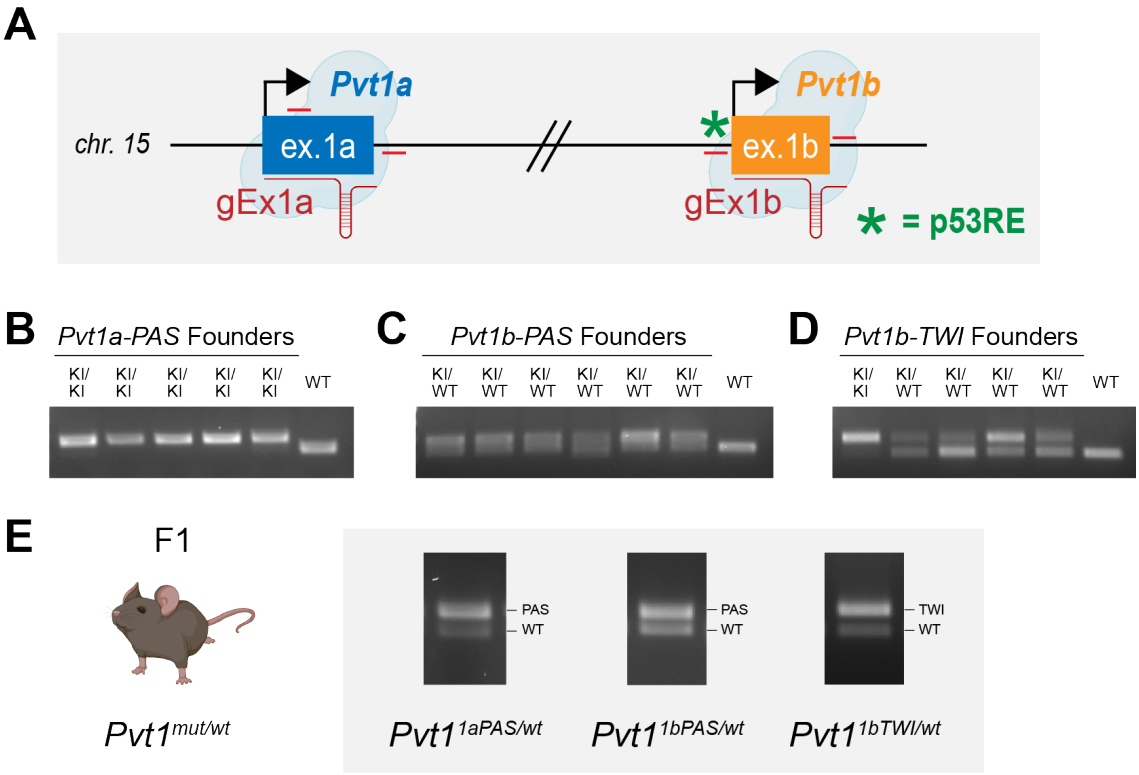

Figure 27. *Pvt1-PAS*, *Pvt1b-PAS*, and *Pvt1b-TWI* genotyping. (A) Schematic of the mouse *Pvt1* locus highlighting *in vivo* CRISPR/Cas9 editing strategy, including guide RNAs used to edit *Pvt1* exon 1a (gEx1a) and exon 1b (gEx1B), respectively. Red lines indicate location of genotyping primers. **(B-D)** Gel images of genotyping results from potential (B) *Pvt1a-PAS*, (C) *Pvt1b-PAS*, and (D) *Pvt1b-TWI* founder mice. Allele combinations of each mouse indicated. PAS = Polyadenylation sequence; TWI = *Twister* ribozyme; KI = Knock-in; WT = Wildtype. Genotyping performed by Nadya Dimitrova. **(E)** Gel images of genotyping results from F1 mice, generated by backcrossing *Pvt1* mutant (mut) founder mice to wildtype (wt) mice to produce heterozygotes. Genotyping results from F1 heterozygous mice with germline transmission of *Pvt1a-PAS*, *Pvt1b-PAS*, and *Pvt1b-TWI* alleles are shown. PAS = Polyadenylation sequence; TWI = *Twister* ribozyme; WT = Wildtype. Matings and genotyping performed by Nadya Dimitrova.

Contributions from N. Dimitrova in (B-E) are described above.

Germline transmission of *Pvt1a-PAS*, *Pvt1b-PAS*, and *Pvt1b-TWI* alleles

As founder mice are subject to mosaicism, it is necessary to deconvolve the altered *Pvt1* alleles and establish germline transmission of successful knock-ins. First, the presence of *Pvt1a-PAS*, *Pvt1b-PAS* and *Pvt1b-TWI* alleles in potential founder mice was confirmed by genotyping of DNA extracted from mouse tail clippings (Figures 27A-D). In multiple independent founders, increases in band size compared to wildtype mice were observed, indicating varying degrees of successful PAS and TWI knock-ins (Figures 27B-D). All five potential *Pvt1a-PAS* founders appeared homozygous for PAS knock-in, six potential *Pvt1b-PAS* founders appeared heterozygous for PAS knock-in, and analysis of potential *Pvt1b-TWI* founders revealed one homozygous and four heterozygous for TWI knock-in (Figures 27B-D).

To establish germline transmission, founder mice were crossed to wildtype C57BL/6J mice. Heterozygous mice constitute generation F1 (filial 1) and are poised to be utilized in future experiments, as they, unlike the founder mice, do not suffer from the challenges associated with mosaicism. Crossing two F1 heterozygous mice together will produce a litter of wildtype mice and mice heterozygous or homozygous for the altered allele, enabling further analyses. Therefore, the establishment of germline transmission for all *Pvt1a-PAS*, *Pvt1b-PAS*, and *Pvt1b-TWI* alleles and production of F1 mice heterozygous for these altered alleles paves the way for downstream analyses (Figure 27E).

Discussion

In summary, this chapter describes the design and generation of the *Pvt1a-PAS*, *Pvt1b-PAS* and *Pvt1b-TWI* mouse models, the combination of which will allow us to better understand and define the functional elements of the *Pvt1* locus and its associated isoforms. Future experiments will focus on characterizing the effects of these mutant alleles on both *Pvt1* transcription and transcript stability, and on neighboring *Myc* expression. Careful analysis of the efficacy of the PAS and TWI alleles *in vivo* and how they influence local regulation of *Myc* will allow us to more precisely describe how *Pvt1b* represses *Myc* transcription, and whether this requires *Pvt1b* RNA production alone, or additionally requires the *Pvt1b* RNA transcript itself. In addition, the inclusion of a *Pvt1a*-specific mouse model, will provide insight into the potentially divergent functions of the *Pvt1a* and *Pvt1b* isoforms and whether pro- and anti-tumorigenic functions coexist in the same locus.

Importantly, the insertion of either the synthetic polyadenylation signal or the *Twister* ribozyme constitutes the addition of a short (<100 bp) sequence into the endogenous *Pvt1* locus without requiring the deletion of any DNA elements. Previous work has demonstrated the role of DNA elements encoded in the *Pvt1* locus in positively and negatively regulating *Myc* expression (Cho et al., 2018; Fulco et al., 2016), suggesting the importance of minimal disruption to the endogenous sequence in *Pvt1* GEMM design considerations. Deletion of over 300 Kb of the *Pvt1* locus *in vivo* resulted in reduced stability of the Myc protein (Tseng et al., 2014), but raised questions about the relative functional contributions of the *Pvt1* RNA itself compared to DNA elements. In comparison, the mouse models

described above do not delete, and should largely avoid the disruption of, DNA elements, providing an opportunity to assess the role of the *Pvt1* RNA and its production more directly.

In addition to helping elucidate molecular mechanism, these three GEMMs will improve our understanding of the importance of *Pvt1* at the organismal level. For example, analysis of developmental timepoints and aging studies will shed light on any physiologic effects observed as a result of *Pvt1* loss, and may suggest new roles in organism growth and homeostasis. While *Pvt1* loss alone may not be sufficient to predispose animals to spontaneous tumor development, crossing the *Pvt1a-PAS*, *Pvt1b-PAS* and *Pvt1b-TWI* models with established cancer models will enable investigation of the importance of *Pvt1* in promoting or suppressing various cancer types. There is much that we still do not understand about the role of *Pvt1b* as a tumor suppressor, For example, previous work has demonstrated that *Pvt1b* loss increases oncogenic *K-ras*-driven tumor growth (see **Chapter 2**, Figure 19). However, it is unknown whether *Pvt1b* loss in combination with loss of another tumor suppressor gene may constitute a second "hit" in the "two-hit hypothesis" framework of cancer development (Knudson, 2001) and be sufficient to drive tumorigenesis. Taken together, these novel GEMMs provide an unprecedented opportunity to further our knowledge of *Pvt1*'s roles in development and disease and inform our understanding of lncRNA function at the organismal level.

Chapter 6: Summary and Perspectives

Here I have identified and characterized *Pvt1b*, revealing a *bona fide* tumor suppressor function embedded in the *Pvt1* locus and mediated by a stress-specific lncRNA isoform. This work substantially contributes to both the fields of cancer biology and lncRNA biology by **(1)** improving our understanding of the regulatory feedback between the p53 tumor suppressor and Myc proto-oncogenic networks and its physiological impacts on cancer initiation and development, and **(2)** uncovering a novel lncRNA isoform-switching mechanism that enables swift locus-specific reverse regulation downstream of a transcriptional activator. In employing diverse LOF and GOF tools in powerful *in vitro* and *in vivo* model systems, I provided critical insight into how, and the extent to which, *Pvt1b* contributes to tumor suppression and growth inhibition downstream of p53 activation.

I found that *Pvt1b* is rapidly induced by genotoxic and oncogenic stress in a p53-dependent manner. The production of the *Pvt1b* RNA, 50 Kb downstream of the *Myc* locus, is necessary and sufficient to repress *Myc* transcription *in cis*, with negative functional consequences on Myc transcriptional activity and cellular proliferation. Importantly, using a genetic LOF model specific to p53 function, I demonstrated that *Pvt1b* inhibition increases *Myc* expression, cellular proliferation, and tumorigenesis under conditions of cellular stress. These observations complicate the long-standing classification of *Pvt1* as an oncogenic lncRNA, which stems from decades of evidence linking its elevated expression to

the increased proliferative and invasive capacities of cancer cells (Cui et al., 2016) as well as increased tumor aggressiveness and poor patient survival (Zhu et al., 2017). In stark contrast to this body of literature, I have unearthed a novel tumor suppressor function for this lncRNA gene, previously hidden in the vast expanse of the complex pro-oncogenic *Pvt1* locus. My work does not contradict the wealth of evidence implicating *Pvt1* as an oncogenic lncRNA, which dates back to its discovery nearly 40 years ago (Cory et al., 1985; Graham et al., 1985). Instead, we can now propose a more nuanced and holistic model for *Pvt1* function that incorporates both tumor suppressive and oncogenic activities as mediated by different elements of the *Pvt1* locus and its associated RNAs.

The study of cancer-associated genetic and epigenetic alterations and their frequencies has implicated both coding and noncoding genes as potential drivers of tumorigenesis (see **Chapter 1**). As such, the identification of recurrent genetic rearrangements involving the lncRNA *Pvt1* locus provided strong impetus for the study of *Pvt1* in the context of cancer (Cory et al., 1985; Graham and Adams, 1986; Graham et al., 1985; Shtivelman et al., 1989). In keeping with co-amplifications of the *Myc-Pvt1* locus observed in tumors (Riquelme et al., 2014), co-gain of *Myc* and *Pvt1* was found to advance cancer progression in a mouse model of breast cancer (Tseng et al., 2014). Importantly, gain of either *Myc* or *Pvt1* alone had more limited effects on tumor growth, suggesting a pro-oncogenic synergy between these two loci, perhaps deriving from a role for the *Pvt1* RNA in promoting Myc protein stability (Tseng et al., 2014). Several additional molecular mechanisms have been proposed to describe the tumor-promoting function of *Pvt1*, including its function as a miRNA sponge (Panda et al., 2017; Zhao et al., 2018), its role in epigenetically

repressing tumor suppressor genes via association with chromatin modifying complexes (Kong et al., 2015), and its ability to stabilize oncoproteins (Xu et al., 2017). Distinct from these proposed oncogenic functions, my work and the work of others demonstrates a function for the *Pvt1* locus in restricting *Myc* expression to limit tumorigenesis, either through DNA elements (Cho et al., 2018; Fulco et al., 2016) or through stress-induced production of the *Pvt1b* RNA, as shown here. The discovery of both activating and repressive DNA elements in the *Pvt1* locus that regulate *Myc* expression calls into question previous assumptions about *Pvt1* function (Fulco et al., 2016), suggesting a more nuanced picture of *Pvt1* regulatory roles.

The tumor suppressive function of *Pvt1* may have operated under the radar of cancer biologists due to the frequent inactivation of the p53 network in tumor development. In such cases where *Pvt1b* is inhibited due to genetic alterations in the *Pvt1* locus or the removal of upstream activating signals, the oncogenic activities of *Pvt1* would be expected to predominate. This may explain the acquisition of extrachromosomal DNA (ecDNA) circles harboring *Myc-Pvt1* co-amplifications in p53-deficient *KPR* cells, as well as the recent observation of a *Myc-Pvt1* ecDNA genomic rearrangement in a human colon cancer cell line, which fused exon 1 of *Pvt1* to exons 2 and 3 of *Myc*, disrupting normal *cis*-regulatory circuits in the locus (Hung et al., 2020). Interestingly, genomic alterations in the region surrounding the *Pvt1* p53RE are common. Several studies have documented translocations specifically involving the first exon of *Pvt1* fused to a number of 3' gene partners across a variety of malignancies (Iwakawa et al., 2013; Nagoshi et al., 2012; Northcott et al., 2012), and various somatic mutations encompassing the

Pvt1 promoter (Cho et al., 2018). In p53-proficient tumors, such genetic aberrations might be expected to either separate the *Myc* locus from *Pvt1b*, likely rendering *Pvt1b* incapable of repressing *Myc in trans*, or potentially destroy the p53RE required for *Pvt1b* expression, thus providing cells with a proliferative advantage. Future studies of *Pvt1* in cancer should therefore consider whether observed *Pvt1*-associated tumor-promoting effects derive from the activation of its oncogenic features, the disruption of its tumor suppressive features, or some combination thereof.

Pvt1 functional studies have largely been performed in cancer backgrounds. However, surprisingly little knowledge has been amassed concerning the function of *Pvt1* under physiological conditions, with the biological relevance of some proposed regulatory activities of *Pvt1* unclear due to their inherent cancer-specificity (Cho et al., 2018). Here I have provided insight into this mystery by demonstrating that *Pvt1b* is a target of the p53 gene expression program, and is therefore linked to normal cellular homeostasis and the cellular response to stress. Researchers long puzzled over the molecular mechanism(s) underpinning the negative regulatory feedback observed between the p53 and Myc pathways, and whether p53 activation played a direct or indirect role in *Myc* repression. Several models to explain this phenomenon have been proposed, including histone de-acetylation at the *Myc* promoter via direct p53 binding (Ho et al., 2005), p53 induction of *Myc*-targeting miRNAs (Christoffersen et al., 2010; Sachdeva et al., 2009), and p53 binding to a distal repressor element in the *Pvt1* locus to reduce *Myc* levels (Porter et al., 2017). *Pvt1* has been previously identified as a p53 target gene in several studies (Allen et al., 2014; Barsotti et al., 2012), but the role of *Pvt1*

within the p53 transcriptional program, if any, was virtually unexplored. Expanding on these findings, I provided evidence that induction of *Pvt1b* by p53 is the primary mediator of p53-dependent *Myc* repression using two distinct LOF systems. Therefore, *Pvt1b* joins a class of p53-regulated lncRNAs responsible for executing important growth-inhibitory functions within the p53-coordinated tumor suppressor response (Dimitrova et al., 2014; Hung et al., 2011; Schmitt et al., 2016). While p53 has been previously implicated in the indirect repression of cell cycle genes through the p53-DREAM (p53–p21–DREAM–E2F/CHR) pathway (*reviewed in* (Engeland, 2018)), to my knowledge *Pvt1b* is the first example of a p53-regulated lncRNA enacting local repression of a growth-promoting gene. Indeed, work from our lab has found that *cis*-acting lncRNA targets of p53 often engage in positive co-regulatory relationships with neighboring loci, while *Pvt1b* is currently the only known outlier in this trend (Tesfaye et al., *In review*). Taken together, I propose a model whereby p53 induces *Pvt1b* to dampen proliferative signaling during potential cancer-initiating events. As such, *Pvt1b* acts at the intersection between two pervasive transcription networks, providing a crucial avenue of communication that allows cells to prioritize tumor suppression over continued growth under conditions of stress.

While I describe a stress-specific role for *Pvt1b* in restricting *Myc* expression, the molecular mechanism underlying this regulatory relationship is still a matter of open investigation. Both the localization of *Pvt1b* near its TSS and the negative effect of *Pvt1b* induction on neighboring *Myc* transcription clearly establish a *cis*-regulatory model for *Pvt1b* function. Moreover, the effects of ASO-mediated depletion or endogenous activation of *Pvt1b* point to a role for the *Pvt1b*

RNA in *Myc* repression, as opposed to DNA elements in the locus. Importantly, p53 is not explicitly required for this effect, as upregulation of *Pvt1b* alone is sufficient to drive *Myc* downregulation. ASO-mediated depletion of *Pvt1b* reduced *Myc* levels, suggesting a role for the RNA molecule itself; however, we cannot formally exclude a role for the act of *Pvt1b* transcription in mediating *Myc* repression, especially in light of recent evidence that ASOs can cause premature transcription termination (Lee and Mendell, 2020). Indeed, my *Pvt1* exon 1b-targeting ASOs act near the *Pvt1b* TSS, rendering it impossible to distinguish between the effects mediated by the *Pvt1b* RNA or its production. However, my findings concerning the role of HDACs in *Myc* repression via p53-dependent deacetylation of the *Myc* promoter, and the potential requirement for *Pvt1b* in this process, while preliminary, strongly suggest a function for the lncRNA molecule. Modulation of epigenetic marks is a common mode by which lncRNAs regulate gene expression, especially *in cis* (see **Chapter 1**). Therefore, we may envision *Pvt1b* as acting in a similar manner to lncRNA *Morrbid*, regulating a neighboring gene via association with chromatin modifying complexes to influence critical cell survival decisions (Kotzin et al., 2016). Notably, the additional evidence that *Pvt1b* binds SHARP (Carter et al., 2020), a transcriptional repressor known to function in *Xist*-mediated histone deacetylation of the X-chromosome (McHugh et al., 2015) and predicted to have *Xist*-like repressive activity via computational analysis (Kirk et al., 2018), may implicate *Pvt1b* as engaging in a similar mechanism of action to *Xist*. Future work will explore this possibility.

Our inability to identify sequences in *Pvt1* exon 1b required for p53-dependent *Myc* repression does not preclude a role for the RNA molecule, or

indeed the existence of sequence-specific function. My attempts to mutate critical *Pvt1b* sequence and/or structural motifs may not have been efficient or pervasive enough to produce effects at the population level. On the other hand, the ~200 bp stretch of nucleotides unique to the *Pvt1b* isoform may be dispensable to its function. The existence of RNA sequence-independent lncRNA mechanisms has been insinuated by the nonspecific interactions between RNAs and PRC2 (Davidovich et al., 2013). Indeed, some proteins can engage in dynamic and promiscuous RNA interactions in the absence of a true RNA-binding domain due to intrinsically disordered regions (IDRs) (Protter et al., 2018). The structural disorder of IDRs also favors liquid-liquid phase separation (LLPS), and IDR-containing proteins have been shown to interact with lncRNAs in the formation of nuclear and cytoplasmic condensates (Garcia-Jove Navarro et al., 2019; Yamazaki et al., 2018). As such, *Pvt1b* may associate with IDR-containing proteins or other RBPs to repress *Myc* in a sequence-independent manner, although this possibility requires further investigation.

Critically, Nadya Dimitrova and Clara Liao demonstrated a role for *Pvt1b* in restricting tumor growth at the organismal level using an autochthonous mouse model of lung cancer. In the KC mouse, tumorigenesis is driven by oncogenic K-ras and is exquisitely sensitive to p53 loss. Tumor-specific *Pvt1b* ablation in this background had profound consequences, producing tumors that were larger than their *Pvt1b*-expressing counterparts and increasing tumor burden almost to the extent observed with p53 loss. Epistasis analysis and the absence of increases in tumor grade in *Pvt1b*-deficient tumors revealed a highly specific and powerful role for *Pvt1b* downstream of p53 activation in curtailing tumor growth. In this way,

Pvt1b acts as a critical barrier to the ability of cancer cells to proliferate unchecked, joining the ranks of a select number of cancer-associated lncRNAs with documented roles in tumorigenesis *in vivo* (Gupta et al., 2010; Gutschner et al., 2013; Mello et al., 2017). While this LOF model does not allow for differentiation between the effects of DNA elements in the *Pvt1* locus and the *Pvt1b* RNA, the *Pvt1a-PAS*, *Pvt1b-PAS*, and *Pvt1b-TWI* mouse models, generated in the course of this work, should distinguish the functional elements of *Pvt1* transcripts with a higher degree of sensitivity.

