Sox2 Is an Oncogenic Driver of Small-Cell Lung Cancer and Promotes the Classic Neuroendocrine Subtype

Ellen Voigt[1,2], Madeline Wallenburg[1,2], Hannah Wollenzien[1,2,3], Ethan Thompson[1,2], Kirtana Kumar[1,2], Joshua Feiner[4], Moira McNally[1,2], Hunter Friesen[1,2], Malini Mukherjee[5], Yohannes Afeworki[5], and Michael S. Kareta[1,2,3,5,6,7]

ABSTRACT

Although many cancer prognoses have improved in the past 50 years due to advancements in treatments, there has been little improvement in therapies for small-cell lung cancer (SCLC). One promising avenue to improve treatment for SCLC is to understand its underlying genetic alterations that drive its formation, growth, and cellular heterogeneity. *RB1* loss is one key driver of SCLC, and *RB1* loss has been associated with an increase in pluripotency factors such as *SOX2*. *SOX2* is highly expressed and amplified in SCLC and has been associated with SCLC growth. Using a genetically engineered mouse model, we have shown that *Sox2* is required for efficient SCLC formation. Furthermore, genome-scale binding assays have indicated that *SOX2* can regulate key SCLC pathways such as *NEUROD1* and *MYC*. These data suggest that *SOX2* can be associated with the switch of SCLC from an *ASCL1* subtype to a *NEUROD1* subtype. Understanding this genetic switch is key to understanding such processes as SCLC progression, cellular heterogeneity, and treatment resistance.

Implications: Understanding the molecular mechanisms of SCLC initiation and development are key to opening new potential therapeutic options for this devastating disease.

Introduction

Small-cell lung cancer (SCLC) is a devastating disease with markedly low survival rates, rapid metastasis, and almost invariable resistance to therapy. Patients who are stricken by this disease face a 6% two-year survival rate, while most will succumb less than a year after diagnosis (1, 2). Despite this alarming statistic, the standard of care for treating SCLC has remained essentially the same for the past 40 years and few innovations have been approved for this disease. First-line treatments still rely primarily on platinum-based chemotherapy that often leads to treatment refractory tumors and poor patient outcomes (3–5). Recently immunotherapy options have been available for SCLC; however, while the results have been encouraging in select individuals, the patient responses have been generally poor (6). Therefore, in the pursuit of new therapies for SCLC, we have sought to understand the genetic factors underlying SCLC dynamics.

On a genetic level, SCLC is both rather simple and complex. It is simple in that the genetic drivers of SCLC are relatively clear. Patients have an almost invariable loss of the tumor suppressors *p53* (*TP53*) and *RB1* (*RB*; refs. 7–9). Intriguingly, established SCLC can be genetically complex considering that, even with almost identical driver mutations, SCLC can be subdivided into four main subtypes defined by the function of key genetic regulators, *ASCL1*, *NEUROD1*, *POU2F3*, and *YAP1* (refs. 10–13, reviewed in: ref. 14). Critically linked to the regulatory networks of the *ASCL1* (SCLC-A) and the *NEUROD1* (SCLC-N) subtypes is the role of the MYC family of oncogenes. MYC (cMYC) is highly expressed and a determining factor for the SCLC-N subtype (15). MYCL (L-Myc) rather, is predominantly expressed in SCLC-A, and is key to SCLC-A growth (7, 11, 16, 17). While MYC family regulation is important to SCLC growth and development (18), how MYC family members are regulated in SCLC is currently unclear (19).

The question of how a tumor with such homogenous driver mutations (*RB1* and *p53* loss) can lead to the diversity of genetic heterogeneity observed in SCLC remains unanswered. One clue to address this question can be found in the nature of the initiating mutations themselves. Beyond its role in regulating the G_1–S checkpoint, RB also plays a multitude of roles in regulating gene expression (20–22). One of the genes regulated by RB is the transcription factor *SOX2* (23). Known primarily as a pluripotency factor, *SOX2* is also a key master regulator of neural and neuroendocrine cell types (24–28). As a master regulator, *SOX2* influences cell identity early and widely in cell fate decisions. Indeed, SOX2 is commonly amplified in SCLC (7). Pulmonary neuroendocrine cells are the predominant cell of origin for SCLC (29), therefore it is possible that *SOX2* upregulation in neuroendocrine cells following *RB1*-loss induces stem or progenitor genetic networks that help to drive oncogenesis. To that end, we generated a conditional knockout mouse in which we could perturb *Sox2* activity in a well characterized SCLC mouse model to assess the consequence of *Sox2*-loss on SCLC formation. Combined with a genome-wide investigation into *SOX2* transcriptional regulation in SCLC, we observed that *SOX2* is indeed required for SCLC

[1]Cancer Biology and Immunotherapies Group, Sanford Research, Sioux Falls, South Dakota. [2]Genetics & Genomics Group, Sanford Research, Sioux Falls, South Dakota. [3]Division of Basic Biomedical Sciences, University of South Dakota, Vermillion, South Dakota. [4]Dakota Wesleyan University, Mitchell, South Dakota. [5]Functional Genomics & Bioinformatics Core, Sanford Research, Sioux Falls, South Dakota. [6]Department of Pediatrics, Sanford School of Medicine, University of South Dakota, Sioux Falls, South Dakota. [7]Department of Chemistry Biochemistry, South Dakota State University, Brookings, South Dakota.

Note: Supplementary data for this article are available at Molecular Cancer Research Online (http://mcr.aacrjournals.org/).

E. Voigt and M. Wallenburg contributed equally to this article.

Corresponding Author: Michael S. Kareta, Cancer Biology and Immunotherapies Group, Sanford Research, 2301 East 60th Street North, Sioux Falls, SD 57104. E-mail: michael.kareta@sanfordhealth.org

Mol Cancer Res 2021;19:2015–25

doi: 10.1158/1541-7786.MCR-20-1006

formation and regulates key genetic regulators of SCLC including *NEUROD1* and members of the *MYC* family.

Materials and Methods

Ethics statement

Mice were maintained according to the guidelines set forth by the NIH and were housed in the Sanford Research Animal Research Center, accredited by Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC) using protocols reviewed and approved by our local Institutional Animal Care and Use Committee (IACUC).

SCLC mouse tumor initiation

We modeled SCLC in the *Rb1*$^{lox/lox}$, *p53*$^{lox/lox}$, *p130*$^{lox/lox}$, *Rosa*^{luc} (RPR2) mouse line (30), which readily develop SCLC after a few months, and added *Sox2* $^{+/+,+/lox}$, or $^{lox/lox}$ alleles (The Jackson Laboratory, stock no. 013093; ref. 31). To study SCLC tumor initiation, we injected Cre-recombinase adenovirus [Ad5–cytomegalovirus (CMV)–Cre, Baylor Vector Development Lab, 0.91 μL of a 5 × 10^{12} pt/mL viral preparation used per mouse] into the mouse lungs by intratracheal intubation to excise the lox-flanked genes (32). The mice were assigned to either a 6-month cohort, a 3-month cohort, or the survival curve. Mouse lungs, livers, and any other metastases were harvested for IHC. Tumors were screened in a blinded manner by an independent pathologist.

