The cardiac lncRNA *Chantico* directly regulates *Cxcl1* chemokine transcription

Johnathon M. Hall¹, Amoldeep S. Kainth^{1,*}, Megan J. Rowton^{3,*}, Emery Lu³, Gabriela Haddad¹, Amanda J. Keplinger¹, Jelena Šćepanović¹, Carlos Perez-Cervantes³, Sunny S.K. Chan⁴, Joseph A. Piccirilli², Ivan P. Moskowitz³, and Alexander J. Ruthenburg^{1,2}✉

¹Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL, USA

²Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL, USA

³Department of Pediatrics, University of Chicago, Chicago, IL, USA

⁴Department of Pediatrics, University of Minnesota, Minneapolis, MN, USA

*these authors contributed equally to this work

✉ e-mail: aruthenburg@uchicago.edu

Summary

In local gene regulation, long noncoding RNA (lncRNA) loci can function at three distinct, but non-mutually exclusive levels: the RNA molecule itself, the process of its transcription, and/or the underlying DNA element. Yet for *cis*-acting lncRNA, these distinctions are particularly challenging with present tools. To address this problem, we developed Omegazymes, catalytic nucleic acid enzymes to specifically target lncRNA for degradation without triggering premature termination of their transcription. We use Omegazymes to selectively target one of a highly refined set of lncRNAs in the mouse cardiomyocytes, demonstrating that the *Chantico* lncRNA molecule directly potentiates the transcription of *Cxcl1*, a neighboring chemokine gene. We find that the *Chantico* locus also acts at the DNA-level, as the binding of an essential cardiac transcription factor to the *Chantico* promoter is necessary for *Cxcl1* transcription in mature cardiomyocytes. This *Chantico* regulation of *Cxcl1* impacts cardiomyocyte signaling to immune cells, potentially regulating tissue-residence and inflammation.

Introduction

The advent of RNA-seq uncovered a wealth of transcription beyond annotated protein-coding genes (Carninci et al., 2005; Djebali et al., 2012; Mortazavi et al., 2008; The FANTOM Consortium et al., 2014). Indeed, there are more than three times the number of well-defined non-coding genes as compared to protein coding genes (Derrien et al., 2012; Iyer et al., 2015; Noguchi et al., 2017). Mutations in some long noncoding RNA (lncRNA) loci can disrupt fundamental biological processes to organismal health and development (Allou et al., 2021; Cho et al., 2018; Gupta et al., 2010; Lewandowski et al., 2019; Plenge et al., 1997; Rinn et al., 2007; Xiang et al., 2014). Among myriad other functions described, a small collection of lncRNAs or transcription through their loci, has been demonstrated to promote transcriptional activation of coding genes at a distance, seemingly by idiosyncratic mechanisms (Bose et al., 2017; Derrien et al., 2012; Di Ruscio et al., 2013; Groff et al., 2016, 2018; Lai et al., 2013; Lam et al., 2014; Li et al., 2013; Noguchi et al., 2017; Rahnamoun et al., 2018; Rinn and Chang, 2020; Wang et al., 2011; Werner and Ruthenburg, 2015; Werner et al., 2017; Yang et al., 2017). Given that more than 80% of the genome is transcribed at some level in cell-type specific manner (Carninci et al., 2005; Derrien et al., 2012; Djebali et al., 2012; The FANTOM Consortium et al., 2014), a salient challenge for the field is now defining which of these RNA species are functional versus transcriptional noise.

Upon finding a functional *cis*-acting lncRNA locus, discerning the molecular mechanism by which it acts presents a major challenge because there are multiple distinct and non-mutually exclusive modalities by which such regulatory elements may function and the toolkit to make these distinctions is limited. First, the underlying regulatory DNA element can influence local chromatin by recruiting factors that promote gene transcription (Engreitz et al., 2016; Paralkar et al., 2016; Yin et al., 2015). Second, the process of transcription through the lncRNA locus can have a regulatory effect on nearby genes by opening up local chromatin (Ebisuya et al., 2008; Engreitz et al., 2016; Holoch et al., 2021; Latos et al., 2012; Petruk et al., 2006). Third, the transcribed lncRNA molecule itself can act locally near its site of production, for example by recruiting regulatory factors to local chromatin (Engreitz et al., 2016; Lai et al., 2013; Li et al., 2013; Rinn et al., 2007; Wang et al., 2011; Yang et al., 2021). In practice, existing tools are insufficient to make these key distinctions. Although the DNA element can be functionally dissected, current methods cannot resolve whether an observed effect is due to the process of transcription at lncRNA loci or the lncRNA molecule. Promoter knockouts of lncRNA loci disrupt all three possible mechanisms (Sauvageau et al., 2013); and although polyadenylation signal

insertions (Engreitz et al., 2016) are less likely to disrupt regulatory DNA, these cause both premature transcription termination and reduction in lncRNA levels. Catalytically dead Cas9-fusion systems (Fulco et al., 2016; Gilbert et al., 2014; Joung et al., 2017; Shariati et al., 2019) can artificially recruit repressive factors, but they also perturb the transcriptional machinery in addition to affecting lncRNA expression levels. Antisense oligonucleotides (ASOs) target the lncRNA molecule directly for degradation, but RNaseH cleavage of the lncRNA leaves a 5' phosphate which is recognized by the transcription terminator Xrn2 to ultimately cause premature transcription termination (Donis-Keller, 1979; Lai et al., 2020; Lee and Mendell, 2020; Maamar et al., 2013; Zamecnik and Stephenson, 1978). Small interfering RNA (siRNA) also induce a 5' phosphate on the cleaved lncRNA (Schwarz et al., 2004) and do not efficiently reach nuclear RNA targets (Maamar et al., 2013). Targeting the 3' end of an RNA being transcribed has been proposed as a work-around (Lai et al., 2020; Lee and Mendell, 2020), but this requires this end of the molecule to be well-defined, as not often the case for low abundance lncRNA (Werner and Ruthenburg, 2015; Werner et al., 2017). Moreover, early sequence elements in the RNA may act locally during the course of transcription (Gorbovytska et al., 2021), making such perturbations that target 3'-termini ineffective tests of RNA molecule function. RNA-cleaving Cas systems may have poor specificity for their targets and leave 5' phosphate groups (Abudayyeh et al., 2017; Özcan et al., 2021), which may also trigger the Xrn2 transcriptional termination pathway. Therefore, the lncRNA field is in critical need of a tool capable of efficiently evaluating the possible molecular mechanisms of lncRNA loci.

Our prior work identified more than 1000 lncRNAs driven by the upstream regulator T-box transcription factor 5 (Tbx5) in mouse cardiomyocytes (Yang et al., 2017). Tbx5 is an essential regulator heart development and cardiac rhythm (Moskowitz et al., 2004; Nadadur et al., 2016) and binds both coding and noncoding regions of the genome (Akerberg et al., 2019; He et al., 2011; Yang et al., 2017). In other contexts, we have biochemically fractionated the nuclear pool of RNA to define a low abundance subclass of lncRNA termed chromatin-enriched RNA (cheRNA), whose presence is a stronger predictor of *cis*-gene transcriptional activity than any other genome scale-measure (Werner 2015), and in several cases, these cheRNA act to potentiate neighboring gene transcription (Werner 2017). Yet neither upstream regulators, nor spatial distribution of lncRNA, are alone sufficient to pinpoint a manageable number of lncRNA candidates for case studies of their mechanisms and functions.

Here we narrow down this large list of *cis*-correlated Tbx5-dependent lncRNA by defining the subset of these lncRNA that are also enriched in the chromatin fraction of mouse

cardiomyocyte nuclei. We identified a Tbx5-dependent cheRNA, Chantico, and investigated its potential to regulate neighboring chemokine gene transcription. To study the potential regulatory function of Chantico, we developed catalytic nucleic acid enzymes (Omegazymes) that can target RNA for degradation without disrupting transcriptional elongation at target loci, and subject these new tools to extensive functional validation. Using Omegazymes, we demonstrate that the Chantico RNA directly activates the nearby *C-X-C motif chemokine ligand 1* (*Cxcl1*) gene. We also find that a tbx5-binding element at the Chantico is also a DNA-level activator of the *Cxcl1* when both are spatially juxtaposed and in accessible chromatin. Finally, we reveal a cellular phenotype resulting from perturbation to the Chantico lncRNA in cardiomyocytes—modulation of chemotactic signaling to macrophages, demonstrating both the power of our pipeline to identify disease-associated lncRNA, and the efficacy of our RNA-targeting tools.

Results

Identification of chromatin-enriched and Tbx5-dependent lncRNA in cardiomyocytes

Although case studies of individual cardiac lncRNA have revealed deep insights into heart development and function (Anderson et al., 2016; Klattenhoff et al., 2013; Micheletti et al., 2017; Ounzain et al., 2015; Ritter et al., 2019; Xue et al., 2016; Yang et al., 2017), choosing a lncRNA to study among the plethora of potential candidates remains an outstanding challenge. To narrow the broad set of lncRNA with potential functional relevance in cardiomyocytes, we determined lncRNA transcripts (Yang et al., 2017) driven by a critical regulator of cardiomyocyte development and mature heart function: the Tbx5 transcription factor (Bruneau et al., 1999, 2001; Moskowitz et al., 2004; Nadadur et al., 2016; Takeuchi et al., 2003). Based on our prior studies of chromatin-enriched lncRNA (cheRNA), which occupy the same nuclear compartment as active transcription complexes and can activate neighboring gene transcription (Werner and Ruthenburg, 2015; Werner et al., 2017; Yang et al., 2017), we hypothesized cheRNA that were also regulated by Tbx5 would be strong candidate transcriptional regulators in cardiomyocytes. To define this set of lncRNAs, we performed nuclear fractionation of HL-1 mouse atrial cardiomyocytes and sequenced RNA from both the chromatin and soluble nuclear fractions to identify chromatin-enriched lncRNAs (Werner and Ruthenburg, 2015; Werner et al., 2017) (**Figure 1A**). We calibrated these RNA-seq libraries with spike-in RNA standards (Devonshire et al., 2010) (**Figure S1A-B**) to estimate the copy number of lncRNA per cardiomyocyte nucleus, with the majority ranging from 0.1 to 10 copies (**Figure 1B**), consistent with our prior estimates of cheRNA abundance in other cell lines (Werner and Ruthenburg, 2015; Werner et al., 2017).

We found that Tbx5-dependent lncRNA are significantly more chromatin-enriched and correlated with neighboring protein-coding gene expression than other annotated lncRNA or coding genes (**Figure 1C, Figure S1C**), suggesting a potential regulatory role in cardiomyocyte transcription. This calibrated ratiometric chromatin-enrichment RNA dataset represents a useful resource for further investigation of the transcriptome in the mouse heart (**Figure S1D-E**).

After defining the set of Tbx5-dependent and chromatin-enriched RNA correlated with neighboring Tbx5-dependent genes (**Figure S2A**), we sought to select those with the highest potential for regulatory functions in cardiomyocytes by culling species that lack independent promoters and well-defined transcriptional directionality. To this end, we filtered out RNA which stem from coding gene promoter upstream transcripts (Preker et al., 2008), transcriptional readthrough (Erokhin et al., 2015), short bidirectionally transcribed ncRNA (Kim et al., 2010; The FANTOM Consortium et al., 2014), and antisense lncRNA (Anderson et al., 2016; Carninci et al., 2005; Yang et al., 2021; Yelin et al., 2003). Lastly, we scored these cheRNA for local binding of the cardiac transcription factors Tbx5, Gata4, and Nkx2-5 (Akerberg et al., 2019; Bruneau et al., 1999, 2001; He et al., 2011; Moskowitz et al., 2004; Nadadur et al., 2016; Shikama et al., 2003; Takeuchi et al., 2003) and p300 histone acetyltransferase (by ChIP-seq peaks (Akerberg et al., 2019; He et al., 2011)), chromatin accessibility (ATAC-seq peaks (Yang et al., 2017)), and nascent transcription (CAGE-seq peaks (Noguchi et al., 2017)) in datasets derived from heart tissue to obtain cheRNA to enrich for candidates regulated by cardiac factors and marked by active transcription. After this stringent selection, we identified eleven cheRNA and confirmed by RT-qPCR that all are chromatin-enriched using previously characterized lncRNA (Ballarino et al., 2018; Yang et al., 2017) as controls (**Figure S2B-C**). Thus, our methodology uncovers a set of candidates (less than 0.2% of the cheRNA detected by RNA-seq of the nuclear fractions) with potential regulatory functions and relevance to cardiomyocyte biology. To validate this approach to nomination of Tbx5-dependent cheRNAs most likely to be functional in cardiomyocytes, we selected one candidate RNA for further study.

The design of tools that show a Tbx5-dependent cheRNA molecule activates neighboring *Cxcl1* gene transcription

Among the validated Tbx5-dependent and chromatin-enriched lncRNA, we identified a cheRNA with robust transcription from the *Cxcl* chemokine locus on the mouse chromosome 5 (**Figure 1D, Figure S3**). Although chemokines have not been investigated as part of the Tbx5 gene regulatory network in cardiomyocytes, studies of ectopic Tbx5 expression in synovial fibroblasts suggest Tbx5 has a role in activating chemokine genes (Karouzakis et al., 2014). The most

proximal gene to this cheRNA that is expressed in cardiomyocytes is *Cxcl1*, which encodes a chemokine involved in homeostatic immune cell function and inflammation in heart tissue (Bachmaier et al., 2014; Bajpai et al., 2018; Boisvert et al., 2006; Wang et al., 2018; Wu et al., 2021). For reasons we present later in this section, we named this cheRNA Chantico (*chromatin-enriched and Tbx5-dependent RNA activator of Cxcl1 transcription*).