An exciting implication from this work is the significant influence *cis*-acting lncRNAs can exert over cellular activities via highly specific local gene regulation. *Pvt1b* represses one transcription factor downstream of another transcription factor, thus acting as a mediator between two gene expression networks to rapidly fine-tune cellular outputs in response to a specific cellular input. To my knowledge, *Pvt1b* is the first example of a lncRNA acting in this manner, although *cis*-regulation of transcription factors by lncRNAs may be far more commonplace (Liu et al., 2018). The closest example I could find of a similar lncRNA mechanism is that of *NANCI* (*Nkx2.1-Associated Noncoding Intergenic RNA*). *NANCI* is induced downstream of Wnt (Wingless/Integrated) signaling, upon which it upregulates its neighbor *NKX2.1* (*NK2 Homeobox 1*) to influence lung epithelial development through the activation of a plethora of NKX2.1 target genes (Herriges et al., 2014). Differing only in whether their *cis*-regulation is activating or repressive, *NANCI* and *Pvt1b* may represent the founding members of a burgeoning class of *cis*-acting lncRNAs that function at the intersection of transcription programs to influence gene expression on a broader scale. Considering the speed with which lncRNAs can

be produced relative to proteins, lncRNAs are uniquely poised to swiftly and dynamically respond to cellular stimuli. *Cis*-acting lncRNAs are even more well-suited for this task because their site of transcription is the same as their site of action. In this way, *Pvt1b* as an expeditious regulatory lever is similar to p53 itself, which is constitutively produced and degraded, and therefore always primed for rapid activation. Global analyses have revealed many lncRNAs expressed in close spatial proximity to genes with roles in transcription regulation including transcription factors and chromatin modifiers (Guttman et al., 2009; Ponjavic et al., 2009). Such lncRNAs are co-expressed with their neighboring protein-coding gene(s) more frequently than expected by chance, suggesting functional regulatory relationships (Ponjavic et al., 2009). Taken together, *cis*-acting lncRNAs may play outsized roles in global gene regulation and are apt candidates for transmitting rapid feedback between cellular pathways. As such, future studies should pay particular attention to lncRNAs adjacent to genes with broad transcription regulatory capabilities.

One compelling discovery from my work is the identification of an isoform-specific function for *Pvt1b*. Traditionally, studies have viewed lncRNA loci as discrete functional units, with many lncRNA genetic LOF models based on deletion of either the entire lncRNA locus, or the promoter, resulting in complete loss of lncRNA expression (see **Chapter 1**). In recent years, our expanding understanding of the numerous and interconnected functional elements of lncRNA genes has prompted the development of more targeted genetic models that attempt to disrupt or enhance specific features of the lncRNA transcript while preserving as much of the endogenous locus as possible (see **Chapter 1**). Increased attention

to how lncRNAs produce a particular regulatory output, and whether these functions are transcript-dependent or transcript-independent has encouraged the use of innovative and thoughtful experimental tools and approaches. That transcript-dependent lncRNA mechanisms can involve varied RNA-DNA, RNA-RNA, and RNA-protein interactions is now well-documented (*reviewed in* (Statello et al., 2020)). While it is understood that the abundance and availability of lncRNA interactors may change in different cellular contexts, little attention has been paid thus far to how alterations in the lncRNA transcript itself might influence function in response to cellular inputs.

The alternative splicing and processing of lncRNAs presents numerous opportunities for the production of transcripts with diverse functions and mechanisms of action. The sequence and/or structural motifs in a lncRNA transcript often dictate function by specifying the molecular interactions in which a lncRNA can engage (*reviewed in* (Zampetaki et al., 2018)). Therefore, it stands to reason that the inclusion or exclusion of defined regions of a lncRNA transcript through alternative transcript initiation, processing, or termination events may expand, restrict, or otherwise transform a lncRNA's regulatory repertoire. Multi-exonic lncRNAs can be spliced in numerous combinations, potentially producing transcripts with different functions due to the combination of specific sequence motifs, or the generation of rare exon-exon junctions, which may influence the set of possible RBPs, or other factors, a lncRNA can bind. Our analysis of splice junctions in nascent *Pvt1* RNAs revealed abundant transcripts produced from the *Pvt1* locus, composed of various combinations of exons. Similar diversity in spliceoforms has been observed for other lncRNAs (Niemczyk et al., 2013). There

are few examples of lncRNA isoforms executing distinct regulatory functions, including a SNP-specific function for *PCAT19* in prostate cancer (Hua et al., 2018) and a role for a long isoform of *CCAT1*, *CCAT1-L*, in transcriptional regulation of *MYC* (Xiang et al., 2014). However, *Pvt1* is unique in that the tumor suppressor function of the stress-induced *Pvt1b* coexists alongside the varied, and potentially oncogenic, activities of the constitutively transcribed *Pvt1a*. Fortunately, advances in the depth and sensitivity of RNA-sequencing technologies may enable more robust identification of lncRNA isoforms moving forward, paving the way for an improved understanding of lncRNA isoform-dependent activities.

Finally, this work has important implications for Myc-based therapeutic interventions in cancer. Myc-driven cancers can regress upon Myc inhibition (Soucek et al., 2002), suggesting its potency as a therapeutic target. However, drug development efforts have been challenged by the absence of a targetable binding pocket on the surface of Myc (Dang et al., 2017). LncRNA perturbation can provide alternative avenues for therapeutic intervention as a way of side-stepping so-called 'undruggable' proteins. This therapeutic perturbation can be accomplished in several ways, with the most common and clinic-ready approaches including: 1) small interfering RNAs (siRNAs) and ASOs to achieve lncRNA degradation and 2) ASOs, often in the form of locked nucleic acids (LNAs), to cause steric disruption of lncRNA function by altering splicing, inhibiting specific lncRNA-binding partner interactions or causing a change in secondary structure formation (Arun et al., 2018). In principle, *cis*-acting lncRNAs make particularly attractive therapeutic targets; with such confined regulatory outputs, their perturbation may be less likely to trigger unwanted off-target effects. Theoretically, *Pvt1b* induction

in p53-deficient tumors may provide a way to limit *Myc* expression at its source. However, current clinic-ready technologies target RNAs for degradation and, in practice, *Myc* repression via *Pvt1b* would instead require its upregulation from the endogenous locus. Targeting lncRNA molecules to specific genomic locations *in vitro* has been made possible by CRISPR (Shechner et al., 2015). However, deployment of this technology *in vivo* is unlikely to happen for many years. Nonetheless, the prospect of modulating *Myc* expression in cancer, via *Pvt1b* or otherwise, is exciting and should be explored further in the future.

Final Remarks

Previous work has often treated *Pvt1* as a simple genetic unit, not always leaving space for complex regulatory functions that may arise from its 300 Kb of genomic information. My distinction between the *Pvt1a* and *Pvt1b* isoforms provides an avenue to re-evaluate *Pvt1* and, by extension, its relationship to *Myc*, in a new light. My findings do not necessarily conflict with the body of literature supporting a synergistic relationship between *Myc* and *Pvt1*. Rather, they raise the possibility that *Pvt1a* and *Pvt1b* may have distinct, and perhaps opposing, functions, with DNA elements in the *Pvt1* locus and *Pvt1a* cooperating with *Myc* to promote its expression and activity, and *Pvt1b* acting as a stress-specific molecular brake for this process. This multiplicity of function may not be restricted to the *Pvt1* locus, suggesting that the lncRNA class as a whole should be re-examined for alternative functions encoded in lncRNA loci. Our work further highlights the potential for *cis*-acting lncRNAs, when expressed in close proximity to transcription factors or other protein-coding genes with widespread influence, to exert profound control over cellular operations. Taken together, the results presented here implicate *Pvt1b* as a central node of communication between the p53 and *Myc* transcription networks, which enacts selective gene repression downstream of a broad transcriptional activator to limit cell growth and perhaps prevent cancer before its onset. Future studies should focus on both identifying the element(s) of *Pvt1b* required for *Myc* repression, and disentangling the intricacies of the *Pvt1* locus to illuminate its varied, and perhaps isoform-specific, functions.

Chapter 7: Materials and Methods

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse strains

All animal work was conducted in accordance with a protocol approved by the Yale University Institutional Animal Care and Use Committee. *K-ras*$^{LSL\text{-}G12D/+}$ *(K)* and *p53*$^{FL/FL}$ *(P)* mice were previously described (Jackson et al., 2005; Jackson et al., 2001) and obtained from the laboratory of T. Jacks (MIT). *Rosa26-Cas9*$^{LSL/LSL}$ *(C)* mice were previously described (Platt et al., 2014) and purchased from Jackson Laboratories (026556). Wild-type (WT) C57BL/6J mice were purchased from Jackson Laboratories (000664). *Pvt1a-PAS*, *Pvt1b-PAS* and *Pvt1b-TWI* mice were generated using CRISPR/Cas9-mediated engineering in C57BL/6J mice in collaboration with Rick Maser and Adam Williams at the Jackson Laboratory for Genomic Medicine. In brief, guide RNA:Cas9 ribonucleoprotein (RNP) complexes and HDR templates were introduced into embryos via electroporation. Mice carrying successful knock-ins were identified by PCR-based genotyping and germline transmission was established by backcrossing once to WT (C57BL/6J) mice. Guide RNA and HDR template sequences can be found in Supplementary Table 1; genotyping primer sequences can be found in Supplementary Table 2.

For irradiation experiments, 4-8 months-old mice were irradiated with 6 Gy of whole body irradiation and sacrificed 6 hours post irradiation. For tumor studies, 3-6 months-old mice were used. Experiments were performed blind to gender and

with an equal distribution of males and females in each experimental group.

Cell culture and drug treatments

WT MEFs were isolated from embryos at E13.5 from timed matings of WT C57BL/6J animals. All MEF experiments were performed at passages 2-10. *KPR8* lung adenocarcinoma cell line of the genotype *K-ras$^{G12D/+}$; p53$^{LSL/LSL}$; Rosa-CreER^{T2}* was previously established from spontaneously arising primary tumors isolated from *K-ras$^{LA2\text{-}G12D/+}$; p53$^{LSL/LSL}$; Rosa-CreER^{T2}* mice, as previously described (Feldser et al., 2010). p53-restorable *p53$^{LSL/LSL}$; Rosa-CreER^{T2}* MEFs were previously described (Ventura et al., 2007). Genotypes and Tam-mediated restoration of p53 expression were validated by genotyping and by qRT-PCR and immunoblotting, respectively. Puromycin-sensitive *KPR8* (*KPR*) and p53-restorable MEF clones were generated by transient transfection with a guide RNA targeting the ORF of puromycin to inactivate the puromycin-resistance gene expressed from the Stop cassette, cloned downstream of a U6 promoter in a BRD004 lentiviral construct (a gift from the Broad Institute, MIT) that co-expresses spCas9 and GFP. Normal human fetal lung fibroblasts were purchased from the NIA Aging Cell Culture Repository (TIG-1, NG06173). Primary MEFs and human fibroblasts were maintained in DMEM (Gibco) supplemented with 15% FBS (F0926, Sigma-Aldrich), 50 U/ml pen/strep (Gibco), 2 mM L-glutamine (Gibco), 0.1 mM non-essential amino acids (Gibco), and 0.055 mM β-mercaptoethanol (Gibco). Cancer cells and 293 viral packaging cells were cultured in DMEM supplemented with 10% FBS, 50 U/ml pen/strep, 2 mM L-glutamine, and 0.1 mM non-essential amino acids. All cell cultures were maintained at 37°C

in a humidified incubator with 5% CO_2. Viral titering was performed in 3TZ cells, a derivative of 3T3 cells, expressing a LSL-LacZ transgene (generously provided by the laboratory of T. Jacks, MIT).

To delete the loxP-STOP-loxP (LSL) cassette preventing p53 expression, cells were treated with 0.5 μM 4-hydroxytamoxifen (Tam, Cayman Chemical Company). To induce DNA damage, cells were treated with 0.5 μM doxorubicin (Doxo, Sigma-Aldrich) or 50 μM etoposide (Etop, Millipore Sigma) for smRNA-FISH studies. To assess protein stability, cells were treated with 50 μg/ml cycloheximide (Chx, Sigma-Aldrich) for the indicated times. To inhibit HDAC3, cells were treated with 14 μM RGFP966 (MedChem Express). To inhibit all HDACs, *KPR* cells were treated with 50 ng/ml trichostatin A (TSA, Sigma-Aldrich) or 50 nM suberoylanilide hydroxamic acid (SAHA, Sigma-Aldrich) and p53-restorable MEFs were treated with 25 ng/ml TSA or 75 nM SAHA, respectively.

Constructs

Mutagenesis of p53REs in cultured cells was performed with a gRNA targeting the p53RE of *Pvt1b* (gΔRE) or Gm26542 (g1 or g2), cloned downstream of a U6 promoter in BRD001 or BRD004 lentiviral constructs (gifts from the Broad Institute, MIT) that co-express spCas9 and either an IRES-driven puromycin-resistance gene or GFP, respectively. Mutagenesis of *Pvt1* exon 1b or the *Pvt1* exon 1b splice site in cultured cells was performed with gRNAs targeting across the *Pvt1* exon 1b sequence (gALT1-11) or a gRNA targeting the 3' end of exon 1b (gdeltaSS), respectively, each cloned into the BRD001 lentiviral construct, as previously

described. Control gRNA targeting dTomato (Con) was used as a negative control. Tumor-specific mutagenesis of p53REs *in vivo* was performed with gRNAs cloned downstream of a U6 promoter in UGPC (U6-gRNA-PGK-Cre) lentiviral vector. UGPC-Con targeting dTomato was used as a negative control. UGPC-p53KO targeting the ORF of p53 was used as a positive control (Xue et al., 2014). For CRISPRa experiments, a lentiviral vector (lenti-SAM-Hygro) was constructed to co-express nuclease-proficient spCas9, a U6-driven 15-mer 'dead RNA' (dRNA) extended by two MS2 loops (dRNA-MS2) (Dahlman et al., 2015), the transcriptional activator domains p65 and HSF1 fused to the MS2-binding protein (MBP), and a hygromycin-resistance gene. All sgRNA and dRNA sequences used in this study can be found in Supplementary Table 1.

Lentivirus was produced in 293 cells by co-transfecting the lentiviral constructs with pCMV-dR8.2 dvpr (Addgene plasmid #8455) and pCMV-VSV-G (Addgene plasmid #8454) viral packaging constructs. Viral containing supernatants supplemented with 4 μg/ml polybrene (Millipore Sigma) were used to infect WT MEFs and *KPR* cells by 2-3 consecutive lentiviral infections, delivered at 24 hour-intervals. Following infections, cells were selected with 5 μg/ml (*KPR*) or 2 μg/ml (MEFs) puromycin (Sigma-Aldrich) or 800 μg/ml hygromycin (Roche). UGPC lentivirus was prepared as above, concentrated by ultracentrifugation, and titered by infecting 3TZ cells and determining the number of viral particles based on the fraction of LacZ-positive cells as previously described (DuPage et al., 2009).

Mutagenesis of the *Pvt1b* and Gm26542 p53REs was confirmed by PCR amplification of the region, subsequent cloning into pCR-Blunt II-TOPO® vector (Invitrogen) and Sanger sequencing. Mutagenesis of the *Pvt1* exon 1b sequence and

the *Pvt1* exon 1b splice site was confirmed by PCR amplification of the region, Sanger sequencing and Tracking of Indels by Decomposition (TIDE) analysis (Brinkman et al., 2014). Primers used in mutagenesis efficiency estimates can be found in Supplementary Table 2.

For overexpression experiments, full-length *Pvt1a* (exon 1a-10) and *Pvt1b* (exon 1b-10) cDNAs were synthesized as gene blocks and cloned into pWZL Hygro retroviral vector (Addgene plasmid #18750). 5 μg of empty vector, *Pvt1a*-, or *Pvt1b*-expressing constructs were transfected into 1-3x10^6 WT MEFs using the Amaxa Mouse/Rat Hepatocyte Nucleofector Kit (Lonza, VPL-1004) and the Nucleofector 2b Device (Lonza). Analysis was performed at 48 hours post transfection. Information about key plasmids used in this work can be found in Supplementary Table 3.

METHOD DETAILS

RNA isolation and qRT-PCR

For RNA-seq and qRT-PCR analysis, RNA was isolated with the RNeasy Mini Kit (Qiagen) and 0.5-1 μg of total RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). SYBR Green PCR master mix (Kapa Biosystems) was used for quantitative PCR in triplicate reactions with primers listed in Supplementary Table 2. Relative RNA expression levels were calculated using the ddCt method compared to GAPDH and normalized to control samples.

Immunoblotting

Cells were collected, counted, and lysed in 2×Laemmli buffer (100 mM Tris-HCl pH6.8, 200 mM DTT, 3% SDS, 20% glycerol) at 0.5-1x10^4 cells/μl. Samples were heated at 95°C for 7 minutes and passed through an insulin syringe. Protein from 1x10^5 cells was separated on 10% SDS-polyacrylamide gels and transferred to nitrocellulose membranes (Bio-Rad). After blocking (5% milk, PBST), membranes were incubated overnight at 4°C in primary antibody, then 1hr at RT in secondary antibody. The following antibodies were used: anti-c-Myc (1:1000, clone Y69, ab32072, Abcam), anti-Hsp90 (1:2500, 610419, BD Transduction Laboratories), anti-Hsp90 (1:1000, 4877S, Cell Signaling Technology), goat anti-mouse secondary antibody (1:50000, 1706516, Bio-Rad), and donkey anti-rabbit secondary antibody (1:50000, 711-035-152, Jackson ImmunoResearch). Protein bands were visualized using Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare). Quantification of Myc and Hsp90 protein levels was performed using the rectangle selection and measure tools in FIJI and Myc levels plotted relative to Hsp90 levels and normalized to negative control in relevant graphs. For cycloheximide experiments, Myc levels were normalized to negative control and half-life of Myc protein was determined using Prism8 software.

Chromatin immunoprecipitation (ChIP)

Cells were harvested by trypsinization, counted, washed once in PBS and crosslinked in 1% methanol-free formaldehyde (Thermo Scientific) diluted in PBS for 10 min at RT. The reaction was stopped by adding glycine to a final concentration of 100 mM and placing the samples on ice for 5 min. Cells were

washed twice in cold PBS and the pellet was frozen and stored at -80°C.

5-10x10^6 nuclei were isolated by incubating the thawed cell pellet in Cell lysis buffer (20 mM Tris-HCl, pH 8.0, 85 mM KCl, 0.5% NP-40), supplemented with protease inhibitors (1 mM PMSF and Mini Complete Protease Inhibitor Cocktail Tablet, Roche) on ice for 10 min. After centrifugation, the supernatant was removed and the nuclei were resuspended in Nuclei lysis buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% SDS supplemented with protease inhibitors) and incubated for 10 min on ice. Next, chromatin was sonicated to 300-500 bp fragment size in an ice-water slurry for 10 cycles (15" ON, 30" OFF) for p53 ChIP and 10-12 cycles (10" ON, 30" OFF) for H3K27ac ChIP using a Bioruptor sonicator (Diagenode). Sonicated lysates were centrifuged at 13K rpm for 20 min and diluted in ChIP dilution buffer (0.01% SDS, 1.1% Triton- X100, 1.1 mM EDTA, 20 mM Tris-HCl, pH 8.0, 167 mM NaCl, supplemented with protease inhibitors). Input aliquots were saved at this point. The sonicated chromatin was precleared with beads (PureProteome Protein G Magnetic Beads, Millipore Sigma) and used to set up chromatin immunoprecipitations with a p53 antibody (P53-CM5P-L, Leica), H3K27ac antibody (ab4729, Abcam), or control IgG (ab46540, Abcam) and incubated overnight at 4°C on a rotator. Beads (PureProteome Protein G Magnetic Beads, Millipore Sigma) were blocked overnight in 1% BSA in PBS supplemented with 20 μg salmon sperm DNA (Invitrogen) per immunoprecipitation reaction. The next day, the blocked beads were added to the immunoprecipitation reactions and samples were incubated on the rotator for an additional hour. Beads were washed once in each of the following washes for 5 min at 4°C on the rotator: Low salt wash (0.1% SDS, 1% Triton-X100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 150

mM NaCl supplemented with protease inhibitors), High salt wash (0.1% SDS, 1% Triton-X100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.0, 500 mM NaCl), LiCl wash (0.25 M LiCl, 1% NP-40, 1% Na deoxycholate, 1 mM EDTA, 20 mM Tris-HCl, pH 8.0), and TE wash (10 mM Tris-HCl, pH 8.0, 1 mM EDTA).

