Elucidating the Molecular Mechanism of Cis-Regulation by the Long Noncoding RNA LincRNA-p21

Lauren Nicole Winkler
Yale University Graduate School of Arts and Sciences, laurennwinkler@gmail.com

Follow this and additional works at: https://elischolar.library.yale.edu/gsas_dissertations

Recommended Citation
Winkler, Lauren Nicole, "Elucidating the Molecular Mechanism of Cis-Regulation by the Long Noncoding RNA LincRNA-p21" (2021). Yale Graduate School of Arts and Sciences Dissertations. 440.
https://elischolar.library.yale.edu/gsas_dissertations/440

This Dissertation is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Graduate School of Arts and Sciences Dissertations by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
Abstract

Elucidating the Molecular Mechanism of Cis-Regulation by the Long Noncoding RNA LincRNA-p21

Lauren Nicole Winkler
2021

Pervasive transcription is a hallmark of mammalian genomes. Although protein-coding genes span only a small fraction of the genome, more than two-thirds is transcribed, yielding thousands of noncoding transcripts whose expression exhibits a tight correlation with cell type, disease state, and other biological phenomena. A subset of these transcripts, termed long noncoding RNAs (lncRNAs) on account of their length (>200 nucleotides) and lack of apparent coding potential, have been shown to play functional roles in processes ranging from immune signaling to organogenesis. In contrast to trans-acting lncRNAs, which may operate in either the nucleus or cytoplasm, cis-acting lncRNAs remain at their site of transcription and regulate the expression of nearby protein-coding genes. These lncRNAs have been proposed to act through three main mechanisms: (1) the RNA molecule may interact with protein factors to enact transcriptional activation or repression; (2) the act of lncRNA transcription may increase the local concentration of RNA polymerase II or chromatin-modifying factors; or (3) DNA elements within a lncRNA locus may directly regulate the expression of both the lncRNA and its neighboring gene. Deconvolving these interlinked mechanisms has proven challenging and necessitates the development and implementation of new experimental techniques. In this work, we used a suite of independent molecular and genetic approaches, including a novel ribozyme-based tool for targeted transcript degradation, to expand our understanding of the molecular “logic” through which cis-acting lncRNAs enact gene regulation.

We focused on the p53-inducible lncRNA LincRNA-p21, which acts in cis to reinforce the expression of the nearby protein-coding gene and key p53 target p21/Cdkn1a. To identify the functional element of cis-regulation at this locus, we generated four mouse strains harboring complementary LincRNA-p21 loss-of-function mutations that allowed us to separately examine the importance of the LincRNA-p21 transcriptional process and the underlying DNA sequence. This parallel genetic approach
demonstrated that full-length *LincRNA-p21* transcription, processing, and accumulation are dispensable for *cis*-regulation and revealed a requirement for a conserved sequence element within exon 1. Further experiments with molecular tools suggested that active transcription through this conserved region promotes *p21* expression, implicating a dual role for the transcriptional process and sequence elements within the locus. This comprehensive functional dissection of a single lncRNA locus attests to the regulatory potential of lncRNA loci and further reveals the biological significance of pervasive genomic transcription.
Elucidating the Molecular Mechanism of Cis-Regulation by the Long Noncoding RNA
LincRNA-p21

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

By
Lauren Nicole Winkler

Dissertation director: Nadya Dimitrova, Ph.D.
December 2021
Table of Contents

Abstract ........................................................................................................................................... i
Table of contents .......................................................................................................................... v
Acknowledgments .......................................................................................................................... vii
Nonstandard abbreviations ......................................................................................................... ix
List of figures and tables ............................................................................................................... xi
Chapter I: Introduction .................................................................................................................. 1
The role of noncoding transcription in gene regulation
  Pervasive noncoding transcription as a hallmark of mammalian genomes ........................................ 1
  The role of cis-acting noncoding RNAs in homeostasis and disease ........................................... 5
  Mechanisms of cis-regulation by noncoding RNAs ..................................................................... 11
  Tools for studying cis-acting noncoding RNAs .......................................................................... 23
Chapter II: Project framework ....................................................................................................... 36
LincRNA-p21: a functional mediator of the p53 pathway
  The p53 tumor suppressor pathway ............................................................................................ 36
  The long noncoding RNA LincRNA-p21 in the p53 tumor suppressor pathway......................... 40
Chapter III: Development and validation of an inducible self-cleaving ribozyme ....................... 51
  Introduction .................................................................................................................................. 51
  Results ......................................................................................................................................... 51
  Data summary and conclusion ..................................................................................................... 55
Chapter IV: Elucidating the molecular mechanism of cis-regulation by LincRNA-p21 ................. 57
  Introduction .................................................................................................................................. 57
  Results ......................................................................................................................................... 59
  Data summary and conclusion ..................................................................................................... 85
Chapter V: Summary and perspectives .......................................................................................... 88
  The regulatory significance of pervasive transcription ................................................................. 88
  The importance of an expanded toolkit for studying noncoding RNA function ......................... 92
  Final remarks ............................................................................................................................... 97
Acknowledgments

As John Donne famously wrote, “no man is an island entire of itself;” indeed, I have leaned heavily on the advice and encouragement of friends, family and colleagues. This thesis would be incomplete without an expression of my deep and heartfelt gratitude for their unfailing support.

I would first like to acknowledge my thesis advisor, Dr. Nadya Dimitrova, for the vision, creativity, and drive that formed the bedrock of my graduate experience. Before even applying to Yale, I heard Nadya speak at the 2015 RNA retreat and immediately knew that I wanted to join her in her bold adventure to chart the frontiers of the noncoding genome. Her innovative approach to long-standing questions has served as both a source of inspiration and as an invitation to grow as a scientist by applying new methods and an inquisitive spirit to my own project. I owe her a debt of immense gratitude for the mentorship and guidance that she has provided during my time in the lab.

I would next like to express my thanks to my thesis committee—Drs. Ronald Breaker, Matt Simon, and Karla Neugebauer—whose trenchant questions challenged me to think deeply about my project and ultimately become a more rigorous scientist. I thank them for generously sharing their expertise in, and enthusiasm for, RNA biology.

My colleagues—now cherished friends—in the Dimitrova lab made the bench feel like a second home. Ephrath Tesfaye, Christiane Olivero, Clara Liao, Elena Martínez-Terroba, Peppe Militello, Rahul Nagvekar, Yee Nip, Sonja Herter, Nina Grexa, Meg Tian, and Ben Shao all helped foster the warm and collaborative spirit that typified the lab. In good times, we shared research questions and personal accomplishments, often over a happy hour drink. And when the scientific process began to feel all-too-overwhelming, they were ready with sound advice, unflappable optimism, and a seemingly limitless supply of memes. My time in the Dimitrova lab was immeasurably richer for having worked alongside such a wonderful group of individuals.

I am grateful to the members of the Department of Molecular, Cellular, and Developmental Biology and the Center for RNA Science and Medicine for fostering a lively scientific environment. It
was a rare privilege to be a part of two communities that embody such a deep commitment to collaboration, intellectual engagement, and innovation.

I have also benefitted from the generous support of the National Institute of Health, whose Training Program in Cellular, Molecular, and Quantitative Biology funded my work and afforded me many opportunities for scientific and professional growth.

Lastly, I would like to express my overwhelming appreciation for my friends and family near and far, who served as a constant source of love and wisdom. Since there is not enough space to express the full measure of my gratitude, I will confine myself to a simple and heartfelt “thank you.” From the bottom of my heart, I look forward to personally telling each of you just how much your support has meant.
## Nonstandard abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>3C</td>
<td>chromosome conformation capture</td>
</tr>
<tr>
<td>ASOs</td>
<td>antisense oligonucleotides</td>
</tr>
<tr>
<td>Cdkn1a</td>
<td>cyclin dependent kinase inhibitor 1a</td>
</tr>
<tr>
<td>CSF</td>
<td>codon substitution frequency</td>
</tr>
<tr>
<td>CTCF</td>
<td>CCCTC binding factor</td>
</tr>
<tr>
<td>DNMT1</td>
<td>DNA methyltransferase 1</td>
</tr>
<tr>
<td>Doxo</td>
<td>doxorubicin</td>
</tr>
<tr>
<td>dRNA</td>
<td>&quot;dead&quot; RNA</td>
</tr>
<tr>
<td>eRNA</td>
<td>enhancer RNA</td>
</tr>
<tr>
<td>FACS</td>
<td>fluorescence-assisted cell sorting</td>
</tr>
<tr>
<td>Gapdh</td>
<td>glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>gRNA</td>
<td>guide RNA</td>
</tr>
<tr>
<td>HDR</td>
<td>homology-directed repair</td>
</tr>
<tr>
<td>HIF-1α</td>
<td>hypoxia inducible factor 1α</td>
</tr>
<tr>
<td>hnRNP-K</td>
<td>heterogeneous nuclear ribonucleoprotein K</td>
</tr>
<tr>
<td>HSF1</td>
<td>heat shock factor 1</td>
</tr>
<tr>
<td>iPSC</td>
<td>induced pluripotent stem cell</td>
</tr>
<tr>
<td>KRAB</td>
<td>Krüppel-associated box</td>
</tr>
<tr>
<td>lincRNA</td>
<td>long intergenic noncoding RNA</td>
</tr>
<tr>
<td>IncRNA</td>
<td>long noncoding RNA</td>
</tr>
<tr>
<td>MEF</td>
<td>mouse embryonic fibroblast</td>
</tr>
<tr>
<td>ncRNA</td>
<td>noncoding RNA</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Name</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>p53RE</td>
<td>p53 responsive element</td>
</tr>
<tr>
<td>PAM</td>
<td>protospacer adjacent motif</td>
</tr>
<tr>
<td>PAS</td>
<td>polyadenylation signal</td>
</tr>
<tr>
<td>Pol II</td>
<td>RNA polymerase II</td>
</tr>
<tr>
<td>PRC1/2</td>
<td>Polycomb repressive complex 1/2</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>reverse transcription-quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>SAM</td>
<td>synergistic activation mediator</td>
</tr>
<tr>
<td>SETDB1</td>
<td>SET Domain Bifurcated Histone Lysine Methyltransferase 1</td>
</tr>
<tr>
<td>smRNA FISH</td>
<td>single molecule RNA fluorescence in situ hybridization</td>
</tr>
<tr>
<td>snRNP</td>
<td>small nuclear ribonucleoprotein</td>
</tr>
<tr>
<td>Srsf3</td>
<td>serine and arginine rich splicing factor 3</td>
</tr>
<tr>
<td>ssDNA</td>
<td>single-stranded DNA</td>
</tr>
<tr>
<td>SV40</td>
<td>Simian virus 40</td>
</tr>
<tr>
<td>Tam</td>
<td>4-hydroxytamoxifen</td>
</tr>
<tr>
<td>TSS</td>
<td>transcription start site</td>
</tr>
<tr>
<td>TTS</td>
<td>transcription termination site</td>
</tr>
<tr>
<td>TWI</td>
<td>Twister</td>
</tr>
<tr>
<td>VHL</td>
<td>Von Hippel-Lindau tumor suppressor</td>
</tr>
<tr>
<td>XRN1/2</td>
<td>Exoribonuclease 1/2</td>
</tr>
</tbody>
</table>
List of Figures and Tables

Figure 1. LncRNAs arise from diverse genomic contexts ................................................................. 3

Figure 2. Molecular approaches for interrogating LncRNA function ................................................. 24

Figure 3. Genetic approaches for interrogating LncRNA function ................................................... 30

Figure 4. Historical characterization of the LincRNA-p21 locus ...................................................... 41

Figure 5. Design of an invertible Twister element for targeted degradation of an exogenous LincRNA-p21 transcript .................................................................................................................. 52

Figure 6. The invertible Twister element destabilizes the exogenous LincRNA-p21 construct .......... 54

Figure 7. The LincRNA-p21 p53RE reinforces p21 expression in a p53-dependent manner in MEFs .............................................................................................................................................. 58

Figure 8. The LincRNA-p21 p53RE does not reinforce p21 expression in a lung adenocarcinoma cell line ........................................................................................................................................ 61

Figure 9. Generation and identification of novel LincRNA-p21 knockdown alleles ......................... 63

Figure 10. Validation of four novel LincRNA-p21 knockdown alleles .............................................. 64

Figure 11. Novel genetic tools abolish RNA accumulation and attenuate transcription and processing at the LincRNA-p21 locus ................................................................................................................. 65

Figure 12. The PAS and Twister mediate the loss of LincRNA-p21 RNA in vivo ............................... 67

Figure 13. The Twister ribozyme attenuates transcription across the LincRNA-p21 locus ............. 68

Figure 14. Single-molecule RNA fluorescence in situ hybridization (smRNA FISH) confirms the loss of RNA signal in novel LincRNA-p21 loss-of-function mutants ........................................................................ 69

Figure 15. p21 does not require splicing or accumulation of the full-length LincRNA-p21 transcript ............................................................................................................................................. 71
Figure 16. Twister does not affect p21 levels in vivo ................................................................. 72

Figure 17. ASO gapmers targeting downstream of the PAS/Twister insertion site abrogate p21 expression in LincRNA-p21PAS/PAS and LincRNA-p21TWI/TWI MEFs ......................................................... 74

Figure 18. Conserved sequence elements within LincRNA-p21 exon 1 reinforce p21 expression..... 76

Figure 19. Stable looping interactions juxtapose the conserved region with the p21 promoter ....... 78

Figure 20. Epigenetic silencing of LincRNA-p21 does not affect p21 levels ..................................... 80

Figure 21. LincRNA-p21 activation promotes p21 expression in ARE p53-restorable MEFs .......... 82

Figure 22. LincRNA-p21 transcriptional inhibition correlates with a reduction in p21 levels....... 84

Table 2.1. Reported roles of LincRNA-p21 in homeostasis and disease ........................................ 47

Supplementary Table 1. Primers and oligos ................................................................................ 106

Supplementary Table 2. Key plasmids used in this work ............................................................... 110
Chapter I: Introduction
The role of noncoding transcription in gene regulation

Pervasive noncoding transcription as a hallmark of mammalian genomes

The advent of modern RNA sequencing technology has given rise to the discovery that mammalian genomes are transcribed into a wealth of RNAs. Although individual cells express no more than 40% of the genome at any one point in time (Djebali et al., 2012), anywhere from 70-90% of all nucleotides are cumulatively transcribed across different cell lines and under disparate environmental conditions (Birney et al., 2007; Djebali et al., 2012). Curiously, protein-coding genes account for only a small fraction of this pervasive transcription: in total, exons make up less than 2% of the genome (International Human Genome Sequencing Consortium, 2004). It therefore follows that the vast majority of transcripts are noncoding.

The noncoding transcriptome encompasses numerous classes of RNAs with diverse mechanisms of biogenesis and function. Historically, efforts to study noncoding transcription have focused on well-defined RNA classes associated with translation (tRNAs and rRNAs), splicing (snRNAs), rRNA biogenesis (snoRNAs), and post-transcriptional gene regulation (miRNAs and siRNAs). However, an expanding number of studies have implicated additional classes of noncoding RNAs (ncRNAs) in cellular processes ranging from genomic imprinting to the development of drug resistance in cancer cells (Joung et al., 2017; Latos et al., 2012).

An emerging class of transcripts known as long ncRNAs (lncRNAs) are operationally defined by their length (>200 nucleotides) and lack of coding potential. The largely arbitrary size threshold serves as a heuristic for distinguishing between IncRNAs and ncRNAs from other functional classes (Quek et al., 2015). The absence of coding potential may be supported by several lines of computational and experimental evidence, including poor conservation of putative open reading frames (ORFs), failure to detect ribosomal interactions, and the inability of mass spectrometry to detect corresponding peptides (Bánfai et al., 2012). Recent studies show that the human genome harbors over 60,000 IncRNAs, more
than double the number of annotated protein-coding genes (Fang et al., 2018; Iyer et al., 2015). Individual members of this transcript class demonstrate varying degrees of RNA processing, including capping, splicing, and polyadenylation, (Djebali et al., 2012) and may localize to either the nucleus or the cytosol (Kapranov et al., 2007), although the majority of lncRNAs remain within the nucleus (Derrien et al., 2012; Djebali et al., 2012). LncRNAs therefore represent a heterogeneous class of transcripts exhibiting variable processing, stability, and localization.

LncRNA loci also show significant diversity in their genomic position and organization with respect to protein coding genes (Fig. 1)(Derrien et al., 2012). A subset of lncRNAs known as long intergenic ncRNAs (lincRNAs) derive from transcriptional units that do not overlap with annotated protein-coding genes or regulatory sequences. Like their coding counterparts, many lincRNAs exhibit the epigenetic hallmarks of active RNA polymerase II (Pol II) transcription, including peaks of promoter-associated H3K4me3 and enrichment of H3K36me3 along their gene bodies (Guttman et al., 2009). On account of this, and because they arise from discrete loci that do not share their sequence with other transcripts, lincRNAs have attracted significant interest as potential mediators of key phenomena in both the nucleus and cytoplasm.

A second category comprises lncRNA loci that overlap partially or entirely with protein-coding genes. This broad class includes lncRNAs transcribed from the sense (s) or antisense (as) strand of the associated coding gene(s). Although s-lncRNAs have received relatively little attention, as-lncRNAs have been widely studied on account of their intriguing genomic architecture. In particular, the spatial and sequential relationship between as-lncRNAs and their overlapping loci suggests that these transcripts may regulate mRNA transcription, stability, or function (Faghihi et al., 2010; Zucchelli et al., 2015). As a testament to the prevalence of as-transcription—and to the potential importance of as-lncRNAs as a genome-wide regulatory mechanism—more than 60% of coding genes are associated with antisense transcripts (Katayama et al., 2005).
Although many s- and as-lncRNAs overlap exons, a subset of these lncRNA loci lie entirely within introns. Like as-lncRNAs, intronic RNAs exhibit a high genomic prevalence, with an estimated 80% of human genes harboring an intronic IncRNA (Louro et al., 2008). Unlike other lncRNAs, however,
the majority of intronic lncRNAs consist of only a single exon (Derrien et al., 2012; Nakaya et al., 2007). LncRNAs from this functional class have been proposed to act to regulate the splicing and stability of their host transcripts (Nakaya et al., 2007).

LncRNAs may also arise as the result of divergent transcription from the promoters of protein-coding genes. Bidirectional transcription is a hallmark of mammalian promoters: nearly 80% of active human promoters show evidence of divergent transcription within a kilobase of their transcription start site (TSS), although divergent transcription most commonly initiates within a 90-120 nucleotide window (Core et al., 2014). The amount of transcriptional initiation at divergent TSSs commonly correlates with the activity of the associated gene (Core et al., 2008), prompting suggestions that divergently-transcribed lncRNAs may promote protein-coding gene expression (Seila et al., 2009). Although some of these lncRNAs are stable, the majority are rapidly degraded by the exosome (Pefanis et al., 2015; Preker et al., 2008), potentially arguing that the RNA products of these loci are largely non-functional.

Enhancer RNAs (eRNAs) comprise the final class of lncRNAs. Over the past decade, a number of transcriptomics studies have confirmed the finding that many enhancers are transcriptionally active (Kim et al., 2010; Santa et al., 2010). Such enhancers are often bidirectionally transcribed by Pol II to yield transcripts of variable lengths; although some enhancer-derived transcribed transcripts fall well short of the arbitrary 200 nucleotide threshold for lncRNA designation (Henriques et al., 2018), others may be up to several kilobases in length (Djebali et al., 2012; Santa et al., 2010). Other features also mark eRNAs as a unique class of lncRNAs. While eRNAs are typically capped, they remain largely unspliced and non-polyadenylated (Andersson et al., 2014; Kim et al., 2010). And like lncRNAs arising from bidirectional promoters, they are often substrates for the RNA exosome (Pefanis et al., 2015). On account of this, eRNAs generally have short half-lives and a low cellular abundance (Li et al., 2013b), calling into question their functionality.

Although transcriptional units from each of these categories have different genomic organizations, lncRNA loci share a few structural commonalities. Firstly, many lncRNAs are located near protein-coding genes that play key roles in transcriptional regulation and development (Cabili et al., 2011;
Guttman et al., 2009). And although many lncRNAs show poor evolutionary conservation across their gene bodies, lncRNA promoters show a similar degree of conservation to the promoters of protein-coding genes (Carninci et al., 2005; Derrien et al., 2012; Guttman et al., 2009). Lastly, many lncRNA loci exhibit preserved synteny with nearby protein-coding genes (Cabili et al., 2011), hinting at the importance of the spatial relationship between the two loci.

As demonstrated above, lncRNAs arise from diverse loci across mammalian genomes. However, the broader significance of this pervasive noncoding transcription remains unclear, in part because our ability to detect noncoding transcripts far outstrips our ability to assess their function. Initial efforts to identify and characterize lncRNAs focused primarily on lincRNAs, largely on account of their similarity to protein-coding genes (Guttman et al., 2009; Wang et al., 2011). Many of these early studies explored the ability of these lncRNA transcripts to act in trans to effect diverse biological outcomes throughout the nucleus and cytoplasm. Intriguingly, recent studies have shown that many lncRNAs exhibit significant chromatin enrichment, remain at or near the site of their transcription, and are co-regulated with nearby protein-coding genes (Werner and Ruthenburg, 2015; Werner et al., 2017). This finding raises the possibility that lncRNAs may act in cis to activate or repress neighboring genes. The following work describes our efforts to investigate the role of the noncoding transcriptome in cis-regulation.

The role of cis-acting noncoding RNAs in homeostasis and disease

As deep sequencing transcriptomics studies have shown, mammalian genomes are pervasively transcribed, yielding thousands of lncRNAs. While some of these transcripts are expressed ubiquitously, many lncRNAs show remarkable tissue, cell type, and disease state specificity (Derrien et al., 2012; Djebari et al., 2012; Iyer et al., 2015). This correlation between lncRNA expression and biological outcome raises the possibility that cis-acting lncRNAs might play critical roles as mediators of development, homeostasis, and disease, and implicated them as a potential source of novel therapeutic targets. Ultimately, large-scale efforts to identify functionally-relevant lncRNAs demonstrate that these loci and their transcripts may enact previously-unappreciated roles in many biological processes. In the
following section, we will highlight some of the cellular and organismic outcomes of \textit{cis}-regulation and -dysregulation by functional noncoding transcriptional units.

\textit{Dosage compensation}

A number of \textit{cis}-acting lncRNAs have been shown to play key roles in X-chromosomal dosage compensation, the process by which eutherian mammals equalize the expression of X-linked genes between the sexes. Foremost among these is \textit{Xist} (X-inactive specific transcript), one of the first lncRNAs to be described (Brown et al., 1991). This X-chromosomal transcript randomly envelopes one of the two X chromosomes in female cells, thereby inducing its heterochromatinization and subsequent silencing of the inactive X chromosome (XI). Although \textit{Xist} expression is essential for establishing X-chromosomal inactivation (XCI) during early embryonic development, its role in XI maintenance remains unclear. Studies have variously reported that the post-XCI loss of \textit{Xist} expression leads to consequences ranging from a mild decrease in viability to a predisposition for certain types of cancer (Yang et al., 2016a; Yildirim et al., 2013). As a testament to the importance of proper \textit{Xist} expression, an additional suite of X-chromosomal lncRNAs work in tandem with \textit{Xist} to support robust XI silencing and protect the active X chromosome (XA; reviewed in Furlan and Rougeulle, 2016). For example, \textit{Jpx} and \textit{Ftx} (five prime to \textit{Xist}), which both escape transcriptional silencing on the XI, appear to reinforce \textit{Xist} expression (Chureau et al., 2011; Tian et al., 2010). Meanwhile, \textit{Tsix} (a lncRNA antisense to \textit{Xist}) and \textit{Xite} both repress \textit{Xist} expression from the XA, thereby ensuring transcriptional silencing of a single X chromosome (Furlan and Rougeulle, 2016). Thus, a network of \textit{cis}-regulatory lncRNAs coordinates dosage compensation in female eutherian cells.

\textit{Imprinting}

\textit{Cis}-acting lncRNAs also play a well-defined role in the establishment and maintenance of epigenetic imprinting at loci throughout the genome. \textit{Airn} (Antisense of Igf2r non-protein-coding RNA), an overlapping as-lncRNA transcribed from the paternal \textit{Igf2r/Airn} locus, silences \textit{Igf2r} (insulin-like
growth factor 2 receptor) in embryonic, extraembryonic, and adult mouse tissue and may additionally repress the nearby *Slc22a2* and *Slc22a3*, (solute carrier 22 member 2 and 3, respectively) genes in some extraembryonic lineages (Latos et al., 2012; Sleutels et al., 2002). Similarly, the as-lncRNA *Kcnq1ot1*, which overlaps the *Kcnq1* (potassium voltage-gated channel subfamily Q member 1) locus, establishes a heterochromatic environment in the paternal *Kcnq1* domain (Pandey et al., 2008). The human and mouse genomes are replete with additional examples of as-lncRNAs that overlap imprinted loci (reviewed in O’Neill, 2005), implicating *cis*-acting lncRNAs as a common mechanism for enacting allele-specific gene expression. Significantly, the loss of genetic imprinting underlies numerous genetic disorders (O’Neill, 2005), suggesting that imprinting-associated lncRNAs such as *Airn* and *Kcnq1ot1* play critical roles during development.

**Cardiac development**

While *Xist* and imprinting-associated lncRNAs enforce allele-specific transcriptional control, other lncRNAs are expressed biallelically and thus play more general roles as activators or repressors of nearby developmental genes. In particular, a large body of literature has illuminated the importance of *cis*-regulatory lncRNAs during cardiac development. The *Hand2* (heart and neural crest derivatives expressed 2) locus is flanked by two lncRNAs, *Upperhand* and *Handsdown* (also known as *Handlr*), whose antagonistic effects maintain *Hand2* expression within a narrow operational window (Anderson et al., 2016; Ritter et al., 2019). While transcription through the *Upperhand* locus establishes a permissive chromatin environment at the *Hand2* promoter, thereby facilitating its transcription, transcription of the *Handsdown* locus sequesters enhancer elements that would otherwise reinforce *Hand2* binding. *Hand2* deficiency is lethal in mice (Anderson et al., 2016), and overexpression of *Hand2* in humans incurs significant developmental defects (Tamura et al., 2013), indicating the importance of maintaining *Hand2* levels within a narrow operational range. The relationship between *Hand* and its neighboring lncRNAs thus highlights the ability of *cis*-acting lncRNAs to “buffer” genes whose over- and underexpression may both have deleterious consequences. As further evidence of their broad functionality, *cis*-acting lncRNAs
have been proposed to regulate additional aspects of heart development. *Charme* and *Fendrr* have been implicated in cardiac myogenesis and mesoderm differentiation, respectively (Ballarino et al., 2018; Grote et al., 2013).

**Immune function**

*Cis*-regulatory lncRNAs have increasingly been recognized as an important component of the immune system, where they underpin processes ranging from cell lineage commitment to immune activation in response to pathogenic threats. As an initial example of the former, the lncRNA *ThymoD* (thymocyte differentiation factor) remolds the local chromatin structure, facilitating interactions between *Bcl11b* (B-cell lymphoma/leukemia 11b) and its intergenic enhancers during T-cell development (Isoda et al., 2017). Strikingly, transcriptional knockdown of this lncRNA predisposes mice to develop lymphoma or leukemia. Dysregulation of *Morbidd* (myeloid RNA regulator of Bim-induced death) has similarly dramatic consequences. This lncRNA regulates the lifespan of myeloid cells by repressing expression of the pro-apoptotic gene *Bim* (also known as *Bcl2l11*) (Kotzin et al., 2016). Notably, knockdown of the *Morbidd* transcript or transcriptional process leads to a significant increase in cell mortality in vitro, while mice deficient for the *Morbidd* locus have fewer eosinophils, neutrophils, and monocytes. In light of its role in myeloid cell development, it is unsurprising that *Morbidd* expression is dysregulated in various human hematological disorders (Kotzin et al., 2016). Together, *Morbidd* and *ThymoD* demonstrate the role of lncRNAs in specifying cell fate during immune cell development.

