

The NANCI–Nkx2.1 gene duplex buffers Nkx2.1 expression to maintain lung development and homeostasis

Michael J. Herriges,¹ David J. Tischfield,^{2,3} Zheng Cui,⁴ Michael P. Morley,⁴ Yumiao Han,⁵ Apoorva Babu,⁴ Su Li,⁴ MinMin Lu,⁴ Isis Cendan,³ Benjamin A. Garcia,⁵ Stewart A. Anderson,^{2,3} and Edward E. Morrisey^{4,6,7,8}

¹Department of Cell and Developmental Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA;

²Neuroscience Graduate Group, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; ³Department of Psychiatry, Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; ⁴Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; ⁵Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; ⁶Penn Center for Pulmonary Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; ⁷Penn Cardiovascular Institute, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; ⁸Penn Institute for Regenerative Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

A subset of long noncoding RNAs (lncRNAs) is spatially correlated with transcription factors (TFs) across the genome, but how these lncRNA–TF gene duplexes regulate tissue development and homeostasis is unclear. We identified a feedback loop within the NANCI (Nkx2.1-associated noncoding intergenic RNA)–Nkx2.1 gene duplex that is essential for buffering Nkx2.1 expression, lung epithelial cell identity, and tissue homeostasis. Within this locus, Nkx2.1 directly inhibits NANCI, while NANCI acts in *cis* to promote Nkx2.1 transcription. Although loss of NANCI alone does not adversely affect lung development, concurrent heterozygous mutations in both NANCI and Nkx2.1 leads to persistent Nkx2.1 deficiency and reprogramming of lung epithelial cells to a posterior endoderm fate. This disruption in the NANCI–Nkx2.1 gene duplex results in a defective perinatal innate immune response, tissue damage, and progressive degeneration of the adult lung. These data point to a mechanism in which lncRNAs act as rheostats within lncRNA–TF gene duplex loci that buffer TF expression, thereby maintaining tissue-specific cellular identity during development and postnatal homeostasis.

[*Keywords:* development; lncRNAs; lung; transcription factors]

Supplemental material is available for this article.

Received February 24, 2017; revised version accepted May 2, 2017.

Long noncoding RNAs (lncRNAs), RNA transcripts ≥ 200 nucleotides that do not encode a functional protein, have been shown to regulate a wide variety of biological processes (Wapinski and Chang 2011; Li and Chang 2014; Meller et al. 2015; Li et al. 2016). Despite growing evidence that lncRNAs play important roles in regulating gene expression, their role in tissue development and homeostasis remains unclear (Li and Chang 2014; Feyder and Goff 2016; Perry and Ulitsky 2016). Recent work from our laboratory and others has identified a class of lncRNA that are spatially located near critical transcription factor (TF) genes (Guttman et al. 2009; Ulitsky et al. 2011; Herriges et al. 2014). These TF–lncRNA duplexes can be further divided into those that reside upstream of and embedded within the regulatory region of TFs or those that reside

downstream from and outside the proximal promoter and regulatory elements of the TF gene. This spatial localization is conserved across species, with mouse and human TF–lncRNA duplexes exhibiting common syntenic locations in the genome, which suggests a critical regulatory function for regulating gene expression.

In both mice and humans, the lncRNA NANCI (Nkx2.1-associated noncoding intergenic RNA) is located downstream from the TF Nkx2.1, a major regulator of pulmonary development and homeostasis (Herriges et al. 2014). Inactivation of Nkx2.1 during development leads to dramatic loss in branching morphogenesis along with defective lung epithelial lineage differentiation (Mino

Corresponding author: emorris@mail.med.upenn.edu

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.298018.117>.

© 2017 Herriges et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

et al. 1995, 1997), while postnatal loss of Nkx2.1 expression results in defects associated with a loss of lung epithelial homeostasis and respiratory failure (Snyder et al. 2013). Despite the importance of maintaining Nkx2.1 expression in the lung, little is known about how its expression is regulated (Herriges and Morrisey 2014).

Here we show that within the NANCI-Nkx2.1 duplex, NANCI plays an essential role in regulating tissue identity by acting as a transcriptional rheostat to buffer Nkx2.1 expression. During lung development and in adult lung homeostasis, NANCI acts in *cis* to positively regulate Nkx2.1, and, in turn, Nkx2.1 directly inhibits NANCI expression. Together, this generates a negative feedback loop between NANCI and Nkx2.1, which buffers against dramatic reductions in Nkx2.1 expression. NANCI expression is also controlled through interactions with Hnrnpab and Hnrnpd, which promote turnover of the NANCI transcript to modulate NANCI expression and, in turn, Nkx2.1 expression. Surprisingly, loss of NANCI expression by itself has a minimal impact on lung development and homeostasis. However, concurrent *trans* mutations of both NANCI and Nkx2.1 disrupt the buffering loop, resulting in persistent Nkx2.1 deficiency. This persistent loss of Nkx2.1 expression leads to defects in

the perinatal innate immune system and progressive lung degeneration due to a loss of lung epithelial cell identity and cellular reprogramming to a posterior endoderm fate. Together, these findings establish a new paradigm for TF-lncRNA duplexes in which lncRNAs act as rheostats to buffer expression of critical TFs to maintain cell fate and normal tissue homeostasis.

Results

Expression of NANCI in a subset of Nkx2.1+ cells during lung, brain, and thyroid development

To fully examine the expression of NANCI in a cell lineage-specific fashion as well as the consequences of loss of NANCI function, we generated a NANCI reporter line (NANCI^{creERT2:RFP}) that also disrupts expression of NANCI (Fig. 1A). The NANCI^{creERT2:RFP} allele contains a polyadenylated cassette encoding a tamoxifen-inducible cre recombinase (creERT2) linked to a TdTomato red fluorescent protein (RFP) that replaces much of exon 1 in the NANCI locus, including the splice donor site. This allows for the isolation and characterization of NANCI-expressing cells, including tamoxifen-inducible lineage tracing.

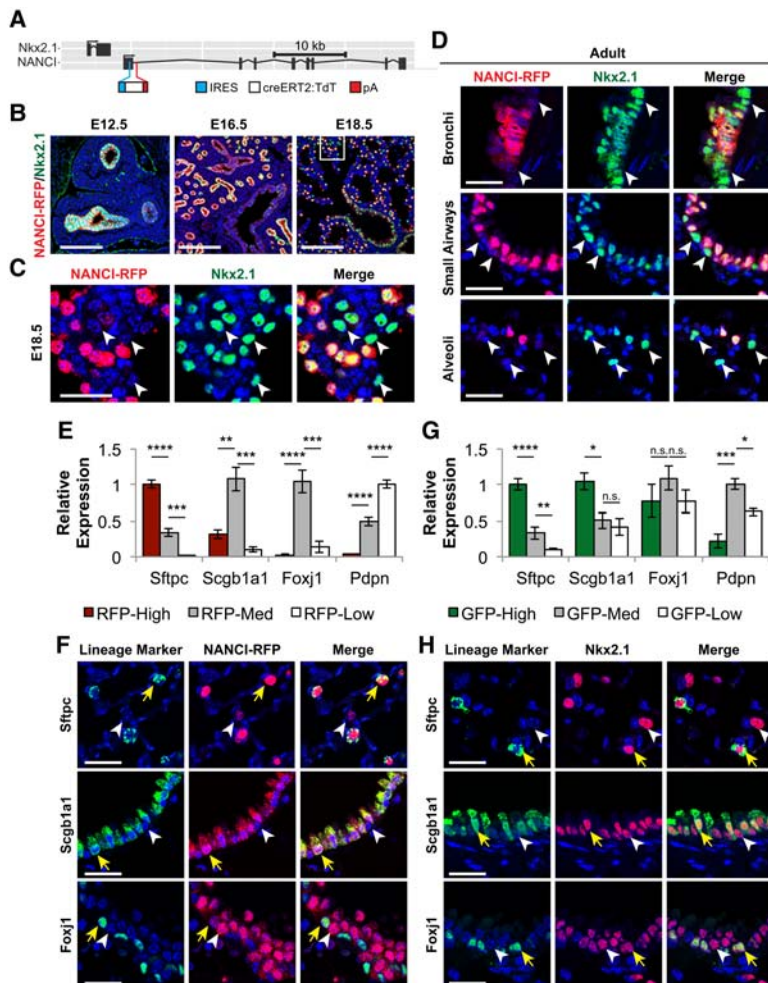


Figure 1. NANCI is expressed in a subset of Nkx2.1⁺ pulmonary epithelial cells. (A) Schematic of the NANCI^{creERT2:RFP} allele. (B) Comparison of NANCI-RFP and Nkx2.1 expression during development. (C) Expression of NANCI-RFP and Nkx2.1 at embryonic day 18.5 (E18.5) showing that some Nkx2.1⁺ cells do not express NANCI-RFP, as noted by arrowheads. (D) Expression of NANCI-RFP and Nkx2.1 in the adult lung showing that some Nkx2.1⁺ cells do not express NANCI, as noted by arrowheads. (E) Alveolar type 2 (AT2) (Sftpc⁺) and secretory (Scgb1a1⁺) cells are enriched in the FACS (fluorescence-activated cell sorting)-isolated RFP-high adult pulmonary epithelial cells, while ciliated (Foxj1⁺) and AT1 (Pdpn⁺) cells are restricted to the RFP-med and RFP-low population. (F) The FACS data were confirmed by immunohistochemistry (IHC) using the cell lineage-specific markers Sftpc (AT2), Scgb1a1 (secretory), and Foxj1 (ciliated), with representative lineage marker-positive (yellow arrow) and lineage marker-negative (white arrowhead) cells displaying different levels of NANCI-RFP. (G) In contrast, if cells are sorted by Nkx2.1-GFP, all four cell types are detectable in the GFP-high population. (H) These results are confirmed by IHC comparing representative lineage marker-positive (yellow arrow) and lineage marker-negative (white arrowhead) cells, which have similar levels of Nkx2.1. Bars: B, 100 μm; C,D,F,H, 25 μm. For E and G, means ± SEM were derived from n = 6 and n = 3 biological replicates, respectively. (*) P < 0.05; (**) P < 0.01; (***) P < 0.001; (****) P < 0.0001; (n.s.) P > 0.05, two-tailed Student's t-test.

The reporter construct was inserted into a region lacking H3K4me1 or DNase hypersensitivity peaks, suggesting that the insertion is not directly influencing a genomic enhancer region (Supplemental Fig. 1A; The ENCODE Project Consortium 2012). The NANCI^{creERT2:RFP} reporter line exhibits broad expression of RFP throughout the lung epithelium, consistent with previous NANCI in situ hybridizations expression patterns (Fig. 1B; Herriges et al. 2014). However, in both embryonic and adult lungs, we identified a subset of Nkx2.1⁺ cells that expressed significantly lower or no detectable RFP (Fig. 1C,D white arrowheads). In contrast, we were unable to find any RFP⁺/Nkx2.1⁻ cells, suggesting that NANCI is not expressed in cells lacking Nkx2.1 expression.

Outside of the lung, NANCI and Nkx2.1 are coexpressed in the forebrain and thyroid (Herriges et al. 2014). To characterize NANCI expression in the brain, we compared RFP and Nkx2.1 expression at embryonic day 12.5 (E12.5) and E18.5. At E12.5, NANCI was expressed in the medial ganglionic eminence (MGE), anterior entopeduncular/preoptic area (PoA), hypothalamus, and developing globus pallidus (Supplemental Fig. 1B). This expression pattern overlapped with that of Nkx2.1. At E18.5, NANCI continued to be expressed in most Nkx2.1⁺ cells of the hypothalamus and globus pallidus. In the thyroid, NANCI expression was restricted to the Nkx2.1⁺ epithelium (Supplemental Fig. 2A). As in the lung, NANCI expression was more variable than Nkx2.1 expression.