After completely removing any remaining liquid from the washes, beads were resuspended in Elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 1% SDS) and incubated at 65°C for 15 min with frequent vortexing to prevent settling. After elution, the beads were pelleted, and the supernatant was transferred to a new tube and incubated overnight at 65°C to reverse the crosslinking. The next day, samples were treated with RNaseA or 2 hours at 37°C, followed by a proteinase K (Roche) treatment for 30 min at 55°C. The DNA was purified by phenol-chloroform extraction and EtOH precipitation. The DNA pellet was air dried, resuspended in 200 µl H_2O and used for quantitative PCR analysis (ChIP-qPCR) using primers listed in Supplementary Table 2.

Single-molecule FISH (smRNA-FISH)

Quasar570 (Q570)- and Quasar670 (Q670)-conjugated Stellaris FISH probes are listed in Supplementary Table 1 (Stellaris, LGC Biosciences). smRNA-FISH was performed according to the manufacturer recommendations. Briefly, cells were grown on coverslips and fixed for 10 min in 4% methanol-free formaldehyde (Thermo Scientific) at RT, followed by PBS washes. Cells were dehydrated overnight at 4°C in 70% EtOH (diluted in DEPC-H_2O) and stored in 70% EtOH for up to a week at 4°C. Coverslips were transferred to a hybridization chamber and equilibrated for 5 min in Wash Buffer A (Stellaris, LGC Biosciences)

prepared with formamide (Millipore Sigma) according to manufacturer's instructions. Cells were incubated overnight at 30°C with the indicated probes diluted 1:50 in Hybridization solution (Stellaris, LGC Biosciences) prepared with formamide according to manufacturer's instructions. The next day, cells were washed 2 times for 30 min at 30°C in Wash Buffer A, incubated in Wash Buffer B (Stellaris, LGC Biosciences) for 5 min at RT, and mounted in antifade reagent (Vectashield Mounting medium with DAPI, Vector Laboratories). The following probesets were used: *Pvt1b (ex.1b)* detecting *Pvt1b* isoform with 10 probes spanning exon 1b, labeled with Q670 and false-colored in red; *Pvt1a (ex.1a)* detecting *Pvt1a* isoform with 11 probes spanning exon 1a, labeled with Q670 and false-colored in red; *Pvt1 (ex.1a-10)* detecting total *Pvt1* with 48 probes spanning exons 1a-10, labeled with Q570 and false-colored in green; *Pvt1 (introns)* detecting nascent *Pvt1a* with 31 probes spanning intron 1 upstream of exon 1b, labeled with Q670 and false-colored in red; and *Myc (intron)* detecting nascent *Myc* with 33 probes spanning intron 1 of *Myc*, labeled with Q570 and false-colored in green. *Pvt1a (ex.1a)* and *Pvt1b (ex.1b)* probesets do not detect at the single molecule level. Images were captured using an Axio Imager 2 microscope system (Zeiss) with a PlanApo 63x 1.4 oil DIC objective lens (Zeiss). For *KPR* cells, z-stacks of 12 planes at 0.5 μm steps were acquired and used to generate maximum intensity projections. For WT MEFs, single plane images were acquired. All images were edited using Adobe Photoshop.

DNA-Fluorescence *in situ* hybridization (FISH)

DNA-FISH was performed as previously described (Chaumeil et al., 2008).

To generate probes, the following BAC clones were used: RP23-55F11 (Myc) and RP24-301E22 (Chr 6) (BACPAC Resources). BAC DNA was purified with a Nucleobond Xtra BAC kit (Takara Bio USA) and nick translated with a nick translation system (Invitrogen) and Alexa Fluor® 488-5-UTP or Alexa Fluor® 594-5-UTP (Invitrogen) following manufacturer instructions. Final probes were ethanol precipitated with 7.5M ammonium acetate and stored in sterile TE at -20°C.

20 ng of nick-translated probe was precipitated with 3 μg of salmon sperm DNA (Invitrogen) and 1 μg of mouse COT1 DNA (Invitrogen) using 1/10th volume of sodium acetate (3M, pH 5.5) and 2.5 volumes of ethanol. Probes were stored overnight at -20°C, then centrifuged at 13K rpm for 30 min at 4°C, washed twice with 70% ethanol, and air dried. Pellets were resuspended in formamide (Millipore Sigma), incubated at 37°C for at least 10 min, and denatured for 7 min at 75°C. After denaturing, an equal volume of 2X hybridization buffer (4X SSC, 20% w/v dextran sulfate, 2 mg/mL BSA, 40 mM RVC) was added and probe-DNA mixtures were pre-annealed for 30 min to 1 hour at 37°C.

Cells were plated on coverslips and fixed in 4% paraformaldehyde in PBS for 10 min at RT, followed by PBS washes. Cells were permeabilized in 0.5% Triton X-100 in PBS for 6 min, washed twice with 70% ethanol and stored in 70% ethanol at -20°C. Cover slips were dehydrated in an ethanol series (80%, 90%, 100%), air dried, and incubated in RNase A diluted in 2X SSC (100 μg/mL) for 1 hour at 37°C. Cover slips were washed three times with 2X SSC for 5 min and incubated in 50 μg/mL pepsin diluted into prewarmed 0.01M HCl for 3 min at 37°C, followed by

two 5 min PBS washes and one in 1X PBS/$MgCL_2$. After washing, cover slips were incubated in 1% formaldehyde (Thermo Scientific) in 1X PBS/$MgCL_2$ for 10 min at RT. Cover slips were next washed in PBS for 5 min and dehydrated in an ethanol series (70%/90%/100%) and air dried. Cover slips were then denatured in prewarmed 50% formamide in 2X SSC for 30 min at 80°C, dehydrated in an ice-cold ethanol series (70%/90%/100%), and incubated with denatured probe DNA overnight at 42°C in a dark chamber humidified with 50% formamide in 2X SSC. Following incubation, cover slips were washed three times with prewarmed 50% formamide in 2X SSC at 42°C for 5 min and three times with prewarmed 2X SSC at 42°C for 5 min. Cover slips were mounted on slides with antifade mounting medium with DAPI (Vector Laboratories) and sealed with nail polish. Single plane images were captured using an Axio Imager 2 microscope system (Zeiss) with a PlanApo 63x 1.4 oil DIC objective lens (Zeiss).

Subcellular fractionation

Subcellular fractionation was performed as previously described (Conrad and Orom, 2017) with slight modifications. Briefly, cells were harvested by trypsinization, rinsed once in PBS and re-suspended in 1 mM EDTA in PBS. $1x10^6$ cells were set aside for whole cell (WC) RNA isolation using TRIzol (Invitrogen) following the manufacturer's protocol. 3 $x10^6$ cells were lysed in 0.4 mL cell lysis buffer (10 mM TrisHCl pH 7.5, 0.15% NP-40, 150 mM NaCl, 100 U/mL RNase-IN (Promega) for 5 min on ice. Lysate was layered on a sucrose cushion (24% w/v sucrose, 150 mM NaCl, 10 mM TrisHCl pH 7.5,100 U/mL RNase-IN) and centrifuged for 10 min at 3,500g, yielding the cleared cytoplasmic fraction

(supernatant) and pelleted nuclei. Nuclear pellets were washed once in PBS supplemented with 1 mM EDTA, re-suspended in 0.25 mL glycerol buffer (50% glycerol, 20 mM Tris-HCl pH 7.5, 75 mM NaCl, 0.5 mM EDTA, 0.85 mM DTT, 100 U/mL RNase-IN), and lysed by the immediate addition of an equal volume nuclear lysis buffer (10 mM HEPES pH 7.6, 7.5 mM $MgCl_2$, 0.2 mM EDTA, 300 mM NaCl, 1% NP-40, 1 mM DTT, 1M Urea, 100 U/mL RNase-IN) with 2 min incubation on ice. Centrifugation for 2 min at 18,800g yielded the nucleoplasmic and chromatin-associated fractions in the supernatant and pellet, respectively. Chromatin pellets were washed once in 1 mM EDTA in PBS and solubilized in 1 mL TRIzol reagent by syringing. RNA was extracted from the cytoplasmic and nucleoplasmic fractions using TRIzol-LS (Invitrogen) and from the chromatin-associated fraction using TRIzol following the manufacturer's protocols. Subcellular RNA enrichment patterns were determined by qRT-PCR, normalizing fraction Ct values to WC Ct values. Cytoplasmically-enriched RNA *Rn7s1* and chromatin-enriched RNA *Kcnq1ot1* served as fractionation quality controls. Primer sequences can be found in Supplementary Table 2.

Antisense knockdown

1 μM *Pvt1*-targeting (ASO1, ASO2, and ASO3) or control (CON) antisense LNA Gapmers (Exiqon, Qiagen) were transfected into 1-3x10^6 MEFs using the Amaxa Mouse/Rat Hepatocyte Nucleofector Kit (Lonza, VPL-1004) and the Nucleofector 2b Device (Lonza). Knockdown of *Pvt1* variants and the corresponding effects on *p21* and *Myc* expression were assayed at 72 hours post-transfection by qRT-PCR following the indicated treatments. The sequences of all

ASOs are listed in Supplementary Table 1.

Chromosome Conformation Capture (3C)

Chromosome conformation capture was performed as described previously with minor modifications (Hagege et al., 2007). Briefly, cells were harvested by trypsinization, counted, washed once in PBS and 5-10x10^6 cells were crosslinked in 1% methanol-free formaldehyde (Thermo Scientific) diluted in PBS for 10 min at RT. The reaction was stopped by adding 1.425 ml of 1 M glycine. Cell pellets were frozen in a bath of dry ice covered in 100% EtOH and stored at -80°C, or were processed immediately. Cells were lysed in 5 ml cell lysis buffer (20 mM TrisHCl pH8.0, 85 nM KCl, 0.5% NP-40, 5 mM $MgCl_2$, 0.1 mM EGTA) including 1x complete protease inhibitor (Roche). Cell nuclei were resuspended in 0.5 ml of 1.2x Cutsmart restriction buffer (New England Biolabs) and SDS was added to each tube to a final concentration of 0.3%. Following extraction with 2% Triton X-100, chromatin was digested overnight at 37°C with 400-800 U BamHI-HF (New England Biolabs). Ligations were performed in a total reaction volume of 6.125 mL of 1.15x ligation buffer (10x Ligation Buffer: 600 mM Tris-HCl pH7.5, 50 mM DTT, 50 mM $MgCl_2$, 10 mM ATP (New England Biolabs) using 100 U of T4 DNA ligase (New England Biolabs) with incubation at 16°C for 4 h, followed by further incubation at RT for 30 min. Reversal of crosslinking was performed by adding 300 μg proteinase K (Roche) followed by incubation at 65°C overnight. DNA was extracted with phenol-chloroform followed by EtOH precipitation. The efficiency of restriction enzyme digestion was examined using qRT-PCR with primer sets spanning BamHI sites. The concentrations of 3C libraries were determined by

qRT-PCR and compared to a genomic DNA reference of known concentration. Samples were subsequently diluted to a concentration of 20 ng/μl and a total of 50 ng was used for each qRT-PCR reaction. Interaction frequencies were determined using a unidirectional primer strategy with an anchor designed against the promoter of *Myc* (A1) and were normalized to a control region in the *Myc-Pvt1* locus. The primer sequences can be found in Supplementary Table 2.

RNA-seq

Total RNA was isolated in three biological replicates. PolyA selection and cDNA library preparation was performed using TruSeq Stranded mRNA Library Prep (Illumina). Paired-end 100 bp sequencing was performed on an Illumina HiSeq 4000 instrument. RNA-seq read files were merged from technical replicates and mapped to the mm10 genome assembly using Tophat (ver 2.0.14) (Trapnell et al., 2009) with gencode (vM10) annotation used as the transcriptome index. Additional transcripts were assembled using stringtie (1.2.4) (Pertea et al., 2015) and reads within exon sequences counted using HTSeq (HTSeq-0.6.1) counts (Anders et al., 2015). The differential expression analysis was performed with EdgeR (3.22.3) (using general linear model settings for biological triplicates with blocked matrix model for paired comparisons) (Robinson et al., 2010). For analysis of Myc targets, the Hallmark Gene Set in the Molecular Signature Database (Broad Institute) (Liberzon et al., 2015) was used and compared to randomly selected and expression matched genes with statistical significance of differential expression determined with a Kolmogorov-Smirnov test.

Transcriptome-wide TimeLapse-seq

At approximately 60% cellular confluence, media was spiked with a final concentration of 100 μM s^4U (Alfa Aesar) and grown in the dark for 1 hour. Cells were rinsed once with PBS, scraped from plates, suspended in 1 mL TRIzol (Invitrogen), and frozen overnight at -80°C. Total RNA was purified and treated with TimeLapse chemistry essentially as described (Schofield et al., 2018) with minor modifications. Briefly, following chloroform extraction and isopropanol precipitation (supplemented with 1 mM DTT) genomic DNA was depleted by treating with TURBO DNase (Invitrogen) and total RNA was extracted with acidic phenol:chloroform:isoamyl alcohol and EtOH precipitation. Isolated total RNA was mixed with 600 mM TFEA, 1 mM EDTA and 100 mM sodium acetate, pH 5.2 in water. A solution of 10 mM $NaIO_4$ was added and the reaction mixture was incubated at 45°C for 1 hr. Chemically treated RNA was purified using Agencourt RNAclean XP beads (1 equivalent volume, Beckman Coulter) according to manufacturer's instructions. Purified material was then incubated in a reducing buffer (10 mM DTT, 100 mM NaCl, 10 mM Tris pH 7.4, 1 mM EDTA) at 37°C for 30 min, followed by a second RNAclean bead purification. For each sample, 10 ng of total RNA input was used to prepare sequencing libraries from the Clontech SMARTer Stranded Total RNA-Seq kit Pico Input (Takara Bio USA) with ribosomal cDNA depletion. Paired-end 100 bp sequencing was performed on an Illumina HiSeq 4000 instrument.

TT-TimeLapse-seq

At approximately 60% cellular confluence, media was spiked with s^4U (1 μM final, Alfa Aesar) and cells were grown in the dark for 5 min. Total RNA and DNA isolation were performed as described above. Total RNA (50 μg) was biotinylated with MTSEA biotin-XX (Biotium), isolation and streptavidin enrichment essentially as described (Schofield et al., 2018). Enriched RNA was chemically treated as described above. Library construction and sequencing were performed essentially as described above.

TimeLapse-seq mutational analysis

Filtering and alignment to the mouse GRCm38.p5 were performed essentially as described previously (Schofield et al., 2018). Briefly, reads were filtered to remove duplicate sequences with FastUniq (Xu et al., 2012), trimmed of adaptor sequences with Cutadapt v1.16 (Martin, 2011) and aligned to GRCm38 using HISAT2 v2.1.0 (Kim et al., 2015a) (with default parameters except -mp 4,2). Reads aligning to transcripts were quantified with HTSeq (Anders et al., 2015) htseq-count. SAMtools v1.5 (Li et al., 2009) was used to collect only read pairs with a mapping quality greater than 2 and concordant alignment (sam FLAG = 147/99 or 83/163). Mutation calling was performed essentially as described previously (Schofield et al., 2018). Briefly, T-to-C mutations were only considered if they met several conditions. Mutations must have a base quality score greater than 40 and be more than 3 nucleotides from the read's end. Sites of likely single-nucleotide polymorphisms (SNPs) and alignment artifacts were identified with bcftools or from sites of high mutation levels in the non-s^4U treated controls and were not

considered in mutation calling. Browser tracks were made using STAR v2.5.3a (Dobin et al., 2013) and normalized across samples using scale factors calculated using RNA-seq reads using edgeR (Robinson et al., 2010) (calcNormFactors using method = 'upperquartile').

Differential expression analysis

Differential expression analysis of transcriptome-wide TimeLapse-seq and TT-TimeLapse-seq data was performed with DESeq2 (Love et al., 2014) essentially as described previously (Schofield et al., 2018). DESeq2 expression analysis was performed on TT-TimeLapse-seq and transcriptome-wide TimeLapse-seq data to determine changes in transcriptional activity and mRNA expression, respectively.

Growth curve and colony assay

To generate growth curves, Con-, ΔRE-, sg1-, or sg2-expressing *KPR* cells were grown in the presence or absence of Tam. Population doublings over indicated time course were plotted as the average of three independent experiments. For colony assays, $4x10^5$ Con- or ΔRE -expressing *KPR* cells were plated in the presence of Tam in 6 cm dishes and monitored for colony formation. Plates were washed with PBS, fixed in 0.5% Crystal Violet; 25% MeOH for 10 minutes and washed in ddH_2O. The average of three biological replicates is shown.

Tumor studies

Lung tumorigenesis was initiated in cohorts of *KC* and *KPC* mice as described in (DuPage et al., 2009) by intratracheal infection with $1x10^5$ pfu UGPC

lentiviruses . Mice were analyzed at 12 or 16 weeks post tumor initiation. For histological analyses, lungs were inflated with 4% paraformaldehyde, and fixed overnight in 4% paraformaldehyde, prior to dehydration in 70% ethanol. Fixed lungs were embedded in paraffin, sectioned, and stained with hematoxylin and eosin (H&E). Tumor burden scored as tumor area relative to total lung area was determined using the freehand selection tool and Measure feature in ImageJ on images acquired with an Axio Imager 2 microscope system (Zeiss) with a PlanApo 10x 0.3 objective lens (Zeiss). Tumor grade was scored as previously described (DuPage et al., 2009; Nikitin et al., 2004).

Immunohistochemistry

Immunohistochemistry on paraffin sections was performed using the ABC Vectastain kit (Vector Labs) with an antibody to pHH3 Serine 10 (9701S, Cell Signaling Technologies). The staining was visualized with DAB (Vector Labs) and slides were counterstained with hematoxylin.

QUANTIFICATION AND STATISTICAL ANALYSIS

In relevant figures, figure legends convey the statistical details of experiments including statistical tests used and type and number (n) of biological replicates, while asterisks define degree of significance as described. All Student's t-tests and Mann-Whitney *U*-tests were analyzed in two sided. All sequencing data were aligned to the mouse genome (GRCm38/mm10). All statistical analyses were performed and graphics were generated using Prism8 software. Tracking of Indels by DEcomposition (TIDE) analysis was used to estimate mutagenesis efficiency.

DATA AND SOFTWARE AVAILABILITY

All software used in this work is listed in Supplementary Table 4. Data generated in (Olivero et al., 2020) are available through Gene Expression Omnibus (GEO) under accession number GEO: GSE126940. Some figures were created using graphics from www.Biorender.com.

