

# Foxa2 and Cdx2 cooperate with Nkx2-1 to inhibit lung adenocarcinoma metastasis

Carman Man-Chung Li,<sup>1</sup> Vasilena Gocheva,<sup>1</sup> Madeleine J. Oudin,<sup>1</sup> Arjun Bhutkar,<sup>1</sup> Shi Yun Wang,<sup>1</sup> Saya R. Date,<sup>1</sup> Sheng Rong Ng,<sup>1</sup> Charles A. Whittaker,<sup>1</sup> Roderick T. Bronson,<sup>2</sup> Eric L. Snyder,<sup>3,4</sup> Frank B. Gertler,<sup>1</sup> and Tyler Jacks<sup>1,5</sup>

<sup>1</sup>David H. Koch Institute for Integrative Cancer Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA; <sup>2</sup>Department of Pathology, Tufts University School of Medicine and Veterinary Medicine, North Grafton, Massachusetts 01536, USA; <sup>3</sup>Department of Pathology, <sup>4</sup>Department of Anatomy, School of Medicine, University of California at San Francisco, San Francisco, California 94143, USA; <sup>5</sup>Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Despite the fact that the majority of lung cancer deaths are due to metastasis, the molecular mechanisms driving metastatic progression are poorly understood. Here, we present evidence that loss of Foxa2 and Cdx2 synergizes with loss of Nkx2-1 to fully activate the metastatic program. These three lineage-specific transcription factors are consistently down-regulated in metastatic cells compared with nonmetastatic cells. Knockdown of these three factors acts synergistically and is sufficient to promote the metastatic potential of nonmetastatic cells to that of naturally arising metastatic cells *in vivo*. Furthermore, silencing of these three transcription factors is sufficient to account for a significant fraction of the gene expression differences between the nonmetastatic and metastatic states in lung adenocarcinoma, including up-regulated expression of the invadopodia component Tks5<sub>long</sub>, the embryonal proto-oncogene Hmga2, and the epithelial-to-mesenchymal mediator Snail. Finally, analyses of tumors from a genetically engineered mouse model and patients show that low expression of Nkx2-1, Foxa2, and Cdx2 strongly correlates with more advanced tumors and worse survival. Our findings reveal that a large part of the complex transcriptional network in metastasis can be controlled by a small number of regulatory nodes that function redundantly, and loss of multiple nodes is required to fully activate the metastatic program.

[*Keywords*: lung adenocarcinoma; metastasis; Nkx2-1; Foxa2; Cdx2; genetically engineered mouse model of cancer]

Supplemental material is available for this article.

Received June 14, 2015; revised version accepted August 13, 2015.

During tumor progression, cancer cells undergo global gene expression alterations through which they acquire the traits that allow them to successfully advance through the multiple steps of the metastatic cascade. These steps include the ability to invade and migrate through surrounding tissues, intravasate into blood vessels, survive in circulation, extravasate at secondary sites, and colonize distant organs (Steeg 2006). A comprehensive understanding of the upstream regulators that orchestrate this metastasis program is lacking.

To understand more fully the molecular mechanisms of tumor progression and metastasis in a well-defined genetic context, our laboratory developed a genetically engineered mouse model of lung adenocarcinoma, a major subtype of lung cancer that is a leading cause of cancer death worldwide. Conditional activation of oncogenic *Kras* and inactivation of *p53* in *Kras*<sup>LSL-G12D/+</sup>; *p53*<sup>fl/fl</sup> (KP) mice by viral delivery of Cre recombinase to lung epithelial cells initiates the development of lung adenocarci-

nomas that closely resemble the pathophysiological features of the human disease, including the capability to metastasize to distant organs (Jackson et al. 2001, 2005). Previously, we found that progression to metastasis in this model was closely associated with decreased expression of the lung lineage transcription factor Nkx2-1, and knockdown of Nkx2-1 in nonmetastatic tumor cells was sufficient to increase their tumor-seeding ability in transplantation experiments (Winslow et al. 2011). Nonetheless, two major lines of evidence indicate that loss of Nkx2-1 alone may not be sufficient for full progression to metastasis. First, knockdown of Nkx2-1 in nonmetastatic lung adenocarcinoma cells does not recapitulate all of the gene expression changes that occur during the transition from a nonmetastatic to a metastatic state (Winslow et al. 2011). Moreover, *Nkx2-1* deletion in KP lung adenocarcinomas was not sufficient to induce the

© 2015 Li 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/>.

Corresponding author: tjacks@mit.edu

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.267393.115>.

metastasis program but instead unmasked a latent gastric differentiation state of the tumor cells (Snyder et al. 2013). These observations indicate that, in addition to Nkx2-1, additional regulatory factors likely exist that govern the program necessary for full acquisition of metastatic potential.

To investigate additional regulators of metastasis, we elected to examine the transcription factors that control the expression of the metastasis mediator Tks5<sub>long</sub>. A critical component of the proteolytic cellular protrusions, invadopodia, Tks5<sub>long</sub> promotes metastasis in a wide variety of cancer types, including lung adenocarcinoma, and its expression is consistently up-regulated in metastatic cells compared with nonmetastatic cells in the KP model (Murphy and Courtneidge 2011; Li et al. 2013). We showed previously that Tks5<sub>long</sub> is critical for promoting invadopodia formation and metastatic progression in transplant and autochthonous mouse models. Moreover, Tks5<sub>long</sub> expression correlates with a more advanced disease stage and poor survival of lung adenocarcinoma patients. Importantly, Tks5<sub>long</sub> is distinct from an invadopodia-inhibiting isoform, Tks5<sub>short</sub>, by the presence of the membrane-binding Phox homology domain and the use of an independent promoter for transcription (Li et al. 2013). A previous study has demonstrated that the protein level of Tks5<sub>long</sub> can be regulated by Src (Cejudo-Martin et al. 2014). However, the transcriptional regulation of Tks5<sub>long</sub> is not well understood.

Here, we explored the transcriptional regulation of Tks5<sub>long</sub> to uncover key regulators of metastasis in lung adenocarcinoma. We identified three transcriptional repressors of Tks5<sub>long</sub>—Nkx2-1, Foxa2, and Cdx2—and subsequently showed that they function collectively as important regulators of a metastasis program in lung adenocarcinoma. While Nkx2-1 and Foxa2 are known for lineage specification and maintenance of the lungs (among other organs), Cdx2 expression is limited to the intestines in normal adult tissues, and its role in lung adenocarcinoma has not been previously explored. Here, we provide evidence that these three transcription factors function cooperatively as critical regulators in suppressing lung adenocarcinoma metastasis.

## Results

### *Nkx2-1, Foxa2, and Cdx2 synergistically suppress the expression of Tks5<sub>long</sub> in nonmetastatic lung adenocarcinoma cells*

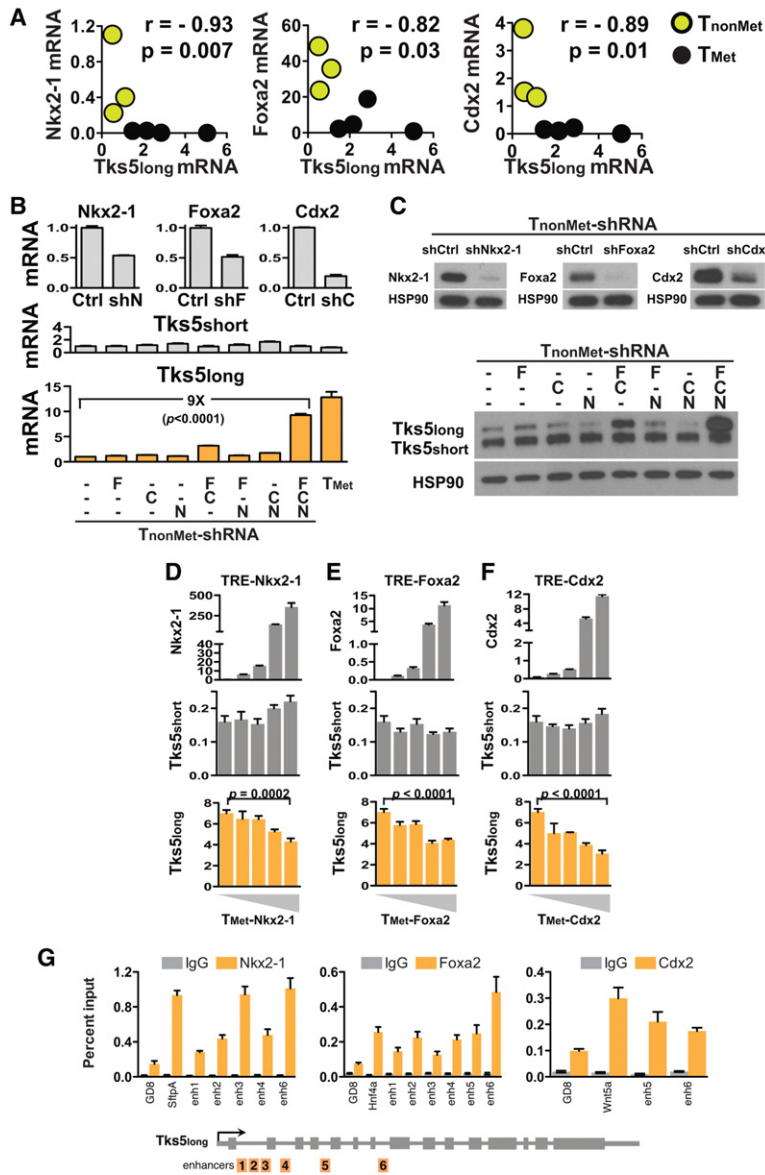
To identify novel mediators of the metastatic program in lung adenocarcinoma, we focused on the transcriptional regulation of Tks5<sub>long</sub>. Tks5<sub>long</sub> has a mechanistically characterized function in promoting metastasis across a wide variety of cancer types, as it mediates the formation of invadopodia, which are proteolytic membrane protrusions that facilitate cellular invasion (Paz et al. 2014). In lung adenocarcinoma, Tks5<sub>long</sub> is critical for promoting metastasis, and increased Tks5<sub>long</sub> expression correlates with poor patient outcomes (Li et al. 2013). Furthermore, Tks5<sub>long</sub> is dramatically and consistently up-regulated in

our collection of metastatic lung adenocarcinoma cells compared with the nonmetastatic cells derived from the KP model (Li et al. 2013), suggesting that its expression is under tight regulation. Importantly, H3K4me3 chromatin immunoprecipitation (ChIP) sequencing (ChIP-seq) analysis showed that Tks5<sub>long</sub> is transcribed from its own promoter independently of the other Tks5 isoform, Tks5<sub>short</sub> (data not shown), suggesting that its increased expression is likely a result of transcriptional regulation and not alternative splicing. Given these data, we hypothesized that the transcriptional regulatory mechanism for Tks5<sub>long</sub> functions as a key switch in regulating a broader metastasis program, which includes many more metastasis-related genes.