NANCI is preferentially expressed in alveolar type 2 (AT2) and secretory epithelial cells in the mature lung

To further explore the expression of NANCI and how it relates to cell lineage identity in the lung, we used fluorescence-activated cell sorting (FACS) to isolate adult lung epithelial cells from NANCI^{creERT2:RFP/+} mice based on RFP expression (Supplemental Fig. 3A). This sorting generated three populations (RFP-high, RFP-med, and RFP-low), which were further analyzed by quantitative real-time PCR (qPCR) to measure RFP, NANCI, and Nkx2.1 expression (Supplemental Fig. 3B). The expression of NANCI as well as Nkx2.1 correlated with RFP expression from the NANCI^{creERT2:RFP/+} line. This suggests that the NANCI^{creERT2:RFP/+} reporter accurately reflects the variations in NANCI expression, which in turn correlates with Nkx2.1 expression.

Gene expression analysis of the RFP-high population reveals that it is composed primarily of Sftpc⁺ AT2 cells and, to a lesser extent, Scgb1a1⁺ secretory cells (Fig. 1E). Foxj1⁺ ciliated cells and Pdpn⁺ AT1 cells are restricted primarily to the RFP-med and RFP-low populations. Analysis of these cell types by immunohistochemistry (IHC) verifies that AT2 and secretory cells (Fig. 1F, yellow arrows) express higher levels of RFP than neighboring epithelial cells (Fig. 1F, white arrowheads). Likewise, ciliated and AT1 cells (Fig. 1F, yellow arrows; Supplemental Fig. 3E, yellow arrows) exhibit lower expression of RFP than neighboring epithelial cells (Fig. 1F, white arrowheads). Although they were not detectable by qPCR, neuroendocrine cells also

express RFP at higher levels than the surrounding epithelial cells (Supplemental Fig. 3E). The majority of goblet and basal cells, found in mainstem bronchi, has negligible RFP expression; however, there were rare RFP-positive basal cells (Supplemental Fig. 3E) FACS of epithelial cells using an Nkx2.1^{GFP} reporter line likewise reveals three distinct populations of Nkx2.1-expressing lung epithelium (Fig. 1G; Supplemental Fig. 3C,D). However, unlike NANCI expression, Nkx2.1 expression does not clearly separate the lung epithelial cell lineages examined by qPCR or IHC (Fig. 1H; Supplemental Fig. 3F). These data suggest that NANCI expression levels, but not Nkx2.1 expression, can delineate epithelial cell lineages in the lung, with the highest levels of NANCI expression found in AT2 and secretory cells.

NANCI expression is essential for normal levels of Nkx2.1 expression

Our previous work suggested that shRNA-mediated knockdown of NANCI leads to decreased Nkx2.1 expression (Herriges et al. 2014). Therefore, we analyzed expression of NANCI and Nkx2.1 in NANCI^{creERT2:RFP/creERT2:RFP} homozygous mice. At E18.5, NANCI^{creERT2:RFP/creERT2:RFP} mice show an almost complete loss of NANCI expression by qPCR and in situ hybridization (Fig. 2A; Supplemental Fig. 4B). This reduced NANCI expression resulted in decreased expression of Nkx2.1 and known target genes, including Abca3, Sftpa, and Sftpc, while genes not directly regulated by Nkx2.1 (Pdpn and Foxj1) were unaffected (Supplemental Fig. 4A). In adult mice, loss of NANCI results in a 24% reduction in Nkx2.1 expression, which did not lead to significant changes in the expression of Nkx2.1 target genes. To determine whether expression of other genes is significantly affected upon loss of NANCI expression, we used RNA sequencing (RNA-seq) to compare gene expression in E18.5 wild-type and NANCI^{creERT2:RFP/creERT2:RFP} lung epithelial cells. Surprisingly, NANCI was the only gene identified as having significantly different expression, suggesting that the moderate effect that loss of NANCI has on Nkx2.1 expression does not result in a dramatic effect on overall lung epithelial gene expression.

Despite reduced Nkx2.1 expression, NANCI^{creERT2:RFP/creERT2:RFP} mutants display no discernable morphological defects at E18.5 or 10 wk of age (Supplemental Fig. 4C). This is consistent with the fact that Nkx2.1^{+/-} mice have similarly reduced Nkx2.1 expression levels but no clear morphological defects at these stages. Thus, while NANCI promotes Nkx2.1 expression during lung development, loss of NANCI alone does not reduce Nkx2.1 expression sufficiently to induce detectable morphological defects.

NANCI and Nkx2.1 form a negative feedback loop that buffers Nkx2.1 expression

To test whether Nkx2.1 regulates NANCI expression, we analyzed NANCI expression in Nkx2.1^{GFP/+} mice, in which one Nkx2.1 promoter drives expression of GFP

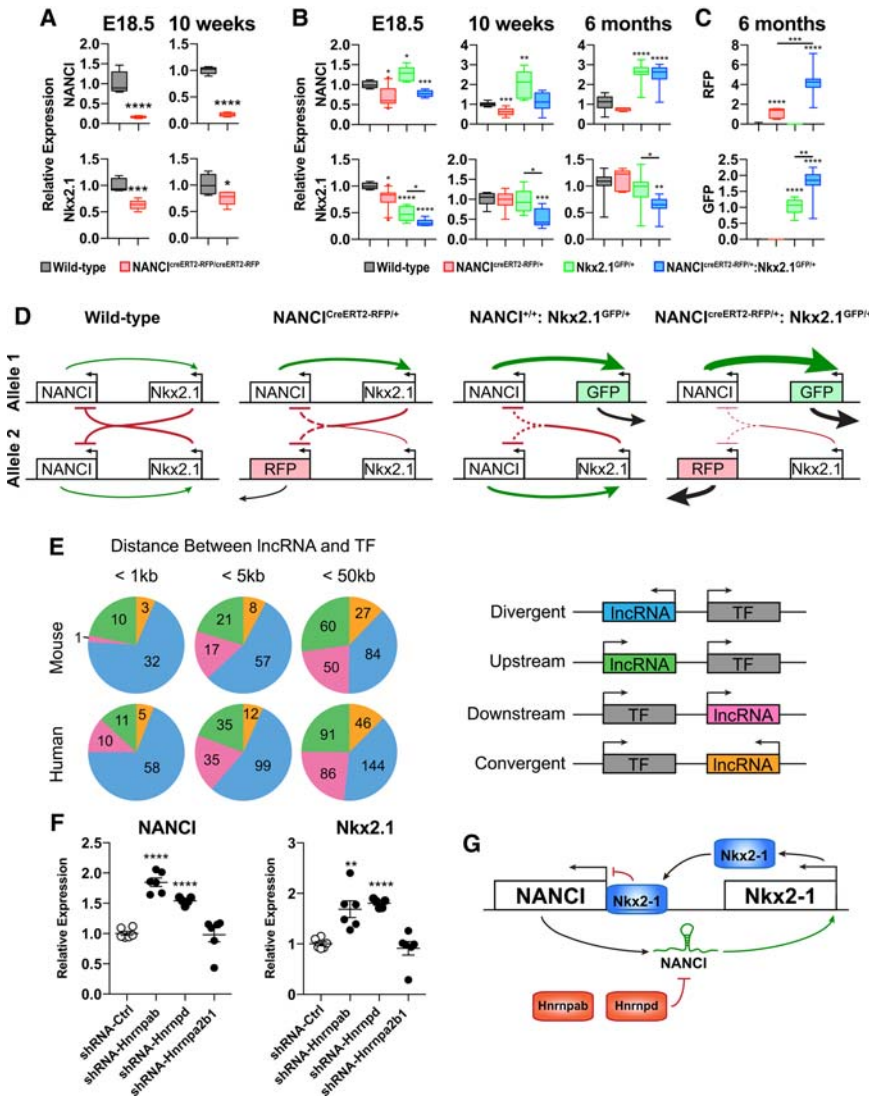


Figure 2. NANCI and Nkx2.1 form a lncRNA-TF duplex that buffers Nkx2.1 expression. (A) NANCI^{creERT2:RFP/creERT2:RFP} mice have reduced expression of NANCI and Nkx2.1 at E18.5 and 10 wk of age. (B) In contrast, Nkx2.1^{GFP/+} mice have increased NANCI expression, indicating that Nkx2.1 inhibits NANCI. Thus, NANCI and Nkx2.1 form a negative feedback loop that buffers Nkx2.1 expression. (C) Increased expression of RFP and GFP in 6-mo-old NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mice. (D) These data suggest a model in which NANCI acts in cis to regulate Nkx2.1 expression. (E) Spatial classification of lncRNA-TF based on genomic orientation of the lncRNA and TF identifies other duplexes with the lncRNA located downstream from the TF, similar to the NANCI-Nkx2.1 duplex. (F,G) Knockdown of Hnrnpab and Hnrnpd, but not Hnrnpa2b1, results in increased expression of NANCI and Nkx2.1, suggesting that Hnrnpab and Hnrnpd inhibit NANCI stability. E18.5 RNA was extracted from EpCAM⁺ lung cells, while 10-wk and 6-mo samples were collected from whole lungs. For A, box and whisker plots for E18.5 and 10 wk were derived from $n = (5,6)$ and $n = (5,5)$ biological replicates, respectively. For B and C, box and whisker plots for E18.5, 10 wk, and 6 mo were derived from $n = (5,11,7,5)$, $n = (6,9,5,8)$, and $n = (9,5,8,8)$ biological replicates, respectively. For F, means \pm SEM were derived from $n = (6,6,5,6)$ biological replicates. (*) $P < 0.05$; (**) $P < 0.01$; (***) $P < 0.001$; (****) $P < 0.0001$, two-tailed Student's *t*-test.

instead of the Nkx2.1 transcript (Longmire et al. 2012). At E18.5, these mice have a 25% increase in expression of NANCI, indicating that Nkx2.1 inhibits NANCI expression (Fig. 2B). Likewise, NANCI expression in Nkx2.1^{GFP/+} mice increases twofold and threefold at 10 wk and 6 mo of age, respectively. This increase in NANCI expression with age correlates with increased relative Nkx2.1 expression such that Nkx2.1^{GFP/+} mice express wild-type levels of Nkx2.1 by 6 mo. Analysis of NANCI and Nkx2.1 expression in the brains of Nkx2.1^{GFP/+} mice indicates that a similar feedback loop exists in the brain (Supplemental Fig. 1C). Furthermore, using chromatin immunoprecipitation (ChIP):qPCR (ChIP:PCR), we found that Nkx2.1 binds to a predicted Nkx2.1-binding site within the NANCI promoter, suggesting that Nkx2.1 directly inhibits NANCI expression (Supplemental Fig. 5C,D). This is consistent with the ability of Nkx2.1 to directly repress expression of genes, including Nrp2 (Nobrega-Pereira et al. 2008). Together, this indicates that NANCI and Nkx2.1 form a negative

feedback loop that can buffer Nkx2.1 expression, especially in situations of a compromised Nkx2.1 locus, such as in Nkx2.1^{GFP/+} mice.

