Title

Nfib Promotes Metastasis through a Widespread Increase in Chromatin Accessibility

Journal

Cell, 166(2)

ISSN

0092-8674

Authors

Denny, Sarah K

Yang, Dian

Chuang, Chen-Hua

et al.

Publication Date

2016-07-01

DOI

10.1016/j.cell.2016.05.052

Peer reviewed

HHS Public Access

Author manuscript

Cell. Author manuscript; available in PMC 2017 July 14.

Published in final edited form as:

Cell. 2016 July 14; 166(2): 328–342. doi:10.1016/j.cell.2016.05.052.

Nfib promotes Metastasis through a Widespread Increase in Chromatin Accessibility

Sarah K. Denny[1,*], **Dian Yang**[2,*], **Chen-Hua Chuang**[3], **Jennifer J. Brady**[3], **Jing Shan Lim**[2], **Barbara M. Grüner**[3], **Shin-Heng Chiou**[3], **Alicia N. Schep**[3], **Jessika Baral**[3], **Cécile Hamard**[7], **Martine Antoine**[7], **Marie Wislez**[7], **Christina S. Kong**[4], **Andrew J. Connolly**[4], **Kwon-Sik Park**[8], **Julien Sage**[2,3,5,6,&], **William J. Greenleaf**[1,3,9,&,#], and **Monte M. Winslow**[2,3,4,5,&,#]

[1] Biophysics Program, Stanford University School of Medicine, Stanford, CA 94305, USA.

[2] Cancer Biology Program, Stanford University School of Medicine, Stanford, CA 94305, USA.

[3] Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA.

[4] Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA.

[5] Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA 94305, USA.

[6] Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94305, USA.

[7]Service de Pneumologie, Hôpital Tenon-APHP, Université Paris 6 Pierre et Marie Curie, Paris, France

[8]Department of Microbiology, Immunology and Cancer Biology, University of Virginia School of Medicine, Charlottesville, VA 22908, USA

[9]Department of Applied Physics, Stanford University, Stanford, CA 94305, USA

SUMMARY

Metastases are the main cause of cancer deaths, but the mechanisms underlying metastatic progression remain poorly understood. We isolated pure populations of cancer cells from primary tumors and metastases from a genetically engineered mouse model of human small cell lung

[#]Correspondence: mwinslow@stanford.edu (MMW) or wjg@stanford.edu (WJG).
[*]Co-first author
[&]Co-senior author

AUTHOR CONTRIBUTIONS

Conceptualization, SKD, DY, JS, MMW, WJG; Formal Analysis, SKD, DY, MW, CSK, AJC, JS, MMW, WJG; Investigation, SKD, DY, CHC, JJB, JSL, BMG, SHC, ANS, JB, CH, MA, CSK, AJC; Resources: MW, CSK, KSP, WJG; Writing – Original Draft: SKD, DY, JS, MMW, WJG; Writing – Reviewing and Editing, all authors; Supervision, MW, JS, MMW, and WJG; Project Administration: JS, MMW, and WJG; Funding Acquisition: JS, MMW, and WJG.

ACCESSION NUMBERS

ATAC-seq, ChIP-seq, and RNA-seq sequencing data can be found in the NCBI Gene Expression Omnibus (GEO) under accession number: GSE81258.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, 7 figures, and 2 tables and can be found with this article online.

cancer (SCLC) to investigate the mechanisms that drive the metastatic spread of this lethal cancer. Genome-wide characterization of chromatin accessibility revealed the opening of large numbers of distal regulatory elements across the genome during metastatic progression. These changes correlate with copy number amplification of the *Nfib* locus, and differentially accessible sites were highly enriched for Nfib transcription factor binding sites. Nfib is necessary and sufficient to increase chromatin accessibility at a large subset of the intergenic regions. Nfib promotes pro-metastatic neuronal gene expression programs and drives the metastatic ability of SCLC cells. The identification of widespread chromatin changes during SCLC progression reveals an unexpected global reprogramming during metastatic progression.

INTRODUCTION

Phenotypic changes that occur during development and disease progression are driven by gene expression changes that are themselves governed by regulatory states encoded within the nucleoprotein structure of chromatin (Voss & Hager 2014). During development and differentiation, tens of thousands of regulatory elements change from inactive to active states (or vice versa), eliciting a concerted transformation of gene expression programs that control cell phenotypes (Zhu et al. 2013). Numerous targeted methods of probing this landscape, from chromatin immunoprecipitation approaches to assays measuring DNA methylation, have produced insight into dimensions of this regulation (Schones & Zhao 2008). Chromatin accessibility, or the genome-wide accounting of loci that are accessible for transcription factor binding, has been identified as perhaps the single most relevant genomic characteristic correlated with biological activity at a specific locus (Thurman et al. 2012).

Recent work has begun to catalog chromatin state changes between normal and cancer cells, and to define the chromatin landscape of several cancer cell lines (Simon et al. 2014; Stergachis et al. 2013). The phenotypic changes associated with metastasis likely require widespread changes in gene expression programs that drive invasion, migration, dissemination, and colonization (Sethi & Kang 2011). However, the specific regulatory changes driving the transition of primary tumors to cells capable of metastatic spread remain largely unexplored.

Small cell lung cancer (SCLC) is a high-grade neuroendocrine carcinoma that accounts for ~15% of all lung cancers and causes over 200,000 deaths worldwide each year (Kalemkerian et al. 2013). The ability of SCLC cells to leave the primary tumor and establish inoperable metastases is a major cause of death and a serious impediment to successful therapy (van Meerbeeck et al. 2011). Molecular analysis of metastatic progression of human cancer is limited by the difficulty in accessing tumor samples at defined stages. This problem is especially true for SCLC, since patients with metastatic disease rarely undergo surgery. Genetically engineered mouse models of human SCLC recapitulate the genetics, histology, therapeutic response, and highly metastatic nature of the human disease (Meuwissen et al. 2003; Schaffer et al. 2010). These models recapitulate cancer progression in a controlled manner and allow isolation of primary tumors and metastases directly from their native microenvironment.

Here we analyzed SCLC cells from primary tumors and metastases to identify global changes in chromatin accessibility during metastatic progression. We uncovered an unexpectedly dramatic increase in accessibility that occurs during malignant progression. We determined that high expression of a single transcription factor, Nfib, alters chromatin state globally and enacts a program of gene expression that promotes multiple steps of the metastatic cascade.

RESULTS

Identification of two distinct chromatin accessibility landscapes within SCLC

To specifically mark cancer cells, we bred a Cre-reporter allele (*R26*^{mTmG}) into the *Trp53*$^{f/f}$;*Rb1*$^{f/f}$;*p130*$^{f/f}$ mouse model of human SCLC (Muzumdar et al. 2007; Schaffer et al. 2010). Adenoviral-Cre inhalation by *Trp53*$^{f/f}$;*Rb1*$^{f/f}$;*p130*$^{f/f}$;*R26*$^{mT/mG}$ (*TKO-mTmG*) mice initiates cancer development. Primary tumors and metastases are Tomato^{negative}, GFP^{positive}, have histologic features of human SCLC, and express markers of SCLC (Figure 1A-B) (Schaffer et al. 2010). These mice develop multifocal metastases in the liver, one of the most common sites of SCLC metastasis in humans (Figure 1B) (Nakazawa et al. 2012; Meuwissen et al. 2003). We isolated cells from individual primary tumors and metastases by fluorescence-activated cell sorting (FACS; Figure 1C, S1A, and S1B).

We used the assay of transposase-accessible chromatin using sequencing (ATAC-seq) to determine the genome-wide chromatin accessibility landscape in primary tumors and liver metastases from *TKO-mTmG* mice (Buenrostro et al. 2013). We isolated cancer cells from one large primary tumor and one liver macro-metastasis from each of four mice. All samples were enriched for reads at transcription start sites (TSSs) and exhibited the expected periodicity of insert length (Figure S1C-E). Hierarchical clustering based on the correlation of accessibility separated the samples into two groups: one containing the majority of primary tumors and the other containing the majority of metastases (Figure 1D). The first principal component of variation also separated the samples into the same two groups and explained 58% of the variance (Figure S1F-G).

We assessed differential accessibility between these two groups, and discovered that ~24% of all accessible regions were >2-fold more accessible in the predominantly metastatic group (Figure 1E, 1F, S1H, and S2A-B). Conversely, only ~0.5% of peaks were >2-fold more accessible in the primary-tumor-enriched group, thus we defined these chromatin states as “hyper-accessible” and “hypo-accessible.”

To put the magnitude of these chromatin differences in context, we re-analyzed our ATAC-seq data and published DNase-seq data from distinct tissues and cell types (Vierstra et al. 2014). Interestingly, the fraction of peaks that differ between the hypo- and hyper-accessible SCLC cell states is smaller than that detected between lung and liver, but greater than between lymphocyte subsets. The nearly unidirectional increase in accessibility during SCLC progression is also a distinct feature of this cell state change (Figure 1G).