SCLC lung and liver IHC

The Sanford Research Histology & Imaging Core performed the IHC for this study. The mouse lungs, livers, and tumors were stained with hematoxylin and eosin (H&E), for SOX2 (Abcam ab92494, 1:100), calcitonin gene-related peptide (CGRP, Sigma C8198, 1:2,000), anti–phospho-histone H3 (pH3, EMD Millipore 06–570, 1:500), cleaved caspase 3 (CC3, Cell Signaling Technology 9664, 1:100), ki67 (Biocare CRM325, 1:100), ASCL1 (Abcam ab74065, 1:500), and MYC (c-MYC, Invitrogen MA1–980, 1:100; Supplementary Table S4). To computationally assess tumor burden and feature characteristics, we digitized each slide using an Aperio VERSA slide scanner. The five images from each sample (H&E and SOX2, CGRP, ki67, pH3, and CC3 IHC) were registered using the Register Virtual Stack Slices Plugin in FIJI/ImageJ (33). We then used CellProfiler (34) to count the tumors and features. The H&E staining was used to identify tumors, then the intensity of IHC staining for the markers SOX2, CGRP, ki67, pH3, and CC3 was determined for the corresponding tumor areas in the other virtual slide images. Registration and CellProfiler scripts are available on the Kareta Lab website (https://research.sanfordhealth.org/researchers-and-labs/kareta-lab).

SCLC cell lines

We used the murine SCLC cell lines KP1 and KP3 (*Rb1*$^{lox/lox}$; *p53*$^{lox/lox}$) and the human SCLC lines NJH29 (H29), NCI-H82 (H82), NCI-H1836 (H1836), and NCI-H209 (H209; refs. 30, 35). The cells were maintained in suspension and cultured in RPMI with 10% bovine growth serum and penicillin/streptomycin. All cell lines regularly tested negative for mycoplasma contamination and validated by IDEXX BioAnalytics.

Lentiviral transduction and cell assays

We made the lentivirus for the short hairpin RNA (shRNA)–mediated knockdown using the third-generation packaging plasmids pMD2.G, pMDLg/pRRE, and pRSV-Rev in 293T cells, transfecting them with polybrene (PB) with a nearly 90% transduction rate. Resulting lentivirus was concentrated using Lenti-X Concentrator (Takara Bio, Inc.) and titered for reproducible transductions. Controls consisted of an empty pSicoR vector or a pSicoR vector containing a shRNA to *Luciferase* (23). Transduced cells were selected for by culture with Puromycin for 5 days. We measured cellular viability after *SOX2* knock down with an alamar blue assay, and the levels of apoptosis with Annexin V staining combined with flow cytometry. qPCR was used to confirm the knock down of *Sox2* in the cells. Cas9-mediated knockdown of SOX2 was achieved by cloning a SOX2 guide RNA (gRNA) sequence (ATTATAAATAC-CGGCCCCGG) into the TLCV2 inducible lentiviral Cas9 vector (36), which was packaged in to lentivirus using the methods above. Transfection was achieved using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's protocol. To enhance transfection efficiency, after adding the transfection mix the cells were processed according to a modified spinfection protocol where they were centrifuged at 940 xg for 2 hours at room temperature. Mock controls were Lipofectamine-treated and spinfected cells that were processed the same but without the presence of the DNA vector. Due to high transfection efficiencies (typically greater than 70%), cells were neither selected nor sorted to minimize stress.

Chromatin immunoprecipitation and cleavage under targets and release using nuclease assays

In preparation for HA-RB1ΔCDK chromatin immunoprecipitation (ChIP), cells were transfected with pCMV-HA-hRb1-delta-CDK (Addgene, catalog no. 58906) using Lipofectamine 3000 (ThermoFisher). ChIP for HA-RB1ΔCDK was performed as previously described (23) with several additional optimizations (37). The alternative swelling buffer was used for cell lysis. Chromatin was sonicated using a ME220 (Covaris, Inc.). ChIP-grade Protein AG magnetic beads (Pierce) were preblocked with BSA and salmon sperm DNA for 15 minutes on a rotating platform at 4°C. The chromatin was precleared before being diluted and incubated with an anti-HA antibody (Sigma H6908, 4 μg) for immunoprecipitation. The antibody-chromatin complexes were incubated with blocked beads for 2 hours at 4°C on a rotating platform prior to washing two times each with low-salt, high-salt, and LiCl wash buffers.

Cleavage Under Targets and Release Using Nuclease (CUT&RUN) assays were carried out according to the protocol (Version 3) published by Janssens and Henikoff (38) which is based on the original protocol developed by Skene and colleagues (39), using the CUTANA pAG-MNase (EpiCypher), and concanavalin-A coated beads (BioMag Plus catalog no. 86057). The optional high-calcium/low-salt conditions were included to prevent premature chromatin release after digestion. Both ChIP and CUT&RUN assays were performed using SOX2 antibodies from both EMD/Millipore (17–656) and R&D Systems (AF2018). ChIP and CUT&RUN libraries were analyzed on an Agilent Bioanalyzer System by the Sanford Research Functional Genomics & Biochemistry Core and sequenced at the Sanford Burnham Prebys Genomics Core. Both ChIP and CUT&RUN reads were aligned to the hg38 genome build using Bowtie 2 version 2.3.4.3 (40) and peaks called using MACS2 version 2.1.2 (41). As described by the authors of CUT&RUN, the top 99.5th percentile of peaks after sorting by *q* values (including peaks with the same *q* value at cutoff) were selected for further analysis (39). HOMER was used for heatmap generation and motif enrichment (42), Diffbind was used for differential peak identification and principal component analysis (PCA) visualization (43), and Ingenuity Pathway Analysis for network analysis (QIAGEN Inc.). Weighted gene coexpression network analysis was performed using

the WGCNA package from Bioconductor (44). RNA sequencing (RNA-seq) data was analyzed using DESeq2 (45).

Data availability

Sequencing data have been uploaded to the Gene Expression Omnibus (GEO) Repository under accession GSE182728.

Results

Sox2 is critical for SCLC tumor initiation

To investigate if *Sox2* is required for the formation of SCLC, we bred a mouse line containing a conditional *Sox2* allele ($Sox2^{lox/lox}$) to the RPR2 [$Rb1^{lox/lox}$; $p53^{lox/lox}$; $Rbl2(p130)^{lox/lox}$] mouse model of SCLC (**Fig. 1A**; refs. 29, 30, 46, 47). With the addition of the conditional *Sox2* allele, we therefore named this line RPR2S. Tumors from RPR2 mice display all the common hallmarks of human SCLC, mainly the same histologic characteristics as scored by an independent pathologist, rapid metastasis, and chemoresistance (30, 47, 48). To overcome the dramatic effects of global *Rb1* and *p53* loss in the mouse, we localized Cre-mediated recombination by an intratracheal instillation of a Cre-expressing adenovirus (Adeno-CMV-Cre-GFP) to target recombination specifically to the lung epithelium (49). As expected, we observed early lesions around 3 months, with a robust tumor burden 6months after Cre-recombination (30).