To identify potential transcriptional regulators of Tks5<sub>long</sub>, we used an existing gene expression profile of a panel of cell lines derived from nonmetastatic and metastatic primary lung adenocarcinomas as well as metastases in the KP model (termed T<sub>nonMet</sub>, T<sub>Met</sub>, and Met cells, respectively) (Winslow et al. 2011). From this data set, we generated a list of transcription factors that (1) were differentially expressed between the collection of T<sub>nonMet</sub> and T<sub>Met</sub>/Met cells and (2) had predicted binding sites in the Tks5<sub>long</sub> locus based on genomic sequence analysis.

This approach neglects potential nontranscriptional regulatory mechanisms of Tks5<sub>long</sub> expression, but we wanted to start by first focusing strictly on transcription. Among the transcription factors that are differentially expressed between T<sub>nonMet</sub> and T<sub>Met</sub>/Met cells, we identified three that meet these criteria: Nkx2-1, Foxa2, and Cdx2. Microarray gene expression profiling and quantitative RT-PCR (qRT-PCR) validation confirmed that Nkx2-1, Foxa2, and Cdx2 are highly expressed in T<sub>nonMet</sub> cells but are partially or completely lost in T<sub>Met</sub>/Met cells (Fig. 1A; Supplemental Fig. S1A,B). The expression pattern of each transcription factor is inversely proportional to Tks5<sub>long</sub>, suggesting that these three factors may suppress Tks5<sub>long</sub> expression (Fig. 1A).

In order to determine whether Nkx2-1, Foxa2, and Cdx2 inhibit Tks5<sub>long</sub> expression, we knocked down the three transcription factors in T<sub>nonMet</sub> cells using shRNAs (referred to here as T<sub>nonMet</sub>-shNFC cells) either singly or in combination and examined the effect on Tks5<sub>long</sub> expression compared with control knockdown using shRNAs against firefly and renilla luciferases. Interestingly, while single or double knockdown of these transcription factors in T<sub>nonMet</sub> cells only moderately affected Tks5<sub>long</sub> expression, triple knockdown of all three transcription factors in T<sub>nonMet</sub>-shNFC cells led to a dramatic increase in Tks5<sub>long</sub> mRNA and protein levels such that the levels were comparable with that of T<sub>Met</sub> cells with the highest Tks5<sub>long</sub> expression (Fig. 1B,C). The effect of the triple knockdown on Tks5<sub>long</sub> expression was synergistic, as the expression levels of Tks5<sub>long</sub> far exceeded that predicted by the additive effects of single knockdown. Importantly, the effect of Nkx2-1, Foxa2, and Cdx2 knockdown was specific to Tks5<sub>long</sub> and did not affect expression of the other Tks5 isoform, Tks5<sub>short</sub> (Fig. 1B,C). We validated these results in an independent T<sub>nonMet</sub> cell line using different shRNAs (Supplemental Fig. S1C). These data suggest



**Figure 1.** Nkx2-1, Foxa2, and Cdx2 synergistically suppress the expression of *Tks5<sub>long</sub>* in nonmetastatic lung adenocarcinoma cells. (A) mRNA levels of Nkx2-1, Foxa2, and Cdx2 in *T<sub>nonMet</sub>* (368T1, 394T4, and 802T4) and *T<sub>Met</sub>* (373T1, 393T3, 393T5, and 482T1) lung adenocarcinoma cell lines anti-correlate with *Tks5<sub>long</sub>* expression as measured by qRT-PCR. (r) Spearman's correlation coefficient; (p) *P*-value. (B, C) Knockdown of Nkx2-1, Foxa2, and Cdx2 in *T<sub>nonMet</sub>* cells (394T4) derepresses *Tks5<sub>long</sub>* expression but not *Tks5<sub>short</sub>*, as measured by qRT-PCR (B) and immunoblotting (C). Lines (—) indicate control hairpins against firefly or renilla luciferase. Data are represented as mean ± SD. The *P*-value was calculated by Student's *t*-test. (D–F) Overexpression of Nkx2-1, Foxa2, and Cdx2 in *T<sub>Met</sub>* cells (393T5) represses *Tks5<sub>long</sub>* expression but not *Tks5<sub>short</sub>*, as measured by qRT-PCR. Data are represented as mean ± SD. The *P*-values were calculated by Student's *t*-test. (G) ChIP-qPCR analysis of the enrichment of Nkx2-1, Foxa2, and Cdx2 binding at the *Tks5<sub>long</sub>* genomic locus. Data are represented as mean ± SEM of at least three independent experiments. *SftpA*, *Hnf4a*, and *Wnt5a* served as positive controls for Nkx2-1, Foxa2, and Cdx2 binding, respectively. (GD8) Negative control mapping to a gene desert region on murine chromosome 8. For each enhancer versus GD8, *P* < 0.05 by Student's *t*-test.

that Nkx2-1, Foxa2, and Cdx2 may suppress *Tks5<sub>long</sub>* expression in a synergistic manner.

To examine whether these transcription factors are sufficient to suppress *Tks5<sub>long</sub>* expression in metastatic cells, we overexpressed each transcription factor in a doxycycline-inducible manner in *T<sub>Met</sub>* cells. Increasing the levels of each transcription factor inhibited *Tks5<sub>long</sub>* expression in a dosage-dependent manner without affecting *Tks5<sub>short</sub>* mRNA levels (Fig. 1D–F). Furthermore, this suppressive effect on *Tks5<sub>long</sub>* expression was additive, as combined overexpression of Nkx2-1 and Foxa2 (or of Nkx2-1 and Cdx2) reduced the mRNA levels of *Tks5<sub>long</sub>* more significantly than one transcription factor alone (Supplemental Fig. S1D).

To determine whether these transcription factors suppress *Tks5<sub>long</sub>* expression by directly binding to its genetic locus, we performed ChIP-qPCR analysis on potential enhancer regions in the *Tks5<sub>long</sub>* locus. We observed binding

of Nkx2-1, Foxa2, and Cdx2 to multiple enhancers in the *Tks5<sub>long</sub>* locus (Fig. 1G), consistent with a direct role in down-regulating gene expression at the genetic locus.

### *Nkx2-1, Foxa2, and Cdx2 cooperate to inhibit metastasis in vivo*

To test whether Nkx2-1, Foxa2, and Cdx2 suppress metastasis, we transplanted *T<sub>nonMet</sub>*-shNFC cells subcutaneously into nude mice and examined their ability to metastasize from the subcutaneous tumor to the lungs over a period of 6 wk. This assay tests a full range of metastatic properties, as it requires that tumor cells invade and intravasate into circulation at the primary site and then extravasate and colonize a distant organ at the metastatic site. Notably, while *T<sub>nonMet</sub>* cells with Nkx2-1 knockdown (*T<sub>nonMet</sub>*-shN cells) formed more lung nodules than control *T<sub>nonMet</sub>* cells, *T<sub>nonMet</sub>*-shNFC cells were

strikingly more metastatic (Fig. 2A,B). The increase in metastatic potential in  $T_{\text{nonMet}}\text{-shNFC}$  cells was more than the additive effect induced by single or double knockdown. Importantly, the knockdown of these transcription factors did not significantly affect the size of the primary tumors in the subcutaneous site (Supplemental Fig. S2). These data suggest that loss of Nkx2-1, Foxa2, and Cdx2 functions cooperatively to promote metastasis in lung adenocarcinoma.

The increased number of lung metastases seen with  $T_{\text{nonMet}}\text{-shNFC}$  cells could be explained by changes in their ability to complete different steps along the metastatic cascade. We first examined how knockdown of Foxa2, Cdx2, and Nkx2-1 affected cell morphology and motility in vivo. We observed that  $T_{\text{nonMet}}\text{-shNFC}$  subcutaneous tumors adopted a mesenchymal morphology similar to that of  $T_{\text{Met}}$  tumors, in contrast to the predominantly epithelial morphology in tumors formed by  $T_{\text{nonMet}}$  cells and the epithelial/slight mesenchymal

morphology in  $T_{\text{nonMet}}\text{-shN}$  tumors (Fig. 3A,B). This change to mesenchymal morphology for  $T_{\text{nonMet}}\text{-shNFC}$  cells occurred only in vivo but not in vitro (Supplemental Fig. S3A), suggesting that it is at least in part induced by non-cell-autonomous factors present in the tumor microenvironment. Consistent with this change to a mesenchymal morphology, qRT-PCR analysis of  $T_{\text{nonMet}}\text{-shNFC}$  tumors showed loss of the epithelial marker Krt19 and a small increase in expression of mesenchymal markers Twist, Snail, Zeb1, and N-cadherin compared with  $T_{\text{nonMet}}$  tumors but no change in Slug, Vimentin, and E-cadherin (Supplemental Fig. S3B; data not shown). These data suggest a partial epithelial-to-mesenchymal transition (EMT), an important step in the metastatic process (Sato et al. 2012; Tsai and Yang 2013), of  $T_{\text{nonMet}}\text{-shNFC}$  cells in vivo.