Differentially accessible regions are in gene deserts, evolutionarily conserved, and shared by other cell types

The differentially accessible regions were largely gene distal, with relatively few promoter-proximal regions exhibiting differential accessibility (Figure 2A and S2C). Large genomic domains (on the order of 1 to 10 megabases) were more likely to have differentially open regions than others (over 90% of the newly open peaks were in these domains, which contain less than 40% of constitutively open peaks, $p<1\times10^{-12}$ by Fisher's exact test) (Figure 2B). Domains enriched for differentially accessible peaks were characterized by significantly lower gene density and late replication timing (Figure 2C-D, and S2D). Gene-poor domains are classically thought to be less important in regulating gene expression. However, the differentially open intergenic regions were enriched for evolutionarily conserved sequences, suggesting that a subset of these regions may be regulatory elements that promote SCLC progression (Figure 2E).

To determine whether these differentially accessible regions are common to other cell types or represent a pattern of accessibility unique to metastatic SCLC, we compared our chromatin accessibility data to DNase-seq from other cell types and tissues (Vierstra et al. 2014). The constitutively open regions were not dramatically enriched in any particular cell type (Figure S2E). However, the differentially open regions had greatest overlap (>50%) with open regions in brain tissues (Figure 2E), suggesting that neuroendocrine SCLC tumor cells may co-opt regulatory elements found in neuronal tissue during cancer progression.

Differentially accessible regions are highly enriched for NFI binding motifs

To identify potential drivers of this dramatic difference in accessibility, we determined the enrichment of transcription factor motifs within the differentially accessible regions, compared to those within constitutively open regions. The most highly enriched motif was the binding site for the NFI family of DNA binding factors (Figure 3A), which can bind either a full site of two inverted 5 bp repeats, or a half site with lower affinity (Meisterernst et al. 1988). *De novo* motif enrichment also assembled the NFI half-site motif, and one or more *de novo*-assembled half sites were present in over 90% of the differentially open regions (Figure 3B and S3A). These data suggest that NFI transcription factors might play an important role in the large-scale chromatin accessibility changes observed during SCLC progression.

The hyper-accessible samples are characterized by Nfib copy number amplification

NFI transcription factors (Nfia, Nfib, Nfic and Nfix) are widely expressed and the DNA-binding domain is distinct from other DNA-binding factors (Gronostajski 2000). We quantified DNA copy number based on ATAC-seq reads outside of accessible regions, and found that the hyper-accessible samples specifically had genomic amplification of the *Nfib* locus (Figure 3C) (Dooley et al. 2011). DNA copy number gains at *Mycl1* or any other loci did not correlate with the hypo- or hyper-accessible states (Supplemental Table 1).

Nfib is highly expressed in invasive primary tumors and metastases, and Nfib^{high} cells specifically have the ability to disseminate

To investigate whether high Nfib expression is associated with metastatic progression of SCLC, we performed immunohistochemical (IHC) staining for Nfib on tumors from *TKO-mTmG* mice at different stages of progression. Early neuroendocrine hyperplasias and most primary lung tumors expressed low or undetectable levels of Nfib (Figure 3D-3E, and S3B). Conversely, invasive SCLC within the pulmonary lymphatic vessels, as well as lymph node, liver, and other distant metastases were >95% Nfib^{high} and and these cells had more advanced cytology (Figure 3D-E and S3B). Rare liver macro-metastases were Nfib^{low} suggesting that other mechanisms can drive metastatic ability. This likely explains why one metastasis in our initial ATAC-seq dataset was in the hypo-accessible group.

Despite the presence of Nfib^{pos} invasive primary tumors and Nfib^{pos} metastases, it remained unclear whether only Nfib^{pos} cells can initially disseminate or whether both Nfib^{neg} and Nfib^{pos} cells can disseminate with selection for Nfib^{pos} cells occurring at a later step of the metastatic cascade. We examined Nfib expression at the single-cell level by immunofluorescence on FACS-purified SCLC cells. As expected, SCLC cells from individual primary tumors were either homogeneously Nfib^{neg} or a mixture of Nfib^{pos} and Nfib^{neg} cells, while cancer cells from metastases were nearly homogeneously Nfib^{pos} (Figure 3F). Disseminated tumor cells in the pleural fluid were also almost universally Nfib^{pos}, suggesting that the Nfib^{pos} cell state has the unique ability to overcome the initial hurdles that limit systemic spread.

NFIB is highly expressed in a subset of human SCLC primary tumors and metastases

NFIB is genomically amplified in 5-15% of human primary SCLC tumors, consistent with high *NFIB* RNA expression in a subset of primary SCLC samples and cell lines (Figure S3C-D) (Dooley et al. 2011; George et al. 2015; Peifer et al. 2012; Rudin et al. 2012; Barretina et al. 2012). By IHC, NFIB was moderately to highly expressed in ~50% of primary human SCLCs, as well as in lymph node and distant metastases, with a trend toward higher expression in distant metastases (Figure S3E-F). Patients with stage IV SCLC had tumors with significantly higher NFIB levels than stage I and stage II patients, and patients with higher NFIB had a trend toward shorter overall and progression-free survival (Figure S3G-I).

Genomic regions containing NFI motifs have increased occupancy in hyper-accessible samples

To approach the question of how increased expression of Nfib could lead to widespread opening of genomic loci, we employed transcription factor footprinting, which uses ATAC-seq reads to infer binding based on the steric hindrance between transcription factor complexes and the Tn5 transposase used to fragment the genome (Buenrostro et al. 2013). We detected an extremely clear footprint of transcription factor occupancy around the aggregated NFI full sites in accessible regions, suggesting that the NFI-DNA interaction is long-lived (Sung et al. 2014) (Figure 3G). None of the other transcription factors motifs enriched in the hyper-accessible state showed clear differential footprinting in the hypo- versus hyper-accessible chromatin state (data not shown).

We observed a clear pattern of ATAC-seq fragment sizes around NFI motif sites, allowing us to use this information in addition to insertion frequency to infer NFI occupancy at individual sites (Figure S3J-K). We divided NFI sites based on their similarity to the consensus NFI position weight matrix (a proxy for biochemical affinity) (Benos et al. 2002), and compared the inferred occupancies of sites within each bin for the hypo- and hyper-accessible chromatin states. The hyper-accessible chromatin state was characterized by increased occupancy of sites, particularly at sites of intermediate affinity, supporting a model in which increased expression of Nfib drives binding of less canonical sites (Figure 3H). NFI sites in differentially accessible regions had even higher occupancy than constitutively open sites, suggesting that they are especially amenable to Nfib binding, beyond what would be expected from their motif quality alone (Figure 3H).

Occupancy of NFI depletes nucleosomes and changes local chromatin architecture

Nucleosomes and transcription factors compete for access to DNA such that increased occupancy of NFI sites may deplete nucleosomes, thereby inducing accessibility changes at NFI sites genome-wide (Bell et al. 2011). By inferring nucleosome occupancy using the insert sizes of the ATAC-seq reads (Schep et al. 2015), we found that differentially open regions were more depleted of nucleosomes around NFI sites in hyper-accessible samples (Figure 3I). The amount of nucleosome depletion was dependent on the occupancy of NFI motifs, such that higher affinity NFI sites were associated with greater depletion of nucleosomes, but lower affinity sites were associated with greatest differential nucleosome depletion between the two chromatin states (Figure S3M-N).

These data support a model in which Nfib stably binds DNA and directly competes with nucleosomes to stabilize an accessible chromatin configuration. These results suggest that Nfib upregulation may initiate and maintain chromatin state transitions, ultimately leading to transcriptional changes associated with a gain in metastatic ability.

SCLC cell lines recapitulate the differential chromatin states associated with Nfib expression

To determine whether SCLC cell lines maintain the chromatin landscapes of the in vivo tumors, we assessed Nfib expression and the chromatin state of six murine SCLC cell lines (Figure 4A-C). The two cell lines with highest Nfib protein expression (KP1 and 16T) had *Nfib* genomic amplification (Figure 4A and 4B). Hierarchical clustering and principal component analysis of the chromatin accessibility landscape of the cell lines clearly separated the $Nfib^{high}$ from $Nfib^{low}$ lines (Figure 4C, S4A), and the $Nfib^{high}$ cell lines had increased accessibility relative to the $Nfib^{low}$ cell lines (Figure 4D-E, S4B). The changes in accessibility were highly correlated with those in ex vivo samples, and the differentially open regions were similarly gene distal and enriched for NFI motifs ($r = 0.72$; Figure 4F, S4C, and S4D-E). Thus cell lines faithfully recapitulate the dramatic chromatin accessibility changes found in vivo and further support the role of Nfib in defining their chromatin states.