By utilizing a breeding strategy that generates all three allelic combinations of *Sox2*: $Sox2^{+/+}$, $Sox2^{+/lox}$, and $Sox2^{lox/lox}$ (Supplementary Table S1), we were able to query if one or both alleles of *Sox2* are involved in SCLC formation. Three and 6 months after Adeno-Cre tumor initiation, $Rb1^{lox/lox}$; $p53^{lox/lox}$; $p130^{lox/lox}$ mice showed a sizeable number of tumor foci displaying the histologic characteristics of SCLC. However, the RPR2S mice had a nearly complete loss of SCLC foci observed at the same timepoint (**Fig. 1B**). To fully characterize these tumors and ensure complete *Sox2* loss in the RPR2S mice, we optimized IHC staining and an unbiased image analysis pipeline using ImageJ and CellProfiler (34) resulting in a thorough statistical analysis of the number and marker expression in the RPR2 tumors compared with the few RPR2S tumors (**Fig. 1C** and **D**; Supplementary Fig. S1). The RPR2 tumors showed typical SCLC histology including high *Cgrp* expression, indicative of a neuroendocrine tumor type, and highly proliferative cells as indicated by ki67 and pH3 staining (ref. 50; Supplementary Fig. S1). At 6 months, there were a handful of very small tumors observed in the $Rb1^{lox/lox}$; $p53^{lox/lox}$; $p130^{lox/lox}$; $Sox2^{lox/lox}$ mice (**Fig. 1D**; Supplementary Fig. S1), although a sizeable number of these showed immunoreactivity to SOX2 antibodies, indicating that they are the result of incomplete Cre function. However, a small minority of SCLC tumors can initiate without *Sox2*, indicating that *Sox2* activity may not be absolutely necessary in some SCLC tumors or tumor subtypes. However, those

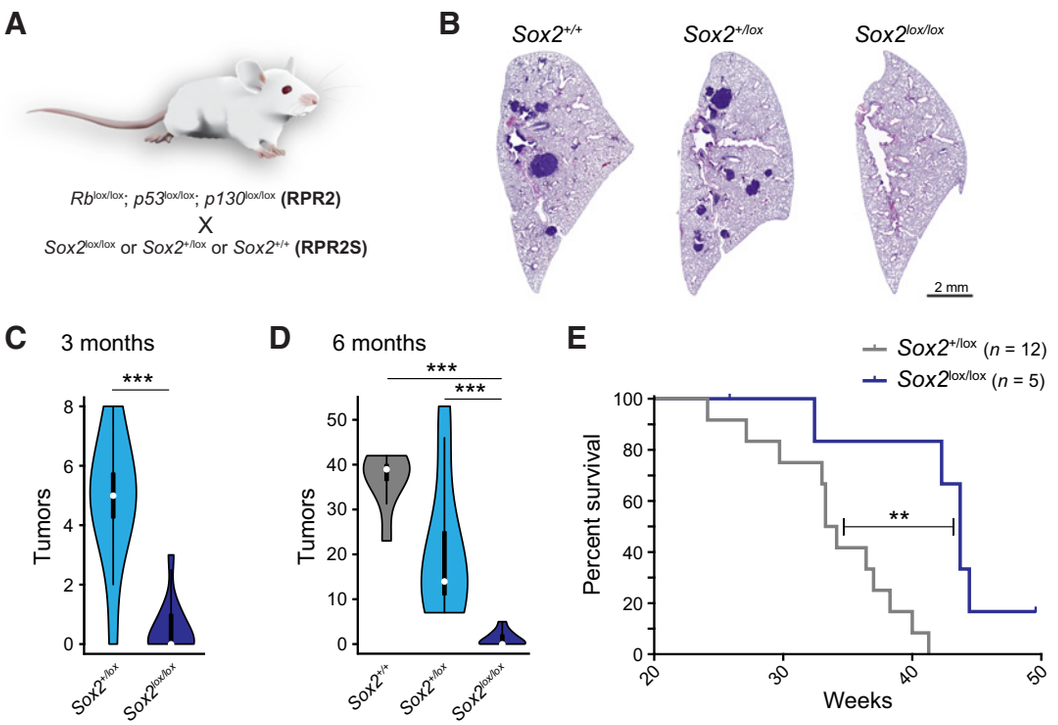

Figure 1.

Sox2 is required for SCLC formation. **A,** Genetically engineered mouse model for the study of *Sox2* in SCLC. **B,** Representative H&E stained lung sections from $Rb1^{lox/lox}$; $p53^{lox/lox}$; $p130^{lox/lox}$; $Sox2^{+/+}$ (left), $Rb1^{lox/lox}$; $p53^{lox/lox}$; $p130^{lox/lox}$; $Sox2^{+/lox}$ (middle), and $Rb1^{lox/lox}$; $p53^{lox/lox}$; $p130^{lox/lox}$; $Sox2^{lox/lox}$ (right) mice, 6 months after Cre recombination. **C,** Number of tumors as indicated by H&E staining 3 months after Cre recombination. **D,** Number of tumors as indicated by H&E staining 6 months after Cre delivery. Numbers of mice used in **C–D** can be found in Supplementary Table S1. **E,** Kaplan–Meier survival curve of SOX2 wild-type (WT) mice ($Sox2^{+/lox}$) compared with $Sox2^{lox/lox}$ mice. Violin plots show median (white dot), interquartile range (box) and the continuous distribution of the data; significance for all panels determined by a two-tailed *t* test where * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.01$.

tumors that grew even when *Sox2* was deleted were markedly smaller in size than the *Sox2*$^{+}$ tumors (Supplementary Fig. S1D). Importantly, we observed a significant lengthening of the lifespan of the *Rb1*$^{lox/lox}$; *p53*$^{lox/lox}$; *p130*$^{lox/lox}$; *Sox2*$^{lox/lox}$ mice (**Fig. 1E**), compared with *Sox2*-expressing controls.

SOX2 is required for the growth of established SCLC lines

The results indicating *Sox2* function in the initiation of SCLC tumors in mice led us to investigate if *Sox2* is required in established tumors. We utilized shRNA-mediated knockdown to reduce *SOX2* expression in both mouse and human SCLC cell lines. We were able to achieve an approximately 60% to 90% knockdown of *SOX2* by qRT-PCR (Supplementary Fig. S2A). We observed that knockdown of *SOX2* in both mouse and human cell lines significantly reduces the growth of these cells in culture compared with mock-transduced cells (**Fig. 2A**), similar to a previously reported *SOX2* knockdown in human SCLC cell lines (7). Concurrent with a loss of cellular viability, we observed an increase in the number of apoptotic cells upon *SOX2* knockdown (Supplementary Fig. S2B). As *RB1*-loss is one of the primary genetic drivers of SCLC (7, 9, 47), and the RB protein can bind to and repress the *Sox2* locus in fibroblasts (23), we set out to investigate if RB is capable of repressing *SOX2* in SCLC to indicate if RB loss in SCLC could be the driver of *SOX2* upregulation. To this end, we overexpressed an *RB1* transgene in human SCLC cell lines in which the CDK phosphorylation sites have been mutated (*RB1ΔCDK*) to render RB resistant to CDK inactivation (51). Overexpression of *RB1ΔCDK* greatly reduced the viability of human SCLC cell lines (**Fig. 2B**; ref. 52). Furthermore, overexpression of RB1 resulted in the repression of *Sox2* (**Fig. 2C**). By ChIP we tested if RB1ΔCDK binds to the promoter or the two known proximal SOX2 enhancers, *SRR1* and *SRR2* (53). Indeed, we do observe significant enrichment of RB1ΔCDK-bound regions at the *SOX2* promoter and the downstream *SRR2* enhancer (**Fig. 2D**). Finally, overexpression of *SOX2-t2a-GFP* rescued the repression of RB1ΔCDK growth-inhibited SCLC cell lines (**Fig. 2E**; Supplementary Fig. S2C). Together these data confirm that *SOX2* is required for SCLC tumor growth and that *SOX2* expression is most likely a consequence of *RB1*-loss.