Many examples attest to the ability of lncRNAs to act in *cis* to regulate transient immune responses. Although rapid and robust immune activation is critical for eliminating threats, immune signaling must be tightly regulated. Recent studies have identified a number of lncRNAs that either enhance gene expression during the early stages of the immune response or help dampen the magnitude of signaling at later timepoints following an immune challenge. For example, the lncRNAs *Pacer* and *LincRNA-Cox2* both regulate the expression of *Ptgs2/Cox2* (prostaglandin endoperoxide synthase 2, also known as cyclooxygenase 2) in response to inflammatory stimuli (Elling et al., 2018; Krawczyk and
Meanwhile, the lncRNA genes *Umlilo* (upstream master lncRNA of the inflammatory chemokine locus) and *Ifng-as1* (interferon gamma antisense transcript 1) have been proposed to prime nearby loci for transcription by modulating the chromatin environment (Fanucchi et al., 2019; Petermann et al., 2019); and while both loci mediate the initial immune response to their respective stimuli, *Ifng-as1* plays an additional role as a long-term potentiator of interferon-γ signaling, which it accomplishes by maintaining a poised epigenetic state around the *Ifng* (interferon gamma) locus (Petermann et al., 2019).

In contrast to *Umlilo* and *Ifng-as1*, which serve as an example of how *cis*-regulatory lncRNAs can support rapid gene expression, other lncRNAs may suppress the immune response. For example, the *Tnf* (tumor necrosis factor) locus harbors numerous lncRNAs that appear to act in tandem to repress the TNF signaling pathway (Shi et al., 2014). *Cis*-acting lncRNAs may thus act to both bolster and suppress the expression of immune genes.

**The p53 tumor suppressor network**

The p53 tumor suppressor network includes many lncRNAs that are induced in response to p53 activation (see Chapter II)(Melo et al., 2013; Sánchez et al., 2014; Tesfaye et al.). Recently, our lab showed that a handful of these lncRNAs are co-regulated with nearby protein-coding genes, suggesting that they mediate p53-dependent gene expression (Tesfaye et al.). Indeed, *cis*-regulatory lncRNAs act as effectors of the p53 pathway under a range of conditions. For example, *LincRNA-p21* reinforces expression of the cell cycle checkpoint inhibitor *p21/Cdkn1a* (cyclin-dependent kinase inhibitor 1a) in response to p53 signaling to promote growth arrest (see Chapters II and IV)(Dimitrova et al., 2014). Similarly, the intergenic lncRNA *LincRNA-Gadd45γ* has been shown to activate the poorly-characterized *Gadd45γ* (growth arrest and DNA damage-inducible 45 gamma) gene (Tesfaye et al.). Although other members of the *Gadd45* gene family play well-established roles in the p53 pathway (Hollander et al., 1993), *Gadd45γ* has not been identified as a direct p53 target. The ability of *LincRNA-Gadd45γ* to activate *Gadd45γ* transcription in a p53-dependent manner provides provocative evidence that *cis*-acting lncRNAs may recruit noncanonical targets into the p53 transcriptional network. Finally, a p53-inducible
isoform of the lncRNA *Pvt1* (plasmacytoma variant 1) mediates crosstalk between the p53 and Myc (myelocytomatosis) networks by downregulating *Myc* transcription in response to p53 activation, thereby acting as a check on cell proliferation following DNA damage or oncogenic signaling (Olivero et al., 2020). Taken together, these examples showcase the potential of *cis*-regulatory lncRNAs to enact precise outcomes in response to p53 signaling.

*Cancer*

As evidenced by the preceding examples, *cis*-acting lncRNAs serve as an additional layer of gene regulation, allowing cells to maintain precise spatiotemporal control of transcription. It is therefore unsurprising that lncRNAs have been implicated in many different types of cancer (Olivero and Dimitrova, 2020). Some tumorigenic lncRNAs appear to act by repressing the activity of nearby tumor suppressors. *Anril* (antisense noncoding RNA in the Ink4 locus), an as-lncRNA overlapping the *Ink4b/Arf/Ink4a* locus, is one such lncRNA that has received significant attention. Experimental findings suggest that *Anril* transcriptionally silences the *Ink4b/Arf/Ink4a* locus, which encodes the tumor suppressor proteins CDKN2A, ARF, and CDKN2B (Yap et al., 2010), thus accounting for the observation that *Anril* overexpression promotes cell proliferation, migration, and invasion (Kong et al., 2018). Clinical evidence further shows that increased *Anril* expression is associated with an aggressive disease progression and poor patient survival. Transcription of another *cis*-regulatory as-lncRNA, *Anrassf1*, negatively regulates the expression of the overlapping, tumor-suppressive *Rassf1* (RAS-associated domain family member 1a) gene. Consistent with this, *Anrassf1* is overexpressed in both breast and prostate cancer cell lines relative to non-tumor controls and is inversely correlated with *Rassf1* levels (Beckedorff et al., 2013).

In keeping with previous examples showing that *cis*-lncRNAs may enact both positive and negative regulation, cancer-associated lncRNAs may also enhance the expression of nearby oncogenes. As one example, *CCAT1-L* (colorectal cancer-associated transcript 1 long isoform), which is transcribed from a super-enhancer upstream of the *Myc* locus, facilitates long-range interactions between *Myc* and
nearby regulatory DNA elements. This, in turn, sustains Myc expression in human colorectal cancer cells (Xiang et al., 2014). A second example, the lncRNA SWINGN, (SWI/SNF interacting Gas6 enhancer noncoding RNA), boosts the transcription of Gas6 (growth arrest specific 6), promoting cellular proliferation and delaying the onset of oncogene-induced senescence (Grossi et al., 2020).

It is important to explicitly note the distinction between a lncRNA’s endogenous function and its pathological role in diseases such as cancer, even in cases where the normal physiological role remains elusive. For example, the Anril locus harbors many cancer-associated SNPs that may alter its structure, abundance, and function (Kong et al., 2018), perhaps accounting for its frequent dysregulation in different types of cancer. The widespread cooption of functional lncRNAs during tumorigenesis serves as further evidence for the involvement of noncoding transcription in vital cellular processes.

Conclusion

Two common observations arise from genome-wide analyses of the noncoding transcriptome. Firstly, lncRNAs are enriched near genes encoding transcription factors and key developmental genes (Mercer et al., 2008; Ulitsky et al., 2011; Wamstad et al., 2012). Secondly, the expression of many chromatin-associated lncRNAs typically correlates with levels of nearby protein-coding genes (Werner et al., 2017). The discovery that hundreds of cis-regulatory lncRNAs coordinate gene expression across a broad range of biological phenomena provides context for these relationships and further helps explain why lncRNAs and protein-coding gene pairs typically exhibit conserved synteny (Ulitsky et al., 2011). Although cis-lncRNAs perform a wide range of roles, functional studies all point to one main conclusion: the unique spatial relationship between lncRNAs and their neighboring genes enables cis-regulatory lncRNAs to enact rapid and robust activation or repression.

Mechanisms of cis-regulation by noncoding RNAs

Functional studies attest to the widespread involvement of lncRNAs in cis-regulation. However, the mechanisms by which lncRNA loci enact cis-regulation are often difficult to identify. As a first
possibility, the lncRNA molecule may itself harbor functional activity. As a second possibility, the act of lncRNA transcription or processing may effect gene regulation independently from the resulting transcript. Finally, many purportedly-functional lncRNAs arise from enhancers, promoters, and other regulatory sequences (Core et al., 2008), giving rise to the prospect that DNA elements within a lncRNA locus may be sufficient to drive expression of both the lncRNA and its neighboring gene(s). In this case, lncRNA transcription and accumulation on the chromatin would represent an incidental byproduct of underlying regulatory processes. Efforts to dissect lncRNA function have implicated these three broad models as overarching paradigms of cis-regulation. Below, we review the many molecular mechanisms through which lncRNAs enact each of these paradigms, with attention where necessary to the methodologies used to demonstrate each function.

The lncRNA molecule

Classical models of lncRNA functionality have emphasized the ability of these RNAs to form complex secondary and tertiary structures capable of interfacing with a wide range of biomolecules (Quinn and Chang, 2016). While this mechanistic paradigm inherently encompasses all trans-acting lncRNAs, it remains unclear whether it is broadly applicable to cis-acting lncRNAs. Nevertheless, there is compelling evidence that a number of lncRNAs enact cis-regulation through their transcripts.

Xist is, perhaps, the archetypal example of a functional cis-regulatory lncRNA molecule. The 17 kilobase Xist transcript comprises many well-studied RNA domains that interact with more than 80 protein factors to enact XCI (Minajigi et al., 2015, reviewed in Pintacuda et al., 2017). Although the mechanistic importance of various Xist-protein interactions at different points during XCI has not been fully resolved, studies have revealed major insights into underlying principles. Firstly, XCI appears to involve several different Xist-protein interactions that mediate distinct outcomes. For example, interactions between Xist and the nuclear lamina protein LBR (Lamin B receptor) appear to target the XI to the nuclear periphery, which in turn facilitates the spread of Xist across the chromatin (Chen et al., 2016). The bivalent DNA- and RNA-binding protein SAF-A (also known as hnRNP-U) and other
DNA/RNA-binding proteins further promote XCI by tethering Xist to the chromatin; this provides an obvious mechanism for the retention and subsequent accumulation of the Xist cloud during the initial stages of transcriptional silencing (Chen et al., 2016; Hasegawa et al., 2010; Jeon and Lee, 2011).

Additional interaction partners such as SHARP (also known as SPEN) bind to the transcript and recruit the histone deacetylase HDAC3, which repressively modifies the chromatin in order to exclude Pol II from the silenced region of the XI (McHugh et al., 2015). Lastly, the Polycomb repressive complexes 1 and 2 (PRC1 and PRC2) have been the subject of ongoing investigation on account of their proposed role in establishing and maintaining a repressive chromatin environment. However, recent work has called into question the specificity and functionality of Xist-PRC1/2 interactions (discussed below) (Brockdorff, 2017). As evidenced by these examples, the complex interplay between Xist and its myriad protein factors enacts multiple functions during XCI. Xist therefore serves as a model for how lncRNAs can act as modular platforms capable of coordinating multi-step, multimeric interactions.

As elegantly illustrated by Xist, chromatin remodeling is a recurring mechanistic theme of lncRNA-mediated cis-regulation. A growing number of functional cis-acting lncRNAs have been shown to interact with chromatin-modifying factors to activate or silence nearby genes. For example, the lncRNA SWINGN associates with the SWI/SNF chromatin remodeling complex, which promotes local chromatin accessibility, resulting in activation of the nearby proto-oncogene Gas6 (Grossi et al., 2020).

Similarly, the LncTCF7 RNA binds SWI/SNF components through a stem-loop structure in its 3’ end, subsequently increasing the occupancy of this protein complex at the neighboring TCF7 promoter (Wang et al., 2015b) and enhancing its expression. As a final example, the Airn lncRNA molecule has been proposed to enact epigenetic silencing of the non-overlapping, paternally-imprinted Slc22a3 locus by recruiting the G9a histone methyltransferase. This leads to the deposition of repressive H3K9me3 marks around the paternal Slc22a3 locus (Nagano et al., 2008). Thus, many transcribed cis-regulatory loci appear to act through their transcripts to modulate the local epigenetic environment.

Functional lncRNA molecules may also establish and maintain the larger chromatin environment: transcripts may themselves act as bridging factors, helping to coordinate 3D chromatin contacts between
sequentially distant loci (Quinn and Chang, 2016). As a first example, the cancer-associated transcript $CCAT1-L$ has been proposed to strengthen pre-existing looping interactions between $Myc$ and its associated super-enhancers through its specific CTCF-binding activity (Xiang et al., 2014). Under this model, the $CCAT1-L$ RNA may act by increasing the local concentration of CTCF available to CTCF sites in the $Myc$ locus, facilitating the establishment and maintenance of chromatin looping. A second example, the lncRNA $Firre$ (functional intergenic repeating RNA element), has been proposed to act in concert with the nuclear matrix factor SAF-A to coordinate long-range chromatin contacts between the $Firre$ locus and several other trans-chromosomal loci (Hacisuleyman et al., 2014). Strikingly, deletion of the $Firre$ locus led to the dissolution of these interchromosomal contacts. While this provides dramatic evidence for the importance of the $Firre$ locus, it raises the question of whether the DNA sequence or act of transcription might instead mediate the role of the $Firre$ locus as a chromatin hub.

LncRNA molecules may also cooperate with the underlying DNA sequence to modulate transcription factor binding at regulatory elements. Many transcription factors exhibit both DNA- and RNA-binding activity (Cassiday and Maher III, 2002), giving rise to the possibility that they may bind both enhancer elements and their resulting transcripts. LncRNAs—and in particularly, eRNAs—may therefore increase the local concentration of these factors. As one example of this, one eRNA has been shown to bind the bivalent transcription factor YY1 ($Yin-Yang 1$), subsequently increasing its concentration at enhancers (Sigova et al., 2015). Other eRNAs have been shown to drive RNA-dependent reporter gene expression (Melo et al., 2013), lending additional support for a model in which RNA-protein interactions augment the activity of activating DNA elements. Alternatively, LncRNAs may compete with DNA elements to sequester transcription factors that either activate or repress the expression of neighboring genes. For example, the activating LncRNA $Pacer$ upregulates $Ptgs2$ in part by binding the repressive NF-KB subunit p50 (Krawczyk and Emerson, 2014). This interaction decreases the local concentration of p50 near the $Cox2$ promoter, subsequently relieving p50-mediated transcriptional repression.
Many of the preceding examples attest to the ability of sequential and structural elements within lncRNA molecules to specifically interact with target proteins. However, there is mounting evidence that functional transcripts may also operate in a sequence-independent manner. Numerous lncRNAs, including Anrassf1, Anril, and Morrbid, mediate local gene silencing by interacting with PRC1/2 (Beckedorff et al., 2013; Kotake et al., 2011; Kotzin et al., 2016). This, in turn, recruits chromatin-modifying complexes that catalyze the deposition of repressive H3K27me3 on target genes. The discovery that PRC2 often interacts with target RNAs in a sequence-independent and length-dependent manner has led some to question the biological significance of these interactions (Davidovich et al., 2013, 2015; Gil and Ulitsky, 2020). However, PRC2’s promiscuous RNA-binding activity does not preclude a functional role and may furthermore explain the prevalence of lncRNA-PRC2 interactions in both cis- and trans-repression (Khalil et al., 2009; Skourti-Stathaki et al., 2019). Taken together, these findings suggest that repressive lncRNA transcripts may enact their regulatory function through a mechanism that is largely independent of their underlying sequence, perhaps helping to explain the observation that lncRNAs typically show poor conservation (Carninci et al., 2005; Derrien et al., 2012).

Like their trans-acting counterparts, cis-regulatory RNAs have the potential to interact with a range of additional biomolecules, including DNA and other RNAs. Examples of lncRNA-DNA interactions abound. R-loops, triple-stranded structures characterized by the presence of a DNA-RNA hybrid, frequently form at lncRNA-protein-coding gene loci, where they have been proposed to serve important regulatory functions (reviewed in Niehrs and Luke, 2020). For example, R-loops involving Anrassf1 recruit PRC2 to the overlapping Rassf1a promoter, ultimately resulting in transcriptional silencing (Beckedorff et al., 2013). Although many R-loops appear to play repressive roles, often through interactions with PRC2 (Skourti-Stathaki et al., 2019), R-loops may also mediate transcriptional activation. The lncRNA Vim-as1, which arises from the Vim (vimentin) locus, forms R-loops that support local chromatin decondensation and permit NF-KB binding at the Vim promoter (Boque-Sastre et al., 2015). As a second mechanism of action, cis-acting lncRNAs may also interact with double-stranded DNA in order to form a triple helix. The human as-lncRNA Khps1, which overlaps the coregulated Sphk1
(sphingosine kinase 1) locus, remains tethered to its locus by its two triplex-forming regions (TFRs) (Postepska-Igielska et al., 2015). These homopurine tracts, which allow the DNA duplex and Khps1 RNA to engage in Hoogsteen base pairing, serve as an “address label” for this RNA, which also recruits the histone acetyltransferase p300/CBP. Thus, lncRNAs may act in part through interactions with the underlying DNA sequence.

While DNA-RNA hybrids appear to be a common theme in lncRNA-mediated cis-regulation, interactions between cis-regulatory lncRNAs and other RNAs are exceedingly rare. As evidence of this, one analysis of the human noncoding transcriptome found that less than 1% of lncRNA-mRNA or lncRNA-pre-mRNA interactions involve cis-lncRNAs (Szcześniak and Makalowska, 2016). Although theoretically free to utilize diverse mechanisms, it appears that cis-regulatory lncRNAs overwhelmingly operate through associations with protein and DNA.

The transcriptional process

As demonstrated above, some lncRNA loci generate transcripts that enact cis-regulation through their sequence and structure. By contrast, other lncRNA loci appear to act through some aspect of the transcriptional process, including co-transcriptional phenomena. Understanding the mechanisms through which lncRNA biogenesis coordinates local gene expression will help illustrate the significance of noncoding transcription throughout the genome.

Transcription:

The intimate spatial relationship between cis-acting lncRNAs and their nearby protein-coding genes raises the possibility that lncRNAs may act by modulating the local epigenetic and transcriptional landscape. Indeed, lncRNAs have been shown to enact both negative and positive gene regulation through their transcriptional process.

Imprinting-associated lncRNAs, including Airn and Kcn1qot1, serve as a model for the ability of lncRNA transcription to enact gene silencing through one of several mechanisms. The 118 kilobase Airn
lncRNA initiates transcription at an antisense promoter within the \( Igf2r \) locus (Lyle et al., 2000). Notably, the insertion of a PAS element within the shared \( Airn/Igf2r \) promoter region abolished \( Airn \) transcription and incurred a near-complete loss of \( Igf2r, Slc22a2, \) and \( Slc22a3 \) silencing, indicating a functional requirement for \( Airn \) transcription or the accumulation of the \( Airn \) transcript (Sleutels et al., 2002). Subsequent work, in which PAS elements were inserted at multiple sites within the \( Airn \) locus, showed that transcription of \( Airn \) through the \( Igf2r \) promoter is both necessary and sufficient to establish repressive methylation of the paternal \( Igf2r \) allele (Latos et al., 2012). Intriguingly, promoter methylation and \( Airn \) transcription are independently able to silence \( Igf2r \) expression (Santoro et al., 2013), demonstrating that lncRNAs may enact cis-regulation through multiple modes of action. This observation additionally highlights the mechanistic importance of transcriptional interference, in which transcription on one strand sterically hinders transcription on the opposite strand.

While it is therefore clear that \( Airn \) silences \( Igf2r \) through its transcriptional process, the mechanism of repression for \( Slc22a2 \) and \( Slc22a3 \) remains elusive. Although genetic imprinting of \( Slc22a2 \) and \( Slc22a3 \) is not predicated on transcriptional overlap, premature \( Airn \) termination disrupts silencing at both loci (Sleutels et al., 2002). A couple of explanations may account for this puzzling observation. As described above, \( Airn \) has been proposed to repress \( Slc22a3 \) through an RNA-based mechanism (Nagano et al., 2008). Alternatively, repressive methylation of the \( Igf2r \) promoter may initiate heterochromatinization of the entire locus, including \( Slc22a2 \) and \( Slc22a3 \) (Sleutels et al., 2002). This has been documented at other imprinted loci, including the \( Kcnq1ot1 \) locus, implicating it as a viable mechanism that lncRNAs can use to silence more distal genes (Thakur et al., 2003). The full extent to which transcription through the \( Airn \) locus enacts silencing thus remains unresolved.

The lncRNA transcriptional process may also modulate the transcriptional and epigenetic environment in order to enact positive regulation of nearby protein-coding genes. In the simplest examples of this phenomenon, transcriptional initiation from or elongation through regulatory elements such as enhancers establishes or maintains a permissive chromatin environment. This is the case with the previously-described lncRNA \( Upperhand \). This lncRNA, which shares a bidirectional promoter with the
cardiac transcription factor gene *Hand2*, spans a 16.5 kilobase locus that includes a known super-enhancer (Anderson et al., 2016). Premature termination of *Upperhand* transcription upstream of this enhancer element decreased the prevalence of H3K4me1 and H3K27ac and reduced the occupancy of the transcription factor Gata4 within this region. Concurrently with this, the loss of *Upperhand* transcription mediated a reduction in Pol II occupancy throughout the gene body—but not promoter—of *Hand2*. Knockdown of the *Upperhand* transcript with antisense oligonucleotides (ASOs) did not affect *Hand2* levels, implicating a model in which the *Upperhand* transcriptional process licenses the associated super-enhancer to promote *Hand2* expression.

In some cases, the act of lncRNA transcription appears to promote the expression of protein-coding genes that are separated by the lncRNA by considerable genomic distances. Transcription through the 11 kilobase locus of the intergenic lncRNA *Maenli* (master activator of engrailed-1 in the limb) has been shown to drive expression of the *En-1* (engrailed 1) gene more than 250 kilobases away (Allou et al., 2021). Curiously, although insertion of a PAS that abolished ~90% of *Maenli* transcription led to a ~90% decrease in *En-1* RNA levels, the insertion of a GFP-PAS element at the same site led to a more moderate reduction in *En-1* expression. Given that this element inhibited the majority *Maenli* transcription while permitting transcription through the ~800 nucleotide GFP reporter, this observation suggests that transcriptional initiation from the *Maenli* locus promotes *En-1* expression in a length-dependent, sequence-independent manner. In further support of this, the deletion of sequences downstream of the PAS insertion site had no effect on *En-1* levels. While it remains unclear how *Maenli* transmits information about its transcriptional state to the *En-1* locus, reported looping interactions between the two genes provide a physical basis for communication.

The act of transcription may even lead to large-scale changes in chromatin organization and nuclear positioning. Indeed, one study showed that antisense transcription through alternate promoters within the *Pcdhα* (protocadherin alpha) locus establishes enhancer/promoter looping interactions that, in turn, drive sense transcription from an individual *Pcdhα* promoter (Canzio et al., 2019). Critically, antisense transcription promotes CpG demethylation at promoter-proximal CTCF sites, enabling CTCF to
bind and form de novo chromatin loops. Additional studies attest to the prevalence of this cis-regulatory mechanism and further reveal the extent to which lncRNAs may regulate nuclear architecture. In addition to coordinating interactions between the Bcl11b locus and its intergenic enhancers, ThymoD also repositions the Bcl11b locus from the lamina to the nuclear interior (Isoda et al., 2017). Like antisense transcription through Pcdha promoters, ThymoD transcription mediates local CpG demethylation to permit CTCF binding and chromatin looping. This transcriptional and looping process additionally causes the entire locus to relocalize to the transcriptionally permissive euchromatic compartment, although the precise molecular mechanism underlying this repositioning remains unclear.

RNA processing events

Many lncRNAs are extensively processed, suggesting that lncRNA processing—and not merely transcription—could play a regulatory role. Indeed, several lncRNAs have a mechanistic requirement for processing events, including both splicing and 3’ end formation.

Although lncRNAs exhibit little overall conservation, the widespread conservation of lncRNA splicing enhancers implicates splicing as a potential mediator of lncRNA functionality (Haerty and Ponting, 2015). Several lines of evidence establish the fundamental plausibility of this mechanism. One study found that deletion of the first 5’ splice site of the murine lncRNA Blustr (bivalent locus (Sfmbt2) is regulated by the splicing and transcription of an RNA) led to a significant reduction in the expression of the nearby protein-coding gene Sfmbt2 (SCM-containing gene with four MBT domains 2)(Engreitz et al., 2016). Curiously, abolishing Blustr splicing at downstream splice sites did not appear to affect Sfmbt2 transcription, pointing to a critical requirement for TSS-proximal splicing. Additional work has shown that active enhancers often yield spliced lncRNAs, suggesting a positive correlation between splicing and enhancer activity (Gil and Ulitsky, 2018). While these reports provide exciting evidence that lncRNA splicing can play an important role in cis-regulation, the underlying mechanism is opaque. TSS-proximal splicing has previously been shown to enhance gene expression by as much as 100-fold (Brinster et al., 1988; Fong and Zhou, 2001), likely because associations between the U snRNP and general transcription
factors increase the local concentration of these factors at the lncRNA promoter, increasing their availability for nearby protein-coding genes (Fong and Zhou, 2001). In light of this, it is unclear whether Sfmbt2 requires splicing per se or whether the requirement for splicing reflects a more general requirement for transcription from the Blustr promoter.

Other features of lncRNA biogenesis have also been shown to contribute to cis-regulation. The intergenic lncRNA A-rod (activating regulator of Dkk1) and the nearby protein-coding gene Dkk1 (Dickkopf homolog 1) serve as an example for how cleavage and polyadenylation may mediate lncRNA function (Ntini et al., 2018). Although A-rod transcription enhances Dkk1 expression, Dkk1 does not require nascent A-rod, which remains tethered to the chromatin. Instead, the dissociation of spliced and polyadenylated A-rod from the chromatin permits it to exploit pre-existing chromatin interactions and localize to the spatially-proximal Dkk1 locus, where it recruits the general transcription factor EBP1. As illustrated by this example, processes such as 3’ end formation may regulate the accessibility of lncRNAs to both protein interaction partners and other loci, thereby acting as another mechanism of cis-regulation (Gil and Ulitsky, 2020).

DNA elements

The genome harbors myriad cis-regulatory elements, including both enhancers and promoters, that are characterized by features such transcription factor binding sites and a low nucleosome occupancy (Melé et al., 2017). The observation that Pol II readily initiates transcription from regions of open chromatin has given rise to the hypothesis that most noncoding transcriptional events are noise and generate non-functional “junk” RNAs (Kung et al., 2013; Struhl, 2007). Under this paradigm, DNA sequence elements within a functional lncRNA locus may be wholly sufficient to account for local regulatory activity; in this case, the transcriptional process and presence of a discrete lncRNA are biological “red herrings” with no functional relevance to cis-regulation at the locus.

Several studies compellingly show the importance of properly evaluating this “null hypothesis.” The intergenic lncRNA LockD (lncRNA downstream of Cdkn1b) is transcribed and processed to yield a
434 nucleotide transcript that is expressed in many mouse tissues (Paralkar et al., 2016). Intriguingly, deletion of the nearly 20 kilobase locus incurred a significant decrease in the expression of the nearby protein-coding gene Cdkn1b (cyclin-dependent kinase inhibitor 1b), implicating LockD as a functional cis-regulatory IncRNA locus. However, the subsequent insertion of a polyadenylation signal (PAS), which induces premature transcriptional termination, into the first exon of LockD did not affect Cdkn1b levels in spite of the near-complete loss of spliced and total LockD expression. This demonstrates the absence of a requirement for transcription at this locus. Similarly, Bendr (Bend4 regulating effects not dependent on the RNA) transcription and IncRNA accumulation have been shown to be dispensable for regulation of the neighboring Bend4 (BEN domain-containing 4) gene (Engreitz et al., 2016). The deletion of the Bendr promoter—but not the loss of Bendr transcription following the insertion of a PAS sequence—attenuated transcription of Bend4. In both cases, the discordant effects of genetic deletions and premature transcriptional termination support the conclusion that the transcriptional process does not mediate cis-regulation at either of these loci.

Underlying DNA sequences thus appear to account for the cis-regulatory capacity of both of these IncRNA loci. The LockD promoter exhibits DNAsse hypersensitivity and binds numerous general transcription factors (Paralkar et al., 2016). Critically, the 20 kilobase LockD deletion, which spanned the promoter, abolished the ability of transcription factors to bind at this locus, likely accounting for its effects on Cdkn1b transcription. Although the Bendr promoter deletion spanned a more modest ~750 nucleotide region, it also encompassed several known transcription factor binding sites, indicating that it likely serves as an enhancer for Bend4 (Engreitz et al., 2016).