Increased direct binding of Nfib to differentially accessible sites

We performed chromatin immunoprecipitation (ChIP-seq) for Nfib on three SCLC cell lines, 16T ($Nfib^{high}$), KP1 ($Nfib^{high}$), and KP22 ($Nfib^{low}$) and compared Nfib occupancy between

the Nfib^{high} and Nfib^{low} cell lines (Figure 4G and 4H). Peaks that were differentially occupied had on average lower motif scores than peaks that were constitutively occupied in the Nfib^{high} and Nfib^{low} cell lines, consistent with increased expression of Nfib enabling binding of lower affinity sites (Figure 4I). Differential accessibility correlated with differential ChIP signal (Figure 4J and S4F), confirming that a high fraction of differential accessibility is attributable to differential occupancy by Nfib. Sites with greater Nfib ChIP signal in the Nfib^{high} lines were specifically enriched for many other transcription factor binding motifs, suggesting that other transcription factors may play a role upstream or downstream of Nfib action (Figure S4G-H).

Nfib maintains the hyper-accessible chromatin state

To assess the requirement for sustained Nfib expression to maintain an open chromatin state and metastatic ability, we stably knocked down *Nfib* in two Nfib^{high} SCLC cell lines (Figure 5A and S5A-S5B). ATAC-seq on shControl and shNfib cell lines uncovered extensive changes in chromatin accessibility (Figure 5B and S5C). Regions that close upon *Nfib* knockdown were generally anticorrelated with accessibility changes between hypo- and hyper-accessible ex vivo samples ($r = -0.55$). Indeed, the majority of regions (82%) that were differentially open in the hyper accessible class relied on Nfib to maintain accessibility (Figure 5C).

The regions that decrease in accessibility in *Nfib* knockdown cells were generally gene distal and highly enriched for NFI sites (Figure S5D-E). The large majority (75%) of these regions showed differential Nfib occupancy based on Nfib ChIP-seq, and on average they had increased nucleosome occupancy at NFI motif sites (Figure 5D). These results demonstrate that Nfib maintains chromatin accessibility and local chromatin architecture.

Nfib is required for metastatic ability in vivo

Given the expression of Nfib in invasive SCLC and metastases, as well as the dramatic changes in global open chromatin structure upon *Nfib* knockdown, we tested whether Nfib is required for metastatic ability. Subcutaneous growth of Nfib^{high} SCLC cell lines led to the formation of spontaneous liver metastases. *Nfib* knockdown did not significantly alter subcutaneous tumor growth, however mice with 16TshNfib tumors had 3-fold fewer liver metastases than mice with 16TshControl tumors (Figure 5E-F). Because the selective pressure of tumor growth and the metastatic process can select for cancer cells that have escaped knockdown, we assessed Nfib expression in the subcutaneous tumors and metastases. In 16TshNfib subcutaneous tumors, 10-20% of cells had lost *Nfib* knockdown and were Nfib^{high} suggesting a small but likely meaningful selective advantage for Nfib expressing cells (Figure 5G). Conversely, >60% of the metastases in mice with 16TshNfib tumors expressed Nfib (Figure 5G). This dramatic selection for Nfib expression in metastases further underscores the strong metastatic fitness advantage conferred by Nfib.

Subcutaneous growth of a second Nfib^{high} SCLC cell line (KP1) was likewise unaffected by *Nfib* knockdown, and KP1shNfib tumors also seeded ~2.5-fold fewer liver metastases than KP1shControl tumors (Figure S5G-H). The Nfib^{pos} cells that escaped *Nfib* knockdown outcompeted Nfib^{neg} cells to an even greater extent within KP1shNfib subcutaneous tumors

than 16TshNfib, and the metastases that formed were almost exclusively Nfib^{pos} (Figure S5I).

We further investigated the role of Nfib in conferring metastatic ability using intravenous transplantation of 16TshControl and 16TshNfib cells. 16TshNfib cells seeded fewer metastases than control cells and more than half of the metastases that formed from 16TshNfib cells had escaped *Nfib* knockdown (Figure 5H-J), indicating that Nfib is also required for the later stages of the metastatic process.

Nfib is required for clonal growth and invasive ability of SCLC cells

Nfib knockdown reduced proliferation but had no consistent impact on cell death under standard culture conditions (Figure S5J-Q). Because these SCLC cell lines grow in culture as floating spheres, the reduced proliferation suggested that Nfib may influence anchorage-independent growth. To directly assess clonal growth in anchorage-independent conditions, we plated shControl and shNfib cells in soft agar. *Nfib* knockdown greatly reduced the ability of SCLC cells to expand into colonies under these conditions (Figure 5K and S5R-S). *Nfib* knockdown also greatly reduced migration (Figure 5L and S5T-U). These results indicate that Nfib controls several cellular phenotypes that likely cooperate to drive metastatic ability.

Nfib is sufficient to open a subset of distal regulatory regions

To determine whether Nfib is also sufficient to drive these changes, we generated an Nfib^{low} cell line with doxycycline-inducible *Nfib* expression (KP22-TRE-Nfib; Figure 6A). ATAC-seq on KP22-TRE-Nfib and KP22-TRE-empty cells after doxycycline treatment indicated that Nfib expression was sufficient to rapidly drive chromatin opening of ~1,800 regions (Figure 6B and S6A). These regions were highly enriched for loci that were more open in the hyper-accessible ex vivo samples, as well as for regions that had reduced accessibility in *Nfib*-knockdown cells (Figure 6C and S6A). The regions that open with Nfib expression were largely gene distal, enriched for NFI motifs, and had high overlap with the differential Nfib ChIP signal in the Nfib^{high} versus the Nfib^{low} lines (60% overlap, $p < 1\times10^{-300}$ by hypergeometric test; Figure S6B-C).

Unlike *Nfib* knockdown, which reduced accessibility at many regions, Nfib overexpression increased accessibility at only a subset of sites, which had slightly higher motif scores and slightly higher overlap with putative regulatory elements in other cell types (Figure 6C and S6D-E). The affinity of NFI factors to DNA has been shown to be reduced by up to 300-fold by the presence of nucleosomes (Blomquist et al. 1996), suggesting NFI is not a classically defined “pioneer” transcription factor capable of evicting nucleosomes directly from nucleosome-bound DNA (Zaret & Mango 2016). The sites sensitive to Nfib induction had modest, but above-background accessibility in the Nfib^{low} cell line prior to induction of Nfib expression (Figure 6D-6E). Those sites that were newly open in the hyper-accessible samples but not sensitive to Nfib overexpression were not significantly open relative to background in this cell line (Figure 6E), suggesting that Nfib sensitivity may require a permissive level of accessibility to allow Nfib-mediated changes in the nearby chromatin

landscape. Further supporting this hypothesis, the hypo-accessible samples had modest accessibility above background in the sites that were differentially open (Figure S6F).

To extrapolate our findings to human SCLC, we tested whether NFIB is sufficient to increase chromatin accessibility in human SCLC cells. Expression of NFIB in two NFIB^{low} human SCLC cell lines increased chromatin accessibility at a subset of regions in both cell lines (Figure S6L-M). Consistent with the effect of Nfib in the mouse SCLC cell lines, NFIB expression predominantly led to increased accessibility at distal regulatory regions, which were enriched for NFI motifs (Figure S6M). Human cell lines with NFIB overexpression showed increased occupancy of NFI sites, to an even greater degree than in mouse cell lines (Figure S6N).

Nfib is sufficient to drive clonal growth in vitro and metastatic ability in vivo

To determine whether Nfib expression is also sufficient to increase clonal growth in cell culture, we quantified the ability of KP22 cells with either inducible or constitutive Nfib expression to form colonies in anchorage-independent conditions. In both cases, Nfib expression greatly increased the clonal growth ability of SCLC cells (Figure 6F and S6P). Nfib expression did not consistently alter cell death in vitro but did increase proliferation under standard suspension growth conditions (data not shown). NFIB expression in human SCLC cells also increased cell growth under standard suspension growth conditions as well as in anchorage-independent conditions (Figure S6Q-R).

Consistent with our functional data from *Nfib* knockdown cells, KP22-Nfib cells formed ~3-fold more liver metastases after intravenous transplantation relative to controls (Figure 6G-I). Notably, the few metastases that formed from KP22 control cells had high Nfib expression (data not shown), suggesting a strong selection for Nfib^{high} cells during the post-intravasation steps of the metastatic cascade. Nfib expression did not change subcutaneous tumor growth, confirming that Nfib has a limited effect on SCLC growth in vivo (Figure S6K, data not shown).

Nfib promotes neuronal gene expression programs

Our results establish that Nfib modulates accessibility of many distal, putative regulatory regions of the genome. To identify what gene expression programs enable the metastatic phenotype of Nfib-high tumors, we examined gene expression changes upon Nfib overexpression and knockdown (Figure S7A). Approximately 500 genes were significantly upregulated with Nfib overexpression and a similar number of genes changed after *Nfib* knockdown (Figure S7B-D).