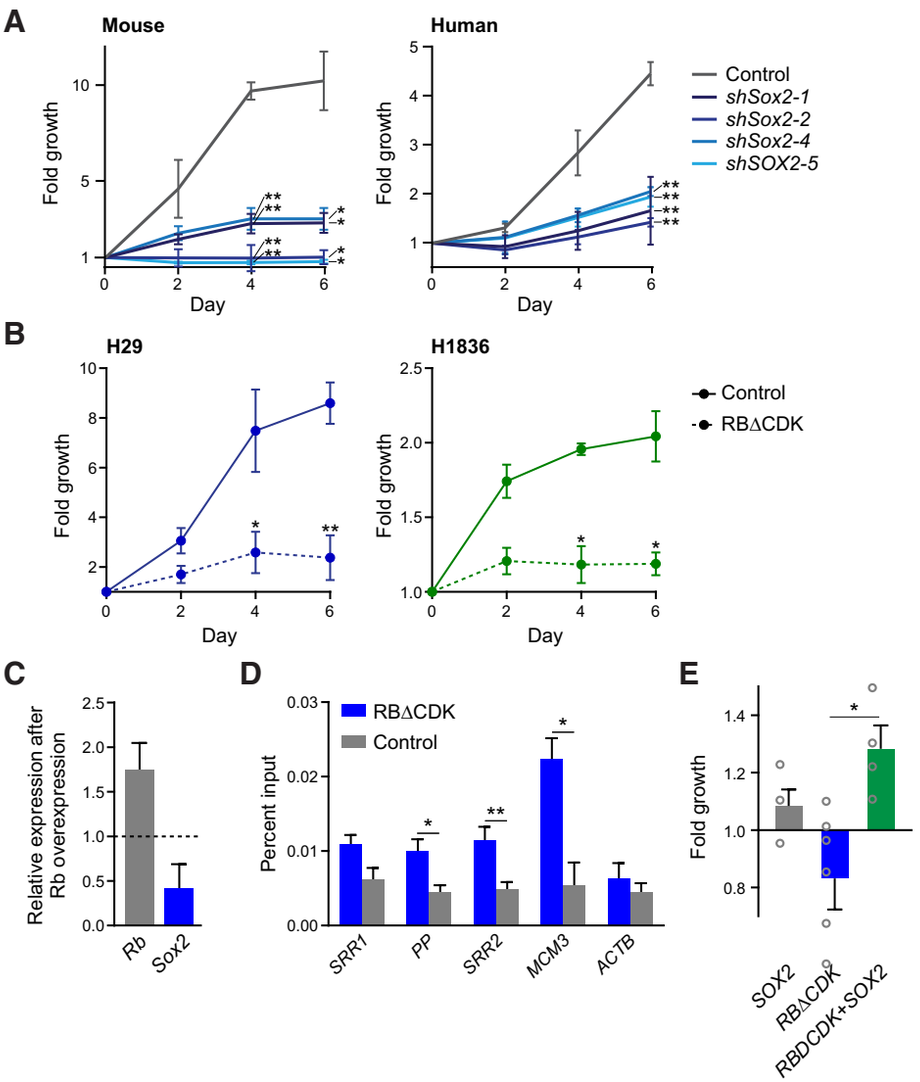

Figure 2.

RB represses SOX2 and is required for SCLC. **A,** Using 3 hairpins designed to murine Sox2, (shSox2–1,2,&4) and one designed to human SOX2 (shSOX2–5) we tested the effect on cellular proliferation by an Alamar Blue assay in KP1, KP3, H29, and H82 cell lines. **B,** Alamar Blue assays of H29 and H1836 cells after transfection with *RB1ΔCDK*. **C,** Expression of *Rb1* and *Sox2* measured by qPCR after transduction of Adeno-Rb1 virus in KP1 and KP3 cells. **D,** ChIP of HA-RBΔCDK or mock-transfected cells (H29 and H1836). Regions tested for ChIP enrichment by qPCR are the *SOX2* proximal promoter (PP), the *SRR1* and *SRR2* enhancers of *SOX2*, *MCM3* promoter as a positive control and *ACTB* promoter as a negative control. **E,** Alamar Blue assay on day 4 to determine the proliferation of H29, H82, H1836, and H209 cells after transfection with *RB1ΔCDK*, and or *SOX2-t2a-GFP*. Proliferation is plotted as the fold change compared with a mock-transfected control. Individual values are notated by gray circles. Bar graphs show mean and SEM, significance for all panels determined by a two-tailed *t* test where * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.01$.

SOX2 regulates key SCLC pathways

To observe the genomic localization of SOX2 in human SCLC cell lines, we performed both ChIP and CUT&RUN (39) using the endogenous SOX2 from both H1836 and H29 cells. While SOX2 ChIP allowed for broad localization studies, we found that SOX2 CUT&RUN was much more sensitive for comparative genomic localization studies due to the lack of chemical cross-linking and the release of SOX2-bound DNA due to SOX2 antibody:ProteinA/G:MNase complexes rather than sonication. We observed a very similar localization of SOX2 in both cell lines (**Fig. 3A**; Supplementary Fig. S3;