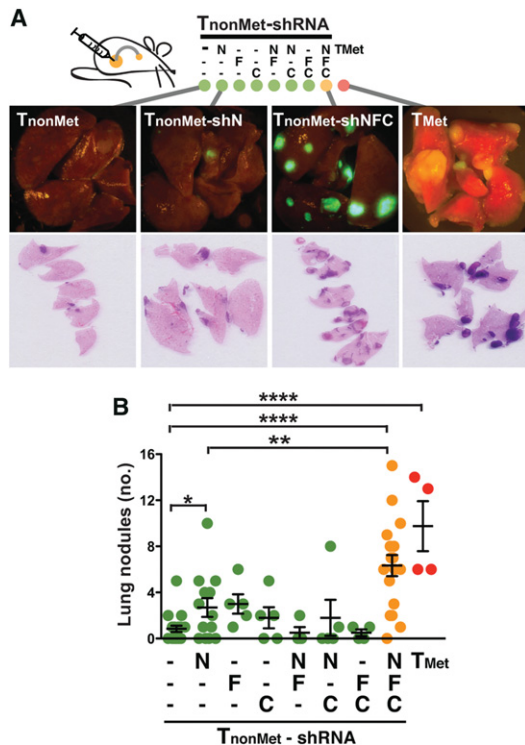
Because mesenchymal morphology is associated with increased motility (Thiery et al. 2009; Tsai and Yang 2013), we performed intravital imaging to monitor migration of GFP-positive cancer cells within subcutaneous tumors.  $T_{\text{nonMet}}\text{-shNFC}$  tumors contained significantly more migratory cells than  $T_{\text{nonMet}}$  and  $T_{\text{nonMet}}\text{-shN}$  cells (Fig. 3C; Supplemental Movies S1–S6). Furthermore, when we measured the chemotactic ability of the tumor cells by performing in vivo fine-needle collection assay using 10% fetal bovine serum (FBS) as a chemoattractant (Wyckoff et al. 2000), we collected a higher number of GFP-positive cancer cells from  $T_{\text{nonMet}}\text{-shNFC}$  subcutaneous tumors than from  $T_{\text{nonMet}}$  and  $T_{\text{nonMet}}\text{-shN}$  tumors (Fig. 3D). Taken together, these data suggest that  $T_{\text{nonMet}}\text{-shNFC}$  cells are more motile in vivo than  $T_{\text{nonMet}}$  and  $T_{\text{nonMet}}\text{-shN}$  cells.

To test whether the enhanced metastatic ability of  $T_{\text{nonMet}}\text{-shNFC}$  cells can also be explained by differences in the colonization of secondary sites during the metastatic cascade, we injected tumor cells intravenously into immunocompromised mice to test their ability to establish tumor nodules upon arriving at the lung capillaries. We observed that  $T_{\text{nonMet}}\text{-shNFC}$  cells and  $T_{\text{nonMet}}\text{-shN}$  cells had equally high colonization capacities compared with  $T_{\text{nonMet}}$  control (Fig. 3E,F), suggesting that while knockdown of Nkx2-1 enhances metastatic colonization, additional inhibition of Foxa2 and Cdx2 does not further contribute to this property.

Taken together, our data support a model in which Nkx2-1, Foxa2, and Cdx2 inhibit metastasis by acting on different steps of the metastatic cascade. While Nkx2-1 inhibits the extravasation/colonization step towards the end of the metastatic cascade, Nkx2-1, Foxa2, and Cdx2 together inhibit migration and invasion in the early steps of metastasis (Fig. 3G). The gain of metastatic ability in  $T_{\text{nonMet}}\text{-shNFC}$  cells compared with  $T_{\text{nonMet}}\text{-shN}$  cells is correlated with an increase in invasion and migration in the primary tumor.

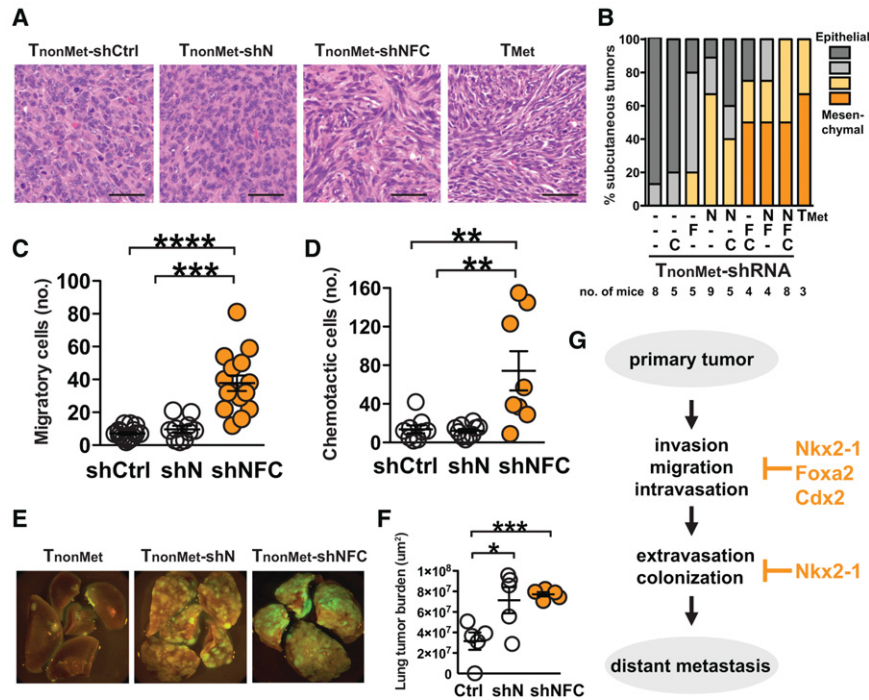
*Nkx2-1, Foxa2, and Cdx2 collectively repress a program of metastasis genes*

Given the increased metastatic ability of  $T_{\text{nonMet}}\text{-shNFC}$  cells compared with  $T_{\text{nonMet}}$  and  $T_{\text{nonMet}}\text{-shN}$  cells, we



**Figure 2.** Foxa2 and Cdx2 synergize with Nkx2-1 in inhibiting metastasis in vivo. (A) Triple knockdown of Nkx2-1, Foxa2, and Cdx2 in a subcutaneous transplantation assay increases the metastatic potential of  $T_{\text{nonMet}}$  cells (394T4)—in comparison with control  $T_{\text{nonMet}}$  cells and  $T_{\text{nonMet}}$  cells with single/double knockdown—to a level similar to that of  $T_{\text{Met}}$  cells (373T1). Representative images of lung metastases are shown. Lines (—) indicate control hairpins against firefly or renilla luciferase. (B) Quantification of metastasis frequencies as determined by numbers of tumor nodules visible on the lung surface. Each circle represents an individual mouse. Data are represented as mean  $\pm$  SEM. The  $P$ -values were calculated by Student's  $t$ -test. (\*)  $P < 0.05$ ; (\*\*)  $P < 0.01$ ; (\*\*\*\*)  $P < 0.0001$ .





**Figure 3.** Foxa2 and Cdx2 synergize with Nkx2-1 in inhibiting migration in vivo. (A,B) Triple knockdown of Nkx2-1, Foxa2, and Cdx2 in T<sub>nonMet</sub> (394T4) subcutaneous tumors induces a mesenchymal morphology similar to T<sub>Met</sub> (373T1) tumors, in contrast to the epithelial T<sub>nonMet</sub>-shCtrl tumors. (A) Representative H&E staining of subcutaneous tumors. Bar, 50  $\mu$ m. (B) Quantification of epithelial and mesenchymal morphology by a pathologist (RT Bronson). Lines (—) indicate control hairpins against firefly or renilla luciferase. (C,D) Knockdown of Nkx2-1, Foxa2, and Cdx2 in T<sub>nonMet</sub> (394T4) subcutaneous tumors enhances migration in vivo compared with T<sub>nonMet</sub>-shCtrl and T<sub>nonMet</sub>-shN tumors, as measured by intravital imaging (C) and fine-needle collection assay (D). Data are represented as mean  $\pm$  SEM. The *P*-values were calculated by Student's *t*-test. (\*\*) *P* < 0.01; (\*\*\*) *P* < 0.001; (\*\*\*\*) *P* < 0.0001. (E,F) T<sub>nonMet</sub>-shNFC and T<sub>nonMet</sub>-shN (394T4) cells show similar colonization ability in the lungs after intravenous transplantation. (E) Representative images of lungs with tumor nodules. (F) Quantification of lung tumor burden. Each circle