Almost all (92%) of the genes that changed significantly with both overexpression and knockdown changed reciprocally (Figure 7A). The genes that were upregulated with Nfib overexpression and downregulated with *Nfib* knockdown were enriched for gene ontology annotations associated with neuronal function, including axon guidance, synapse organization, and regulation of nervous system development (Figure 7B and S7E).

Gene expression changes are associated with distal regions that become accessible upon Nfib upregulation

Interestingly, we observed Nfib-driven neuronal-associated gene expression programs and a high fraction of Nfib-driven chromatin accessibility changes overlap with putative regulatory elements found in brain tissue (>70%; Figure 7C and S6E). Thus, two distinct datasets (gene expression and chromatin accessibility) link neuronal programs to the metastatic ability of this neuroendocrine cancer.

Nfib appears to drive chromatin accessibility at largely gene-distal regions of the genome. Examining 100 kb windows around genes that were upregulated with Nfib overexpression and downregulated with *Nfib* knockdown uncovered a strong enrichment (>3-fold) for gene-distal regions that increased in accessibility with Nfib overexpression (Figure 7D). Having only an Nfib ChIP peak at the promoter (in any cell line) did not have a significant effect on gene expression with Nfib overexpression (Figure 7E), indicating that direct action at the promoters is not the primary mode of gene regulation by Nfib. Collectively, our data support a model in which Nfib stabilizes the open chromatin state at distal regulatory elements in SCLC cells, thereby inducing a gene expression program related to neuronal-associated processes that drive metastatic ability (Figure 7F and 7G).

DISCUSSION

Nfib drives multiple steps of the metastatic process

Metastasis is a low-probability, multistep process in which cancer cells from primary tumors must invade the local tissue, disseminate, survive in circulation, extravasate into a secondary site, and expand into a tissue-destructive metastasis. The early spread and high frequency of metastasis in SCLC patients could suggest that SCLC cells inherently possess the capacity to metastasize (Hou et al. 2012). In contrast, our work shows that SCLC tumors can gain metastatic ability through a dramatic remodeling of their chromatin state (Figure 7G). Our results highlight the power of combining genome-wide molecular approaches with genetic model systems to uncover hidden variations in cellular states.

Our data demonstrate that a single transcription factor, Nfib, promotes the ability of SCLC cells to perform several of the requisite steps to metastasize, including invasion, dissemination, and clonal growth. We confirmed moderate to high NFIB expression in human SCLC and showed that NFIB overexpression increases anchorage-independent growth and induces chromatin changes in human SCLC cell lines (Dooley et al. 2011). These data suggest that a subset of human SCLC likely use NFIB-driven changes in chromatin state as their route for metastasis, although alternative pathways likely exist. A more complete understanding of all pathways by which human SCLC gains metastatic ability, NFIB-driven or otherwise, will allow better patient stratification and the improved division of SCLC into subtypes based on their pro-metastatic programs.

The stable and highly metastatic SCLC state drives rampant metastatic spread

Across many cancer types, loss of differentiation correlates with aggressiveness and poor patient outcome; however, in SCLC, the link between dedifferentiation and aggressiveness

has remained unclear. Amplification of pluripotency factors drives aggressive SCLC (Rudin et al. 2012; Lin et al. 2012), and inhibition of a histone lysine demethylase, which is known to maintain pluripotency in human embryonic stem cells, delays SCLC growth (Mohammad et al. 2015). Conversely, lineage differentiation pathways driven by Ascl1 and NeuroD1 also promote SCLC tumorigenesis and progression (Augustyn et al. 2014; Osborne et al. 2013). Our data suggest an alternate model, in which a partial transdifferentiation driven by Nfib creates a neuronal program of gene expression that drives metastatic ability (Figure 7B, 7G, and S7E).

Unlike the epithelial-mesenchymal transition, which is proposed to create a transient, highly metastatic state (Yang & Weinberg 2008), Nfib amplification, overexpression, and the resulting chromatin changes create a stable and highly metastatic state that is maintained in metastases. This finding may shed light on why SCLC is so widely metastatic in patients: established Nfib-driven metastases remain in a highly metastatic state and may continuously seed additional secondary metastases (Hou et al. 2012).

Mechanism of targeting Nfib to functional sites

Nfib's inability to bind nucleosome-bound DNA suggests that it does not act as a canonical pioneer factor (Zaret & Mango 2016; Blomquist et al. 1996). Our data suggest that Nfib-sensitive sites exist in a “preset” configuration that, upon upregulation of Nfib, allows binding of these motifs. On average, these preset sites are characterized by modest, but above-background accessibility and an absence of Nfib binding (Figure 6D-E). It is interesting that such large numbers of preset sites exist in SCLC cells, which allow widespread changes in chromatin accessibility upon upregulation of Nfib. The ease with which Nfib can reconfigure the chromatin state of SCLC cells indicates that the initial cancerous state may be prone to conversion towards a metastatic state.

Several mechanisms could establish this preset architecture. Foxa transcription factors are pioneer factors that interact with NFI in other cell types, and Foxa motifs are enriched around the NFI sites that are more bound in the $Nfib^{high}$ cell lines (Figure S4J) (Grabowska et al. 2014). A fraction of sites, therefore, may rely on Foxa factors to maintain a permissive chromatin state around NFI sites, marking them for opening upon Nfib upregulation. Histone turnover in active regulatory regions may also enable sensitivity to Nfib expression, possibly through the action of chromatin remodelers (Hebbar & Archer 2003; Deal et al. 2010).

Molecular mechanism of chromatin opening and gene activation

How does Nfib establish an open chromatin architecture after binding? Nfib gives especially strong “footprinting” signal in our ATAC-seq data, a feature associated with a slow DNA binding off-rate (Sung et al. 2014). Furthermore, the high affinity of NFI to its consensus full site (Meisterernst et al. 1988) implies a long half-life that is similar in magnitude to the longest-lived transcription factors (Sung et al. 2016). This distinguishes Nfib from other transcription factors that are thought to have short residence time on DNA and are insufficient to compete with nucleosomes or other transcription factors (Sung et al. 2016; Voss & Hager 2014). The long residence time of Nfib on DNA likely allows the adaptation

of nearby chromatin and enforces a prototypical chromatin architecture with positioned nucleosomes. This architecture is supported by the depleted nucleosome occupancy and positioned proximal nucleosomes around NFI sites in the hyper-accessible state (Figure 3I). Many additional motifs are enriched around newly open sites, and binding at these motifs may help stabilize an open structure (Figure 3A, S4J, S6B). Our data does not rule out Nfib-induced destabilization of higher-order nucleosome interactions (Alevizopoulos et al. 1995) or recruitment of chromatin remodelers (Hebbar & Archer 2003), but rather suggests that these mechanisms could complement the direct nucleosome depletion and positioning effects of Nfib.

Our data indicate that Nfib establishes a pro-metastatic gene expression program through stabilizing chromatin accessibility at distal regulatory elements. Increased Nfib occupancy in the absence of increased accessibility is only weakly associated with increased gene expression, while Nfib-driven increases in distal accessibility is associated with upregulation of nearby genes (Figure 7D and S7E-F). Nfib has diverse functions and has been shown to both promote cell differentiation during development and maintain populations of stem cells in adult tissues (Harris et al. 2015; Gronostajski 2000; Chang et al. 2013). An implication of our study is that maintenance of open chromatin architecture by Nfib may promote varied functional outcomes through combinatorial binding of nearby transcription factors.

Nfib regulates neuronal guidance and migration pathways

Our data suggest that incorporating features of a neuronal-like program into the SCLC state promotes metastatic ability. Nfib is important for brain development and neuronal migration, and several of the neuronal-associated Nfib-regulated genes and pathways (including Cxcr4, the Eph receptors, Semaphorins, Netrin-1 and Slit/Robo families) have been implicated in invasion and metastasis of other cancer types (Steele-Perkins et al. 2005; Betancourt et al. 2014; Burger et al. 2003; Vanharanta et al. 2013; Chen 2012; Foley et al. 2015; Ko et al. 2014; Zhou et al. 2011). Our results suggest that overlapping components of a neuronal program with the transformed SCLC state can drive the insidious metastatic ability of this cancer type.

Epigenetic machinery has recently emerged as a promising therapeutic target, and preclinical results from modulating components of these pathways have been encouraging (Mohammad et al. 2015). Progression to pro-metastatic epigenetic states may create collateral therapeutic vulnerabilities, or provide direct targets to limit the development of initial metastases or halt the propagation of secondary metastases. Given the dramatic remodeling of the gene regulatory state driven by a single factor in SCLC, we speculate that other major cancer types may also gain metastatic proclivity through large-scale remodeling of their chromatin state.

EXPERIMENTAL PROCEDURES

Mouse model

All experiments were performed in accordance with Stanford University Animal Care and Use Committee guidelines. *Trp53*^{flox}, *Rb1*^{flox}, *p130*^{flox}, and *R26*^{mTmG} mice have been

described (Schaffer, 2010; Muzumdar et al. 2007). Tumors were initiated by inhalation of Adeno-CMV-Cre (University of Iowa Vector Core).