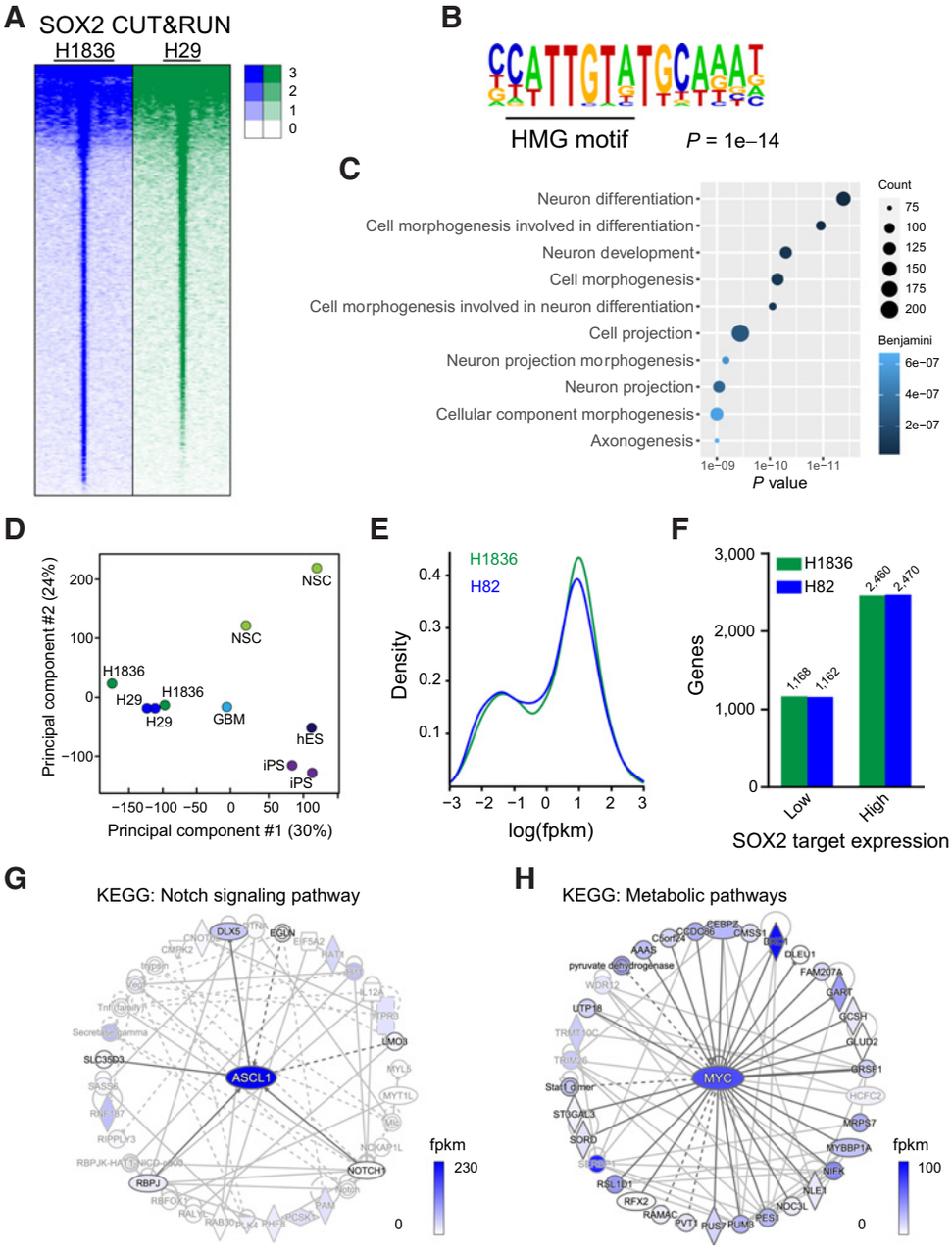

Figure 3.

SOX2 regulates key SCLC pathways. **A,** SOX2 CUT&RUN heatmap from H1836 and H29 cell lines. Each row represents the normalized read counts at all peaks identified for SOX2 binding. **B,** *De novo* motif identification discovers an HMG domain as the most prominent motif in the SOX2 peaks. **C,** Top 10 Gene ontology terms enriched at the genes associated with the SOX2 peaks. **D,** PCA plot of other human SOX2 genome binding profiles. Studies include samples from iPS cells, ES cells, NSCs, and iPS-derived NSCs, and glioblastoma (GBM). Datasets include GSE69479, GSE81900, GSE49405, GSE23839 (57–60). **E,** Density plot of the log(fpkm) values of all genes associated with a SOX2 peak. **F,** Number of genes in **E** that are predicted to be part of the low- or high-expression group after Gaussian mixed model clustering (Supplementary Fig. S4). **G** and **H,** WGCNA identified two networks that include *ASCL1* and *MYC*. Color scale reflects the relative expression of each gene in the network from the expression profiles available in the CCLE. hES, human embryonic stem cells; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Supplementary Table S3). Unbiased motif enrichment of the SOX2 peaks identified an HMG binding domain as the most highly enriched motif (**Fig. 3B**). The high mobility group (HMG) domain is the DNA-binding domain of the SOX family of proteins therefore, the presence of HMG motifs validates the specificity of the SOX2 localization (54). As expected for a neuroendocrine tumor, and with the known role of *SOX2* in the regulation of neurogenesis (55, 56), the top ontology terms for the SOX2 adjacent genes were related to neural development and function (**Fig. 3C**). To assess if the binding topology of SOX2 in SCLC is similar to other *SOX2*-expressing cells, we compared the binding similarity by read counts for SOX2 datasets from human embryonic stem (ES) cells, induced pluripotent stem (iPS) cells, neural stem cells (NSC), and glioblastoma (57–60). We observe that SOX2 binding in SCLC is distinct from both NSCs and pluripotent cells (ES and iPS cells). The closest binding profile to SCLC was glioblastoma therefore the function of SOX2 in cancer may be distinct from its role in normal cellular development (**Fig. 3D**).

The genes that are bound by SOX2 appear to show a biphasic distribution of high or low expression, indicating that they are either upregulated or repressed by SOX2 (**Fig. 3E** and **F**; Supplementary Fig. S4). Indeed, SOX2 can either repress or transactivate target genes based upon the cofactors recruited (54, 61, 62), and it appears these two roles of SOX2 are maintained in SCLC. To better describe the genetic networks that are regulated by *SOX2* in SCLC we performed a weighted gene cexpression network analysis (WGCNA) to identify the gene networks coexpressed with *SOX2* using the SCLC cell lines in the Cancer Cell Line Encyclopedia (CCLE; refs. 63, 64). The WCGNA analysis identified multiple modules that are coexpressed with *SOX2* (Supplementary Figs. S5 and S6). The most highly upregulated module with *SOX2* contained *ASCL1*, a known regulator of classic SCLC (refs. 10, 11, 14; **Fig. 3G**). The most downregulated module identified contained a MYC network, which is associated with the variant state of SCLC (ref. 15; **Fig. 3H**).

SOX2 regulates SCLC subtype–specific specification

To further investigate the result that high levels of *SOX2* favor *ASCL1* gene modules and is anticorrelated with *MYC* gene modules (**Fig. 3G**) we investigated if SOX2 expression favors the ASCL1 SCLC subtype. We performed unbiased clustering of CCLE SCLC cell lines based on their expression of *ASCL1*, *NEUROD1*, *YAP1*, *POU2F3*, *MYC*, and *MYCL* (**Fig. 4A**). The cell lines generally clustered by subtype and *SOX2* specifically clustered with the *ASCL1* subtype. As it is unclear if the ASCL1-SOX2 module (**Fig. 3G**) is due to direct SOX2 regulation of ASCL1 or a correlation due to high SOX2 levels in the SCLC-A subtype cell lines (Supplementary Fig. S6), we set out to determine if the regulation of the SCLC subtype-specific factors *ASCL1* and *NEUROD1* is directly regulated by SOX2. Overexpression of *SOX2-t2a-GFP* in two SCLC-A (H1836 & H209) and two SCLC-N (H29 & H82) cell lines does not appear to perturb *ASCL1* levels, but does result in significant downregulation of *NEUROD1* (**Fig. 4B**; Supplementary Fig. S7). We observed similar changes at the protein level, although levels of NEUROD1 were marked lower in H1836 and H209 cells (**Fig. 4D**). We then used an inducible Cas9-mediated knockdown of *SOX2* rather than shRNA-mediated knockdown to observe the rapid effects of target gene expression after Cas9 induction, which results in significant SOX2 knockdown (Supplementary Fig. S8A). In contrast to *SOX2* overexpression, we observed a significant upregulation of *NEUROD1* (**Fig. 4C**; Supplementary Fig. S8A). To test if regulation of *NEUROD1* by SOX2 is direct we performed ChIP of SOX2. We observed significant binding of SOX2 at the *NEUROD1* and *MYC* promoters (**Fig. 4E**; Supplementary Fig. S8B). Therefore, it appears that SOX2 does not directly regulate *ASCL1*; however, it is associated with the progression of SCLC tumors to the *NEUROD1* state.