Supplementary Tables

Supplementary Table 1: ASO, dRNA, gRNA, HDR template, and smRNA FISH probe information

ASOs		
Name	**Target**	**Sequence**
Con	N/A	GCTCCCTTCAATCCAA
ASO1	Pvt1 exon 1b	GTAACTAGCACACATC
ASO2	Pvt1 exon 1b	TTTGCTCCTTCTAAAT
ASO3	Pvt1 exon 1b	GAGTCCATGTGACGTT
SpCas9 dRNAs		
Name	**Target**	**Sequence**
Con	dTomato	gCGAGTTCGAGATCGA
A1	Pvt1a TSS	GGAGATCGGGGACAC
A2	Pvt1b TSS	gATGGTCATAGCTAGT
SpCas9 gRNAs		
Name	**Target**	**Sequence**
Con	dTomato	GGCCACGAGTTCGAGATCGA
RE	Pvt1b p53RE	gATATGGGCAGTGACAAGTTT
p53	p53 ORF	GTGTAATAGCTCCTGCATGG
sg1	Gm26542 p53RE	gTCTGAGGCCTGGGACTTGCC
sg2	Gm26542 p53RE	GGACTTGCTCAGTTCTTGGA
sgALT1	Pvt1 exon 1b	gAAACACAAACGCTTTCCCAC
sgALT2	Pvt1 exon 1b	gTTCTTAAAGCTCTAGCCAGT
sgALT3	Pvt1 exon 1b	gAAGTCCCACTTGGAGCTCCA
sgALT4	Pvt1 exon 1b	gTCTATCCTTGGAGCTCCAAG
sgALT5	Pvt1 exon 1b	GACTTCTTAAAAGATTTAGA
sgALT6	Pvt1 exon 1b	gTTAGAAGGAGCAAAGCTGTC
sgALT7	Pvt1 exon 1b	gAGGAAATCAGAAACGTCACA
sgALT8	Pvt1 exon 1b	gCGTCACATGGACTCCATGAC
sgALT9	Pvt1 exon 1b	GACTGGGAAAAACCTCGTGG
sgALT10	Pvt1 exon 1b	GGTGGCCTGCTCTCAGTGCT
sgALT11	Pvt1 exon 1b	GATGTGTGCTAGTTACATCT
sgdeltaSS	Pvt1 exon 1b 3' end	GTGTGCTAGTTACATCTCGG
Puro	Puromycin resistance gene	gCGGGTGGCGAGGCGCACCGT
***in vivo* SpCas9 gRNAs**		
Name	**Target**	**Sequence**
sgEx1a	Pvt1 exon 1a	gCTGGTCAAGCGGGCTCGGCA
sgEx1b	Pvt1 exon 1b	gTCTATCCTTGGAGCTCCAAG
HDR Templates		
Name	**Target**	**Sequence**
Pvt1a-PAS	Pvt1 exon 1a	CCGGGGCTGCCAACATCCTTTCC ACGCGGATATCCACTCGGGGGG CTCTGGGAATGCTAAGTTCGTAG CTTCTCTTCATCCTGGCCTTGCCA CACAAAAAACCAACACACAGATC

		TAATGAAAATAAAGATCTTTTAT TCGAGCCCGCTTGACCAGTGGGT CCATGTGCTCGGCGGCCA
Pvt1b-PAS	Pvt1 exon 1b	CACGAGGTTTTTCCCAGTCATGG AGTCCATGTGACGTTTCTGATTT CCTGACAGCTTTGCTCCTTCTAA ATCTTTTAAGAAGTCCCACTTCA CACAAAAAACCAACACACAGATC TAATGAAAATAAAGATCTTTTAT TGGAGCTCCAAGGATAGAAACA CAAACGCTTTCCCACTGGC
Pvt1b-TWI	Pvt1 exon 1b	CGAGGTTTTTCCCAGTCATGGAG TCCATGTGACGTTTCTGATTTCCT GACAGCTTTGCTCCTTCTAAATCT TTTAAGAAGTCCCACTTCAATAAA GCTGCATTAATGCCGCGCTATCGC GACATTACTCTGCTATTTTTGCGG GCTTGTAACCGCTTTATTGGGAG CTCCAAGGATAGAAACACAAACG CTTTCCCACTG
smRNA FISH Probes		
Name	**3' Modification**	**Sequence**
PVT1_exons_1	Quasar 570	TCTGGGAATGCTAAGTTCGT
PVT1_exons_2	Quasar 570	CCATGTGACGTTTCTGATTT
PVT1_exons_3	Quasar 570	ACACATCCAAGCACTGAGAG
PVT1_exons_4	Quasar 570	AAATCAGACCTCCGAGATGT
PVT1_exons_5	Quasar 570	TTCAGGAAGTCTCCAGAGAG
PVT1_exons_6	Quasar 570	CAGAATTACTCCCCAGGAAA
PVT1_exons_7	Quasar 570	GGGTAGAGATACAATCCTCT
PVT1_exons_8	Quasar 570	GCTCTCAGAAACACTGCATT
PVT1_exons_9	Quasar 570	CTGGTTCTTCTGAGAGACTG
PVT1_exons_10	Quasar 570	AGGCATCTCACAGCAAAGTA
PVT1_exons_11	Quasar 570	TTATCACATTAGAGGACCCG
PVT1_exons_12	Quasar 570	ACTTGGCATCTCTTAAGTCA
PVT1_exons_13	Quasar 570	AGACTTCCATCTTTGCTATT
PVT1_exons_14	Quasar 570	CAGCTGTCTTATAGGATTGC
PVT1_exons_15	Quasar 570	TCTTAGGGTCAGTATCATGG
PVT1_exons_16	Quasar 570	AGTATTCTAGCTTGGAGCTA
PVT1_exons_17	Quasar 570	TTGTCACTCCATTTGGCAAA
PVT1_exons_18	Quasar 570	TGCTTAAAGACCACAGAGGC
PVT1_exons_19	Quasar 570	ATTGCTTTGGGTATTTTGGT
PVT1_exons_20	Quasar 570	AATGTCTACTTGTTGGCCAA
PVT1_exons_21	Quasar 570	TAGCAGAGTGGTTCAAAGGC
PVT1_exons_22	Quasar 570	AGAATTTCAGAGGGCACTCG
PVT1_exons_23	Quasar 570	AGACTTAGGGCATACAGGTA
PVT1_exons_24	Quasar 570	CTGGATTCTGTAGCTATTCT
PVT1_exons_25	Quasar 570	TAAAGCATCCAGGGCAGAAC
PVT1_exons_26	Quasar 570	TGACTCCTGTTGGAAAACCA
PVT1_exons_27	Quasar 570	CACGCTCATGTCCTTTAATA
PVT1_exons_28	Quasar 570	GTTTAGCACTATCCATCTTT
PVT1_exons_29	Quasar 570	TTGCTCTCCTTATGAAGAGG
PVT1_exons_30	Quasar 570	GAAACCTTAAGCATGAGCCA
PVT1_exons_31	Quasar 570	AGTGCACTCTTATACGTCAC
PVT1_exons_32	Quasar 570	ATCTTAAGATGGCTTGGACC

PVT1_exons_33	Quasar 570	GAGATTCGGAACTGACAGGC
PVT1_exons_34	Quasar 570	AAAGAGAACGTGTCCCTTGT
PVT1_exons_35	Quasar 570	TCTATTGTAGGTTGTTCCTG
PVT1_exons_36	Quasar 570	AAATCCAGGCTACTTCTCAG
PVT1_exons_37	Quasar 570	GCCTCCAGAGAAAACGATGA
PVT1_exons_38	Quasar 570	CAGGGCTCATGAGAACAGAG
PVT1_exons_39	Quasar 570	CTTACCAGGAGAAGCATCAT
PVT1_exons_40	Quasar 570	CAGCACATAGAACACAGGCA
PVT1_exons_41	Quasar 570	GAAGATTGTGCCAGGAACTC
PVT1_exons_42	Quasar 570	CAGATCCTGGTTTAGAACGG
PVT1_exons_43	Quasar 570	CTGTCATCTTCTCTTCTTTG
PVT1_exons_44	Quasar 570	TCCTTAATGTGCTACCACAA
PVT1_exons_45	Quasar 570	GGATTCTACTTCACCATAGG
PVT1_exons_46	Quasar 570	CCAAGGCATTATGAAGTGCA
PVT1_exons_47	Quasar 570	AAAGTGTCTCAGGGAATCCT
PVT1_exons_48	Quasar 570	TCAGTAAGTCACAGCTGTGA
PVT1_intron_AE1up_1	Quasar 670	TTCCAGGGGATAAACTTGGA
PVT1_intron_AE1up_2	Quasar 670	AATGCAAAAGCCACTTTCCT
PVT1_intron_AE1up_3	Quasar 670	AAGGTTAACACGCGCTCGTG
PVT1_intron_AE1up_4	Quasar 670	TCGAGTCTAGTGATGAGGAA
PVT1_intron_AE1up_5	Quasar 670	ACACCCAAACTCTCTGGCAA
PVT1_intron_AE1up_6	Quasar 670	TAGAGGCCATCCTGGGAAAT
PVT1_intron_AE1up_7	Quasar 670	GCATAAATCCAGAATTACCT
PVT1_intron_AE1up_8	Quasar 670	CTGAGGAAATGGGCTCTTGA
PVT1_intron_AE1up_9	Quasar 670	CAAATCTGCGCTGATTGCAG
PVT1_intron_AE1up_10	Quasar 670	TCGTAAATGAGGCCTCCAAA
PVT1_intron_AE1up_11	Quasar 670	GACTAGACTCAGACTTCCAG
PVT1_intron_AE1up_12	Quasar 670	AAGGATGGAGGGAGCATCAC
PVT1_intron_AE1up_13	Quasar 670	GTTTTAGGAGATCACCTTCT
PVT1_intron_AE1up_14	Quasar 670	GCACAGAAAGTTTCCTGACA
PVT1_intron_AE1up_15	Quasar 670	CTTCCACGAACACAGGAACG
PVT1_intron_AE1up_16	Quasar 670	TAGCAAGGATGAAGGCGTGG
PVT1_intron_AE1up_17	Quasar 670	GTTAAAGCAACAAGCTATCC
PVT1_intron_AE1up_18	Quasar 670	TAGCCAAGAAAGGGCCAATC
PVT1_intron_AE1up_19	Quasar 670	CACTCATAGGTACAGCAGAA
PVT1_intron_AE1up_20	Quasar 670	TGGAAGTCTGCACAGTTCTC
PVT1_intron_AE1up_21	Quasar 670	CACATGTAGCTTCATGGCTG
PVT1_intron_AE1up_22	Quasar 670	ATCCATGATGTGTCTACACA
PVT1_intron_AE1up_23	Quasar 670	CAGATTATCACCCACTAGTA
PVT1_intron_AE1up_24	Quasar 670	GAACGTTCTGGAGAGCTCAA
PVT1_intron_AE1up_25	Quasar 670	GATTTCTCTCCTTAAGCTTC
PVT1_intron_AE1up_26	Quasar 670	TTCTCCCTATACTCTCTTAA
PVT1_intron_AE1up_27	Quasar 670	GTCAAATGACAACAACCCCT
PVT1_intron_AE1up_28	Quasar 670	AACAGAGACCTGCATCCTTA
PVT1_intron_AE1up_29	Quasar 670	CAACATCCTACCACATGCAC
PVT1_intron_AE1up_30	Quasar 670	GTGAACAAGCCCAAACTTGT
PVT1_intron_AE1up_31	Quasar 670	CTAGCTATGACCATAGGACT
PVT1_Exon1a_1	Quasar 670	GTGGCCCGTGACGTCACG
PVT1_Exon1a_2	Quasar 670	TGGTAGAGCGCGGGGCTG
PVT1_Exon1a_3	Quasar 670	CGGCCACACGCGCTCTGC
PVT1_Exon1a_4	Quasar 670	AGTGGGTCCATGTGCTCG
PVT1_Exon1a_5	Quasar 670	TTGCCGAGCCCGCTTGAC
PVT1_Exon1a_6	Quasar 670	CGTAGCTTCTCTTCATCC
PVT1_Exon1a_7	Quasar 670	GCTCTGGGAATGCTAAGT

PVT1_Exon1a_8	Quasar 670	GCGGATATCCACTCGGGG
PVT1_Exon1a_9	Quasar 670	CTGCCAACATCCTTTCCA
PVT1_Exon1a_10	Quasar 670	GACTCCGAGGTCACCGGG
PVT1_Exon1a_11	Quasar 670	AGAGGGTGGATCCAGCCG
PVT1_Exon1b_1	Quasar 670	GGCTAGAGCTTTAAGAAG
PVT1_Exon1b_2	Quasar 670	ACACAAACGCTTTCCCAC
PVT1_Exon1b_3	Quasar 670	TGGAGCTCCAAGGATAGA
PVT1_Exon1b_4	Quasar 670	TCTTTTAAGAAGTCCCAC
PVT1_Exon1b_5	Quasar 670	CAGCTTTGCTCCTTCTAA
PVT1_Exon1b_6	Quasar 670	GACGTTTCTGATTTCCTG
PVT1_Exon1b_7	Quasar 670	CCAGTCATGGAGTCCATG
PVT1_Exon1b_8	Quasar 670	GGCCACCACGAGGTTTTT
PVT1_Exon1b_9	Quasar 670	CATCCAAGCACTGAGAGC
PVT1_Exon1b_10	Quasar 670	TCCGAGATGTAACTAGCA
c-Myc intron_1	Quasar 570	AAAGACCACCAGATCTGTGC
c-Myc intron_2	Quasar 570	TAACCGGCCGCTACATTCAA
c-Myc intron_3	Quasar 570	CCCCAACTACTCTTGAGAAA
c-Myc intron_4	Quasar 570	CATCTTGACAAGTCGCTCTA
c-Myc intron_5	Quasar 570	CGCTTCAAAATGCATCCCGG
c-Myc intron_6	Quasar 570	CCCATAGTAACCTCGGGAAC
c-Myc intron_7	Quasar 570	AAGCAAGAATGTCCAACCGG
c-Myc intron_8	Quasar 570	CCCTCAAAGGACACATATCA
c-Myc intron_9	Quasar 570	GATTCCAAGGGCTTTCTTTG
c-Myc intron_10	Quasar 570	TAATCCCTTCTCCAAAGACC
c-Myc intron_11	Quasar 570	TCTCGCTCCCAAACGCAAAA
c-Myc intron_12	Quasar 570	GGTAAGTCAGAAGCTACGGA
c-Myc intron_13	Quasar 570	TTTAAATGCCCTCTCAGAGA
c-Myc intron_14	Quasar 570	GTCAGAAATGCACCAAGCTG
c-Myc intron_15	Quasar 570	TTAAAAGGCTCAGGGACGGG
c-Myc intron_16	Quasar 570	GGGGGTCAGGCTTAAATTTT
c-Myc intron_17	Quasar 570	CCAACATCAAGTCCTAGTGC
c-Myc intron_18	Quasar 570	AATTTTGCTTCTCCTCACTG
c-Myc intron_19	Quasar 570	TCAACGAATCGGTCACATCC
c-Myc intron_20	Quasar 570	CAGTCTTCCTAGCAATTCAG
c-Myc intron_21	Quasar 570	TTACGGAACCGCTCAGATCA
c-Myc intron_22	Quasar 570	TACACTCTAAACCGCGACGC
c-Myc intron_23	Quasar 570	ATAATAAGAGACACCTCCCT
c-Myc intron_24	Quasar 570	GCTATCACAAGCCTCTCGAA
c-Myc intron_25	Quasar 570	TGGAGGAGAGAGCTCAGTCT
c-Myc intron_26	Quasar 570	CTTTTCTTTCCGATTGCTGA
c-Myc intron_27	Quasar 570	AAGGAGAAAGGCGAGAGGCG
c-Myc intron_28	Quasar 570	CTAAGAGCCGAGGCGCAAAG
c-Myc intron_29	Quasar 570	GAGGCGACTGTAGGGAATAC
c-Myc intron_30	Quasar 570	TCCTTCGAGCAGGGACTTAG
c-Myc intron_31	Quasar 570	TACTATCAGTGACGCTCGTC
c-Myc intron_32	Quasar 570	AGGCATGCACTCTTTTACTC
c-Myc intron_33	Quasar 570	GAGTTATCCAGCTCTGGTTG

Supplementary Table 2: qRT-PCR and PCR primer sequences

qRT-PCR primers (mouse)		
Gene	**Forward sequence**	**Reverse sequence**
Pvt1a (ex1a-2)	ACTTAGCATTCCCAGAGCCC	TGGAGGGCATCTTCTTACCG
Pvt1b (ex1b-2)	CCATGACTGGGAAAAACCTCG	TGGAGGGCATCTTCTTACCG
Pvt1b nascent (ex1b-int)	CCATGACTGGGAAAAACCTCG	CCAGCACAATAGCCCACAATG
Pvt1 (ex4-5)	CTGGGACACTGCCTGATTGA	TCCTTCTGGAACGCTTAAAGG
Myc	TTCATCTGCGATCCTGACGAC	CACTGAGGGGTCAATGCACTC
p21	TCCACAGCGATATCCAGACA	GGACATCACCAGGATTGGAC
Rn7s1	CTGTAGTGCGCTATGCCGA	GTTCACCCCTCCTTAGGCAA
Kcnq1ot1	GGCCAGAAGCAGAGGTGATT	CCGAGCCGTAACTGCAAAAC
Gm26542	CCTTGGCTGACACCCGAACC	CCGAGTTCGAGCGCGTCTTC
Gapdh	AGCTTGTCATCAACGGGAAG	TTTGATGTTAGTGGGGTCTCG
Pvt1 p53RE (ChIP)	GGCTAAGGATGCAGGTCTC	AAACGCTTTCCCACTGGCTA
Myc promoter (ChIP)	CGCGAGCAAGAGAAAATGGTC	CTTTGGGAACTCGGGAGGG
qRT-PCR primers (human)		
Gene	**Forward sequence**	**Reverse sequence**
PVT1a (ex1a-2)	TTCCAGTGGATTTCCTTGCGG	CTGACAGGCACAGCCATCTTG
PVT1b (ex1b-2)	GCACAAGGGCCCAACTGGA	CTGACAGGCACAGCCATCTTG
MYC	TCGGATTCTCTGCTCTCCTCG	AGGTGATCCAGACTCTGACCT
3C qRT-PCR primers		
Name	**Sequence**	
BamHI-016	GTGCTCAGCTCCATCCTGCAAG GACAATTC	
BamHI-015	GAGGGAACAAAATACTCATGG GAGGAGATAC	
BamHI-008	TTGAAAAAGCCTGCTAGACAGT CCCTGGTG	
BamHI-007	TGCACTTCCTTGAGAAGCTGGT AGGATACC	
BamHI001	CACAGGAGGAACATCAGGAGA CCCAAATTC	
BamHI002	TCAGCTGCCGGGTCCGACTCG CCTCAC	
BamHI003	GCAGTGAGGAGAAGCAAAATT GGGACAGGG	
BamHI008	TGAGTGACACCAACATCCTGGA GCCTCAG	
BamHI009	AACCAGGCTACTCTAACTCTCT CTGCTCAG	
BamHI013	CTCCTCTCAGGCTTGATGCCCC TTCATTG	
BamHI014	ACGAACTTAGCATTCCCAGAGC CCCCCG	
BamHI015	TGGAAAGGTGTCTCTTTTCTGT TGTTTCTGG	
BamHI016	GGCACCGACTGGGCATGTATC CTGCTTG	
BamHI017	GTCTCAGTGCAGCAGCCCTTGA GTGAAAG	

BamHI018	TGCATTGCCACATTCCAGATTG TCACCTTG	
BamHI019	CAGTAGCAGAGAGCATAAGCC TTTGTCTCC	
BamHI020	GAGGTATGAATGTAAACATTG TACACATACTG	
BamHI022	GTTAATTGGGTGTTCTAGCTCT GGAAAATGC	
BamHI024	CTATTTTGCCCCTTTGTTCCCT GTTCTATCC	
BamHI028	TGTAGATCTCAACAGATGAACC CAGGGGAC	
BamHI029	GCAACACTTGATGACTTGACCA AATAAACAGC	
BamHI030	AGTTCCAAAGGTGAAAAAGCT GTATAATCGTC	
BamHI031	CTGCCTCTCAGCTCACGGCCAC TGTGTC	
BamHI032	CTTGAGAGCCTGCATATCCTTT GAGCAGAT	
BamHI033	CCCCAATCCTTTTCTCTACTCC ATACCCAC	
BamHI036	AGGCAGGGCTGGAGTTTTGTT CTGTTTGTG	
BamHI042	AATTGAGAAACCACCCGATAGT AACCTGGG	
BamHI050	CACAAGAGACAGCTACATCTG GGTCCTTTC	
Pvt1 3'E.1 BamHI003	CCTGTCTCCTCCCCCATCCTGA TAGTAC	
Pvt1 3'E.1 BamHI004	TGGTATGAGTATCCAAAGACAT TGAGGACTC	
Pvt1 3'E.2 BamHI003	ACTAGAGTATGTCTGCCTTTTG TGTGGGAC	
Pvt1 3'E.2 BamHI005	TCTTTGAGTTCATTTGTAAGGG TATTTCCAGC	
BamHIctrl	ATTAAAGGTGGAGTGAGACAT CAGAGGTGG	
RT-PCR primers		
Name	**Forward sequence**	**Reverse sequence**
Pvt1 (ex1a-5)	TGGATATCCGCGTGGAAAGG	TCCTTCTGGAACGCTTAAAGG
Pvt1 (ex1b-5)	CTCTAGCCAGTGGGAAAGCG	TCCTTCTGGAACGCTTAAAGG
Genotyping PCR primers		
Name	**Forward sequence**	**Reverse sequence**
Pvt1a PAS	TACCAGGCAGAGCGCGTG	CTGGGCTCCAGAGTTTCCA
Pvt1b PAS/TWI	ACTTGCACAGTCCTATGGTCA	CGTAAGGCACATCCTCACCT
Mutagenesis efficiency PCR primers		
Name	**Forward sequence**	**Reverse sequence**
Pvt1 p53RE	GAAGTGCATGTGGTAGGATG	GCACATCCTCACCTCCGAGA
Pvt1 exon 1b	GCCTGTTTTGCATATGGGCAG	ACAAGGCAGTCCCATACAGTC
Pvt1 exon 1b 3' end	CAGTGGGAAAGCGTTTGTG	AGCAAGAAACAGCCACCCTT

Supplementary Table 3: Key plasmids and recombinant DNA used in this work

Name	Source	Identifier
pCMV-dR8.2 dvpr	(Stewart et al., 2003)	Addgene #8455
pCMV-VSV-G	(Stewart et al., 2003)	Addgene #8454
pWZL Hygro	S. Lowe, unpublished	Addgene #18750
BRD001	Broad Institute	N/A
BRD004	Broad Institute	N/A
UGPC	(Olivero et al., 2020)	N/A
lenti-SAM-hygro	(Olivero et al., 2020)	N/A
Myc BAC	BACPAC Resources Center	Cat#RP23-55F11
Chr 6 BAC	BACPAC Resources Center	Cat#RP24-301E22

Supplementary Table 4: Software and algorithms used in this work

Resource	Source	Location
GraphPad Prism, version 8.2.1 for MacOS	N/A	www.graphpad.com
FIJI	N/A	https://imagej.net/Downloads
TIDE	(Brinkman et al., 2014)	http://shinyapps.datacurators.nl/tide/
Biorender	N/A	www.biorender.com
Tophat (v2.0.14)	(Trapnell et al., 2009)	http://ccb.jhu.edu/software/tophat/index.shtml
stringtie (v1.2.4)	(Pertea et al., 2015)	https://ccb.jhu.edu/software/stringtie/
HTSeq (v0.6.1)	(Anders et al., 2015)	https://htseq.readthedocs.io/en/release_0.11.1/
EdgeR (v3.22.3)	(Robinson et al., 2010)	https://bioconductor.org/packages/release/bioc/html/edgeR.html
FastUniq	(Xu et al., 2012)	http://sourceforge.net/projects/fastuniq/
Cutadapt (v1.16)	(Martin, 2011)	https://cutadapt.readthedocs.io/en/stable/
HISAT2 (v2.1.0)	(Kim et al., 2015a)	https://ccb.jhu.edu/software/hisat2/index.shtml
SAMtools (v1.5)	(Li et al., 2009)	http://samtools.sourceforge.net
STAR (v2.5.3a)	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
DEseq2	(Love et al., 2014)	https://bioconductor.org/packages/release/bioc/html/DESeq2.html

References

Adams, J.M., Harris, A.W., Pinkert, C.A., Corcoran, L.M., Alexander, W.S., Cory, S., Palmiter, R.D., and Brinster, R.L. (1985). The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature *318*, 533-538.