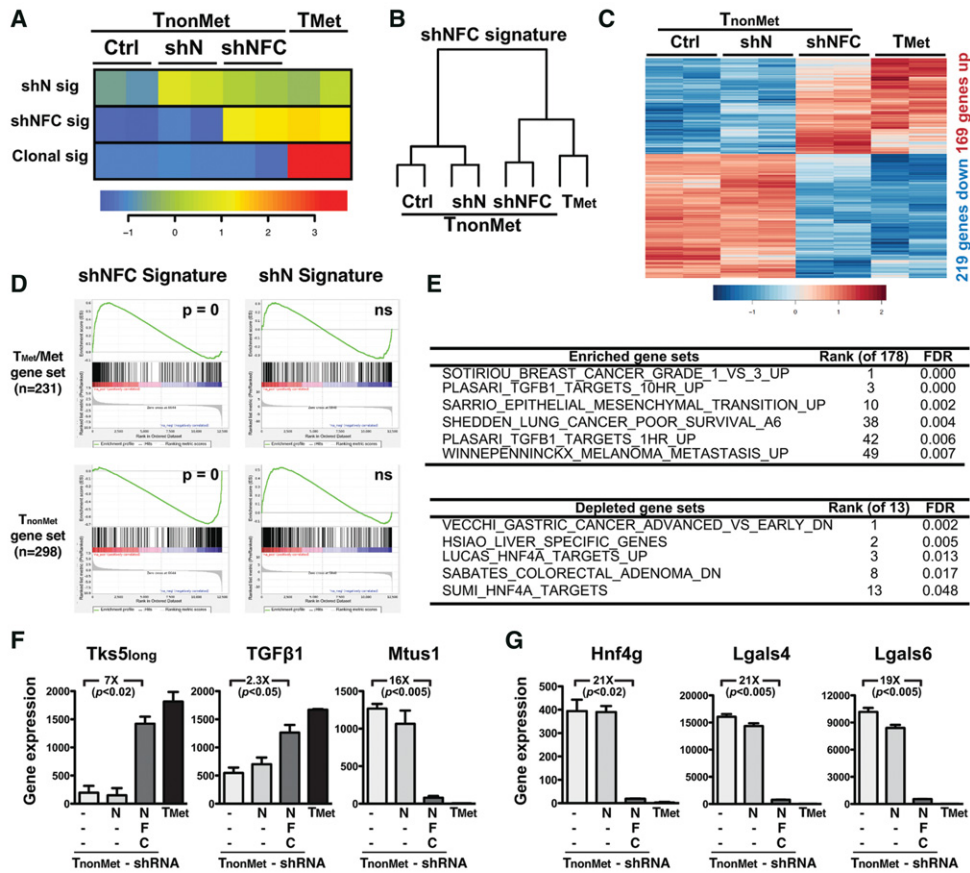
represents an individual mouse. Data are represented as mean  $\pm$  SEM. The *P*-values were calculated by Student's *t*-test. (\*) *P* < 0.05; (\*\*\*) *P* < 0.001. (G) Model for distinct roles of Nkx2-1, Foxa2, and Cdx2 in inhibiting metastatic progression.

next investigated whether a network of metastasis-related genes might be differentially regulated upon knockdown of Nkx2-1, Foxa2, and Cdx2. Transcription factors generally regulate a broad network of target genes, often genes with similar functions. Therefore, we hypothesized that, in addition to suppressing the expression of Tks5<sub>long</sub> (a critical mediator of metastasis), Nkx2-1, Foxa2, and Cdx2 may also regulate the expression of other metastasis-related genes. Thus, we performed RNA sequencing (RNA-seq) on T<sub>nonMet</sub>, T<sub>nonMet</sub>-shN, T<sub>nonMet</sub>-shNFC, and T<sub>Met</sub> cells and used an unsupervised blind source separation strategy using independent component analysis (ICA) (see the Supplemental Material; data not shown). to elucidate statistically independent gene expression signatures that characterize the transcriptomes of these cells. This high-resolution approach allowed us to identify two statistically significant and biologically relevant signatures that are separate from a "clonal signature" that embodies the background identity of T<sub>Met</sub> versus T<sub>nonMet</sub>-derived cells (Fig. 4A): (1) a signature differentiating Nkx2-1-high samples (T<sub>nonMet</sub>) from all Nkx2-1-low samples (T<sub>nonMet</sub>-shN, T<sub>nonMet</sub>-shNFC, and T<sub>Met</sub> cells), which we termed a "shN signature," and (2) a signature clustering T<sub>nonMet</sub>-shNFC/T<sub>Met</sub> cells away from T<sub>nonMet</sub>/T<sub>nonMet</sub>-shN cells, which we designated as the "shNFC signature."

Our analysis provides several lines of evidence that T<sub>nonMet</sub>-shNFC cells significantly recapitulated the major metastasis-related gene expression patterns associated with T<sub>Met</sub> cells. First, hierarchical clustering based on the top and bottom second percentile of genes identi-

fied in the shNFC signature showed clustering of T<sub>nonMet</sub>-shNFC cells with T<sub>Met</sub> cells, and this cluster segregated away from T<sub>nonMet</sub> and T<sub>nonMet</sub>-shN cells (Fig. 4B,C), indicating that the identified gene expression pattern robustly supports the signature. Highly similar clustering patterns were also obtained in analysis of the top and bottom first percentile and fifth percentile of genes in the shNFC signature (data not shown). Second, clustering based on global gene expression profiles after depletion of the background clonal signature showed that T<sub>nonMet</sub>-shNFC cells were substantially more closely related to T<sub>Met</sub> cells than T<sub>nonMet</sub>-shN and T<sub>nonMet</sub> cells (Supplemental Fig. S4A). Finally, the shNFC signature was highly enriched for the "T<sub>Met</sub>/Met" gene set and depleted for the "T<sub>nonMet</sub>" gene set (Winslow et al. 2011) when analyzed using gene set enrichment analysis (GSEA) (Fig. 4D; Subramanian et al. 2005). In contrast, the shN signature did not show a similar enrichment or depletion pattern of T<sub>Met</sub>/Met or T<sub>nonMet</sub> gene sets (Fig. 4D).

Of the 388 genes identified by the shNFC signature, 169 genes were up-regulated (among which is Tks5<sub>long</sub>), and 219 genes were down-regulated in the T<sub>nonMet</sub>-shNFC/T<sub>Met</sub> cells compared with the T<sub>nonMet</sub>/T<sub>nonMet</sub>-shN cells (Fig. 4C). Interestingly, GSEA analyses using publicly available gene sets in the Molecular Signatures Database (Subramanian et al. 2005) revealed that the shNFC signature is significantly enriched for gene sets that represent poor patient prognosis, metastasis/EMT, and TGF $\beta$  targets (Fig. 4E). Furthermore, the shNFC signature is significantly depleted for gene sets related to gastrointestinal/



**Figure 4.** Nkx2-1, Foxa2, and Cdx2 together repress a program of metastasis genes. (A) RNA-seq gene expression analysis of 394T4  $T_{nonMet}$ ,  $T_{nonMet}$ -shN,  $T_{nonMet}$ -shNFC, and 373T1  $T_{Met}$  cells reveals two statistically significant and biologically relevant signatures: an shN signature and an shNFC signature, both separate from a clonal signature that embodies the background identity of  $T_{Met}$ - versus  $T_{nonMet}$ -derived cells. (B) Dendrogram showing sample relationships via clustering based on the top and bottom second percentile of genes in the shNFC signature. (C) Differentially expressed genes that drive the shNFC signature distinguishing  $T_{nonMet}$ -shNFC/ $T_{Met}$  cells from  $T_{nonMet}$ / $T_{nonMet}$ -shN cells. (D) GSEA (gene set enrichment analysis) reveals that the shNFC signature is highly enriched for  $T_{Met}$ /Met genes and depleted for  $T_{nonMet}$  genes, whereas the shN signature does not show similar enrichment. (ns) Not significant. (E) GSEA using the Molecular Signatures Database (MSigDB) curated gene set collection shows that the shNFC signature is enriched for gene sets associated with metastasis and poor prognosis and depleted for gene sets associated with Hnf4a-related gastrointestinal/hepatic differentiation. (F) Examples of prometastatic (*Tks5<sub>long</sub>* and *TGFβ*) and anti-metastatic (*Mtus1*) genes that were identified in the shNFC signature by RNA-seq. Data are represented as mean  $\pm$  SEM. The *P*-values were calculated by Student's *t*-test. (G) Examples of gastrointestinal differentiation genes (*Hnf4g*, *Lgals4*, and *Lgals6*) that were identified in the shNFC signature by RNA-seq. Data are represented as mean  $\pm$  SEM. The *P*-values were calculated by Student's *t*-test.

liver-related genes, reflecting the gene expression changes induced upon knockdown of Foxa2 and Cdx2 (Fig. 4E). Many of the gene expression changes identified in the shNFC signature that were relevant for metastasis or gastrointestinal differentiation were validated by qRT-PCR and immunoblotting (Fig. 4F,G; Supplemental Fig. S4B; data not shown).

To directly answer the question of what fraction of the gene expression differences between our collection of  $T_{nonMet}$  and  $T_{Met}$ /Met cells was recapitulated by combined knockdown of Nkx2-1, Foxa2, and Cdx2 in  $T_{nonMet}$  394T4 cells, we performed targeted pairwise differential analysis. We found that a large fraction (32%) of the genes that were differentially expressed between  $T_{nonMet}$  and  $T_{Met}$ /Met cells also showed significant gene expression alterations by twofold or more in comparing 394T4  $T_{nonMet}$

versus  $T_{nonMet}$ -shNFC cells ( $P = 2.22 \times 10^{-16}$ , hypergeometric test), suggesting that reduced expression of the three transcription factors can explain about one-third of the gene expression changes in metastatic progression. In contrast, only 9% of the genes that were differentially expressed between  $T_{nonMet}$  and  $T_{Met}$ /Met cells were found to be altered when comparing 394T4  $T_{nonMet}$  versus  $T_{nonMet}$ -shN cells ( $P = 2.22 \times 10^{-16}$ , hypergeometric test).

Taken together, these data argue that the  $T_{nonMet}$ -shNFC gene expression program driven by loss of Nkx2-1, Foxa2, and Cdx2 significantly recapitulated the characteristics of  $T_{Met}$  cells. These findings are consistent with data from our in vivo metastasis assays that  $T_{nonMet}$ -shNFC cells are more metastatic than  $T_{nonMet}$ -shN cells and further support that Foxa2 and Cdx2 cooperate with Nkx2-1 to regulate a network of metastasis-related genes.