SOX2 directly regulates *MYC* and *MYCL* in the ASCL1 and NEUROD1 SCLC subtypes

With the observation that SOX2 potentially regulates MYC networks in SCLC (**Fig. 3H**), we investigated if SOX2 directly regulates the MYC family in SCLC. We observed binding of SOX2 to both *MYC* and *MYCL* in SCLC from the CUT&RUN data (**Figs. 3A** and **5A**). Both *MYC* and *MYCL* are expressed in SCLC, with *MYCL* predominantly expressed in the SCLC-A subtype and *MYC* expressed in the SCLC-N subtype (14–16). Interestingly, SOX2 appears bound at *MYCL* in H1836 cells, which are of the SCLC-A subtype and is bound at *MYC* in H29 cells, which are of the SCLC-N subtype (Supplementary Fig. S7; ref. 65). This is consistent with a role for SOX2 to activate these genes in their respective SCLC subtype. Overexpression of *SOX2* in both SCLC-A and SCLC-N cells further supports a role for SOX2 in the regulation of *MYC* and *MYCL*. When *SOX2-t2a-GFP* is transfected into the SCLC-A cell lines H1836 and H209, we observed a downregulation of *MYC* at both the mRNA (**Fig. 5B**) and protein levels (**Fig. 5C** and **D**). Rather, in the SCLC-N lines H29 and H82, there is significant downregulation of *MYCL* upon *SOX2* overexpression (**Fig. 5B**) and an apparent, but not significant increase in the protein levels of MYC ($P = 0.0724$), perhaps not reaching significance due to the already elevated levels of MYC in these cell lines (Supplementary Fig. S7). This indicates that overexpression of *SOX2*, in contrast to normal levels of expression (**Fig. 5A**), is repressive at either *MYC* or *MYCL* yet still favoring *MYCL* expression in the SCLC-A subtype and *MYC* expression in SCLC-N. We tested for either ASCL1, NEUROD1 or MYC expression in the tumors from the RPR2S mice, and observed that $Sox2^{+}$ tumors display high ASCL1 and low NEUROD1/MYC staining, which is expected as the RPR2 mice predominantly form tumors of the SCLC-A subtype (15). However, the few $Sox2^{lox/lox}$ tumors showed reduced ASCL1 staining and increased NEUROD1/MYC immunoreactivity (**Fig. 5E**). ASCL1, NEUROD1, and MYC staining showed nuclear localization consistent with SCLC cells and not infiltrating cells (Supplementary Fig. S9). Blinded scoring of the tumors as either $ASCL1^{+}$ $NEUROD1^{+}$, or MYC^{+} showed a significant increase in the number of $NEUROD1^{+}/MYC^{+}$ tumors from the $Sox2^{lox/lox}$ mice (**Fig. 5F**; Supplementary Table S2). Therefore, it appears that SOX2 favors the formation of an SCLC-A subtype.

Discussion

There have been a few indications that *SOX2* may be a key factor in SCLC; however, its role in SCLC has so far been obscure. Rudin and colleagues showed that *SOX2* is amplified in approximately 27% of patients with SCLC and cell lines, and that knockdown of *SOX2* can impair growth of SCLC cell lines (7). We have previously shown that *RB1* loss, one of the two driver mutations required for SCLC initiation, can result in *SOX2* upregulation (23). *SOX2* has been observed to be misregulated in various cancers of the epithelium (66). As SCLC is a cancer that rises from the lung epithelium, predominantly from pulmonary neuroendocrine cells which themselves express *SOX2* during development, it seemed reasonable that *SOX2* may indeed be a driver of SCLC (29, 67). However, the role for *SOX2* in SCLC initiation and its mechanism in SCLC was unclear.