Taken together, these data suggest that DNA elements—and not the act of transcription or the activity of the RNA—enact cis-regulation at these two IncRNA loci. As an important note, although these loci appear to function as enhancer-like elements, neither exhibits the epigenetic signature of a “classical” enhancer (Calo and Wysocka, 2013). While enhancers are typically enriched for H3K4me1 relative to H3K4me3, the promoters of LockD and Bendr both have prominent H3K4me3 peaks. Moreover, while enhancer transcription is often bidirectional and commonly gives rise to unspliced, non-polyadenylated
transcripts (Andersson et al., 2014; Kim et al., 2010), the LockD and Bendr loci are both unidirectionally transcribed to yield processed RNAs (Engreitz et al., 2016; Paralkar et al., 2016). This raises the possibility that many lncRNA loci generate non-functional RNAs and instead act primarily as enhancer-like elements. The existence of such a class of transcribed DNA elements with noncanonical chromatin marks indicates that a gene-like epigenetic signature does not establish a functional requirement for transcription. These examples therefore illustrate the importance of developing new heuristics for distinguishing between enhancers and functional lncRNAs. Additionally, they demonstrate the necessity of properly addressing the possibility that DNA elements may be sufficient to account for the cis-regulatory activity of functional lncRNA loci.

**Conclusion**

As the preceding examples illustrate, cis-regulatory lncRNA loci act through a variety of mechanisms to modulate the expression of neighboring genes. Careful mechanistic studies have identified numerous examples of bona fide lncRNAs that act either post- or co-transcriptionally. The existence of these lncRNAs serves as tantalizing evidence that pervasive genomic transcription may represent more than transcriptional noise. However, the observation that many enhancer-like elements are incidentally transcribed to yield dispensable transcripts with many of the traditional hallmarks of functional lncRNAs demonstrates the importance of experimentally validating putative lncRNA loci. Future work will likely identify lncRNAs with additional modes of action. This will simultaneously expand our understanding of the role of noncoding transcription and challenge our ability to resolve new and complex mechanistic phenomena.
Tools for studying cis-acting noncoding RNAs

The unique relationship between cis-acting lncRNA loci and their gene products poses major experimental challenges. The ever-expanding repertoire of molecular and genetic tools provides many exciting avenues for manipulating gene expression; however, the majority of these strategies disrupt multiple elements of lncRNA biogenesis and function, making their effects difficult to evaluate. It is therefore imperative to understand both the advantages and experimental caveats of each approach when investigating the function of cis-acting lncRNAs (Bassett et al., 2014; Kopp and Mendell, 2018).

Molecular

Molecular tools provide an opportunity for manipulating transcription and RNA accumulation without altering the underlying DNA sequence (Fig. 2). These approaches are alluring on account of the relative ease and rapidity with which they can be implemented; however, many have been shown to incur significant off-target effects, casting doubt on their reliability (Goudarzi et al., 2019; Stojic et al., 2018). As a result, researchers must take care to use multiple approaches or targeting sequences to confirm the specificity of any putative phenotype.

RNA interference (RNAi):

RNAi exploits the ability of small interfering RNAs (siRNAs) to mediate the homology-dependent cleavage and degradation of target transcripts. In this approach, siRNA “guide strands” associate with components of the RNA-induced silencing complex (RISC), which includes the endonuclease Ago2 (Argonaute-2). Ago2, in turn, cleaves within the target RNA, yielding an upstream fragment that can be degraded by the RNA exosome (Lima et al., 2016) as well as a downstream fragment that may be degraded by 5’-3’ exonucleases such as XRN1 or XRN2 (Lima et al., 2016).

Although RNAi efficiently depletes cytoplasmic transcripts, its activity in the nucleus is more controversial (Lennox and Behlke, 2016). Several studies have reported that RNAi factors are present and
Figure 2. Molecular approaches for interrogating lncRNA function. A number of molecular tools may perturb lncRNA transcription, processing, accumulation, and function without altering the underlying DNA sequence. RNAi may attenuate transcription or prevent transcript accumulation, although its efficacy for knocking down nuclear—and in particular, chromatin-enriched—transcripts remains controversial. ASO gapmers act co- or post-transcriptionally to trigger RNase H-dependent RNA degradation. Alternatively, non-gapmer ASOs may be used to inhibit secondary structure formation or prevent RNA-protein interactions without incurring RNA knockdown. Catalytically inactive Cas9/CRISPR-based tools can be used to block transcription or occlude transcription factor binding sites within a locus. Alternatively, Cas9 may be fused to activating (CRISPRa) or repressive (CRISPRi) effector domains in order to modulate local transcription. Lastly, programmable RNA-targeting enzymes such as Cas13 may be used to either destabilize target transcripts or impair RNA processing and function.
active in the nucleus (Gagnon et al., 2014; Robb et al., 2005), where they have been proposed to mediate both post-transcriptional RNA knockdown and transcriptional silencing (Morris et al., 2004). However, other studies have observed lower knockdown efficacies for nuclear targets, casting doubt on the robustness of this approach (Bassett et al., 2014; Lennox and Behlke, 2016; Lima et al., 2016; Stojic et al., 2018). It is therefore difficult to evaluate data from studies that use RNAi to knock down cis-acting lncRNAs.

**Antisense oligonucleotides (ASOs):**

“Gapmer” ASOs represent a well-established molecular approach for knocking down both nuclear and cytoplasmic RNAs (Lennox and Behlke, 2016). These single-stranded oligonucleotides typically comprise a central DNA sequence flanked by nuclease-resistant nucleotides such as locked nucleic acids. ASOs bind complementary sequences within their target RNAs, forming an RNA:DNA heteroduplex. RNAse H subsequently cleaves the RNA strand of this hybrid complex (Wu et al., 2004), yielding unprotected RNA fragments that are rapidly degraded by the RNA exosome (Lima et al., 2016) and the 5'-end surveillance factors XRN1 and XRN2 (Hori et al., 2015; Lima et al., 2016). Although gapmers can effect post-transcriptional RNA depletion, as evidenced by their ability to act in the cytoplasm (Liang et al., 2017), recent work reveals that they may also promote the co-transcriptional cleavage and degradation of target RNAs (Lai et al., 2020; Lee and Mendell, 2020). Of particular significance for studies on cis-acting lncRNAs, ASOs targeting sequences proximal to the TSS may mediate a significant reduction in transcription upstream of the targeting site (Lee and Mendell, 2020), making it difficult to distinguish between the effects of RNA knockdown and changes in the local transcriptional landscape. These studies therefore underscore the need for caution when interpreting the results of gapmer knockdown experiments.

A second class of ASOs can inhibit RNA processing and function without triggering RNAse H-dependent RNA degradation. For example, morpholino oligomers targeting splice sites may impair splicing and impede the accumulation of mature RNA species (Almada et al., 2013; Ulitsky et al., 2011).
Alternatively, ASOs targeting splicing enhancers or repressors may promote exon skipping or retention. Lastly, these ASOs may be used to protect microRNA targets (Choi et al., 2007) or sterically block RNA-protein interactions (Ntini et al., 2018). Together, these tools can help unravel the post-transcriptional role of noncoding RNAs.

**CRISPR/Cas-based approaches:**

Since the initial characterization of the CRISPR/Cas9 system in 2013, researchers have reported many innovative applications for CRISPR-associated (Cas) proteins. In addition to their revolutionary role in gene editing (discussed below), these proteins have been shown to serve as a programmable platform for gene regulation. Cas9-based tools have proven particularly versatile as both negative and positive effectors of transcription. However, ongoing research also implicates other Cas proteins as exciting tools for regulating RNA processing, editing, and stability.

**Loss-of-function:**

Researchers have implemented several CRISPR/Cas-based loss-of-function approaches for studying both coding and noncoding loci. The simplest of these approaches exploits the ability of Cas9 to sterically block transcriptional elongation when targeted downstream of a TSS (Gilbert et al., 2013; Qi et al., 2013). To avoid cleaving the DNA target, researchers can either use catalytically dead Cas9 (dCas9) in conjunction with a standard guide RNA (gRNA) or use active Cas9 in tandem with a 14–15 nucleotide “dead” gRNA (dRNA) that permits target binding without incurring nuclease activity (Dahlman et al., 2015). This so-called “transcriptional roadblock” approach has been shown to yield a 60–80% decrease in gene expression (Gilbert et al., 2013).

Alternatively, Cas9-fusion proteins can be used to modulate the epigenetic landscape around a TSS (CRISPRi). When fused to Cas9 or dCas9, the repressive Krüppel-associated box (KRAB) recruits histone-methylating and -deacetylating complexes to a region centered around its binding site, enacting local repression (Thakore et al., 2015). Although this strategy appears to have minimal off-target activity
(Moses et al., 2020; Stojic et al., 2018), its consequences for neighboring gene expression remain unclear: studies have reached mixed conclusions regarding the potential of KRAB to induce heterochromatin spreading over regions as large as 20 kilobases (Groner et al., 2010; Thakore et al., 2015). Since many cis-acting lncRNAs are located in close proximity to their neighboring genes, this raises the unfortunate possibility that KRAB may directly silence both targeted and nearby genes. It may therefore be necessary to monitor local H3K9me3 levels when investigating cis-regulation over short distances.

Other Cas proteins may enact repression through their interactions with target RNAs. Cas13, the archetypal CRISPR Class 2 Type VI protein, binds RNAs in a programmable RNA guide-dependent manner (Konermann et al., 2018). Cas13 proteins mediate the cleavage of target transcripts and have been reported to mediate a specific 50-90% decrease in RNA levels (Abudayyeh et al., 2017; Konermann et al., 2018). Intriguingly, the Cas13 family proteins Cas13a and Cas13b have been shown to generate upstream and downstream fragments with a 2’,3’-cyclic phosphate and a 5’-hydroxyl, respectively (Gootenberg et al., 2018). This raises the tantalizing possibility that co-transcriptional cleavage may not incur co-transcriptional RNA degradation by XRN2, which exhibits a preference for substrates with 5’-phosphate groups (West et al., 2004). Like Cas9, Cas13 can be rendered catalytically inactive (dCas13) via point mutations within each of its two endonuclease domains. This observation implicates CRISPR/Cas13 as a programmable RNA-binding platform. Indeed, researchers have successfully used dCas13 to restore normal splicing of Smn2 (survival motor neuron 2) in fibroblasts from patients with spinal muscular atrophy, a neuromuscular disorder caused by aberrant Smn2 splicing (Du et al., 2020). Additionally, dCas13-fusion proteins show promise as tools for RNA editing. For example, Cas13-ADAR (adenosine deaminase acting on RNA) fusion proteins can mediate A-to-I editing, facilitating the site-specific modification of target transcripts (Cox et al., 2017).

While the identification and implementation of Cas13 family proteins opens up an exciting new chapter in our ability to regulate RNA expression, two major obstacles remain unaddressed. Firstly, the principles governing the design of efficient gRNAs are murky. Although some groups claim to have developed rules for rational gRNA design (Bandaru et al., 2020; Wessels et al., 2020), the paucity of more
recent studies suggest that this problem requires additional attention. Secondly, and perhaps more concerning, RNA guide-matching can trigger Cas13d to cleave “bystander” RNAs (Abudayyeh et al., 2016; Konermann et al., 2018; Wang et al., 2019a). However, the extent of this so-called “collateral effect” remains unclear, with one study even reporting that CRISPR/Cas13 collateral activity is virtually undetectable in mouse and zebrafish embryos (Kushawah et al., 2020). Additional work will doubtlessly unlock the potential of these tools to enhance our understanding of *cis*-acting lncRNA function.

Gain-of-function:

By definition, *cis*-acting lncRNAs regulate gene expression in a locus-specific manner. Common gain-of-function approaches such as transgene overexpression are therefore ill-suited for studying the function of lncRNAs whose transcription, processing, or accumulation operates locally. CRISPR/Cas9-based technology provides two major strategies for bypassing this problem.

CRISPR/Cas9-based gene activation (CRISPRa) uses Cas9 or dCas9 to recruit transcriptional activators to a target site. Similar to CRISPRi, Cas9 may be directly fused to transcriptional activators such as VP16, p65, or RTA (Chavez et al., 2015; Gilbert et al., 2013). While genome-wide screens demonstrate the fundamental viability of this approach (Gilbert et al., 2014), other strategies may facilitate more robust transcriptional activation. For example, the synergistic activation mediator (SAM) system uses a chimeric gRNA harboring two MS2 stem-loops to recruit the MS2 bacteriophage coat protein. This, in turn, is fused to the bipartite p65-HSF1 transcriptional activator (Konermann et al., 2015). Since the MS2 coat protein binds its stem-loop as a dimer, this approach allows for one Cas9 complex to bind up to four activator complexes, thereby enhancing its ability to upregulate target genes. Importantly, CRISPRa appears to mediate efficient transcriptional activation only when targeted within ~200 nucleotides of a TSS (Gilbert et al., 2014; Joung et al., 2017; Konermann et al., 2015). While this inherently limits the sequence space available for targeting, it suggests that CRISPRa may be a viable approach for studying lncRNAs that are located in close proximity to their target genes.
A second class of Cas9-based tools uses dCas9 to recruit exogenous RNAs to a lncRNA locus. In this approach, termed CRISPR-display, the RNA guide is fused to an exogenous RNA sequence; this, in turn, facilitates the delivery of the RNA cargo to a locus of interest (Shechner et al., 2015). Since the chimeric gRNA is exogenously expressed, this strategy provides an opportunity to divorce the locus-specific accumulation of lncRNA molecules from features such as transcription or DNA elements. Though exciting, this technique has not been widely adopted, making it difficult to evaluate its experimental value.

Genetic

In contrast to protein-coding genes, whose products may be compromised by small insertions and deletions that disrupt ORFs, lncRNAs exhibit less obvious sequence requirements, making it difficult to genetically ablate lncRNA expression without dramatically altering the underlying DNA sequence. Given the unique considerations and challenges posed by each lncRNA locus, researchers have developed a number of genetic approaches for modulating lncRNA biogenesis and function both in vitro and in vivo (Fig. 3). The merits and pitfalls of each approach are discussed below.

Knockout strategies:

A number of studies have used largescale deletion-based approaches to establish the functionality of lncRNA loci (Dimitrova et al., 2014; Engreitz et al., 2016; Hacisuleyman et al., 2014; Li et al., 2013a; Schorderet and Duboule, 2011). This approach entails the genetic ablation of either an entire lncRNA locus or a smaller region encompassing the promoter. While this strategy can provide exciting preliminary evidence for lncRNA functionality, it cannot distinguish the role of transcription and the resulting RNA from the role of regulatory DNA elements within the locus (Bassett et al., 2014). Moreover, large deletions have the potential to disrupt the spacing between protein-coding genes and regulatory elements outside the locus. This, in turn, may mediate the dysregulation of genes whose activity is not subject to cis-regulation by the deleted locus.
Figure 3. Genetic approaches for interrogating lncRNA function. Genetic techniques facilitate the dissection of lncRNA activity. Large-scale deletion approaches, which abrogate transcription and RNA accumulation by destroying the entire locus or promoter region, can establish a functional requirement for a lncRNA locus but cannot specifically test the significance of the transcriptional process. Similarly, CRISPR/Cas9 mutagenesis experiments may reveal minimal sequence elements that drive both lncRNA and protein-coding gene expression but cannot distinguish between DNA elements that act directly and elements that act through lncRNA biogenesis and accumulation. Reporter gene substitution, which entails the substitution of a reporter gene such as lacZ for all or part of a lncRNA locus, disrupts DNA elements and the RNA sequence while permitting some degree of transcription. Locus inversion approaches, which leave critical DNA sequences intact, may be used to test the functionality of transcription or RNA sequence elements. Similarly, genetic knock-in strategies interrogate the role of transcription and RNA accumulation without destroying the underlying DNA sequence. PAS sequences prematurely terminate transcription, while self-cleaving ribozymes promote RNA destabilization.
To avoid this potential caveat, other studies have used CRISPR/Cas9 to achieve precise mutagenesis of known regulatory elements such as transcription factor binding sites and 5’ splice sites (Engreitz et al., 2016; Korkmaz et al., 2016; Li et al., 2017; Olivero et al., 2020). While this approach minimally disrupts the chromatin neighborhood surrounding the lncRNA locus, it may attenuate or even abolish lncRNA transcription and accumulation. In this case, it may still be difficult to distinguish between a requirement for sequence elements and a requirement for lncRNA transcription (Bassett et al., 2014).

Transcript substitution:

Transcript substitution entails the partial or complete replacement of a lncRNA locus with a reporter gene such as LacZ or GFP. While the reporter gene ostensibly recapitulates the expression of the lncRNA in different cell types and biological conditions (Goff et al., 2015; Groff et al., 2016; Sauvageau et al., 2013), the absence of key DNA sequence elements may lead to aberrant reporter gene expression (Bassett et al., 2014). Moreover, although this strategy permits transcription, sequence differences, including the absence of splice sites in some reporter substitution approaches, may alter the transcriptional kinetics of the locus (Brinster et al., 1988; Engreitz et al., 2016; Suter et al., 2011). If nearby genes are sensitive to lncRNA transcription or processing, then this approach may yield a phenotype even if the DNA sequence does not per se harbor cis-regulatory activity.

Locus inversion:

Inversion-based strategies have the advantage of disrupting lncRNA transcription and function without deleting key DNA sequence elements (Bassett et al., 2014). Some researchers have inverted promoter sequences in order to abolish lncRNA transcription (Pease et al., 2013). Others have inverted regions that harbor putative RNA elements (Senner et al., 2011). While such approaches leave the DNA sequence intact, large inversions may disrupt the spacing between regulatory elements and their target
genes (Delpretti et al., 2013). Additionally, the inherent bidirectionality of promoters may limit the utility of locus inversion as an approach for abrogating transcription. These considerations underscore the need for caution when interpreting results from locus inversion experiments.

**Knock-in strategies:**

With a few exceptions, the genetic approaches outlined above perturb both the DNA sequence and lncRNA biogenesis. Genetic knock-in strategies, in which a sequence element is inserted into an otherwise-intact locus, provide a means of altering lncRNA transcription and stability without deleting potentially critical DNA sequence elements.

**Polyadenylation signals:**

Eukaryotic PAS elements represent a well-established strategy for inducing premature transcriptional termination both *in vitro* and *in vivo*. These elements, which comprise a canonical A(A/U)UAAA hexamer located 30 nucleotides upstream of a GU-rich tract, stimulate the 3’ end-processing machinery to cleave and polyadenylate nascent transcripts (Millevoi and Vagner, 2010). Similar to RNAse H-mediated cleavage, this yields an unprotected downstream fragment that is progressively degraded by the nuclear 5’ end surveillance factor XRN2. XRN2 displaces RNA polymerase II from the DNA template, abolishing transcription. This process typically occurs within two kilobases of the PAS, although transcription may occasionally continue for more than ten kilobases (Core et al., 2008).

A number of PAS sequences have been shown to successfully suppress transcription from endogenous loci. The synthetic rabbit β-globin PAS has previously been used to induce premature transcriptional termination *in vitro* and appears to be most efficient when inserted within close proximity to a TSS (Engreitz et al., 2016; Levitt et al., 1989). The short length (49 nucleotides) of this sequence makes it particularly attractive as a tool for knocking down lncRNA expression, since it minimally disrupts the spacing of DNA elements within a locus. Longer elements comprising three tandem PAS
sequences may prove more appropriate for inducing transcriptional termination at regions distal from a TSS, although they will inherently have more of an effect on the spacing of the locus.

As an important consideration, premature transcriptional termination, particularly termination upstream of the first intron, may have significant consequences for transcriptional initiation (Brinster et al., 1988; Engreitz et al., 2016). It has long been known that the U1 small nuclear ribonucleoprotein (snRNP) can bind general transcription factors (Fong and Zhou, 2001; Kwek et al., 2002). Cotranscriptional splicing may therefore recruit Pol II and enhance transcriptional initiation and elongation. On account of the tight coupling between splicing and transcription, PAS insertion may attenuate transcription from a lncRNA TSS in addition to suppressing transcription downstream of the insertion site (Engreitz et al., 2016).

Self-cleaving ribozymes:

Self-cleaving ribozymes represent an emerging class of tools for dissecting endogenous lncRNA function. These naturally-occurring RNA sequences form unique secondary and tertiary structures that enable them to enact site-specific phosphodiester scission (Jimenez et al., 2015). Unlike the cleavage events mediated by RNase H and the cleavage and polyadenylation machinery, ribozyme cleavage yields an upstream fragment with a 2′-3′ cyclic phosphate as well as a downstream fragment with a 5′-hydroxyl group (Roth et al., 2014). In mammalian cells, the upstream fragment likely undergoes rapid degradation by the RNA exosome; the fate of the downstream fragment is less obvious. While the nuclear exosome readily degrades RNAs with 2′-3′ cyclic phosphates (Zinder et al., 2016), 5′-3′ exoribonucleases such as XRN2 exhibit a stringent requirement for RNAs with 5′-phosphate groups (Doamekpor et al., 2020; Jinek et al., 2011; Mathy et al., 2007; West et al., 2004). This raises the possibility that fast-acting ribozymes such as the Hepatitis Delta ribozyme and Twister may cleave lncRNA transcripts without incurring transcriptional termination by XRN2.

The functional consequences of co-transcriptional ribozyme cleavage remain murky. While self-cleaving ribozymes do not appear to affect mRNA stability in the context of short reporter gene constructs
(Bird et al., 2005; Fong et al., 2009; NOTT et al., 2003; West et al., 2004), a handful of pioneering studies show that they can be used to destabilize endogenous lncRNAs (Camblong et al., 2009; Tuck and Bühler, 2021; Tuck et al., 2018; Wery et al., 2018a). The Hammerhead and Hepatitis Delta ribozymes have both been shown to suppress lncRNA accumulation in mouse and yeast cells, although the knockdown efficiency varies considerably across lncRNA targets (Camblong et al., 2009; Tuck and Bühler, 2021; Tuck et al., 2018). The reason for the observed discrepancies in efficiency between mRNAs and lncRNAs and between individual lncRNAs is presently unclear. It is possible that factors such as the distance of the ribozyme from a TSS or transcription termination site (TTS) may affect the stability of the nascent transcript; additionally, the effects of ribozyme cleavage may be more pronounced in the context of endogenous loci, which are often many times longer than reporter genes and therefore require more time to generate a full-length, processed transcript. On the whole, the experimental variability in ribozyme efficiency highlights the lack of clear criteria for determining optimal ribozyme placement. As a further challenge, the molecular effects of co-transcriptional ribozyme cleavage on lncRNA transcription and processing are still enigmatic. Subsequently, it remains difficult to determine the mechanistic basis for any changes to nearby gene expression.

**Conclusion**

In summary, researchers have developed an extensive set of molecular and genetic tools for manipulating the expression of cis-acting lncRNAs. On account of their idiosyncrasy, however, cis-acting lncRNA loci do not lend themselves to a one-size-fits-all approach. The merits of each tool must be evaluated in the context of each locus; and depending on the unique considerations posed by each gene neighborhood, individual tools may prove insufficient for distinguishing between competing mechanisms. In this case, multiple orthogonal strategies may help to resolve outstanding questions at technically challenging loci. In this work, we used complementary molecular and genetic strategies to identify the functional element of cis-regulation at the LincRNA-p21 locus. Additionally, we set out to rigorously
characterize the molecular effects of ribozyme cleavage on transcription in the context of our locus. We present specific experimental details in Chapters III and IV.
Chapter II: Project framework
LincRNA-p21: a functional mediator of the p53 pathway

The p53 tumor suppressor pathway

The p53 pathway lies at the heart of the cellular response to stress. On account of its central role in response to diverse biological stressors, the tumor suppressor protein p53 has earned the appellation “guardian of the genome” and secured a vaunted spot at the top of the list of the most frequently studied human genes (Dolgin, 2017). Following exposure to adverse stimuli such as DNA damage, oncogenic signaling, oxidative stress, and nutrient deprivation (Sermeus and Michiels, 2011), p53 activates complex transcriptional programs that promote various functional outcomes. These consequences include cell cycle arrest, DNA damage repair, senescence, and even apoptosis (Vousden and Prives, 2009). Although the mechanisms through which p53 specifies cell fate in response to disparate stimuli remain murky, the centrality of p53 in each of these pathways is readily apparent. As a grim testament to its importance, both humans and mice deficient for p53 show a propensity to develop a broad range of cancers (Donehower et al., 1992; Olivier et al., 2010). Moreover, the p53 protein is inactivated in more than 50% of cancers, marking it as the gene most frequently inactivated during tumorigenesis (Kastenhuber and Lowe, 2017). A poor patient prognosis typically accompanies p53 inactivation (Olivier et al., 2010), further demonstrating the broad importance of the p53 transcriptional network.

Like p53 loss-of-function, p53 overexpression has deleterious cellular and organismal effects (Kastenhuber and Lowe, 2017). Cells therefore employ several regulatory mechanisms to escape the pathological consequences of inappropriate p53 signaling. Under conditions of homeostasis, the constitutively-translated p53 protein undergoes rapid degradation, primarily as a result of its interactions with MDM2 (mouse double minute 2). This E3 ubiquitin ligase binds and ubiquitylates p53, mediating its degradation by the proteasome and subsequently inhibiting p53 accumulation (Haupt et al., 1997; Kubbutat et al., 1997; Kussie et al., 1996). Other protein factors also exert modest repressive effects on p53 signaling. For example, the MDM2 homolog MDM4 (mouse double minute 4), which lacks ubiquitin
ligase activity, binds to and occludes p53’s transactivation domain, preventing it from interacting with transcriptional co-activators (Toledo and Wahl, 2007). Together, these repressors enact a strict regulatory paradigm that suppresses signaling under homeostatic conditions while leaving p53 poised for rapid activation in response to cellular stressors.

Such stressors may activate p53 expression through different context-dependent mechanisms. Oncogenic signaling, for example, activates the tumor suppressor ARF (alternative reading frame of CDKN2A), which binds and destabilizes MDM2 (Zhang et al., 1998). This relieves p53 from constitutive repression, permitting its accumulation within the cell. Alternatively, the presence of DNA damage may mediate the post-translational modification of p53, altogether abrogating its ability to bind to MDM2 (Shieh et al., 1997). Such post-translational modifications both enhance p53’s stability and contribute to its ability to effect transactivation.

Once stabilized, p53 translocates into the nucleus, where it induces the expression of myriad transcriptional targets. To accomplish this, the p53 protein forms tetramers that bind to and “scan” genomic DNA for p53 responsive elements (p53REs)(Tafvizi et al., 2011). These genomic elements comprise two decameric half-sites, separated by a spacer of 0-13 nucleotides, with the consensus sequence 5’-RRRCWWGYYY-3’ (El-Deiry et al., 1992). As a pioneer transcription factor, p53 readily binds to p53REs in diverse chromatin environments (Laptenko et al., 2011; Nili et al., 2010; Sammons et al., 2015). Once bound to the chromatin, p53 may recruit chromatin remodeling factors, histone acetyltransferases, or general transcription factors (Dornan et al., 2003; Lee et al., 2002; Zilfou et al., 2001). This, in turn, creates a transcriptionally permissive epigenetic environment and facilitates the expression of nearby gene targets.