To further test this buffering feedback loop, we generated mice with concurrent in trans mutations of NANCI and Nkx2.1 (NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+}). If NANCI acts in cis, NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mice should display a further reduction in Nkx2.1 expression. Indeed, at E18.5 and 10 wk and 6 mo of age, NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mice had a significant reduction in Nkx2.1 expression relative to the other genotypes (Fig. 2B). Analysis of AT2-enriched and AT2-depleted epithelial cell populations in adult NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mice indicates that Nkx2.1 is decreased in both cell populations (Fig. 5A,B). This suggests that reduced Nkx2.1 expression reflects decreased gene expression in multiple epithelial cell types and not just a loss of Nkx2.1-high AT2 cells. At 6 mo, NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mice also have increased expression of GFP, which is expressed from the Nkx2.1 locus in cis to the remaining

intact NNCI locus (Fig. 2C). This is consistent with NNCI acting in *cis*, but not in *trans*, to promote expression of the neighboring Nkx2.1 gene (Fig. 2D). Thus, NNCI and Nkx2.1 act in concert as a buffering system to control the expression of this gene duplex. Examination of TF-associated lncRNAs in the mouse and human genomes shows that there is a unique subset of lncRNAs that is located downstream from TFs in a duplex formation similar to NNCI–Nkx2.1 (Fig. 2E). These downstream sense lncRNAs likely act through a different mechanism than divergent lncRNAs, which often overlap the regulatory regions of their neighboring coding gene (Supplemental Table 1; Guttman et al. 2009; Sauvageau et al. 2013; Herriges et al. 2014; Akerman et al. 2016; Anderson et al. 2016). Thus, our data suggest that convergent downstream sense TF–lncRNA gene duplexes regulate gene expression through a complex in *cis* buffering loop, and this spatial arrangement may serve as a paradigm for similar TF–lncRNA duplex loci, including those located near *Foxa2* and *Neurog1* (Supplemental Table 1; Herriges et al. 2014).

hnRNPa and *hnRNpd* regulate Nkx2.1 expression through their ability to modulate NNCI transcript stability

To identify proteins that interact with NNCI and regulate its expression and/or function, we used mass spectrometry to identify NNCI–protein interactions in a murine lung epithelial cell line. Forty-four proteins were found to be enriched consistently in the NNCI pull-down relative to pull-down with a control RNA transcript of similar length (Supplemental Table 2; Supplemental Fig. 5E). This included *Hnrnpab*, *Hnrnpd*, and *Hnrnpa2b1*, which were verified by Western blots (Supplemental Fig. 5F). *Hnrnpab* and *Hnrnpa2b1* have been shown to mediate the function of the lncRNA *lincRNA-Cox2* (Carpenter et al. 2013). In contrast, *Hnrnpd* promotes RNA degradation (Brewer 1991; Zhang et al. 1993; White et al. 2017), and *Hnrnpab*, a member of the same subfamily, may serve a similar function (Akindahansi et al. 2005; Czapinski et al. 2005). To test the ability of these proteins to regulate NNCI activity, we used shRNAs to knock down their expression in MLE12 cells (Supplemental Fig. 5G). Knockdown of *Hnrnpab* and *Hnrnpd*, but not *Hnrnpa2b1*, resulted in increased expression of NNCI and Nkx2.1 (Fig. 2F). Together, this suggests that *Hnrnpab* and *Hnrnpd* regulate NNCI transcript stability and, in turn, Nkx2.1 expression (Fig. 2G).

The NNCI–Nkx2.1 buffering loop is essential for lung epithelial development and the perinatal innate immune response

To determine the effects of disrupting the NNCI–Nkx2.1 buffering loop during development, we analyzed pulmonary morphology and gene expression at E18.5 and postnatal day 4 (P4). At E18.5, both Nkx2.1^{GFP/+} and NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants have mild morphological defects in the developing alveolar space, as in-

dicated by extended distal airspaces that lacked much of the alveolar septation found in wild-type mice (Fig. 3A). Moreover, at this time point, both Nkx2.1^{GFP/+} and NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants have reduced expression of markers for several lung epithelial cell lineages, indicative of delayed pulmonary development (Fig. 3B,C). Nkx2.1^{GFP/+} and NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants exhibit ectopic expression of *Hnf4a*, a TF expressed in gastric epithelium, consistent with previous work suggesting that Nkx2.1 represses a gastric transcriptional program (Snyder et al. 2013).

By P4, Nkx2.1^{GFP/+} mice express normal levels of many lung epithelial markers except *Sftpc*, suggesting compensatory mechanisms for maintaining Nkx2.1 expression in the perinatal period (Fig. 3D–F). In contrast, the NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants exhibit alveolar simplification and decreased expression of lineage markers for AT2, secretory, multiciliated, and AT1 cells (Fig. 3F). At P4, *Hnf4a* expression had increased dramatically in NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants, suggesting increased activation of ectopic posterior endoderm transcriptional pathways in these mutants due to persistent Nkx2.1 deficiency. These data suggest that without NNCI in *cis*, the functional Nkx2.1 allele is unable to normalize perinatal Nkx2.1 expression.

In addition to alveolar simplification and developmental delay, NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants at P4 display a localized severe perinatal inflammatory response. This response is exemplified by engorged distal airspaces filled with a mixture of proteinaceous debris and immune cells (six out of 13 by flow or histological analysis) (Fig. 4A). Flow analysis indicates a dramatic influx of neutrophils in NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants at P4 (Fig. 4B). These mutants display gaps in the epithelial barrier in both the proximal and distal airways (Fig. 4C, white arrowheads) as well as E-cadherin-positive cells in the airway lumen (Fig. 4C, yellow arrows). This suggests a loss of epithelial barrier integrity, which likely contributes to the severe inflammatory response.

To better characterize the causes of this perinatal phenotype, we used RNA-seq to compare gene expression in wild-type, Nkx2.1^{GFP/+}, and NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} lung epithelium at E18.5, prior to evidence of an immunological response (Supplemental Fig. 6A). By gene ontology (GO) analysis, NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants displayed altered expression of several gene categories, including increased expression of genes involved in mitosis and decreased expression of genes involved in regulation of immune response (Fig. 4D; Supplemental Table 3). Increased mitosis was verified by Ki67 staining, which indicated increased proliferation of AT2 and secretory cells but not AT1 or ciliated cells (Supplemental Fig. 6B). Combined with the reduced expression of lung epithelial lineage markers, these data suggest that the lung epithelium in NNCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants is more proliferative but immature compared with wild-type controls or Nkx2.1^{GFP/+} mutants.

In the immune response categories, expression of colony-stimulating factor 2 (*Csf2*) was significantly

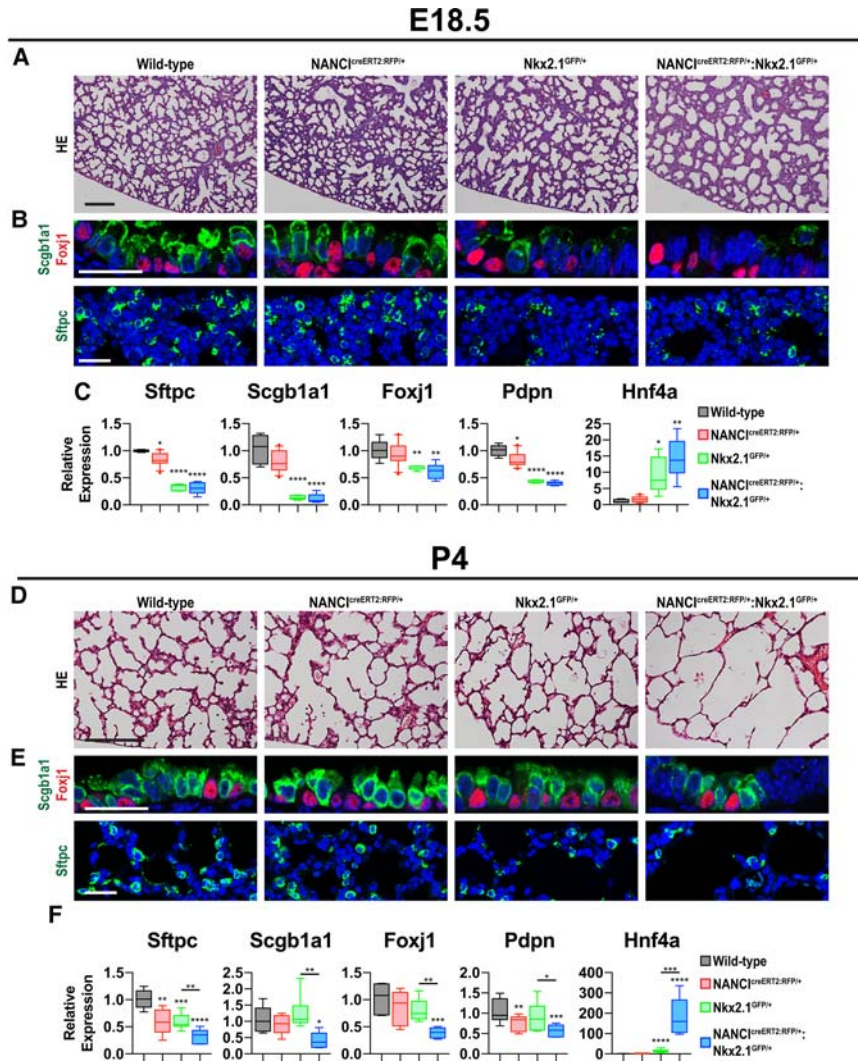


Figure 3. Disruption of the NANCI-Nkx2.1 buffering loop leads to persistent developmental delay. (A) At E18.5, NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mice have a mild morphological defect in the developing alveolar space, as noted by thickened septae in hematoxylin and eosin (HE)-stained tissue. (B) IHC for cell type-specific markers showing reduced expression of secretory (Scgb1a1) and AT2 (Sftpc) cell markers. (C) At E18.5, both Nkx2.1^{GFP/+} and NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mice exhibit decreased expression of several cell type-specific markers and ectopic expression of the posterior endoderm marker Hnf4a. (D) By P4, only NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mice exhibit significant alveolar simplification in HE-stained tissue. (E) IHC for cell type-specific markers showing reduced expression of secretory (Scgb1a1) and AT2 (Sftpc) cell markers at P4. (F) At P4, only NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} lungs exhibit reduced expression of the pulmonary cell type markers and ectopic expression of Hnf4a. In contrast, Nkx2.1^{GFP/+} mice have largely normalized development. Bars: A,D, 200 μm; B,E, 25 μm. RNA was extracted from EpCAM⁺ lung cells. For C and F, box and whisker plots were derived from n = (5,11,7,5) and n = (6,5,7,5) biological replicates, respectively. (*) P < 0.05; (**) P < 0.01; (***) P < 0.001; (****) P < 0.0001, two-tailed Student's *t*-test.

decreased in NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants. *Csf2* is a key regulator of epithelial barrier function and a cytokine that promotes granulocyte and monocyte maturation (Overgaard et al. 2015; Whitsett and Alenghat 2015). *Csf2* deficiency leads to pulmonary alveolar proteinosis (PAP), in which the pulmonary airspace is filled with noncellular debris, similar to the phenotype observed in NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants at P4. qPCR verified that *Csf2* is significantly down-regulated in Nkx2.1^{GFP/+} and NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants at E18.5 but only in NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants at P4 (Fig. 4E). Nkx2.1 binds to the *Csf2* promoter, suggesting that Nkx2.1 directly activates *Csf2* expression in the maturing pulmonary epithelium (Fig. 4G,H).