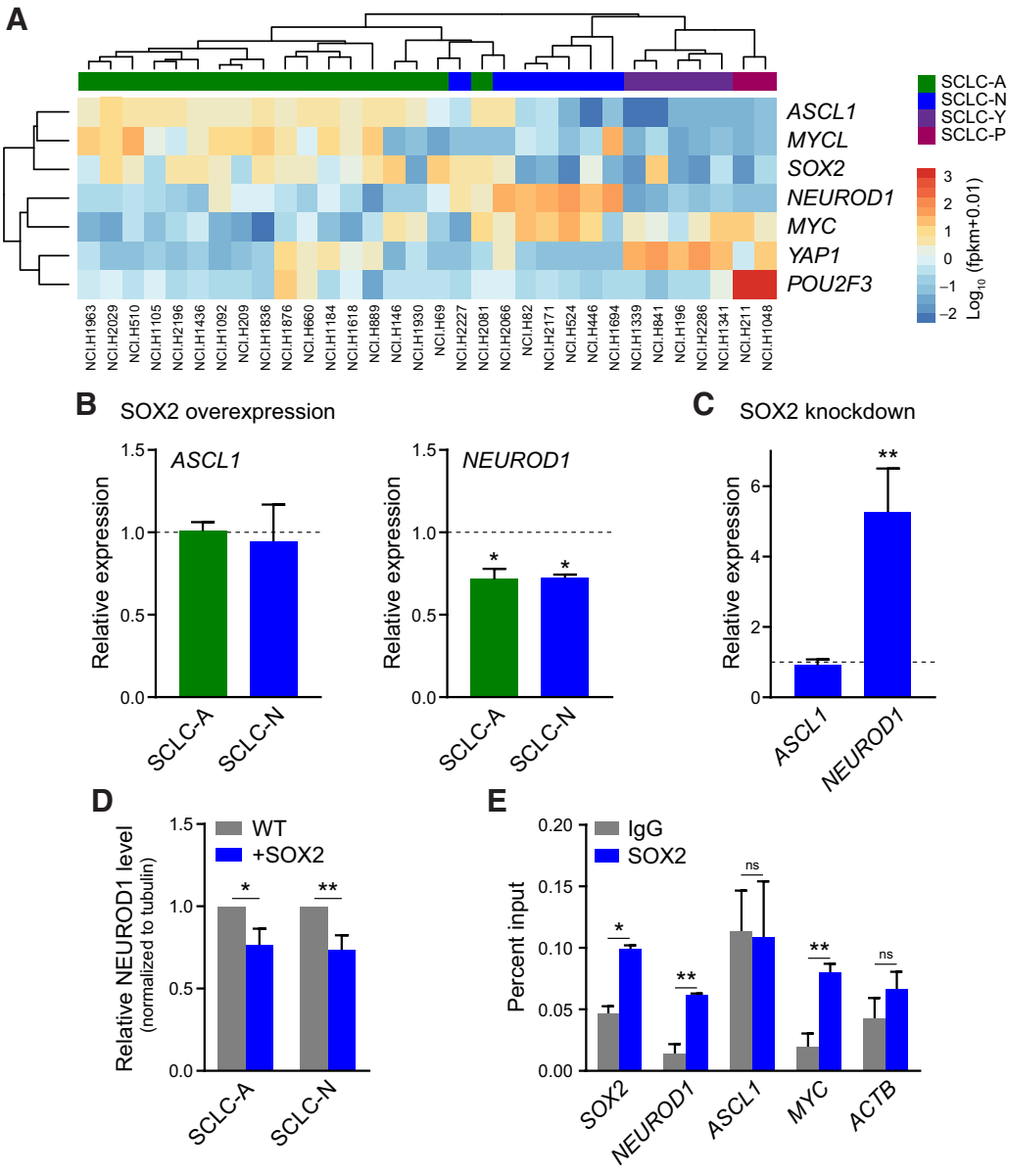

Figure 4.

SOX2 partially regulates *NEUROD1*. **A,** Heatmap of the log transformed fpkm values from SCLC cell lines from CCLE. Cell lines are clustered independently from *SOX2* expression. **B,** Transfection of H1836 and H209 (SCLC-A) and H29 and H82 (SCLC-N) with *SOX-t2a-GFP*. qPCR of *ASCL1* and *NEUROD1* are shown. **C,** qPCR of *ASCL1* and *NEUROD1* are shown upon Cas9-mediated knockdown of *SOX2*. **D,** Quantitation of NEUROD1 protein levels as assessed by Western blotting, $n = 3$. **E,** ChIP of SOX2 or an IgG control assessed by qPCR at *SOX2*, *NEUROD1*, *ASCL1*, *MYC*, and *ACTB* as a negative control. Bar graphs show mean and SEM, significance for all panels determined by a two-tailed *t* test where * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.01$.

To that end, we generated a genetically engineered mouse model of SCLC based on the RPR2 [$Rb1^{lox/lox}$; $p53^{lox/lox}$; $Rbl2(p130)^{lox/lox}$] line, where we introduced a conditional $Sox2^{lox/lox}$ allele (named the RPR2S line). We observed that deletion of *Sox2* in these mice greatly hampers the formation of SCLC tumors. The requirement of SOX2 in SCLC formation was not completely penetrant, however, as there were a handful of small tumors that developed in the absence of *Sox2*. These tumors had properties similar to the SCLC-N subtype as they showed low levels of ASCL1 and high NEUROD1 and MYC. Therefore, SOX2 may be required primarily for SCLC-A type tumors, which are the primary subtype of the RPR2 line, and that any escapees were able to activate *Neurod1* subtype networks to compensate and/or bypass the *Ascl1* state.

To assess the function of SOX2, we assessed its genomic localization and observed that SOX2 primarily binds to genes involved in neurogenesis, where neural gene signatures are commonly found in SCLC (10, 11). Intriguingly, the genes bound by SOX2 did not strictly overlap with SOX2 binding profiles in either pluripotent cells (ES and iPS cells) or NSCs. Rather the SOX2 binding profile was most similar to glioblastoma multiforme, indicating that SOX2 may share a more

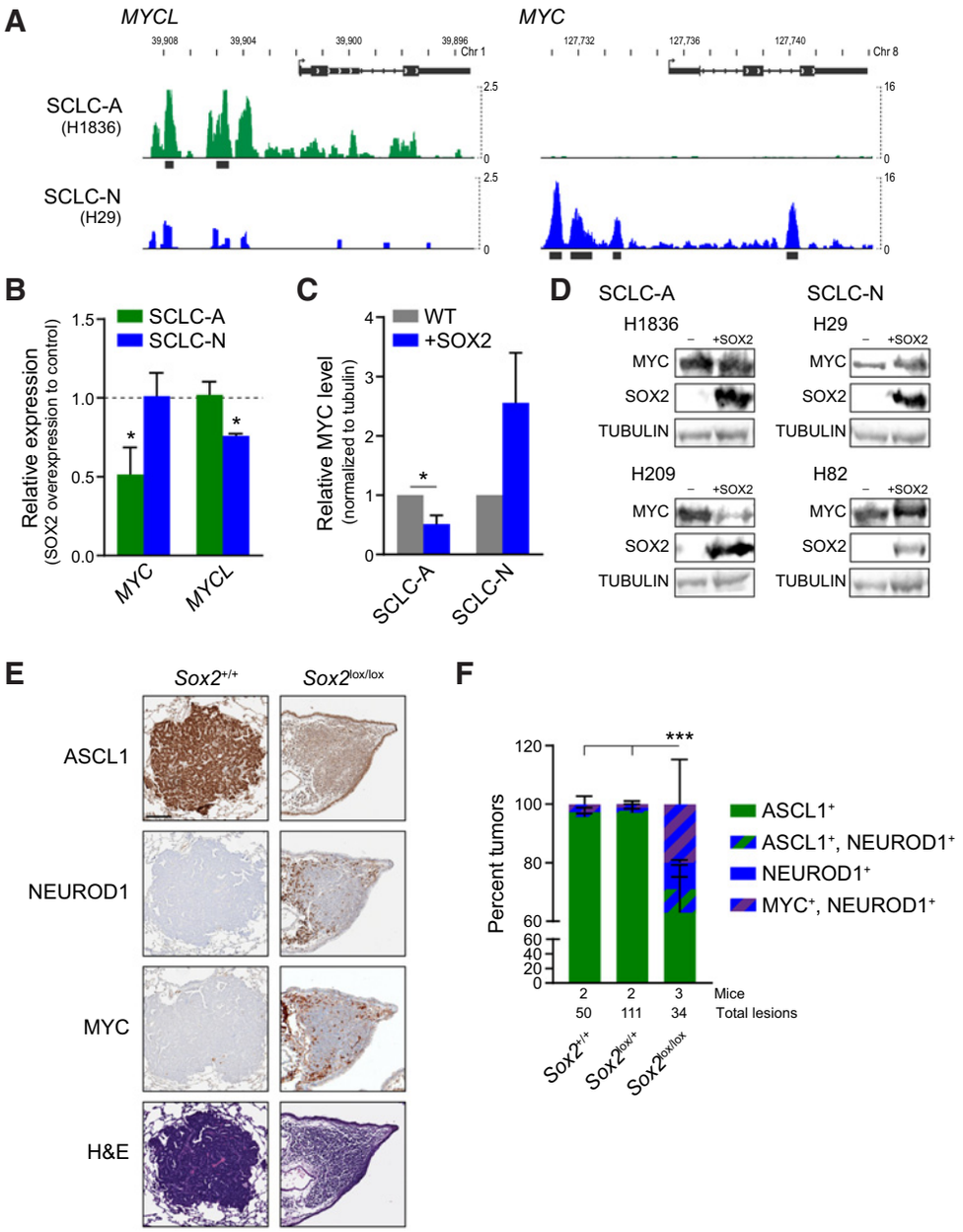

Figure 5.