*The inhibitory effect of Nkx2-1, Foxa2, and Cdx2 on metastasis depends on the activation of Tks5<sub>long</sub>, Hmga2, and Snail expression*

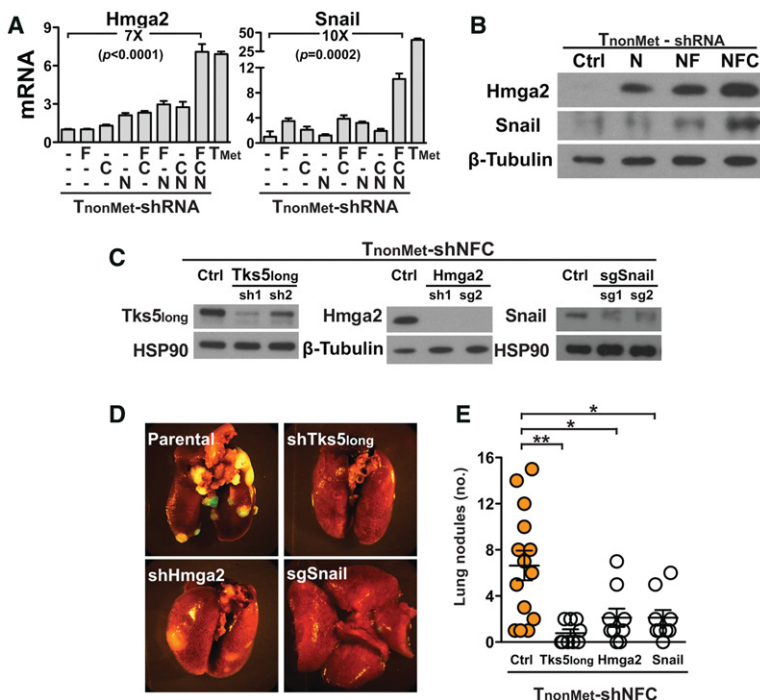
Given the large number of genes activated in the shNFC signature, we then examined whether some of these genes might be functionally required for promoting the metastatic capacity of T<sub>nonMet</sub>-shNFC cells. In addition to Tks5<sub>long</sub>, we elected to examine the embryonal proto-oncogene Hmga2 and the EMT transcription factor Snail, as these metastasis-promoting genes are also significantly up-regulated in T<sub>Met</sub>/Met cells compared with T<sub>nonMet</sub> cells in microarray-based gene expression analysis (Supplemental Fig. S5A; Winslow et al. 2011). T<sub>nonMet</sub>-shNFC cells exhibited a significant increase in Hmga2 and Snail expression at the levels of both mRNA and protein, exceeding the levels in T<sub>nonMet</sub> and T<sub>nonMet</sub> single/double-knockdown cells (Fig. 5A,B). The activation effect was synergistic, as the expression levels of Hmga2 and Snail were much higher than predicted by the additive effects of single knockdown. To test whether increased expression of Tks5<sub>long</sub>, Hmga2, and Snail are required for increased metastatic ability of T<sub>nonMet</sub>-shNFC cells, we knocked down these three genes by shRNAs or CRISPR/Cas sgRNAs (Fig. 5C). Decreased expression of Tks5<sub>long</sub>, Hmga2, or Snail individually impaired the metastatic ability of T<sub>nonMet</sub>-shNFC cells without affecting the size of the primary subcutaneous tumors (Fig. 5D,E; Supplemental Fig. S5E). Importantly, Hmga2 and Snail knockdown did not affect the expression of Tks5<sub>long</sub> or of each other (Supplemental Fig. S5C,D), suggesting that they each contributed individually to increasing metastatic potential. These data support our hypothesis that Nkx2-1, Foxa2, and Cdx2 may function as key regulators for a net-

work of metastasis-related genes that include Tks5<sub>long</sub>, Hmga2, and Snail.

*The endogenous expression pattern of Nkx2-1, Foxa2, and Cdx2 correlates with tumor progression in vivo*

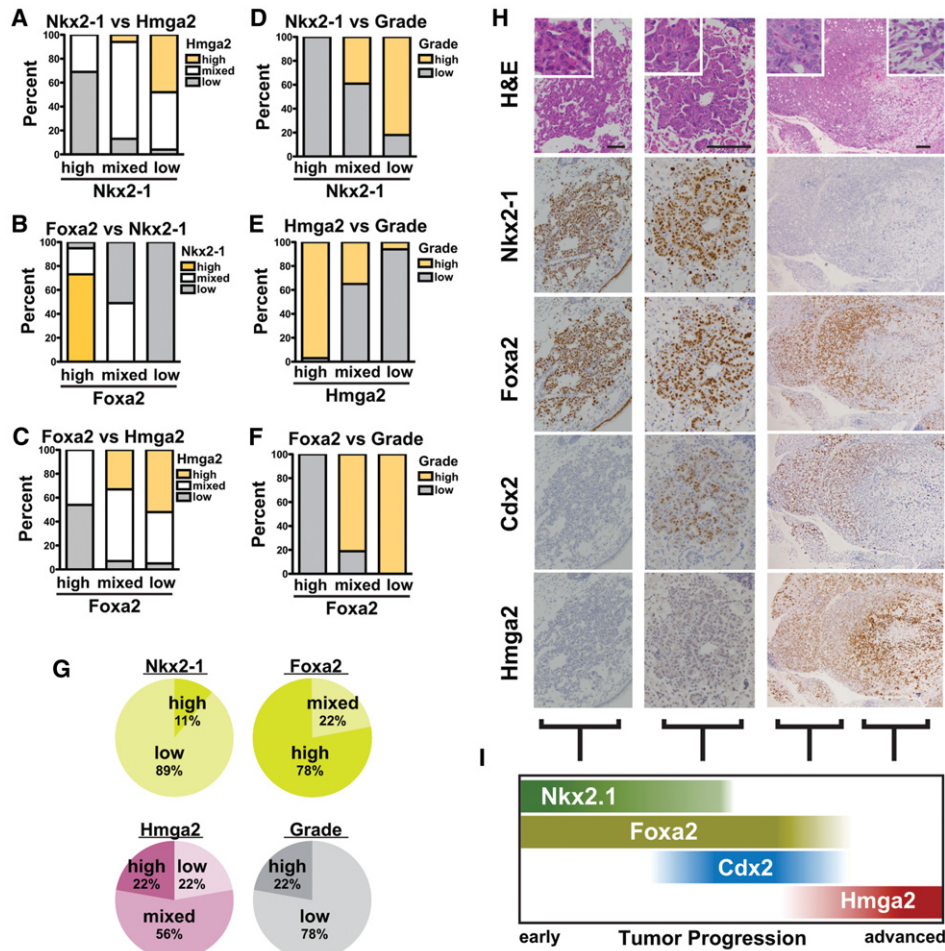
To further characterize the suppressive roles of Nkx2-1, Foxa2, and Cdx2 during tumor progression, we examined their endogenous expression in the KP model of lung adenocarcinoma. This animal model provides a well-defined genetic context and a stereotypic temporal pattern of histologic progression, which facilitate the identification of patterns of gene expression alterations that accompany tumor progression. We analyzed 195 tumor regions from mice ranging from 17 to 33 wk post-initiation and scored them as low grade (grades 1–3) or high grade (grade 4, poorly differentiated) based on nucleocytoplasmic morphology, tumor architecture, and the presence of stromal invasion. Consistent with previous findings (Winslow et al. 2011), expression of Nkx2-1 and Hmga2 anti-correlated with each other in these tumors (Fig. 6A), and Nkx2-1<sup>low</sup> and Hmga2<sup>high</sup> tumors were mostly high grade (Fig. 6D,E).

The pattern of Foxa2 expression was highly similar to that of Nkx2-1 and was anti-correlated with Hmga2 (Fig. 6B,C). The inverse correlation of Foxa2 expression to tumor grade was more striking than Nkx2-1: Foxa2<sup>high</sup> expression was invariably associated with low-grade tumors (117 of 117; 100%), while Foxa2<sup>low</sup> tumors were consistently high grade (21 of 21; 100%) (Fig. 6F). These data suggest that Foxa2, similar to Nkx2-1, marks an early state of tumor progression, and loss of Foxa2 expression is a stringent diagnostic marker of high-grade tumors.



**Figure 5.** The inhibitory effect of Nkx2-1, Foxa2, and Cdx2 on metastasis depends on the activation of Tks5<sub>long</sub>, Hmga2, and Snail expression. (A,B) Combined knockdown of Nkx2-1, Foxa2, and Cdx2 in T<sub>nonMet</sub> cells (394T4) derepresses the expression of Hmga2 and Snail, as analyzed by qRT-PCR (A) and immunoblotting (B). Lines (—) indicate control hairpins against firefly or renilla luciferase. Data are represented as mean ± SD. The P-values were calculated by Student's *t*-test. (C–E) Knockdown of Tks5<sub>long</sub>, Hmga2, or Snail dampens the metastatic ability of 394T4 T<sub>nonMet</sub>-shNFC cells after subcutaneous transplantation. (C) Validation of knockdown by immunoblotting. (D) Representative images of lungs with tumor nodules. (E) Quantification of lung tumor nodules. Each circle represents an individual mouse. Control includes T<sub>nonMet</sub>-shNFC cells (*n* = 10 mice) and T<sub>nonMet</sub>-shNFC-sgRosa cells (*n* = 4 mice); Tks5<sub>long</sub> knockdown was generated by hairpins shTks5<sub>long</sub>#1 (*n* = 5 mice) and shTks5<sub>long</sub>#2 (*n* = 4 mice), Hmga2 knockdown was generated by hairpin shHmga2#1 (*n* = 5 mice) and sgRNA sgHmga2#2 (*n* = 4 mice), and Snail knockdown was generated by sgRNAs sgSnail#1 (*n* = 5 mice) and sgSnail#2 (*n* = 4 mice). Data are represented as mean ± SEM. The P-values were calculated by Student's *t*-test. (\*) *P* < 0.05; (\*\*) *P* < 0.01.





**Figure 6.** The endogenous expression pattern of Nkx2-1, Foxa2, Cdx2, and Hmga2 correlates with tumor progression in an autochthonous model of lung adenocarcinoma. (A–C) Pairwise correlation of Nkx2-1, Foxa2, and Hmga2 expression in KP lung adenocarcinomas. (D–F) Correlation of Nkx2-1, Foxa2, and Hmga2 expression in KP lung adenocarcinomas with tumor grades. Tumor regions were scored as low grade (grades 1–3) or high grade (grade 4, poorly differentiated) based on nucleocytoplasmic morphology, tumor architecture, and the presence of stromal invasion. (G) High expression of Cdx2 in KP lung adenocarcinomas is frequently associated with low expression of Nkx2-1, high expression of Foxa2, low–medium expression of Hmga2, and low-grade histology. (H) Representative H&E and immunohistochemical stainings of KP lung adenocarcinomas. Bar, 150  $\mu$ m. The insets show the nucleocytoplasmic morphology of the tumor cells. (I) Model summarizing expression changes of Nkx2-1, Foxa2, Cdx2, and Hmga2 in lung adenocarcinomas.