To test whether the *Chantico* locus regulates *Cxcl1*, we first used antisense oligonucleotides (ASOs) to knockdown the cheRNA (**Figure 2A**). ASO knockdown of the Chantico cheRNA resulted in modest but significant decreases in *Cxcl1* transcript levels (**Figure 2B**), suggesting that *Chantico* regulates *Cxcl1*. However, ASO experiments cannot resolve whether *Cxcl1* regulation occurs by destruction of the Chantico cheRNA or by premature *Chantico* transcription termination induced by these reagents (Lai et al., 2020; Lee and Mendell, 2020), precluding unambiguous mechanistic delineation of the role of the Chantico RNA molecule from that of the *Chantico* locus.

To unequivocally determine the possible effect of Chantico cheRNA on *Cxcl1*, we developed a method capable of distinguishing between the mechanistic roles of transcription through noncoding loci and the cheRNA molecule in *cis*-gene regulation. RNaseH cleavage of the ASO-DNA:RNA heteroduplex leaves a 5' phosphate, which is recognized by the exoribonuclease Xrn2 (Jinek et al., 2011), resulting in premature transcription termination (Lai et al., 2020; Lee and Mendell, 2020) along with RNA destruction (**Figure 2A**). We hypothesized that the 10-23 class of RNA-cleaving DNAzymes, which leave a 5' hydroxyl after RNA cleavage (Santoro and Joyce, 1998) might be adapted to specifically target *cis*-acting cheRNA without prematurely terminating their transcription (**Figure S4A**, right panel). We first tested a composite nucleic acid enzyme following a previous design (Schubert et al., 2003) with 2'-O-Me ribonucleotides in place of 2'-deoxyribonucleotides at tolerated positions in the catalytic core that notably retained deoxynucleotides in the homology arms. With this composite version of the DNAzyme targeting *Chantico*, we observed similar effects on the neighboring *Cxcl1* gene transcription as with ASOs (**Figure S4A**). Yet the interpretation of this result remained uncertain as this reagent could act directly through its own enzymatic activity or trigger the RNase H/Xrn2 pathway by the patch of DNA-RNA heteroduplex formed upon target recognition. To resolve this problem, we designed a next generation of composite nucleic acid enzymes, which we term "Omegazymes", by supplanting all 2'- deoxyribonucleotides with 2'-O-Me ribonucleotides in the homology arms (**Figure 2C**).

As transcription of the *PPIB* gene was recently shown to be prematurely terminated upon ASO treatment in HCT116 cells (Lee and Mendell, 2020), we sought to compare Omegazyme

activity in this context (**Figure 2D**). All Omegazymes we tested significantly reduced PPIB mRNA levels and pre-mRNA levels upstream of the target cleavage site, demonstrating robust targeting efficiencies as compared to ASOs (**Figure 2E, Figure S4B**). To determine whether Omegazyme targeting impairs continued transcription elongation, we performed ChIP-qPCR for RNA Polymerase (Pol) II in the *PPIB* gene body at several sites downstream of where Omegazymes could target the nascent transcript. Unlike previous studies of the same gene in the same cell line with ASOs (Lee and Mendell, 2020), Omegazyme-mediated RNA degradation does not perturb Pol II occupancy downstream of pre-mRNA cleavage sites (**Figure 2F, Figure S4C**), nor are pre-mRNA regions downstream of the target site disproportionately lost (**Figure S4B**), consistent with evasion of Xrn2-degradation and normal elongation. Remarkably, the depletion effect size on PPIB introns upstream of the targeting site was greater than that downstream of the targeting site (**Figure S4B**), potentially enabling the use of a panel of Omegazymes spanning a *cis*-acting lncRNA to perform functional domain dissection *in situ*. We detect no significant off-target effects on very similar mRNAs, suggesting that the Omegazyme activity is highly specific (**Figure 2G-H**). Collectively, these data suggest that Omegazymes may be ideal tools to specifically interrogate the function of RNA molecules without the potential transcriptional consequences of all other available methodologies (Engreitz et al., 2016; Fulco et al., 2016; Lai et al., 2020; Lee and Mendell, 2020; Schwarz et al., 2004).

Next, we applied these Omegazyme tools to investigate the function of the Chantico cheRNA in HL-1 cardiomyocytes (**Figure 3A**). We ablated the Chantico cheRNA with several distinct Omegazymes and observed a marked reduction in *Cxcl1* transcription (**Figure 3B-C**). A control Chantico-targeting Omegazyme with catalytic core mutations that correspond to those that inactivate the 10-23 DNAzyme (Zaborowska et al., 2002) did not reduce Chantico or *Cxcl1* transcript levels, consistent with the interpretation that the Omegazymes function via target cleavage. As further confirmation of the Chantico RNA-level effect, we observed a similar reduction of Chantico and *Cxcl1* RNA from Chantico-cleaving Omegazyme treatment with and without concomitant Xrn2 knockdown (**Figure 3D, Figure S5**). This further supports the conclusion that Omegazyme-mediated target depletion is independent of Xrn2-mediated Pol II termination. Taken together, these data not only establish Omegazymes as a powerful new tool for distinguishing the molecular mechanisms of *cis*-acting cheRNAs but also demonstrate an essential role for the Chantico RNA molecule in activating *Cxcl1* transcription.

Tbx5 binding to the *Chantico* promoter in cardiomyocytes is essential to the activation of *Cxcl1* in *cis*

Having established a critical role for the Tbx5-dependent Chantico RNA molecule in transcriptional activation of *Cxcl1* in mouse cardiomyocytes, we next investigated the mechanism of this activation in greater detail. Analysis of our prior mouse atrial heart ncRNA datasets (Yang et al., 2017) suggest that Chantico transcription is dependent on the cardiac transcription factor Tbx5. Yet, whether Tbx5 directly regulates *Chantico* and which cardiac cell type contributed the signal to this analysis remained uncertain, so we first validated the direct function of Tbx5 in driving transcription from a minimal *Chantico* promoter in cardiomyocytes by luciferase assays (**Figure 4A**). When the 8 bp Tbx5 consensus binding site is deleted from this *Chantico* promoter, we observe a significant reduction in transcription of luciferase, establishing Tbx5 as a direct Chantico regulator. In this system, luciferase transcription is completely abrogated by the insertion of three strong polyadenylation signal elements between the *Chantico* promoter and luciferase gene, indicating that transcription of the luciferase gene in this system is driven solely by the *Chantico* promoter. We similarly validated the promoter activity of the Tbx5-dependent lncRNA IRENE-div (Salamon et al., 2020) (**Figure S2D-E**). These results demonstrate that direct Tbx5 binding promotes transcription from the *Chantico* promoter element.

Next, we investigated whether Tbx5 binding to the *Chantico* promoter can influence the expression of luciferase from the *Cxcl1* promoter in *cis* with a tethered enhancer assay (Zabidi et al., 2015). Indeed, we observe a significant increase in luciferase expression under the *Cxcl1* promoter when the *Chantico* promoter element is inserted in *cis* at an enhancer position (2 kb apart in this luciferase vector versus 58 kb at the endogenous locus), relative to expression under the *Cxcl1* promoter alone (**Figure 4B-C**). The Chantico promoter increased luciferase expression when cloned in either the forward or reverse orientation, which is a hallmark of enhancer elements (Banerji et al., 1981; Zabidi et al., 2015). This increase is entirely abrogated by the deletion of the Tbx5 consensus binding site within the *Chantico* promoter element in the enhancer position (**Figure 4B**). Although the 1127 bp *Chantico* promoter element contains less than 6% of the 19.8 kb region expressed endogenously, a small portion of the 5' end of the Chantico cheRNA could be transcribed from this element and affect transcription at the *Cxcl1* promoter. Thus, to query whether this enhancer activity is derived solely from Tbx5 binding, as opposed to additional regulatory sequence or the 5' end of the Chantico cheRNA, we evaluated the sufficiency of the 8 bp Tbx5 consensus binding sequence from the *Chantico* promoter in enhancer assays (**Figure 4C**). Remarkably, the insertion of just one Tbx5 consensus binding site at the *cis* enhancer position is sufficient for *Cxcl1* promoter activation slightly beyond what we observe from the whole *Chantico* promoter, suggesting the Tbx5 consensus binding site is

the only critical piece of the *Chantico* promoter element acting as an enhancer in this context. Moreover, the activation of the *Cxcl1* promoter in this context appears to be specific, as its replacement with a minimal *Tbp* promoter renders the system insensitive to the *Chantico* element (**Figure 4D**). Altogether, these data indicate that when the *Chantico* promoter and *Cxcl1* promoter are artificially juxtaposed within this vector reporter system, Tbx5 binding at the *Chantico* promoter is necessary and sufficient to stimulate transcription specifically from the *Cxcl1* promoter.

As there is clear *cis*-enhancer activity of this *Chantico* promoter element in this minimal heterologous reporter system, we sought to evaluate whether it could also act on the endogenous *Cxcl1* promoter in *trans* (when these DNA elements are unlinked). We observe no significant changes to *Cxcl1* RNA levels when the *Chantico* element is transiently transfected (**Figure 4E**), suggesting that endogenous *Cxcl1* gene activation requires the *Chantico* promoter element to act in *cis*. Collectively, these luciferase assay and Omegazyme targeting experiments suggest a model whereby *Chantico* acts at the RNA-level to bring its promoter element to the *Cxcl1* locus, where it exerts a DNA-level effect through bound Tbx5 in cardiomyocytes (**Figure 4F**).

Next, we investigated the role of Tbx5 in endogenous *Chantico*-*Cxcl1* gene regulation. To validate that Tbx5 regulates endogenous *Chantico* and *Cxcl1*, we induced transient overexpression of *Tbx5* in cardiomyocytes. We observe substantial increases in *Chantico* and *Cxcl1* RNA upon robust *Tbx5* overexpression in cardiomyocytes (**Figure 5A**). In contrast, we do not detect any *Chantico* or *Cxcl1* RNA in mouse embryonic stem cells (mESCs), with or without ectopic expression of *Tbx5* (**Figure 5B**), indicating that the ability of Tbx5 to activate *Chantico* or *Cxcl1* is specific to the differentiated cardiomyocyte state. Similar overexpression experiments in cardiomyocytes for other cardiac transcription factors Gata4 and Nkx2-5 resulted in no apparent changes in *Chantico* or *Cxcl1* transcript levels (**Figure S6A**), suggesting that *Chantico* and *Cxcl1* are specifically responsive to Tbx5 in cardiomyocytes.

Since the influence of Tbx5 in *Chantico*-*Cxcl1* gene regulation is absent in naïve mESCs but critical in mature cardiomyocytes, we evaluated the regulation of *Chantico* and *Cxcl1* through discrete stages of cardiac differentiation. To accomplish this, we directed the differentiation of mouse embryonic stem cells (mESCs) into cardiomyocytes (mESC-CMs) (Kattman et al., 2011; Rowton et al., 2020) and measured chromatin accessibility by ATAC-seq (Buenrostro et al., 2013) every 24 hours during differentiation and RNA abundance by RT-qPCR (**Figure 5C-D**). We observed low levels of *Chantico* transcription until day 8 (D8) of differentiation, roughly concurrent with inaccessible chromatin at the *Chantico* promoter. High *Chantico* expression at

the later stages of differentiation correlates with the emergence of accessible chromatin at the *Chantico* promoter. Importantly, increases in *Chantico* expression at the later stages of differentiation coincide with the highest *Cxcl1* expression levels, although the *Cxcl1* promoter maintains accessible chromatin through every stage of mESC-CM differentiation. Taken together, these data support the conclusion that a Tbx5 binding site at the *Chantico* promoter that becomes accessible in cardiomyocytes is essential for Tbx5-mediated *Cxcl1* expression at this stage. We also observe that Tbx5 occupancy is enriched at cardiomyocyte-specific ATAC peaks (**Figure S6B**), suggesting that a general mechanism of establishing the Tbx5 gene regulatory network throughout cardiac differentiation involves increasing the accessibility of Tbx5-binding sites.

***Cxcl1* gene regulation by *Chantico* can influence cardiomyocyte signaling to immune cells**

Next, we sought to test the functional role of the Tbx5-dependent *Chantico*-*Cxcl1* axis in cardiomyocytes. Recent data suggest that tissue-resident immune cells are critical for ongoing function of heart tissue (Hulsmans et al., 2017; Nicolás-Ávila et al., 2020; Simões et al., 2020), but little is known about the signaling between cardiac cells and these immune cell populations. We hypothesize that the *Cxcl1* chemokine gene regulation by *Chantico* ensures basal level secretion of chemokine ligand, which may recruit immune cells to the cardiomyocytes. To investigate *Cxcl1* as a potential *Chantico*-modulated signal from cardiomyocyte to immune cells, we first performed an ELISA for secreted *Cxcl1* protein in the media of cultured HL-1 cardiomyocytes treated with a *Chantico*-cleaving Omegazyme 1 (**Figure 6A**). Compared to the control, we detected significantly lower levels of *Cxcl1* chemokine secretion (**Figure 6B**). Conversely, we measure significantly elevated *Cxcl1* secretion in cells transiently overexpressing Tbx5. This trend in *Cxcl1* protein secretion demonstrates that both Tbx5 and *Chantico* cheRNA influence the secretion of chemokine ligand from cardiomyocytes. Finally, we performed a Boyden chamber transwell migration assay to test if cardiomyocyte chemokine secretion can attract macrophages (**Figure 6A**). In the lower compartments, we placed cardiomyocyte cell culture supernatant harvested from *Chantico*-cleaving Omegazyme knockdown versus control non-targeting Omegazyme treatment after 1 hour of secretion. Then, we seeded RAW 264.7 mouse macrophages in the upper compartment and allowed the macrophages to migrate for 1 hour. Secreted media from the *Chantico*-cleaving Omegazyme treatment decreased macrophage chemotaxis as compared to control (**Figure 6C**). These data

suggest that transcriptional regulation of *Cxcl1*-mediated signaling by the *Chantico* cheRNA in cardiomyocytes may be an important contributor to macrophage residence in the cardiac niche.