ATAC-seq library preparation and differential accessibility

5×10^4 cells from ex vivo tumors or directly from cell culture. ATAC-seq libraries were generated as described (Buenrostro et al. 2013). Peaks were called on the merged set of all ex vivo ATAC-seq reads using MACS2 (Zhang et al. 2008), and filtered to remove putative copy number altered regions. Number of reads/peak was determined for each sample using bedtools multicov, and the relative sequencing depth was estimated using a set of 'housekeeping' peaks at transcription start sites of genes that were uniformly expressed across *TKO-mTmG* tumors. Differential accessibility was assessed using DESeq2 (Love et al. 2014). Unless otherwise stated, regions were called differentially accessible if the absolute value of the log2 fold change was >0.5 at an FDR <0.1. Visualizations of insertion tracks were smoothed by 150 bp sliding windows (20 bp step size).

Transcription factor occupancy

For each motif site, the distribution of fragments as a function of fragment size and midpoint position relative to the motif center was fit to a mixture model that finds the optimal mixture parameter between two generating distributions: 1) the distribution from an ideal "bound" model modulated by local sequence bias, and 2) the distribution from an "unbound" model in which fragments are distributed based only on the local sequence bias. Local sequence bias was modeled (Schep et al. 2015), and the ideal 'bound' model was found by aggregating reads around motif sites with high motif scores (>7.6) and divided by the expectation given sequence bias at those sites.

Nfib Knockdown and Overexpression

Stable Nfib knockdown cell lines were generated using lentiviral pLKO/PuroR vectors. Knockdown was confirmed by qRT-PCR and western blotting. Inducible and constitutive Nfib expressing human and mouse cells were generated using lentiviral vectors.

Transplantation and cell culture assays

For intravenous transplantation 2×10^4 cells were injected into the lateral tail vein of NOD/SCID/γc (NSG) mice. For subcutaneous injection, 5×10^4 cells were resuspended in 100μl PBS and mixed with 100μl matrigel (Corning, 356231) with 4 injection sites per mouse. Cell culture assays were performed using standard protocols.

RNA-seq expression analysis

RNA-sequencing libraries were prepared using the Illumina TruSeq v2 kit, according to manufacturer's instructions. RNA-seq reads were separately aligned to the mouse genome (mm10) using TopHat. Read counts within merged exons (RefSeq) were found and compared using DESeq2. Size factors were estimated from cqnnorm with exon length and average GC content used as covariates in the normalization.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank Pauline Chu for technical assistance; Sean Dolan and Alexandra Orantes for administrative support; David Feldser, Gerald Crabtree, Namita Bisaria, and members of the Winslow, Greenleaf, and Sage laboratories for helpful comments. We thank the Stanford Shared FACS Facility. This work was supported by a Stanford Cancer Institute Cancer Biology Seed Grant (to MMW, WJG, and JS) and in part by the Stanford Cancer Institute support grant (NIH P30-CA124435), NIH grant P50HG007735, Baxter Foundation Faculty Scholar Grant, the Rita Allen Foundation, and the Human Frontier Science Program. SKD was supported by the Stanford Biophysics training grant (T32 GM008294) and by the NSF GRFP. DY was supported by a Stanford Graduate Fellowship and by a TRDRP Dissertation Award (24DT-0001). C-HC was funded by an American Lung Association Fellowship. JJB was supported by NIH F32-CA189659. BMG is supported by a fellowship from the Hope Funds for Cancer Research (HFCR-15-06-07). J.S is the Harriet and Mary Zelencik Scientist in Children's Cancer and Blood Diseases. WJG is scientific co-founder of Epinomics.

REFERENCES

Alevizopoulos A, et al. A proline-rich TGF-beta-responsive transcriptional activator interacts with histone H3. Genes & development. 1995; 9(24):3051–66. [PubMed: 8543151]

Augustyn A, et al. ASCL1 is a lineage oncogene providing therapeutic targets for high-grade neuroendocrine lung cancers. Proceedings of the National Academy of Sciences of the United States of America. 2014; 111(41):14788–93. [PubMed: 25267614]

Barretina J, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483(7391):603–7. [PubMed: 22460905]

Bell O, et al. Determinants and dynamics of genome accessibility. Nature reviews. Genetics. 2011; 12(8):554–64.

Benos PV, Bulyk ML, Stormo GD. Additivity in protein-DNA interactions: how good an approximation is it? Nucleic acids research. 2002; 30(20):4442–51. [PubMed: 12384591]

Betancourt J, Katzman S, Chen B. Nuclear factor one B regulates neural stem cell differentiation and axonal projection of corticofugal neurons. The Journal of comparative neurology. 2014; 522(1):6–35. [PubMed: 23749646]

Blomquist P, Li Q, Wrange O. The affinity of nuclear factor 1 for its DNA site is drastically reduced by nucleosome organization irrespective of its rotational or translational position. The Journal of biological chemistry. 1996; 271(1):153–9. [PubMed: 8550551]

Buenrostro JD, et al. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature methods. 2013; 10(12):1213–8. [PubMed: 24097267]

Burger M, et al. Functional expression of CXCR4 (CD184) on small-cell lung cancer cells mediates migration, integrin activation, and adhesion to stromal cells. Oncogene. 2003; 22(50):8093–101. [PubMed: 14603250]

Chang C-Y, et al. NFIB is a governor of epithelial-melanocyte stem cell behaviour in a shared niche. Nature. 2013; 495(7439):98–102. [PubMed: 23389444]

Chen J. Regulation of tumor initiation and metastatic progression by Eph receptor tyrosine kinases. Advances in cancer research. 2012; 114:1–20. [PubMed: 22588054]

Deal RB, Henikoff JG, Henikoff S. Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. Science. 2010; 328(5982):1161–4. [PubMed: 20508129]

Dooley AL, et al. Nuclear factor I/B is an oncogene in small cell lung cancer. Genes & development. 2011; 25(14):1470–5. [PubMed: 21764851]

Fidler IJ. The pathogenesis of cancer metastasis: the “seed and soil” hypothesis revisited. Nature reviews. Cancer. 2003; 3(6):453–8.

Foley K, et al. Semaphorin 3D autocrine signaling mediates the metastatic role of annexin A2 in pancreatic cancer. Science signaling. 2015; 8(388):ra77. [PubMed: 26243191]

George J, et al. Comprehensive genomic profiles of small cell lung cancer. Nature. 2015; 524(7563): 47–53. [PubMed: 26168399]

Grabowska MM, et al. NFI Transcription Factors Interact with FOXA1 to Regulate Prostate Specific Gene Expression. Molecular endocrinology. 2014; 28(6):949–64. [PubMed: 24801505]

Gronostajski RM. Roles of the NFI/CTF gene family in transcription and development. Gene. 2000; 249(1-2):31–45. [PubMed: 10831836]

Harris L, et al. Nuclear factor one transcription factors: Divergent functions in developmental versus adult stem cell populations. Developmental dynamics. 2015; 244(3):227–38. [PubMed: 25156673]

Hebbar PB, Archer TK. Nuclear Factor 1 Is Required for Both Hormone-Dependent Chromatin Remodeling and Transcriptional Activation of the Mouse Mammary Tumor Virus Promoter. Molecular and Cellular Biology. 2003; 23(3):887–898. [PubMed: 12529394]

Hou JM, et al. Clinical significance and molecular characteristics of circulating tumor cells and circulating tumor microemboli in patients with small-cell lung cancer. Journal of Clinical Oncology. 2012; 30(5):525–532. [PubMed: 22253462]

Kalemkerian GP, et al. Small cell lung cancer. Journal of the National Comprehensive Cancer Network. 2013; 11(1):78–98. [PubMed: 23307984]

Ko SY, Blatch GL, Dass CR. Netrin-1 as a potential target for metastatic cancer: focus on colorectal cancer. Cancer metastasis reviews. 2014; 33(1):101–13. [PubMed: 24338005]

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014; 15(12):550. [PubMed: 25516281]

van Meerbeeck JP, Fennell DA, De Ruysscher DKM. Small-cell lung cancer. Lancet. 2011; 378(9804): 1741–55. [PubMed: 21565397]

Meisterernst M, et al. A quantitative analysis of nuclear factor I/DNA interactions. Nucleic acids research. 1988; 16(10):4419–35. [PubMed: 3380685]

Meuwissen R, et al. Induction of small cell lung cancer by somatic inactivation of both Trp53 and Rb1 in a conditional mouse model. Cancer Cell. 2003; 4(3):181–189. [PubMed: 14522252]

Mohammad HP, et al. A DNA Hypomethylation Signature Predicts Antitumor Activity of LSD1 Inhibitors in SCLC. Cancer cell. 2015; 28(1):57–69. [PubMed: 26175415]

Muzumdar MD, et al. A global double-fluorescent Cre reporter mouse. Genesis. 2007; 45(9):593–605. [PubMed: 17868096]

Nakazawa K, et al. Specific organ metastases and survival in small cell lung cancer. Oncology Letters. 2012; 4(4):617–620. [PubMed: 23205072]

Osborne JK, et al. NeuroD1 regulates survival and migration of neuroendocrine lung carcinomas via signaling molecules TrkB and NCAM. Proceedings of the National Academy of Sciences. 2013; 110(16):6524–9.