Despite reduced expression of *Csf2*, perinatal NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutant lungs contain a large population of neutrophils, suggesting up-regulation of other cytokines that recruit neutrophils. Using our RNA-seq data as well as qPCR, we found that the cytokine secreted phosphoprotein 1 (*Spp1*), also known as osteopon-

tin, was up-regulated in NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants at E18.5 (Fig. 4F). *Spp1* promotes recruitment of neutrophils, consistent with the increased neutrophil population in NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutant lungs (Zissler et al. 2016). However, we did not observe Nkx2.1 binding to the *Spp1* promoter, suggesting that up-regulation of *Spp1* is a secondary response to the barrier defects in these lungs (Fig. 4G,H; Supplemental Table 4). This is consistent with previous Nkx2.1 ChIP-seq (ChIP combined with high-throughput sequencing) data showing that Nkx2.1 binds to the *Csf2* promoter but not the *Spp1* promoter (Snyder et al. 2013). Taken together, these data suggest that the perinatal phenotype observed in NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants results in part from multiple defects, including *Csf2* deficiency, elevated *Spp1* expression, and a loss of epithelial barrier integrity.

Despite the profound perinatal phenotype seen in NANCI^{creERT2:RFP/+};Nkx2.1^{GFP/+} mutants, these animals survive to at least 4 wk of age, when the lung reaches maturity. Unlike the earlier time points, at 4 wk of age, Nkx2.1 expression is not significantly different between

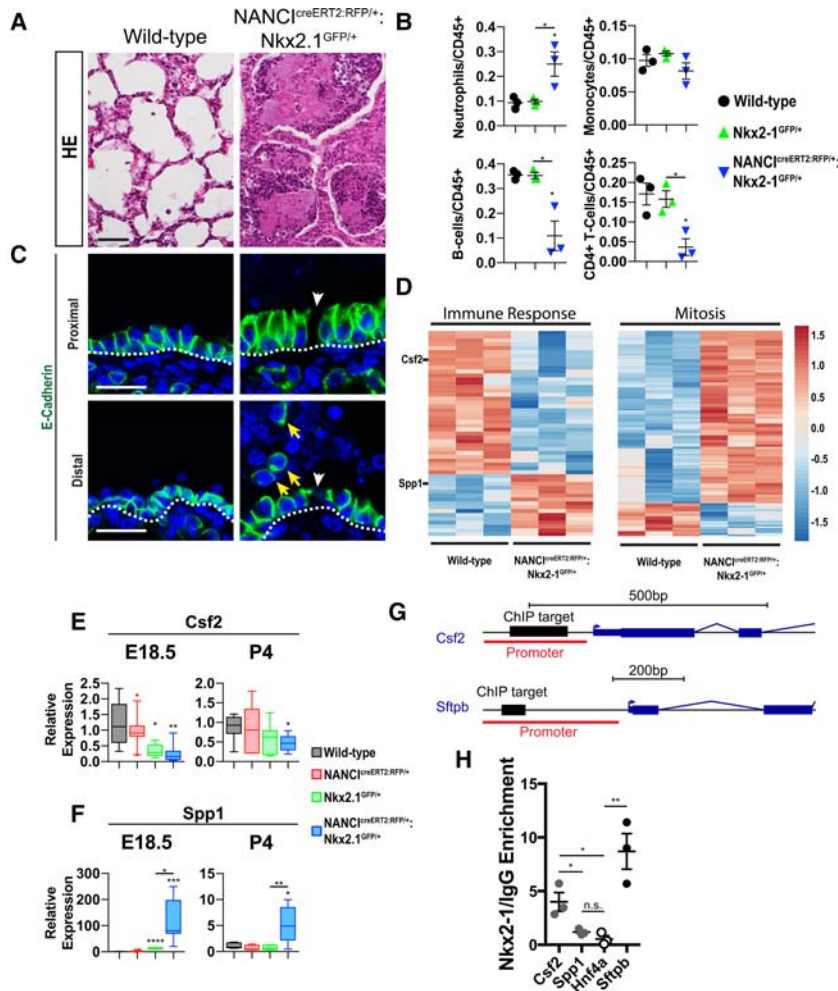


Figure 4. Severe perinatal immune response in P4 $\text{Nkx2.1}^{\text{GFP/+}}$; $\text{Nkx2.1}^{\text{GFP/+}}$ mice. (A) A subset of $\text{Nkx2.1}^{\text{GFP/+}}$; $\text{Nkx2.1}^{\text{GFP/+}}$ mice displays a localized and severe innate immune response at P4 (three out of seven by HE and three out of six by flow analysis). (B) Flow analysis indicates that these mice have an increased infiltration of neutrophils but not monocytes, B cells, or helper T cells. (C) $\text{Nkx2.1}^{\text{GFP/+}}$ lungs have defects in epithelial barrier integrity as well as E-cadherin⁺ cells in the airway lumen (white arrowheads and yellow arrows, respectively). (D) Comparison of $\text{Nkx2.1}^{\text{GFP/+}}$; $\text{Nkx2.1}^{\text{GFP/+}}$ and wild-type mice by RNA-seq indicates that the former have decreased expression of immune response genes but increased expression of genes involved in the mitotic cell cycle. (E) $\text{Nkx2.1}^{\text{GFP/+}}$ lungs express lower levels of colony-stimulating factor 2 (*Csf2*). (F) $\text{Nkx2.1}^{\text{GFP/+}}$ lungs express higher levels of secreted phosphoprotein 1 (*Spp1*). (G) Diagram of Nkx2.1-enriched ChIP targets for *Csf2* and *Sftpb*, with promoter region defined on H3K4me3 marks from ENCODE Project mouse lung tracks. (H) Nkx2.1 binds to the promoter of *Csf2* but not *Sftpb*. Data for the controls (*Hnf4a* and *Sftpb*) in *I* are the same as in Supplemental Figure 5D. Bars: A, 50 μm ; C, 25 μm . RNA for D–F was extracted from EpCAM⁺ lung cells. For B and H, means \pm SEM were derived from $n = 3$ biological replicates. For E and F, E18.5 and P4 box and whisker plots were derived from $n = (5, 11, 7, 5)$ and $n = (6, 5, 7, 5)$, respectively. (*) $P < 0.05$; (**) $P < 0.01$; (***) $P < 0.001$; (****) $P < 0.0001$, two-tailed Student's *t*-test.

$\text{Nkx2.1}^{\text{GFP/+}}$ and $\text{Nkx2.1}^{\text{GFP/+}}$; $\text{Nkx2.1}^{\text{GFP/+}}$ lungs (Supplemental Fig. 7A). This increase in Nkx2.1 expression is not accompanied by increased expression in NNCI, suggesting that a NNCI-independent mechanism is responsible for the normalization of Nkx2.1 expression. By 4 wk of age, the exogenous immune cells and PAP phenotype have also resolved, and lung morphology is similar to that of wild-type lungs (Supplemental Fig. 7B). Furthermore, expression of pulmonary cell-type markers has largely normalized to wild-type levels in these mutants (Supplemental Fig. 7C). At this stage, *Csf2* expression has increased and *Spp1* expression has decreased in $\text{Nkx2.1}^{\text{GFP/+}}$; $\text{Nkx2.1}^{\text{GFP/+}}$ lungs, counterbalancing the earlier altered expression of these cytokines and correlating with a resolution of the innate immune response observed in the perinatal period.

Disruption in the NNCI–Nkx2.1 gene duplex leads to a failure of pulmonary homeostasis and reprogramming of lung epithelium to a gastric endoderm fate

While $\text{Nkx2.1}^{\text{GFP/+}}$; $\text{Nkx2.1}^{\text{GFP/+}}$ mutants were relatively normal at 4 wk of age, adult 10-wk-old mutants

began to exhibit inflammation of the large airways as well as basal and mucous cell metaplasia (Fig. 5A). This phenotype was accompanied by Nkx2.1 deficiency, consistent with the known role of Nkx2.1 in inhibiting airway inflammation (Maeda et al. 2011). We also observed reduced expression of *Scgb1a1*, a marker of the airway secretory cell lineage (Fig. 5G). In contrast to the perinatal period, the inflammatory reaction at 10 wk was not associated with significant alterations in *Csf2* or *Spp1* expression (Fig. 5G).

Alveolar simplification and bronchiolization of the alveoli were observed in $\text{Nkx2.1}^{\text{GFP/+}}$; $\text{Nkx2.1}^{\text{GFP/+}}$ mutants at 10 wk (Fig. 5A, black arrowheads). Bronchiolization of the alveoli is characterized by cells resembling bronchiolar epithelium lining normal or thickened alveolar walls (Wang et al. 2009). These bronchiolized distal alveoli consist primarily of Sox2⁺/*Scgb1a1*⁺ cells, suggesting that they represent airway secretory cells (Fig. 5B). While we did not observe any Sox2⁺/*Pdnp*⁺ AT1 cells or Sox2⁺/*Foxj1*⁺ multiciliated epithelial cells, there were rare Sox2⁺/*Sftpc*⁺ and *Scgb1a1*⁺/*Sftpc*⁺ cells (Fig. 5C–F). At 10 wk of age, there is also a resurgence of ectopic *Hnf4a* expression in $\text{Nkx2.1}^{\text{GFP/+}}$; $\text{Nkx2.1}^{\text{GFP/+}}$ mutants