Adriaens, C., Standaert, L., Barra, J., Latil, M., Verfaillie, A., Kalev, P., Boeckx, B., Wijnhoven, P.W., Radaelli, E., Vermi, W., *et al.* (2016). p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. Nat Med *22*, 861-868.

Alitalo, K., Schwab, M., Lin, C.C., Varmus, H.E., and Bishop, J.M. (1983). Homogeneously staining chromosomal regions contain amplified copies of an abundantly expressed cellular oncogene (c-myc) in malignant neuroendocrine cells from a human colon carcinoma. Proc Natl Acad Sci U S A *80*, 1707-1711.

Allen, M.A., Andrysik, Z., Dengler, V.L., Mellert, H.S., Guarnieri, A., Freeman, J.A., Sullivan, K.D., Galbraith, M.D., Luo, X., Kraus, W.L., *et al.* (2014). Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms. Elife *3*, e02200.

Amandio, A.R., Necsulea, A., Joye, E., Mascrez, B., and Duboule, D. (2016). Hotair Is Dispensible for Mouse Development. PLoS Genet *12*, e1006232.

Amati, B., Brooks, M.W., Levy, N., Littlewood, T.D., Evan, G.I., and Land, H. (1993). Oncogenic activity of the c-Myc protein requires dimerization with Max. Cell *72*, 233-245.

Amodio, N., Raimondi, L., Juli, G., Stamato, M.A., Caracciolo, D., Tagliaferri, P., and Tassone, P. (2018). MALAT1: a druggable long non-coding RNA for targeted anti-cancer approaches. J Hematol Oncol *11*, 63.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Ariyoshi, M., and Schwabe, J.W. (2003). A conserved structural motif reveals the essential transcriptional repression function of Spen proteins and their role in developmental signaling. Genes Dev *17*, 1909-1920.

Arun, G., Diermeier, S., Akerman, M., Chang, K.C., Wilkinson, J.E., Hearn, S., Kim, Y., MacLeod, A.R., Krainer, A.R., Norton, L., *et al.* (2016). Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. Genes Dev *30*, 34-51.

Arun, G., Diermeier, S.D., and Spector, D.L. (2018). Therapeutic Targeting of Long Non-Coding RNAs in Cancer. Trends Mol Med *24*, 257-277.

Arun, G., and Spector, D.L. (2019). MALAT1 long non-coding RNA and breast cancer. RNA Biol *16*, 860-863.

Attardi, L.D., de Vries, A., and Jacks, T. (2004). Activation of the p53-dependent G1 checkpoint response in mouse embryo fibroblasts depends on the specific DNA damage inducer. Oncogene *23*, 973-980.

Barsotti, A.M., Beckerman, R., Laptenko, O., Huppi, K., Caplen, N.J., and Prives, C. (2012). p53-Dependent induction of PVT1 and miR-1204. J Biol Chem *287*, 2509-2519.

Bassett, A.R., Akhtar, A., Barlow, D.P., Bird, A.P., Brockdorff, N., Duboule, D., Ephrussi, A., Ferguson-Smith, A.C., Gingeras, T.R., Haerty, W.*, et al.* (2014). Considerations when investigating lncRNA function in vivo. Elife *3*, e03058.

Beckerman, R., and Prives, C. (2010). Transcriptional regulation by p53. Cold Spring Harb Perspect Biol *2*, a000935.

Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M.*, et al.* (2010). The landscape of somatic copy-number alteration across human cancers. Nature *463*, 899-905.

Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S.*, et al.* (2004). Global identification of human transcribed sequences with genome tiling arrays. Science *306*, 2242-2246.

Bertucci, F., Lagarde, A., Ferrari, A., Finetti, P., Charafe-Jauffret, E., Van Laere, S., Adelaide, J., Viens, P., Thomas, G., Birnbaum, D.*, et al.* (2012). 8q24 Cancer risk allele associated with major metastatic risk in inflammatory breast cancer. PLoS One *7*, e37943.

Bester, A.C., Lee, J.D., Chavez, A., Lee, Y.R., Nachmani, D., Vora, S., Victor, J., Sauvageau, M., Monteleone, E., Rinn, J.L.*, et al.* (2018). An Integrated Genome-wide CRISPRa Approach to Functionalize lncRNAs in Drug Resistance. Cell *173*, 649-664 e620.

Blackwell, T.K., Kretzner, L., Blackwood, E.M., Eisenman, R.N., and Weintraub, H. (1990). Sequence-specific DNA binding by the c-Myc protein. Science *250*, 1149-1151.

Blume, C.J., Hotz-Wagenblatt, A., Hullein, J., Sellner, L., Jethwa, A., Stolz, T., Slabicki, M., Lee, K., Sharathchandra, A., Benner, A.*, et al.* (2015). p53-

dependent non-coding RNA networks in chronic lymphocytic leukemia. Leukemia *29*, 2015-2023.

Botti, G., Collina, F., Scognamiglio, G., Aquino, G., Cerrone, M., Liguori, G., Gigantino, V., Malzone, M.G., and Cantile, M. (2018). LncRNA HOTAIR Polymorphisms Association with Cancer Susceptibility in Different Tumor Types. Curr Drug Targets *19*, 1220-1226.

Bradner, J.E., Hnisz, D., and Young, R.A. (2017). Transcriptional Addiction in Cancer. Cell *168*, 629-643.

Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res *42*, e168.

Brockdorff, N., Bowness, J.S., and Wei, G. (2020). Progress toward understanding chromosome silencing by Xist RNA. Genes Dev *34*, 733-744.

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature *349*, 38-44.

Brugarolas, J., Chandrasekaran, C., Gordon, J.I., Beach, D., Jacks, T., and Hannon, G.J. (1995). Radiation-induced cell cycle arrest compromised by p21 deficiency. Nature *377*, 552-557.

Cabili, M.N., Dunagin, M.C., McClanahan, P.D., Biaesch, A., Padovan-Merhar, O., Regev, A., Rinn, J.L., and Raj, A. (2015). Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol *16*, 20.

Cai, L., Chang, H., Fang, Y., and Li, G. (2016). A Comprehensive Characterization of the Function of LincRNAs in Transcriptional Regulation Through Long-Range Chromatin Interactions. Sci Rep *6*, 36572.

Camblong, J., Beyrouthy, N., Guffanti, E., Schlaepfer, G., Steinmetz, L.M., and Stutz, F. (2009). Trans-acting antisense RNAs mediate transcriptional gene cosuppression in S. cerevisiae. Genes Dev *23*, 1534-1545.

Cao, L., Zhang, P., Li, J., and Wu, M. (2017). LAST, a c-Myc-inducible long noncoding RNA, cooperates with CNBP to promote CCND1 mRNA stability in human cells. Elife *6*.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., *et al.* (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559-1563.

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., *et al.* (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature *491*, 454-457.

Carter, A.C., Xu, J., Nakamoto, M.Y., Wei, Y., Zarnegar, B.J., Shi, Q., Broughton, J.P., Ransom, R.C., Salhotra, A., Nagaraja, S.D., *et al.* (2020). Spen links RNA-mediated endogenous retrovirus silencing and X chromosome inactivation. Elife *9*.

Cech, T.R., and Steitz, J.A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. Cell *157*, 77-94.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., *et al.* (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov *2*, 401-404.

Chakravarty, D., Sboner, A., Nair, S.S., Giannopoulou, E., Li, R., Hennig, S., Mosquera, J.M., Pauwels, J., Park, K., Kossai, M., *et al.* (2014). The oestrogen receptor alpha-regulated lncRNA NEAT1 is a critical modulator of prostate cancer. Nat Commun *5*, 5383.

Chaudhary, R., Gryder, B., Woods, W.S., Subramanian, M., Jones, M.F., Li, X.L., Jenkins, L.M., Shabalina, S.A., Mo, M., Dasso, M., *et al.* (2017). Prosurvival long noncoding RNA PINCR regulates a subset of p53 targets in human colorectal cancer cells by binding to Matrin 3. Elife *6*.

Chaumeil, J., Augui, S., Chow, J.C., and Heard, E. (2008). Combined immunofluorescence, RNA fluorescent in situ hybridization, and DNA fluorescent in situ hybridization to study chromatin changes, transcriptional activity, nuclear organization, and X-chromosome inactivation. Methods Mol Biol *463*, 297-308.

Chen, X., Lun, L., Hou, H., Tian, R., Zhang, H., and Zhang, Y. (2017). The Value of lncRNA HULC as a Prognostic Factor for Survival of Cancer Outcome: A Meta-Analysis. Cell Physiol Biochem *41*, 1424-1434.

Cho, S.W., Xu, J., Sun, R., Mumbach, M.R., Carter, A.C., Chen, Y.G., Yost, K.E., Kim, J., He, J., Nevins, S.A., *et al.* (2018). Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element. Cell *173*, 1398-1412 e1322.

Christoffersen, N.R., Shalgi, R., Frankel, L.B., Leucci, E., Lees, M., Klausen, M., Pilpel, Y., Nielsen, F.C., Oren, M., and Lund, A.H. (2010). p53-independent upregulation of miR-34a during oncogene-induced senescence represses MYC. Cell Death Differ *17*, 236-245.

Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., and Chang, H.Y. (2015). Systematic discovery of Xist RNA binding proteins. Cell *161*, 404-416.

Clark, M.B., Johnston, R.L., Inostroza-Ponta, M., Fox, A.H., Fortini, E., Moscato, P., Dinger, M.E., and Mattick, J.S. (2012). Genome-wide analysis of long noncoding RNA stability. Genome Res *22*, 885-898.

Clarke, A.R., Gledhill, S., Hooper, M.L., Bird, C.C., and Wyllie, A.H. (1994). p53 dependence of early apoptotic and proliferative responses within the mouse intestinal epithelium following gamma-irradiation. Oncogene *9*, 1767-1773.

Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H., Chess, A., and Lawrence, J.B. (2009). An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol Cell *33*, 717-726.

Collins, S., and Groudine, M. (1982). Amplification of endogenous myc-related DNA sequences in a human myeloid leukaemia cell line. Nature *298*, 679-681.

Colognori, D., Sunwoo, H., Kriz, A.J., Wang, C.Y., and Lee, J.T. (2019). Xist Deletional Analysis Reveals an Interdependency between Xist RNA and Polycomb Complexes for Spreading along the Inactive X. Mol Cell *74*, 101-117 e110.

Colombo, T., Farina, L., Macino, G., and Paci, P. (2015). PVT1: a rising star among oncogenic long noncoding RNAs. Biomed Res Int *2015*, 304208.

Concordet, J.P., and Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res *46*, W242-W245.

Conrad, T., and Orom, U.A. (2017). Cellular Fractionation and Isolation of Chromatin-Associated RNA. Methods Mol Biol *1468*, 1-9.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845-1848.

Cory, S., Graham, M., Webb, E., Corcoran, L., and Adams, J.M. (1985). Variant (6;15) translocations in murine plasmacytomas involve a chromosome 15 locus at least 72 kb from the c-myc oncogene. EMBO J *4*, 675-681.

Cui, M., You, L., Ren, X., Zhao, W., Liao, Q., and Zhao, Y. (2016). Long non-coding RNA PVT1 and cancer. Biochem Biophys Res Commun *471*, 10-14.

Cunnington, M.S., Santibanez Koref, M., Mayosi, B.M., Burn, J., and Keavney, B. (2010). Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression. PLoS Genet *6*, e1000899.

Dahlman, J.E., Abudayyeh, O.O., Joung, J., Gootenberg, J.S., Zhang, F., and Konermann, S. (2015). Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. Nat Biotechnol *33*, 1159-1161.

Dang, C.V. (2012). MYC on the path to cancer. Cell *149*, 22-35.

Dang, C.V., Reddy, E.P., Shokat, K.M., and Soucek, L. (2017). Drugging the 'undruggable' cancer targets. Nat Rev Cancer *17*, 502-508.

Davidovich, C., Zheng, L., Goodrich, K.J., and Cech, T.R. (2013). Promiscuous RNA binding by Polycomb repressive complex 2. Nat Struct Mol Biol *20*, 1250-1257.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G.*, et al.* (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res *22*, 1775-1789.

Diermeier, S.D., Chang, K.C., Freier, S.M., Song, J., El Demerdash, O., Krasnitz, A., Rigo, F., Bennett, C.F., and Spector, D.L. (2016). Mammary Tumor-Associated RNAs Impact Tumor Cell Proliferation, Invasion, and Migration. Cell Rep *17*, 261-274.

Dimitrova, N., Zamudio, J.R., Jong, R.M., Soukup, D., Resnick, R., Sarma, K., Ward, A.J., Raj, A., Lee, J.T., Sharp, P.A.*, et al.* (2014). LincRNA-p21 activates p21 in cis to promote Polycomb target gene expression and to enforce the G1/S checkpoint. Mol Cell *54*, 777-790.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F.*, et al.* (2012). Landscape of transcription in human cells. Nature *489*, 101-108.

Doamekpor, S.K., Gozdek, A., Kwasnik, A., Kufel, J., and Tong, L. (2020). A novel 5'-hydroxyl dinucleotide hydrolase activity for the DXO/Rai1 family of enzymes. Nucleic Acids Res *48*, 349-358.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Jr., Butel, J.S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature *356*, 215-221.

Dong, X., Gao, W., Lv, X., Wang, Y., Wu, Q., Yang, Z., Mao, G., and Xing, W. (2020). Association between lncRNA GAS5, MEG3, and PCAT-1 Polymorphisms and Cancer Risk: A Meta-Analysis. Dis Markers *2020*, 6723487.

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

Easton, D.F., and Eeles, R.A. (2008). Genome-wide association studies in cancer. Hum Mol Genet *17*, R109-115.

Eissmann, M., Gutschner, T., Hammerle, M., Gunther, S., Caudron-Herger, M., Gross, M., Schirmacher, P., Rippe, K., Braun, T., Zornig, M.*, et al.* (2012). Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. RNA Biol *9*, 1076-1087.

el-Deiry, W.S., Kern, S.E., Pietenpol, J.A., Kinzler, K.W., and Vogelstein, B. (1992). Definition of a consensus binding site for p53. Nat Genet *1*, 45-49.

Elling, R., Robinson, E.K., Shapleigh, B., Liapis, S.C., Covarrubias, S., Katzman, S., Groff, A.F., Jiang, Z., Agarwal, S., Motwani, M.*, et al.* (2018). Genetic Models Reveal cis and trans Immune-Regulatory Activities for lincRNA-Cox2. Cell Rep *25*, 1511-1524 e1516.

Engeland, K. (2018). Cell cycle arrest through indirect transcriptional repression by p53: I have a DREAM. Cell Death Differ *25*, 114-132.

Engreitz, J.M., Haines, J.E., Perez, E.M., Munson, G., Chen, J., Kane, M., McDonel, P.E., Guttman, M., and Lander, E.S. (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. Nature *539*, 452-455.

Fang, H., Bonora, G., Lewandowski, J.P., Thakur, J., Filippova, G.N., Henikoff, S., Shendure, J., Duan, Z., Rinn, J.L., Deng, X.*, et al.* (2020). Trans- and cis-acting effects of Firre on epigenetic features of the inactive X chromosome. Nat Commun *11*, 6053.

Fang, S., Zhang, L., Guo, J., Niu, Y., Wu, Y., Li, H., Zhao, L., Li, X., Teng, X., Sun, X.*, et al.* (2018). NONCODEV5: a comprehensive annotation database for long non-coding RNAs. Nucleic Acids Res *46*, D308-D314.

Feldser, D.M., Kostova, K.K., Winslow, M.M., Taylor, S.E., Cashman, C., Whittaker, C.A., Sanchez-Rivera, F.J., Resnick, R., Bronson, R., Hemann, M.T.*, et al.* (2010). Stage-specific sensitivity to p53 restoration during lung cancer progression. Nature *468*, 572-575.

Ferbeyre, G., de Stanchina, E., Lin, A.W., Querido, E., McCurrach, M.E., Hannon, G.J., and Lowe, S.W. (2002). Oncogenic ras and p53 cooperate to induce cellular senescence. Mol Cell Biol *22*, 3497-3508.

Flockhart, R.J., Webster, D.E., Qu, K., Mascarenhas, N., Kovalski, J., Kretz, M., and Khavari, P.A. (2012). BRAFV600E remodels the melanocyte transcriptome and induces BANCR to regulate melanoma cell migration. Genome Res *22*, 1006-1014.

Fradet, Y., Saad, F., Aprikian, A., Dessureault, J., Elhilali, M., Trudel, C., Masse, B., Piche, L., and Chypre, C. (2004). uPM3, a new molecular urine test for the detection of prostate cancer. Urology *64*, 311-315; discussion 315-316.

Freedman, M.L., Monteiro, A.N., Gayther, S.A., Coetzee, G.A., Risch, A., Plass, C., Casey, G., De Biasi, M., Carlson, C., Duggan, D.*, et al.* (2011). Principles for the post-GWAS functional characterization of cancer risk loci. Nat Genet *43*, 513-518.

Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M., Cleary, B., Lander, E.S., and Engreitz, J.M. (2016). Systematic mapping of functional enhancer-promoter connections with CRISPR interference. Science *354*, 769-773.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E.*, et al.* (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal *6*, pl1.

Gao, P., Xia, J.H., Sipeky, C., Dong, X.M., Zhang, Q., Yang, Y., Zhang, P., Cruz, S.P., Zhang, K., Zhu, J.*, et al.* (2018). Biology and Clinical Implications of the 19q13 Aggressive Prostate Cancer Susceptibility Locus. Cell *174*, 576-589 e518.

Garcia-Closas, M., Hall, P., Nevanlinna, H., Pooley, K., Morrison, J., Richesson, D.A., Bojesen, S.E., Nordestgaard, B.G., Axelsson, C.K., Arias, J.I.*, et al.* (2008). Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. PLoS Genet *4*, e1000054.

Garcia-Jove Navarro, M., Kashida, S., Chouaib, R., Souquere, S., Pierron, G., Weil, D., and Gueroui, Z. (2019). RNA is a critical element for the sizing and the composition of phase-separated RNA-protein condensates. Nat Commun *10*, 3230.

Ge, Y., He, Y., Jiang, M., Luo, D., Huan, X., Wang, W., Zhang, D., Yang, L., and Zhou, J. (2017). Polymorphisms in lncRNA PTENP1 and the Risk of Gastric Cancer in a Chinese Population. Dis Markers *2017*, 6807452.

Ghoussaini, M., Song, H., Koessler, T., Al Olama, A.A., Kote-Jarai, Z., Driver, K.E., Pooley, K.A., Ramus, S.J., Kjaer, S.K., Hogdall, E.*, et al.* (2008). Multiple loci with different cancer specificities within the 8q24 gene desert. J Natl Cancer Inst *100*, 962-966.

Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. Nat Rev Genet *21*, 102-117.

Graham, L.D., Pedersen, S.K., Brown, G.S., Ho, T., Kassir, Z., Moynihan, A.T., Vizgoft, E.K., Dunne, R., Pimlott, L., Young, G.P.*, et al.* (2011). Colorectal Neoplasia Differentially Expressed (CRNDE), a Novel Gene with Elevated Expression in Colorectal Adenomas and Adenocarcinomas. Genes Cancer *2*, 829-840.

Graham, M., and Adams, J.M. (1986). Chromosome 8 breakpoint far 3' of the c-myc oncogene in a Burkitt's lymphoma 2;8 variant translocation is equivalent to the murine pvt-1 locus. EMBO J *5*, 2845-2851.

Graham, M., Adams, J.M., and Cory, S. (1985). Murine T lymphomas with retroviral inserts in the chromosomal 15 locus for plasmacytoma variant translocations. Nature *314*, 740-743.

Grisanzio, C., and Freedman, M.L. (2010). Chromosome 8q24-Associated Cancers and MYC. Genes Cancer *1*, 555-559.

Groff, A.F., Barutcu, A.R., Lewandowski, J.P., and Rinn, J.L. (2018). Enhancers in the Peril lincRNA locus regulate distant but not local genes. Genome Biol *19*, 219.

Groff, A.F., Sanchez-Gomez, D.B., Soruco, M.M.L., Gerhardinger, C., Barutcu, A.R., Li, E., Elcavage, L., Plana, O., Sanchez, L.V., Lee, J.C.*, et al.* (2016). In Vivo Characterization of Linc-p21 Reveals Functional cis-Regulatory DNA Elements. Cell Rep *16*, 2178-2186.

Gu, B., Swigut, T., Spencley, A., Bauer, M.R., Chung, M., Meyer, T., and Wysocka, J. (2018). Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. Science *359*, 1050-1055.

Guan, Y., Kuo, W.L., Stilwell, J.L., Takano, H., Lapuk, A.V., Fridlyand, J., Mao, J.H., Yu, M., Miller, M.A., Santos, J.L.*, et al.* (2007). Amplification of PVT1 contributes to the pathophysiology of ovarian and breast cancer. Clin Cancer Res *13*, 5745-5755.

Guo, H., Ahmed, M., Zhang, F., Yao, C.Q., Li, S., Liang, Y., Hua, J., Soares, F., Sun, Y., Langstein, J., *et al.* (2016). Modulation of long noncoding RNAs by risk SNPs underlying genetic predispositions to prostate cancer. Nat Genet *48*, 1142-1150.

Guo, J., Hao, C., Wang, C., and Li, L. (2018). Long noncoding RNA PVT1 modulates hepatocellular carcinoma cell proliferation and apoptosis by recruiting EZH2. Cancer Cell Int *18*, 98.

Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., *et al.* (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature *464*, 1071-1076.

Gutschner, T., and Diederichs, S. (2012). The hallmarks of cancer: a long non-coding RNA point of view. RNA Biol *9*, 703-719.

Gutschner, T., Hammerle, M., Eissmann, M., Hsu, J., Kim, Y., Hung, G., Revenko, A., Arun, G., Stentrup, M., Gross, M., *et al.* (2013). The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. Cancer Res *73*, 1180-1189.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., *et al.* (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223-227.

Hacisuleyman, E., Goff, L.A., Trapnell, C., Williams, A., Henao-Mejia, J., Sun, L., McClanahan, P., Hendrickson, D.G., Sauvageau, M., Kelley, D.R., *et al.* (2014). Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. Nat Struct Mol Biol *21*, 198-206.

Hagege, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forne, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat Protoc *2*, 1722-1733.

Haiman, C.A., Patterson, N., Freedman, M.L., Myers, S.R., Pike, M.C., Waliszewska, A., Neubauer, J., Tandon, A., Schirmer, C., McDonald, G.J., *et al.* (2007). Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet *39*, 638-644.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57-70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646-674.

Harms, K.L., and Chen, X. (2007). Histone deacetylase 2 modulates p53 transcriptional activities through regulation of p53-DNA binding activity. Cancer Res *67*, 3145-3152.

Hashemi, M., Moazeni-Roodi, A., Sarabandi, S., Karami, S., and Ghavami, S. (2019). Association between genetic polymorphisms of long noncoding RNA H19 and cancer risk: a meta-analysis. J Genet *98*.

Hayward, W.S., Neel, B.G., and Astrin, S.M. (1981). Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. Nature *290*, 475-480.

Herriges, M.J., Swarr, D.T., Morley, M.P., Rathi, K.S., Peng, T., Stewart, K.M., and Morrisey, E.E. (2014). Long noncoding RNAs are spatially correlated with transcription factors and regulate lung development. Genes Dev *28*, 1363-1379.

Hessels, D., Klein Gunnewiek, J.M., van Oort, I., Karthaus, H.F., van Leenders, G.J., van Balken, B., Kiemeney, L.A., Witjes, J.A., and Schalken, J.A. (2003). DD3(PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. Eur Urol *44*, 8-15; discussion 15-16.

Hibi, K., Nakamura, H., Hirai, A., Fujikake, Y., Kasai, Y., Akiyama, S., Ito, K., and Takagi, H. (1996). Loss of H19 imprinting in esophageal cancer. Cancer Res *56*, 480-482.

Ho, J.S., Ma, W., Mao, D.Y., and Benchimol, S. (2005). p53-Dependent transcriptional repression of c-myc is required for G1 cell cycle arrest. Mol Cell Biol *25*, 7423-7431.

Hoadley, K.A., Yau, C., Hinoue, T., Wolf, D.M., Lazar, A.J., Drill, E., Shen, R., Taylor, A.M., Cherniack, A.D., Thorsson, V.*, et al.* (2018). Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell *173*, 291-304 e296.

Hosono, Y., Niknafs, Y.S., Prensner, J.R., Iyer, M.K., Dhanasekaran, S.M., Mehra, R., Pitchiaya, S., Tien, J., Escara-Wilke, J., Poliakov, A.*, et al.* (2017). Oncogenic Role of THOR, a Conserved Cancer/Testis Long Non-coding RNA. Cell *171*, 1559-1572 e1520.

Hu, W.L., Jin, L., Xu, A., Wang, Y.F., Thorne, R.F., Zhang, X.D., and Wu, M. (2018). GUARDIN is a p53-responsive long non-coding RNA that is essential for genomic stability. Nat Cell Biol *20*, 492-502.

Hu, X., Feng, Y., Zhang, D., Zhao, S.D., Hu, Z., Greshock, J., Zhang, Y., Yang, L., Zhong, X., Wang, L.P.*, et al.* (2014a). A functional genomic approach identifies FAL1 as an oncogenic long noncoding RNA that associates with BMI1 and represses p21 expression in cancer. Cancer Cell *26*, 344-357.

Hu, Y., Wang, J., Qian, J., Kong, X., Tang, J., Wang, Y., Chen, H., Hong, J., Zou, W., Chen, Y., *et al.* (2014b). Long noncoding RNA GAPLINC regulates CD44-dependent cell invasiveness and associates with poor prognosis of gastric cancer. Cancer Res *74*, 6890-6902.

Hua, J.T., Ahmed, M., Guo, H., Zhang, Y., Chen, S., Soares, F., Lu, J., Zhou, S., Wang, M., Li, H., *et al.* (2018). Risk SNP-Mediated Promoter-Enhancer Switching Drives Prostate Cancer through lncRNA PCAT19. Cell *174*, 564-575 e518.

Huang, D., Chen, J., Yang, L., Ouyang, Q., Li, J., Lao, L., Zhao, J., Liu, J., Lu, Y., Xing, Y., *et al.* (2018a). NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. Nat Immunol *19*, 1112-1125.

Huang, X., Zhang, W., and Shao, Z. (2018b). Association between long non-coding RNA polymorphisms and cancer risk: a meta-analysis. Biosci Rep *38*.

Huarte, M. (2015). The emerging role of lncRNAs in cancer. Nat Med *21*, 1253-1261.

Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., *et al.* (2010). A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell *142*, 409-419.

Hung, K.L., Yost, K.E., Xie, L., Wu, S., Lange, J.T., Duffy, C.V., Kraft, K., Tang, J., Shi, Q., Rose, J.C., *et al.* (2020). EcDNA hubs driver cooperative intermolecular oncogene expression. bioRxiv.

Hung, T., Wang, Y., Lin, M.F., Koegel, A.K., Kotake, Y., Grant, G.D., Horlings, H.M., Shah, N., Umbricht, C., Wang, P., *et al.* (2011). Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet *43*, 621-629.

Huppi, K., Pitt, J.J., Wahlberg, B.M., and Caplen, N.J. (2012). The 8q24 gene desert: an oasis of non-coding transcriptional activity. Front Genet *3*, 69.

Hutchinson, J.N., Ensminger, A.W., Clemson, C.M., Lynch, C.R., Lawrence, J.B., and Chess, A. (2007). A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC Genomics *8*, 39.

Idogawa, M., Ohashi, T., Sasaki, Y., Nakase, H., and Tokino, T. (2017). Long non-coding RNA NEAT1 is a transcriptional target of p53 and modulates p53-induced transactivation and tumor-suppressor function. Int J Cancer *140*, 2785-2791.

International Human Genome Sequencing, C. (2004). Finishing the euchromatic sequence of the human genome. Nature *431*, 931-945.

Isoda, T., Moore, A.J., He, Z., Chandra, V., Aida, M., Denholtz, M., Piet van Hamburg, J., Fisch, K.M., Chang, A.N., Fahl, S.P., *et al.* (2017). Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. Cell *171*, 103-119 e118.

Ito, A., Kawaguchi, Y., Lai, C.H., Kovacs, J.J., Higashimoto, Y., Appella, E., and Yao, T.P. (2002). MDM2-HDAC1-mediated deacetylation of p53 is required for its degradation. EMBO J *21*, 6236-6245.

Iwakawa, R., Takenaka, M., Kohno, T., Shimada, Y., Totoki, Y., Shibata, T., Tsuta, K., Nishikawa, R., Noguchi, M., Sato-Otsubo, A., *et al.* (2013). Genome-wide identification of genes with amplification and/or fusion in small cell lung cancer. Genes Chromosomes Cancer *52*, 802-816.

Iyer, M.K., Niknafs, Y.S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., Barrette, T.R., Prensner, J.R., Evans, J.R., Zhao, S., *et al.* (2015). The landscape of long noncoding RNAs in the human transcriptome. Nat Genet *47*, 199-208.

Jacks, T., Remington, L., Williams, B.O., Schmitt, E.M., Halachmi, S., Bronson, R.T., and Weinberg, R.A. (1994). Tumor spectrum analysis in p53-mutant mice. Curr Biol *4*, 1-7.

Jackson, E.L., Olive, K.P., Tuveson, D.A., Bronson, R., Crowley, D., Brown, M., and Jacks, T. (2005). The differential effects of mutant p53 alleles on advanced murine lung cancer. Cancer Res *65*, 10280-10288.

Jackson, E.L., Willis, N., Mercer, K., Bronson, R.T., Crowley, D., Montoya, R., Jacks, T., and Tuveson, D.A. (2001). Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. Genes Dev *15*, 3243-3248.

Jadaliha, M., Zong, X., Malakar, P., Ray, T., Singh, D.K., Freier, S.M., Jensen, T., Prasanth, S.G., Karni, R., Ray, P.S., *et al.* (2016). Functional and prognostic significance of long non-coding RNA MALAT1 as a metastasis driver in ER negative lymph node negative breast cancer. Oncotarget *7*, 40418-40436.

Jain, A.K., Xi, Y., McCarthy, R., Allton, K., Akdemir, K.C., Patel, L.R., Aronow, B., Lin, C., Li, W., Yang, L., *et al.* (2016). LncPRESS1 Is a p53-Regulated LncRNA that Safeguards Pluripotency by Disrupting SIRT6-Mediated De-acetylation of Histone H3K56. Mol Cell *64*, 967-981.

Jandura, A., and Krause, H.M. (2017). The New RNA World: Growing Evidence for Long Noncoding RNA Functionality. Trends Genet *33*, 665-676.

Ji, P., Diederichs, S., Wang, W., Boing, S., Metzger, R., Schneider, P.M., Tidow, N., Brandt, B., Buerger, H., Bulk, E., *et al.* (2003). MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene *22*, 8031-8041.

Jiang, D., Brady, C.A., Johnson, T.M., Lee, E.Y., Park, E.J., Scott, M.P., and Attardi, L.D. (2011). Full p53 transcriptional activation potential is dispensable for tumor suppression in diverse lineages. Proc Natl Acad Sci U S A *108*, 17123-17128.

Jiang, H., Li, T., Qu, Y., Wang, X., Li, B., Song, J., Sun, X., Tang, Y., Wan, J., Yu, Y., *et al.* (2018). Long non-coding RNA SNHG15 interacts with and stabilizes transcription factor Slug and promotes colon cancer progression. Cancer Lett *425*, 78-87.

Jimenez, R.M., Polanco, J.A., and Luptak, A. (2015). Chemistry and Biology of Self-Cleaving Ribozymes. Trends Biochem Sci *40*, 648-661.

Jinek, M., Coyle, S.M., and Doudna, J.A. (2011). Coupled 5' nucleotide recognition and processivity in Xrn1-mediated mRNA decay. Mol Cell *41*, 600-608.

Jones, S.N., Hancock, A.R., Vogel, H., Donehower, L.A., and Bradley, A. (1998). Overexpression of Mdm2 in mice reveals a p53-independent role for Mdm2 in tumorigenesis. Proc Natl Acad Sci U S A *95*, 15608-15612.

Joung, J., Engreitz, J.M., Konermann, S., Abudayyeh, O.O., Verdine, V.K., Aguet, F., Gootenberg, J.S., Sanjana, N.E., Wright, J.B., Fulco, C.P., *et al.* (2017). Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. Nature *548*, 343-346.

Kamijo, T., Bodner, S., van de Kamp, E., Randle, D.H., and Sherr, C.J. (1999). Tumor spectrum in ARF-deficient mice. Cancer Res *59*, 2217-2222.

Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., *et al.* (2013). Mutational landscape and significance across 12 major cancer types. Nature *502*, 333-339.

Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermuller, J., Hofacker, I.L., *et al.* (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science *316*, 1484-1488.

Kastenhuber, E.R., and Lowe, S.W. (2017). Putting p53 in Context. Cell *170*, 1062-1078.

Kawakami, T., Okamoto, K., Sugihara, H., Hattori, T., Reeve, A.E., Ogawa, O., and Okada, Y. (2003). The roles of supernumerical X chromosomes and XIST expression in testicular germ cell tumors. J Urol *169*, 1546-1552.

Khaitan, D., Dinger, M.E., Mazar, J., Crawford, J., Smith, M.A., Mattick, J.S., and Perera, R.J. (2011). The melanoma-upregulated long noncoding RNA SPRY4-IT1 modulates apoptosis and invasion. Cancer Res *71*, 3852-3862.

Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., *et al.* (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A *106*, 11667-11672.

Khorshidi, H.R., Taheri, M., Noroozi, R., Sarrafzadeh, S., Sayad, A., and Ghafouri-Fard, S. (2017). ANRIL Genetic Variants in Iranian Breast Cancer Patients. Cell J *19*, 72-78.

Kim, D., Langmead, B., and Salzberg, S.L. (2015a). HISAT: a fast spliced aligner with low memory requirements. Nat Methods *12*, 357-360.

Kim, H.P., Cho, G.A., Han, S.W., Shin, J.Y., Jeong, E.G., Song, S.H., Lee, W.C., Lee, K.H., Bang, D., Seo, J.S., *et al.* (2014a). Novel fusion transcripts in human gastric cancer revealed by transcriptome analysis. Oncogene *33*, 5434-5441.

Kim, J., Piao, H.L., Kim, B.J., Yao, F., Han, Z., Wang, Y., Xiao, Z., Siverly, A.N., Lawhon, S.E., Ton, B.N., *et al.* (2018). Long noncoding RNA MALAT1 suppresses breast cancer metastasis. Nat Genet *50*, 1705-1715.

Kim, T., Cui, R., Jeon, Y.J., Lee, J.H., Lee, J.H., Sim, H., Park, J.K., Fadda, P., Tili, E., Nakanishi, H., *et al.* (2014b). Long-range interaction and correlation between MYC enhancer and oncogenic long noncoding RNA CARLo-5. Proc Natl Acad Sci U S A *111*, 4173-4178.

Kim, T., Jeon, Y.J., Cui, R., Lee, J.H., Peng, Y., Kim, S.H., Tili, E., Alder, H., and Croce, C.M. (2015b). Role of MYC-regulated long noncoding RNAs in cell cycle regulation and tumorigenesis. J Natl Cancer Inst *107*.

Kirk, J.M., Kim, S.O., Inoue, K., Smola, M.J., Lee, D.M., Schertzer, M.D., Wooten, J.S., Baker, A.R., Sprague, D., Collins, D.W., *et al.* (2018). Functional classification of long non-coding RNAs by k-mer content. Nat Genet *50*, 1474-1482.

Knudson, A.G. (2001). Two genetic hits (more or less) to cancer. Nat Rev Cancer *1*, 157-162.

Kondo, M., Suzuki, H., Ueda, R., Osada, H., Takagi, K., Takahashi, T., and Takahashi, T. (1995). Frequent loss of imprinting of the H19 gene is often associated with its overexpression in human lung cancers. Oncogene *10*, 1193-1198.

Kong, R., Zhang, E.B., Yin, D.D., You, L.H., Xu, T.P., Chen, W.M., Xia, R., Wan, L., Sun, M., Wang, Z.X., *et al.* (2015). Long noncoding RNA PVT1 indicates a poor prognosis of gastric cancer and promotes cell proliferation through epigenetically regulating p15 and p16. Mol Cancer *14*, 82.

Kong, Y., Hsieh, C.H., and Alonso, L.C. (2018). ANRIL: A lncRNA at the CDKN2A/B Locus With Roles in Cancer and Metabolic Disease. Front Endocrinol (Lausanne) *9*, 405.

Kopp, F., and Mendell, J.T. (2018). Functional Classification and Experimental Dissection of Long Noncoding RNAs. Cell *172*, 393-407.

Kotzin, J.J., Spencer, S.P., McCright, S.J., Kumar, D.B.U., Collet, M.A., Mowel, W.K., Elliott, E.N., Uyar, A., Makiya, M.A., Dunagin, M.C., *et al.* (2016). The long non-coding RNA Morrbid regulates Bim and short-lived myeloid cell lifespan. Nature *537*, 239-243.

Kress, T.R., Sabo, A., and Amati, B. (2015). MYC: connecting selective transcriptional control to global RNA production. Nat Rev Cancer *15*, 593-607.

Kretz, M., Siprashvili, Z., Chu, C., Webster, D.E., Zehnder, A., Qu, K., Lee, C.S., Flockhart, R.J., Groff, A.F., Chow, J., *et al.* (2013). Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature *493*, 231-235.

Lai, F., Damle, S.S., Ling, K.K., and Rigo, F. (2020). Directed RNase H Cleavage of Nascent Transcripts Causes Transcription Termination. Mol Cell *77*, 1032-1043 e1034.

Lai, F., Orom, U.A., Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A., and Shiekhattar, R. (2013). Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature *494*, 497-501.

Latos, P.A., Pauler, F.M., Koerner, M.V., Senergin, H.B., Hudson, Q.J., Stocsits, R.R., Allhoff, W., Stricker, S.H., Klement, R.M., Warczok, K.E., *et al.* (2012). Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. Science *338*, 1469-1472.

Lee, J.S., and Mendell, J.T. (2020). Antisense-Mediated Transcript Knockdown Triggers Premature Transcription Termination. Mol Cell *77*, 1044-1054 e1043.

Lee, J.T. (2012). Epigenetic regulation by long noncoding RNAs. Science *338*, 1435-1439.

Lee, S., Kopp, F., Chang, T.C., Sataluri, A., Chen, B., Sivakumar, S., Yu, H., Xie, Y., and Mendell, J.T. (2016). Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. Cell *164*, 69-80.

Lefever, S., Anckaert, J., Volders, P.J., Luypaert, M., Vandesompele, J., and Mestdagh, P. (2017). decodeRNA- predicting non-coding RNA functions using guilt-by-association. Database (Oxford) *2017*.

Leucci, E., Vendramin, R., Spinazzi, M., Laurette, P., Fiers, M., Wouters, J., Radaelli, E., Eyckerman, S., Leonelli, C., Vanderheyden, K.*, et al.* (2016). Melanoma addiction to the long non-coding RNA SAMMSON. Nature *531*, 518-522.

Leveille, N., Melo, C.A., Rooijers, K., Diaz-Lagares, A., Melo, S.A., Korkmaz, G., Lopes, R., Moqadam, F.A., Maia, A.R., Wijchers, P.J.*, et al.* (2015). Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a lncRNA. Nat Commun *6*, 6520.