Peifer M, et al. Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nature genetics. 2012; 44(10):1104–10. [PubMed: 22941188]

Rudin CM, et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nature genetics. 2012; 44(10):1111–6. [PubMed: 22941189]

Schaffer BE, et al. Loss of p130 accelerates tumor development in a mouse model for human small-cell lung carcinoma. Cancer research. 2010; 70(10):3877–83. [PubMed: 20406986]

Schep AN, et al. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. Genome research. 2015; 25(11):1757–70. [PubMed: 26314830]

Schones DE, Zhao K. Genome-wide approaches to studying chromatin modifications. Nature reviews. Genetics. 2008; 9(3):179–91.

Sethi N, Kang Y. Unravelling the complexity of metastasis - molecular understanding and targeted therapies. Nature reviews. Cancer. 2011; 11(10):735–48. [PubMed: 21941285]

Simon JM, et al. Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. Genome research. 2014; 24(2): 241–50. [PubMed: 24158655]

Steele-Perkins G, et al. The transcription factor gene Nfib is essential for both lung maturation and brain development. Molecular and cellular biology. 2005; 25(2):685–98. [PubMed: 15632069]

Stergachis AB, et al. Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. Cell. 2013; 154(4):888–903. [PubMed: 23953118]

Sung M-H, et al. DNase footprint signatures are dictated by factor dynamics and DNA sequence. Molecular cell. 2014; 56(2):275–85. [PubMed: 25242143]

Sung M-H, Baek S, Hager GL. Genome-wide footprinting: ready for prime time? Nature methods. 2016; 13(3):222–228. [PubMed: 26914206]

Thurman RE, et al. The accessible chromatin landscape of the human genome. Nature. 2012; 489(7414):75–82. [PubMed: 22955617]

Vanharanta S, et al. Epigenetic expansion of VHL-HIF signal output drives multiorgan metastasis in renal cancer. Nature medicine. 2013; 19(1):50–6.

Vierstra J, et al. Mouse regulatory DNA landscapes reveal global principles of cis-regulatory evolution. Science. 2014; 346(6212):1007–12. [PubMed: 25411453]

Voss TC, Hager GL. Dynamic regulation of transcriptional states by chromatin and transcription factors. Nature reviews. Genetics. 2014; 15(2):69–81.

Yang J, Weinberg RA. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Developmental cell. 2008; 14(6):818–29. [PubMed: 18539112]

Zaret KS, Mango SE. Pioneer transcription factors, chromatin dynamics, and cell fate control. Current opinion in genetics & development. 2016; 37:76–81. [PubMed: 26826681]

Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). Genome biology. 2008; 9(9):R137. [PubMed: 18798982]

Zhou W-J, et al. Slit-Robo signaling induces malignant transformation through Hakai-mediated E-cadherin degradation during colorectal epithelial cell carcinogenesis. Cell research. 2011; 21(4): 609–26. [PubMed: 21283129]

Zhu J, et al. Genome-wide Chromatin State Transitions Associated with Developmental and Environmental Cues. Cell. 2013; 152(3):642–54. [PubMed: 23333102]

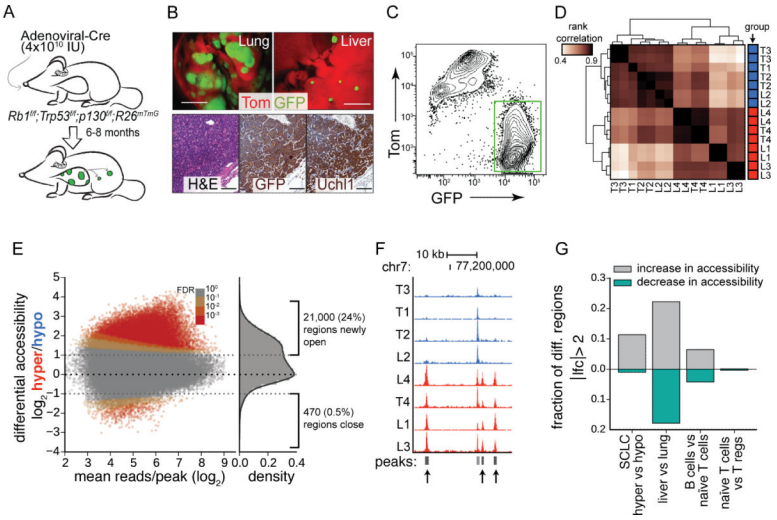

Figure 1. Small cell lung cancer exists in two distinct chromatin accessibility states

(A) Genetically engineered mouse model of SCLC.

(B) GFP^{positive} primary tumors and metastases within Tomato (Tom)^{positive} normal tissues (top). GFP and Uchl1 IHC and H&E are shown. Top scale bars = 5 mm. Bottom scale bars = 100 μm.

(C) Representative FACS plot of a dissociated tumor. FSC/SSC gated, Lineage^{negative}, viable (DAPI^{negative}) cells are shown. GFP^{postive}Tom^{negative} SCLC cells are indicated.

(D) Correlation of chromatin accessibility of 8 samples by ATAC-seq analysis. T = tumor, L = liver metastasis.

(E) Differential accessibility (log_2 fold change in reads per accessible region) plotted against the mean reads per region. FDR is false discovery rate that the absolute value of log_2 fold change is >1.

(F) Insertion tracks of samples at an example locus on chromosome 7. Differentially open regions are marked with arrows.

(G) Fraction of total regions that are differentially accessible in multiple tissue and cell type comparisons. lfc = log_2 fold change.

See also Figure S1.

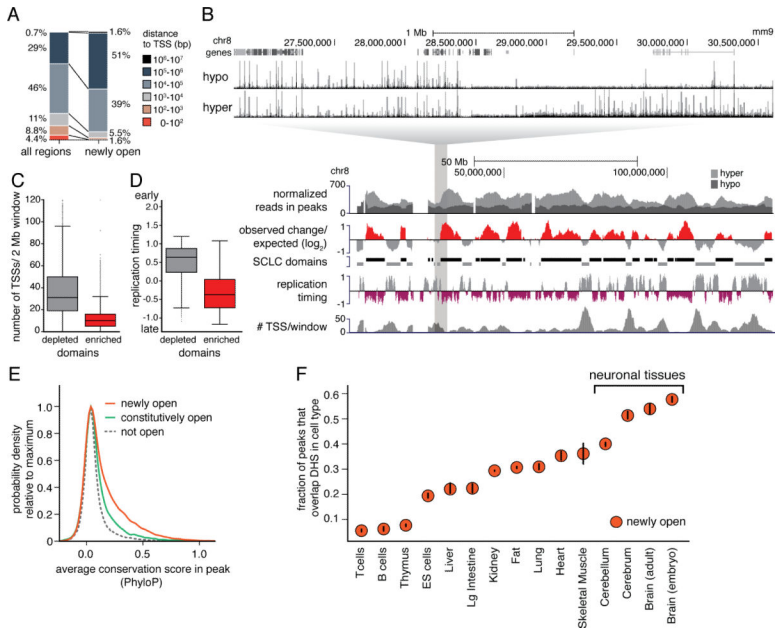

Figure 2. Changes in chromatin accessibility are gene distal, late replicating, evolutionarily conserved, and enriched for regulatory regions in neuronal tissue

(A) Distance to closest transcription start sites (TSSs) of all accessible regions and differentially open regions.

(B) Insertion tracks of merged hypo- and hyper-accessible samples. (Bottom) Reads in distal peaks are binned into 2 Mb windows. Examining differential read counts in windows identifies domains that are enriched for differentially open regions (black bars) and depleted for differentially open regions (grey bars, lower). Replication timing (ES-TT2 cells) and number of TSSs per window is also shown.

(C) Number of TSSs per 2 Mb window in domains either enriched or depleted for differentially open regions in SCLC. Enriched domains have fewer TSSs ($p < 1\times10^{-100}$ by Mann-Whitney U test).

(D) Average replication timing per 2 Mb window (ES-TT2 cells) in domains either enriched or depleted for differentially open regions in SCLC. Enriched domains have significantly later (more negative) replication timing ($p < 1\times10^{-100}$ by Mann-Whitney U test).

(E) Average sequence conservation (phyloP) in differentially open intergenic regions, constitutively open intergenic regions, and closed regions (5 kb downstream of any accessible region, same window size). Differentially open, intergenic regions are twice as likely to have higher average sequence conservation (>0.2) than constitutively open, intergenic regions ($p = 1\times10^{-114}$ by Fisher's exact test).