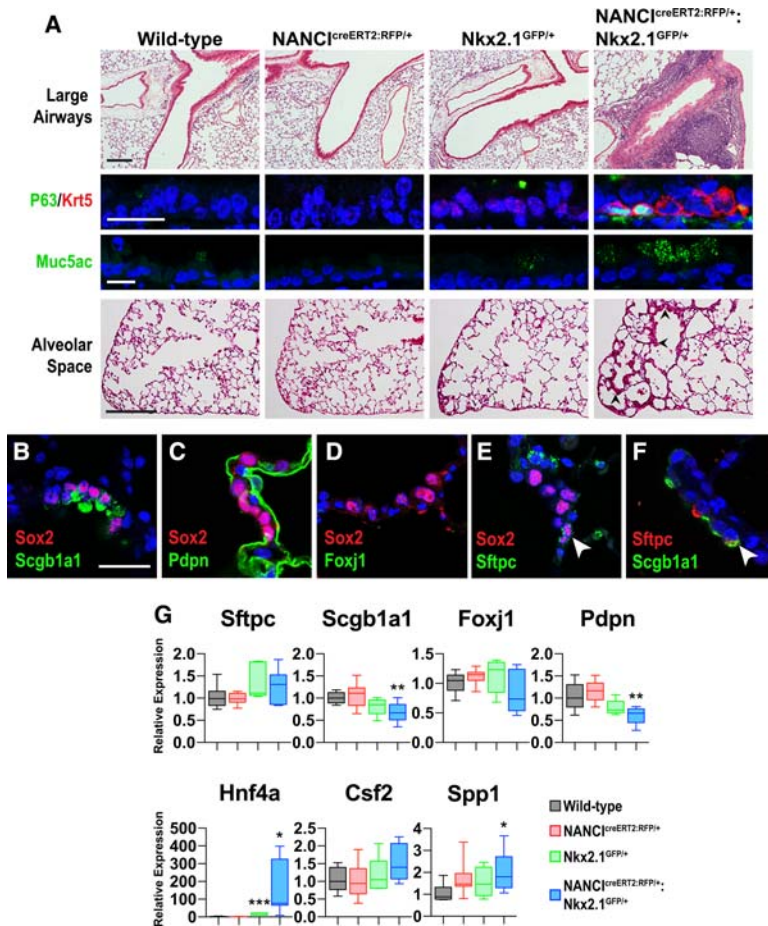
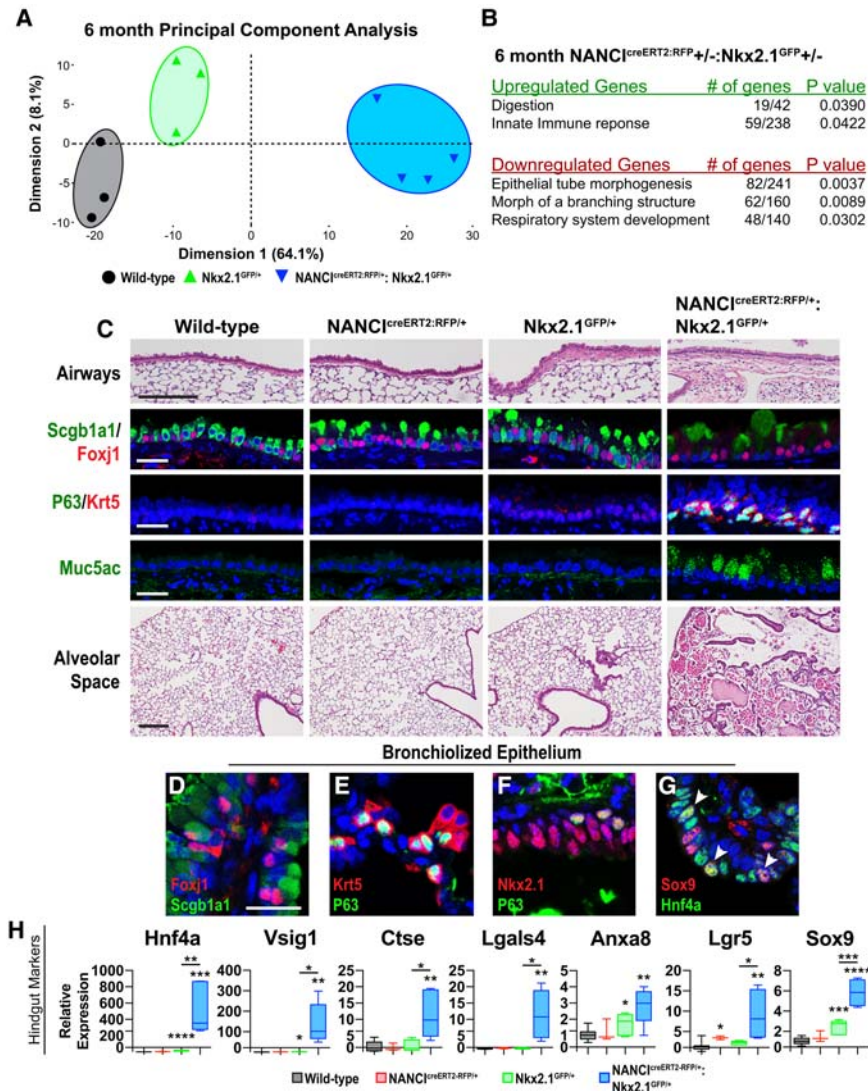


Figure 5. Disruption of the Nanci-Nkx2.1 duplex buffering loop leads to failure of pulmonary homeostasis. (A) At 10 wk of age, Nanci^{creERT2:RFP/+}; Nkx2.1^{GFP/+} mice exhibit airway inflammation (HE staining; top row) as well as basal (p63 and Krt5 IHC; second row) and mucous (Muc5ac IHC; third row) cell metaplasia. (Fourth row) These mice also show signs of alveolar simplification and bronchiolization of the alveoli (HE staining). (B) Bronchiolized cells in the distal lungs of Nanci^{creERT2:RFP/+}; Nkx2.1^{GFP/+} mice at 10 wk are primarily Sox2⁺/Scgb1a1⁺. (C,D) These do not express Pdpn or Foxj1. (E,F) A small subset of these cells is Sox2⁺/Sftpc⁺ or Scgb1a1⁺/Sftpc⁺ (white arrowheads). (G) qPCR for lineage markers confirms a reduction in secretory (Scgb1a1) and AT1 (Pdpn) cells but an increase in Hnf4a⁺ cells. Bars: A, 200 μm; B–F, 25 μm. RNA was extracted from whole-lung samples. For G, box and whisker plots were derived from n = (6,9,5,8). (*) P < 0.05; (**) P < 0.01; (***) P < 0.001, two-tailed Student's t-test.

(Fig. 5G). Pdpn and Scgb1a1 expression is decreased in Nanci^{creERT2:RFP/+}; Nkx2.1^{GFP/+} mutants at 10 wk, indicating disruption in the Nanci-Nkx2.1 regulatory gene duplex and defects in AT1 and secretory cell homeostasis (Fig. 5G).

Six-month-old Nanci^{creERT2:RFP/+}; Nkx2.1^{GFP/+} mutants exhibit further progression of the phenotypes observed at earlier stages. RNA-seq analysis of the lung epithelial cells at 6 mo of age revealed decreased expression of genes involved in respiratory and epithelial development but increased expression of genes involved in digestion, confirming the reprogramming of pulmonary epithelial cells to a posterior endoderm fate (Fig. 6A,B; Supplemental Fig. 8A,B). At this time, airway inflammation has resolved, but there is a dramatic loss of pulmonary secretory cells and ectopic expansion of goblet (Muc5ac⁺) and basal (Krt5⁺/Trp63⁺) cells (Fig. 6C; Supplemental Fig. 8C). The bronchiolization of the alveolar region has progressed concomitantly with a sharp decrease in expression of the AT2 marker Sftpc (Fig. 6C; Supplemental Fig. 8B). The bronchiolized alveolar region contains epithelium bearing markers of secretory and ciliated cells of the proximal airways as well as ectopic Krt5⁺/Trp63⁺ cells in the distal airspaces, suggesting either loss of alveolar epithelium and extension of airway

basal cells into the alveolar space or a reprogramming of alveolar epithelium into a basal cell phenotype (Fig. 6D–F; Supplemental Fig. 8C). Importantly, this bronchiolized epithelium expresses high levels of the Hnf4a P2 isoform, but not the P1 isoform, characteristic of stomach epithelial cells (Supplemental Fig. 8D; Czaplinski et al. 2005). This ectopic Hnf4a expression is not found in the airways of the lung, suggesting that the gastric reprogramming is specific to the bronchiolized alveoli (Supplemental Fig. 8E). Importantly, these bronchiolized regions contain Sox9⁺/Hnf4a⁺ cells, resembling gastric stem cells (Fig. 6G,H). Additional markers of gastric endoderm, including Lgr5, Vsig1, Lgals4, and Ctse, are also dramatically increased in Nanci^{creERT2:RFP/+}; Nkx2.1^{GFP/+} mutants at 6 mo (Fig. 6H). The Nanci^{creERT2:RFP/+}; Nkx2.1^{GFP/+} peripheral epithelium is highly apoptotic and degenerative at this stage (Supplemental Fig. 8F). This correlates with increased expression of immune response genes and recruitment of macrophages to clear cellular debris in the airways (Fig. 6B; Supplemental Fig. 8G,H). These molecular and cellular changes result in a dramatic loss of alveolar structure in Nanci^{creERT2:RFP/+}; Nkx2.1^{GFP/+} mutants at this time point (Fig. 6C). These phenotypes are in contrast to all of the other genotypes examined, including Nkx2.1^{GFP/+} mice, whose gene expression profile and



morphology are indistinguishable from those of wild-type mice (Fig. 6C; Supplemental Fig. 8B–D,F). At these time points, the adult brain showed no major morphological defects (data not shown). However, the thyroids of both Nanci^{CreERT2:RFP/+};*Nkx2.1*^{GFP/+} and *Nkx2.1*^{GFP/+} mutants exhibit similar morphological defects as well as elevated serum thyroid-stimulating hormone (TSH) levels, consistent with hypothyroidism due to *Nkx2.1* deficiency (Supplemental Fig. 2B,C; Kusakabe et al. 2006).

Examination of cell proliferation in Nanci^{CreERT2:RFP/+};*Nkx2.1*^{GFP/+} mutant lungs at 6 mo reveals that the Scgb1a1⁺ secretory cells found in the distal bronchiolized alveoli are highly proliferative (Supplemental Fig. 8I,K). In contrast, Sftpc⁺ AT2 cells in these mutants do not exhibit an increase in proliferation, suggesting that they are not responding to the loss of homeostasis by reactivating a pro-growth regenerative program (Supplemental Fig. 8J, K). These data indicate that loss of the Nanci–*Nkx2.1* rheostat function on the gene duplex leads to a cell lineage-specific loss of lung epithelial homeostasis caused in part by a reprogramming of distal alveolar epithelium to-

ward a more posterior endoderm fate. This loss of lung epithelial cell identity also appears to reduce the regenerative potential of resident AT2 cells, possibly due to the loss of their native cell fate. Taken together, these studies reveal that Nanci acts as an important rheostat of *Nkx2.1* expression that is necessary for maintaining lung epithelial cell identity and function.

Discussion

Several reports have shown that lncRNAs regulate protein-coding genes in *cis*; however, few have revealed the in vivo consequences of this regulation (Feyder and Goff 2016; Kotzin et al. 2016; Quinn and Chang 2016). Genetic inactivation of lncRNAs has often resulted in no apparent phenotype, leading to the conclusion that many lncRNAs have little impact on gene regulation and tissue development or that their loss is compensated for by other mechanisms (Eissmann et al. 2012; Zhang et al. 2012; Sauvageau et al. 2013; Kohtz 2014). The present study

Figure 6. Progressive pulmonary degeneration and lineage reprogramming in 6-month-old Nanci^{CreERT2:RFP/+};*Nkx2.1*^{GFP/+} mice. (A) Principal component analysis of RNA-seq data from lung epithelial cells of 6-month-old wild-type, *Nkx2.1*^{GFP/+}, and Nanci^{CreERT2:RFP/+};*Nkx2.1*^{GFP/+} mice. (B) GO analysis of these samples indicates that Nanci^{CreERT2:RFP/+};*Nkx2.1*^{GFP/+} mice have up-regulation of genes involved in digestion and innate immune response but decreased expression of genes involved in epithelial and respiratory morphogenesis. The second column of this table reflects genes affected out of total genes in each category. (C, top row) By 6 mo, inflammation around the large airways has receded, leaving behind fibrous tissue (HE staining). The airway epithelium has reduced expression of the secretory cell marker Scgb1a1 but expanded expression of basal (p63 and Krt5) and goblet (Muc5ac) cell markers. (Bottom row) Normal alveolar architecture is highly disrupted in Nanci^{CreERT2:RFP/+};*Nkx2.1*^{GFP/+} mice (HE staining). (D–G) The distal bronchiolized epithelium also includes ciliated cells (D), basal-like cells (E,F), and Sox9⁺/Hnf4a⁺ cells (G). (H) qPCR for gastric markers. Bars: C, 200 μ m; D–G, 25 μ m. RNA for RNA-seq was extracted from purified epithelial cells, while RNA for qPCR was extracted from whole-lung samples. For H, box and whisker plots were derived from $n = (8,3,5,5)$ biological replicates. (*) $P < 0.05$; (**) $P < 0.01$; (***) $P < 0.001$; (****) $P < 0.0001$, two-tailed Student's *t*-test.

shows that while the loss of the lncRNA NANCI alone does not affect overall lung development or homeostasis in a significant manner, NANCI is required to maintain Nkx2.1 expression. Loss of NANCI's rheostat function in situations of compromised Nkx2.1 expression leads to profound innate immune defects and a reprogramming of lung epithelium toward a posterior endoderm phenotype. The overall effect on adult tissue homeostasis is dramatic and provides critical insight into how lncRNAs function as transcriptional rheostats, which are essential for buffering neighboring protein-coding genes, including master TFs such as Nkx2.1.