Discussion

We identified candidate *Tbx5*-dependent cheRNA through the intersection of two RNA-seq experiments: *Tbx5*-dependent noncoding RNA-seq in dissected mouse atrial tissue (Yang et al., 2017) and chromatin-enriched RNA-seq in cultured mouse cardiomyocytes. This intersection focused our attention on less than 10% of the previously annotated *Tbx5*-dependent noncoding RNA and less than 0.2% of the cheRNA detected with ratiometric interrogation of the chromatin RNA pool. With this refined set of candidates most likely to function as *Tbx5*-dependent regulators of neighboring gene transcription, we sought an unambiguous way to test the functionality of these RNA species. To this end, we developed Omegazymes to target cheRNA for degradation without causing premature transcription termination. Since the underlying DNA element and process of transcription at the cheRNA locus are not perturbed with Omegazyme treatment, we were able to conclusively assign a function to the *Chantico* cheRNA molecule in the transcriptional activation of its neighboring *Cxcl1* gene. We leveraged the dependence of *Chantico* on its upstream regulator *Tbx5* to study the mechanism of *Cxcl1* regulation in further detail. Specifically, we used *Tbx5* overexpression to drive further *Chantico*-*Cxcl1* gene regulation and dissected the roles for *Tbx5* binding to an accessible consensus site at the *Chantico* promoter in cardiomyocytes. Thus, the *Chantico* locus acts as both the source of a functional RNA and at the DNA-level as a binding platform for *Tbx5*, both of which appear to promote transcription of *Cxcl1* in cardiomyocytes. Finally, we applied Omegazymes to demonstrate that *Chantico* control of *Cxcl1* expression can markedly impact chemotactic signaling to macrophages in a cultured model, suggesting this regulatory circuit may be important for homeostatic maintenance of immune cells in the cardiac cell niche.

Advantages, limitations, and extensions of Omegazymes

The genetic study of lncRNA mechanism is challenging. It is hard to unambiguously distinguish any genetic alteration as acting at the level of the RNA molecule produced, as opposed to perturbation of the DNA element from which it is derived (e.g., disruption of a transcription factor binding site) (Bassett et al., 2014), or even the act of transcription itself being important (independent of the RNA molecule produced) (Ebisuya et al., 2008; Engreitz et al., 2016; Holoch et al., 2021; Latos et al., 2012; Petruk et al., 2006). CRISPRi and related technologies are powerful for disrupting transcription without alteration of the underlying DNA elements (Gilbert et al., 2014), however the latter, and more important distinction between the act of

transcription and the RNA molecule itself has proven more challenging. RNAi is generally ineffective against the nuclear pool of noncoding RNA as the RISC machinery is excluded from the nucleus (Bassett et al., 2014; Liang et al., 2011; Maamar et al., 2013), and recent work suggests that antisense oligonucleotides (ASOs) which act to target RNase H activity in the nucleus can impact both the RNA molecule and its continued transcription (Lai et al., 2020; Lee and Mendell, 2020; Maranon and Wilusz, 2020). Specifically, ASO-guided RNase H cleavage produces a 5'-phosphate which is critical to the loading of XRN2, a processive 5' to 3' exonuclease that can terminate transcription if the cleavage occurs during elongation (West et al., 2004). To remedy this current technology gap for examining the specific effects of noncoding RNA molecules within the nucleus, we have pioneered the development of mixed 2'-O-methyl/DNAzymes to target these RNAs selectively without impact on transcription. We adapted the 10-23 DNAzyme (Santoro and Joyce, 1997), to have no DNA in the target complementarity segments, thereby achieving cleavage while preventing RNase H-triggered activity from occurring. Critically, these nucleic acid enzymes, which we term "Omegazymes" yield 2'-3'-cyclic phosphates and unphosphorylated 5' ends, which preclude the loading of Xrn2, thereby specifically targeting the RNA molecule without transcriptional elongation effects.

Omegazymes present a non-invasive means of interrogating RNA molecular function that can be rapidly applied to any system. Indeed, Omegazymes also only require a purine-uridine motif in the target RNA sequence, making nearly any RNA region targetable. In our limited deployment of these reagents, Omegazymes yield generally larger knockdown effect sizes than ASOs. For example, we observe a greater degree of knockdown with the Chantico-cleaving Omegazyme #1 compared to the Chantico-targeting ASO #2, which target nearly the same sequence of the Chantico cheRNA (12 out of 14 nt overlapping). Additionally, PPIB-targeting Omegazymes display more potent knockdowns than previously observed with ASOs (Lee and Mendell, 2020). In our limited application of Omegazymes to PPIB, we note that we do not detect any apparent difference in the efficiency of targeting exon or intron as has been observed for ASOs (Lai et al., 2020, compare Omegazymes 1-4 to 5-10, **Figure 2E, S4C**).

Although ASOs targeting the 3' end of an RNA may escape apparent premature transcription termination (Lai et al., 2020; Lee and Mendell, 2020), transcription-level effects remain a formal possibility for this approach that can be difficult for the experimenter to rule out. In cases of genes that are transcribed at a low level, such as *Chantico*, such evaluation by ChIP is extremely challenging due to signal/noise limitations. This 3' end design restriction can be problematic for studies of *cis*-acting lncRNA in other ways: *i.*) The variety of alternative TSSes coupled with a variety of splicing and processing outcomes for lncRNA can lead to multiple

lncRNA isoforms (Carninci et al., 2005; Cho et al., 2018; Lagarde et al., 2017; Schlackow et al., 2017), many of which will not be uniquely targetable when restricted to 3'-targeting. For instance, targeting the 3' end of the longer isoforms of lncRNA *COOLAIR* with an ASO may leave Class II transcripts, which have 3' ends far upstream of other *COOLAIR* isoforms, unperturbed despite their influence on *FLC* gene regulation (Marquardt et al., 2014). *ii.*) Ambiguity about the precise 3' end of a low abundance transcript due to sequencing depth limitations (Werner and Ruthenburg, 2015; Werner et al., 2017), renders the design of ASOs to target this region challenging. *iii.*) When constrained to targeting the 3' end the functional probing of a *cis*-acting lncRNA that acts co-transcriptionally before the 3' end targeting site is accessible may be impossible. For instance, the 3' end of the lncRNA may be inaccessible from tethering to chromatin by a stalled RNA polymerase II prior to reaching the 3' end of the transcriptional unit (Werner and Ruthenburg, 2015; Werner et al., 2017). *iv.*) As the 5' end of the lncRNA is typically the most conserved (Ruiz-Orera and Albà, 2019) and can be critical for mediating protein interactions (Gorbovytska et al., 2021; Somarowthu et al., 2015), cleavage too distal to this end may have blunted impact as a functional perturbation. *v.*) The 3' end of a lncRNA may be structured (Brown et al., 2014), limiting ASO knockdown efficacy. Unconstrained by the 3'-end targeting requirement for unambiguous interpretation of ASO consequences, Megazymes can also be used to selectively degrade virtually any RNA isoform.

Despite the advantages afforded by Megazymes, there are some considerations and possible limitations of their use. Although we do not observe significant changes to mRNA levels that have slight complementarity mismatches with the Megazyme targeting sequence, even when distal to the cleavage site, we have not yet systematically explored the potential off-target effects. The 2-O-methyl groups in the Megazyme design may increase target affinity and non-specific RNA binding compared to the unmodified 10-23 DNAzyme (Schubert et al., 2003; Tsourkas et al., 2002). We have also arbitrarily chosen lengths of homology arms to match those of prior DNAzyme work (Schubert et al., 2003) and have not optimized their length/predicted annealing temperature to accommodate for distinct hybridization energetics. Nor have we established a position weight matrix for each homology arm that could enable design of more efficient and specific Megazymes. Certain sequences in the homology arms may disfavor the ideal fold for RNA cleavage by forming structures with portions of the catalytic core and thereby inhibiting catalysis. As with ASO and siRNA experiments, the interpretation of Megazyme experiments as representing on-target cleavage is strengthened by the recapitulation of the observed effect with one or more additional Megazymes designed against

distinct sequences in the same target. Omegazymes also share similar limitations to ASOs in that they require transfection, do not completely ablate target RNA levels, and likely display limited duration of activity in treated cells.

In future work, we envision that Omegazymes could be improved with additional modifications and design considerations. Incorporating different types of nucleic acid modifications and varying which positions are modified in the 10-23 DNAzyme design may improve intracellular catalytic activity (Schubert et al., 2003; Wang et al., 2021), delivery to certain tissues (Nagata et al., 2021), extend intracellular half-lives, and/or targeting as well as catalytic efficiency of these nucleic acid enzymes. For instance, adding 2'-fluoroarabino modifications to positions G2 and U8 in the catalytic core of the 10-23 DNAzyme increased catalytic activity by about 50% relative to the unmodified DNAzyme (Wang et al., 2021). Since Omegazymes show strong and specific knockdown of their targets, these could be suitable for large-scale functional screens of lncRNA as akin to those performed with ASOs (Ramilowski et al., 2020). Altogether, we reason future studies of modification to the 10-23 DNAzyme scaffold and sequence determinants of catalytic activity may yield next-generation catalytic nucleic acid enzymes with further improved performance.

In the two cases we have examined, we find that Omegazyme knockdown significantly reduces levels of nascent RNA upstream of the cleavage site while leaving downstream RNA largely intact, likely due to the 3' to 5' exonuclease activity of the exosome, the predominant exonuclease in the nucleus (Ibrahim et al., 2008; Lubas et al., 2015; Preker et al., 2008; Schneider and Tollervey, 2013). Thus, Omegazymes might be used to dissect the functional requirement of regions within the target lncRNA by comparing the Omegazyme targeting different regions of the RNA to a sensitive functional readout. Omegazymes could also be applied side-by-side with ASOs to assign function exclusively to the process of transcription in the case that an observed effect is altered with ASOs but not Omegazymes targeting the same RNA to similar extents.

Chantico acts locally to regulate *Cxcl1* gene transcription

Using Omegazymes, we establish that the Tbx5-dependent cheRNA Chantico directly regulates the nearby gene *Cxcl1*. This function of local gene activation is consistent with ASO and siRNA experiment-based models of estrogen receptor α -dependent enhancer RNA (Li et al., 2013), noncoding RNA-a (Lai et al., 2013), and our prior studies of cheRNA-gene pairs (Werner et al., 2017; Yang et al., 2017). Although all these studies are unable to exclude transcription-level effects, they are interpreted as cases where both the noncoding RNA transcript and its

underlying DNA promote contributing to local gene activation via the RNA spatially apposing the two DNA elements in some fashion. Indeed, transcription at active enhancers (Hsieh et al., 2020; Hua et al., 2021; Isoda et al., 2017) and scaffold lncRNA (Creamer et al., 2021; Hacisuleyman et al., 2014; Quinodoz et al., 2021) have been directly implicated in partitioning enhancers and gene promoters into domains of chromatin interaction. Consistent with this, we observe the cheRNA *Chantico* is required for endogenous *Cxcl1* activation and that a Tbx5-bound *Chantico* promoter can enhance *Cxcl1* transcription when these elements are spatially constrained together. Although it is formally possible that the small portion of *Chantico* produced by the *Chantico* promoter element we use in our luciferase assays plays a role in this process, we favor the interpretation that physical linkage of these two DNA elements nearby in the same plasmid as opposed to the normal 57.7 kb that interleave them in the genome circumvents the need for *Chantico* RNA in these assays. Although Tbx5 occupies both the *Chantico* and the *Cxcl1* promoters (He et al., 2011), we only observe a consensus binding site (Ghosh, 2001) for Tbx5 at the *Chantico* promoter. These data support a model wherein Tbx5 bound at the *Chantico* promoter may contact the *Cxcl1* promoter for activation in a *Chantico* cheRNA-dependent manner, where the RNA may be playing a role of bridging these two DNA elements through space.

Implications of cardiomyocyte-to-immune cell signaling

Here, our systematic approach to functional lncRNA discovery revealed that transcription and secretion of the *Cxcl1* chemokine in cardiomyocytes is regulated by the *Chantico* lncRNA, which is in turn, regulated by the cardiac transcription factor Tbx5 (**Figure 6D**). These types of cardiomyocyte-to-immune cell signals must be carefully modulated to preserve robust heart function and development without inducing excessive inflammation. For instance, macrophages facilitate electrical conduction (Hulsmans et al., 2017), mitochondrial homeostasis (Nicolás-Ávila et al., 2020), and tissue repair (Simões et al., 2020) for the heart. Conversely, inflammation from the excessive recruitment of monocytes and other immune cells can contribute to diseases like autoimmune myocarditis (Barin et al., 2012). Our systematic approach to lncRNA identification yielded lncRNA with biological relevance to cardiomyocytes; we anticipate that the Omegazyme tools we establish here will enable further mechanistic interrogation of transcriptional regulation in cardiomyocytes and beyond.

Acknowledgements

We thank Mikayla Marchuk and Pieter Faber for assistance with sequencing at the University of Chicago Genomics Facility. We are grateful to the Anindita Basu, Yingming Zhao, and Ralph Weichselbaum laboratories for assistance with cell line acquisition. We are thankful for sequencing data deposited by ENCODE, the FANTOM5 consortium, and the William Pu laboratory. We thank Claire Kokontis, Julia Meng, Gönen Memişoğlu, and Rohan Shah for manuscript feedback. J.M.H. received the American Heart Association Predoctoral Fellowship (20PRE35180107) and support from the NIH Genetics & Regulation Training Grant (T32 GM07197). G.H. and A.J.K. were supported by the NIH Genetics & Regulation Training Grant (T32 GM07197), and G.H. was also supported by the Genetic Mechanisms and Evolution Training Grant (T32 GM139782). A.J.K. was supported by an NSF Graduate Research Fellowship. This work is supported by grants to A.J.R. and I.P.M. from the NIH (R01-HL148719-01) and an AHA Collaborative Sciences Award (17CSA33610126).