Intriguingly, p53 frequently associates with response elements in intergenic regions, including enhancers (Léveillé et al., 2015; Tesfaye et al.; Younger et al., 2015). Many of these intergenic p53REs appear to be functional in spite of their large genomic distance from protein-coding genes. Recently, p53 has attracted considerable attention as a potential mediator of long-range chromatin looping (Link et al., 2013; Melo et al., 2013). p53 primarily binds each p53RE as a tetramer (Stenger et al., 1992, 1994; Wang
et al., 1995), and p53 tetramers may further oligomerize to promote looping interactions between p53 sites separated by over a kilobase (Stenger et al., 1994). Notably, many p53-bound enhancers are located in close spatial—if not genomic—proximity to p53-responsive genes, raising the possibility that p53 binding might promote chromatin reorganization. However, 3D chromatin interactions between p53-bound enhancers and promoters are largely p53-insensitive (Link et al., 2013; Melo et al., 2013), suggesting that p53 exploits pre-established chromatin contacts rather than specifying de novo interactions. These reports, which indicate that p53 is unlikely to be a major determinant of chromatin architecture, thus provide a plausible alternative mechanism through which distal p53REs can regulate their target genes.

p53 binding induces the expression of numerous transcriptional targets. Historically, efforts to investigate the effects of p53 signaling and dysregulation have focused primarily on protein-coding genes, including the cell cycle arrest mediator p21 (also known as Cdkn1a and p21WAF1/Cip1). Although among the first p53 target genes to be discovered, p21 remains an active subject of investigation on account of its outsized role in p53-mediated transcriptional repression (Fischer, 2017). Upon exposure to stress, p53 activates p21 transcription through two p53REs within the p21 promoter (El-Deiry et al., 1994; Espinosa and Emerson, 2001). Once generated, p21 protein may either remain in the cytoplasm or translocate into the nucleus. There, it enacts its tumor suppressive effects by inhibiting the ability of cyclin-dependent kinases (CDKs) to phosphorylate Rb (retinoblastoma) family proteins (Brugarolas et al., 1995; Harper et al., 1995; Wade Harper et al., 1993; Xiong et al., 1993). Since this phosphorylation prevents Rb family proteins from repressively associating with E2 family (E2F) transcription factors and blocking the transcription of key cell cycle progression genes (Harbour and Dean, 2000), p21 activity supports transcriptional repression by the Rb-E2F complex. Using a similar mechanism, p21 also promotes the assembly of the DREAM (dimerization partner, Rb-like, E2F, and multivulval class B) complex, which downregulates a broad class of target genes. Additionally, p21 sequesters PCNA, preventing it from interacting with DNA polymerase δ and initiating DNA replication (Chen et al., 1995; Luo et al., 1995).
These parallel mechanisms have important cellular consequences. Replication inhibition and transcriptional repression of cell cycle genes both impede cell cycle progression from G1 into S phase (Harper et al., 1995). Critically, this prevents cells from replicating damaged DNA and instead provides them with a critical opportunity to repair compromised templates (Shaltiel et al., 2015). Moreover, the DREAM complex likely underlies p21’s ability to regulate the G2/M checkpoint (Bunz et al., 1998). In some cases, p21 activation can also lead to permanent growth arrest (i.e., cellular senescence). How p21 levels influence cells’ decisions to enter transient growth arrest or permanently senesce remains an active area of investigation (Hsu et al., 2019).

Over the past decade, noncoding RNAs have been implicated as key effectors of the p53 pathway and have generated significant excitement on account of their proposed role in outcome specificity (Vousden and Prives, 2009). The discovery of p53-regulated miRNAs dramatically enhanced our understanding of the downstream mechanisms of p53 signaling. And more recently, transcriptomics studies have implicated lncRNAs as a novel class of p53 targets in both humans and mice (Sánchez et al., 2014; Tesfaye et al.). Like protein-coding genes, these p53-inducible lncRNAs may show either universal or context-specific expression, as elegantly illustrated by one recent study, which used RNA profiling to examine the consequences of p53 expression in a panel of p53-deficient mouse cancer cell lines. Notably, cancer cell lines that exhibited similar functional outcomes (i.e., senescence or apoptosis) in response to p53 activation showed similar p53-dependent lncRNA transcriptional signatures (Tesfaye et al.). This intriguing correlation between lncRNA expression and functional outcome raises the possibility that lncRNAs may help govern context-specific p53 outcomes, adding a new regulatory dimension to the p53 tumor suppressor network. Indeed, a growing body of work has demonstrated the ability of p53-responsive lncRNAs enact diverse functional consequences (see Chapter I)(Dimitrova et al., 2014; Olivero et al., 2020; Tesfaye et al.). The remainder of this work will largely focus on the role of the p53-inducible lncRNA LincRNA-p21 in transcriptional activation of the nearby p21 gene.
The long noncoding RNA *LincRNA-p21* in the p53 tumor suppressor pathway

The lncRNA *LincRNA-p21* has garnered intense interest on account of its enigmatic role within the p53 pathway. This lncRNA, which in mice spans a ~20 kilobase locus in the intergenic region between the *Srsf3* (serine and arginine rich splicing factor 3) gene and the *p21* cell cycle checkpoint inhibitor, was named for its close proximity to the *p21* locus (Fig. 4). Although the human *LincRNA-p21* locus occupies a smaller genomic footprint (approximately four kilobases), the human and mouse *LincRNA-p21* loci share several important commonalities. Notably, in both species, the *LincRNA-p21* and *p21* transcriptional units constitute a non-overlapping head-to-head pair separated by a distance of just over ten kilobases. Moreover, although *LincRNA-p21* exhibits only modest conservation at the sequence level, mouse and human *LincRNA-p21* share a ~200 nucleotide region of homology at their 5’ ends (Chillón and Pyle, 2016). In keeping with the finding that lncRNA loci often harbor conserved transcription factor binding sites, this region includes binding sites for both p53 and hnRNP-K (Chillón and Pyle, 2016; Wu et al., 2014). Taken together, these conserved features seem to attest to the importance of the spatial relationship at this locus and hint at a common regulatory role in humans and mice.

*LincRNA-p21* was first identified in a screen for murine lincRNAs with high levels of H3K4me3 and H3K36me3 in their promoters and gene bodies, respectively (Guttman et al., 2009). This so-called H3K4me3-H3K36me3 signature, a hallmark of active Pol II transcription, suggested that the *LincRNA-p21* locus might be transcribed into a lncRNA. (Importantly, its CSF score strongly argues against the possibility that it could be subsequently translated (Guttman et al., 2009)). Indeed, the mouse *LincRNA-p21* locus gives rise to a 3.1 kilobase RNA that is capped, spliced, and polyadenylated (Guttman et al., 2009; Huarte et al., 2010). Meanwhile, the human locus has been variously proposed to encode a spliced transcript or two single-exon isoforms: a short, polyadenylated transcript and a longer, non-polyadenylated transcript (Chillón and Pyle, 2016; Yoon et al., 2012). Consistent with the observation that its promoter harbors a p53RE, DNA damage was shown to induce *LincRNA-p21* expression in the context of wildtype, but not p53-deficient, mouse embryonic fibroblasts (MEFs) (Guttman et al., 2009).
Figure 4. Historical characterization of the *LincRNA-p21* locus. Schematic of the murine *LincRNA-p21* locus highlighting molecular and genetic approaches that established its *cis*-regulatory function. Red arrows indicate the location of *loxP* sites used to generate conditional *LincRNA-p21* knockout cells, while white boxes show the relative position of two ASO gapmers targeting the *LincRNA-p21* intron (ASO-I) and second exon (ASO-E2) (Dimitrova et al., 2014). The *lacZ* substitution strategy is represented by the blue box (Groff et al., 2016; Sauvageau et al., 2013). The *LincRNA-p21* and *p21* p53REs are indicated by red asterisks (*).

This observation provided tantalizing preliminary evidence that *LincRNA-p21* could play a role in the p53 tumor suppressor pathway.

Subsequent studies in both mouse and human cells implicated *LincRNA-p21* as a *trans*-acting lncRNA but failed to reach a consensus on its mechanism or function within the p53 pathway (reviewed in Table 2.1). These early studies, which overwhelmingly favored RNAi-based knockdown and transgene overexpression strategies (Bao et al., 2015; Huarte et al., 2010; Yang et al., 2014a; Yoon et al., 2012), suggested the involvement of *LincRNA-p21* in processes ranging from angiogenesis and glycolysis to apoptosis and cell cycle arrest. These studies further proposed wide-ranging and occasionally contradictory mechanisms for *LincRNA-p21* function. Initial experiments in p53-restorable MEFs and murine tumor cells supported a model in which *LincRNA-p21* RNA binds hnRNP-K in order to promote apoptosis, an outcome accomplished by enacting p53-dependent transcriptional repression of numerous target genes (Huarte et al., 2010). Under this paradigm, *LincRNA-p21* functions in *trans*, as it modulates the localization of hnRNP-K across the genome. A second study, this time in mouse pre-iPSCs, proposed that *LincRNA-p21* inhibits cellular reprogramming by associating with hnRNP-K in order to recruit the SETDB1 and DNMT1 methyltransferases to the promoters of pluripotency genes (Bao et al., 2015). Curiously, the authors did not detect any changes in apoptotic potential following *LincRNA-p21* knockdown.
Experimental forays into the role of human *LincRNA-p21* spurred the development of addition models for *LincRNA-p21* function. One study proposed that *LincRNA-p21* RNA acts as a cytoplasmic repressor of target mRNA translation (Yoon et al., 2012). Although the authors showed that *LincRNA-p21* associates with several of its proposed mRNA targets and can also bind polysomes in HeLa cells, it remains unclear how this IncRNA might be enacting repression. A second study reported entirely unrelated conclusions: Yang et al. found that *LincRNA-p21* RNA inhibits HIF-1α degradation, thereby promoting glycolysis under conditions of hypoxia (Yang et al., 2014a). Interestingly, they identified a hypoxia response element within the human *LincRNA-p21* promoter and demonstrated that *LincRNA-p21* is induced in response to hypoxic stress. They further observed that *LincRNA-p21* RNA blocks the association of HIF-1α with the VHL E3 ubiquitin ligase, which mediates the constitutive degradation of HIF-1α under normoxic conditions. On account of this, they posited that *LincRNA-p21* promotes glycolysis under hypoxic conditions and implicated it as a novel IncRNA player in the Warburg effect (Yang et al., 2014a).

While these findings are certainly intriguing, the failure of early studies to converge on a mechanism for *LincRNA-p21* function casts doubt on the robustness of their conclusions. Notably, all of these studies used RNAi to deplete *LincRNA-p21* RNA levels, raising the possibility that off-target effects could underlie the diverse phenotypes observed in these studies (Dimitrova et al., 2014). As an additional point of inconsistency, the ascribed cellular localization of *LincRNA-p21* RNA varies dramatically. While the findings of Huarte et al. and Bao et al. imply that *LincRNA-p21* RNA is largely chromatin-bound, Yoon et al. reported that this RNA exhibits extensive cytoplasmic localization in HeLa cells (Bao et al., 2015; Huarte et al., 2010; Yoon et al., 2012). While this discrepancy may reflect largescale differences in IncRNA processing between mice and humans (Guo et al., 2020), it may also point to a lack of methodological robustness.

Indeed, more recent findings have cast doubt on the results of these initial studies and instead demonstrated that *LincRNA-p21* functions as a *cis*-regulatory locus. In contrast with the RNAi-based approaches favored in previous work, these subsequent studies made notable use of genetic tools to
elucidate the mechanism and function of this locus. In one study, researchers generated mice in which the entire LincRNA-p21 promoter, including the p53RE, had been deleted (Fig. 4)(Dimitrova et al., 2014). Intriguingly, the resulting LincRNA-p21-deficient MEFs exhibited a specific 30-50% decrease in p21 expression, suggesting that LincRNA-p21 reinforces p21 levels. The authors built on these results, which they confirmed using ASOs targeting LincRNA-p21 RNA (Fig. 4), to show that the LincRNA-p21 transcriptional process or transcript acts synergistically with hnRNP-K to stabilize p53 binding at the p21 promoter; they thus provided a clear mechanism for the observed role of LincRNA-p21 in G1/S cell cycle checkpoint inhibition (Table 2.1). In a separate study, researchers used a different genetic approach to further implicate a cis-regulatory role for LincRNA-p21. Groff et al. genetically engineered mice in which the entire LincRNA-p21 locus was replaced with a lacZ reporter cassette (Fig. 4)(Groff et al., 2016; Sauvageau et al., 2013). In keeping with Dimitrova et al., they observed a significant decrease in p21 expression across multiple adult and embryonic mouse tissues; however, they noted that many genes in the four megabase window surrounding the LincRNA-p21 locus were also dysregulated. Moreover, they found that their genetic replacement led to a decrease in p21 levels even in tissues lacking LincRNA-p21 expression, leading them to conclude that DNA elements within the LincRNA-p21 locus must act as local enhancers of gene regulation.

In contrast to earlier work on LincRNA-p21, in which each study seemingly revealed a new function and mechanism, subsequent reports have largely confirmed a cis-regulatory function for LincRNA-p21. In particular, two recent studies used CRISPR/Cas9-mediated mutagenesis to disrupt DNA sequence elements within the LincRNA-p21 locus of mouse and human cancer cells (Korkmaz et al., 2016; Tesfaye et al.). In both studies, the authors opted to target the p53RE, which has previously been shown to drive the expression of a synthetic reporter construct (Groff et al., 2016). In keeping with the role of LincRNA-p21 in the p53 pathway, p53RE mutagenesis abrogated lncRNA transcription (Korkmaz et al., 2016; Tesfaye et al.). Strikingly, mutagenesis also incurred a specific decrease in p21 expression, consistent with the findings of the two previous studies.
Collectively, these four studies imply that the *LincRNA-p21* locus harbors *cis*-regulatory potential. In line with this, and belying its alleged *trans* activity, the murine *LincRNA-p21* transcript is present at ~8 molecules per cell and largely remains at its site of transcription in MEFs (Dimitrova et al., 2014). However, it is unclear what aspect of the *LincRNA-p21* locus—DNA elements, the transcriptional process, or the full-length transcript—enacts *cis*-regulation. The *LincRNA-p21* p53RE certainly harbors *cis*-regulatory activity, as it is both necessary for *LincRNA-p21/p21* expression and sufficient to drive reporter gene expression (Groff et al., 2016; Korkmaz et al., 2016; Tesfaye et al.). Other elements in the human *LincRNA-p21* promoter region also appear to co-regulate the two genes (Korkmaz et al., 2016). However, given that mutagenesis of these elements abolishes *LincRNA-p21* transcription, it is unclear whether they are acting directly or regulating *p21* by promoting *LincRNA-p21* transcription and RNA accumulation. ASO-mediated RNA knockdown experiments provide the strongest evidence that *LincRNA-p21* transcription or RNA accumulation contributes to the *cis*-regulation of *p21*, although these experiments cannot distinguish between the act of transcription and the activity of the mature transcript (see Chapter I)(Dimitrova et al., 2014; Lai et al., 2020; Lee and Mendell, 2020).

Although Groff *et al.* disputed this mechanism on the basis of their observation that *LincRNA-p21* was not transcribed in many dysregulated tissues (Goff et al., 2015; Groff et al., 2016), it is worth noting that their strategy substituted a short reporter construct for the entire ~20 kilobase *LincRNA-p21* locus (Groff et al., 2016; Sauvageau et al., 2013); such a large deletion may have disrupted the chromatin architecture of the locus, perturbing the expression of genes that are not regulated by *LincRNA-p21 per se* and disrupting *p21* levels in tissues that do not exhibit a dependence on *LincRNA-p21*. In support of this, embryonic *p21* expression, which this model extensively alters, does not uniformly require p53 signaling (Gartel et al., 2004; Macleod et al., 1995). This suggests that the authors could have disrupted both *LincRNA-p21*-dependent and -independent gene regulation at this locus. Moreover, it is clear from their data that their *lacZ* reporter does not faithfully recapitulate *LincRNA-p21* transcription in all tissue types, casting scrutiny on their claim that their approach decouples transcription from the underlying DNA sequence. In summary, the literature seems to tentatively support a model in which DNA elements within
the LincRNA-p21 locus drive transcription of the LincRNA-p21 lncRNA, which acts in turn to reinforce p21 expression. However, the field lacks a precise mechanistic understanding of this process.

Intriguingly, a number of studies have demonstrated that LincRNA-p21 RNA binds hnRNP-K (Bao et al., 2015; Dimitrova et al., 2014; Huarte et al., 2010; Yoon et al., 2012). This observation, which has been replicated across studies ascribing both cis and trans functions to LincRNA-p21, may help to shed some light on the function of this lncRNA. The heterogeneous nuclear ribonucleoprotein hnRNP-K belongs to a class of nucleic acid-binding proteins implicated in diverse aspects of transcription and RNA processing. In their study, Huarte et al. proposed that hnRNP-K forms a complex with LincRNA-p21 RNA in order to repress target genes (Huate et al., 2010). This was predicated in part on the observation that, in LincRNA-p21-proficient cells, hnRNP-K bound the promoters of genes that were derepressed following LincRNA-p21 knockdown. However, hnRNP-K is known to be enriched at actively transcribed genes (Li et al., 2019), where it has been proposed to interact with Pol II in order to regulate the biogenesis and metabolism of nascent transcripts (Li et al., 2019). Huarte et al. additionally implicated hnRNP-K as a co-mediator of p53-dependent apoptosis. This conclusion is surprising in light of other work showing that hnRNP-K has an anti-apoptotic effect (Revil et al., 2009; Xiao et al., 2013; Yang et al., 2014b). In summary, the literature on hnRNP-K is not consistent with its purported role in LincRNA-p21 mediated trans-repression.

In keeping with hnRNP-K’s more broadly accepted role as a transcriptional activator, Dimitrova et al. set out a model in which hnRNP-K binds nascent or processed LincRNA-p21 in order to stabilize p53 binding at the p21 promoter. In support of this, hnRNP-K has previously been proposed to serve as a coactivator for p53 target genes, including p21 (Moumen et al., 2005, 2013). Indeed, hnRNP-K knockdown strongly inhibits p21 expression following DNA damage (Moumen et al., 2005), meaning that genetic ablation and ASO-mediated knockdown of LincRNA-p21 recapitulate some of the effects of hnRNP-K depletion. Taken together, the biology of hnRNP-K appears to align with a model in which LincRNA-p21 acts in concert with hnRNP-K to promote p21 expression under conditions of p53 activation.
In summary, the molecular mechanism of *cis*-regulation at the *LincRNA-p21* locus remains unclear. In keeping with the growing consensus that lncRNAs play modest, if necessary, roles in fine-tuning cellular processes (Gil and Ulitsky, 2020; Schor et al., 2018; Swarr et al., 2019), recent studies seem to be converging on the consensus that *LincRNA-p21* acts in *cis* to reinforce local gene expression. Moreover, the evidence seems to suggest that some feature of the *LincRNA-p21* transcriptional process or transcript harbors functional activity. As an important consideration for the field, Dimitrova *et al.* used two orthogonal knockdown approaches to identify an apparent role for the transcriptional process, demonstrating the importance of using multiple experimental strategies for studying lncRNA function. Although other work provides valuable insight into the involvement of sequence elements at the *LincRNA-p21* locus, each study implemented a single genetic knockdown strategy that perturbed multiple aspects of lncRNA biogenesis and function and therefore could not resolve a regulatory mechanism. A more rigorous understanding of the mechanism of *cis*-regulation at this locus therefore requires the judicious application of multiple complementary genetic and molecular approaches (see Chapters III and IV).
Table 2.1. Reported roles of *LincRNA-p21* in homeostasis and disease

<table>
<thead>
<tr>
<th>Species</th>
<th>Method(s)</th>
<th>Function</th>
<th>Mechanism</th>
<th>Cis or trans?</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>Promoter deletion, ASOs</td>
<td>Reinforce G1/S cell cycle arrest by promoting <em>p21</em> expression</td>
<td>Nascent or mature RNA recruits hnRNP-K to stabilize p53 at the <em>p21</em> promoter</td>
<td>Cis</td>
<td>(Dimitrova et al., 2014)</td>
</tr>
<tr>
<td>m</td>
<td>Reporter gene substitution</td>
<td>Enhance the expression of multiple neighboring genes, including <em>p21</em></td>
<td>Multiple enhancer elements within the <em>LincRNA-p21</em> locus</td>
<td>Cis</td>
<td>(Groff et al., 2016)</td>
</tr>
<tr>
<td>h</td>
<td>p53RE mutagenesis</td>
<td>Reinforce G1/S cell cycle arrest by promoting <em>p21</em> expression</td>
<td>Unclear</td>
<td>Cis</td>
<td>(Korkmaz et al., 2016)</td>
</tr>
<tr>
<td>m</td>
<td>p53RE mutagenesis</td>
<td>Reinforce G1/S cell cycle arrest by promoting <em>p21</em> expression</td>
<td>Unclear</td>
<td>Cis</td>
<td>(Tesfaye et al.)</td>
</tr>
<tr>
<td>m</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td><em>LincRNA-p21</em> RNA associates with hnRNP-K to enact repression at the promoters of target genes</td>
<td>Trans</td>
<td>(Huarte et al., 2010)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi</td>
<td>Downregulate pro-oncogenic, prosurvival genes</td>
<td><em>LincRNA-p21</em> RNA associates with polysomes to repress the translation of target mRNAs</td>
<td>Trans</td>
<td>(Yoon et al., 2012)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi</td>
<td>Repress NF-κB signaling</td>
<td><em>LincRNA-p21</em> RNA sequesters RELA mRNA</td>
<td>Trans</td>
<td>(Spurlock et al., 2014)</td>
</tr>
<tr>
<td>h,m</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis and repress proliferation</td>
<td><em>LincRNA-p21</em> RNA blocks interactions between p53 and MDM2</td>
<td>Trans</td>
<td>(Wu et al., 2014)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi</td>
<td>Promote glycolysis under conditions of hypoxia</td>
<td><em>LincRNA-p21</em> RNA blocks interactions between HIF-1α and the VHL E3 ligase</td>
<td>Trans</td>
<td>(Yang et al., 2014a)</td>
</tr>
<tr>
<td>m</td>
<td>RNAi, transgene overexpression</td>
<td>Repress cellular reprogramming</td>
<td><em>LincRNA-p21</em> RNA associates with hnRNP-K to maintain repression at pluripotency gene promoters</td>
<td>Trans</td>
<td>(Bao et al., 2015)</td>
</tr>
<tr>
<td>h,m</td>
<td>Transgene overexpression</td>
<td>Promote apoptosis and cell cycle progression</td>
<td><em>LincRNA-p21</em> RNA acts as a sponge for miR-130b</td>
<td>Trans</td>
<td>(He et al., 2015)</td>
</tr>
<tr>
<td>h</td>
<td>Transgene overexpression</td>
<td>Promote PTEN expression</td>
<td><em>LincRNA-p21</em> RNA acts as a sponge for miR-181b</td>
<td>Trans</td>
<td>(Yu et al., 2016)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Suppress metastasis</td>
<td><em>LincRNA-p21</em> RNA associates with EZH2 to repress target genes</td>
<td>Trans</td>
<td>(Liu et al., 2017)</td>
</tr>
<tr>
<td>m</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis by inducing TGF-β signaling</td>
<td><em>LincRNA-p21</em> RNA acts as a sponge for miR-30</td>
<td>Trans</td>
<td>(Tu et al., 2017)</td>
</tr>
<tr>
<td>Cell type</td>
<td>Method</td>
<td>Effect</td>
<td>Mechanism</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Inhibit the Wnt/β-catenin signaling pathway</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-17-5p</td>
<td>Trans</td>
<td>(Yu et al., 2017)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Inhibit proliferation</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-130b</td>
<td>Trans</td>
<td>(Han and Liu, 2018)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-451</td>
<td>Trans</td>
<td>(Tang et al., 2018)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis by upregulating α-synuclein</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-1277-5p</td>
<td>Trans</td>
<td>(Xu et al., 2018)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi</td>
<td>Promote inflammation by upregulating PKC-δ</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-181a/b/c/d</td>
<td>Trans</td>
<td>(Ye et al., 2018)</td>
</tr>
<tr>
<td>m</td>
<td>RNAi, transgene overexpression</td>
<td>Promote proliferation</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-18b</td>
<td>Trans</td>
<td>(Zhang et al., 2019)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-17-5p</td>
<td>Trans</td>
<td>(Ao et al., 2019)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Promote G1/S cell cycle arrest and apoptosis</td>
<td>LincRNA-p21 RNA binds to STAT3 to inhibit its phosphorylation</td>
<td>Trans</td>
<td>(Jin et al., 2019)</td>
</tr>
<tr>
<td>m</td>
<td>Transgene overexpression</td>
<td>Inhibit apoptosis</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-221</td>
<td>Trans</td>
<td>(Sun et al., 2019)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Inhibit proliferation</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-34c</td>
<td>Trans</td>
<td>(Yang et al., 2019b)</td>
</tr>
<tr>
<td>m</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td>LincRNA-p21 RNA binds the CHOP protein</td>
<td>Trans</td>
<td>(Ge et al., 2020)</td>
</tr>
<tr>
<td>m</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td>LincRNA-p21 RNA acts as a sponge for miR-17-5p</td>
<td>Trans</td>
<td>(Wang et al., 2020)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis by inhibiting the Wnt/β-catenin signaling pathway</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Wang et al., 2014)</td>
</tr>
<tr>
<td>h,m</td>
<td>RNAi</td>
<td>Promote apoptosis</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Hall et al., 2015)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Ning et al., 2015)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Tran et al., 2015)</td>
</tr>
<tr>
<td>h</td>
<td>Transgene overexpression</td>
<td>Inhibit cancer stemness and tumorigenicity</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Wang et al., 2015a)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis by inducing ER stress</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Yang et al., 2015)</td>
</tr>
<tr>
<td>m</td>
<td>Transgene overexpression</td>
<td>Reinforce G1/S cell cycle arrest by promoting p21 expression</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Zheng et al., 2015)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi</td>
<td>Promote angiogenesis</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Castellano et al., 2016)</td>
</tr>
<tr>
<td>r</td>
<td>RNAi</td>
<td>Inhibit proliferation</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Han et al., 2016)</td>
</tr>
<tr>
<td>RNAi/Transgene Overexpression</td>
<td>Effect</td>
<td>Unclear</td>
<td>Unclear</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Inhibit invasion and metastasis by repressing Notch signaling</td>
<td></td>
<td></td>
<td>(Jia et al., 2016)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Inhibit Wnt/β-catenin signaling</td>
<td></td>
<td></td>
<td>(Yang et al., 2016b)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Promote proliferation</td>
<td></td>
<td></td>
<td>(Zhou et al., 2016)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Inhibit invasion and metastasis by repressing YAP signaling</td>
<td></td>
<td></td>
<td>(Chen et al., 2017)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Inhibit invasion and metastasis by repressing miR-9</td>
<td></td>
<td></td>
<td>(Ding et al., 2017)</td>
<td></td>
</tr>
<tr>
<td>Transgene overexpression</td>
<td>Reinforce G1/S cell cycle arrest by promoting p21 expression</td>
<td></td>
<td></td>
<td>(Peng et al., 2017)</td>
<td></td>
</tr>
<tr>
<td>RNAi</td>
<td>Promote survival under conditions of hypoxia</td>
<td></td>
<td></td>
<td>(Shen et al., 2017)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Inhibit proliferation by repressing glycolysis</td>
<td></td>
<td></td>
<td>(Wang et al., 2017a)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td></td>
<td>Stabilize p53 binding at promoters of target genes</td>
<td>(Wang et al., 2017b)</td>
<td></td>
</tr>
<tr>
<td>RNAi</td>
<td>Promote senescence by inhibiting the Wnt/β-catenin signaling pathway</td>
<td></td>
<td></td>
<td>(Xia et al., 2017)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td></td>
<td></td>
<td>(Zhou and Chen, 2017)</td>
<td></td>
</tr>
<tr>
<td>RNAi</td>
<td>Inhibit angiogenesis by repressing NF-KB signaling</td>
<td></td>
<td></td>
<td>(Cui et al., 2018)</td>
<td></td>
</tr>
<tr>
<td>RNAi</td>
<td>Promote proliferation and inhibit apoptosis</td>
<td></td>
<td></td>
<td>(Meng et al., 2018)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Promote apoptosis under conditions of hypoxia</td>
<td></td>
<td></td>
<td>(Xia et al., 2018)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Promote senescence</td>
<td></td>
<td></td>
<td>(Xie et al., 2018)</td>
<td></td>
</tr>
<tr>
<td>Transgene overexpression</td>
<td>Inhibit proliferation by repressing the Wnt/β-catenin signaling pathway</td>
<td></td>
<td></td>
<td>(Chen et al., 2019)</td>
<td></td>
</tr>
<tr>
<td>RNAi, transgene overexpression</td>
<td>Inhibit tumorigenesis</td>
<td></td>
<td></td>
<td>(Geng et al., 2019)</td>
<td></td>
</tr>
<tr>
<td>RNAi</td>
<td>Promote senescence by reinforcing p21 expression</td>
<td></td>
<td></td>
<td>(Guo et al., 2019)</td>
<td></td>
</tr>
<tr>
<td>Transgene overexpression</td>
<td>Promote apoptosis</td>
<td></td>
<td></td>
<td>(Hu et al., 2019)</td>
<td></td>
</tr>
<tr>
<td>RNAi</td>
<td>Promote apoptosis</td>
<td></td>
<td></td>
<td>(Wang et al., 2019b)</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Inhibit apoptosis</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Yang et al., 2019b)</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------</td>
<td>------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------------------</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Inhibit apoptosis</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Yang et al., 2019a)</td>
</tr>
<tr>
<td>h</td>
<td>Transgene overexpression</td>
<td>Inhibit invasion and metastasis by repressing VEGF and HIF-1α</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Ye et al., 2019)</td>
</tr>
<tr>
<td>h</td>
<td>Transgene overexpression</td>
<td>Promote G1/S cell cycle arrest</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Zhang et al., 2019)</td>
</tr>
<tr>
<td>h</td>
<td>RNAi, transgene overexpression</td>
<td>Inhibit proliferation</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Zhou et al., 2019)</td>
</tr>
<tr>
<td>m</td>
<td>Transgene overexpression</td>
<td>Inhibit apoptosis</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Yang et al., 2020)</td>
</tr>
<tr>
<td>m</td>
<td>CRISPRi, transgene overexpression</td>
<td>Promote apoptosis</td>
<td>Unclear</td>
<td>Unclear</td>
<td>(Chen et al., 2021)</td>
</tr>
</tbody>
</table>
Chapter III: Development and validation of an inducible self-cleaving ribozyme

Introduction

*Cis*-acting lncRNAs exhibit regulatory functions that are uniquely dependent on their position within the genome (see Chapter I). Given the intimate relationship between DNA sequence elements, the transcriptional process, and the accumulation of mature transcripts, molecular dissection of *cis*-regulatory lncRNA function presents complex technical challenges. Although the current suite of molecular and genetic tools offers a large number of strategies for manipulating different elements of lncRNA biology (see Chapter I), the ever-expanding mechanistic diversity of *cis*-acting lncRNAs necessitates the constant development of new experimental approaches. In this chapter, we describe the development and validation of an inducible self-cleaving ribozyme as a tool for specific RNA destabilization. We demonstrate that the active—but not inert—form of this ribozyme mediates the depletion of an exogenous *LincRNA-p21* construct in multiple murine cell lines, illustrating its suitability as a new tool for lncRNA loss-of-function.