Previous studies have revealed that lncRNAs are found near TFs more often than other categories of protein-coding genes (Guttman et al. 2009; Ulitsky et al. 2011; Herriges et al. 2014). Interestingly, mutations in TFs underlie many human congenital haploinsufficient diseases (Dang et al. 2008). In particular, NKX2.1 haploinsufficiency causes a rare but often lethal disorder called brain–lung–thyroid syndrome (BLTS) (Hamvas et al. 2013). Symptoms in BLTS patients include respiratory failure along with serious thyroid and neural dysfunction. However, there is a wide variation in the phenotypes observed in BLTS patients, with many lacking a notable phenotype in any one of the tissues that express NKX2.1. In addition to BLTS, alterations in NKX2.1 have been associated with different forms of lung cancer (Yang et al. 2012; Yamaguchi et al. 2013). Given the critical role for TFs such as NKX2.1 in both development and tissue homeostasis, exquisite control of their expression levels is essential. This fine control likely requires multiple modes of regulation, but the role of TF-associated lncRNAs in this system is largely unknown.

While lncRNA expression has been shown to be highly tissue-specific, characterization of in vivo expression patterns has been hampered by the lack of reporters and the inability to use standard antibody protein detection techniques (Cabili et al. 2011). We overcame these limitations by using a fluorescent and lineage tracing reporter system to characterize in detail the expression of NANCI and compare it at cellular resolution with its neighboring TF and gene duplex partner, Nkx2.1. We found that NANCI expression, while highly similar to Nkx2.1, shows cell lineage-specific differences of expression in embryonic and adult lungs. The highest levels of NANCI expression are restricted to AT2 and secretory lineages, suggesting that NANCI plays a preferential role in the development and homeostatic maintenance of these epithelial lineages. NANCI-null mice have reduced expression of Nkx2.1, verifying previous studies indicating that NANCI promotes Nkx2.1 expression (Herriges et al. 2014). Surprisingly, Nkx2.1 appears to repress NANCI expression through directly binding to the NANCI promoter, similar to how Nkx2.1 directly represses expression of Nrp2, suggesting a feedback loop within the NANCI–Nkx2.1 gene duplex (Fig. 7A; Nobrega-Pereira et al. 2008). The net effect of this loop is to buffer against drastic changes in Nkx2.1 expression, as illustrated by the normalization of Nkx2.1 expression in aging Nkx2.1^{GFP/+} mice. Previous studies suggest that NANCI may regulate Nkx2.1 expression through the recruitment of Wdr5 and the Trithorax complex; however, further work is needed to verify these findings (Herriges et al. 2014). In addition to the NANCI–Nkx2.1 feedback loop itself, the proteins Hnrnpab and Hnrnpd inhibit expression of NANCI, adding an additional check on the expression of NANCI and Nkx2.1. Thus,

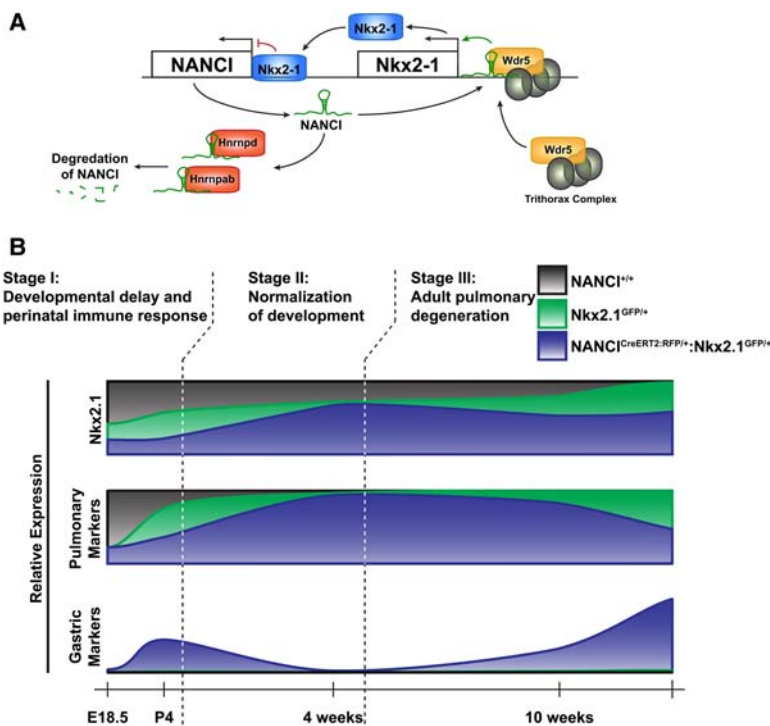


Figure 7. Disruption of the NANCI–Nkx2.1 feedback loop leads to three stages of phenotypic progression in NANCI^{CreERT2:RFP/+};Nkx2.1^{GFP/+} mice. (A) NANCI and Nkx2.1 form a negative feedback loop that buffers Nkx2.1 expression. In addition, the proteins Hnrnpab and Hnrnpd interact with NANCI and inhibit its expression, potentially acting as another check on NANCI and Nkx2.1 expression. Previous work also suggests that Wdr5 and the Trithorax complex may mediate NANCI's ability to regulate Nkx2.1. (B) In this study, we saw three phenotypic stages in the NANCI^{CreERT2:RFP/+};Nkx2.1^{GFP/+} mice, all of which are dependent on Nkx2.1 expression. First, persistent Nkx2.1 deficiency leads to a perinatal developmental delay and ectopic expression of Hnf4a, leading to an innate immune response (stage I). Over the next 4 wk, Nkx2.1 expression is up-regulated, inducing a normalization of development (stage II). Finally, in the adult lung, persistent Nkx2.1 deficiency leads to degeneration of pulmonary epithelial lineages and a resurgence of a gastric transcriptional program (stage III).

our studies point to a new model for lncRNA regulatory function in which lncRNAs act as rheostats that buffer expression of neighboring TFs. This model suggests that lncRNAs in such lncRNA–TF duplexes are not an absolute requirement for expression of neighboring TFs but instead act as fine-tuning rheostats that ensure proper TF expression in response to genetic or environmental stress.

Using *in vivo* loss-of-function models for Nanci and Nkx2.1, we show that loss of Nanci expression leads to reduced Nkx2.1 expression but not major morphological defects. This suggests as yet unknown compensatory mechanisms for regulating Nkx2.1 expression in the absence of Nanci. However, by the end of gestation, both Nkx2.1^{GFP/+} and Nanci^{creERT2:RFP/+};Nkx2.1^{GFP/+} mice show signs of pulmonary developmental delay (Fig. 7B). Postnatal Nkx2.1^{GFP/+} mice up-regulate Nanci in a compensatory fashion, which increases Nkx2.1 expression back to wild-type levels and restores normal pulmonary maturation. In contrast, concurrent *in trans* heterozygous loss of both Nanci and Nkx2.1 leads to persistent Nkx2.1 deficiency at all stages of life. At P4, this deficiency begins to reprogram lung epithelial cell identity, causing ectopic expression of a gastric transcriptional program and delayed development of pulmonary epithelial cells necessary for proper lung homeostasis. The persistent Nkx2.1 deficiency in Nanci^{creERT2:RFP/+};Nkx2.1^{GFP/+} aged mutants leads to a severe pulmonary degenerative phenotype by 6 mo, as the normal alveolar structure is replaced by ectopically reprogrammed airway and gastric-like cells. These reprogrammed lung epithelial cells cannot maintain lung alveolar epithelial cell identity and are unlikely to be able to efficiently perform gas exchange. This progressive and degenerative phenotype indicates that while a complete loss of Nanci can be compensated for in the presence of two intact Nkx2.1 alleles, further decrease in Nkx2.1 expression due to deletion of one allele results in a failure to maintain expression of this critical TF. This in turn results in a loss of adult lung epithelial cell fate.

There have been several lncRNAs identified near important TFs, including Gata6, Foxa2, Foxf1, and eHand (Guttman et al. 2009; Sauvageau et al. 2013; HERRIGES et al. 2014; Anderson et al. 2016). These syntenically conserved lncRNAs are found either upstream of or downstream from their neighboring TFs. However, it has been unclear whether the relative localizations or orientations reflect unique functions. The lncRNA Upperhand, which is located 150 base pairs upstream of and antisense to the TF Hand2, has been shown to play an almost essential role for expression of the TF Hand2 (Anderson et al. 2016). The lncRNA Fendrr similarly shares a divergent promoter region with the TF Foxf1 but, like Nanci, is not absolutely necessary for expression of its neighboring TF (Sauvageau et al. 2013). In contrast, Nanci is located downstream from and sense to Nkx2.1. This orientation is observed in a significant subset of lncRNAs associated with crucial TFs such as Foxa2 and Neurog1 (Supplemental Table 1; HERRIGES et al. 2014). The downstream orientation of these lncRNAs away from the upstream regulatory regions of neighboring coding genes suggests a mode of ac-

tion different from those that share divergent promoters with the neighboring TF. The Nanci–Nkx2.1 duplex provides a novel model for how this subset of lncRNAs functions as rheostats to buffer expression of neighboring TF genes. This category of lncRNAs, including Nanci, is likely critical for maintaining TF expression in response to genetic or environmental stress. Thus, it will be important to assess their function *in trans* to their neighboring TFs using compound mutants such as described here. Given the diverse functions identified thus far for lncRNAs, it is essential that their functions be assessed fully and carefully *in vivo* to define their contribution to tissue development and homeostasis. Analysis and comparison of these lncRNAs will help to reveal the common and unique mechanisms necessary for their ability to act as buffering rheostats for TF expression.

Materials and methods

Animals

The Nkx2.1^{GFP} mouse line was a generous gift from Darrell Kotton (Boston University), and its construction has been described previously (Longmire et al. 2012). The Nanci^{creERT2:RFP} mouse was constructed by inserting an expression cassette consisting of an IRES sequence, creERT2 cDNA linked to a 2A self-cleaving peptide sequence linked to a TdTomato cDNA, and a polyadenylated stop site into the first splice donor site of Nanci. Full details on the construction of this reporter line are available on request. All animal studies were performed under guidance of the University of Pennsylvania Institutional Animal Care and Use Committee.

Isolation of adult epithelial cells by FACS

Adult epithelial cells were collected as described previously (Frank et al. 2016). Nanci^{creERT2:TdTomato} and Nkx2.1^{GFP/+} adult lungs were harvested and processed into single-cell suspensions using a dispase (Collaborative Biosciences)/collagenase (Life Technologies)/DNase solution. Epithelial cells were sorted from the single-cell suspension using either a MoFlo Astrios EQ (Beckman Coulter) or FACSJazz (BD Biosciences) flow cytometer with antibody staining for CD31-PECy7 (eBioscience), CD45-PECy7 (eBioscience), and EpCAM-APC (eBioscience). Following negative selection for CD31 and CD45, epithelial cells were isolated by positive selection for EpCAM, separated based on RFP or GFP expression, and sorted into Trizol LS.

AT2-enriched and AT2-depleted populations were isolated from 18-wk-old mice using a modified version of previously published methods (Alder et al. 2015). In short, pulmonary cells were digested and sorted as described above using antibody staining for Sca-1-APC (eBioscience), CD34-APC (BioLegend), Pdpn-eFluor660 (eBioscience), CD31-PECy7 (eBioscience), CD45-PECy7 (eBioscience), and EpCAM-eFluor780 (eBioscience). AT2-enriched (PECy7⁻, eFluor780⁺, and APC/eFluor660⁻) and AT2-depleted (PECy7⁻, eFluor780⁺, and APC/eFluor660⁺) cells were sorted into Trizol LS.