Author Contributions

A.J.R. and J.M.H. designed the study and wrote the paper with input from all other authors. J.A.P. suggested the use of catalytic nucleic acids for RNA knockdown, which J.M.H. developed and implemented. A.J.K. performed nuclear fractionation and RT-qPCR for Figure S2C. G.H. performed the luciferase experiment for Figure S2E and developed the Omegazyme design pipeline. S.S.K.C. generated the iTBX5 mESC line. M.J.R. and E. L. maintained and differentiated mESCs-CMs and performed ATAC-seq time series experiments with oversight from I.P.M. C. P-C. analyzed ATAC-seq time series data with oversight from I.P.M. J.Š. performed mESC culture and transfection for Figure 5B. A.S.K. performed ChIP-qPCR experiments. J.M.H. performed experiments and analysis for all other presented data.

Figure 1 | Nuclear Fractionation of HL-1 mouse atrial cardiomyocytes reveals a cheRNA transcribed from the mouse C-X-C chemokine locus.

(A) Schematic of calibrated chromatin-enriched RNA-seq.

(B, C) Violin-boxplots of $\log_{10}$ copies per cell and $\log_{10}$ chromatin-enrichment for protein-coding RNA (n = 17,913), Gencode annotated lncRNA (Frankish et al., 2019) (n = 2654), and Tbx5-dependent noncoding RNA present in the annotated lncRNA set and discovered in Yang et al. 2017. (n = 849). *** $p < 2 \times 10^{-16}$, Wilcoxon test.

(D) The mouse C-X-C chemokine locus with average (dark lines) chromatin-enriched RNA and soluble nuclear RNA contoured over their sites of origin (mean FPKM, dark lines $\pm$ S.D., lighter color, n=3 independent experiments) corresponding to + and - strand (above and below origin, respectively). Transcription factor (Tbx5, Gata4, Nkx2-5) and p300 transcriptional co-activator ChIP peaks (He et al., 2011) from HL-1 cardiomyocytes are also shown. The putative Chantico promoter is highlighted in yellow across all tracks. The chromatin-enriched and Tbx5-dependent RNA activator of Cxcl1 transcription (*Chantico*) and protein-coding genes are annotated above sequencing data.

Figure 2 | The Development of Omegazymes, efficient RNA-depletion nucleic acid enzymes that do not disrupt transcription of target RNA.

(A) Diagram of the chemical mechanism of ASO-directed RNA cleavage by RNase H and a cartoon of the consequences of this cleavage if it occurs during transcription. Notably the 5'-phosphate product of ASO-guided RNase H cleavage can recruit the Xrn2 nuclease to terminate transcriptional elongation (Jinek et al., 2011; Lai et al., 2020; Lee and Mendell, 2020) ASOs are frequently used to functionally interrogate long noncoding RNAs (lncRNAs).

(B) RT-qPCR of Chantico, Cxcl1 mRNA, and Cxcl1 pre-mRNA levels in HL-1 cardiomyocytes with two distinct Chantico-targeting ASOs (LNA-GapmeR) compared to a non-targeting ASO control. ($n \geq 3$, * $p < 0.05$, ** $p < 0.005$, unpaired student's t-test).

(C) Schematic of Omegazyme composition and the mechanism of target RNA cleavage that by analogy to the parental 10-23 scaffold should afford 2',3'-cyclic- phosphate and unphosphorylated 5'-hydroxyl products (Santoro and Joyce, 1998; Schubert et al., 2003) S. W. and Joyce, G. F. (1997). Cartoon of transcription-concurrent Omegazyme activity-- the polymerase-proximal cleavage product of the Omegazyme product (free 5'-hydroxyl) would not be anticipated to recruit Xrn2, thus this cleavage could escape premature termination observed with ASOs.

(D) Schematic of *PPIB* gene structure shown to be susceptible to ASO-mediated transcriptional termination (Lee and Mendell, 2020) ASOs are frequently used to functionally interrogate long noncoding RNAs (lncRNAs, with Omegazyme/ ASO target sites for the corresponding pre-mRNA and amplicons used in panels E-F indicated.

(E) RT-qPCR of *PPIB* RNA from HCT116 cells transfected with indicated Omegazymes ($n \geq 3$, ns = not significant, * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, unpaired student's t-test against the non-targeting control Omegazyme control).

(F) Pol II ChIP-qPCR throughout the *PPIB* gene body upon Omegazyme treatment displays no apparent elongation deficits ($n = 4$, ns = not significant, unpaired student's t-test).

(G) Schematic of target mismatches queried in panel H.

(H) RT-qPCR of transcripts with 1, 2, or 3 mismatches with an Omegazyme used in HCT116 cells. no significant variation of off-target expression by unpaired student's t-test.

Figure 3 | Omegazyme tools demonstrate that the Chantico cheRNA molecule functionally contributes to Cxcl1 gene activation.

(A) Schematic of *Chantico* with Omegazyme/ASO target sites for the corresponding cheRNA indicated.

(B) RT-qPCR of Chantico cheRNA (at amplicon B), Cxcl1 pre-mRNA, and Cxcl1 mRNA (exon-spanning primer set) in HL-1 cardiomyocytes with catalytically dead Omegazyme (with the same targeting arms as Chantico Omegazyme 1, with inactivating mutations in the loop (Zaborowska et al., 2002)) or Chantico-cleaving Omegazyme 1 treatment as compared to a non-targeting Omegazyme control (n = 3, * p < 0.05, ** p < 0.005, unpaired student's t-test).

(C) Similar experiment to previous panel, comparing two additional Omegazymes to a non-targeting control (n = 3, * p < 0.05, unpaired student's t-test).

(D) RT-qPCR of Chantico lncRNA (at amplicon B), Cxcl1 mRNA, and Cxcl1 pre-mRNA in HL-1 cardiomyocytes with concomitant Omegazyme and siRNA treatment targeting Xrn2 (n = 3, ns = not significant, * p < 0.05, ** p < 0.005, unpaired student's t-test).

Figure 4 | Tbx5 binding at the *Chantico* promoter is essential for enhancer activity on the *Cxcl1* promoter.

(A) Ratiometric luciferase assays in HL-1 cardiomyocytes transfected with the putative *Chantico* promoter (1127 bp, delimited by ATAC peak in Figure 4C) and variations as indicated directly driving firefly luciferase transcription, normalized to the minimal promoter negative control vector and to co-transfected *Renilla* luciferase (n = 3 independent replicates, ** $p < 0.005$, unpaired student's t-test).

(B) Luciferase assays as in panel A, but with the minimal *Cxcl1* promoter (from 250 bp upstream to the TSS) driving luciferase transcription with variation at the *cis* enhancer position, ranging from no inserted element, to the putative *Chantico* promoter element in two different orientations, or with the 8 bp *Tbx5* consensus site deleted (n = 3, ns = not significant, *** $p < 0.0005$, unpaired student's t-test).

(C) Luciferase assays as in panel B, with the additional vectors containing one or two 8 bp *Tbx5* consensus binding sites inserted at the *cis* enhancer position (n = 3, *** $p < 0.0005$, unpaired student's t-test).

- (D) Luciferase assays as in panel B, with the substitution of the *Tbp* promoter for the *Cxcl1* promoter with and without the Chantico promoter in the enhancer position ($n \geq 3$, *** $p < 0.0005$, unpaired student's t-test).
- (E) RT-qPCR of *Cxcl1* pre-mRNA and *Cxcl1* mRNA from HL-1 cardiomyocytes transfected with the putative Chantico promoter or pUC19 vector control ($n = 3$, ns = not significant, unpaired student's t-test).
- (F) Model of DNA-level and RNA-level activity of *Chantico* on the *Cxcl1* gene.

Figure 5 | Accessible chromatin at the *Chantico* promoter is required for *Cxcl1* gene activation by Tbx5.

(A) RT-qPCR of Chantico, Cxcl1 pre-mRNA, Cxcl1 mRNA, and Tbx5 mRNA levels relative to eIF3s5 housekeeping gene in HL-1 cardiomyocytes transfected with control mCherry expression or Tbx5 expression vectors (n=3). ns = not significant, * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, unpaired student's t-test.

(B) Same experiments as panel A, but in E14 mESCs (n = 2).

(C) Chromatin accessibility of the mouse chemokine locus by ATAC-seq is shown for the time course (2i stage to Day 15, scale: 0-16 fold-enrichment) of directed differentiation of mESCs to cardiomyocytes and compared to HL-1 cardiomyocytes (scale 0-11 fold-enrichment for all). The putative Chantico promoter element used for luciferase assays is highlighted in yellow.

(D) RT-qPCR of Chantico and Cxcl1 mRNA levels relative to the eIF3s5 housekeeping gene at indicated days in the differentiation of mESCs to cardiomyocytes (n = 2).

Figure 6 | Cxcl1 gene regulation by Chantico in cardiomyocytes can influence macrophage recruitment.

(A) Schematic of experimental set-up: 48 hours post-transfection of HL-1 cardiomyocytes, cell culture media was replaced and media with secreted protein was used in an ELISA for Cxcl1 protein or Boyden chamber transwell migration assays.

(B) ELISA for Cxcl1 protein secretion at 15 minutes and 1 hour time points normalized to calibration curve of Cxcl1 (n = 3, * $p < 0.05$, ** $p < 0.005$, unpaired student's t-test) upon Chantico knockdown with Omegazyme #1, or Tbx5 overexpression by transient transfection.

(C) Boyden chamber transwell migration assay is shown to quantify RAW 264.7 mouse macrophage chemotaxis in response to HL-1 cell culture media with secreted protein collected at the 1 hour time point as in B (n = 3, * $p < 0.05$, unpaired student's t-test).

(D), Model of Cxcl1 transcriptional activation at the DNA and RNA-level by the Chantico locus in mature cardiomyocytes, but not mESCs, that results in increased macrophage chemoattraction.

Figure S1, related to Figure 1. Chromatin enrichment and abundance signatures of distinct classes of RNA.

- (A) Linear regression analysis of $\log_{10}$ total (CPE + SNE) FPKM vs $\log_{10}$ attomoles of ERCC spike-in RNA added, which is used to estimate copies per cell transcripts in RNA-seq data.
- (B) Linear regression analysis of $\log_{10}$ SNE FPKM vs $\log_{10}$ CPE FPKM of ERCC spike-in RNA standards, which is used to calculate a scalar correction constant.
- (C) $\log_{10}$ SNE FPKM of nearest protein-coding gene for protein-coding, Gencode annotated lncRNA (Frankish et al., 2019), *cis*-correlated Tbx5-dependent ncRNA (Yang et al., 2017), and novel cheRNA. ns = not significant, **** $p < 10^{-8}$, Wilcoxon test.
- (D) Linear regression analysis of $\log_{10}$ total chromatin-enrichment (CPE / SNE) vs $\log_{10}$ total (CPE + SNE) FPKM for all transcripts (excluding spike-in RNA). Each dot represents a single RNA species measured in the cheRNA-seq libraries and alpha transparency values are set to 0.025.
- (E) $\log_{10}$ total (CPE + SNE) FPKM at the same scale is shown on each x-axis. $\log_{10}$ fold chromatin enrichment (CPE FPKM / SNE FPKM) at the same scale is shown on each y-axis. Each plot represents a different class of RNA from the gencode annotation or Tbx5-dependent ncRNA annotation. Data points for protein-coding genes were adjusted to an alpha transparency value of 0.15 in ggplot2 to avoid over-plotting. Each dot represents a single RNA species measured in the cheRNA-seq libraries.

Figure S2, related to Figure 1. Filtering scheme for and validation of Tbx5-dependent, chromatin-enriched lncRNA.

(A) Venn diagram (left) intersecting the 1,577 Tbx5-dependent ncRNA correlated with neighboring Tbx5-dependent gene expression (Yang et al., 2017), novel cheRNA, and the 11 candidates identified after stringent filtering for independent

promoters and clear transcription directionality (right).

(B) Genome browser view of the Charme (Ballarino et al., 2018) locus. Three biological replicates of RNA-seq data from the chromatin fraction and soluble nuclear fraction are shown in addition to Tbx5 ChIP (He et al., 2011) and CAGE-seq (Noguchi et al., 2017) peaks.

(C) Log₁₀ fold chromatin-enrichment in HL-1 cells measured by RT-qPCR with RNA from the chromatin and soluble nuclear fractions.

(D) View of the IRENE-div locus (Salamon et al., 2020). Three biological replicates of RNA-seq data from the chromatin fraction and soluble nuclear fraction are shown in addition to Tbx5 ChIP (He et al., 2011) and CAGE-seq peaks (Noguchi et al., 2017). The putative promoter region of IRENE-div cloned for the luciferase assay in panel E is highlighted in purple.

(E) Luciferase assay analysis in transfected HL-1 cardiomyocytes of the IRENE-div putative promoter compared to a minimal promoter control driving firefly luciferase expression (n = 3, ns = not significant, * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$, unpaired student's t-test).

Figure S4, related to Figure 2. Omegazymes can target RNA for degradation with minimal off-target effects.

(A) RT-qPCR of Chantico (2 primer sets), Cxcl1 pre-mRNA, and Cxcl1 mRNA in HL-1 cardiomyocytes transfected with transfection reagent alone, or Chantico-cleaving DNAzyme based on the DV15E4 scaffold (Schubert et al., 2003), as compared to non-targeting DNAzyme ($n \geq 2$, ns = not significant, * = $p < 0.05$, unpaired student's t-test).

(B) Pol II ChIP-qPCR throughout the PPIB gene body upon Omegazyme treatment normalized to input ($n = 4$, ns = not significant, unpaired student's t-test).

(C) RT-qPCR of each intron in the PPIB pre-mRNA in HCT116 cells transfected with non-targeting control or PPIB-targeting Omegazymes ($n \geq 3$, ns = not significant, * $p < 0.05$, ** $p < 0.005$, unpaired student's t-test).

Figure S5, related to Figure 3. Xrn2 is efficiently targeted in the conditions used in Figure 3D. RT-qPCR of Xrn2 relative to eIF3s5 under the four conditions displayed in Figure 3D.

Figure S6, related to Figure 5. Tbx5 is enriched at cardiac-specific ATAC peaks and is the main driver of Chantico/Cxcl1 transcription in cardiomyocytes.