Results

*Design of a synthetic LincRNA-p21 construct with an inducible Twister element*

The self-cleaving ribozyme Twister presented an attractive candidate for targeted transcript degradation on account of several factors. Firstly, Twister family ribozymes exhibit faster cleavage kinetics than other commonly-used ribozyme tools, including the minimal Hammerhead and Hepatitis Delta ribozymes (Diegelman-Parente and Bevilacqua, 2002; Roth et al., 2014; Stage-Zimmermann and Uhlenbeck, 1998). Secondly, Twister ribozyme cleavage does not require the presence of specific divalent metal ions (Panja et al., 2017; Roth et al., 2014), suggesting that Twister might function in a variety of cellular environments. Finally, a dinucleotide inversion renders this ribozyme catalytically inactive,
providing a control for the possibility that Twister might be acting through some mechanism other than its intrinsic auto-scissile activity (Fig. 5A)(Roth et al., 2014).

Christiane Olivero (C. Olivero) therefore designed an invertible Twister element that was inserted into an exogenous LincRNA-p21 overexpression construct (Fig. 5B). In its constitutively inactive form, this element features the inverted Twister or catalytically dead mutant sequence flanked by two pairs of noncomplementary loxP sites. Cre recombinase can subsequently induce the stable inversion of this genetic element, restoring Twister to its active sense orientation. This novel genetic tool was inserted into an intronless LincRNA-p21 construct in a retroviral expression vector. Its position within exon 1 was predicted to minimally disrupt the secondary structure of the host transcript, permitting efficient transcription and folding prior to ribozyme cleavage (Fig. 5C).

**Cre recombinase restores Twister to its active orientation**

In order to investigate the viability of our genetic tool, we first asked whether Cre expression could efficiently restore the inverted Twister sequence to its sense orientation. To test this, we stably infected each construct into Rosa26-CreERT2 MEFs and lung adenocarcinoma cells and isolated genomic DNA from untreated cells as well as cells treated with the estrogen precursor 4-hydroxytamoxifen, which promotes translocation of the constitutively-expressed CreER<sup>T2</sup> fusion protein into the nucleus, where it
can mediate recombination. PCR genotyping suggested that 4-hydroxytamoxifen treatment stably restored the Twister and mutant sequences to their sense orientation (Fig. 6A). To establish the specificity of our genotyping approach, we incubated PCR amplicons from untreated and 4-hydroxytamoxifen-treated cells with the restriction enzyme XhoI. Enzymatic digestion yielded fragments of the expected sizes in both untreated and treated cells, confirming that our approach was fundamentally successful (Fig. 6B).

Figure 6. The invertible Twister element destabilizes the exogenous LincRNA-p21 construct. (A) PCR genotyping of lung adenocarcinoma cells stably expressing the indicated constructs and cultured for 24 hours in the absence or presence of 4-hydroxytamoxifen to induce recombination. Expected bands: 402 nucleotides (inverted); 274 nucleotides (reverted); 86 nucleotides (endogenous). The asterisk (*) indicates a nonspecific product. EV=empty vector; MUT=catalytically inactive Twister mutant; TWI=environmental Twister sequence. (B) Enzymatic digestion of PCR-genotyping products from MEFs treated as described in (A) confirms the identity of the products. Expected bands: 402 nucleotides (-XhoI, inverted); 274 nucleotides (-XhoI, reverted); 229 (*) and 173 (**) nucleotides (+XhoI, inverted); 87 (†) and 187 (‡) nucleotides (+XhoI, reverted). M=catalytically inactive Twister mutant; T=environmental Twister sequence. (C) PCR analysis of cDNA from lung adenocarcinoma cells described in (A) using primers spanning the Twister cleavage site. Expected bands: 402 nucleotides (inverted); 274 nucleotides (reverted). (D) RT-qPCR analysis of cDNA from MEFs described in (B) using primers upstream (US) or downstream (DS) of the Twister cleavage site. LincRNA-p21 levels were calculated relative to Gapdh and normalized to untreated cells expressing LincRNA-p21 harboring the catalytically dead Twister sequence.
Twister cleavage inhibits the accumulation of the exogenous LincRNA-p21 transcript

Ultimately, we wished to examine the functional consequences of Twister restoration. In service of this, we first asked whether Twister mediated LincRNA-p21 RNA cleavage in its sense, but not antisense, orientation. To qualitatively assess the extent of cleavage in transcripts arising from each sample, we performed gel electrophoresis on RT-PCR products generated using primers spanning the Twister cleavage site. Consistent with our expectation, the restored Twister sequence showed a significant reduction in signal relative to the restored mutant sequence (Fig. 6C). Of note, antisense Twister did not appear to self-cleave in spite of its near-palindromic sequence. These results demonstrate that Twister exhibits specific autocleavage activity only when present in its sense orientation.

Next, we sought to understand how ribozyme cleavage affects the accumulation of the resulting LincRNA-p21 RNA fragments. To test this, we used RT-qPCR to analyze LincRNA-p21 expression before and after Twister restoration (Fig. 6D). Transcripts harboring the antisense Twister sequence were expressed at similar levels to transcripts harboring either orientation of the catalytically dead mutant sequence. By contrast, the presence of the restored Twister sequence reduced exogenous LincRNA-p21 levels by ~75% suggesting that Twister cleavage specifically depletes both the upstream and downstream products. Taken together, these results support a model in which Twister’s auto-scissile ability mediates the knockdown of host transcripts in diverse cellular environments.

Data summary and conclusion

These data firmly establish the inducible Twister ribozyme as a viable tool for targeted transcript depletion. Using an invertible genetic system, we showed that Twister self-cleaves and destabilizes host transcripts only when restored to its sense orientation. These findings demonstrate the broad applicability of both inducible and constitutively active Twister as tools for studying the role of IncRNAs in diverse functional and genomic contexts. The inducible Twister element described here has the potential to reveal the temporal consequences of RNA depletion on local gene regulation, facilitating efforts to investigate the complex relationship between IncRNA expression and changes in the regulatory landscape of nearby
protein-coding genes. Additionally, Twister failed to mediate transcript cleavage and destabilization in its antisense orientation, suggesting that it may represent a powerful strategy for dissecting the function of antisense lncRNAs that overlap with protein-coding genes. Finally, the ability to replicate these findings in both MEFs and cancer cells indicates that Twister cleavage is robust across a variety of cellular environments.

Although we have shown that our approach is fundamentally feasible, the effects of ribozyme cleavage are likely context specific and may vary based on the position of the ribozyme relative to the TSS, TTS, or splice sites (Bird et al., 2005; NOTT et al., 2003). Additionally, features such as the secondary structure of the host transcript may affect the efficiency of autocleavage and the accessibility of upstream and downstream cleavage products to exonucleases (Akiyama et al., 2016). As a result, the effects of ribozyme placement on transcript cleavage and destabilization will have to be validated on a case-by-case basis. Later in this work, we will discuss our efforts to characterize the effects of co-transcriptional Twister cleavage on the transcriptional landscape of the endogenous LincRNA-p21 locus (see Chapter IV).
Chapter IV: Elucidating the molecular mechanism of cis-regulation by LincRNA-p21

Material from this chapter is being prepared for publication as follows: Winkler, L.N., Zimmer, J.T., Williams, A., Simon, M.D., and Dimitrova, N.M. “Functional elements of the cis-regulatory LincRNA-p21.” Co-author contributions are specified in figure legends and the main text.

Introduction

While it is clear that many IncRNA loci harbor cis-regulatory potential, the mechanisms through which they effect activation and repression remain largely uncharacterized. In particular, the limitations of conventional molecular and genetic approaches have hindered attempts to deconvolve the effects of (1) the mature RNA, (2) transcriptional and co-transcriptional processes, and (3) DNA elements within the IncRNA locus (Bassett et al., 2014; Kopp and Mendell, 2018). Due to the intrinsically interlinked nature of these phenomena, it remains unclear whether the majority of genomic transcription is functional or simply reflects the stochastic association of Pol II with regions of open chromatin near regulatory DNA elements (Engreitz et al., 2016; Paralkar et al., 2016; Struhl, 2007).

The p53-inducible IncRNA LincRNA-p21 has previously been shown to act in cis to reinforce the expression of the nearby p53 target and G1/S checkpoint inhibitor p21/Cdkn1a (Dimitrova et al., 2014; Groff et al., 2016; Korkmaz et al., 2016; Tesfaye et al.). However, the molecular mechanism underlying this cis-regulatory relationship remains unclear (see Chapter II). Specifically, the use of genetic approaches that both abolish LincRNA-p21 transcription and destroy DNA elements within the locus has confounded efforts to elucidate the molecular grammar through which LincRNA-p21 communicates complex regulatory information to p21. This chapter describes our efforts to resolve this controversy by implementing a powerful suite of genetic and molecular techniques to manipulate different aspects of LincRNA-p21 biogenesis. Using complementary experimental approaches, we show that full-length transcription, splicing, and RNA accumulation are dispensable for cis-regulation. By
Figure 7. The LincRNA-p21 p53RE reinforces p21 expression in a p53-dependent manner in MEFs (A) Schematic of p53 expression restoration and stabilization in p53<sup>LSL/LSL</sup>: Rosa26-CreER<sup>12</sup> MEFs. (B) Top: schematic of LincRNA-p21 p53RE mutagenesis showing the location of the p53RE (bold), the PAM site (gray box), and Cas9 cleavage site (black arrow). Bottom: representative p53RE mutations identified by Sanger sequencing in cells described in (A). Mutagenesis rate=100%. (C and D) RT-qPCR analysis of (C) LincRNA-p21 and (D) p21 expression in p53RE restorable MEFs that
were treated with 0.5 µM 4-hydroxytamoxifen to restore p53 expression 24 hours prior to treatment with 24 hour 0.1 or 0.5 µM doxorubicin. Data represent the mean ± SEM (n=3 biological replicates); *p<0.05, ***p<0.001, paired t-test. For all samples, LincRNA-p21 and p21 levels were calculated relative to Gapdh. (E) Representative immunoblot for p21 protein in MEFs treated as described above. α-tubulin served as a loading control. (F) Proliferation of the indicated p53-restorable MEFs cultured in the absence or presence of 0.5 µM 4-hydroxytamoxifen. Data represent the mean ± SEM (n=3 biological replicates) n.s., not significant, p<0.0001, paired t-test. (G) Top: schematic of LincRNA-p21 p53RE mutagenesis showing the location of the p53RE (bold), the PAM site (gray box), and Cas9 cleavage site (black arrow). Bottom: representative p53RE mutations identified by Sanger sequencing in primary MEFs. Mutagenesis rate=90%. (H and I) RT-qPCR analysis for spliced LincRNA-p21 and p21 RNA in the indicated MEFs. Data represent the mean ± SEM (n=4 biological replicates); *p<0.05, **p<0.01, paired t-test. LincRNA-p21 and p21 levels were calculated relative to Gapdh and normalized to Con.

contrast, we implicate transcription through a minimal sequence element in exon 1 of LincRNA-p21 as a key mediator of regulatory activity at the locus. Our findings unveil the complex molecular grammar of cis-regulation at the LincRNA-p21 locus and further support a functional role for pervasive genomic transcription.

Results

The LinRNA-p21 p53RE reinforces p21 expression

The p53-inducible IncRNA LincRNA-p21 harbors a p53RE within its first exon. (Fig. 3)(Huart et al., 2010). Since p53 binding within noncoding transcriptional units has previously been shown to activate the expression of distant p53 targets (Melo et al., 2013), we wondered whether this DNA element might reinforce p21 expression in response to p53 activation by genotoxic or oncogenic stress. In order to test this, we used CRISPR/Cas9 to mutagenize the p53RE in a panel of murine cell lines. Under this approach, we stably infected cells with lentiviral constructs co-expressing spCas9 and either a nontargeting control gRNA (Con) or a gRNA targeting the p53RE (ΔRE). We reasoned that CRISPR/Cas9-mediated mutagenesis of the p53RE in ΔRE cells would abolish p53 binding at the LincRNA-p21 promoter, leading to a reduction in LincRNA-p21 and p21 transcription only in the presence of p53 signaling.

We first performed CRISPR/Cas9-mediated mutagenesis in p53-restorable MEFs, a powerful genetic system that enabled us to directly test the role of p53 in the cis-regulatory relationship between
In this system, an excisable STOP cassette (LoxP-STOP-LoxP; LSL) in the first intron of p53 renders cells functionally p53 deficient (Ventura et al., 2007). These cells also express a Cre-ER\textsuperscript{T2} fusion protein that mediates the deletion of the LSL upon treatment with 4-hydroxytamoxifen, permitting p53 expression. Genotoxic stress, including DNA damage induced by doxorubicin, further stabilizes p53, leading to the induction of its target genes (Fig. 7A). To confirm p53RE mutagenesis in ΔRE MEFs, we performed Sanger sequencing on gDNA from a bulk population of cells. We observed a 100% mutagenesis rate in ΔRE cells, which showed a corresponding loss in LincRNA-p21 induction following p53 restoration and treatment with doxorubicin (Fig. 7B-C). Consistent with a p53-dependent cis-regulatory role for the p53RE, p53-proficient ΔRE MEFs exhibited a significant decrease in p21 levels in both the absence of DNA damage and following treatment with low levels of doxorubicin (Fig. 7D-E). p53-deficient ΔRE MEFs, however, showed normal p21 expression in both the absence and presence of doxorubicin. In keeping with p21’s role as a cell cycle checkpoint inhibitor, p53-proficient ΔRE MEFs proliferated faster than p53-proficient Con cells (Fig. 7F). By contrast, p53-deficient ΔRE and Con MEFs proliferated at similar rates, further indicating that the p53RE acts as a cis-regulatory element for LincRNA-p21 and p21 only in the presence of p53 signaling.

We next validated these results in primary wildtype E13.5 MEFs. Following mutagenesis, ΔRE primary MEFs showed a 90% mutagenesis rate as well as a 48% reduction in LincRNA-p21 expression relative to Con primary cells (Fig. 7G-H). In keeping with previous findings, we observed that ΔRE cells also exhibited a 27% decrease in p21 expression (Fig. 7I). Taken together, these results indicate that the LincRNA-p21 p53RE promotes the expression of LincRNA-p21 and p21 in a p53-dependent manner in the presence of genotoxic stress.

Finally, we performed p53RE mutagenesis in a p53-restorable lung adenocarcinoma cell line, a model for investigating p53 signaling in response to oncogenic stress. Like the p53-restorable MEFs, these cells harbor a LSL element that inhibits p53 expression, and constitutively express Cre-ER\textsuperscript{T2}. Additionally, these cells express the oncogenic G12D Kras allele (Fig. 8A). Using Sanger sequencing, we observed a 100% mutagenesis rate for the p53RE in these cells (Fig. B). Although loss of the p53RE
mediated a ~75% reduction in LincRNA-p21 levels in the presence of p53 signaling, we did not observe any changes in p21 expression (Fig. These findings suggest that this cell line is not a tractable model for studying the cis-regulatory relationship between LincRNA-p21 and p21 (Fig 8C-E).

**Generation and initial characterization of four novel LincRNA-p21 loss-of-function mouse models**

Our mutagenesis experiments clearly show that the LincRNA-p21 p53RE acts in cis to activate p21 expression in response to p53 signaling. Since p53RE mutagenesis also abrogates LincRNA-p21 expression, we wondered whether the p53RE acts directly or mediates its cis-regulatory effects through the LincRNA-p21 transcriptional process or RNA. We therefore set out to investigate the regulatory
contributions of each of these elements. To accomplish this, we implemented multiple complementary genetic strategies for knocking down LincRNA-p21 transcription (see Chapter II). In collaboration with Adam Williams (A. Williams; the Jackson Laboratory for Genomic Medicine), we used CRISPR/Cas9-mediated genome engineering to insert a 49 nucleotide synthetic PAS or the 74 nucleotide Twister ribozyme (TWI; see Chapter III) into exon 1 of the endogenous LincRNA-p21 locus of C57BL/6J mice (Fig. 9A)(Engreitz et al., 2016; Levitt et al., 1989; Roth et al., 2014). Briefly, A. Williams electroporated wildtype blastocysts with Cas9, a gRNA targeting the LincRNA-p21 locus, and a <200 nucleotide ssDNA donor template comprising the knock-in sequence flanked by two homology arms (Richardson et al., 2016). Under this strategy, the ssDNA donor template mediated the homology-directed repair (HDR) of the double-strand break, thus inserting the knock-in sequence into exon 1 of LincRNA-p21 (Williams et al., 2016).

This process generated chimeric mice, some of which harbored the correct PAS or Twister insertion (Fig. 9B-C). Additionally, we identified two mouse lines with small deletions that partially overlap LincRNA-p21 exon 1. These novel alleles comprise a 151 nucleotide deletion (Δ151) that spans the exon 1-intron junction and a 127 nucleotide deletion (Δ127) that falls entirely within exon 1 (Fig. 9D). Critically, these deletions leave the canonical p53RE intact. Sanger sequencing further confirmed the identity of each allele (Fig. 9E). All four founders were back-crossed once with wildtype C57BL/6J mice, yielding heterozygotes carrying each allele of interest. These offspring were viable and did not display any abnormalities, consistent with previous findings (Fig. 10)(Dimitrova et al., 2014).
Figure 9. Generation and identification of novel LincRNA-p21 knockdown alleles. (A) Schematic of the LincRNA-p21 Exon 1 showing the relative positions of the gRNA, PAS and Twister donor templates, genotyping primers (black arrows), and p53RE (red asterisk). (B and C) PCR genotyping of chimeric mice generated using the knock-in strategy described in (A). The tag numbers of the Twister (B) and PAS (C) founder mice are indicated in red. Expected bands: 558 nucleotides (wildtype); 632 nucleotides (Twister); 607 nucleotides (PAS). (C) The tag numbers of two additional mice with 127 and 151 nucleotide Exon 1 deletions are shown in red and orange. Expected bands: 431 nucleotides (Δ127); 407 nucleotides (Δ151). (D) Schematic showing the location of the novel 127 and 151 nucleotide deletions relative to Exon 1, including a conserved region. The p53RE is indicated in red. (E) Sanger sequencing across the PCR amplicons generated as described in (B) and (C) confirms the identity of each allele.
Figure 10. Validation of four novel *LincRNA-p21* knockdown alleles. (A) Left, mating strategy for backcrossing chimeric founder mice (red and gray) with wildtype C57BL/6J mice (white) in order to generate heterozygotes (red and white stripes) carrying each *LincRNA-p21* knockdown allele. Right, mating strategy for generating wildtype (white), heterozygous (red and white stripes), and homozygous (red) animals and MEFs. (B) PCR genotyping confirms that heterozygote matings described in (A) right yield offspring of all possible genotypes. Expected bands: 253 nucleotides (PAS), 204 nucleotides (wildtype); 278 nucleotides (Twister), 204 nucleotides (wildtype); 177 nucleotides (Δ151), 328 nucleotides (wildtype); 203 nucleotides (Δ127), 330 nucleotides (wildtype). (C) Genotype frequencies for offspring from the indicated *LincRNA-p21* heterozygote matings. The expected numbers derive from a Mendelian distribution.
Figure 11. Novel genetic tools abolish RNA accumulation and attenuate transcription and processing at the LincRNA-p21 locus. (A) Left, schematic of the murine LincRNA-p21 locus highlighting the relative positions of the LincRNA-p21 p53RE (*), the PAS/Twister insertion site, and primers for detecting total LincRNA-p21 RNA. Right, schematic showing the position of the novel
deletions within exon 1. (B-E) RT-qPCR analysis of spliced LincRNA-p21 RNA in MEFs of the indicated genotypes cultured for 8 hours in the absence or presence of 0.5 µM doxorubicin. Data represent the mean ± SEM (PAS, n=13; TWI, n=7; Δ151, n=3; Δ127, n=18; biological replicates). **p<0.01, ***p<0.001, and ****p<0.0001, paired t-test. LincRNA-p21 levels were calculated relative to Gapdh and normalized to untreated wildtype cells. (F-I) RT-qPCR detection of total LincRNA-p21 RNA in MEFs of the indicated genotypes cultured for 24 hours in the absence or presence of 0.5 µM doxorubicin. Data represent the mean ± SEM (PAS, n=3 biological replicates; TWI, n=3 biological replicates; Δ151, n=3 technical replicates; Δ127, n=4 biological replicates). For each primer pair, LincRNA-p21 levels were calculated relative to Gapdh and normalized to untreated wildtype cells.

*p21 does not require full-length transcription, splicing, or accumulation of LincRNA-p21 RNA*

To characterize the effects of our genetic tools and deletion alleles on LincRNA-p21 biogenesis, we first analyzed the expression of spliced LincRNA-p21 RNA in primary MEFs. We observed that both knock-in and deletion MEFs exhibited a near-complete loss of mature LincRNA-p21 accumulation in both the absence and presence of doxorubicin (Fig. 11B-E). These results indicate that spliced LincRNA-p21 RNA is largely absent from our mutant cells; however, we also wished to understand how each mutation affects the expression of total (spliced and unspliced) LincRNA-p21. We therefore monitored steady-state RNA expression across the LincRNA-p21 locus and observed that all four of our loss-of-function alleles mediate a significant loss of total LincRNA-p21 RNA (Fig. 11F-I).

We subsequently confirmed that the PAS and Twister had similar effects on LincRNA-p21 expression in vitro and in vivo. We compared spliced LincRNA-p21 RNA levels across a panel of tissues from adult wildtype and knock-in mice (Fig. 12). Both knock-in mouse strains showed a marked loss of LincRNA-p21 RNA in all tissues, recapitulating the effects of a previously-described LincRNA-p21 promoter deletion mouse model (LincRNA-p21−/−)(Dimitrova et al., 2014). These results indicate that our genetic tools are active in vivo and mediate the loss of spliced LincRNA-p21 RNA in a variety of different cellular contexts.
Figure 12. The PAS and Twister mediate the loss of LincRNA-p21 RNA in vivo. (A) RT-qPCR analysis of spliced LincRNA-p21 levels in the indicated tissues from LincRNA-p21^{PAS/PAS} and LincRNA-p21^{+/+} mice. Data represent the mean ± SEM (n=3 technical replicates). (B) RT-qPCR analysis of spliced LincRNA-p21 levels in the indicated tissues from LincRNA-p21^{TWI/TWI} and LincRNA-p21^{+/+} mice. Data represent the mean ± SEM (n=3 biological replicates). (C) RT-qPCR analysis of spliced LincRNA-p21 levels in the indicated tissues from LincRNA-p21^{-/-} and LincRNA-p21^{+/+} mice. Data represent the mean ± SEM (n=3 technical replicates). For all samples, LincRNA-p21 levels are expressed relative to Gapdh.
Since self-cleaving ribozymes are an emerging strategy for destabilizing endogenous lncRNAs (Camblong et al., 2009; Tuck and Bühler, 2021; Tuck et al., 2018; Wery et al., 2018b), it was not clear \textit{a priori} how Twister cleavage would affect \textit{LincRNA-p21} transcription. To address this key question, Joshua Zimmer (J. Zimmer; Yale University) used transient transcriptome (TT)-qPCR to monitor transcription at multiple sites across the \textit{LincRNA-p21} locus in metabolically-labeled MEFs (Fig. 13).

Strikingly, Twister appeared to suppress transcription from the \textit{LincRNA-p21} promoter by ~60% and incurred a ~90% reduction in transcription near the TTs. Based on these results, we concluded that Twister mediates the co-transcriptional knockdown of the \textit{LincRNA-p21} RNA. Collectively, our steady-state and TT-qPCR experiments define the scope of our loss-of-function alleles on \textit{LincRNA-p21} transcription, processing, and accumulation \textit{in vitro} and \textit{in vivo}.

We next sought an alternative approach to confirm that our novel genetic models deplete \textit{LincRNA-p21} expression. We therefore used single-molecule RNA fluorescence \textit{in situ} hybridization (smRNA-FISH) to visualize intronic and exonic \textit{LincRNA-p21} RNA in wildtype and knockdown MEFs. In keeping with the results of previous experiments, we observed a substantial loss of foci across all of our knockdown cell lines (Fig. 14). Interestingly, although the majority of \textit{LincRNA-p21}^{TWI/TWI} cells had no foci, a modest ~15% of scored cells at least one intronic or exonic focus (Fig. 14B). This corroborates
Figure 14. Single-molecule RNA fluorescence in situ hybridization (smRNA FISH) confirms the loss of RNA signal in novel LincRNA-p21 loss-of-function mutants. (A) smRNA FISH detection of LincRNA-p21 RNA with intronic (green) and exonic (red) probes in indicated MEFs treated for 24 hours with 25 µM etoposide. DNA was counterstained with DAPI. (B-E) Frequency distribution for the number of intronic (left) and exonic (right) foci in each nucleus of the indicated genotype. MEFs were treated as described in (A).

our finding that Twister cleavage permits some degree of transcription through the LincRNA-p21 locus.