Pulmonary qPCR

E18.5 and P4 lungs were harvested and processed into single-cell suspensions using a dispase (Collaborative Biosciences)/collagenase (Life Technologies)/DNase solution. Epithelial cells were enriched using an EpCAM antibody (eBioscience, 17-5791-82)

and Dynabeads (Invitrogen, sheep anti-rat IgG) following the manufacturer's instructions. Adult lungs (4 wk, 10 wk, and 6 mo) were harvested, and total RNA was isolated using Trizol (Invitrogen) per the manufacturer's protocol. cDNA was synthesized from total RNA by using the SuperScript strand synthesis system (Invitrogen). qPCR was performed using the SYBR Green system (Applied Biosystems) with primers listed in Supplemental Table 5. GAPDH expression values were used to control for RNA quality and quantity.

Forebrain qPCR

For analysis of forebrain transcription, brain regions caudal to the expression of Nkx2.1 at the junction of the hypothalamus were removed, and the remaining forebrains were placed in TRIzol reagent (Invitrogen). RNA was extracted using the PureLink RNA minikit (Invitrogen). cDNA was synthesized using the VILO cDNA synthesis kit (Invitrogen). qPCR was conducted using TaqMan gene expression assays (Applied Biosystems). Each sample was run in triplicate on the same plate on a Stratagene MX3005P real-time PCR machine following the manufacturer's recommended protocol.

RNA-seq analysis

E18.5 lung epithelial cells from wild-type, $\text{NANC1}^{\text{creERT2:RFP/creERT2:RFP}}$, $\text{Nkx2.1}^{\text{GFP/+}}$, and $\text{NANC1}^{\text{creERT2:RFP/+}};\text{Nkx2.1}^{\text{GFP/+}}$ lungs (three samples per genotype) were collected as described for qPCR, and RNA was isolated using Trizol. The University of Pennsylvania Next-Generation Sequencing Core performed library preparation and RNA-seq for these samples. Six-month epithelial cells were sorted into Trizol as described for FACS; however, epithelial cells were not sorted into distinct populations based on RFP or GFP expression. Samples were collected from three wild-type, three $\text{Nkx2.1}^{\text{GFP/+}}$, and four $\text{NANC1}^{\text{creERT2:RFP/+}};\text{Nkx2.1}^{\text{GFP/+}}$ lungs. Library preparation and sequencing for these samples were performed at GeneWiz, LLC.

Fastq files were assessed for quality control using the FastQC program. Fastq files were aligned against the mouse reference genome (mm9) using the STAR aligner (Dobin et al. 2013). Duplicate reads were flagged using the MarkDuplicates program from Picard tools. Per-gene read counts for Ensembl (version 67) gene annotations were computed using the Rsubread R package, with duplicate reads removed. Gene counts represented as counts per million (CPM) were first normalized using the TMM method in the edgeR R package, and genes with 25% of samples with a CPM >1 were removed and deemed low expressed. The data were transformed using the VOOM function from the limma R package (Law et al. 2014). Differential gene expression was performed using a linear model with the limma package. Heat maps and principal component analysis plots were generated in R. Heat maps in Figure 4D were generated from genes expressed at significantly different levels in $\text{NANC1}^{\text{creERT2:RFP/+}};\text{Nkx2.1}^{\text{GFP/+}}$ mice, including *Csf2*, *Spp1*, and genes in GO categories associated with mitosis (GO:0000087 and GO:0007067) or immune response (GO:0002819, GO:0002712, GO:0002697, and GO:0042742). GO enrichment analysis for the biological process was performed using R package GAGE. The Gene Expression Omnibus accession number for the RNA-seq data is GSE92304.

TF-lncRNA duplex identification

Gene annotation and position information were obtained from the GENCODE project for humans (version 25) and mice (version

12) (Harrow et al. 2012). Genes annotated with transcript_type of "lincRNA" and "bidirectional_promoter_lincRNA" were used to define lncRNAs. The distance between each protein-coding gene and the closest nonoverlapping lncRNA was computed using the closest function in the BedTools suite. Protein-coding genes were annotated and filtered as TFs using the AnimalTFDB 2.0 database (Zhang et al. 2015).

Flow cytometric analysis of immune populations

Flow cytometry was used to identify immune cells in P4 lungs. P4 lungs were harvested and processed into a single-cell suspension as described for FACS. Cells were stained with Aqua Live/Dead viability stain to exclude dead cells. Cells were then stained with one of two standard panels of antibodies listed in Supplemental Table 5. After staining for 30 min on ice, cells were washed in PBS and run for data acquisition with a BD LSR11 cytometer using the FACS Diva software (BD Bioscience). Data were analyzed using FlowJo. For identification of neutrophils and Ly6c^+ monocytes, we used the following strategy. After exclusion of dead cells, all other cells expressing the common leukocyte marker CD45 were identified. Cd11b^+ myeloid cells were then divided into $\text{Ly6g}^{\text{High}}$ neutrophils and Ly6c^+ monocytes. For identification of B cells (B220^+) and CD4^+ T cells, we used the following strategy. After exclusion of dead cells, all other cells expressing the common leukocyte marker CD45 were identified, and lymphocytes were selected using their side scatter and forward scatter profiles. For identification of CD4^+ T cells, we gated the population CD4^+ cells. B cells were identified as those cells that were $\text{CD4}^-/\text{CD8}^-$ as well as B220^+ .

TSH test

Serum was collected from 10-wk-old mice and analyzed for TSH levels by ELISA following the manufacturer's instructions (Thyroid-Stimulating Hormone kit, Cloud-Clone Corp., ABIN415519).

ChIP:PCR

ChIP:PCR was performed as described previously (Snyder et al. 2013). Adult lung epithelial cells were isolated from a single-cell suspension using an EpCAM antibody (eBioscience, 17-5791-82) in conjunction with anti-APC microbeads (Miltenyi Biotec, 130-090-855) and LS columns (Miltenyi Biotec, 130-042-401). These cells were sonicated using a Bioruptor for 8 min (30 sec on, 30 sec off), and Nkx2.1-bound DNA was purified using an Nkx2.1 antibody (Millipore, 07-601) and the MAGnify ChIP kit following the manufacturer's instructions. qPCR was performed on the isolated samples using primers listed in Supplemental Table 4 and normalized to samples isolated using a rabbit IgG antibody. The mouse genome (mm9; human hg19) was scanned for the occurrence of Nkx2-1 motifs using the FIMO program from the MEME suite (Grant et al. 2011).

Sample preparation for mass spectrometry

The MLE12 cell line was obtained from American Type Culture Collection and cultured in the HITES medium (Ham's F12 [Mediatech], 0.005 mg/mL insulin, 0.01 mg/mL transferrin, 30 nM sodium selenite, 10 nM hydrocortisone, 10 nM β -estradiol, 10 mM HEPES, 2 mM L-glutamine) with 2% FBS and 1% penicillin/streptomycin. One week before collection, half of the MLE12 cells were switched to SILAC (stable isotope labeling by amino acids in cell culture) medium containing labeled lysine and arginine. RNA pull-down assay was performed as described

previously (Tsai et al. 2010). Proteins were pulled down using either biotinylated NNCI or biotinylated IRES-GFP-Puro mRNA, a processed transcript of similar length to NNCI. In each of three experiments, one transcript was used to pull down proteins from SILAC-labeled nuclear extract, and the alternate transcript was used to pull down proteins from unlabeled nuclear extract. To avoid protein enrichment due to the SILAC labeling, we alternated which transcript was used to pull down from the SILAC-labeled nuclear extract. The eluted samples were run in a 12% Bis-Tris gel (Life Technologies), and resolved proteins were visualized by silver staining. The entire lanes were then excised, diced into 1-mm³ pieces taking precautions against keratin contamination, and alternatively washed in water followed by 100 mM NaHCO₃ and dehydration in acetonitrile. The samples were reduced using 5 mM dithiothreitol (Sigma) at 55°C and alkylated by 14 mM iodoacetamide (Sigma) at room temperature in the dark. This was followed by in-gel digestion using a 1:100 ratio of sequencing-grade modified trypsin (Promega) to protein ratio overnight at 37°C. These peptides were extracted in the solution outside the gel pieces. These were subsequently desalted using C18 reversed-phase cartridges (Sep-Pak, Waters) and dried to remove organic solvents. The dried peptides were resuspended in 0.1% acetic acid for liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis and loaded onto an Easy-nLC system (Thermo Fisher Scientific) coupled online with an Orbitrap Elite mass spectrometer (Thermo Scientific). Paired SILAC-labeled and unlabeled samples were combined-loaded on a pulled tip fused silica column with a 100- μ m inner diameter packed in-house with 12 cm of 3 μ M C18 resin (Reprosil-Pur C18-AQ) using solvents: 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). A gradient of 95 min was set for peptide elution from 2%–35% buffer B. The mass spectrometer was operated in positive ion mode and the data-dependent mode with dynamic exclusion enabled (repeat count: 1; exclusion duration: 0.5 min). Full MS scan (m/z 350–1650) was collected at a resolution of 60,000 at an AGC target value of 1×10^6 . MS2 scans of the most intense peptide ions was performed using CAD (normalized collision energy = 35%; isolation width = 2 m/z) at an AGC target value of 1×10^4 or HCD (normalized collision energy = 36%; isolation width = 2 m/z ; resolution = 15,000) at an AGC target value of 5×10^4 . Ions with a charge state of 1 and a rejection list of common contaminant ions (exclusion width = 10 ppm) were excluded from the analysis.

Proteomics data analysis

MS data were analyzed using Proteome Discoverer 1.4 (Thermo Scientific) for protein identification/quantification. The mouse UniProt database was used for Mascot searches for assigning peptide sequences and identification of proteins. The search was performed for precursor ion mass tolerance of 10 ppm, fragment ion mass tolerance of 0.6 Da, trypsin digestion, and two missed cleavages. Dynamic modification sets were carbamidomethylation of cysteine and oxidation of methionine. Common contaminants such as keratin and trypsin were eliminated, and only peptides with a false discovery rate <1% were considered.

Western blot

The MS results were verified by Western blot analysis of proteins isolated using NNCI, NNCI antisense, and IRES-GFP-Puro transcripts. Antibodies were used to test the enrichment of Hnrnpab (rabbit, 1:100; Sigma Aldrich), Hnrnpa2b1 (mouse, 1:200; Santa Cruz Biotechnology), and Hnrnpd (rabbit, 1:1000; Millipore).

In vitro lentiviral shRNA knockdown

Lentiviral work was performed as described previously (Herriges et al. 2014) with the exception that shRNA oligos (Supplemental Table 5) were annealed into the pLKO.1-Neo vector generously provided by the Winslow laboratory. MLE12 cells were cultured for 6 d in medium containing 50 μ g/mL geneticin to select for infected cells before collection and isolation of total RNA.