(A) RT-qPCR of Chantico, Cxcl1 mRNA, and Cxcl1 pre-mRNA levels in HL-1 cardiomyocytes upon transient Gata4 or Nkx2-5 overexpression treatments compared to a transient expression of mCherry negative control. (n = 3. ns = not significant compared to negative control, unpaired student's t-test).

(B), Fold enrichment of Tbx5 ChIP over input (Akerberg et al., 2019) centered on ATAC peaks present in the mESC cell state (2i + Day 0 ATAC peaks) versus cardiac-specific ATAC peaks [(Day 10 + Day 15) – (2i + Day 0)].

Resource Availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Alexander Ruthenburg (aruthenburg@uchicago.edu)

Materials availability

Plasmids generated in this study have been deposited to Addgene: pChantico-Firefly (163901), pChantico-Tdel-Firefly (163902), pChantico-pAS-Firefly (163903), pCxcl1-Firefly (163896), pCxcl1-Firefly-Chantico (163897), pCxcl1-Firefly-ocitnahC (163898), pCxcl1-Firefly-Chantico-Tdel (163900), pGL4-23-NegCtrl (163904), pCxcl1-Firefly-1-Tbx5-site (177832), pCxcl1-Firefly-2-Tbx5-sites (177833), pTbp-Firefly (177835), pTbp-Firefly-Chantico (177836), and pUC19-Chantico (170659).

Data and code availability

- Chromatin-enriched RNA-seq (GSE164126) and ATAC-seq (GSE183767) data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. This paper also analyzes existing, publicly available data; and these accession numbers for the datasets are listed in the key resources table. All data used for plotting in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental Model and Subject Details

Mammalian cell culture

HL-1 mouse cardiomyocytes were grown in Claycomb medium (Sigma 51800C) supplemented with 10% pre-screened fetal bovine serum (Sigma TMS-016), 2 mM L-glutamine (Sigma G7513), and 0.1 mM norepinephrine (Sigma A0937) according to the distributor's instructions (Claycomb et al., 1998). Confluent cell cultures were passaged between 1:3 and 1:6. The HL-1 cell line originated from an AT-1 subcutaneous tumor in a female mouse.

E14Tg2a.4 cells (BayGenomics, RRID:MMRRC_015890-UCD) were cultured in feeder-free conditions on 0.1% gelatin coated dishes. Cells were kept at subconfluence in DMEM, high

glucose (Invitrogen 11995) supplemented with 15% FBS (Gibco 25140), L-glutamine (Gibco 25030), non-essential amino acids (Gibco 11140-035), pen/strep (Gibco 15140), 0.1mM 2-mercaptoethanol, LIF (R&D Systems) and 2i (CHIR99021, Stemgent 04-0004, 3 μ M final concentration and PD0325901, Stemgent 04-0006, 1 μ M final concentration). E14Tg2a.4 are male. Cells were transfected 5h post plating using the Xfect mESC transfection reagent (Takara 631320) following the manufacturer's instructions, using 40 μ g of DNA per 10 cm dish. Cells were collected 48h post transfection and subjected to subsequent lysis with TRIzol reagent (Life Technologies 15596026).

A2Lox.cre mES cells were maintained and differentiated into cardiomyocytes as previously described (Kattman et al., 2011; Rowton et al., 2020). A2Lox.cre mES cells derive from E14Tg2a.4, which are male.

RAW264.7 Gamma NO(-) macrophages were grown in Dulbecco's Modified Eagle Medium: Nutrient Mixture F12 Ham (DMEM/F12) (Life Technologies 11320033) supplemented with 10% FBESsence (Seradigm 3100-500). RAW264.7 Gamma NO(-) are male.

HCT116 human colon carcinoma cells were grown in Dulbecco's Modified Eagle Medium: Nutrient Mixture F12 Ham (DMEM/F12) (Life Technologies 11320033) supplemented with 10% FBESsence (Seradigm 3100-500). HCT116 cells are male. Prior to transfection, HCT116 cells were washed with PBS and incubated for 30 minutes with Opti-MEM (Life Technologies 31985070).

All cell lines used in this study were purchased directly from the distributor and cultured within 1 year of receipt or validated by ATCC STR profiling. All cell cultures were grown in humidified incubators at 37°C with 5% CO₂ and tested negative for mycoplasma (InvivoGen #rep-pt1).

Method Details

Nuclear fractionation and RNA-seq

HL-1 cells from confluent T75 flasks were detached with TrypLE Express (Gibco #12605010), washed with PBS, and pelleted by centrifugation for 5 min at 400 x g. Cell pellets were immediately resuspended in Buffer A (10 mM HEPES-KOH pH 7.5, 10 mM KCl, 10% glycerol, 340 mM sucrose, 4mM MgCl₂, 1 mM DTT, 1 x Protease Inhibitor Cocktail (PIC) [1mM AEBSF, 0.8 μ M aprotinin, 20 μ M leupeptin, 15 μ M pepstatin A, 40 μ M bestatin, 15 μ M E-64]) and nuclear

fractions were separated as described previously (Werner and Ruthenburg, 2015; Werner et al., 2017). 1 μ L of a 1:10 dilution of ERCC RNA Spike-in standards (Life Technologies 4456740_3674355015) was added to each nuclear fraction prior to addition of TRIzol reagent (Life Technologies 15596026). RNA from each fraction was extracted as described previously (Werner and Ruthenburg, 2015; Werner et al., 2017) with on-column DNase digestion (Zymo R1017) and subsequent Ribo-zero Gold rRNA depletion (Illumina MRZG12324). Then, 100 ng rRNA-depleted RNA from each nuclear fraction was used with the NEBNext Ultra II Directional library kit (NEB E7760S) to prepare RNA-seq libraries. RNA-seq libraries of nuclear fractions (chromatin pellet extract and soluble nuclear extract) from three independent cultures were sequenced through the University of Chicago Genomics Core Facility on the Illumina HiSeq 4000 to obtain 50bp single reads.

RNA-seq and bioinformatic analysis

RNA-seq reads were aligned to the mm10 reference genome concatenated with ERCC standards using hisat2 version 2.1.0 (Kim et al., 2019). Transcriptomes of the chromatin pellet extract and soluble nuclear extract were assembled using cufflinks version 2.2.1 and cuffmerge version 2.2.1 (Trapnell et al., 2010) with gencode gene, gencode lincRNA, Tbx5-dependent ncRNA, and ERCC annotations as a guide. Novel transcripts were detected with Bedtools version 2.27.0 intersect commands (Quinlan and Hall, 2010). To keep shorter Tbx5-dependent lincRNA but remove transcripts which may represent incomplete transcript assembly, we lowered our previous length threshold from 1000 to 500 nt. Then, we concatenated novel transcripts onto the reference guide GTF for gene expression analysis with cuffdiff version 2.2.1 (Trapnell et al., 2010). Linear regression analysis of the ERCC spike-in RNA standard FPKM values was performed to determine a scalar correction constant to be applied to FPKM values in the chromatin pellet extract (Figure S1B). $\log_{10}$ total FPKM and $\log_{10}$ chromatin-enrichment were then calculated by: $\log_{10}(\text{corrected CPE FPKM} + \text{SNE FPKM})$ and $\log_{10}(\text{corrected CPE FPKM} / \text{SNE FPKM})$, respectively. Cell counting and linear regression analysis of $\log_{10}$ total FPKM vs $\log_{10}$ attomoles of spike-in RNA were used to estimate the number of copies per cell for each transcript. The Euler venn diagram for intersecting sets of transcripts was generated in R (eulerr). Bedtools version 2.2.7.0 closest commands (Quinlan and Hall, 2010) were used to identify the closest protein-coding genes. RNA-seq bedgraph files of nuclear fractions for each of the three independent cultures were generated with Deeptools version 3.5.0 bamCoverage (Ramírez et al., 2016), filtering for RNA strand and normalizing using FPKM (50 bp bin size). Then, awk was used to compute average FPKM, average FPKM plus standard deviation, and

average FPKM minus standard deviation for chromatin pellet extract and soluble nuclear extract libraries. These were used for visualization in genome browser view to show average FPKM $\pm$ standard deviation for the chromatin-enriched RNA-seq and soluble nuclear-enriched RNA-seq libraries on the forward and reverse strand.

RNA knockdown experiments in HL-1 cells

HL-1 cardiomyocytes were transfected 16 hours after passaging, when cells reached 60 to 80% confluence. About 30 minutes prior to transfections, HL-1 cell culture medium was replaced with supplemented Claycomb medium lacking norepinephrine. Then, HL-1 cells in 6-well plates were transfected with 15 μ L Lipofectamine 2000 reagent (Life Technologies 11668019) complexes with nucleic acid. At 8 to 16 hours post-transfection, cell culture medium was replaced with fully supplemented Claycomb medium. Antisense LNA GapmeRs (ASOs) were designed against Chantico through the QIAGEN design tool. HL-1 cells were transfected with ASOs at a 50 nM final concentration in 6-well plates. The Chantico-cleaving DNAzyme and OMegazyme 1 were designed to target the same region of Chantico as the most effective ASO. The catalytically dead version of the Chantico-cleaving OMegazyme 1 was designed with mutations in the catalytic core that are shown to ablate catalytic activity (Zaborowska et al., 2002). HL-1 cells were transfected with the OMegazyme at 100 nM final concentration. For Xrn2 knockdown, HL-1 cells were transfected with 40 nM ON-TARGETplus non-targeting siRNA #1 or Mouse Xrn2 siRNA SMARTpool (Horizon Discovery); and following 48 hours of incubation, cells were transfected with OMegazymes. For both ASO and OMegazyme knockdown RT-qPCR experiments, cells were lysed with TRIzol reagent (Life Technologies #15596026) 48 hours post-transfection.

HCT116 colorectal cancer cells were transfected with 5 μ L Lipofectamine 2000 reagent (Life Technologies 11668019) in complex with OMegazymes for 100 nM final OMegazyme concentration in 6-well plates. At 4 hours post-transfection, medium was replaced with DMEM/F12 supplemented with 10% FBESsence and additional 100 nM OMegazyme without transfection reagent to continue passive uptake of oligonucleotides through 48 hours, at which point cells were lysed with TRIzol reagent (Life Technologies 15596026).

In all RNA knockdown experiments, TRIzol lysates were used for subsequent RNA extraction, purification, and cDNA synthesis. All ASO, siRNA, and OMegazyme sequences are available in the Supplementary Table 1.

OMegazyme design and specificity analysis

Target regions of RNA were input into the PFRED v1.0 (Sciabola et al., 2021) software for design of 15 oligonucleotide ASO sequences. ASO designs from PFRED v1.0 were filtered for oligonucleotides that target an A-U or G-U at the center of the complementary RNA sequence, which is optimal for cleavage by the 10-23 DNAzyme (Cairns et al., 2003). Filtered ASOs were subsequently sorted by predicted efficacy scores and input into NCBI BLAST to select sequences with the fewest predicted off-target hits. Sequences with a perfect 15/15 match to multiple RNA targets were discarded. The first and last 7 nucleotides of the top remaining ASO sequences were used as the left and right homology arms respectively in a 10-23 DNAzyme-based design. The catalytic cores of these 10-23 DNAzymes (nt 8-22) were modified with 2'-O-Me groups at tolerated positions, leaving some 2'-deoxynucleotides (Figure 2A), as in the DH5E and DV15E4 designs of prior work (Schubert et al., 2003). Additionally, we modified every nucleotide of the left and right homology arms with 2'OMe groups, hypothesizing these substitutions would eliminate possible RNaseH cleavage of OMegazymes binding their target RNA (Sazani et al., 2001). 2'-O-methyl modifications in the catalytic core and flanking 2-3 nt were demonstrated to improve intracellular stability and improve target affinity without severely obstructing catalytic activity (Schubert et al., 2003). We refer to these fully modified nucleic acids as OMegazymes. PPIB-cleaving OMegazymes 1-10 and Chantico-cleaving OMegazymes 2 and 3 were designed through this pipeline and were synthesized by IDT. Top off-target hits with 1, 2, or 3 mismatches in the complementary targeting sequence for several OMegazymes were measured by RT-qPCR of cDNA from the same experiments that demonstrated effective PPIB mRNA knockdown to determine the tolerance for mismatches in RNA targeting and specificity of individual OMegazymes.