Levine, A.J., and Oren, M. (2009). The first 30 years of p53: growing ever more complex. Nat Rev Cancer *9*, 749-758.

Levitt, N., Briggs, D., Gil, A., and Proudfoot, N.J. (1989). Definition of an efficient synthetic poly(A) site. Genes Dev *3*, 1019-1025.

Levy, N., Yonish-Rouach, E., Oren, M., and Kimchi, A. (1993). Complementation by wild-type p53 of interleukin-6 effects on M1 cells: induction of cell cycle exit and cooperativity with c-myc suppression. Mol Cell Biol *13*, 7942-7952.

Lewandowski, J.P., Lee, J.C., Hwang, T., Sunwoo, H., Goldstein, J.M., Groff, A.F., Chang, N.P., Mallard, W., Williams, A., Henao-Meija, J.*, et al.* (2019). The Firre locus produces a trans-acting RNA molecule that functions in hematopoiesis. Nat Commun *10*, 5137.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Li, L., Liu, B., Wapinski, O.L., Tsai, M.C., Qu, K., Zhang, J., Carlson, J.C., Lin, M., Fang, F., Gupta, R.A.*, et al.* (2013). Targeted disruption of Hotair leads to homeotic transformation and gene derepression. Cell Rep *5*, 3-12.

Li, L., van Breugel, P.C., Loayza-Puch, F., Ugalde, A.P., Korkmaz, G., Messika-Gold, N., Han, R., Lopes, R., Barbera, E.P., Teunissen, H.*, et al.* (2018). LncRNA-OIS1 regulates DPP4 activation to modulate senescence induced by RAS. Nucleic Acids Res *46*, 4213-4227.

Li, T., Kon, N., Jiang, L., Tan, M., Ludwig, T., Zhao, Y., Baer, R., and Gu, W. (2012). Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. Cell *149*, 1269-1283.

Li, X.L., Subramanian, M., Jones, M.F., Chaudhary, R., Singh, D.K., Zong, X., Gryder, B., Sindri, S., Mo, M., Schetter, A., *et al.* (2017). Long Noncoding RNA PURPL Suppresses Basal p53 Levels and Promotes Tumorigenicity in Colorectal Cancer. Cell Rep *20*, 2408-2423.

Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst *1*, 417-425.

Lieberman, J. (2018). Tapping the RNA world for therapeutics. Nat Struct Mol Biol *25*, 357-364.

Lin, C.Y., Loven, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. Cell *151*, 56-67.

Ling, H., Spizzo, R., Atlasi, Y., Nicoloso, M., Shimizu, M., Redis, R.S., Nishida, N., Gafa, R., Song, J., Guo, Z., *et al.* (2013). CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res *23*, 1446-1461.

Liu, B., Sun, L., Liu, Q., Gong, C., Yao, Y., Lv, X., Lin, L., Yao, H., Su, F., Li, D., *et al.* (2015). A cytoplasmic NF-kappaB interacting long noncoding RNA blocks IkappaB phosphorylation and suppresses breast cancer metastasis. Cancer Cell *27*, 370-381.

Liu, S.J., Horlbeck, M.A., Cho, S.W., Birk, H.S., Malatesta, M., He, D., Attenello, F.J., Villalta, J.E., Cho, M.Y., Chen, Y., *et al.* (2017). CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science *355*.

Liu, S.J., Malatesta, M., Lien, B.V., Saha, P., Thombare, S.S., Hong, S.J., Pedraza, L., Koontz, M., Seo, K., Horlbeck, M.A., *et al.* (2020). CRISPRi-based radiation modifier screen identifies long non-coding RNA therapeutic targets in glioma. Genome Biol *21*, 83.

Liu, X., Li, D., Zhang, W., Guo, M., and Zhan, Q. (2012). Long non-coding RNA gadd7 interacts with TDP-43 and regulates Cdk6 mRNA decay. EMBO J *31*, 4415-4427.

Liu, Z., Dai, J., and Shen, H. (2018). Systematic analysis reveals long noncoding RNAs regulating neighboring transcription factors in human cancers. Biochim Biophys Acta Mol Basis Dis *1864*, 2785-2792.

Loda, A., and Heard, E. (2019). Xist RNA in action: Past, present, and future. PLoS Genet *15*, e1008333.

Loewer, S., Cabili, M.N., Guttman, M., Loh, Y.H., Thomas, K., Park, I.H., Garber, M., Curran, M., Onder, T., Agarwal, S., *et al.* (2010). Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nat Genet *42*, 1113-1117.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Lowe, S.W., Jacks, T., Housman, D.E., and Ruley, H.E. (1994). Abrogation of oncogene-associated apoptosis allows transformation of p53-deficient cells. Proc Natl Acad Sci U S A *91*, 2026-2030.

Lu, D., Luo, P., Wang, Q., Ye, Y., and Wang, B. (2017). lncRNA PVT1 in cancer: A review and meta-analysis. Clin Chim Acta *474*, 1-7.

Lu, Y., Hu, Z., Mangala, L.S., Stine, Z.E., Hu, X., Jiang, D., Xiang, Y., Zhang, Y., Pradeep, S., Rodriguez-Aguayo, C., *et al.* (2018). MYC Targeted Long Noncoding RNA DANCR Promotes Cancer in Part by Reducing p21 Levels. Cancer Res *78*, 64-74.

Majumdar, G., Adris, P., Bhargava, N., Chen, H., and Raghow, R. (2012). Pan-histone deacetylase inhibitors regulate signaling pathways involved in proliferative and pro-inflammatory mechanisms in H9c2 cells. BMC Genomics *13*, 709.

Marin-Bejar, O., Marchese, F.P., Athie, A., Sanchez, Y., Gonzalez, J., Segura, V., Huang, L., Moreno, I., Navarro, A., Monzo, M., *et al.* (2013). Pint lincRNA connects the p53 pathway with epigenetic silencing by the Polycomb repressive complex 2. Genome Biol *14*, R104.

Marin-Bejar, O., Mas, A.M., Gonzalez, J., Martinez, D., Athie, A., Morales, X., Galduroz, M., Raimondi, I., Grossi, E., Guo, S., *et al.* (2017). The human lncRNA LINC-PINT inhibits tumor cell invasion through a highly conserved sequence element. Genome Biol *18*, 202.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal *17*, 10-12.

Mathy, N., Benard, L., Pellegrini, O., Daou, R., Wen, T., and Condon, C. (2007). 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. Cell *129*, 681-692.

Mattick, J.S. (2004). RNA regulation: a new genetics? Nat Rev Genet *5*, 316-323.

McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., *et al.* (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature *521*, 232-236.

Meijer, D., van Agthoven, T., Bosma, P.T., Nooter, K., and Dorssers, L.C. (2006). Functional screen for genes responsible for tamoxifen resistance in human breast cancer cells. Mol Cancer Res *4*, 379-386.

Mello, S.S., Sinow, C., Raj, N., Mazur, P.K., Bieging-Rolett, K., Broz, D.K., Imam, J.F.C., Vogel, H., Wood, L.D., Sage, J., *et al.* (2017). Neat1 is a p53-inducible lincRNA essential for transformation suppression. Genes Dev *31*, 1095-1108.

Melo, C.A., Drost, J., Wijchers, P.J., van de Werken, H., de Wit, E., Oude Vrielink, J.A., Elkon, R., Melo, S.A., Leveille, N., Kalluri, R., *et al.* (2013). eRNAs are required for p53-dependent enhancer activity and gene transcription. Mol Cell *49*, 524-535.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat Genet *36*, 1073-1078.

Mercer, T.R., Dinger, M.E., and Mattick, J.S. (2009). Long non-coding RNAs: insights into functions. Nat Rev Genet *10*, 155-159.

Meyer, K.B., Maia, A.T., O'Reilly, M., Ghoussaini, M., Prathalingam, R., Porter-Gill, P., Ambs, S., Prokunina-Olsson, L., Carroll, J., and Ponder, B.A. (2011). A functional variant at a prostate cancer predisposition locus at 8q24 is associated with PVT1 expression. PLoS Genet *7*, e1002165.

Millevoi, S., and Vagner, S. (2010). Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. Nucleic Acids Res *38*, 2757-2774.

Milowich, D., Le Mercier, M., De Neve, N., Sandras, F., Roumeguere, T., Decaestecker, C., Salmon, I., and Rorive, S. (2015). Diagnostic value of the UCA1 test for bladder cancer detection: a clinical study. Springerplus *4*, 349.

Moindrot, B., Cerase, A., Coker, H., Masui, O., Grijzenhout, A., Pintacuda, G., Schermelleh, L., Nesterova, T.B., and Brockdorff, N. (2015). A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. Cell Rep *12*, 562-572.

Monfort, A., Di Minin, G., Postlmayr, A., Freimann, R., Arieti, F., Thore, S., and Wutz, A. (2015). Identification of Spen as a Crucial Factor for Xist Function through Forward Genetic Screening in Haploid Embryonic Stem Cells. Cell Rep *12*, 554-561.

Monte, M., Simonatto, M., Peche, L.Y., Bublik, D.R., Gobessi, S., Pierotti, M.A., Rodolfo, M., and Schneider, C. (2006). MAGE-A tumor antigens target p53 transactivation function through histone deacetylase recruitment and confer resistance to chemotherapeutic agents. Proc Natl Acad Sci U S A *103*, 11160-11165.

Mount, S.M. (1982). A catalogue of splice junction sequences. Nucleic Acids Res *10*, 459-472.

Murphy, D.J., Junttila, M.R., Pouyet, L., Karnezis, A., Shchors, K., Bui, D.A., Brown-Swigart, L., Johnson, L., and Evan, G.I. (2008). Distinct thresholds govern Myc's biological output in vivo. Cancer Cell *14*, 447-457.

Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science *322*, 1717-1720.

Nagoshi, H., Taki, T., Hanamura, I., Nitta, M., Otsuki, T., Nishida, K., Okuda, K., Sakamoto, N., Kobayashi, S., Yamamoto-Sugitani, M.*, et al.* (2012). Frequent PVT1 rearrangement and novel chimeric genes PVT1-NBEA and PVT1-WWOX occur in multiple myeloma with 8q24 abnormality. Cancer Res *72*, 4954-4962.

Nakagawa, S., Ip, J.Y., Shioi, G., Tripathi, V., Zong, X., Hirose, T., and Prasanth, K.V. (2012). Malat1 is not an essential component of nuclear speckles in mice. RNA *18*, 1487-1499.

Nakagawa, S., Naganuma, T., Shioi, G., and Hirose, T. (2011). Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. J Cell Biol *193*, 31-39.

Narita, N., Fujieda, S., Kimura, Y., Ito, Y., Imoto, Y., Ogi, K., Takahashi, N., Tanaka, T., Tsuzuki, H., Yamada, T.*, et al.* (2010). Suppression of histone deacetylase 3 (HDAC3) enhances apoptosis induced by paclitaxel in human maxillary cancer cells in vitro and in vivo. Biochem Biophys Res Commun *396*, 310-316.

Nassar, D., Latil, M., Boeckx, B., Lambrechts, D., and Blanpain, C. (2015). Genomic landscape of carcinogen-induced and genetically induced mouse skin squamous cell carcinoma. Nat Med *21*, 946-954.

Niemczyk, M., Ito, Y., Huddleston, J., Git, A., Abu-Amero, S., Caldas, C., Moore, G.E., Stojic, L., and Murrell, A. (2013). Imprinted chromatin around DIRAS3 regulates alternative splicing of GNG12-AS1, a long noncoding RNA. Am J Hum Genet *93*, 224-235.

Nikitin, A.Y., Alcaraz, A., Anver, M.R., Bronson, R.T., Cardiff, R.D., Dixon, D., Fraire, A.E., Gabrielson, E.W., Gunning, W.T., Haines, D.C.*, et al.* (2004). Classification of proliferative pulmonary lesions of the mouse: recommendations of the mouse models of human cancers consortium. Cancer Res *64*, 2307-2316.

Nissan, A., Stojadinovic, A., Mitrani-Rosenbaum, S., Halle, D., Grinbaum, R., Roistacher, M., Bochem, A., Dayanc, B.E., Ritter, G., Gomceli, I.*, et al.*

(2012). Colon cancer associated transcript-1: a novel RNA expressed in malignant and pre-malignant human tissues. Int J Cancer *130*, 1598-1606.

Northcott, P.A., Shih, D.J., Peacock, J., Garzia, L., Morrissy, A.S., Zichner, T., Stutz, A.M., Korshunov, A., Reimand, J., Schumacher, S.E.*, et al.* (2012). Subgroup-specific structural variation across 1,000 medulloblastoma genomes. Nature *488*, 49-56.

Olivero, C.E., Martinez-Terroba, E., Zimmer, J., Liao, C., Tesfaye, E., Hooshdaran, N., Schofield, J.A., Bendor, J., Fang, D., Simon, M.D.*, et al.* (2020). p53 Activates the Long Noncoding RNA Pvt1b to Inhibit Myc and Suppress Tumorigenesis. Mol Cell *77*, 761-774 e768.

Olivier, M., Hollstein, M., and Hainaut, P. (2010). TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb Perspect Biol *2*, a001008.

Panda, A.C., Grammatikakis, I., Kim, K.M., De, S., Martindale, J.L., Munk, R., Yang, X., Abdelmohsen, K., and Gorospe, M. (2017). Identification of senescence-associated circular RNAs (SAC-RNAs) reveals senescence suppressor CircPVT1. Nucleic Acids Res *45*, 4021-4035.

Pandey, G.K., Mitra, S., Subhash, S., Hertwig, F., Kanduri, M., Mishra, K., Fransson, S., Ganeshram, A., Mondal, T., Bandaru, S.*, et al.* (2014). The risk-associated long noncoding RNA NBAT-1 controls neuroblastoma progression by regulating cell proliferation and neuronal differentiation. Cancer Cell *26*, 722-737.

Panzitt, K., Tschernatsch, M.M., Guelly, C., Moustafa, T., Stradner, M., Strohmaier, H.M., Buck, C.R., Denk, H., Schroeder, R., Trauner, M.*, et al.* (2007). Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. Gastroenterology *132*, 330-342.

Paralkar, V.R., Taborda, C.C., Huang, P., Yao, Y., Kossenkov, A.V., Prasad, R., Luan, J., Davies, J.O., Hughes, J.R., Hardison, R.C.*, et al.* (2016). Unlinking an lncRNA from Its Associated cis Element. Mol Cell *62*, 104-110.

Paris, R., Henry, R.E., Stephens, S.J., McBryde, M., and Espinosa, J.M. (2008). Multiple p53-independent gene silencing mechanisms define the cellular response to p53 activation. Cell Cycle *7*, 2427-2433.

Pasic, I., Shlien, A., Durbin, A.D., Stavropoulos, D.J., Baskin, B., Ray, P.N., Novokmet, A., and Malkin, D. (2010). Recurrent focal copy-number changes and loss of heterozygosity implicate two noncoding RNAs and one tumor suppressor gene at chromosome 3q13.31 in osteosarcoma. Cancer Res *70*, 160-171.

Pasmant, E., Laurendeau, I., Heron, D., Vidaud, M., Vidaud, D., and Bieche, I. (2007). Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. Cancer Res *67*, 3963-3969.

Payne, G.S., Bishop, J.M., and Varmus, H.E. (1982). Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. Nature *295*, 209-214.

Perna, D., Faga, G., Verrecchia, A., Gorski, M.M., Barozzi, I., Narang, V., Khng, J., Lim, K.C., Sung, W.K., Sanges, R., *et al.* (2012). Genome-wide mapping of Myc binding and gene regulation in serum-stimulated fibroblasts. Oncogene *31*, 1695-1709.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol *33*, 290-295.

Petrovics, G., Zhang, W., Makarem, M., Street, J.P., Connelly, R., Sun, L., Sesterhenn, I.A., Srikantan, V., Moul, J.W., and Srivastava, S. (2004). Elevated expression of PCGEM1, a prostate-specific gene with cell growth-promoting function, is associated with high-risk prostate cancer patients. Oncogene *23*, 605-611.

Platt, R.J., Chen, S., Zhou, Y., Yim, M.J., Swiech, L., Kempton, H.R., Dahlman, J.E., Parnas, O., Eisenhaure, T.M., Jovanovic, M., *et al.* (2014). CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell *159*, 440-455.

Ponjavic, J., Oliver, P.L., Lunter, G., and Ponting, C.P. (2009). Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. PLoS Genet *5*, e1000617.

Porter, J.R., Fisher, B.E., Baranello, L., Liu, J.C., Kambach, D.M., Nie, Z., Koh, W.S., Luo, J., Stommel, J.M., Levens, D., *et al.* (2017). Global Inhibition with Specific Activation: How p53 and MYC Redistribute the Transcriptome in the DNA Double-Strand Break Response. Mol Cell *67*, 1013-1025 e1019.

Prensner, J.R., and Chinnaiyan, A.M. (2011). The emergence of lncRNAs in cancer biology. Cancer Discov *1*, 391-407.

Prensner, J.R., Iyer, M.K., Balbin, O.A., Dhanasekaran, S.M., Cao, Q., Brenner, J.C., Laxman, B., Asangani, I.A., Grasso, C.S., Kominsky, H.D., *et al.* (2011). Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. Nat Biotechnol *29*, 742-749.

Protter, D.S.W., Rao, B.S., Van Treeck, B., Lin, Y., Mizoue, L., Rosen, M.K., and Parker, R. (2018). Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. Cell Rep *22*, 1401-1412.

Qian, X., Zhao, J., Yeung, P.Y., Zhang, Q.C., and Kwok, C.K. (2019). Revealing lncRNA Structures and Interactions by Sequencing-Based Approaches. Trends Biochem Sci *44*, 33-52.

Quinn, J.J., and Chang, H.Y. (2016). Unique features of long non-coding RNA biogenesis and function. Nat Rev Genet *17*, 47-62.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432-445.

Rainier, S., Johnson, L.A., Dobry, C.J., Ping, A.J., Grundy, P.E., and Feinberg, A.P. (1993). Relaxation of imprinted genes in human cancer. Nature *362*, 747-749.

Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L., and Corn, J.E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol *34*, 339-344.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu Rev Biochem *81*, 145-166.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E.*, et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Riquelme, E., Suraokar, M.B., Rodriguez, J., Mino, B., Lin, H.Y., Rice, D.C., Tsao, A., and Wistuba, II (2014). Frequent coamplification and cooperation between C-MYC and PVT1 oncogenes promote malignant pleural mesothelioma. J Thorac Oncol *9*, 998-1007.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Roth, A., Weinberg, Z., Chen, A.G., Kim, P.B., Ames, T.D., and Breaker, R.R. (2014). A widespread self-cleaving ribozyme class is revealed by bioinformatics. Nat Chem Biol *10*, 56-60.

Sabo, A., Kress, T.R., Pelizzola, M., de Pretis, S., Gorski, M.M., Tesi, A., Morelli, M.J., Bora, P., Doni, M., Verrecchia, A.*, et al.* (2014). Selective

transcriptional regulation by Myc in cellular growth control and lymphomagenesis. Nature *511*, 488-492.

Sachdeva, M., Zhu, S., Wu, F., Wu, H., Walia, V., Kumar, S., Elble, R., Watabe, K., and Mo, Y.Y. (2009). p53 represses c-Myc through induction of the tumor suppressor miR-145. Proc Natl Acad Sci U S A *106*, 3207-3212.

Sahakyan, A., Yang, Y., and Plath, K. (2018). The Role of Xist in X-Chromosome Dosage Compensation. Trends Cell Biol *28*, 999-1013.

Salzman, J., Gawad, C., Wang, P.L., Lacayo, N., and Brown, P.O. (2012). Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. PLoS One *7*, e30733.

Sanchez, Y., Segura, V., Marin-Bejar, O., Athie, A., Marchese, F.P., Gonzalez, J., Bujanda, L., Guo, S., Matheu, A., and Huarte, M. (2014). Genome-wide analysis of the human p53 transcriptional network unveils a lncRNA tumour suppressor signature. Nat Commun *5*, 5812.

Sasakawa, Y., Naoe, Y., Inoue, T., Sasakawa, T., Matsuo, M., Manda, T., and Mutoh, S. (2003). Effects of FK228, a novel histone deacetylase inhibitor, on tumor growth and expression of p21 and c-myc genes in vivo. Cancer Lett *195*, 161-168.

Sauvageau, M., Goff, L.A., Lodato, S., Bonev, B., Groff, A.F., Gerhardinger, C., Sanchez-Gomez, D.B., Hacisuleyman, E., Li, E., Spence, M.*, et al.* (2013). Multiple knockout mouse models reveal lincRNAs are required for life and brain development. Elife *2*, e01749.