(F) Overlap of differentially accessible regions with DNase hypersensitive sites (DHS) from other cell types. Mean overlap with DHS peak calls are shown. Bars represent 95% confidence intervals.

See also Figure S2.

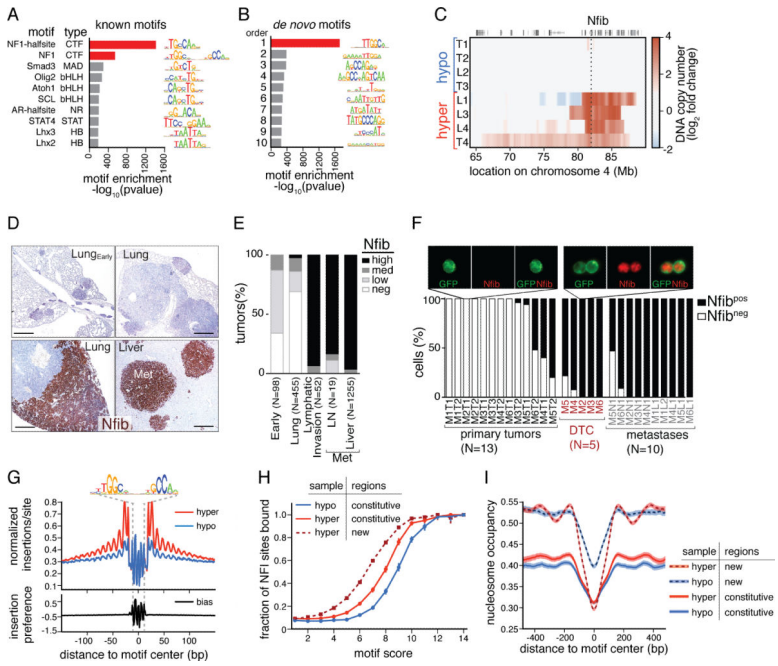

Figure 3. Increased Nfib expression in invasive SCLC leads to increased Nfib binding in differentially open regions and occupancy of less canonical sites

(A, B) Motif enrichment in newly open regions compared to other accessible regions. (A) Top known motif enrichments and (B) de novo motif enrichments.

(C) Copy number amplification of the *Nfib* locus inferred from ATAC-seq.

(D) Representative IHC for Nfib on tumors at different stages of SCLC progression from *TKO-mTmG* mice. Lung^{Early} are neuroendocrine hyperplasias. Scale bars = 500 μm.

(E) Nfib expression at different stages of SCLC progression. Number of tumors in each group is indicated. The percent of Nfib^{high} tumors in early hyperplasias and lung tumors versus lymph node (LN) and liver metastases is significantly different ($p < 1\times10^{-16}$ by Fisher's exact test).

(F) Immunofluorescence for Nfib on FACS-isolated GFP^{pos} cancer cells from primary tumors, disseminated tumor cells (DTCs), and metastases. A representative SCLC cell from a primary tumor and DTCs from the pleural cavity are shown. The percent of Nfib^{high} DTCs versus Nfib^{high} lung tumor cells is significantly different ($p < 0.0001$).

(G) (Top) ATAC-seq footprint at the NFI full site. Insertions per site are normalized to have the same average number of insertions 200-500 bp away from motif. (Bottom) Modeled insertion bias of Tn5 around NFI full sites.

(H) Occupancy of NFI full sites in the merged hyper- and hypo-accessible samples in bins of motif score (log odds similarity to the consensus motif). Error bars represent 95% confidence intervals.

(I) Nucleosome occupancy around NFI full sites in differentially open and constitutively open regions in hypo- and hyper-accessible samples. Shaded areas represent 95% confidence intervals.

See also Figure S3 and Table S1.

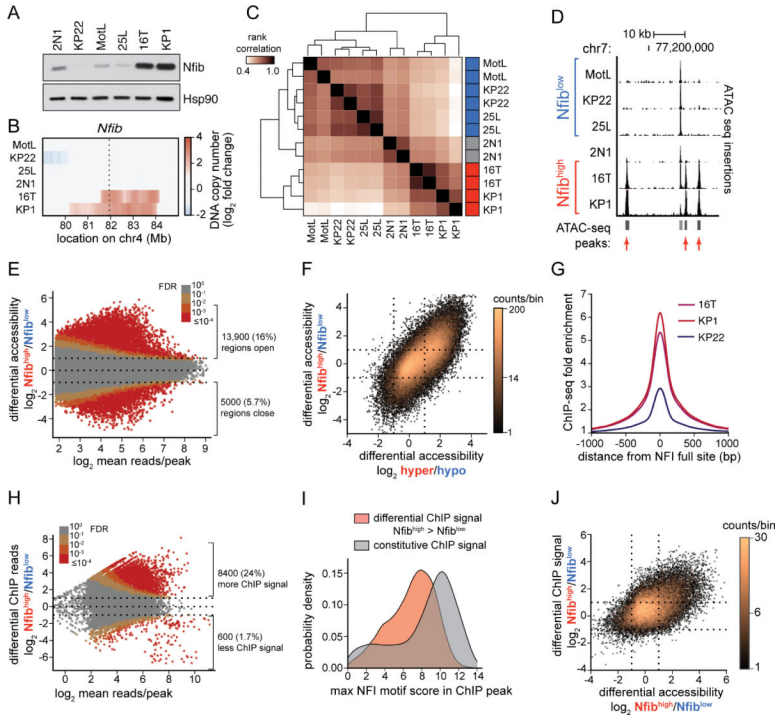

Figure 4. Cell lines confirm relationship between Nfib expression, Nfib binding by ChIP, and chromatin accessibility state

(A) Immunoblot analysis of Nfib expression in 6 murine SCLC cell lines. Hsp90 shows loading.

(B) Copy number amplifications on chromosome 4 from ATAC-seq.

(C) Correlation of chromatin accessibility across cell lines and technical replicates. Nfib^{high} and Nfib^{low} cell lines are indicated in red and blue, respectively.

(D) ATAC-seq insertion tracks of cell lines. Differentially accessible regions are indicated with arrows.

(E) Differential accessibility (log_2 fold change in reads per accessible region) between Nfib^{high} and Nfib^{low} cell lines, plotted against the mean reads per region. FDR is false discovery rate that the absolute value of log_2 fold change is >0.5. Brackets indicate number of significantly changed peaks.

(F) Correlation of differential accessibility between Nfib^{high} and Nfib^{low} cell lines and hypo- and hyper-accessible ex vivo samples (r = 0.72).

(G) Fold enrichment above input of Nfib ChIP around NFI motif sites for three cell lines.

(H) Log_2 fold change in Nfib ChIP-seq reads per merged ChIP peak versus the mean number of reads per peak. FDR is for absolute value of log_2 fold change > 0.5.

(I) Distribution of motif scores of sites within Nfib ChIP peaks, either those that gave more signal in the Nfib^{high} than the Nfib^{low} cell lines, or those that were not significantly different (constitutive). The maximum scoring NFI full site within each ChIP peak was used. Constitutive peaks have higher motif score than differential peaks (Mann-Whitney $p < 1\times10^{-100}$).

(J) Correlation of differential ChIP signal in accessible regions (from ATAC-seq) and differential accessibility of Nfib^{high} and Nfib^{low} cell lines (r = 0.35).

See also Figure S4.

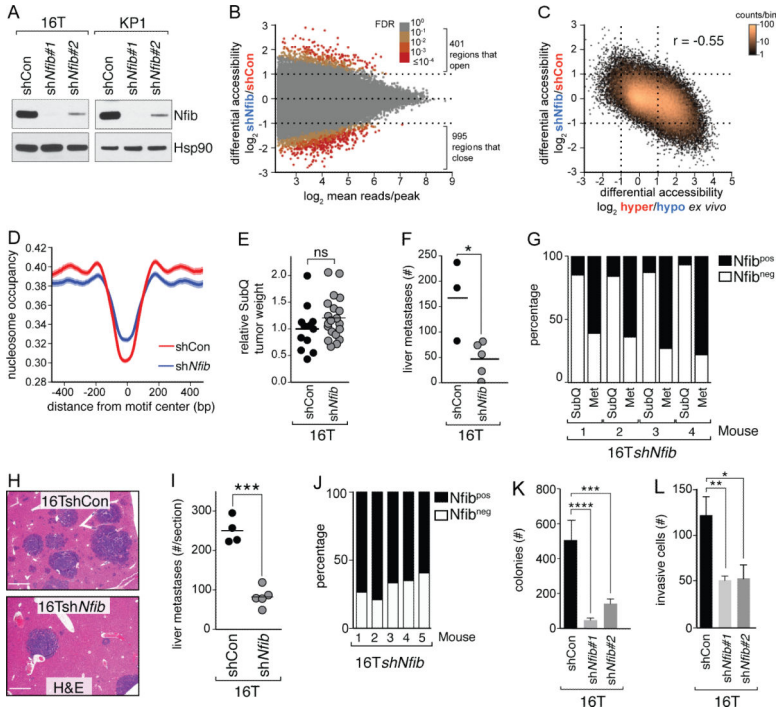

Figure 5. Nfib maintains chromatin accessibility at a subset of regions and is required for metastatic ability

(A) Immunoblot for Nfib in two Nfib^{high} SCLC cell lines +/− *Nfib* knockdown (shNfib). Hsp90 shows loading. shCon = shControl.