Chromatin immunoprecipitation

ChIP was carried out with 20 million cells. Following appropriate treatment, growth media was replaced with DMEM/F12 containing 1% formaldehyde (Pierce #28906) and incubated at room temperature for 10 mins. Excess formaldehyde was quenched with 0.75 M Tris (pH 7.5) for 5 mins at room temperature. After washing twice with 5 ml PBS, cells were scraped and transferred to 15 ml LoBind conical tubes (Eppendorf). Cells were harvested by centrifugation at 500g for 5 mins. The cell pellet was washed twice with PBS. Cells were lysed by resuspending the pellet in 10 ml lysis buffer I (140 mM NaCl, 1 mM EDTA, 50 mM HEPES-KOH pH 8.0, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, 1X Protease Inhibitor Cocktail (PIC; 1mM AEBSF,

0.8 μ M aprotinin, 20 μ M leupeptin, 15 μ M pepstatin A, 40 μ M bestatin, 15 μ M E-64)) for 10 mins at 4°C with constant mixing on nutator. The nuclei were harvested by centrifugation at 500g for 5 mins. Nuclei were then resuspended in 10 ml lysis buffer II (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1X PIC) and incubated for 10 mins at 4°C with constant mixing on nutator. The lysed nuclei, collected by spinning at 1200g for 10 mins, were resuspended in 1 ml sonication buffer (150 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.1% SDS, 1X PIC). The chromatin was fragmented using Covaris S220 (duty cycle: 20%; intensity: 8; cycle/burst: 200) for 8 mins. Sonicated chromatin was centrifuge clarified at 10,000g for 10 mins and diluted 1:1 with dilution buffer (280 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl pH 8.0, 1.0% Triton X-100, 0.1% SDS, 1X PIC). The chromatin was pre-cleared with 50 μ l Protein G Dynabeads for 1 hr at 4°C. Equal amount of chromatin used for immunoprecipitation and “beads only” control; 10 % of chromatin was set aside as input. For Pol II ChIP, samples were incubated with 4 μ g of Pol II antibody (Active Motif, 4H8) at 4°C overnight. Antibody-chromatin complexes were immobilized on Protein G Dynabeads for 16 hrs at 4°C. These were then washed sequentially with wash buffer 1 (140 mM NaCl, 1 mM EDTA, 50 mM Tris pH 8.0, 1.0% Triton X-100, 0.1% SDS, 1X PIC), wash buffer 2 (250 mM LiCl, 1 mM EDTA, 10 mM Tris pH 8.0, 0.5% NP-40, 0.1% SDS, 1X PIC) and 1X TE. Immunoprecipitated chromatin was eluted by incubating beads with elution buffer (50 mM Tris pH 8.0, 10 mM EDTA, 1% SDS) at 65 °C for 30 mins. RNA was removed by 100 μ g RNaseA (Qiagen, # 19101) treatment at 37°C for 1 hr. Crosslinks were reversed with overnight digestion with 2 mg/ml proteinase K (Invitrogen, # 25530015) at 65°C in the presence of 1 % SDS. The ChIP template DNA was then extracted by two rounds of phenol-chloroform-isoamyl alcohol extraction and one round of chloroform:isoamyl alcohol extraction to remove phenol. DNA was precipitated by adding 2.5 volumes of ethanol in the presence of sodium acetate (300 mM, pH 5.2) and glycogen (15 μ g/ml), and then quantified by qPCR.

RT-qPCR

For RNA samples, chloroform extraction was performed on TRIzol lysates followed by in-column DNase digestion (Zymo R1017) for RNA purification. 1 μ g of purified RNA was used in cDNA synthesis with MMLV HP Reverse Transcriptase (Lucigen RT80125K) and 100 ng random hexamers (IDT). RT-PCR reactions contained a mixture of cDNA (1.75% of total cDNA sample for human, or 7% of cDNA for mouse), 250 nM specific primers (screened for single peaks by thermal melt curve analysis and >85% amplification efficiency by serial dilution), and 1X PowerUp SYBR Green Master Mix (Applied Biosystems). These reactions (20 μ L human; 15 μ L

mouse) were measured using a Bio-Rad CFX96 or CFX384 system. The average of three technical measurements for was taken for each independent replicate. Each independent replicate was normalized by the ΔC_t method (Schmittgen and Livak, 2008) to eIF3s5 mRNA (mouse) or GAPDH mRNA (human) due to the stability of these mRNA (Kouadjo et al., 2007) relative to spike-in RNA standards (Figure S2C) and abundance relative to target amplicons. Relative abundance was calculated by $2^{\Delta C_t}$. The average relative abundance among independent replicates of the control treatment were used to calculate fold-change over control. When displayed, bars represent the average fold change over control among independent replicates and error bars represent the standard deviation of mean of the individual replicates. Individual replicates are displayed as circles overlaid on the bars for the corresponding condition.

For ChIP samples, qPCR reactions (20 μ L) contained a mixture of DNA, 200 nM specific primers, and 1X PowerUp SYBR Green Master Mix (Applied Biosystems). These reactions were measured using a CFX384 system. Each sample was measured in technical duplicates and normalized to the primer pair targeting the *GAPDH* promoter, accounting for primer pair amplification efficiencies. Primer pair-efficiencies were calculated from a standard curve obtained from serial dilution of genomic DNA.

All RT-qPCR primer sequences are available in the Supplementary Table 1.

Cardiac differentiation of mESCs and ATAC-seq

A2Lox.cre mES cells were maintained and differentiated into cardiomyocytes as previously described (Kattman et al., 2011; Rowton et al., 2020). RNA was harvested from two replicate samples every 24 hours of the differentiation from the mESC (2i) stage to D8, in addition to D10 and D15 (cardiomyocyte stage). The assay for transposase-accessible chromatin (ATAC-seq) was performed as previously described (Buenrostro et al., 2013). Briefly, 150,000 cells were collected from two replicate samples every 24 hours of the differentiation from the 2i stage to D8, in addition to D10 and D15. Samples were then lysed, and transposition was carried out at 37C for 30 minutes with Illumina's Nextera DNA Library Prep kit (Illumina 15028212). Libraries were generated from transposed DNA and sequenced on an Illumina HiSeq4000 instrument in the University of Chicago's Genomic Facility.

Adapter sequences were clipped from reads using Cutadapt version 2.10 (Martin, 2011) (cutadapt -a CTGTCTCTTATACACATCT -A CTGTCTCTTATACACATCT -j 4 -m 10) , then aligned to UCSC mouse genome mm10 with bowtie2 version 2.3.2 with default settings. Low quality reads ($q < 30$) and reads mapping to blacklisted regions were removed using samtools version 1.5 (Danecek et al., 2021). Duplicated reads were marked and removed with picard tool MarkDuplicates version 2.8.1 (Broad Institute, 2019). Accessible regions were identified using MACS2 version 2.1.1 (Zhang et al., 2008) grouping replicates together. (macs2 callpeak -f BAMPE -g mm --keep-dup all -B --SPMR). Fold enrichment tracks were generated with MACS2 bdgcmp (-m FE). Bedgraph files from MACS2 were converted to bigwig format with UCSC tool bedGraphToBigWig version 4 (Kent et al., 2010).

ATAC-seq peaks from the 2i and Day 0 stages were combined with bedtools to yield a list of mESC ATAC-seq peaks. These mESC ATAC-seq peaks were subtracted from a list of ATAC-seq peaks combined from Day 10 and Day 15 of the differentiation to yield a list of cardiac-specific ATAC-seq peaks. Tbx5 fold enrichment from published ChIP-seq data (Akerberg et al., 2019) was displayed over the mESC and cardiac-specific ATAC-seq peak centers +/- 500 bp with Deeptools version 3.5.0 (Ramírez et al., 2016).

Luciferase assays

HL-1 cardiomyocytes were co-transfected in 6-well plates with a *Firefly* luciferase vector (0.7 pmol) and a *Renilla* luciferase vector (50 ng). *Firefly* luciferase vectors were engineered to test various conditions and are available through Addgene. Cell lysates were collected 48 hours post-transfection. Lysates were stored at -20°C or used immediately in a Dual-Glo luciferase assay (Promega #E2920) with the SpectraMax ID5 microplate reader. Relative luciferase activity was calculated by: *Firefly* luciferase activity / *Renilla* luciferase activity. Then, relative luciferase data for each transfection condition were normalized to a concurrent negative control luciferase vector (minimal promoter driving *Firefly* luciferase) transfection to determine fold change in luciferase activity over the negative control.

ELISA for secreted Cxcl1

At 48 hours post-transfection, HL-1 cells in a 6-well culture plate were washed twice with PBS and given 1 mL of supplemented Claycomb medium. Cells were returned to incubation to allow secretion for the following time points: 15 minutes, 1 hour, and 12 hours. Then, media was collected and used to measure secreted Cxcl1 protein on an ELISA plate according to the

KC/Cxcl1 ELISA kit manufacturer's instructions (Sigma-Aldrich #RAB0117). Absorbance values at 450 nm were measured with a Tecan plate reader. A dilution series of purified Cxcl1 protein (included in kit) was used to measure a standard curve. Regression analysis was used to calculate total picograms of secreted Cxcl1 protein.

Boyden chamber transwell migration assay

Boyden chamber transwell migration assays were set-up according to the cell migration assay kit manufacturer's instructions (Sigma #ECM507), with RAW264.7 macrophages seeded in the upper chamber (30,000 per chamber) and cardiomyocyte cell culture supernatants from 1 hour of secretion placed in the lower chamber. Migrated cells were detached and lysed with provided solutions in the kit, followed by quantification of nucleic acid by measuring CyQUANT GR dye (Molecular Probes) fluorescence intensity with a Tecan plate reader. Signal from supplemented Claycomb medium (unused for cell culture) in the lower chamber was subtracted from the signal for tested conditions to account for unstimulated migratory events; then, fold change over a mock transfection control was calculated.

Quantification and Statistical Analysis

RNA-seq was performed with three independent biological replicates. RT-qPCR, ChIP qPCR, luciferase assays, ELISA, and Boyden chamber transwell migration assays are reported as the average of independent transfections $\pm$ standard deviation with individual transfections overlaid as circles on bar graphs. P values are calculated by unpaired student's t-test unless specified otherwise (ns = not significant, * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$, **** = $p < 0.00005$). All RT-qPCR primer sets were tested with a thermal melt curve to ensure single peaks, supporting their specificity for the target molecule. Violin-boxplots in **Figure 1B,C** were generated in R with ggplot2; boxplots represent the median and interquartile range and p -values were calculated here with a Wilcoxon signed rank test.

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
RNA Pol II antibody	Active Motif	Clone 4H8; RRID: AB_2732926
Chemicals, peptides, and recombinant proteins		
Lipofectamine 2000	Life Technologies	11668019
Xfect mESC transfection reagent	Takara	631320
ERCC RNA Spike-in standards	Life Technologies	4456740_3674355015
Pierce™ 16% Formaldehyde (w/v), Methanol-free	Thermo Scientific	28906
TRIzol reagent	Life Technologies	15596026
TURBO DNase	Invitrogen	AM2239
Protein G Dynabeads	Invitrogen	10004D
RNaseA	Qiagen	19101
Proteinase K	Invitrogen	25530015
MMLV HP Reverse Transcriptase	Lucigen	RT80125K
Critical commercial assays		
KC/Cxcl1 ELISA kit	Sigma-Aldrich	RAB0117
QCM Chemotaxis Cell Migration Assay, 24-well (5 µm)	Sigma-Aldrich	ECM507
RNA Clean & Concentrator Kit	Zymo	R1017
Ribo-zero Gold rRNA depletion kit	Illumina	MRZG12324
NEBNext Ultra II Directional library kit	NEB	E7760S
Nextera DNA Library Prep kit	Illumina	15028212
Dual-Glo Luciferase Assay kit	Promega	E2920
PowerUp SYBR Green Master Mix	Applied Biosystems	A25741
Deposited data		
cheRNA-seq datasets in HL-1 mouse cardiomyocytes	This paper	GEO: GSE164126
ATAC-seq datasets in A2Lox.cre mESCs to mESC-CMs	This paper	GEO: GSE183767
Tbx5-dependent RNA-seq datasets in dissected mouse atrial tissue	Yang et al., 2017	GEO: GSE129503
ChIP-seq datasets in HL-1 mouse cardiomyocytes	He et al., 2011	GEO: GSE21529
ChIP-seq datasets in dissected heart tissue	Akerberg et al., 2019	GEO: GSE124008
Mouse ENCODE consortium sequencing data	Stamatoyannopoulos et al., 2012	https://www.encodeproject.org/
Experimental models: Cell lines		
Mouse: HL-1 cardiac muscle cell line	Sigma-Aldrich	SCC065
Human: HCT116 colon carcinoma cell line	Sigma-Aldrich	91091005-1VL
Mouse: RAW264.7 Gamma No(-) macrophage cell line	ATCC	50-188-594FP
Mouse: A2Lox.cre embryonic stem cell line	This paper	iTBX5 mESC
Mouse: E14Tg2a.4 embryonic stem cell line	BayGenomics	RRID: MMRRC_015890-UCD
Oligonucleotides		
Primers – See Table S1	This paper	Table S1
DNAzymes – See Table S1	This paper	Table S1
Omegazymes – See Table S1	This paper	Table S1
Antisense oligonucleotides – See Table S1	This paper	Table S1

siRNA – See Table S1	This paper	Table S1
NEBNext Multiplex Oligos for Illumina (Index Primers Set 1)	NEB	E7335S
ReadyMade Random hexamers	IDT	51-01-18-27
Recombinant DNA		
pUC19	Addgene	5005
pChantico-Firefly	This paper	Addgene: 163901
pChantico-Tdel-Firefly	This paper	Addgene: 163902
pChantico-pAS-Firefly	This paper	Addgene: 163903
pCycl1-Firefly	This paper	Addgene: 163896
pCycl1-Firefly-Chantico	This paper	Addgene: 163897
pCycl1-Firefly-ocitnahC	This paper	Addgene: 163898
pCycl1-Firefly-Chantico-Tdel	This paper	Addgene: 163900
pGL4-23-NegCtrl	This paper	Addgene: 163904
pCycl1-Firefly-1-Tbx5-site	This paper	Addgene: 177832
pCycl1-Firefly-2-Tbx5-sites	This paper	Addgene: 177833
pTbp-Firefly	This paper	Addgene: 177835
pTbp-Firefly-Chantico	This paper	Addgene: 177836
pUC19-Chantico	This paper	Addgene: 170659
Software and algorithms		
R version 4.0.1 – “See Things Now”	The R Foundation for Statistical Computing	https://www.r-project.org/
Hisat2 version 2.1.0	Kim, et al., 2019	https://github.com/DaehwanKimLab/hisat2
Cufflinks version 2.2.1	Trapnell et al., 2010	https://github.com/cole-trapnell-lab/cufflinks
Bedtools version 2.27.0	Quinlan and Hall, 2010.	https://github.com/arq5x/bedtools2/releases/tag/v2.27.0
Deeptools version 3.5.0	Ramírez et al., 2016	https://github.com/deeptools/deepTools
Samtools version 1.5	Danecek et al., 2021	https://github.com/samtools
Cutadapt version 2.1	Martin, 2019	https://github.com/marcelm/cutadapt
Bowtie2 version 2.3.2	Langmead and Salzberg, 2012	https://github.com/BenLangmead/bowtie2
MACS2	Zheng et al., 2008	https://github.com/macs3-project/MACS
Picard	Picard toolkit. Broad Institute. 2019	https://github.com/broadinstitute/picard
UCSC tools	Kent et al., 2010	https://genome.ucsc.edu/util.html