Interestingly, loss of Foxa2 expression in high-grade tumors often lagged behind loss of Nkx2-1 (see examples in Fig. 6H, third column; Supplemental Fig. S6A). These observations indicate that loss of Foxa2 expression in tumor progression does not occur concurrently with loss of Nkx2-1 but happens after Nkx2-1 expression is lost (Fig. 6I).

Cdx2 staining was detected in a significant fraction of these lung adenocarcinomas (35 of 195; ~20%), albeit at a lower frequency than the staining of Nkx2-1 (139 of 195; ~70%) and Foxa2 (174 of 195; ~90%). This lower frequency of Cdx2 staining suggests that Cdx2 expression may represent a transient state during tumor progression that is only detectible when a large number of tumors are analyzed. Importantly, the Cdx2<sup>high</sup> tumors showed a consistent pattern that argues that Cdx2 marks an intermediate state of tumor progression that is temporally

situated after loss of Nkx2-1 but before loss of Foxa2 and full activation of Hmga2 (Fig. 6G). First, the vast majority of Cdx2<sup>high</sup> tumor areas were Nkx2-1<sup>low</sup> (eight of nine, 89%), indicating a strong anti-correlation between Cdx2 and Nkx2-1 expression. Second, tumor sections with high Cdx2 levels were typically Foxa2<sup>high</sup> (seven of nine, 87%). In fact, all of the Nkx2-1<sup>low</sup> Foxa2<sup>high</sup> tumor regions were invariably Cdx2<sup>high</sup> (six of six, 100%), and we did not observe any Nkx2-1<sup>low</sup>, Foxa2<sup>low</sup>, and Cdx2<sup>high</sup> regions. Finally, the majority of Cdx2<sup>high</sup> tumor regions were Hmga2<sup>mixed</sup> (five of nine, 56%), whereas only a small fraction was Hmga2<sup>high</sup> (two of nine; 22%) or Hmga2<sup>low</sup> (two of nine; 22%). Even though some of these Cdx2<sup>high</sup> regions were found adjacent to Hmga2<sup>high</sup>, high-grade, and poorly differentiated areas, Cdx2<sup>high</sup> regions themselves were invariably well/moderately differentiated and never part of the poorly differentiated regions (Fig. 6H, third column;



Supplemental Fig. S6A). Collectively, these data strongly suggest that *Cdx2* marks an intermediate state of tumor progression (Fig. 6I).

An alternative model that could explain these observations is that *Cdx2*-expressing tumors represent a “dead end” differentiation state that will never progress to advanced metastatic tumors. However, multiple lines of evidence argue against this model. First, within a single tumor, well/moderately differentiated tumor regions with strong *Cdx2* expression and weak *Hmga2* expression were frequently associated with adjacent poorly differentiated regions that exhibit a reciprocal expression pattern of low *Cdx2* staining and intense *Hmga2* staining (Fig. 6H, third column; Supplemental Fig. S6A). Second, cell line-based experiments showed that knockdown of *Nkx2-1* in  $T_{\text{nonMet}}$  cells derepressed *Cdx2* mRNA and protein levels, while knockdown of *Foxa2* in  $T_{\text{nonMet}}$ -shN cells reduced *Cdx2* levels (Supplemental Fig. S6B,C), suggesting that *Cdx2* expression in these tumor cells is plastic and can be regulated by changes in expression of *Nkx2-1* and *Foxa2*. Furthermore, ChIP-qPCR analysis detected binding of *Nkx2-1* and *Foxa2* to an enhancer downstream from the genomic locus of *Cdx2* (Supplemental Fig. S6D). Based on these findings, we propose a model for the regulation of *Cdx2* expression in lung adenocarcinoma in which transcription of *Cdx2* is inhibited by binding of *Nkx2-1* to a nearby enhancer region. Upon loss of *Nkx2-1*, expression of *Cdx2* is derepressed in a manner dependent on *Foxa2* binding to the same enhancer (Supplemental Fig. S6E).

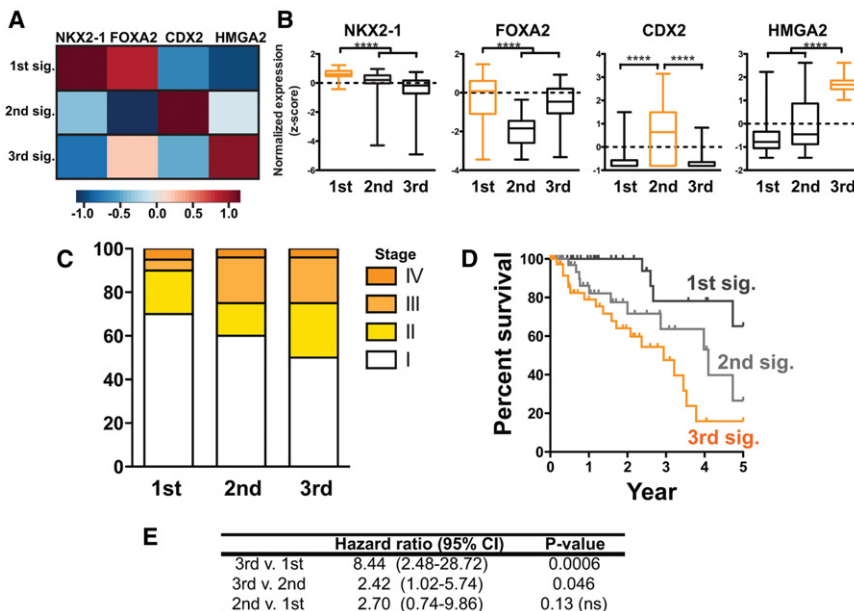
Taken together, these observations suggest that lung adenocarcinoma progress from early tumors that express *Nkx2-1* and *Foxa2* (Fig. 6I). Over time, *Nkx2-1* is silenced, leading to activation of *Cdx2* and partial activation of metastasis-promoting genes such as *Hmga2* in at least a subset of tumors. Finally, suppression of *Foxa2* leads to reduced *Cdx2* expression, and the combined loss of *Nkx2-1*,

*Foxa2*, and *Cdx2* leads to complete derepression of *Hmga2* and other metastasis-promoting genes, resulting in full acquisition of metastatic potential.

*Nkx2-1, Foxa2, and Cdx2 gene expression signatures predict clinical outcomes of lung adenocarcinoma patients*

We next asked whether our observations could provide prognostic information relevant to human lung adenocarcinomas. To this end, we analyzed RNA-seq expression data for 488 lung adenocarcinoma primary tumors from patients with stage I–IV disease from The Cancer Genome Atlas (TCGA, <http://cancergenome.nih.gov>). Unsupervised signature analysis (see the Supplemental Material) of the expression patterns of *NKX2-1*, *FOXA2*, *CDX2*, and *HMGA2* in these tumors revealed three gene expression signatures (Fig. 7A,B). Interestingly, these signatures closely correlated with the expression patterns of *Nkx2-1*, *Foxa2*, *Cdx2*, and *Hmga2* in the KP mouse model of lung adenocarcinoma. The first signature is driven by high expression of *NKX2-1* and *FOXA2* as well as low expression of *HMGA2*, similar to the early tumors in mice. The second signature is characterized by high expression of *CDX2*, similar to tumors in the intermediate state. Finally, the third signature is characterized by high expression of *HMGA2*, similar to advanced tumors in our mouse model. While the gene expression patterns in these signatures are consistent with our observations in the KP model, we note that there is a small increase of *FOXA2* levels in the third signature compared with the second signature. This difference from our KP model is not surprising given the larger heterogeneity and complexity that exist in human lung adenocarcinomas compared with the genetically engineered model.

Importantly, these signatures also strongly correlated with clinical outcomes (Fig. 7C,D). The *NKX2-1*/



**Figure 7.** *NKX2-1*, *FOXA2*, *CDX2*, and *HMGA2* gene expression signatures predict clinical outcomes of lung adenocarcinoma patients. (A) Heat map showing the expression patterns of *NKX2-1*, *FOXA2*, *CDX2*, and *HMGA2* in signatures identified in TCGA human primary lung adenocarcinoma ( $n = 488$ ) using ICA. (B) Box and whisker plots of standardized expression levels of *NKX2-1*, *FOXA2*, *CDX2*, and *HMGA2* in each signature. The horizontal dashed line reflects the mean expression level across all tumors. The  $P$ -values were calculated by Student’s  $t$ -test. (\*\*\*\*)  $P < 0.0001$ . (C–E) Analysis of the top 10th percentile of the 488 lung adenocarcinoma patients in each signature identified by ICA shows that the three signatures correlate with disease stage (C) and overall survival ( $P = 0.0006$  by log-rank test) (D). (E) Multivariate Cox proportional hazard regression analysis of overall survival after adjustments for gender, age, and stage.

FOXA2-driven signature and the CDX2-driven signature were associated with favorable and intermediate disease stage and patient survival, respectively. In contrast, the HMGA2-driven signature was associated with the worst stage and survival. Furthermore, the prognostic values of these four gene pattern-derived signatures were more powerful than analysis using a single gene alone (data not shown). Finally, a multivariate Cox proportional hazard regression analysis showed that these signatures of gene expression pattern were significant prognostic factors independent of gender, age, and disease stage (Fig. 7E; Supplemental Fig. S7A). These results provide further support to our proposed model that early-stage tumors are marked by NKX2-1 and FOXA2 expression, while intermediate tumors acquire expression of CDX2, and most advanced tumors acquire expression of HMGA2. Interestingly, these signatures are prognostic for survival outcome even for patients with early-stage disease (i.e., stages I and II), suggesting that the signatures could be informative for identifying high-risk patients with micro-metastatic disease who may benefit from adjuvant treatment after surgical tumor resection (Supplemental Fig. S7B). Taken together, these data argue that our proposed model for lung adenocarcinoma progression is highly relevant for human cancer. Our findings provide important information for the prognosis of lung adenocarcinoma patients and may inform the future development of therapeutic strategies for this highly metastatic disease.