Pulmonary and thyroid histology

Postnatal lungs were inflated with 2% paraformaldehyde under constant pressure of 25 cm (P4) or 30 cm (4 wk, 10 wk, or 6 mo) water and allowed to fix overnight. E18.5 lungs and dissected thyroids were fixed overnight in 2% paraformaldehyde. Tissue was embedded in paraffin and sectioned. Hematoxylin and eosin staining was performed to examine tissue morphology. In situ hybridization was performed as described previously (Herriges et al. 2014). IHC was used to detect protein expression using the following antibodies on paraffin sections: cleaved Caspase 3 (rabbit, 1:25; Biocare), E-cadherin (rabbit, 1:100; Cell Signaling), F4/80 (rat, 1:50; Biologend), GFP (goat, 1:100; Abcam), Hnf4a (goat, 1:40; Santa Cruz Biotechnology), Hnf4a (rabbit, 1:50; Cell Signaling), Ki67 (rabbit, 1:50; Abcam), Krt5 (rabbit, 1:1500; Covance), Muc5ac (mouse, 1:100; Abcam), Nkx2.1 (rabbit, 1:50; Santa Cruz Biotechnology), Pgp9.5 (mouse, 1:100; RDI), Pdpn (mouse, 1:50; Hybridoma Bank), P63 (mouse, 1:50; Santa Cruz Biotechnology), RFP (rabbit, 1:250; Rockland), Scgb1a1 (goat, 1:20; Santa Cruz Biotechnology), Sftpc (goat, 1:50; Santa Cruz Biotechnology), and Sox9 (rabbit, 1:100; Santa Cruz Biotechnology).

Forebrain histology

Embryonic mice were dissected on ice-cold PBS and fixed overnight at 4°C in RNase-free 4% PFA in 1 \times PBS. Embryos were then washed twice in 1 \times PBS and cryoprotected by immersion through sucrose (15% and 30% sucrose in 1 \times PBS; solutions were changed once embryos had sunk). Afterward, they were embedded in freezing compound, sectioned coronally at 12 μ m on a cryostat (Leica), and collected onto slides (Superfrost slides, Fisher Scientific). After allowing the sections to dry for 30 min, brain sections were preincubated for 1 h in blocking buffer (5% bovine serum albumin with 0.1% Triton X-100). Sections were then incubated overnight at 4°C with the following primary antibodies: mouse Nkx2.1 (mouse, 1:200; Abcam) and RFP (rabbit, 1:1000; Rockland).

Statistical analysis

Statistical analysis was performed on the data using Prism 7. Two-tailed Student's *t*-test was used for comparison of two experimental groups. Statistical data were considered significant if $P < 0.05$.

Acknowledgments

E.E.M. received funding from the National Institutes of Health (HL122993, HL087825, and HL132349). M.J.H. received funding from National Institutes of Health 5 F31 HL129683-02 and American Heart Association Predoctoral Fellowship 15PRE25080178. Y.H. and B.A.G. received funding from US Department of Defence grant W81XWH-113-1-0426.

References

- Akerman I, Tu Z, Beucher A, Rolando DM, Sauty-Colace C, Benazra M, Nakic N, Yang J, Wang H, Pasquali L, et al. 2016. Human pancreatic β cell lncRNAs control cell-specific regulatory networks. *Cell Metab* **25**: 400–411.
- Akindahunsi AA, Bandiera A, Manzini G. 2005. Vertebrate 2xRBD hnRNP proteins: a comparative analysis of genome, mRNA and protein sequences. *Comput Biol Chem* **29**: 13–23.
- Alder JK, Barkauskas CE, Limjunyawong N, Stanley SE, Kembou F, Tuder RM, Hogan BL, Mitzner W, Armanios M. 2015. Telomere dysfunction causes alveolar stem cell failure. *Proc Natl Acad Sci* **112**: 5099–5104.
- Anderson KM, Anderson DM, McAnally JR, Shelton JM, Bassel-Duby R, Olson EN. 2016. Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. *Nature* **539**: 433–436.
- Brewer G. 1991. An A + U-rich element RNA-binding factor regulates c-myc mRNA stability in vitro. *Mol Cell Biol* **11**: 2460–2466.
- Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* **25**: 1915–1927.
- Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL, Byron M, Monks B, Henry-Bezy M, Lawrence JB, et al. 2013. A long noncoding RNA mediates both activation and repression of immune response genes. *Science* **341**: 789–792.
- The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57–74.
- Czaplinski K, Kocher T, Schelder M, Segref A, Wilm M, Mattaj JW. 2005. Identification of 40LoVe, a *Xenopus* hnRNP D family protein involved in localizing a TGF- β -related mRNA during oogenesis. *Dev Cell* **8**: 505–515.
- Dang VT, Kassahn KS, Marcos AE, Ragan MA. 2008. Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. *Eur J Hum Genet* **16**: 1350–1357.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.
- Eissmann M, Gutschner T, Hammerle M, Gunther S, Caudron-Herger M, Gross M, Schirmacher P, Rippe K, Braun T, Zornig M, et al. 2012. Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. *RNA Biol* **9**: 1076–1087.
- Feyder M, Goff LA. 2016. Investigating long noncoding RNAs using animal models. *J Clin Invest* **126**: 2783–2791.
- Frank DB, Peng T, Zepp JA, Snitow M, Vincent TL, Penkala IJ, Cui Z, Herriges MJ, Morley MP, Zhou S, et al. 2016. Emergence of a wave of Wnt signaling that regulates lung alveologenesis by controlling epithelial self-renewal and differentiation. *Cell Rep* **17**: 2312–2325.
- Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**: 1017–1018.
- Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, et al. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**: 223–227.
- Hamvas A, Deterding RR, Wert SE, White FV, Dishop MK, Alfano DN, Halbower AC, Planer B, Stephan MJ, Uchida DA, et al. 2013. Heterogeneous pulmonary phenotypes associated with mutations in the thyroid transcription factor gene NKX2-1. *Chest* **144**: 794–804.
- Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, et al. 2012. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* **22**: 1760–1774.
- Herriges M, Morrisey EE. 2014. Lung development: orchestrating the generation and regeneration of a complex organ. *Development* **141**: 502–513.
- Herriges MJ, Swarr DT, Morley MP, Rathi KS, Peng T, Stewart KM, Morrisey EE. 2014. Long noncoding RNAs are spatially correlated with transcription factors and regulate lung development. *Genes Dev* **28**: 1363–1379.
- Kohtz JD. 2014. Long non-coding RNAs learn the importance of being in vivo. *Front Genet* **5**: 45.
- Kotzin JJ, Spencer SP, McCright SJ, Kumar DB, Collet MA, Mowel WK, Elliott EN, Uyar A, Makiya MA, Dunagin MC, et al. 2016. The long non-coding RNA Morrbid regulates Bim and short-lived myeloid cell lifespan. *Nature* **537**: 239–243.
- Kusakabe T, Kawaguchi A, Hoshi N, Kawaguchi R, Hoshi S, Kimura S. 2006. Thyroid-specific enhancer-binding protein/NKX2.1 is required for the maintenance of ordered architecture and function of the differentiated thyroid. *Mol Endocrinol* **20**: 1796–1809.
- Law CW, Chen Y, Shi W, Smyth GK. 2014. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**: R29.
- Li L, Chang HY. 2014. Physiological roles of long noncoding RNAs: insight from knockout mice. *Trends Cell Biol* **24**: 594–602.
- Li Z, Shen J, Chan MT, Wu WK. 2016. TUG1: a pivotal oncogenic long non-coding RNA of human cancers. *Cell Prolif* **49**: 471–475.
- Longmire TA, Ikonomou L, Hawkins F, Christodoulou C, Cao Y, Jean JC, Kwok LW, Mou H, Rajagopal J, Shen SS, et al. 2012. Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. *Cell Stem Cell* **10**: 398–411.
- Maeda Y, Chen G, Xu Y, Haitchi HM, Du L, Keiser AR, Howarth PH, Davies DE, Holgate ST, Whitsett JA. 2011. Airway epithelial transcription factor NK2 homeobox 1 inhibits mucous cell metaplasia and Th2 inflammation. *Am J Respir Crit Care Med* **184**: 421–429.
- Meller VH, Joshi SS, Deshpande N. 2015. Modulation of chromatin by noncoding RNA. *Annu Rev Genet* **49**: 673–695.
- Minoo P, Hamdan H, Bu D, Warburton D, Stepanik P, deLemos R. 1995. TTF-1 regulates lung epithelial morphogenesis. *Dev Biol* **172**: 694–698.
- Minoo P, Li C, Liu HB, Hamdan H, deLemos R. 1997. TTF-1 is an epithelial morphoregulatory transcriptional factor. *Chest* **111**: 135S–137S.
- Nobrega-Pereira S, Kessaris N, Du T, Kimura S, Anderson SA, Marin O. 2008. Postmitotic Nkx2-1 controls the migration of telencephalic interneurons by direct repression of guidance receptors. *Neuron* **59**: 733–745.
- Overgaard CE, Schlingmann B, Dorsainvil White S, Ward C, Fan X, Swarnakar S, Brown LA, Guidot DM, Koval M. 2015. The relative balance of GM-CSF and TGF- β 1 regulates lung epithelial barrier function. *Am J Physiol Lung Cell Mol Physiol* **308**: L1212–L1223.
- Perry RB, Ulitsky I. 2016. The functions of long noncoding RNAs in development and stem cells. *Development* **143**: 3882–3894.
- Quinn JJ, Chang HY. 2016. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* **17**: 47–62.
- Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, Sanchez-Gomez DB, 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.
- Snyder EL, Watanabe H, Magendantz M, Hoersch S, Chen TA, Wang DG, Crowley D, Whittaker CA, Meyerson M, Kimura S, et al. 2013. Nkx2-1 represses a latent gastric differentiation program in lung adenocarcinoma. *Mol Cell* **50**: 185–199.
- Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY. 2010. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**: 689–693.
- Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. 2011. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**: 1537–1550.
- Wang XY, Demelash A, Kim H, Jensen-Taubman S, Dakir el H, Ozbun L, Birrer MJ, Linnoila RI. 2009. Matrilysin-1 mediates bronchiolization of alveoli, a potential premalignant change in lung cancer. *Am J Pathol* **175**: 592–604.
- Wapinski O, Chang HY. 2011. Long noncoding RNAs and human disease. *Trends Cell Biol* **21**: 354–361.
- White EJ, Matsangos AE, Wilson GM. 2017. AUF1 regulation of coding and noncoding RNA. *Wiley Interdiscip Rev RNA* **8**: e1393.
- Whitsett JA, Alenghat T. 2015. Respiratory epithelial cells orchestrate pulmonary innate immunity. *Nat Immunol* **16**: 27–35.
- Yamaguchi T, Hosono Y, Yanagisawa K, Takahashi T. 2013. NKX2-1/TTF-1: an enigmatic oncogene that functions as a double-edged sword for cancer cell survival and progression. *Cancer Cell* **23**: 718–723.
- Yang L, Lin M, Ruan WJ, Dong LL, Chen EG, Wu XH, Ying KJ. 2012. Nkx2-1: a novel tumor biomarker of lung cancer. *J Zhejiang Univ Sci B* **13**: 855–866.
- Zhang W, Wagner BJ, Ehrenman K, Schaefer AW, DeMaria CT, Crater D, DeHaven K, Long L, Brewer G. 1993. Purification, characterization, and cDNA cloning of an AU-rich element RNA-binding protein, AUF1. *Mol Cell Biol* **13**: 7652–7665.
- Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, Xiao X, Booth CJ, Wu J, Zhang C, et al. 2012. The lincRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep* **2**: 111–123.
- Zhang HM, Liu T, Liu CJ, Song S, Zhang X, Liu W, Jia H, Xue Y, Guo AY. 2015. AnimalTFDB 2.0: a resource for expression, prediction and functional study of animal transcription factors. *Nucleic Acids Res* **43**: D76–D81.
- Zissler UM, Esser-von Bieren J, Jakwerth CA, Chaker AM, Schmidt-Weber CB. 2016. Current and future biomarkers in allergic asthma. *Allergy* **71**: 475–494.