References

- Abudayyeh, O.O., Gootenberg, J.S., Essletzbichler, P., Han, S., Joung, J., Belanto, J.J., Verdine, V., Cox, D.B.T., Kellner, M.J., Regev, A., et al. (2017). RNA targeting with CRISPR–Cas13. *Nature* 550, 280–284.
- Akerberg, B.N., Gu, F., VanDusen, N.J., Zhang, X., Dong, R., Li, K., Zhang, B., Zhou, B., Sethi, I., Ma, Q., et al. (2019). A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers. *Nature Communications* 10, 4907.
- Allou, L., Balzano, S., Magg, A., Quinodoz, M., Royer-Bertrand, B., Schöpflin, R., Chan, W.-L., Speck-Martins, C.E., Carvalho, D.R., Farage, L., et al. (2021). Non-coding deletions identify Maenli lncRNA as a limb-specific En1 regulator. *Nature* 592, 93–98.
- Anderson, K.M., Anderson, D.M., McAnally, J.R., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016). Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. *Nature* 539, 433–436.
- Bachmaier, K., Toya, S., and Malik, A.B. (2014). Therapeutic Administration of the Chemokine CXCL1/KC Abrogates Autoimmune Inflammatory Heart Disease. *PLOS ONE* 9, e89647.
- Bajpai, G., Schneider, C., Wong, N., Bredemeyer, A., Hulsmans, M., Nahrendorf, M., Epelman, S., Kreisel, D., Liu, Y., Itoh, A., et al. (2018). The Human Heart Contains Distinct Macrophage Subsets with Divergent Origins and Functions. *Nat Med* 24, 1234–1245.
- Ballarino, M., Cipriano, A., Tita, R., Santini, T., Desideri, F., Morlando, M., Colantoni, A., Carrieri, C., Nicoletti, C., Musarò, A., et al. (2018). Deficiency in the nuclear long noncoding RNA Charmc causes myogenic defects and heart remodeling in mice. *The EMBO Journal* 37, e99697.
- Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299–308.
- Barin, J.G., Rose, N.R., and Čiháková, D. (2012). Macrophage diversity in cardiac inflammation: a review. *Immunobiology* 217, 468–475.
- Bassett, A.R., Akhtar, A., Barlow, D.P., Bird, A.P., Brockdorff, N., Duboule, D., Ephrussi, A., Ferguson-Smith, A.C., Gingeras, T.R., Haerty, W., et al. (2014). Considerations when investigating lncRNA function in vivo. *Elife* 3, e03058.
- Boisvert, W.A., Rose, D.M., Johnson, K.A., Fuentes, M.E., Lira, S.A., Curtiss, L.K., and Terkeltaub, R.A. (2006). Up-Regulated Expression of the CXCR2 Ligand KC/GRO-α in Atherosclerotic Lesions Plays a Central Role in Macrophage Accumulation and Lesion Progression. *The American Journal of Pathology* 168, 1385–1395.
- Bose, D.A., Donahue, G., Reinberg, D., Shiekhata, R., Bonasio, R., and Berger, S.L. (2017). RNA Binding to CBP Stimulates Histone Acetylation and Transcription. *Cell* 168, 135-149.e22.
- Brown, J.A., Bulkley, D., Wang, J., Valenstein, M.L., Yario, T.A., Steitz, T.A., and Steitz, J.A. (2014). Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix. *Nat Struct Mol Biol* 21, 633–640.

- Bruneau, B.G., Logan, M., Davis, N., Levi, T., Tabin, C.J., Seidman, J.G., and Seidman, C.E. (1999). Chamber-Specific Cardiac Expression of Tbx5 and Heart Defects in Holt–Oram Syndrome. *Developmental Biology* 211, 100–108.
- Bruneau, B.G., Nemer, G., Schmitt, J.P., Charron, F., Robitaille, L., Caron, S., Conner, D.A., Gessler, M., Nemer, M., Seidman, C.E., et al. (2001). A Murine Model of Holt-Oram Syndrome Defines Roles of the T-Box Transcription Factor Tbx5 in Cardiogenesis and Disease. *Cell* 106, 709–721.
- Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10, 1213–1218.
- Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., et al. (2005). The transcriptional landscape of the mammalian genome. *Science* 309, 1559–1563.
- Cho, S.W., Xu, J., Sun, R., Mumbach, M.R., Carter, A.C., Chen, Y.G., Yost, K.E., Kim, J., He, J., Nevins, S.A., et al. (2018). Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element. *Cell* 173, 1398-1412.e22.
- Creamer, K.M., Kolpa, H.J., and Lawrence, J.B. (2021). Nascent RNA scaffolds contribute to chromosome territory architecture and counter chromatin compaction. *Molecular Cell* 81, 3509-3525.e5.
- Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, giab008.
- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789.
- Devonshire, A.S., Elaswarapu, R., and Foy, C.A. (2010). Evaluation of external RNA controls for the standardisation of gene expression biomarker measurements. *BMC Genomics* 11, 662.
- Di Ruscio, A., Ebralidze, A.K., Benoukraf, T., Amabile, G., Goff, L.A., Terragni, J., Figueroa, M.E., De Figueiredo Pontes, L.L., Alberich-Jorda, M., Zhang, P., et al. (2013). DNMT1-interacting RNAs block gene-specific DNA methylation. *Nature* 503, 371–376.
- Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. *Nature* 489, 101–108.
- Donis-Keller, H. (1979). Site specific enzymatic cleavage of RNA. *Nucleic Acids Res* 7, 179–192.
- Ebisuya, M., Yamamoto, T., Nakajima, M., and Nishida, E. (2008). Ripples from neighbouring transcription. *Nat Cell Biol* 10, 1106–1113.

- Engreitz, J.M., Haines, J.E., Perez, E.M., Munson, G., Chen, J., Kane, M., McDonel, P.E., Guttman, M., and Lander, E.S. (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* 539, 452–455.
- Erokhin, M., Elizar'ev, P., Parshikov, A., Schedl, P., Georgiev, P., and Chetverina, D. (2015). Transcriptional read-through is not sufficient to induce an epigenetic switch in the silencing activity of Polycomb response elements. *Proc Natl Acad Sci USA* 112, 14930–14935.
- Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 47, D766–D773.
- Fulco, C.P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S.R., Perez, E.M., Kane, M., Cleary, B., Lander, E.S., and Engreitz, J.M. (2016). Systematic mapping of functional enhancer–promoter connections with CRISPR interference. *Science* 354, 769–773.
- Ghosh, T.K. (2001). Characterization of the TBX5 binding site and analysis of mutations that cause Holt-Oram syndrome. *Human Molecular Genetics* 10, 1983–1994.
- Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159, 647–661.
- Gorbovytska, V., Kim, S.-K., Kuybu, F., Götze, M., Um, D., Kang, K., Pittroff, A., Schneider, L.-M., Leitner, A., Kim, T.-K., et al. (2021). Enhancer RNAs stimulate Pol II pause release by harnessing multivalent interactions to NELF. *BioRxiv* <https://doi.org/10.1101/2021.04.25.441328>
- Groff, A.F., Sanchez-Gomez, D.B., Soruco, M.M.L., Gerhardinger, C., Barutcu, A.R., Li, E., Elcavage, L., Plana, O., Sanchez, L.V., Lee, J.C., et al. (2016). In Vivo Characterization of Linc-p21 Reveals Functional cis-Regulatory DNA Elements. *Cell Reports* 16, 2178–2186.
- Groff, A.F., Barutcu, A.R., Lewandowski, J.P., and Rinn, J.L. (2018). Enhancers in the Peril lincRNA locus regulate distant but not local genes. *Genome Biol* 19, 219.
- Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.-C., Hung, T., Argani, P., Rinn, J.L., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464, 1071–1076.
- Hacisuleyman, E., Goff, L.A., Trapnell, C., Williams, A., Henao-Mejia, J., Sun, L., McClanahan, P., Hendrickson, D.G., Sauvageau, M., Kelley, D.R., et al. (2014). Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat Struct Mol Biol* 21, 198–206.
- He, A., Kong, S.W., Ma, Q., and Pu, W.T. (2011). Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proceedings of the National Academy of Sciences* 108, 5632–5637.
- Holoch, D., Wassef, M., Lövkist, C., Zielinski, D., Aflaki, S., Lombard, B., Héry, T., Loew, D., Howard, M., and Margueron, R. (2021). A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in mammals. *Nat Genet.* 53,1686-1697.

- Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R., and Darzacq, X. (2020). Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. *Molecular Cell* 78, 539-553.
- Hua, P., Badat, M., Hanssen, L.L.P., Hentges, L.D., Crump, N., Downes, D.J., Jeziorska, D.M., Oudelaar, A.M., Schwessinger, R., Taylor, S., et al. (2021). Defining genome architecture at base-pair resolution. *Nature* 595, 125–129.
- Hulsmans, M., Clauss, S., Xiao, L., Aguirre, A.D., King, K.R., Hanley, A., Hucker, W.J., Wülfers, E.M., Seemann, G., Courties, G., et al. (2017). Macrophages Facilitate Electrical Conduction in the Heart. *Cell* 169, 510-522.e20.
- Ibrahim, H., Wilusz, J., and Wilusz, C.J. (2008). RNA recognition by 3'-to-5' exonucleases: The substrate perspective. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1779, 256–265.
- Isoda, T., Moore, A.J., He, Z., Chandra, V., Aida, M., Denholtz, M., Piet van Hamburg, J., Fisch, K.M., Chang, A.N., Fahl, S.P., et al. (2017). Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. *Cell* 171, 103-119.e18.
- Iyer, M.K., Niknafs, Y.S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., Barrette, T.R., Prensner, J.R., Evans, J.R., Zhao, S., et al. (2015). The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet* 47, 199–208.
- Jinek, M., Coyle, S.M., and Doudna, J.A. (2011). Coupled 5' Nucleotide Recognition and Processivity in Xrn1-Mediated mRNA Decay. *Molecular Cell* 41, 600–608.
- Joung, J., Engreitz, J.M., Konermann, S., Abudayyeh, O.O., Verdine, V.K., Aguet, F., Gootenberg, J.S., Sanjana, N.E., Wright, J.B., Fulco, C.P., et al. (2017). Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* 548, 343–346.
- Karouzakis, E., Trenkmann, M., Gay, R.E., Michel, B.A., Gay, S., and Neidhart, M. (2014). Epigenome Analysis Reveals TBX5 as a Novel Transcription Factor Involved in the Activation of Rheumatoid Arthritis Synovial Fibroblasts. *The Journal of Immunology* 193, 4945–4951.
- Kattman, S.J., Witty, A.D., Gagliardi, M., Dubois, N.C., Niapour, M., Hotta, A., Ellis, J., and Keller, G. (2011). Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell Stem Cell* 8, 228–240.
- Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* 26, 2204–2207.
- Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37, 907–915.
- Kim, T.-K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187.

- Klattenhoff, C.A., Scheuermann, J.C., Surface, L.E., Bradley, R.K., Fields, P.A., Steinhauser, M.L., Ding, H., Butty, V.L., Torrey, L., Haas, S., et al. (2013). Braveheart, a Long Noncoding RNA Required for Cardiovascular Lineage Commitment. *Cell* 152, 570–583.
- Lagarde, J., Uszczyńska-Ratajczak, B., Carbonell, S., Pérez-Lluch, S., Abad, A., Davis, C., Gingeras, T.R., Frankish, A., Harrow, J., Guigo, R., et al. (2017). High-throughput annotation of full-length long noncoding RNAs with capture long-read sequencing. *Nat Genet* 49, 1731–1740.
- Lai, F., Orom, U.A., Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A., and Shiekhattar, R. (2013). Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* 494, 497–501.
- Lai, F., Damle, S.S., Ling, K.K., and Rigo, F. (2020). Directed RNase H Cleavage of Nascent Transcripts Causes Transcription Termination. *Molecular Cell* 77, 1032-1043.e4.
- Lam, M.T.Y., Li, W., Rosenfeld, M.G., and Glass, C.K. (2014). Enhancer RNAs and regulated transcriptional programs. *Trends Biochem Sci* 39, 170–182.
- Latos, P.A., Pauler, F.M., Koerner, M.V., Şenergin, H.B., Hudson, Q.J., Stocsits, R.R., Allhoff, W., Stricker, S.H., Klement, R.M., Warczak, K.E., et al. (2012). Airn Transcriptional Overlap, But Not Its lncRNA Products, Induces Imprinted Igf2r Silencing. *Science* 338, 1469–1472.
- Lee, J.-S., and Mendell, J.T. (2020). Antisense-Mediated Transcript Knockdown Triggers Premature Transcription Termination. *Molecular Cell* 77, 1044-1054.e3.
- Lewandowski, J.P., Lee, J.C., Hwang, T., Sunwoo, H., Goldstein, J.M., Groff, A.F., Chang, N.P., Mallard, W., Williams, A., Henao-Meija, J., et al. (2019). The Firre locus produces a trans-acting RNA molecule that functions in hematopoiesis. *Nat Commun* 10, 5137.
- Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498, 516–520.
- Liang, X., Vickers, T.A., Guo, S., and Crooke, S.T. (2011). Efficient and specific knockdown of small non-coding RNAs in mammalian cells and in mice. *Nucleic Acids Res* 39, e13.
- Lubas, M., Andersen, P.R., Schein, A., Dziembowski, A., Kudla, G., and Jensen, T.H. (2015). The Human Nuclear Exosome Targeting Complex Is Loaded onto Newly Synthesized RNA to Direct Early Ribonucleolysis. *Cell Reports* 10, 178–192.
- Maamar, H., Cabili, M.N., Rinn, J., and Raj, A. (2013). linc-HOXA1 is a noncoding RNA that represses Hoxa1 transcription in cis. *Genes Dev.* 27, 1260–1271.
- Maranon, D.G., and Wilusz, J. (2020). Mind the Gapmer: Implications of Co-transcriptional Cleavage by Antisense Oligonucleotides. *Molecular Cell* 77, 932–933.
- Marquardt, S., Raitskin, O., Wu, Z., Liu, F., Sun, Q., and Dean, C. (2014). Functional Consequences of Splicing of the Antisense Transcript COOLAIR on FLC Transcription. *Molecular Cell* 54, 156–165.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* 17, 10–12.