## Discussion

In this study, we present evidence from the autochthonous KP model of lung adenocarcinoma, derivative cell lines, and human patients that together demonstrate that Nkx2-1, Foxa2, and Cdx2 function collectively to suppress metastatic progression of lung adenocarcinoma. We found that loss of Nkx2-1, Foxa2, and Cdx2 profoundly enhanced metastasis in a subcutaneous transplantation model compared with loss of a single factor alone. In particular, these three factors regulated metastasis by acting on multiple steps of the metastatic cascade: Loss of Nkx2-1 alone promoted colonization at distant sites, whereas combined loss of Nkx2-1, Foxa2, and Cdx2 cooperated to promote tumor cell migration. The observed increase in metastatic ability upon loss of Nkx2-1, Foxa2, and Cdx2 can be explained by activation of a network of transcriptional targets that account for a large and significant fraction of the gene expression alterations between  $T_{\text{nonMet}}$  and  $T_{\text{Met/Met}}$  cells. These target genes include the invadopodia component *Tks5<sub>long</sub>*, the embryonal proto-oncogene *Hmga2*, and the EMT mediator *Snail*, all of which functionally contribute to metastasis. Furthermore, analysis of autochthonous lung tumors in the KP model revealed a strong correlation between loss of expression of these three factors and tumor progression to an advanced, poorly differentiated, and metastatic state. Finally, in human lung adenocarcinoma, the expression patterns of NKX2-1, FOXA2, CDX2, and HMGA2 predicted tumor differentiation states that significantly correlat-

ed with disease stage and survival outcome. Thus, our results are highly relevant for the human disease and provide prognostic information for lung adenocarcinoma patients.

Combining these data, we propose a model for lung adenocarcinoma progression (Fig. 6I). This model starts with high Nkx2-1 and Foxa2 expression that restrains tumors in a well-differentiated, nonmetastatic state. Subsequent loss of Nkx2-1 expression, presumably as a result of extracellular stimulatory signals or intracellular stochastic alterations in gene expression, leads to activation of Cdx2, shifting the cells to an aberrant, albeit still differentiated, state. Finally, suppression of Foxa2 leads to loss of Cdx2, and these alterations synergize with silencing of Nkx2-1 to induce a dedifferentiated stem-like state in the tumor cells, leading to activation of a metastasis program.

Our findings provide two important conclusions. First, Nkx2-1, Foxa2, and Cdx2 can function as key regulatory nodes in transcriptional regulation of the metastasis program. Our data strongly suggest that a relatively small number of factors could be responsible for the vast amount of gene expression alterations in metastasis. It is very likely that only a subset of the gene expression changes induced by loss of Nkx2-1, Foxa2, and Cdx2 is a direct target of these three factors, while other targets may be indirectly controlled via secondary transcription factors downstream from Nkx2-1, Foxa2, and Cdx2. The fact that suppression of Nkx2-1, Foxa2, and Cdx2 is sufficient to derepress this network of metastasis-related genes also suggests that the transcriptional activators for a subset of these target genes may be expressed in tumors and therefore available to act upon loss of these three transcriptional suppressors. As such, while there are multiple pathways regulating the metastasis program, Nkx2-1, Foxa2, and Cdx2 function as three central regulatory nodes in governing transcriptional programs that together restrain tumor metastasis.

Second, the roles of Nkx2-1, Foxa2, and Cdx2 in regulating metastasis underscore the intricate links between dysregulated differentiation and metastasis progression. Striking overlaps exist between genes involved in development and cancer (Kho et al. 2004; Liu et al. 2006; Ben-Porath et al. 2008; Kopantzev et al. 2008). All three transcription factors in this study are well-established developmental regulators. Nkx2-1 and Foxa2 are important developmental regulators of the lungs as well as other endoderm-derived organs (Kimura et al. 1996; Zhou et al. 1997; Minoo et al. 1999; Wan et al. 2005). Nkx2-1 is a homeodomain-containing transcription factor essential for differentiation of the lungs during early embryogenesis (Kimura et al. 1996). In the lungs, Nkx2-1 is expressed in all epithelial cells in early pulmonary development but becomes progressively restricted to alveolar type II and Club cells in adults (Minoo et al. 1999), where Nkx2-1 activates expression of pulmonary-specific genes, including *SftpA*, *SftpB*, *SftpC*, and *CCSP* (Minoo et al. 1999). Foxa2 is a forkhead transcription factor that is expressed in the endoderm and cooperates with its paralog, Foxa1, in mediating organogenesis of the lungs as well as the stomach, intestine, liver, pancreas, and other organs (Kaestner

2010). Foxa2 is important for alveolarization of the lungs during development (Wan et al. 2004). Adult lungs express Foxa2 in the bronchiolar epithelium and alveolar type II cells (Besnard et al. 2004). Interestingly, while Foxa2 expression and function overlap largely with Foxa1, we did not see differential expression of Foxa1 between  $T_{\text{nonMet}}$  and  $T_{\text{Met}}/\text{Met}$  cells (Supplemental Fig. S8). In contrast to Nkx2-1 and Foxa2, the expression of Cdx2 is not appreciably expressed in normal embryonic or adult lungs. Instead, as a member of the Caudal-type homeobox protein, Cdx2 is required for intestine morphogenesis during embryonic development and is expressed in the small and large intestines in adults (Beck and Stringer 2010). Cdx2 also functions in an earlier stage in development for trophoblast formation and axial patterning (Beck and Stringer 2010). While detection of Cdx2 expression in the KP lung adenocarcinoma is perhaps surprising, it is unlikely to be an artifact of the mouse model, as other studies on human lung adenocarcinomas have reported expression of Cdx2 in a subset of patients, although the role of Cdx2 in this context was not previously understood (Yatabe et al. 2004; Grimminger et al. 2009).

Our findings reflect the redundant nature of the cellular differentiation program in impeding tumor progression. The lungs and the intestines are developmentally related, as they are both derived from the developing gut tube. We showed that, during progression of lung adenocarcinoma, partial loss of the lung differentiation program by silencing of Nkx2-1 can lead to aberrant activation of an alternative differentiation program of the intestine that is driven by Cdx2 in at least a subset of these tumors. The activation of the latent intestinal program may serve as a redundant mechanism in the cells to restrain tumor dedifferentiation and metastatic progression. Subsequent loss of both the pulmonary and intestinal differentiation programs upon silencing of Nkx2-1, Foxa2, and Cdx2 expression is required for full progression to a cellular state of more primitive differentiation and higher metastatic potential. This hypothesis about redundancy is further corroborated by the results from a previous study in our laboratory, which showed that deletion of Nkx2-1 alone in the KP model of lung adenocarcinoma was not sufficient to promote metastasis (Snyder et al. 2013). In fact, loss of Nkx2-1 expression early at tumor initiation led to activation of a gastric differentiation program driven by Hnf4 $\alpha$ , Foxa1, and Foxa2 in these cancer cells. In this study, loss of Nkx2-1 at a later stage of tumor progression, as opposed to at tumor initiation, activated the expression of Cdx2 instead of Hnf4 $\alpha$ . This difference may reflect the different cellular states of early and late tumors and the existence of multiple possible alternative differentiation states in the progression of the same tumor type. Interestingly, the aberrant activation of Cdx2 has also been identified to induce intestinal differentiation in other tumor types, including gastric cancer, esophageal cancer, nasal adenocarcinoma, pancreatic cancer, and ovarian cancer, and in some cases has been shown to associate with favorable prognosis (Mizoshita et al. 2003; Yuasa 2003; Guo et al. 2004; Matsumoto et al. 2004). These studies, together with our findings, strongly suggest that the activation

of a Cdx2-driven alternative differentiation program in tumors may be a common phenomenon in the evolution of cancer development and may serve as a mechanism to restrain malignant progression.

Our study has shown that the developmental transcription factors Foxa2 and Cdx2 function cooperatively with Nkx2-1 as important regulators in inhibiting metastasis of lung adenocarcinoma. These data provide strong evidence for the important roles of active and latent developmental regulators in restraining the programs of tumor dedifferentiation and metastatic progression. Our findings also demonstrate that while the pathways in the metastasis program are diverse and complex, they can be controlled by a small number of transcription factors that together govern a network of downstream targets to regulate metastasis.

## Materials and methods

### *Autochthonous K-ras<sup>G12D/WT</sup>; p53<sup>-/-</sup> lung tumors and derivative cell lines*

Lung tumors were initiated via intratracheal delivery of Lenti-Cre or Adeno-Cre in *K-ras<sup>LSL-G12D/WT</sup>; p53<sup>flox/flox</sup>* mice as described previously (DuPage et al. 2009). The Massachusetts Institute of Technology Institutional Animal Care and Use Committee approved all animal studies and procedures.

$T_{\text{nonMet}}$ ,  $T_{\text{Met}}$ , and Met cell lines were generated previously using Lenti-Cre-initiated primary tumors and metastases harvested at 6–14 mo post-infection (Winslow et al. 2011). All cell lines were cultured in complete medium (DMEM with 10% FBS, 50 U/mL penicillin, 50 mg/mL streptomycin). Five  $T_{\text{nonMet}}$  cell lines (368T1, 393T1, 394T4, 802T4, and 2557T1), six  $T_{\text{Met}}$  cell lines (373T1, 373T2, 389T2, 393T3, 393T5, and 482T1) and five Met cell lines (373N1, 393N1, 393M1, 482N1, and 482M1) were used for subsequent gene expression analysis and/or functional experiments in this study.