(B) Differential accessibility with shNfib#1 and control in combined 16T and KP1 cell lines.

(C) Correlation of differential accessibility with *Nfib* knockdown and differential accessibility of hyper- and hypo-accessible ex vivo samples (r = −0.55).

(D) Nucleosome occupancy around NFI sites in shNfib and control cells. Shaded regions are 95% confidence intervals.

(E) Subcutaneous (SubQ) tumor weight of 16T cells +/− *Nfib* knockdown 4 weeks after transplantation. Each dot represents a tumor and the line indicates the mean. ns = not significant.

(F) Number of liver metastases from SubQ tumors of 16T cells +/− *Nfib* knockdown. Scale bars = 500 μm. Each dot represents a mouse and the line indicates the mean. *p < 0.025.

(G) Percent of SubQ tumor area that expresses Nfib (Nfib^{pos}) and percent of liver metastases (Met) that are Nfib^{pos}. Paired analysis of Nfib^{pos} areas in SubQ and Met tumors is significant (p < 0.0025).

(H-I) Representative H&E images and quantification of liver metastases 3 weeks after intravenous transplantation. Scale bars = 500 μm. Each dot represents a mouse and the bar is the mean. ***p < 0.001.

(J) Percent of Nfib-positive liver metastases (assessed by IHC) after intravenous transplantation of 16TshNfib cells.

(K) Anchorage-independent growth of Nfib^{high} SCLC cell lines +/− *Nfib* knockdown. Mean +/− SD is shown. ****p < 0.0001, ***p < 0.001.

(L) Matrigel migration assay of Nfib^{high} SCLC cell lines +/− *Nfib* knockdown. Mean +/− SD is shown. **p < 0.01 *p < 0.02.

See also Figure S5.

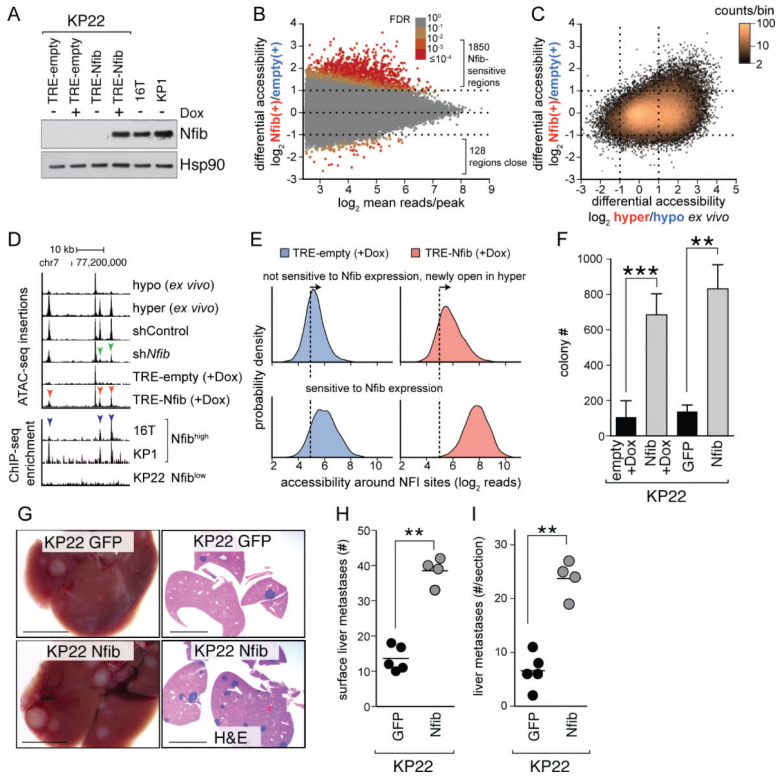

Figure 6. Nfib is sufficient to open a subset of sensitive regions and to drive metastatic ability

(A) Expression of Nfib by immunoblot in an Nfib^{low} cell line (KP22) with doxycycline (Dox)-inducible expression of Nfib. Two Nfib^{high} cell lines 16T and KP1 are shown. Hsp90 shows loading.

(B) Differential accessibility in Nfib-expressing and control samples.

(C) Correlation of differential accessibility with Nfib expression and differential accessibility of hyper- and hypo-accessible ex vivo samples (r = 0.26).

(D) (Top) ATAC-seq insertions and (Bottom) Nfib ChIP enrichment above input at an example locus. Green arrows highlight regions that close with *Nfib* knockdown, red arrows highlight regions that open with Nfib expression, and blue arrows show regions with differential ChIP signal.

(E) Distribution of accessibility (log2 read count) in 1000 bp windows around NFI motif sites +/− Nfib induction. Panels show motif sites that are specifically more open in hyper-accessible ex vivo samples or that are sensitive to Nfib overexpression in KP22 cell line. The dashed line indicates the average number of reads around motif sites that were not accessible.

(F) Colony formation of KP22 cells with induced (Nfib+Dox) and constitutive Nfib expression in anchorage-independent conditions. Mean +/− SD is shown. ***p < 0.0005 **p < 0.001.

(G) Light (left) and H&E (right) image of liver metastases of KP22 +/− Nfib expression after intravenous transplantation. Scale bars = 5 mm.

(H,I) Number of surface liver metastasis (H) and liver metastases quantified by histology (I). **p < 0.001

See also Figure S6.

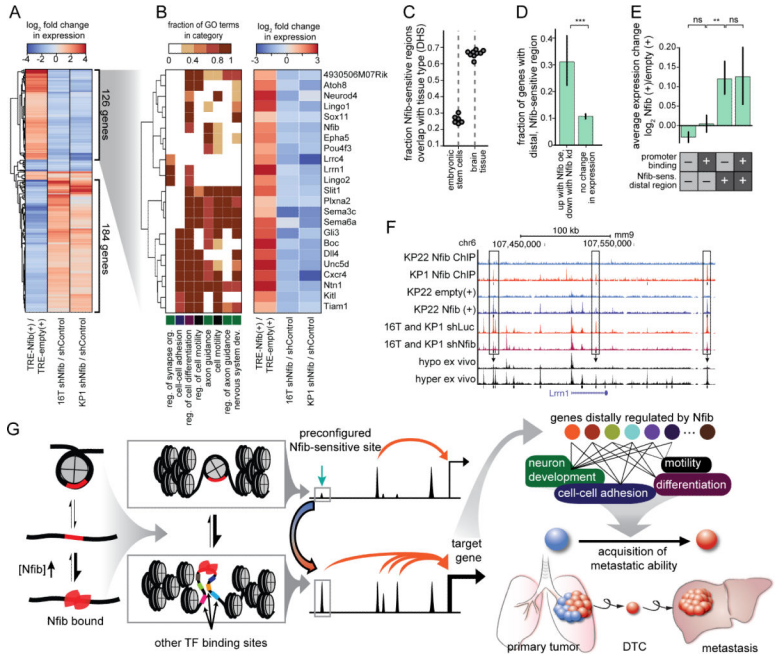

Figure 7. Gene expression changes in response to Nfib promote neuronal state

(A) Log_2 fold change of all genes that significantly change (absolute value of log2 fold change > 0.5 at FDR < 0.1) with both knockdown and overexpression by RNA-seq.

(B) GO terms (merged into categories) and log_2 fold change in expression of a subset of genes.

(C) Overlap of Nfib-sensitive regions with DNase hypersensitive sites (DHS) in other tissues. Dots represent multiple ES cell lines plus technical replicates, or adult and embryonic brain plus technical replicates.

(D) Fraction of genes that have at least one distal, Nfib-sensitive region within 100 kb of transcription start site. oe=overexpression, kd=knockdown. Error bars are 95% confidence intervals. ***$p < 1\times10^{-10}$ by Fisher's exact test.

(E) Average change in expression with Nfib overexpression of 4 gene categories. Promoter binding = Genes with promoter-proximal ChIP peak within 1 kb of TSS. Nfib-sens. distal region = Genes with a distal, Nfib-sensitive region within 100 kb of the TSS that opens with Nfib overexpression. Error bars are 95% confidence intervals. ns=not significant, **$p < 1\times10^{-7}$ by Mann-Whitney U test.

(F) Example locus showing Nfib ChIP (top two panels) and ATAC-seq data. Black squares below each set of tracks represent significantly changed regions.

(G) Model for Nfib-dependent metastatic progression.

See also Figure S7 and Table S2.