- Micheletti, R., Plaisance, I., Abraham, B.J., Sarre, A., Ting, C.-C., Alexanian, M., Maric, D., Maison, D., Nemir, M., Young, R.A., et al. (2017). The long noncoding RNA Wisper controls cardiac fibrosis and remodeling. *Sci. Transl. Med.* 9, eaai9118.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5, 621–628.
- Moskowitz, I.P.G., Pizard, A., Patel, V.V., Bruneau, B.G., Kim, J.B., Kupersmidt, S., Roden, D., Berul, C.I., Seidman, C.E., and Seidman, J.G. (2004). The T-Box transcription factor Tbx5 is required for the patterning and maturation of the murine cardiac conduction system. *Development* 131, 4107–4116.
- Nadadur, R.D., Broman, M.T., Boukens, B., Mazurek, S.R., Yang, X., van den Boogaard, M., Bekeny, J., Gadek, M., Ward, T., Zhang, M., et al. (2016). Pitx2 modulates a Tbx5 -dependent gene regulatory network to maintain atrial rhythm. *Sci. Transl. Med.* 8, 354ra115-354ra115.
- Nagata, T., Dwyer, C.A., Yoshida-Tanaka, K., Ihara, K., Ohyagi, M., Kaburagi, H., Miyata, H., Ebihara, S., Yoshioka, K., Ishii, T., et al. (2021). Cholesterol-functionalized DNA/RNA heteroduplexes cross the blood–brain barrier and knock down genes in the rodent CNS. *Nat Biotechnol.* 39, 1529-1536.
- Nicolás-Ávila, J.A., Lechuga-Vieco, A.V., Esteban-Martínez, L., Sánchez-Díaz, M., Díaz-García, E., Santiago, D.J., Rubio-Ponce, A., Li, J.L., Balachander, A., Quintana, J.A., et al. (2020). A Network of Macrophages Supports Mitochondrial Homeostasis in the Heart. *Cell* 183, 94-109.
- Noguchi, S., Arakawa, T., Fukuda, S., Furuno, M., Hasegawa, A., Hori, F., Ishikawa-Kato, S., Kaida, K., Kaiho, A., Kanamori-Katayama, M., et al. (2017). FANTOM5 CAGE profiles of human and mouse samples. *Sci Data* 4, 1–10.
- Ounzain, S., Micheletti, R., Arnan, C., Plaisance, I., Cecchi, D., Schroen, B., Reverter, F., Alexanian, M., Gonzales, C., Ng, S.Y., et al. (2015). CARMEN, a human super enhancer-associated long noncoding RNA controlling cardiac specification, differentiation and homeostasis. *Journal of Molecular and Cellular Cardiology* 89, 98–112.
- Özcan, A., Krajewski, R., Ioannidi, E., Lee, B., Gardner, A., Makarova, K.S., Koonin, E.V., Abudayyeh, O.O., and Gootenberg, J.S. (2021). Programmable RNA targeting with the single-protein CRISPR effector Cas7-11. *Nature* 597, 720–725.
- Paralkar, V.R., Taborda, C.C., Huang, P., Yao, Y., Kossenkov, A.V., Prasad, R., Luan, J., Davies, J.O.J., Hughes, J.R., Hardison, R.C., et al. (2016). Unlinking an lncRNA from Its Associated cis Element. *Molecular Cell* 62, 104–110.
- Petruk, S., Sedkov, Y., Riley, K.M., Hodgson, J., Schweisguth, F., Hirose, S., Jaynes, J.B., Brock, H.W., and Mazo, A. (2006). Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx in cis by transcriptional interference. *Cell* 127, 1209–1221.
- Plenge, R.M., Hendrich, B.D., Schwartz, C., Arena, J.F., Naumova, A., Sapienza, C., Winter, R.M., and Willard, H.F. (1997). A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. *Nat Genet* 17, 353–356.

- Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., and Jensen, T.H. (2008). RNA Exosome Depletion Reveals Transcription Upstream of Active Human Promoters. *Science* 322, 1851–1854.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842.
- Quinodoz, S.A., Jachowicz, J.W., Bhat, P., Ollikainen, N., Banerjee, A.K., Goronzy, I.N., Blanco, M.R., Chovanec, P., Chow, A., Markaki, Y., et al. (2021). RNA promotes the formation of spatial compartments in the nucleus. *Cell* 184, 5775–5790.
- Rahnamoun, H., Lee, J., Sun, Z., Lu, H., Ramsey, K.M., Komives, E.A., and Lauberth, S.M. (2018). RNAs interact with BRD4 to promote enhanced chromatin engagement and transcription activation. *Nat Struct Mol Biol* 25, 687–697.
- Ramilowski, J.A., Yip, C.W., Agrawal, S., Chang, J.-C., Ciani, Y., Kulakovskiy, I.V., Mendez, M., Ooi, J.L.C., Ouyang, J.F., Parkinson, N., et al. (2020). Functional annotation of human long noncoding RNAs via molecular phenotyping. *Genome Res.* 30, 1060–1072.
- Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44, W160–165.
- Rinn, J.L., and Chang, H.Y. (2020). Long Noncoding RNAs: Molecular Modalities to Organismal Functions. *Annu Rev Biochem* 89, 283–308.
- Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Bruggmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., et al. (2007). Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* 129, 1311–1323.
- Ritter, N., Ali, T., Kopitchinski, N., Schuster, P., Beisaw, A., Hendrix, D.A., Schulz, M.H., Müller-McNicoll, M., Dimmeler, S., and Grote, P. (2019). The lncRNA Locus Handsdown Regulates Cardiac Gene Programs and Is Essential for Early Mouse Development. *Developmental Cell* 50, 644–657.e8.
- Rowton, M., Perez-Cervantes, C., Rydeen, A., Hur, S., Jacobs-Li, J., Deng, N., Lu, E., Guzzetta, A., Steimle, J.D., Hoffmann, A., et al. (2020). Hedgehog signaling activates a heterochronic gene regulatory network to control differentiation timing across lineages. *BioRxiv*. <https://doi.org/10.1101/270751>
- Ruiz-Orera, J., and Albà, M.M. (2019). Conserved regions in long non-coding RNAs contain abundant translation and protein–RNA interaction signatures. *NAR Genomics and Bioinformatics* 1, e2.
- Salamon, I., Serio, S., Bianco, S., Pagiatakis, C., Crasto, S., Chiariello, A.M., Conte, M., Cattaneo, P., Fiorillo, L., Felicetta, A., et al. (2020). Divergent Transcription of the Nkx2-5 Locus Generates Two Enhancer RNAs with Opposing Functions. *IScience* 23, 101539.
- Santoro, S.W., and Joyce, G.F. (1997). A general purpose RNA-cleaving DNA enzyme. *PNAS* 94, 4262–4266.

- Santoro, S.W., and Joyce, G.F. (1998). Mechanism and Utility of an RNA-Cleaving DNA Enzyme †. *Biochemistry* 37, 13330–13342.
- Sauvageau, M., Goff, L.A., Lodato, S., Bonev, B., Groff, A.F., Gerhardinger, C., Sanchez-Gomez, D.B., Hacisuleyman, E., Li, E., Spence, M., et al. (2013). Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *ELife* 2, e01749.
- Schlackow, M., Nojima, T., Gomes, T., Dhir, A., Carmo-Fonseca, M., and Proudfoot, N.J. (2017). Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. *Molecular Cell* 65, 25–38.
- Schneider, C., and Tollervey, D. (2013). Threading the barrel of the RNA exosome. *Trends in Biochemical Sciences* 38, 485–493.
- Schubert, S., Gül, D.C., Grunert, H.-P., Zeichhardt, H., Erdmann, V.A., and Kurreck, J. (2003). RNA cleaving ‘10-23’ DNAzymes with enhanced stability and activity. *Nucleic Acids Res* 31, 5982–5992.
- Schwarz, D.S., Tomari, Y., and Zamore, P.D. (2004). The RNA-Induced Silencing Complex Is a Mg²⁺-Dependent Endonuclease. *Current Biology* 14, 787–791.
- Shariati, S.A., Dominguez, A., Xie, S., Wernig, M., Qi, L.S., and Skotheim, J.M. (2019). Reversible Disruption of Specific Transcription Factor-DNA Interactions Using CRISPR/Cas9. *Mol Cell* 74, 622-633.
- Shikama, N., Lutz, W., Kretzschmar, R., Sauter, N., Roth, J.-F., Marino, S., Wittwer, J., Scheidweiler, A., and Eckner, R. (2003). Essential function of p300 acetyltransferase activity in heart, lung and small intestine formation. *The EMBO Journal* 22, 5175–5185.
- Simões, F.C., Cahill, T.J., Kenyon, A., Gavriouchkina, D., Vieira, J.M., Sun, X., Pezzolla, D., Ravaut, C., Masmanian, E., Weinberger, M., et al. (2020). Macrophages directly contribute collagen to scar formation during zebrafish heart regeneration and mouse heart repair. *Nature Communications* 11, 600.
- Somarowthu, S., Legiewicz, M., Chillón, I., Marcia, M., Liu, F., and Pyle, A.M. (2015). HOTAIR forms an intricate and modular secondary structure. *Mol Cell* 58, 353–361.
- Stamatoyannopoulos, J.A., Snyder, M., Hardison, R., Ren, B., Gingeras, T., Gilbert, D.M., Groudine, M., Bender, M., Kaul, R., Canfield, T., et al. (2012). An encyclopedia of mouse DNA elements (Mouse ENCODE). *Genome Biology* 13, 418.
- Takeuchi, J.K., Ohgi, M., Koshiba-Takeuchi, K., Shiratori, H., Sakaki, I., Ogura, K., Saijoh, Y., and Ogura, T. (2003). Tbx5 specifies the left/right ventricles and ventricular septum position during cardiogenesis. *Development* 130, 5953–5964.
- The FANTOM Consortium, Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., et al. (2014). An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455–461.
- Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq

reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28, 511–515.

Tsourkas, A., Behlke, M.A., and Bao, G. (2002). Hybridization of 2'-O-methyl and 2'-deoxy molecular beacons to RNA and DNA targets. *Nucleic Acids Res* 30, 5168–5174.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472, 120–124.

Wang, L., Zhang, Y.-L., Lin, Q.-Y., Liu, Y., Guan, X.-M., Ma, X.-L., Cao, H.-J., Liu, Y., Bai, J., Xia, Y.-L., et al. (2018). CXCL1–CXCR2 axis mediates angiotensin II-induced cardiac hypertrophy and remodelling through regulation of monocyte infiltration. *European Heart Journal* 39, 1818–1831.

Wang, Y., Nguyen, K., Spitale, R.C., and Chaput, J.C. (2021). A biologically stable DNAzyme that efficiently silences gene expression in cells. *Nat. Chem.* 13, 319–326.

Werner, M.S., and Ruthenburg, A.J. (2015). Nuclear Fractionation Reveals Thousands of Chromatin-Tethered Noncoding RNAs Adjacent to Active Genes. *Cell Reports* 12, 1089–1098.

Werner, M.S., Sullivan, M.A., Shah, R.N., Nadadur, R.D., Grzybowski, A.T., Galat, V., Moskowitz, I.P., and Ruthenburg, A.J. (2017). Chromatin-enriched lncRNAs can act as cell-type specific activators of proximal gene transcription. *Nat Struct Mol Biol* 24, 596–603.

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

Wu, C.-L., Yin, R., Wang, S.-N., and Ying, R. (2021). A Review of CXCL1 in Cardiac Fibrosis. *Frontiers in Cardiovascular Medicine* 8, 359.

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

Xue, Z., Hennelly, S., Doyle, B., Gulati, A.A., Novikova, I.V., Sanbonmatsu, K.Y., and Boyer, L.A. (2016). A G-Rich Motif in the lncRNA Braveheart Interacts with a Zinc-Finger Transcription Factor to Specify the Cardiovascular Lineage. *Molecular Cell* 64, 37–50.

Yang, F., Tanasa, B., Micheletti, R., Ohgi, K.A., Aggarwal, A.K., and Rosenfeld, M.G. (2021). Shape of promoter antisense RNAs regulates ligand-induced transcription activation. *Nature* 595, 444–449.

Yang, X.H., Nadadur, R.D., Hilvering, C.R., Bianchi, V., Werner, M., Mazurek, S.R., Gadek, M., Shen, K.M., Goldman, J.A., Tyan, L., et al. (2017). Transcription-factor-dependent enhancer transcription defines a gene regulatory network for cardiac rhythm. *ELife* 6, e31683.

Yelin, R., Dahary, D., Sorek, R., Levanon, E.Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R., et al. (2003). Widespread occurrence of antisense transcription in the human genome. *Nat Biotechnol* 21, 379–386.

- Yin, Y., Yan, P., Lu, J., Song, G., Zhu, Y., Li, Z., Zhao, Y., Shen, B., Huang, X., Zhu, H., et al. (2015). Opposing Roles for the lncRNA Haunt and Its Genomic Locus in Regulating HOXA Gene Activation during Embryonic Stem Cell Differentiation. *Cell Stem Cell* 16, 504–516.
- Zabidi, M.A., Arnold, C.D., Schernhuber, K., Pagani, M., Rath, M., Frank, O., and Stark, A. (2015). Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* 518, 556–559.
- Zaborowska, Z., Fürste, J.P., Erdmann, V.A., and Kurreck, J. (2002). Sequence Requirements in the Catalytic Core of the “10-23” DNA Enzyme. *J. Biol. Chem.* 277, 40617–40622.
- Zamecnik, P.C., and Stephenson, M.L. (1978). Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc Natl Acad Sci USA* 75, 280–284.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* 9, R137.