Ultimately, we wished to test the role of LincRNA-p21 biogenesis in regulating p21 expression. To accomplish this, we examined p21 expression in wildtype and littermate LincRNA-p21TWI/TWI, LincRNA-p21PAS/PAS, or LincRNA-p21Δ151/Δ151 MEFs. Curiously, neither premature transcriptional termination (the PAS) nor the loss of full-length transcription (TWI), canonical splicing (Δ151), or RNA accumulation (all three models) affected p21 RNA or protein levels in either the absence or presence of doxorubicin (Fig. 15A-G). Consistent with these results, LincRNA-p21PAS/PAS and LincRNA-p21TWI/TWI MEFs did not display increased proliferation relative to wildtype littermate controls, further suggesting that p21 levels are unaltered (Fig. 15H-I). Together, these results reveal that p21 transcription is insensitive to full-length LincRNA-p21 transcription, processing, and accumulation.

Organismal data largely validates our in vitro findings. Although LincRNA-p21−/− animals show a pronounced decrease in p21 RNA levels across a panel of tissues relative to wildtype controls, we could not detect any significant differences in p21 expression between the majority of wildtype and LincRNA-p21TWI/TWI tissue samples (Fig. 16). Curiously, we did observe a statistically significant decrease in p21 expression in brain samples from LincRNA-p21TWI/TWI animals. Since LincRNA-p21-deficient mice have previously been reported to clasp their hindlimbs when lifted by the tail (Sauvageau et al., 2013), we wondered whether this brain-specific decrease in p21 levels might mediate a similar phenotype in LincRNA-p21TWI/TWI mice. However, we failed to observe hindlimb clamping or other behavioral abnormalities in LincRNA-p21TWI/TWI mice or any of our other mutants (data not shown).
Figure 15. *p21* does not require splicing or accumulation of the full-length LincRNA-*p21* transcript. (A-C) RT-qPCR analysis of *p21* mRNA levels in MEFs of the indicated genotypes cultured for 8 hours in the absence or presence of 0.5 µM doxorubicin. Data represent the mean ± SEM (PAS, n=11; TWI, n=9; Δ151, n=3; biological replicates); n.s., not significant, paired t-test. For all samples, *p21* levels were calculated relative to *Gapdh* and normalized to untreated wildtype cells. (D and E) Representative immunoblot for *p21* and *p53* protein from MEFs treated as described in (A and B). HSP90 served as a loading control. (F and G) Densitometric analysis of *p21* protein expression in cells described in (D and E). *p21* levels were normalized to HSP90. Data represent the mean ± SEM (PAS, n=8; TWI, n=7; biological replicates); n.s., not significant, paired t-test. (H and I) Proliferation of the indicated pairs of MEFs isolated from littermates. Data are representative of n=2 independent growth curve experiments.
Figure 16. Twister does not affect p21 levels in vivo. (A) RT-qPCR analysis of p21 levels in the indicated tissues from LincRNA-p21TWI/TWI and LincRNA-p21+/+ mice. Data represent the mean ± SEM (n=3 biological replicates); *p<0.05, paired t-test. (B) RT-qPCR analysis of p21 levels in the indicated tissues from LincRNA-p21−/− and LincRNA-p21+/+ mice. Data represent the mean ± SEM (n=3 technical replicates). For all samples, p21 levels are expressed relative to Gapdh.

ASOs targeting LincRNA-p21 likely incur nonspecific effects on p21 expression.

Our results, which show that transcription, processing, and accumulation of full-length LincRNA-p21 RNA does not participate in cis-regulation, conflict with previous findings supporting a role for these processes (Dimitrova et al., 2014). This functionality was identified using ASO gapmers targeting sequences in the intron (ASO-I) and second exon (ASO-E2), eight and 19 kilobases downstream of the PAS/Twister insertion site, respectively (Fig. 4). Consistent with prior work, we found that these ASOs mediated a >90% reduction in the expression of spliced LincRNA-p21 expression and a concurrent 30-40% decrease in p21 levels (Fig. 17A-B)(Dimitrova et al., 2014). To further understand the effects of these ASOs on LincRNA-p21 expression, we monitored steady-state RNA levels across the locus (Fig. 17C). Similar to the PAS and Twister, both ASOs mediated a significant loss of RNA from exon 2.

Although we did not observe a significant knockdown in total LincRNA-p21 levels using primers
targeting the exon 1-intron junction or intron, we noted a puzzling decrease in LincRNA-p21 expression from the upstream TSS.

We next investigated how ASO cleavage affects p21 levels in knock-in MEFs. As expected, ASO treatment significantly abrogated LincRNA-p21 expression in wildtype control MEFs (Fig. 17D, F). We could not detect additional knockdown in our mutant cells, which already showed a near-complete loss of spliced LincRNA-p21 expression. In all samples, however, ASO-I and ASO-E2 mediated a reduction in p21 levels relative to the nontargeting control ASO (Fig. 17E, G). This perplexing result points to one of two conclusions: either a small degree of transcription through the LincRNA-p21 locus is sufficient to reinforce p21 expression or ASO-I and ASO-E2 mediate a nonspecific decrease in p21 expression.
Figure 17. ASO gapmers targeting downstream of the PAS/Twister insertion site abrogate p21 expression in LincRNA-p21\textsuperscript{PAS/PAS} and LincRNA-p21\textsuperscript{TWI/TWI} MEFs. (A and B) RT-qPCR analysis of (A) LincRNA-p21 and (B) p21 expression in wildtype MEFs nucleofected with the indicated ASOs. Data represent the mean ± SEM (n=3 biological replicates), *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, paired t-test. (C) RT-qPCR for total LincRNA-p21 RNA using the indicated primer pairs. Red arrows indicate the location of each ASO relative to the primer sites. Data represent the mean ± SEM (n=3 biological replicates), n.s., not significant, *p<0.05, **p<0.01, ***p<0.001, paired t-test. (D and E) RT-qPCR analysis of (D) LincRNA-p21 and (E) p21 expression in LincRNA-p21\textsuperscript{PAS/PAS} and wildtype MEFs nucleofected with the indicated ASOs. Data represent the mean (n=3 technical replicates). (F and G) RT-qPCR analysis of (F) LincRNA-p21 and (G) p21 levels in LincRNA-p21\textsuperscript{TWI/TWI} and wildtype MEFs nucleofected with the indicated ASOs. Data represent the mean (n=3 technical replicates).
A conserved sequence element reinforces p21 expression in cis

Since transcription, processing, and accumulation of the full length LincRNA-p21 RNA appear to be dispensable for cis-regulation, we questioned whether other elements of the locus might promote p21 expression, as has been shown for human LincRNA-p21 (Korkmaz et al., 2016). Although LincRNA-p21 exhibits little overall sequence conservation, a ~200 nucleotide region within exon 1 is moderately conserved and shares significant homology with its counterpart in other mammals (Chillón and Pyle, 2016) (Fig. 18A). We hypothesized that this region, which includes the p53RE, might harbor additional cis-regulatory potential.

In order to test this, we investigated p21 expression in primary LincRNA-p21Δ127/Δ127 MEFs, which harbor a deletion that spans the majority of the conserved region without disrupting the canonical p53RE (Fig. 9D). Strikingly, deletion of the conserved region mediated significant decreases in p21 RNA and protein levels in both the absence and presence of doxorubicin (Fig. 18B-D). In keeping with the role of p21 in cell cycle progression, the reduction in p21 expression had clear functional consequences: we observed that mutant MEFs showed enhanced proliferation relative to wildtype littermate controls (Fig. 18E)(Dimitrova et al., 2014). Together, these results indicate a requirement for sequence elements beyond the canonical p53RE, demonstrating that the LincRNA-p21 locus is not merely a distal p53RE for p21.
Figure 18. A conserved sequence element within LincRNA-p21 Exon 1 reinforces p21 expression. (A). Sequence alignment of the mouse conserved element to homologous regions in rat, human, chimpanzee, pangolin, dog, and pig. Nucleotides with 100% conservation across these species are highlighted in yellow. The position of the LincRNA-p21 p53RE is indicated. (B) RT-qPCR analysis for p21 RNA in wildtype and LincRNA-p21Δ127/Δ127 MEFs cultured for 8 hours in the absence or presence of 0.5 µM doxorubicin. Data represent the mean ± SEM (n=18 biological replicates); *p<0.05, ****p<0.0001, paired t-test. For all samples, p21 levels were calculated relative to Gapdh and
normalized to untreated wildtype cells. (F) Representative immunoblot of p21 and p53 protein levels in cells described in (E). HSP90 was used as a loading control. (G) Densitometric analysis of p21 protein expression in cells described in (E and F). p21 levels were normalized to HSP90. Data represent the mean ± SEM (n=6 biological replicates); **p<0.01, ***p<0.001, paired t-test. (H) Proliferation of littermate wildtype and LincRNA-p21Δ127/Δ127 MEFs. Data are representative of n=3 independent growth curve experiments.

**Stable looping interactions juxtapose the conserved region with the promoter of p21**

We next desired to understand how the conserved sequence element reinforces p21 expression. Many lncRNAs have been proposed to promote chromatin remodeling or stabilize long-range looping interactions between promoters and regulatory elements (Canzio et al., 2019; Isoda et al., 2017; Melé and Rinn, 2016; Quinn and Chang, 2016). Additionally, previous work has shown that many distal p53REs harbor lncRNAs that are transcribed in a p53-dependent manner and participate in looping interactions that juxtapose the p53-bound response element with the promoter of a protein-coding gene (Melo et al., 2013). Since the LincRNA-p21 and p21 loci have been shown to physically associate (Groff et al., 2016; Sahlén et al., 2015), we hypothesized that the conserved region might establish or maintain the local chromatin architecture, thereby facilitating communication between the two loci. To test this, we used chromosome conformation capture (3C) to investigate the 3D organization of the LincRNA-p21/p21 locus in our LincRNA-p21Δ127/Δ127 and wildtype control MEFs in both the presence and absence of doxorubicin. Notably, although the promoter regions of LincRNA-p21 and p21 associate, providing a mechanism for long-range communication between the two loci, neither LincRNA-p21 transcription nor deletion of the conserved region leads to a reorganization of the 3D architecture of the locus (Fig. 19).

In further support of this conclusion, we recapitulated these results in lung adenocarcinoma cells with intact (Con) or mutagenized (ΔRE) p53REs (Fig. 8). In this system, we observed looping interactions between the promoters of LincRNA-p21 and p21 (Fig. 19). These looping interactions were not sensitive to the presence of a functional p53 pathway, since tamoxifen treatment did not alter contact strength in either cell population. Our results also show that the LincRNA-p21/p21 locus of lung
Figure 19. Stable looping interactions juxtapose the conserved region with the \textit{p21} promoter. (A and B) Taqman-3C analysis of the \textit{LincRNA-p21}/\textit{p21} locus. Chromatin was digested using the restriction enzyme BglII and 3C libraries were evaluated using unidirectional primers spanning the locus. The signal intensity for long-range interactions is shown relative to the signal of a control amplicon within the \textit{LincRNA-p21} locus that does not span a restriction site. (A) MEFs of the indicated cell lines were cultured for 24 hours in the absence or presence of 0.5 µM doxorubicin. Data represent the mean ± SEM (n=3 biological replicates); n.s., not significant, paired t-test. (B) Lung
adenocarcinoma cells of the indicated genotypes were cultured for 24 hours in the absence or presence of 0.5 µM 4-hydroxytamoxifen. Data represent the mean ± SEM (n=3 technical replicates). (C) Schematic showing the location of BglII sites, the LincRNA-p21 anchor, and the Taqman probe within the LincRNA-p21/p21 locus.

adenocarcinoma cells exhibits normal chromatin architecture, demonstrating that aberrant looping cannot explain the absence of a clear phenotype in ΔRE cells.

Our data therefore implicate the existence of robust chromatin interactions between the promoters of LincRNA-p21 and p21. In accordance with reports that topological domains are largely consistent across different cell types (Dixon et al., 2012), the presence of oncogenic signaling in the lung adenocarcinoma cells did not appear to affect the architecture of the locus. Moreover, looping did not appear to require LincRNA-p21 transcription, the region absent from the Δ127 allele, the p53RE, or p53 signaling. Although these results do not reveal the role of the conserved sequence element, they provide a clear mechanistic basis for the exchange of regulatory information between the LincRNA-p21 and p21 promoters.

Investigating the requirement for transcription through exon 1 of LincRNA-p21

Although full-length LinRNA-p21 transcription is dispensable for cis-regulation, we wondered whether low levels of transcription through exon 1, particularly the conserved region, could still play a role in reinforcing p21 expression. To address this, we used several molecular approaches to down- and upregulate LincRNA-p21 expression. We first turned to CRISPRi, which we hoped would enforce epigenetic silencing of the LinRNA-p21 locus and incur a decrease in p21 expression (Gilbert et al., 2014). We implemented a doxycycline-inducible dCas9-KRAB system, which should only repress LinRNA-p21 in the presence of doxycycline. Upon doxycycline treatment, p53-restorable MEFs stably infected with the dCas9-KRAB expression system showed robust dCas9-Krab RNA induction, indicating
Figure 20. Epigenetic silencing of LincRNA-p21 does not affect p21 levels. (A) Schematic of LincRNA-p21 exon 1 showing the relative positions of RNA guides for CRISPRi. The p53RE is indicated by the red box. (B) RT-qPCR analysis of dCas9-KRAB expression in p53-restorable MEFs stably infected with a doxycycline-inducible dCas9-KRAB expression vector and treated with the indicated concentration of doxycycline. Data represent the mean of n=3 technical replicates. For all samples, dCas9-KRAB levels were calculated relative to Gapdh and normalized to levels in untreated control cells. (C) RT-qPCR analysis of LincRNA-p21 expression in p53-restorable MEFs expressing dCas9-KRAB and the indicated gRNA, cultured for 24 hours in the absence or presence of 0.5 µM 4-hydroxytamoxifen and cultured for an additional 24 hours in the absence or presence of 0.5 µM doxorubicin. Data represent the mean ± SD (n=3 technical replicates). For all samples, LincRNA-p21 levels were calculated relative to Gapdh and normalized to levels in untreated control cells. (D and E) RT-qPCR analysis of LincRNA-p21 and p21 expression. p53-restorable MEFs stably expressing dCas9-KRAB and the indicated gRNA were cultured for >1 week in the presence of 2.0 µg mL⁻¹ doxycycline, cultured for 24 hours in the absence or presence of 0.5 µM 4-hydroxytamoxifen, and cultured for an additional 24 hours in the absence or presence of 0.5 µM doxorubicin. Data represent the mean ± SEM (n=3 biological replicates), n.s., not significant, paired t-test. For all samples, LincRNA-p21 and p21 levels were calculated relative to Gapdh.
the viability of this approach (Fig. 20B). We stably infected these MEFs with constructs expressing one of nine RNA guides targeting a 400 nucleotide window centered on the *LincRNA-p21* TSS (Fig. 19A). Cells were cultured for over a week in order to allow KRAB sufficient time to enact epigenetic repression (Kearns et al., 2014).

We observed that guides I-6, I-10, I-12, I-13, and I-14 mediated the knockdown of *LincRNA-p21* in both p53-deficient and p53-proficient MEFs (Fig. 20C). However, we chose to focus I-6, I-10, and I-12, as I-13 and I-14 both occlude the p53RE. In subsequent experiments, these three guides reduced *LincRNA-p21* levels only modestly and did not appear to perturb *p21* expression (Fig. 20D-E). Our results are therefore inconclusive. The lack of an effect on *p21* expression may indicate that *LincRNA-p21* transcription is dispensable for cis-regulation; alternatively, these data may support a model in which a moderate amount of transcription from the *LincRNA-p21* locus is sufficient to reinforce *p21* activity.

We next wondered whether increased transcription from the *LincRNA-p21* locus might upregulate *p21* expression. To test this, we performed CRISPRa in clonal p53-restorable MEFs with mutagenized (ΔRE) or wildtype (Con) p53REs (Fig. 21A). Elena Martínez-Terroba (E. Martínez-Terroba) designed lentiviral constructs harboring the SAM system and one of five dRNA guides targeting a 200 nucleotide window upstream of the *LincRNA-p21* TSS (Fig. 21A). These constructs were stably infected into ΔRE and Con MEFs. In ΔRE cells, dRNA A4 upregulated *LincRNA-p21* expression by 400-600% in both the absence and presence of p53 signaling (Fig. 21B). This increase in *LincRNA-p21* transcription correlated with a significant increase in *p21* levels at low levels of stress: p53-deficient cells exhibited a 1500% upregulation in *p21* expression, while p53-restored, unstressed cells showed a 200% increase in *p21* levels (Fig. 21C). Unexpectedly, CRISPRa did not increase *LincRNA-p21* levels in Con MEFs, perhaps suggesting that *LincRNA-p21* transcription is constrained to a narrow dynamic range in these cells (Fig. 21D). Although several dRNAs mediate significant increases in *p21* levels in p53-deficient and p53-proficient, unstressed Con cells, their effect size is relatively modest compared to the transcriptional activation observed in ΔRE cells (Fig. 21E). This may suggest that the increased *p21* expression in the ΔRE cells is a specific consequence of *LincRNA-p21* transcriptional activation. If
Figure 21. *LincRNA-p21* activation promotes *p21* expression in ΔRE p53-restorable MEFs. (A) Schematic of *LincRNA-p21* exon 1 showing the relative positions of dRNA guides for CRISPRa as well as the sequence of each allele in clonal ΔRE cells. The p53RE is indicated by the red box. (B and C) RT-qPCR analysis of (B) *LincRNA-p21* and (C) *p21* expression in clonal ΔRE cells stably infected with Cas9, the indicated chimeric dRNA guide harboring two MS2 stem-loops, and the HSF-p64-MS2 binding protein fusion. Data represent the mean ± SEM (n=2 biological replicates), n.s., not significant, *p*<0.05, **p**<0.01, paired t-test. For all samples, *LincRNA-p21* and *p21* levels were calculated relative to *Gapdh* and normalized to untreated dCon cells. (D and E) RT-qPCR analysis of (D) *LincRNA-p21* and (E) *p21* expression in clonal Con cells stably infected with Cas9, the indicated chimeric dRNA guide harboring two MS2 stem-loops, and the HSF-p64-MS2 binding protein fusion. Data represent the mean ± SEM (n=2 biological replicates), n.s., not significant, *p*<0.05, **p**<0.01, paired t-test. For all samples, *LincRNA-p21* and *p21* levels were calculated relative to *Gapdh* and normalized to untreated dCon cells.

E. Martínez-Terroba designed and cloned the constructs described in (A)
correct, this interpretation might indicate that $p21$ is insensitive to $\text{LincRNA-p21}$ levels after they exceed a critical threshold.

We also sought an alternative approach for addressing the role of transcription through the conserved region of exon 1 in the context of primary MEFs. To accomplish this, we used dRNA guides to target Cas9 to exon 1 of $\text{LincRNA-p21}$ and occlude transcription without engaging its endonuclease activity (Fig. 22A) (Dahlman et al., 2015; Qi et al., 2013). Although a dRNA targeting a region immediately downstream of the TSS (d2) did not significantly reduce the expression of spliced $\text{LincRNA-p21}$, likely due to the presence of a second TSS downstream of the p53RE, it did mediate an ~80% loss of transcription upstream of the p53RE (Fig. 22B). Intriguingly, both this dRNA and a positive control dRNA targeting the p53RE (d1) suppressed $p21$ RNA levels by 21% and 35%, respectively (Fig. 22B). These results, in conjunction with our data showing that Twister permits ~40% of wildtype transcription through exon 1 (Fig. 22C), raised the exciting possibility that a small amount of active transcription may be essential for cis-regulation, hinting at a dual requirement for both transcription and the conserved sequence element.
Figure 22. LincRNA-p21 transcriptional inhibition correlates with a reduction in p21 levels. (A) Schematic showing the position of each dRNA guide and amplicon relative to LincRNA-p21 Exon 1. (B) RT-qPCR analysis of spliced and total LincRNA-p21 RNA levels in MEFs co-expressing Cas9 and the indicated dRNA guide. Data represent the mean ± SEM (Spliced LincRNA-p21 n=4; E1_1 n=3; E1_2 n=3; biological replicates); *p<0.05, **p<0.01, ***p<0.001. For each primer pair, LincRNA-p21 levels were calculated relative to Gapdh and normalized to cells expressing dCon. (C) RT-qPCR analysis of p21 RNA expression in the MEFs described in (B-D). Data represent the mean ± SEM (n=4 biological replicates); *p<0.05, ***p<0.001, paired t-test. p21 levels were calculated relative to Gapdh and normalized to cells expressing dCon.
Data summary and conclusion

Although recent genetic studies have converged on the consensus that *LincRNA-p21* acts locally to reinforce *p21* expression (Dimitrova et al., 2014; Groff et al., 2016; Korkmaz et al., 2016), its mechanism of action has remained the subject of scrutiny. To elucidate the molecular logic underlying this *cis*-regulatory relationship, we performed an extensive genetic and molecular dissection of the *LincRNA-p21* locus. As a result of our efforts, which we report in this chapter, we exclude full-length transcription, transcript processing, and RNA accumulation as effecters of *cis*-regulation; instead, we provide evidence that transcription through a conserved region within exon 1 of *LincRNA-p21* supports *p21* expression. We show that it likely accomplishes this by exploiting stable chromatin interactions between the promoters of both loci, thereby providing a mechanistic basis for the exchange of information about *LincRNA-p21*’s regulatory state.

This finding is consistent with several known mechanisms of *cis*-regulatory lncRNA function. LncRNAs have been proposed to recruit transcription factors with dual DNA/RNA binding activity (Sigova et al., 2015), and the *LincRNA-p21* transcript has previously been shown to interact with the transcriptional activator hnRNP-K (Dimitrova et al., 2014; Huarte et al., 2010; Yoon et al., 2012). It is therefore possible that the conserved sequence element is transcribed into an RNA scaffold for protein factors that act, in turn, to reinforce *p21* expression. Alternatively, the process of transcriptional initiation may promote transcription factor binding by “licensing” enhancer elements within the conserved region (Scruggs et al., 2015). Indeed, transcription has been observed to promote epigenetic remodeling at enhancers under diverse biological contexts (Anderson et al., 2016; Kaikkonen et al., 2013), although the functional significance of this relationship remains unclear.

Our data support several key findings from previous genetic studies of both mouse and human *LincRNA-p21*. In particular, we confirm reports that enhancer-like DNA elements within exon 1 harbor *cis*-regulatory activity (Groff et al., 2016; Korkmaz et al., 2016) and show that the locus is a functional driver of *p21* expression (Dimitrova et al., 2014; Groff et al., 2016; Korkmaz et al., 2016; Tesfaye et al.). As an important addition to previous literature, which showed that the conserved region in exon 1 is
sufficient to drive the expression of an exogenous reporter construct, we demonstrated that this region is necessary for cis-regulation at the endogenous LincRNA-p21 locus (Groff et al., 2016). However, our conclusions also diverge from those of previous genetic studies on a key point. Notably, we observed a dual requirement for both the conserved region and the transcriptional process. This stands in contrast to the findings of one study, which reported that LincRNA-p21 operates primarily as a DNA element whose transcription is dispensable for its function (Groff et al., 2016). It is thus important to note that, while we used multiple loss-of-function strategies that minimally disrupted the LincRNA-p21 sequence, Groff et al. replaced the entire locus with a lacZ reporter (Groff et al., 2016; Sauvageau et al., 2013). Their data, which show a reduction in p21 levels across multiple different tissue types, including tissues in which mature LincRNA-p21 is absent, are therefore also consistent with a model in which transcription through sequence elements in exon 1—but not the mature transcript—contributes to cis-regulation. Our results also stand in contrast to work that used ASOs to identify a role for either the full-length transcriptional process or mature RNA. The same ASOs mediated a significant reduction in p21 levels in mutant MEFs, which overwhelmingly lack LincRNA-p21 RNA, perhaps implicating nonspecific effects.

As part of this work, we sought to develop complementary genetic strategies to disrupt biogenesis of the endogenous LincRNA-p21 RNA. Self-cleaving ribozymes represent an emerging strategy for interrogating cis-acting IncRNA function (Camblong et al., 2009; Tuck et al., 2018; Wery et al., 2018b). Although tools such as the Hammerhead and Hepatitis Delta ribozymes have been extensively used to study RNA processing in the context of short transgenic constructs (Bird et al., 2005; Eckner et al., 1991; West et al., 2004), the effects of co-transcriptional ribozyme cleavage on endogenous IncRNA transcription remain largely uncharacterized, making it difficult to evaluate the conclusions of existing work. In this study, we used TT-qPCR to reveal that Twister mediates the co-transcriptional knockdown of the nascent LineRNA-p21 RNA. This was surprising in light of reports that co-transcriptional ribozyme cleavage does not destabilize transcription (Fong et al., 2009; West et al., 2004) and that downstream ribozyme cleavage products are poor substrates for XRN2-mediated transcriptional termination (Doamekpor et al., 2020; Jinek et al., 2011; Roth et al., 2014). Our data highlight the need to investigate
the effects of these promising genetic tools on the transcriptional and epigenetic landscape of endogenous IncRNAs and develop robust “design rules” for their implementation.

Ultimately, our work helps to illuminate the biological significance of the relationship between LinRNA-p21 and p21. LncRNAs have been proposed to specify cell fate by “fine tuning” the expression of key protein-coding genes (Cabili et al., 2011; Gil and Ulitsky, 2020; Guttman et al., 2009). In keeping with this paradigm, an elegant study recently showed that early p21 expression dynamics specify cell fate in response to genotoxic stress (Hsu et al., 2019). Given that LinRNA-p21 is detectably induced as early as one hour after p53 stabilization in human colon cancer cells (Allen et al., 2014), this lncRNA may represent a mechanism for modulating p21 levels during the initial p53 transcriptional response. This would explain our intriguing observation that deletion of the conserved sequence element and dRNA-mediated transcriptional inhibition decreased p21 expression in primary MEFs cultured in the absence of doxorubicin treatment. Since ambient oxygen levels have been shown to induce oxidative stress (Jagannathan et al., 2016), which in turn can activate the p53 pathway (Kastenhuber and Lowe, 2017), we propose that LinRNA-p21 may help activate p21 levels past a critical threshold in response to low levels of p53 signaling (Dimitrova et al., 2014).

In summary, our data demonstrate that full-length lncRNA transcription, processing, and accumulation are dispensable for cis-regulation at the LinRNA-p21 locus. By systematically dissecting the overlapping layers of LinRNA-p21 biogenesis, we reveal a dual requirement for transcriptional initiation and sequence elements embedded in exon 1. Our results showcase the power of complementary genetic and molecular tools to illuminate new facets of cis-regulation and, in doing so, enhance our understanding of the noncoding transcription that pervades our genome.
Chapter V: Summary and perspectives

The regulatory significance of pervasive transcription

Our work adds to the growing body of evidence that the noncoding transcriptome is often a functional driver—and not a byproduct—of diverse regulatory programs. Unexpectedly, we found that the process of active transcription through exon 1 of LincRNA-p21 is necessary to reinforce p21 expression. By contrast, full-length transcription, splicing, and RNA accumulation are dispensable. Moreover, we demonstrated that a conserved region, which itself enhances LincRNA-p21 transcription, also contributes to cis-regulation. Notably, this region harbors predicted binding sites for a number of transcription factors, including CEBPα/β, EGR4, and Sp1. In support of the functional relevance of these predicted sites, both empirical and computational data indicate that the corresponding transcription factors bind homologous sequences within the human locus. Interestingly, Sp1 has previously been shown to interact with p53 (Koutsodontis et al., 2001), raising the possibility that transcription factors bound within the Δ127 region might cooperate with DNA-bound p53 in order to activate LincRNA-p21 and p21 expression. Nonetheless, it remains unclear how the transcriptional process and conserved region might cooperate to enact their cis-regulatory function. Here, we envision several mechanistic possibilities with distinct consequences for our understanding of cis-regulation at this locus and throughout the genome.