Schmitt, A.M., Garcia, J.T., Hung, T., Flynn, R.A., Shen, Y., Qu, K., Payumo, A.Y., Peres-da-Silva, A., Broz, D.K., Baum, R.*, et al.* (2016). An inducible long noncoding RNA amplifies DNA damage signaling. Nat Genet *48*, 1370-1376.

Schofield, J.A., Duffy, E.E., Kiefer, L., Sullivan, M.C., and Simon, M.D. (2018). TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding. Nat Methods *15*, 221-225.

Schorderet, P., and Duboule, D. (2011). Structural and functional differences in the long non-coding RNA hotair in mouse and human. PLoS Genet *7*, e1002071.

Seoane, J., Le, H.V., and Massague, J. (2002). Myc suppression of the p21(Cip1) Cdk inhibitor influences the outcome of the p53 response to DNA damage. Nature *419*, 729-734.

Sharma, S., and Munger, K. (2020). Expression of the Long Noncoding RNA DINO in Human Papillomavirus-Positive Cervical Cancer Cells Reactivates the Dormant TP53 Tumor Suppressor through ATM/CHK2 Signaling. mBio *11*.

Shechner, D.M., Hacisuleyman, E., Younger, S.T., and Rinn, J.L. (2015). Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. Nat Methods *12*, 664-670.

Shen, L., Shi, Q., and Wang, W. (2018). Double agents: genes with both oncogenic and tumor-suppressor functions. Oncogenesis *7*, 25.

Shi, X.S., Li, J., Yang, R.H., Zhao, G.R., Zhou, H.P., Zeng, W.X., and Zhou, M. (2015). Correlation of increased MALAT1 expression with pathological features and prognosis in cancer patients: a meta-analysis. Genet Mol Res *14*, 18808-18819.

Shima, H., Kida, K., Adachi, S., Yamada, A., Sugae, S., Narui, K., Miyagi, Y., Nishi, M., Ryo, A., Murata, S.*, et al.* (2018). Lnc RNA H19 is associated with poor prognosis in breast cancer patients and promotes cancer stemness. Breast Cancer Res Treat *170*, 507-516.

Shtivelman, E., Henglein, B., Groitl, P., Lipp, M., and Bishop, J.M. (1989). Identification of a human transcription unit affected by the variant chromosomal translocations 2;8 and 8;22 of Burkitt lymphoma. Proc Natl Acad Sci U S A *86*, 3257-3260.

Sirchia, S.M., Ramoscelli, L., Grati, F.R., Barbera, F., Coradini, D., Rossella, F., Porta, G., Lesma, E., Ruggeri, A., Radice, P.*, et al.* (2005). Loss of the inactive X chromosome and replication of the active X in BRCA1-defective and wild-type breast cancer cells. Cancer Res *65*, 2139-2146.

Sleutels, F., Zwart, R., and Barlow, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature *415*, 810-813.

Soucek, L., Jucker, R., Panacchia, L., Ricordy, R., Tato, F., and Nasi, S. (2002). Omomyc, a potential Myc dominant negative, enhances Myc-induced apoptosis. Cancer Res *62*, 3507-3510.

Srikantan, V., Zou, Z., Petrovics, G., Xu, L., Augustus, M., Davis, L., Livezey, J.R., Connell, T., Sesterhenn, I.A., Yoshino, K.*, et al.* (2000). PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. Proc Natl Acad Sci U S A *97*, 12216-12221.

Statello, L., Guo, C.J., Chen, L.L., and Huarte, M. (2020). Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol.

Stewart, S.A., Dykxhoorn, D.M., Palliser, D., Mizuno, H., Yu, E.Y., An, D.S., Sabatini, D.M., Chen, I.S., Hahn, W.C., Sharp, P.A.*, et al.* (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. RNA *9*, 493-501.

Sun, Y., and Ma, L. (2019). New Insights into Long Non-Coding RNA MALAT1 in Cancer and Metastasis. Cancers (Basel) *11*.

Sunwoo, H., Dinger, M.E., Wilusz, J.E., Amaral, P.P., Mattick, J.S., and Spector, D.L. (2009). MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome Res *19*, 347-359.

Swerdlow, A.J., Schoemaker, M.J., Higgins, C.D., Wright, A.F., Jacobs, P.A., and Group, U.K.C.C. (2005). Cancer incidence and mortality in men with Klinefelter syndrome: a cohort study. J Natl Cancer Inst *97*, 1204-1210.

Szafron, L.M., Balcerak, A., Grzybowska, E.A., Pienkowska-Grela, B., Felisiak-Golabek, A., Podgorska, A., Kulesza, M., Nowak, N., Pomorski, P., Wysocki, J., *et al.* (2015). The Novel Gene CRNDE Encodes a Nuclear Peptide (CRNDEP) Which Is Overexpressed in Highly Proliferating Tissues. PLoS One *10*, e0127475.

Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., and Leder, P. (1982). Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc Natl Acad Sci U S A *79*, 7837-7841.

Tesfaye, E., Bendor, J., Martínez-Terroba, E. Winkler, L., Olivero, C., Zamudio, J., Dimitrova, N. In review. The p53 transcriptional response across tumor types reveals core and senescence-specific signatures modulated by long noncoding RNAs.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., *et al.* (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell *39*, 925-938.

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science *329*, 689-693.

Tseng, Y.Y., and Bagchi, A. (2015). The PVT1-MYC duet in cancer. Mol Cell Oncol *2*, e974467.

Tseng, Y.Y., Moriarity, B.S., Gong, W., Akiyama, R., Tiwari, A., Kawakami, H., Ronning, P., Reuland, B., Guenther, K., Beadnell, T.C., *et al.* (2014). PVT1 dependence in cancer with MYC copy-number increase. Nature *512*, 82-86.

Tuck, A.C., and Buhler, M. (2021). Long Non-coding RNA Depletion Using Self-Cleaving Ribozymes. Methods Mol Biol *2167*, 287-301.

Tuck, A.C., Natarajan, K.N., Rice, G.M., Borawski, J., Mohn, F., Rankova, A., Flemr, M., Wenger, A., Nutiu, R., Teichmann, S., *et al.* (2018). Distinctive features of lincRNA gene expression suggest widespread RNA-independent functions. Life Sci Alliance *1*, e201800124.

Vendramin, R., Verheyden, Y., Ishikawa, H., Goedert, L., Nicolas, E., Saraf, K., Armaos, A., Delli Ponti, R., Izumikawa, K., Mestdagh, P., *et al.* (2018). SAMMSON fosters cancer cell fitness by concertedly enhancing mitochondrial and cytosolic translation. Nat Struct Mol Biol *25*, 1035-1046.

Ventura, A., Kirsch, D.G., McLaughlin, M.E., Tuveson, D.A., Grimm, J., Lintault, L., Newman, J., Reczek, E.E., Weissleder, R., and Jacks, T. (2007). Restoration of p53 function leads to tumour regression in vivo. Nature *445*, 661-665.

Visel, A., Zhu, Y., May, D., Afzal, V., Gong, E., Attanasio, C., Blow, M.J., Cohen, J.C., Rubin, E.M., and Pennacchio, L.A. (2010). Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. Nature *464*, 409-412.

Volders, P.J., Lefever, S., Baute, S., Nuytens, J., Vanderheyden, K., Menten, B., Mestdagh, P., and Vandesompele, J. (2018). Targeted Genomic Screen Reveals Focal Long Non-Coding RNA Copy Number Alterations in Cancer Cell Lines. Noncoding RNA *4*.

Vousden, K.H., and Prives, C. (2009). Blinded by the Light: The Growing Complexity of p53. Cell *137*, 413-431.

Walz, S., Lorenzin, F., Morton, J., Wiese, K.E., von Eyss, B., Herold, S., Rycak, L., Dumay-Odelot, H., Karim, S., Bartkuhn, M., *et al.* (2014). Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles. Nature *511*, 483-487.

Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Mol Cell *43*, 904-914.

Wang, Z., Yang, B., Zhang, M., Guo, W., Wu, Z., Wang, Y., Jia, L., Li, S., Cancer Genome Atlas Research, N., Xie, W., *et al.* (2018). lncRNA Epigenetic Landscape Analysis Identifies EPIC1 as an Oncogenic lncRNA that Interacts with MYC and Promotes Cell-Cycle Progression in Cancer. Cancer Cell *33*, 706-720 e709.

Wapinski, O., and Chang, H.Y. (2011). Long noncoding RNAs and human disease. Trends Cell Biol *21*, 354-361.

West, J.A., Davis, C.P., Sunwoo, H., Simon, M.D., Sadreyev, R.I., Wang, P.I., Tolstorukov, M.Y., and Kingston, R.E. (2014). The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. Mol Cell *55*, 791-802.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522-525.

Wilusz, J.E., Freier, S.M., and Spector, D.L. (2008). 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. Cell *135*, 919-932.

Winkler, L., Zimmer, J., Williams, A., Simon, M.D., and Dimitrova, N. In preparation. Narrowing down the functional element of a transcribed cis-regulatory locus.

Wu, C.L., Wang, Y., Jin, B., Chen, H., Xie, B.S., and Mao, Z.B. (2015). Senescence-associated Long Non-coding RNA (SALNR) Delays Oncogene-induced Senescence through NF90 Regulation. J Biol Chem *290*, 30175-30192.

Wu, L., Xu, Q., Zhang, H., Li, M., Zhu, C., Jiang, M., Sang, X., Zhao, Y., Sun, Q., and Zhao, H. (2016). A new avenue for obtaining insight into the functional characteristics of long noncoding RNAs associated with estrogen receptor signaling. Sci Rep *6*, 31716.

Wu, W., Bhagat, T.D., Yang, X., Song, J.H., Cheng, Y., Agarwal, R., Abraham, J.M., Ibrahim, S., Bartenstein, M., Hussain, Z.*, et al.* (2013). Hypomethylation of noncoding DNA regions and overexpression of the long noncoding RNA, AFAP1-AS1, in Barrett's esophagus and esophageal adenocarcinoma. Gastroenterology *144*, 956-966 e954.

Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat Genet *30*, 167-174.

Xiang, J.F., Yin, Q.F., Chen, T., Zhang, Y., Zhang, X.O., Wu, Z., Zhang, S., Wang, H.B., Ge, J., Lu, X.*, et al.* (2014). Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. Cell Res *24*, 513-531.

Xie, H., Ma, H., and Zhou, D. (2013). Plasma HULC as a promising novel biomarker for the detection of hepatocellular carcinoma. Biomed Res Int *2013*, 136106.

Xing, F., Liu, Y., Wu, S.Y., Wu, K., Sharma, S., Mo, Y.Y., Feng, J., Sanders, S., Jin, G., Singh, R.*, et al.* (2018). Loss of XIST in Breast Cancer Activates MSN-c-Met and Reprograms Microglia via Exosomal miRNA to Promote Brain Metastasis. Cancer Res *78*, 4316-4330.

Xing, Z., Lin, A., Li, C., Liang, K., Wang, S., Liu, Y., Park, P.K., Qin, L., Wei, Y., Hawke, D.H., *et al.* (2014). lncRNA directs cooperative epigenetic regulation downstream of chemokine signals. Cell *159*, 1110-1125.

Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., Chen, J., and Chen, S. (2012). FastUniq: a fast de novo duplicates removal tool for paired short reads. PLoS One *7*, e52249.

Xu, M.D., Wang, Y., Weng, W., Wei, P., Qi, P., Zhang, Q., Tan, C., Ni, S.J., Dong, L., Yang, Y., *et al.* (2017). A Positive Feedback Loop of lncRNA-PVT1 and FOXM1 Facilitates Gastric Cancer Growth and Invasion. Clin Cancer Res *23*, 2071-2080.

Xu, W.S., Parmigiani, R.B., and Marks, P.A. (2007). Histone deacetylase inhibitors: molecular mechanisms of action. Oncogene *26*, 5541-5552.

Xu, Y., Voelter-Mahlknecht, S., and Mahlknecht, U. (2005). The histone deacetylase inhibitor suberoylanilide hydroxamic acid down-regulates expression levels of Bcr-abl, c-Myc and HDAC3 in chronic myeloid leukemia cell lines. Int J Mol Med *15*, 169-172.

Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N.S., Cai, W., Yang, G., Bronson, R., Crowley, D.G., *et al.* (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. Nature *514*, 380-384.

Yamazaki, T., Souquere, S., Chujo, T., Kobelke, S., Chong, Y.S., Fox, A.H., Bond, C.S., Nakagawa, S., Pierron, G., and Hirose, T. (2018). Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. Mol Cell *70*, 1038-1053 e1037.

Yan, X., Hu, Z., Feng, Y., Hu, X., Yuan, J., Zhao, S.D., Zhang, Y., Yang, L., Shan, W., He, Q., *et al.* (2015). Comprehensive Genomic Characterization of Long Non-coding RNAs across Human Cancers. Cancer Cell *28*, 529-540.

Yang, C., Li, Z., Li, Y., Xu, R., Wang, Y., Tian, Y., and Chen, W. (2017). Long non-coding RNA NEAT1 overexpression is associated with poor prognosis in cancer patients: a systematic review and meta-analysis. Oncotarget *8*, 2672-2680.

Yap, K., Mukhina, S., Zhang, G., Tan, J.S.C., Ong, H.S., and Makeyev, E.V. (2018). A Short Tandem Repeat-Enriched RNA Assembles a Nuclear Compartment to Control Alternative Splicing and Promote Cell Survival. Mol Cell *72*, 525-540 e513.

Yap, K.L., Li, S., Munoz-Cabello, A.M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M.J., and Zhou, M.M. (2010). Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol Cell *38*, 662-674.

Yeager, M., Orr, N., Hayes, R.B., Jacobs, K.B., Kraft, P., Wacholder, S., Minichiello, M.J., Fearnhead, P., Yu, K., Chatterjee, N., *et al.* (2007). Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet *39*, 645-649.

Yildirim, E., Kirby, J.E., Brown, D.E., Mercier, F.E., Sadreyev, R.I., Scadden, D.T., and Lee, J.T. (2013). Xist RNA is a potent suppressor of hematologic cancer in mice. Cell *152*, 727-742.

Yin, Q.F., Yang, L., Zhang, Y., Xiang, J.F., Wu, Y.W., Carmichael, G.G., and Chen, L.L. (2012). Long noncoding RNAs with snoRNA ends. Mol Cell *48*, 219-230.

You, S.H., Lim, H.W., Sun, Z., Broache, M., Won, K.J., and Lazar, M.A. (2013). Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 in vivo. Nat Struct Mol Biol *20*, 182-187.

Younger, S.T., Kenzelmann-Broz, D., Jung, H., Attardi, L.D., and Rinn, J.L. (2015). Integrative genomic analysis reveals widespread enhancer regulation by p53 in response to DNA damage. Nucleic Acids Res *43*, 4447-4462.

Yu, N., Chen, P., Wang, Q., Liang, M., Qiu, J., Zhou, P., Yang, M., Yang, P., Wu, Y., Han, X., *et al.* (2020). Histone deacetylase inhibitors differentially regulate c-Myc expression in retinoblastoma cells. Oncol Lett *19*, 460-468.

Zampetaki, A., Albrecht, A., and Steinhofel, K. (2018). Long Non-coding RNA Structure and Function: Is There a Link? Front Physiol *9*, 1201.

Zeng, Y., Wang, T., Liu, Y., Su, Z., Lu, P., Chen, X., and Hu, D. (2017). LncRNA PVT1 as an effective biomarker for cancer diagnosis and detection based on transcriptome data and meta-analysis. Oncotarget *8*, 75455-75466.

Zhang, A., Zhou, N., Huang, J., Liu, Q., Fukuda, K., Ma, D., Lu, Z., Bai, C., Watabe, K., and Mo, Y.Y. (2013). The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage. Cell Res *23*, 340-350.

Zhang, B., Arun, G., Mao, Y.S., Lazar, Z., Hung, G., Bhattacharjee, G., Xiao, X., Booth, C.J., Wu, J., Zhang, C., *et al.* (2012a). The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. Cell Rep *2*, 111-123.

Zhang, D., Zhang, G., Hu, X., Wu, L., Feng, Y., He, S., Zhang, Y., Hu, Z., Yang, L., Tian, T., *et al.* (2017). Oncogenic RAS Regulates Long Noncoding RNA Orilnc1 in Human Cancer. Cancer Res *77*, 3745-3757.

Zhang, J., Zhang, B., Wang, T., and Wang, H. (2015). LncRNA MALAT1 overexpression is an unfavorable prognostic factor in human cancer: evidence from a meta-analysis. Int J Clin Exp Med *8*, 5499-5505.

Zhang, M., Zhao, K., Xu, X., Yang, Y., Yan, S., Wei, P., Liu, H., Xu, J., Xiao, F., Zhou, H., *et al.* (2018a). A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. Nat Commun *9*, 4475.

Zhang, T.J., Zhou, J.D., Zhang, W., Lin, J., Ma, J.C., Wen, X.M., Yuan, Q., Li, X.X., Xu, Z.J., and Qian, J. (2018b). H19 overexpression promotes leukemogenesis and predicts unfavorable prognosis in acute myeloid leukemia. Clin Epigenetics *10*, 47.

Zhang, X., Chen, Q., He, C., Mao, W., Zhang, L., Xu, X., Zhu, J., and Chen, B. (2012b). Polymorphisms on 8q24 are associated with lung cancer risk and survival in Han Chinese. PLoS One *7*, e41930.

Zhang, Y., Pitchiaya, S., Cieslik, M., Niknafs, Y.S., Tien, J.C., Hosono, Y., Iyer, M.K., Yazdani, S., Subramaniam, S., Shukla, S.K., *et al.* (2018c). Analysis of the androgen receptor-regulated lncRNA landscape identifies a role for ARLNC1 in prostate cancer progression. Nat Genet *50*, 814-824.

Zhang, Y., Yang, G., and Luo, Y. (2019). Long non-coding RNA PVT1 promotes glioma cell proliferation and invasion by targeting miR-200a. Exp Ther Med *17*, 1337-1345.

Zhao, L., Kong, H., Sun, H., Chen, Z., Chen, B., and Zhou, M. (2018). LncRNA-PVT1 promotes pancreatic cancer cells proliferation and migration through acting as a molecular sponge to regulate miR-448. J Cell Physiol *233*, 4044-4055.

Zhao, X., Wei, X., Zhao, L., Shi, L., Cheng, J., Kang, S., Zhang, H., Zhang, J., Li, L., Zhang, H., *et al.* (2016). The rs6983267 SNP and long non-coding RNA CARLo-5 are associated with endometrial carcinoma. Environ Mol Mutagen *57*, 508-515.

Zheng, X., Hu, H., and Li, S. (2016). High expression of lncRNA PVT1 promotes invasion by inducing epithelial-to-mesenchymal transition in esophageal cancer. Oncol Lett *12*, 2357-2362.

Zhou, C.C., Yang, F., Yuan, S.X., Ma, J.Z., Liu, F., Yuan, J.H., Bi, F.R., Lin, K.Y., Yin, J.H., Cao, G.W., *et al.* (2016). Systemic genome screening identifies the outcome associated focal loss of long noncoding RNA PRAL in hepatocellular carcinoma. Hepatology *63*, 850-863.

Zhou, Y., Sheng, B., Xia, Q., Guan, X., and Zhang, Y. (2017). Association of long non-coding RNA H19 and microRNA-21 expression with the biological features and prognosis of non-small cell lung cancer. Cancer Gene Ther *24*, 317-324.

Zhu, P., Wu, J., Wang, Y., Zhu, X., Lu, T., Liu, B., He, L., Ye, B., Wang, S., Meng, S.*, et al.* (2018). LncGata6 maintains stemness of intestinal stem cells and promotes intestinal tumorigenesis. Nat Cell Biol *20*, 1134-1144.

Zhu, S., Li, W., Liu, J., Chen, C.H., Liao, Q., Xu, P., Xu, H., Xiao, T., Cao, Z., Peng, J.*, et al.* (2016). Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. Nat Biotechnol *34*, 1279-1286.

Zhu, S., Shuai, P., Yang, C., Zhang, Y., Zhong, S., Liu, X., Chen, K., Ran, Q., Yang, H., and Zhou, Y. (2017). Prognostic value of long non-coding RNA PVT1 as a novel biomarker in various cancers: a meta-analysis. Oncotarget *8*, 113174-113184.

Zindy, F., Eischen, C.M., Randle, D.H., Kamijo, T., Cleveland, J.L., Sherr, C.J., and Roussel, M.F. (1998). Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. Genes Dev *12*, 2424-2433.

Zylicz, J.J., Bousard, A., Zumer, K., Dossin, F., Mohammad, E., da Rocha, S.T., Schwalb, B., Syx, L., Dingli, F., Loew, D.*, et al.* (2019). The Implication of Early Chromatin Changes in X Chromosome Inactivation. Cell *176*, 182-197 e123.