### *Transplantation assays for metastasis*

For subcutaneous transplantation,  $5 \times 10^4$  cells resuspended in 100  $\mu\text{L}$  of PBS were injected under the skin on the hind flank of nude mice. Mice were analyzed 6.5 wk after injection. For intravenous transplantation,  $5 \times 10^4$  cells resuspended in 100  $\mu\text{L}$  of PBS were injected into the lateral tail vein. Animals were analyzed 2.5 wk after injection.

### *Intravital imaging*

Multiphoton imaging of GFP-labeled tumors was performed as described previously (Wyckoff et al. 2011). Briefly, subcutaneous tumors at 5–6 wk after injection were exposed by skin flap surgery performed on anesthetized animals. Tumors were imaged with an Olympus FV1000 multiphoton microscope using a 25 $\times$ , 1.05 NA water immersion objective with correction lens. Thirty-minute time-lapse movies were analyzed for numbers of migratory GFP-positive tumor cells using ImageJ. Three mice were used per condition, with four to seven fields imaged per mouse.

### *In vivo fine-needle collection assay*

The in vivo invasion assay was performed as previously described (Wyckoff et al. 2000). In brief, four to six catheterized



microneedles held in place by micromanipulators were inserted into the primary tumor of an anesthetized mouse. Needles contained a mixture of 10% Matrigel and 0.01 mM EDTA with L-15 medium  $\pm$  10% FBS. After 4 h, the contents of the needle were extruded, and the total number of tumor cells that migrated into each needle was quantified using DAPI. Three mice were used per condition.

## Acknowledgments

We thank Phillip Sharp, David Barbie, Mandar Muzumdar, and Talya Dayton for critical reading of the manuscript. We also thank Hideo Watanabe for assistance with ChIP-PCR, and Monte Winslow, David Feldser, Nadya Dimitrova, Nikhil Joshi, Wen Xue, Thales Papagiannakopoulos, Francisco Sanchez-Rivera, Tuomas Tammela, Kim Mercer, Kim Dorans, and the entire Jacks laboratory for helpful discussions and experimental assistance. We acknowledge the Swanson Biotechnology Center—and especially Jeff Wyckoff, Eliza Vasile, Denise Crowley, Kathleen Cormier, Michael Brown, and Michele Griffin—for technical support. This work was partially supported by the Cancer Center Support Grant (CCSG) P30-CA14051 from the National Cancer Institute, grants from the Howard Hughes Medical Institute and the National Institutes of Health (5-U01-CA84306) to T.J., Department of Defense Breast Cancer Research Program grant W81XWH-12-1-0031 to M.J.O, and funds from the Ludwig Center at Massachusetts Institute of Technology to F.B.G. T.J. is a Howard Hughes Investigator and a Daniel K. Ludwig Scholar.

## References

- Beck F, Stringer EJ. 2010. The role of Cdx genes in the gut and in axial development. *Biochem Soc Trans* **38**: 353–357.
- Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA. 2008. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* **40**: 499–507.
- Besnard V, Wert SE, Hull WM, Whittsett JA. 2004. Immunohistochemical localization of Foxa1 and Foxa2 in mouse embryos and adult tissues. *Gene Expr Patterns* **5**: 193–208.
- Cejudo-Martin P, Yuen A, Vlahovich N, Lock P, Courtneidge SA, Diaz B. 2014. Genetic disruption of the Sh3pxd2a gene reveals an essential role in mouse development and the existence of a novel isoform of Tks5. *PLoS One* **9**: e107674.
- DuPage M, Dooley AL, Jacks T. 2009. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nat Protoc* **4**: 1064–1072.
- Grimminger P, Ling FC, Neiss S, Vallböhmer D, Lurje G, Schneider PM, Hölscher AH, Metzger R, Brabender J. 2009. The role of the homeobox genes BFT and CDX2 in the pathogenesis of non-small cell lung cancer. *Anticancer Res* **29**: 1281–1286.
- Guo RJ, Suh ER, Lynch JP. 2004. The role of Cdx proteins in intestinal development and cancer. *Cancer Biol Ther* **3**: 593–601.
- Jackson E, Willis N, Mercer K, Bronson R, Crowley D, Montoya R, Jacks T, Tuveson D. 2001. Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Gene Dev* **15**: 3243–3248.
- Jackson E, Olive K, Tuveson D, Bronson R, Crowley D, Brown M, Jacks T. 2005. The differential effects of mutant p53 alleles on advanced murine lung cancer. *Cancer Res* **65**: 10280–10288.
- Kaestner KH. 2010. The FoxA factors in organogenesis and differentiation. *Curr Opin Genet Dev* **20**: 527–532.
- Kho AT, Zhao Q, Cai Z, Butte AJ, Kim JYH, Pomeroy SL, Rowitch DH, Kohane IS. 2004. Conserved mechanisms across development and tumorigenesis revealed by a mouse development perspective of human cancers. *Gene Dev* **18**: 629–640.
- Kimura S, Hara Y, Pineau T, Fernandez-Salguero P, Fox CH, Ward JM, Gonzalez FJ. 1996. The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. *Gene Dev* **10**: 60–69.
- Kopantzev EP, Monastyrskaya GS, Vinogradova TV, Zinovyeva MV, Kostina MB, Filyukova OB, Tonevitsky AG, Sukhikh GT, Sverdlov ED. 2008. Differences in gene expression levels between early and later stages of human lung development are opposite to those between normal lung tissue and non-small lung cell carcinoma. *Lung Cancer* **62**: 23–34.
- Li CM-C, Chen G, Dayton TL, Kim-Kiselak C, Hoersch S, Whittaker CA, Bronson RT, Beer DG, Winslow MM, Jacks T. 2013. Differential Tks5 isoform expression contributes to metastatic invasion of lung adenocarcinoma. *Gene Dev* **27**: 1557–1567.
- Liu H, Kho AT, Kohane IS, Sun Y. 2006. Predicting survival within the lung cancer histopathological hierarchy using a multi-scale genomic model of development. *PLoS Med* **3**: e232.
- Matsumoto K, Mizoshita T, Tsukamoto T, Ogasawara N, Hirata A, Shimizu Y, Haneda M, Yamao K, Tatematsu M. 2004. Cdx2 expression in pancreatic tumors: relationship with prognosis of invasive ductal carcinomas. *Oncol Rep* **12**: 1239–1243.
- Minoo P, Su G, Drum H, Bringas P, Kimura S. 1999. Defects in tracheoesophageal and lung morphogenesis in Nkx2.1<sup>-/-</sup> mouse embryos. *Dev Biol* **209**: 60–71.
- Mizoshita T, Tsukamoto T, Nakanishi H, Inada K-I, Ogasawara N, Joh T, Itoh M, Yamamura Y, Tatematsu M. 2003. Expression of Cdx2 and the phenotype of advanced gastric cancers: relationship with prognosis. *J Cancer Res Clin* **129**: 727–734.
- Murphy DA, Courtneidge SA. 2011. The ‘ins’ and ‘outs’ of podosomes and invadopodia: characteristics, formation and function. *Nat Rev Mol Cell Biol* **12**: 413–426.
- Paz H, Pathak N, Yang J. 2014. Invading one step at a time: the role of invadopodia in tumor metastasis. *Oncogene* **33**: 4193–4202.
- Sato M, Shames DS, Hasegawa Y. 2012. Emerging evidence of epithelial-to-mesenchymal transition in lung carcinogenesis. *Respirology* **17**: 1048–1059.
- 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.
- Steeg PS. 2006. Tumor metastasis: mechanistic insights and clinical challenges. *Nat Med* **12**: 895–904.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* **102**: 15545–15550.
- Thiery JP, Acloque H, Huang RYJ, Nieto MA. 2009. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**: 871–890.
- Tsai JH, Yang J. 2013. Epithelial-mesenchymal plasticity in carcinoma metastasis. *Gene Dev* **27**: 2192–2206.
- Wan H, Kaestner KH, Ang S-L, Ikegami M, Finkelman FD, Stahlman MT, Fulkerson PC, Rothenberg ME, Whittsett JA. 2004.

- Foxa2 regulates alveolarization and goblet cell hyperplasia. *Development* **131**: 953–964.
- Wan H, Dingle S, Xu Y, Besnard V, Kaestner KH, Ang S-L, Wert S, Stahlman MT, Whitsett JA. 2005. Compensatory roles of Foxa1 and Foxa2 during lung morphogenesis. *J Biol Chem* **280**: 13809–13816.
- Winslow MM, Dayton TL, Verhaak RGW, Kim-Kiselak C, Snyder EL, Feldser DM, Hubbard DD, Dupage MJ, Whittaker CA, Hoersch S, et al. 2011. Suppression of lung adenocarcinoma progression by Nkx2-1. *Nature* **473**: 101–104.
- Wyckoff JB, Segall JE, Condeelis JS. 2000. The collection of the motile population of cells from a living tumor. *Cancer Res* **60**: 5401–5404.
- Wyckoff J, Gligorijevic B, Entenberg D, Segall J, Condeelis J. 2011. High-resolution multiphoton imaging of tumors in vivo. *Cold Spring Harb Protoc* **2011**: 1167–1184.
- Yatabe Y, Koga T, Mitsudomi T, Takahashi T. 2004. CK20 expression, CDX2 expression, K-ras mutation, and goblet cell morphology in a subset of lung adenocarcinomas. *J Pathol* **203**: 645–652.
- Yuasa Y. 2003. Control of gut differentiation and intestinal-type gastric carcinogenesis. *Nat Rev Cancer* **3**: 592–600.
- Zhou L, Dey CR, Wert SE, Yan C, Costa RH, Whitsett JA. 1997. Hepatocyte nuclear factor-3 $\beta$  limits cellular diversity in the developing respiratory epithelium and alters lung morphogenesis in vivo. *Dev Dyn* **210**: 305–314.