Model 1: transcription per se

As a first possibility, the act transcription per se might be sufficient to drive p21 expression. Under this model, the p53RE and adjacent conserved region would function as an enhancer-like element for LincRNA-p21 but not p21. The LincRNA-p21 transcriptional process would act, in turn, to reinforce p21 levels. Since the identity of the transcribed sequence is irrelevant in this scenario, CRISPRa in ΔRE or Δ127 cells would likely upregulate p21 expression, demonstrating that the transcriptional process is sufficient for cis-regulation in the absence of regulatory DNA elements.
This mechanistic paradigm has been described for other lncRNAs, including both Blustr and Maenli (Allou et al., 2021; Engreitz et al., 2016). Transcription is essential to the cis-regulatory activity of both loci, as evidenced by the functional consequences of premature transcriptional termination. However, the underlying sequences appear to be largely dispensable, since intron and exon deletions do not dysregulate neighboring genes. Blustr and Maenli thus serve as an example for how local transcription can enact cis-regulation in a sequence-independent manner.

Model 2: enhancer licensing through transcription

As a second model, LincRNA-p21 transcription may license the conserved region to act as a distal p21 enhancer. Under this paradigm, the conserved region would serve as a cis-regulatory DNA element for both LincRNA-p21 and p21. The transcriptional process would thus confer an additional layer of regulatory specificity by modulating the local epigenetic state in order to promote transcription factor binding.

LincRNA-p21 transcription could promote transcription factor binding through one of several mechanisms. Firstly, it could mediate nucleosome repositioning or disassembly, leading to the formation of a nucleosome-free region around the LincRNA-p21 promoter. Indeed, lncRNA transcription has been shown to reduce the local nucleosome density at transcription factor binding sites, permitting binding by cognate transcription factors (Hirota et al., 2008; Senmatsu et al., 2019). Although this phenomenon has largely been documented in yeast, a few studies suggest that transcription may also contribute to the formation of nucleosome-free regions at mammalian promoters (Scruggs et al., 2015; Seila et al., 2009). Since many transcription factors are unable to bind within regions of high nucleosome occupancy (Zhu et al., 2018), active transcription could thus expose previously-occluded binding sites within the conserved region.

As an alternative, transcription could create a more permissive epigenetic environment around the LincRNA-p21 promoter. This, in turn, could promote transcription factor binding within the region. LncRNA transcription has previously been shown to modulate the local epigenetic state. For example,
Upperhand transcription establishes and maintains high levels of H3K27ac across its associated super-enhancer (Anderson et al., 2016), although it remains unclear whether H3K27ac levels have functional consequences for enhancer activity (Zhang et al., 2020). Interestingly, regions of stable H3K4 methylation, which are associated with transcriptionally active promoters, have been shown to be more accessible to transcription factors (Huang et al., 2019). Taken together, these findings demonstrate the viability of a model in which transcription creates a more permissive epigenetic environment for transcription factor binding at the LincRNA-p21 locus.

Model 3: functional sequence elements in the nascent RNA

If correct, the preceding models would point to the conclusion that the LincRNA-p21 transcript is dispensable for cis-regulation and instead identify a sequence-independent role for transcription. By contrast, a third model, in which the 5’ end of the nascent RNA molecule acts as a recruitment platform for activating protein factors, implicates LincRNA-p21 sequence and structure as a key mediator of cis-regulation. LincRNA-p21 has previously been shown to bind hnRNP-K via a structurally-conserved ~200 nucleotide sequence at its 5’ end (Huang et al., 2010). Intriguingly, the conserved region described in our work harbors a predicted binding site for hnRNP-K, suggesting that it could serve as a binding platform for this and other proteins. This mode of action has been documented for other lncRNAs, including Khps1, which recruits p300/CBP to its locus (Postepska-Igielska et al., 2015). As discussed previously (see Chapter I), the Khps1 RNA interacts with the DNA duplex to form a triple helix, demonstrating the requirement for a tether to retain the Khps1/p300/CBP complex on the chromatin (Postepska-Igielska et al., 2015). Under our final model of LincRNA-p21 function, the nascent 5’ end would remain tethered to the chromatin by elongating Pol II. Active transcription through exon 1 would thus maintain a pool of 5’ ends capable of recruiting protein factors to the lncRNA promoter.
The majority of LincRNA-p21 transcription is nonfunctional

One important outcome of our work is the observation that the vast majority of LincRNA-p21 transcription is nonfunctional. Although the LincRNA-p21 locus spans over 20 kilobases, only transcription through exon 1 contributes to cis-regulation. Our results therefore suggest that a single transcriptional initiation event may mediate both functional and nonfunctional transcription. In particular, if transcription through a regulatory element such as the conserved region within exon 1 contributes to cis-regulation, then downstream transcription and RNA processing may be dispensable so long as their absence does not impair the kinetics of upstream processes. Notably, this has been shown to hold true for other cis-acting IncRNAs, including 118 kilobase IncRNA Airn. Although transcription through the ~30 kilobase region overlapping the Igf2r locus is required for gene silencing, premature transcriptional termination outside of this window has no functional consequences (Latos et al., 2012). Thus, the majority of the Airn transcriptional unit is nonfunctional, since it does not contribute to Igf2r imprinting. Similarly, full-length transcription of Blustr has been shown to be inessential for Sfmbt2 expression. Although Sfmbt2 levels exhibit some dependence on the length of the Blustr transcriptional unit, transcription through the final ~ten kilobases of this 26 kilobase locus does not appear to contribute to its cis-regulatory function (Engreitz et al., 2016). Our work thus adds to the mounting body of evidence that many functionally transcribed IncRNAs do not act through full-length transcription.

One question that arises from our work is why IncRNAs such as LincRNA-p21, Airn, and Blustr have such large genomic footprints if their full-length transcription is dispensable for cis-regulation. The answer to this question may be rooted in the evolutionary dynamics of eukaryotic genomes. In humans, and likely in other mammals, the majority of the genome evolves nearly neutrally, meaning that it is not subject to strong selective pressure. This weak selective regime may permit the emergence of cryptic TSSs that give rise to nonfunctional “junk” transcripts (Palazzo and Koonin, 2020). Importantly, transcription and RNA degradation impose a negligible energetic burden on eukaryotic cells (Lynch and Marinov, 2015), suggesting that the majority of “junk” IncRNAs are unlikely to incur a significant fitness cost. Over time, neutral mutations may allow nonfunctional IncRNA loci to explore the sequence space
and acquire novel functions within pre-existing cellular pathways (Palazzo and Koonin, 2020). This functionality may not require the lncRNA transcript or even full-length transcription; in the absence of selective pressure against unnecessarily long transcriptional units, and without neutral evolution that results in a shortened transcriptional unit, the transcript will likely retain its initial length. Under this paradigm, the transcription of a multi-kilobase, multi-exonic locus may thus reflect the evolutionary process that generated a functional transcriptional event from an ancestral “junk” transcript.

Our finding that transcription through the majority of the LincRNA-p21 locus is inessential has broad implications for our ability to successfully identify other functionally transcribed lncRNAs. In particular, our findings show that parallel genetic and molecular approaches may be required to establish lncRNA functionality. Although both the PAS and Twister abrogated LincRNA-p21 expression in primary MEFs, neither mediated a reduction in p21 levels. In other experimental contexts, this might have been interpreted as evidence that the transcriptional process does not contribute to cis-regulation, supporting a model in which the conserved region acts as an enhancer-like DNA element. However, transcriptional roadblock experiments indicated that transcription through exon 1, which Twister (and likely the PAS) permits, harbors cis-regulatory activity. Our results demonstrate the need for caution when interpreting negative results from lncRNA knockdown experiments and underscore the importance of careful tool design and placement.

The importance of an expanded toolkit for studying noncoding RNA function

Cis-acting lncRNAs have a unique capacity to enact gene regulation through multiple features of their locus, transcription, and RNA. This mechanistic diversity poses considerable challenges for identifying the functional unit of cis-regulation at uncharacterized lncRNA loci. In particular, the lncRNA field requires tools that can deconvolve the closely-interlinked effects of transcription, co-transcriptional processes such as splicing, and the mature RNA transcript. As Orgel’s rule famously states, “evolution is cleverer than you are.” In keeping with this, the kaleidoscopic repertoire of naturally-occurring genetic elements provides ample inspiration for new tools. Many of these elements have the potential to act in cis.
to specifically regulate the processing, stability, and even sequence of host RNAs. This makes them attractive candidates for development into a new generation of tools.

The current tools for modulating transcription (discussed in Chapter I) are both limited and limiting. When implemented successfully, PAS sequences entirely abolish transcription from a locus. While this approach is invaluable for establishing a requirement for lncRNA transcription, it is less useful for addressing how transcription mediates gene regulation. In order to examine the temporal relationship between lncRNA transcription and changes in the local chromatin environment, studies have often turned to molecular approaches such as the reversible transcriptional inhibitor DRB (5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole). This permits the investigation of the effects of abolishing and restoring transcription at a locus of interest. This approach has the potential to reveal intriguing correlations between molecular events at two loci. However, because it disrupts global Pol II transcription, it cannot specifically test the effects a lncRNA on the local epigenetic state. Novel genetic tools may help circumvent the caveats of current approaches.

RNA-based transcriptional attenuation is one solution for specifically and reversibly inhibiting lncRNA transcription. This approach is predicated on the ability of structured RNA sequences to constitutively attenuate the transcription of host genes. Short RNA sequences, termed small transcription-activating RNAs (STARS), may bind to a complementary sequence within the attenuator, disrupting its structure and permitting transcription through the locus (Chappell et al., 2015). Although this approach was pioneered in bacterial cells, several lines of evidence hint at its applicability as a tool for studying eukaryotic gene regulation. In particular, early work on the SV40 virus revealed the presence of two hairpin elements that competently block transcriptional elongation, demonstrating the ability of structured RNA elements to block Pol II transcription in mammalian cells (Resnekov et al., 1989). Subsequent work showed that short DNA sequences—conceptually akin to STARS—could release this transcriptional block (Kessler and Aloni, 1989). Taken together, these studies provide preliminary support for the viability of a specific and reversible RNA-based transcriptional attenuator.
Other genetic elements present attractive opportunities for modulating lncRNA processing, particularly splicing. Riboswitches can reversibly inhibit splicing by sequestering the 5’ splice site, 3’ splice site, or branch point. As a proof of principle, insertion of the theophylline riboswitch into either the branch point or 3’ splice site of a reporter gene has been shown to suppress splicing in the presence of theophylline (Kim et al., 2005, 2008). It is important to note that such an approach may not prevent splicing entirely and may instead induce exon skipping (Vogel et al., 2018). However, this may prove useful for testing the functionality of the skipped sequence.

An additional suite of tools may alter RNA stability either co- or post-transcriptionally. Here, self-cleaving ribozymes again demonstrate their versatility, this time as tools for mediating context-specific transcript cleavage. As a first example, small molecule-inducible self-cleaving ribozymes may be used to regulate RNA stability. In this approach, aptamers grafted to ribozymes permit cleavage in the absence of a small molecule such as theophylline or xanthine (Xiang et al., 2019). Following ligand binding, however, the aptamer inhibits the ribozyme’s self-cleaving ability. This, in turn, permits the accumulation of the host transcript. Such an approach may prove particularly useful for elucidating the role of lncRNAs in the initiation and long-term maintenance of subcellular structures such as paraspeckles. Split ribozymes may also enact context-specific transcript cleavage. For example, a self-cleaving ribozyme spanning an intron may self-cleave following a splicing event that removes the intervening intron. Such an approach would facilitate spliceform-specific transcript degradation and could serve as a viable approach for specifically destabilizing processed transcripts. As an alternative to self-cleaving ribozymes, UTRs may be used to destabilize transcripts. For example, the c-fos 3’ UTR has previously been used to destabilize the Kcnq1ot1 RNA (Pandey et al., 2008), demonstrating its fundamental feasibility as an approach for regulating lncRNA functionality.

While the genetic elements discussed above provide several different strategies for destabilizing host transcripts, other elements may enhance RNA stability. For example, elements for nuclear expression (ENEs) protect host transcripts from degradation by 3’-5’ exonucleases by sequestering the poly(A) tail (Conrad et al., 2006). This increased stability may result in functional lncRNA upregulation. As an
important note, this approach is specific to RNAs with poly(A) tails. Since many *cis*-acting IncRNAs remain tethered to the chromatin by Pol II (Werner and Ruthenburg, 2015), its applicability may be limited. Lariat-capping ribozymes are another attractive option for both cleaving and selectively stabilizing host transcripts. The *Didymium* lariat-capping ribozyme self-cleaves to yield a downstream fragment featuring a 2′,5′-phosphodiester bond between the first and third nucleotides (Krogh et al., 2017). In yeast, this ribozyme efficiently cleaves host transcripts and stabilizes the downstream RNA product, suggesting that it may present a new strategy for specifically interrogating RNA sequence requirements while minimally disrupting the DNA sequence. RNA-based genetic tools therefore provide many exciting avenues for modulating IncRNA expression *in vitro* and *in vivo*.

Even as the IncRNA field develops promising new approaches, it must remain committed to understanding the mechanisms and caveats of existing tools. Our work reveals two surprising findings that, taken together, demonstrate the need to rigorously evaluate current molecular and genetic approaches. Firstly, we observed that ASOs targeting the *LincRNA-p21* transcript attenuated *p21* transcription in *LincRNA-p21*\textsuperscript{PAS/PAS} and *LincRNA-p21*\textsuperscript{TWI/TWI} MEFs despite the overwhelming absence of spliced RNA in both cell lines. Since Twister cleavage permits a small degree of transcription into exon 2, it is possible that ASO-mediated degradation of the remaining transcript further suppresses transcription from the promoter, thus preventing *LincRNA-p21* from enacting *cis*-regulation. However, it is unclear whether this explanation could also apply to *LincRNA-p21*\textsuperscript{PAS/PAS} MEFs. Given that transcription may continue for more than ten kilobases after a PAS (Core et al., 2008), it is formally possible that the remaining transcription, which ASO-I and ASO-E2 deplete, is sufficient to reinforce *p21* expression. However, our results may also indicate nonspecific or off-target effects of ASOs, demonstrating the need to better understand this technology.

As a second outcome, we found that Twister mediates a ~60 decrease in nascent RNA expression at the *LincRNA-p21* TSS (see Chapter IV). This observation is puzzling in light of reports that ribozyme cleavage likely does not incur transcriptional knockdown by XRN2, the primary effector of co-transcriptional termination (West et al., 2004). Since self-cleaving ribozymes have garnered increasing
interest as tools for elucidating previously-unappreciated aspects of lncRNA biology, (Camblong et al., 2009; Tuck and Bühler, 2021; Tuck et al., 2018; Wery et al., 2018b), it is imperative to understand their effects on the transcription and metabolism of host RNAs.

There are several key possibilities, each with disparate implications for this tool family. Firstly, although exoribonucleases such as XRN2 exhibit a strong preference for 5’-monophorylated substrates (Doamekpor et al., 2020; Jinek et al., 2011; Mathy et al., 2007; Pellegrini et al., 2008), it is possible that XRN2 eventually, if inefficiently, metabolizes substrates with 5’-hydroxyl groups, generating a 5’-phosphorylated intermediate that is subsequently amenable to rapid degradation.

As a second possibility, other enzymes may modify the 5’ ends of downstream cleavage products to promote their reentry into the XRN2 degradation pathway. In yeast, for example, RNA 5’-kinases have been proposed to promote the degradation of 5’-hydroxylated RNAs by the XRN2 homolog XRN1 (Braglia et al., 2010; Peach et al., 2015). Provocatively, many yeast 5’-hydroxylated RNA termini arise from genomic regions with the consensus sequence 5’-CAUU|A-3’, a perfect match for the Twister cleavage site (Peach et al., 2015). Since mammalian cells harbor many RNA 5’-kinases (Shuman and Hurwitz, 1979; Weitzer and Martinez, 2007), 5’ end phosphorylation may represent a general mechanism for permitting the reentry of non-phosphorylated substrates into the major 5’ end surveillance pathway. Alternatively, distributive exonucleases such as decapping and exoribonuclease (DXO) family proteins may hydrolyze the 5’-terminal nucleotide, generating a 5’-phosphorylated product that is subsequently degraded by XRN2 (Doamekpor et al., 2020). Similarly, endonuclease cleavage within the host transcript may generate additional RNA fragments with 3’-hydroxyls and 5’-monophosphate groups; these may serve as substrates for the nuclear exosome and XRN2, respectively.

As a final possibility, another nuclear 5’-3’ exoribonuclease may degrade downstream cleavage products. Interestingly, the yeast exonuclease Rrp17, which acts processively to trim the 5’-hydroxylated ends of pre-rRNAs, has a mammalian homolog (NoI12) that likely also mediates 5’-3’ RNA degradation (Oeffinger et al., 2009; Sloan et al., 2013). This suggests that non-XRN2 exonucleases could potentially recognize and degrade the LincRNA-p21 transcript following Twister cleavage.
The mechanism of Twister-mediated RNA degradation has considerable implications for our understanding of the “design rules” governing the placement of Twister and similar self-cleaving ribozymes within host transcripts. If, as discussed above, XRN2 inefficiently metabolizes the first nucleotide of substrates with 5’-hydroxyl groups, or if other enzymes modify the 5’ end of the cleavage product such that it becomes a suitable substrate for XRN2, then this initial event may serve as the rate-limiting step for RNA degradation. Since kinetic competition between XRN2 and Pol II defines the region of transcriptional termination (Fong et al., 2015), then the extent of transcription will depend in part on the distance traversed by Pol II during the rate limiting step. Alternatively, if other 5’-3’ exonucleases progressively degrade cleaved transcripts, then their kinetics will determine how ribozyme cleavage affects transcription. A thorough understanding of how Twister and other self-cleaving ribozymes effect targeted transcript degradation may therefore prove critical for selecting optimal ribozyme insertion sites and evaluating experimental results.

Although cis-regulatory IncRNAs pose many unique challenges, the IncRNA field appears well-prepared to surmount them. The steady development and implementation of novel tools, including self-cleaving ribozymes such as Twister, has given rise to an expansive array of approaches for addressing increasingly mechanistic questions about IncRNA functionality. Continuing innovation, including the adoption of new tools such as those described above, will further expand the suite of techniques available for experimental application. We predict that this innovation, coupled with rigorous efforts to understand the molecular effects of existing tools, will enable the IncRNA field to further mature.

Final remarks

Mammalian genomes harbor numerous transcribed regulatory loci that play an increasingly recognized role as mediators of local gene expression. However, the molecular mechanisms through which these loci enact cis-regulation have remained largely unclear. Using a comprehensive genetic and molecular approach, we showed that active transcription through exon 1 of LincRNA-p21 is sufficient to
drive expression of the nearby *p21* gene. Two major lessons arise from our work. Firstly, we echo the observation that *cis*-regulatory lncRNA loci can harbor multiple functional elements. In light of this, experimental efforts to dissect a lncRNA’s mechanism of action must make use of multiple complementary approaches. Secondly, our work adds to the growing body of literature recognizing transcription or the nascent RNA as an important mechanism of *cis*-activation (Allou et al., 2021; Anderson et al., 2016; Canzio et al., 2019; Engreitz et al., 2016; Isoda et al., 2017). Like *LincRNA-p21*, many functionally transcribed lncRNAs have large genomic footprints; for many of these lncRNAs, transcription through this vast genomic distance is central to their *cis*-regulatory activity (Allou et al., 2021; Anderson et al., 2016; Engreitz et al., 2014; Isoda et al., 2017). By contrast, our data demonstrate that full-length *LincRNA-p21* transcription is dispensable, pointing to a regulatory role for TSS-proximal transcription at this locus. In narrowing down the functional element of *cis*-regulation at the *LincRNA-p21* locus, we expand the range of mechanistic possibilities for functional *cis*-regulatory loci and demonstrate the genomic significance of pervasive noncoding transcription.
Chapter VI: Materials and methods

Animal handling and genotyping

*LincRNA-p21* knockdown mice were generated using CRISPR/Cas9-mediated engineering in C57BL/6J mice at the Jackson Laboratory for Genomic Medicine. Briefly, blastocysts were electroporated with Cas9, a guide RNA targeting *LincRNA-p21* and one of the two HDR templates described in Supplementary Table 1. Mice carrying each allele of interest were identified by PCR-based genotyping using primers described in Supplementary Table 1 followed by Sanger sequencing and were subsequently back-crossed once with wildtype C57BL/6J mice. For whole-tissue RNA expression analysis, 2-5 month-old female mice were used. All studies and procedures were conducted with the approval of the Yale University Institutional Animal Care and Use Committee.

Cell lines and treatments

Primary MEFs were isolated from E13.5 embryos resulting from timed heterozygote matings. MEFs were genotyped using PCR primers described in Supplementary Table 1. All experiments were performed between passages 2 and 10. p53-restorable (*p53*^LSL/LSL; *Rosa26-CreER^T2) MEFs and puromycin-resistant lung adenocarcinoma cells from *K-ras^LA2-G12D/+; p53^LSL/LSL; Rosa26-CreER^T2* mice have previously been described (Feldser et al., 2010; Olivero et al., 2020; Ventura et al., 2007).

Primary MEFs were maintained in DMEM (Gibco) supplemented with 15% fetal bovine serum (FBS), 50 U mL^−1^ penicillin-streptomycin, 2 mM L-glutamine, 0.1 mM non-essential amino acids, and 0.055 mM β-mercaptoethanol, while p53-restorable MEFs and lung adenocarcinoma cells were maintained in DMEM supplemented with 10% FBS, 50 U mL^−1^ penicillin-streptomycin, 2 mM L-glutamine, and 0.1 mM non-essential amino acids. All cells were maintained at 37°C in a humidified incubator with 5% CO₂.

In order to induce DNA damage, all cells were treated with 0.5 μM doxorubicin (Sigma-Aldrich) or 25 μM etoposide (Sigma-Aldrich). To delete the loxP-STOP-loxP cassette preventing p53 expression in p53-restorable MEFs and lung adenocarcinoma cells, cells were treated with 0.5 μM 4-
hydroxytamoxifen (Cayman Chemical Company) for 24 hours prior to treatment with doxorubicin. Tetracycline-dependent gene expression (pHAGE-TRE-dCas9-KRAB; Addgene #50917) was induced by treating cells with 2.0 µg mL⁻¹ for the time specified in the figure legend.

**Constructs**

The viral constructs, gRNAs/dRNAs, and shRNA sequences used in this work can be found in Supplementary Tables 1 and 2.

**P53RE mutagenesis and dRNA-mediated transcriptional inhibition**

CRISPR/Cas9-mediated mutagenesis and transcriptional inhibition were accomplished using guide sequences targeting the LincRNA-p21 locus. Guides were cloned downstream of the hU6 promoter of a lentiviral construct that also expresses spCas9 and a puromycin resistance gene (Brd001) or GFP selection marker (Brd004). Both lentiviral vectors were gifts from the Broad Institute (MIT).

**CRISPRi**

To perform CRISPRi, p53-restorable MEFs were serially infected with constructs expressing dCas9-KRAB and gRNAs targeting the LincRNA-p21 TSS. Cells were first infected with the lentiviral construct pHAGE-TRE-dCas9-KRAB (Addgene #50917), a doxycycline-inducible expression system harboring dCas9 from *S. pyogenes* fused to the KRAB repressor. Successfully infected cells, which exhibit neomycin resistance, were subsequently infected with a lentiviral vector expressing gRNAs under the control of the hU6 promoter and harboring a zeocin selection marker (Addgene #61427).

**CRISPRa**

CRISPRa was accomplished by cloning 14-15 nucleotide “dead” guides downstream of the hU6 promoter of a lentiviral vector co-expressing HSF-p64-MS2 binding protein and a hygromycin resistance gene (generated by E. Martínez-Terroba)(Olivero et al., 2020). The chimeric gRNA backbone includes
MS2 stem-loops that facilitate interactions between the guide and the HSF-p64-MS2 binding protein module. This system was introduced into stably-infected p53-restorable ΔRE-spCas9-GFP or Con-spCas9-GFP cells.

**Sh-CEBPB**

To knock down CEBPB expression, primary MEFs were stably infected with TRC2-pLKO lentiviral constructs co-expressing sh-CEBPB and a puromycin resistance gene (TRCN0000231407 and TRCN0000231411; MilliporeSigma).

**Viral transduction and infection**

To generate lentivirus, HEK293 cells were co-transfected with pCMV-Δ8.2 (Addgene #8455) and pCMV-VSV-G (Addgene #8454), and the specified lentiviral construct. Viral media was harvested at 48, 72, and 96 hours post-transfection, supplemented with 4 µg mL\(^{-1}\) polybrene, and applied directly to cells. For retroviral production, Phoenix cells were transfected with retroviral constructs. Retroviral media was harvested at four consecutive 12-hour intervals, supplemented with 4 µg mL\(^{-1}\) polybrene, and used directly to infect cells. Following infection, cells were selected with puromycin (2 µg mL\(^{-1}\) for MEFs and 5 µg mL\(^{-1}\) for lung adenocarcinoma cells), neomycin (1 mg mL\(^{-1}\)), zeocin (500 µg mL\(^{-1}\)), or hygromycin (400 µg mL\(^{-1}\)). Cells stably expressing a GFP marker were selected using a FACS Aria III sorter with FACS DIVA software (BD Biosciences).

**Antisense oligonucleotide (ASO) knockdown**

ASO treatment was performed as previously described (Dimitrova et al., 2014), using the 5-10-5 MOE “gapmer” oligonucleotides (Ionis Pharmaceuticals) listed in Supplementary Table 1. Briefly, 2x10^6 primary MEFs were nucleofected twice using the Amaza Mouse/Rat Hepatocyte Nucleofector Kit (Lonza) and Nucleofector 2b Device (Lonza) with a total of 1 µM of targeting or control nontargeting ASO at 48 hour intervals. Samples were harvested 48 hours after the final nucleofection.
RNA isolation and RT-qPCR

Total RNA was isolated following the protocol for TRIzol (Thermo Fisher Scientific) or the RNeasy Mini Kit with or without DNAse I digestion (Qiagen). 0.5-1.0 µg RNA was reverse-transcribed using the High Capacity cDNA Reverse Transcriptase Kit (Applied Biosystems). For qPCR, SYBR Green mastermix was used in conjunction with the primers listed in Supplementary Table 1. Expression was calculated relative to GAPDH and normalized to the corresponding control sample.

Transient transcriptome (TT)-qPCR

Primary MEFs cultured in the absence or presence of doxorubicin were treated with 1 mM 4-thiouridine (Alfa Aesar) and cultured for 5 minutes in the dark. Samples were placed on ice, rinsed with PBS, scrape-harvested in PBS, suspended in TRIzol, and frozen at -80°C. Total RNA was isolated using chloroform extraction and isopropanol precipitation and treated with TURBO DNase to deplete genomic DNA. Purified RNA was then extracted using phenol-chloroform-isoamyl alcohol followed by ethanol precipitation. 50 µg of total RNA was biotinylated using MTSEA biotin-XX (Biotium) and enriched using streptavidin as previously described (Duffy et al., 2015). The resulting nascent RNA was reverse-transcribed using the SuperScript VILO cDNA Synthesis Kit (Thermo Fisher Scientific) and analyzed via RT-qPCR using iTaq Universal SYBR Green Supermix (Bio-Rad) in conjunction with the primers listed in Supplementary Table 1. Expression was calculated relative to GAPDH and normalized to the corresponding control sample.

Protein isolation and immunoblotting

Whole-cell protein samples were prepared by resuspending cell pellets in 2X Laemmli buffer (100mM Tris-HCl pH 6.8, 200mM DTT, 3% SDS, 20% glycerol) at a concentration of 10⁵ cells/µL. All samples were boiled at 95°C for 7 minutes and passed through an insulin syringe. Protein samples from 10⁵ cells were separated by electrophoresis on a 10% or 12% SDS-polyacrylamide gel and transferred to a
0.2 µm nitrocellulose membrane. Immunoblotting was performed using the following antibodies: p21 (clone F-5, sc-6246, Santa Cruz Biotechnology; 1:250), p53 (clone CM5, NCL-L-p53-CM5p, Leica; 1:1000), and loading control HSP90 (clone C45G5, #4877S, Cell Signaling Technology; 1:1000) or β-actin. Goat α-mouse (1706516, Bio-Rad; 1:50000) and donkey α-rabbit (711-035-152, Jackson ImmunoResearch; 1:50000) secondary antibodies were used in conjunction with Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare) to visualize protein bands. Densitometry analysis was performed by normalizing background-adjusted band intensities to the loading control using Fiji (ImageJ).

**Genomic DNA isolation for genotyping and Sanger sequencing**

Cells were lysed in genomic DNA lysis buffer (100 mM Tris HCl pH 8.0, 50 mM EDTA pH 8.0, 0.2% SDS, 200 mM NaCl) supplemented with 1 mg mL⁻¹ proteinase K at 55°C. Genomic DNA was extracted by isopropanol precipitation and resuspended in Tris-EDTA buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA pH 8.0). PCR genotyping was performed using the primers described in Supplementary Table 1. To assess p53RE mutagenesis, a region containing the p53RE was PCR-amplified from genomic DNA isolated from a heterogenous cell population, cloned into the pCR-Blunt II-TOPO vector (450245, Invitrogen), and analyzed via Sanger sequencing.

**Single-molecule RNA (smRNA) FISH**

SmRNA-FISH was performed as previously described (Dimitrova et al., 2014). Briefly, primary MEFs were seeded on coverslips and cultured for 24 hours in the presence or absence of etoposide before being fixed in 4% methanol-free formaldehyde in DEPC-treated water for 10 minutes at room temperature. After being washed twice in PBS, cells were dehydrated overnight in 70% ethanol at 4°C and stored for up to a week. Dehydrated coverslips were transferred to a hybridization chamber and equilibrated in Wash Buffer A for 5 minutes (LGC, Biosearch Technologies). Samples were incubated overnight at 30°C in hybridization buffer containing a 1:50 dilution of Stellaris probes conjugated to
Quasar 570 or Quasar 670 dye (Supplementary Table 1). The following day, samples were washed twice with Wash Buffer A for 30 minutes at 30°C, washed once with Wash Buffer B for 5 minutes at room temperature, and mounted in VECTASHIELD Antifade Mounting Medium with DAPI (Vector Laboratories). Samples were imaged using the Axio Imager 2 microscope system (Zeiss) with a PlanApo 63x 1.4 oil DIC objective lens (Zeiss).

**Chromosome conformation capture (3C)**

3C was performed as previously described (Hagège et al., 2007; Naumova et al., 2012), with some modifications. Briefly, 5-10x10⁶ cells were harvested by trypsinization, rinsed once in 1X PBS, and resuspended in 10 mL 1% formaldehyde in PBS. Samples were rotated end-over-end for 10 minutes at room temperature and quenched with 570 µL ice cold 2.5 M glycine. Samples were incubated for five minutes at room temperature and placed on ice for 15 minutes, with periodic inversion. Samples were spun down at 225g for 8 minutes at 4°C and washed once in 10 mL ice cold 1X PBS. Samples were flash frozen in a bath of dry ice and ethanol and stored for up to six months at -80°C.

For cell lysis, frozen samples were thawed on ice and resuspended in 5 mL ice cold lysis buffer (20 mM Tris HCl pH 8.0, 85 mM KCl, 0.5% NP-40, and 1X complete protease inhibitor tablet [Roche]) for two hours, with periodic inversion. Samples were spun down at 400g for 5 minutes at 4°C and resuspended in 0.5 mL 1.2X NEB Buffer 3.1 with 0.3% SDS. Next, samples were rotated end-over-end for 1 hour at 37°C, quenched with 2% Triton X-100, and again rotated end-over-end for 1 hour at 37°C. For chromatin digestion, samples were treated with 400 U BglII (NEB) and rotated end-over-end overnight at 37°C.

To inactivate the restriction enzyme, the SDS concentration was brought up to 1.6% and samples were shaken at 900 RPM for 25 minutes at 65°C. Next, samples were diluted by adding 6.125 mL 1.15X ligation buffer (10X ligation buffer: 600 mM Tris HCl pH 7.5, 50 mM DTT, 50 mM MgCl₂, 10 mM ATP) and Triton X-100 to a final concentration of 1%. After incubating for 1 hour at 37°C with gentle
shaking, 100 U of T4 DNA ligase (NEB) was added to each sample. All samples were incubated for 4 hours at 16°C followed by 30 minutes at room temperature.

To reverse the crosslinking, samples were treated with 300 μg Proteinase K and incubated overnight at 65°C with end-over-end rotation. The next morning, samples were treated with 300 μg RNAse A (Qiagen) and incubated for 45 minutes at 37°C. DNA was isolated using phenol-chloroform followed by ethanol precipitation. Purified DNA was resuspended in 150 μL 10 mM Tris pH 7.5. RT-qPCR was used to determine the concentration of each 3C library relative to a standard curve obtained by running serial dilutions of a control template. Libraries were diluted to a final concentration of 50 ng μL⁻¹. Relative interaction frequencies were measured by performing TaqMan qPCR on 50 ng 3C library using unidirectional primers in tandem with an anchor primer and fluorescent TaqMan probe targeting the LincRNA-p21 promoter (Supplementary Table 1). Interactions were normalized to a control region in the LincRNA-p21 locus.

To determine the digestion efficiency, chromatin samples from before and after BglII digestion were diluted in 500 μL Proteinase K buffer (5 mM EDTA, pH 8.0; 10 mM Tris-HCl, pH 8.0; 0.5% SDS) and de-crosslinked for 30 minutes at 65°C in the presence of 20 μg Proteinase K. Samples were treated with 1 μg RNAse A and incubated for 2 hours at 37 °C. Finally, DNA was isolated using phenol-chloroform followed by ethanol precipitation. Purified DNA was resuspended in 60-120 μL ultrapure water, and digestion efficiency was measured by performing qPCR with primer sets spanning BglII sites.

**Statistical analysis**

Statistical tests and sample sizes are included in each figure legend. All tests were performed in GraphPad Prism 7.0. The threshold for statistical significance was p<0.05, with *p<0.05, **p<0.01, ***p<0.001 and ****p<0.0001.
### Supplementary Table 1. Primers and oligos

<table>
<thead>
<tr>
<th>RT-qPCR primers</th>
<th>Target</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gapdh</td>
<td>AGCTTGTCATCAACGGGAAG</td>
<td>TTTGATGTTAGTGCGGTGTCTCG</td>
<td></td>
</tr>
<tr>
<td>Spliced LincRNA-p21</td>
<td>CCTGTCCCACTCGCTTTTC</td>
<td>GGAACCTGGGACGAGGAATGTC</td>
<td></td>
</tr>
<tr>
<td>LincRNA-p21 E1_1</td>
<td>GAGCCCCCATAGCCACAACCTC</td>
<td>AGTGGGTGTGTCACCTCTTG</td>
<td></td>
</tr>
<tr>
<td>LincRNA-p21 E1_2</td>
<td>CCTGTCCCACTCGCTTTTC</td>
<td>CCCACATGCTAGCTGTCATTG</td>
<td></td>
</tr>
<tr>
<td>LincRNA-p21 E2_1</td>
<td>GAGCTCTCATGCTCAGAGAAC</td>
<td>CCGTGACAGCAAGAATGCTTG</td>
<td></td>
</tr>
<tr>
<td>LincRNA-p21 E2_2</td>
<td>CCAACGTTAAGCCACCAAGGGA</td>
<td>CTTCATGGAAGGCAAGCAACG</td>
<td></td>
</tr>
<tr>
<td>p21</td>
<td>TCCACAGCGATATCCAGACA</td>
<td>GGACATCAACAGGATGGAC</td>
<td></td>
</tr>
<tr>
<td>p53</td>
<td>ACGCTTTCTCCAGAAGACTGG</td>
<td>ACGCGACCTGAGGCTGATA</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TT-qPCR primers</th>
<th>Target</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>GTGGAAGGGCTCATGATGTAGG</td>
<td>GGACACCTAAAGAAGGGTCC</td>
<td></td>
</tr>
<tr>
<td>LincRNA-p21 E1_1</td>
<td>GAGCCCCATAGCCACAACCTC</td>
<td>AGTGGGTGTGTCACCTCTG</td>
<td></td>
</tr>
<tr>
<td>LincRNA-p21 E1_2</td>
<td>CCTGTCCCACTCGCTTTTC</td>
<td>CCCACATGCTAGCTGTCATTG</td>
<td></td>
</tr>
<tr>
<td>LincRNA-p21 E2_1</td>
<td>GAGCTCTCATGCTCAGAGAAC</td>
<td>CCGTGACAGCAAGAATGCTTG</td>
<td></td>
</tr>
<tr>
<td>LincRNA-p21 E2_2</td>
<td>CCAACGTTAAGCCACCAAGGGA</td>
<td>CTTCATGGAAGGCAAGCAACG</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genotyping PCR primers</th>
<th>Target</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWI, PAS alleles</td>
<td>GAAGTGAACCACCCACTCAGC</td>
<td>CCCACATGCTAGCTGTCATTG</td>
<td></td>
</tr>
<tr>
<td>Δ127 allele</td>
<td>GAGCCCCATAGCCACAACCTC</td>
<td>CCCACATGCTAGCTGTCATTG</td>
<td></td>
</tr>
<tr>
<td>Δ151 allele</td>
<td>GAGCCCCATAGCCACAACCTC</td>
<td>CTCCTGTGGGAGAGTG</td>
<td></td>
</tr>
<tr>
<td>p53RE</td>
<td>GAGCCCCATAGCCACAACCTC</td>
<td>CTCCTGTGGGAGAGTG</td>
<td></td>
</tr>
</tbody>
</table>

| Primers for TaqMan RT-qPCR (3C) | | |
|----------------------------------|-----------------------|
| **TaqMan probe:**                | 5’-FAM-CCATCTCTCCAGCCCTCAAGAAGAT-BHQR-3’ |
| **Target**                       | **Sequence**          |
| Anchor (LincRNA-p21 promoter)    | CACACATGCAAAACACTCTTAA |
| Fragment 1                       | AAGAAGACTTCTGTAGGCTAGTC |
| Fragment 2                       | GTTGCCAGTATGGGTTCCTCC |
| Fragment 3                       | CTTCAGTTCTGCACACTAG |
| Fragment 4                       | AAGGTATTCTTTTGGTCAATG |
| Fragment 5                       | GTCTGGGTGGAGAAACTC |
| Fragment 6                       | TTCGTGCTGATCCATGTTG |
| Fragment 7                       | GCCCGTGGGACATTCTTAC |
| Fragment 8                       | ATATCTTCGTTCAGACACTCC |
| Fragment 9                       | GTAAGTACACTGTAGCTGTCCT |
| Fragment 10                      | CTCTTCACACACACAGAG |
## ASOs

<table>
<thead>
<tr>
<th>Target</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (non-targeting)</td>
<td>CCTTCCCTGAAGGTTCCTCC</td>
</tr>
<tr>
<td>LincRNA-p21 intron (ASO-I)</td>
<td>GCCATCCTCGGTACTGAGGT</td>
</tr>
<tr>
<td>LincRNA-p21 exon 2 (ASO-E2)</td>
<td>TCAGTTGACCCGGGCAAGGT</td>
</tr>
</tbody>
</table>

## RNA FISH probes

### LincRNA-p21 exonic probes (conjugated to Quasar670)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>lincRNA-p21_1</td>
<td>aggttgcacccagaagtcttg</td>
</tr>
<tr>
<td>lincRNA-p21_2</td>
<td>tccaagtgtttcctacactg</td>
</tr>
<tr>
<td>lincRNA-p21_3</td>
<td>gaaatgctcattccaccatc</td>
</tr>
<tr>
<td>lincRNA-p21_4</td>
<td>ctgaattgagaactggaaga</td>
</tr>
<tr>
<td>lincRNA-p21_5</td>
<td>ttgttgctgttgacacacaca</td>
</tr>
<tr>
<td>lincRNA-p21_6</td>
<td>ttctgtggtgtaggcag</td>
</tr>
<tr>
<td>lincRNA-p21_7</td>
<td>actgaaagggctatctecac</td>
</tr>
<tr>
<td>lincRNA-p21_8</td>
<td>cgacagtaatgtccatggg</td>
</tr>
<tr>
<td>lincRNA-p21_9</td>
<td>tgccttcataacacaggaggc</td>
</tr>
<tr>
<td>lincRNA-p21_10</td>
<td>atttgatttgggtgtgctg</td>
</tr>
<tr>
<td>lincRNA-p21_11</td>
<td>tgaagaggagtcacaggatt</td>
</tr>
<tr>
<td>lincRNA-p21_12</td>
<td>ataagttcaatgtctctgc</td>
</tr>
<tr>
<td>lincRNA-p21_13</td>
<td>gcaggtttctgttgaggaagt</td>
</tr>
<tr>
<td>lincRNA-p21_14</td>
<td>acacacacaccttcttggagc</td>
</tr>
<tr>
<td>lincRNA-p21_15</td>
<td>ttgacaacacttggcaactg</td>
</tr>
<tr>
<td>lincRNA-p21_16</td>
<td>gcagagtctgttcttatttt</td>
</tr>
<tr>
<td>lincRNA-p21_17</td>
<td>tcaagttggttaggcaaca</td>
</tr>
<tr>
<td>lincRNA-p21_18</td>
<td>aaagatggagacaggaagaact</td>
</tr>
<tr>
<td>lincRNA-p21_19</td>
<td>aagcttcagtcagctgttagt</td>
</tr>
<tr>
<td>lincRNA-p21_20</td>
<td>cctcagaggttagaactccat</td>
</tr>
<tr>
<td>lincRNA-p21_21</td>
<td>tactagttgagaagccat</td>
</tr>
<tr>
<td>lincRNA-p21_22</td>
<td>gagacacctctcaagcaact</td>
</tr>
<tr>
<td>lincRNA-p21_23</td>
<td>ttacagatgggggaacttgg</td>
</tr>
<tr>
<td>lincRNA-p21_24</td>
<td>tttcttctcttccaggtg</td>
</tr>
<tr>
<td>lincRNA-p21_25</td>
<td>tagagatagtagccaggat</td>
</tr>
<tr>
<td>lincRNA-p21_26</td>
<td>agttctgatagcactttactt</td>
</tr>
<tr>
<td>lincRNA-p21_27</td>
<td>attcttggaacacctttat</td>
</tr>
<tr>
<td>lincRNA-p21_28</td>
<td>ttctcttgagcactgagact</td>
</tr>
<tr>
<td>lincRNA-p21_29</td>
<td>ttcccctgacagcaagta</td>
</tr>
<tr>
<td>lincRNA-p21_30</td>
<td>aggtttagagagatgacaca</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>lincRNA-p21_31</td>
<td>atccattctcagactgaggt</td>
</tr>
<tr>
<td>lincRNA-p21_32</td>
<td>tgcaagtaatgagtccag</td>
</tr>
<tr>
<td>lincRNA-p21_33</td>
<td>agatgtatgccctgctgt</td>
</tr>
<tr>
<td>lincRNA-p21_34</td>
<td>agctacgtagtgaaaacag</td>
</tr>
<tr>
<td>lincRNA-p21_35</td>
<td>ttctctctacagagcagg</td>
</tr>
<tr>
<td>lincRNA-p21_36</td>
<td>ctctggtctggaactcata</td>
</tr>
<tr>
<td>lincRNA-p21_37</td>
<td>ttattgtggtctctctgct</td>
</tr>
<tr>
<td>lincRNA-p21_38</td>
<td>tcctggctcaatcgatacgt</td>
</tr>
<tr>
<td>lincRNA-p21_39</td>
<td>tgctctctctctctctgca</td>
</tr>
<tr>
<td>lincRNA-p21_40</td>
<td>agatgttggtctctctctctctc</td>
</tr>
<tr>
<td>lincRNA-p21_41</td>
<td>ttgaaaccagaggtgaact</td>
</tr>
<tr>
<td>lincRNA-p21_42</td>
<td>aaagaagaaaagggggaagg</td>
</tr>
<tr>
<td>lincRNA-p21_43</td>
<td>agcttccaggttctctaaag</td>
</tr>
<tr>
<td>lincRNA-p21_44</td>
<td>agaagtcacagagccacaa</td>
</tr>
<tr>
<td>lincRNA-p21_45</td>
<td>ggtcaactgtttcaaagagatgt</td>
</tr>
<tr>
<td>lincRNA-p21_46</td>
<td>atgaaagtcttttgctctcagaa</td>
</tr>
<tr>
<td>lincRNA-p21_47</td>
<td>ggctgacagaatggaactcag</td>
</tr>
<tr>
<td>lincRNA-p21_48</td>
<td>tcaagcttaagaggtcttgag</td>
</tr>
<tr>
<td><strong>LincRNA-p21 intronic probes (conjugated to Quasar570)</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intron lincRNA-p21_1</td>
<td>tgcgtgaggttctctaaaac</td>
</tr>
<tr>
<td>Intron lincRNA-p21_2</td>
<td>tcctggaaggtattacgga</td>
</tr>
<tr>
<td>Intron lincRNA-p21_3</td>
<td>gttatggtggcttgtcaacctac</td>
</tr>
<tr>
<td>Intron lincRNA-p21_4</td>
<td>ggtcaactgtttcacaacga</td>
</tr>
<tr>
<td>Intron lincRNA-p21_5</td>
<td>cacctctctctggaataggt</td>
</tr>
<tr>
<td>Intron lincRNA-p21_6</td>
<td>agtagcacaaggtgaaccttg</td>
</tr>
<tr>
<td>Intron lincRNA-p21_7</td>
<td>ctctgccaggtctcttaaag</td>
</tr>
<tr>
<td>Intron lincRNA-p21_8</td>
<td>attgctgggatgaccaacac</td>
</tr>
<tr>
<td>Intron lincRNA-p21_9</td>
<td>gtttttgacgacattgaca</td>
</tr>
<tr>
<td>Intron lincRNA-p21_10</td>
<td>atttaggcagcaattgaca</td>
</tr>
<tr>
<td>Intron lincRNA-p21_11</td>
<td>ageggcataaaattgtctga</td>
</tr>
<tr>
<td>Intron lincRNA-p21_12</td>
<td>gctctcagcagcttaaagt</td>
</tr>
<tr>
<td>Intron lincRNA-p21_13</td>
<td>cactcgcaatgaacacaagt</td>
</tr>
<tr>
<td>Intron lincRNA-p21_14</td>
<td>tctcttgggatttaagctg</td>
</tr>
<tr>
<td>Intron lincRNA-p21_15</td>
<td>tgcgtgaagttttagtgggc</td>
</tr>
<tr>
<td>Intron lincRNA-p21_16</td>
<td>taaagctcactgagctact</td>
</tr>
<tr>
<td>Intron lincRNA-p21_17</td>
<td>tcaaaatcataagccacggt</td>
</tr>
<tr>
<td>Intron lincRNA-p21_18</td>
<td>getcaactgtgctattagcc</td>
</tr>
<tr>
<td>Intron lincRNA-p21_19</td>
<td>cacagagccagggataaacc</td>
</tr>
<tr>
<td>Intron lincRNA-p21_20</td>
<td>acaagaaactttggaacatga</td>
</tr>
<tr>
<td>Intron lincRNA-p21_21</td>
<td>caggtgctactgtgcttctc</td>
</tr>
<tr>
<td>Intron lincRNA-p21_22</td>
<td>tagagctactgtgtctctc</td>
</tr>
<tr>
<td>Intron lincRNA-p21_23</td>
<td>gacagtgttaaactcagct</td>
</tr>
<tr>
<td>Intron lincRNA-p21_24</td>
<td>tctggaactcaggttcttgc</td>
</tr>
<tr>
<td>Intron lincRNA-p21_25</td>
<td>ttatagggtcactctagggc</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Intron lincRNA-p21_26</td>
<td>gttggcttcctgtgaataa</td>
</tr>
<tr>
<td>Intron lincRNA-p21_27</td>
<td>ctcattgtacggtcactc</td>
</tr>
<tr>
<td>Intron lincRNA-p21_28</td>
<td>ctaaatgcactgttcctt</td>
</tr>
<tr>
<td>Intron lincRNA-p21_29</td>
<td>gctgaaattataccagggg</td>
</tr>
<tr>
<td>Intron lincRNA-p21_30</td>
<td>ctaaaatgcacctgtgtcctt</td>
</tr>
<tr>
<td>Intron lincRNA-p21_31</td>
<td>agctgtttaaggcacatcaa</td>
</tr>
<tr>
<td>Intron lincRNA-p21_32</td>
<td>ttctacgctcttaatgggg</td>
</tr>
<tr>
<td>Intron lincRNA-p21_33</td>
<td>tcaaggttggagaagtctc</td>
</tr>
<tr>
<td>Intron lincRNA-p21_34</td>
<td>cctttgccctcttgactag</td>
</tr>
<tr>
<td>Intron lincRNA-p21_35</td>
<td>agctctaaataactcagaaga</td>
</tr>
<tr>
<td>Intron lincRNA-p21_36</td>
<td>ctttcattccgtcctaatggg</td>
</tr>
<tr>
<td>Intron lincRNA-p21_37</td>
<td>tcagatctgggagaagtctc</td>
</tr>
<tr>
<td>Intron lincRNA-p21_38</td>
<td>cctttgaccccttgcactag</td>
</tr>
<tr>
<td>Intron lincRNA-p21_39</td>
<td>agctcctaaactccagaaga</td>
</tr>
<tr>
<td>Intron lincRNA-p21_40</td>
<td>ctggttacgtaacacat</td>
</tr>
<tr>
<td>Intron lincRNA-p21_41</td>
<td>ctcaggttttggcaggaga</td>
</tr>
<tr>
<td>Intron lincRNA-p21_42</td>
<td>cacagcagtaacagcttta</td>
</tr>
<tr>
<td>Intron lincRNA-p21_43</td>
<td>cactttcttgactttagg</td>
</tr>
<tr>
<td>Intron lincRNA-p21_44</td>
<td>cagtcacgtgcacaaat</td>
</tr>
<tr>
<td>Intron lincRNA-p21_45</td>
<td>gctggcagttggagcttct</td>
</tr>
<tr>
<td>Intron lincRNA-p21_46</td>
<td>atttccttaaactcgtcttg</td>
</tr>
<tr>
<td>Intron lincRNA-p21_47</td>
<td>aagcctacatccttgagctg</td>
</tr>
<tr>
<td>Intron lincRNA-p21_48</td>
<td>gaaaggaaggccctttagg</td>
</tr>
</tbody>
</table>

**sgRNAs and dRNAs**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Guide name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>In vivo targeting for TWI/PAS insertion</strong></td>
<td>N/A</td>
<td>GGGACAGGCCAGGCCAGCTG</td>
</tr>
<tr>
<td><strong>LincRNA-p21 p53RE mutagenesis</strong></td>
<td>sgCon</td>
<td>GGCCACGAGTTCGAGATCGA</td>
</tr>
<tr>
<td><strong>LincRNA-p21 p53RE mutagenesis</strong></td>
<td>sgRE</td>
<td>GATGCAACCGAAGGCAAGCC</td>
</tr>
<tr>
<td><strong>LincRNA-p21 transcriptional inhibition</strong></td>
<td>dCon</td>
<td>gCGAGTTCGAGATCGA</td>
</tr>
<tr>
<td><strong>LincRNA-p21 transcriptional inhibition</strong></td>
<td>d1</td>
<td>gCTGCCGCTTTGCCC</td>
</tr>
<tr>
<td><strong>LincRNA-p21 transcriptional inhibition</strong></td>
<td>d2</td>
<td>GGGCGTGGGAGGTGT</td>
</tr>
<tr>
<td><strong>LincRNA-p21 CRISPRa</strong></td>
<td>A1</td>
<td>GGGGGCGCTCTGTTC</td>
</tr>
<tr>
<td><strong>LincRNA-p21 CRISPRa</strong></td>
<td>A2</td>
<td>GGGCTTTAATCAAC</td>
</tr>
<tr>
<td><strong>LincRNA-p21 CRISPRa</strong></td>
<td>A3</td>
<td>GCCATCCGTTGGCAAC</td>
</tr>
<tr>
<td><strong>LincRNA-p21 CRISPRa</strong></td>
<td>A4</td>
<td>GGTGGAAGAGTCCG</td>
</tr>
<tr>
<td><strong>LincRNA-p21 CRISPRa</strong></td>
<td>A5</td>
<td>AGTTTGTTACTGTGCA</td>
</tr>
<tr>
<td><strong>LincRNA-p21 CRISPRi</strong></td>
<td>I-6</td>
<td>CGGAGACTTGGGCGAGCCGCG</td>
</tr>
<tr>
<td><strong>LincRNA-p21 CRISPRi</strong></td>
<td>I-7</td>
<td>gAGTTTGTTACTGTGCAAGGAT</td>
</tr>
</tbody>
</table>
**LincRNA-p21 CRISPRi**  
I-8  
gTCCCAGAACAGAGCGCC

**LincRNA-p21 CRISPRi**  
I-9  
gCTGTTCTGGGAGATTAATC

**LincRNA-p21 CRISPRi**  
I-10  
GCTGCAGGTGAAGAGTCCG

**LincRNA-p21 CRISPRi**  
I-11  
GTCAGGACCACCTCAGCGC

**LincRNA-p21 CRISPRi**  
I-12  
GCTATGGGCTCAGGCGACT

**LincRNA-p21 CRISPRi**  
I-13  
gCTGGGAGATGATGCAACCGA

**LincRNA-p21 CRISPRi**  
I-14  
gACCAGGCAAGCCCGGGA

### HDR templates

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAS_HDR_template</td>
<td>ggtcagaagtgaacccacactcagectggaaaaaccagtaattatatctcaagacccaggaacttggacaacctcagAATAAAAGATCTTTTATTTCATTAGATCTGTGTGGTTTTTTTTGTTGGTGGTggctgctggctgcctcactcgettcctggtgcctgcc</td>
</tr>
<tr>
<td>TWI_HDR_template</td>
<td>tgcagaagtgaacccacactcagectggaaaaaccagtaattatatctcaagacccaggaacttggacaacctcagCAATAAAGCGGTTACAAGCCCGCAAAAATAGCAGAGTAATGTCGCGCATAGCGCCGCGATATTACAGCAGCTTTATTTCGGTggctgctggctgcctcactcgettcctggtgcctgcc</td>
</tr>
</tbody>
</table>

**Supplementary Table 2. Key plasmids used in this work**

<table>
<thead>
<tr>
<th>Name</th>
<th>Details</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCMV-dR8.2</td>
<td>Lentiviral packaging vector</td>
<td>Addgene #8455</td>
</tr>
<tr>
<td>pCMV-VSV-G</td>
<td>Lentiviral packaging vector</td>
<td>Addgene #8454</td>
</tr>
<tr>
<td>Brd001</td>
<td>Lentiviral; spCas9; hU6-gRNA; Puro</td>
<td>Broad Institute</td>
</tr>
<tr>
<td>Brd004</td>
<td>Lentiviral; spCas9; hU6-gRNA; GFP</td>
<td>Broad Institute</td>
</tr>
<tr>
<td>Lenti SAM Hygro</td>
<td>Lentiviral; p65-HSF1-MBP; hU6-gRNA-MS2; Hygro</td>
<td>Generated by E. Martínez-Terroba</td>
</tr>
<tr>
<td>MSCV puro</td>
<td>Retroviral; puro</td>
<td></td>
</tr>
<tr>
<td>pHAGE-TRE-dCas9-KRAB</td>
<td>Lentiviral; doxy-inducible dCas9-KRAB; Neo</td>
<td>Addgene #50917</td>
</tr>
<tr>
<td>Lenti sgRNA(MS2)_zeo backbone</td>
<td>Lentiviral; hU6-gRNA; Hygro</td>
<td>Addgene #61427</td>
</tr>
<tr>
<td>pCR-Blunt II-TOPO vector</td>
<td>Blunt DNA cloning vector</td>
<td>Invitrogen (K280002)</td>
</tr>
</tbody>
</table>
References


117