SOX2 is a regulator of *MYC* and *MYCL* in SCLC. **A,** Gene track showing SOX2 CUT&RUN at the *MYC and MYCL* loci in H1836 and H29 cells. Blue/green track represents the normalized read maps across the loci, the black bars under the track represent regions where significant peaks were called. **B,** qPCR of *MYC* and *MYCL* in H1836 and H209 (SCLC-A) and H29 and H82 (SCLC-N) plotted as a $\log_2$ ratio of *SOX2* overexpressed cells to the mock control. Values greater than 1 indicate higher expression upon *SOX2* overexpression. Significance determined by a two-tailed *t* test. **C,** Quantitation of MYC protein levels as assessed by Western blotting as in (**D**), $n = 3$. **D,** Western blot of MYC, SOX2, and TUBULIN after *SOX2* overexpression or mock-transfected cells as a control. **E,** IHC of ASCL1, NEUROD1, and MYC in murine SCLC tumors. Representative tumors shown. Scale bar, 100 μm. **F,** Quantification of tumors scored as ASCL1$^+$, NEUROD1$^+$, or MYC$^+$ expressing in **E**. Number of tumors and their staining classifications are notated in Supplementary Table S2. Significance assessed by ANOVA. Bar graphs show mean and SEM, significance identified where $^* = P < 0.05$, $^{**} = P < 0.01$, $^{***} = P < 0.01$.

common function amongst cancer than its well studied functions in development. This is perhaps unexpected as SOX2 has been described as a pioneer factor that is able to bind its target DNA sequences regardless of any regional heterochromatin, and therefore should be able to regulate target sequences in a wide assortment of donor cells (68). Rather we observe that the cellular context does impart some level of regulation on the broader SOX2 network. This is particularly relevant considering that SCLC can arise from a few different cell types on the lung epithelium and can influence the resulting SCLC subtype (18, 29, 69). It is possible that the few

NEUROD1^{+}/MYC^{+} lesions observed in the *Sox2*$^{lox/lox}$ mice are a result of tumors initiating from a nonneuroendocrine lineage. Finally, what cell-type–specific factors may be constraining SOX2 function will be of particular importance towards understanding SOX2 regulation in SCLC, and potentially provide novel avenues for therapeutic targeting SCLC, and perhaps other SOX2-driven cancers.

We observed two regulator networks that correlate with *SOX2* expression in SCLC. The first is ASCL1 that is required for SCLC formation in the RPR2 mouse model, and indeed is localized at *SOX2* indicating a direct role in SOX2 regulation (11). Consistent with ASCL1 lying upstream of *SOX2* in established SCLC cell lines, we observe that neither overexpression nor knockdown of *SOX2* alters *ASCL1* expression. This prompts the question of how *ASCL1* can lie upstream of *SOX2* if *SOX2* upregulation is a direct consequence of *RB1* loss, one of the two SCLC driver mutations. It could be that *RB1* loss promotes the derepression of *SOX2*, but *ASCL1* activity is required for full *SOX2* transactivation and subsequent tumor development. Intriguingly, ASCL1 and SOX2 have been found at similar enhancer regions (70), therefore the regulation of these two factors may not be strictly linear. Further investigation into the genetic networks at play in early SCLC tumors will be required to address these questions.

With the potential link between *SOX2* activity and *ASCL1*, we also investigated the other neuroendocrine SCLC subtype specific factor, *NEUROD1*. SOX2 has been found to regulate *Neurod1* in neural progenitor cells, where it functions to maintain an epigenetically permissive state at the *Neurod1* promoter (71). Conversely, in neural stem cells of the adult hippocampus, it was observed that SOX2 binds to the *Neurod1* promoter and silences *Neurod1* expression (72). In SCLC, we observe that SOX2 overexpression leads to *NEUROD1* silencing, while basal levels of SOX2 appear to be associated with activation or attenuation of the levels of activated *NEUROD1*. This regulation appears direct as we observe SOX2 bound at the *NEUROD1* promoter by ChIP, although binding at *NEUROD1* was unclear in the CUT&RUN data. It is possible that these two techniques may recognize different SOX2 protein complexes due to their differing methods to assess DNA localization. As *MYC* is a target of NEUROD1 (11), SOX2 loss could then promote a maintenance of the SCLC-N subtype network.

We also uncovered a role of SOX2 in the regulation of *MYC* and *MYCL* in SCLC. We observe that endogenous levels of SOX2 appear associated with activation as SOX2 was found at *MYCL* in SCLC-A subtype cell lines while it was bound at *MYC* in SCLC-N cell lines. Yet, in contrast we observe that overexpression of SOX2 enhanced repression of *MYC* and *MYCL* in SCLC-A and SCLC-N, respectively. As was shown for SOX2 in embryonic stem cells (61), we also observe that SOX2 can be associated with both gene activation and gene silencing. The alternating functions of SOX2 of both gene activation or repression most likely reflect differing SOX2 protein complexes that are assembled in a context-specific manner, with tight stoichiometric regulation of the endogenous activating complex so that overexpressed SOX2 favors the formation of a more promiscuous repressive complex. Further investigation into the SOX2 protein interactome in SCLC and specifically in different SCLC subtypes will be required to delineate the mechanistic function of SOX2 on different gene targets. SOX2, while typically oncogenic in the lung (73, 74), can indeed act as a tumor suppressor when overexpressed in multiple cancer types (75) indicating cell type–specific roles. Consequently, SOX2 may possess differing functions, either favoring transcriptional activation or silencing in different cells within a single SCLC tumor, or tumors that arise from alternative cells of origin as SCLC is indeed a heterogeneous tumor comprised of multiple cell types responsible for tumor propagation and treatment resistance (19, 76–79). Further investigation into the mechanism of SOX2 activity in these different cell types may shed additional light on the development of SCLC heterogeneity and treatment resistance.

Together we have illustrated that *SOX2* is strongly favorable to SCLC formation in the RPR2 SCLC mouse model. SOX2 serves to regulate *NEUROD1* expression and is associated with the switch from *MYCL* to *MYC* expression, although further investigations into its regulatory mechanisms of this switch are required. ASCL1 is the predominant network controlling SCLC activity in the early tumor; however, during tumor progression there is a switch to the *NEUROD1* state, driven in part by MYC and is linked with poorer patient outcomes (18, 79). Our data indicates that *SOX2* is associated with this process by the concurrent regulation of *NEUROD1*, *MYC*, and *MYCL*. Understanding the genetic networks that underlie this switch during SCLC tumor progression will add to the explanation of such processes as treatment resistance, and ultimately lead to improved therapies to treat this devastating disease.

Authors' Disclosures

No disclosures were reported.

Authors' Contributions

E. Voigt: Conceptualization, resources, data curation, formal analysis, validation, investigation, visualization, methodology, writing–original draft, writing–review and editing. **M. Wallenburg:** Conceptualization, resources, data curation, formal analysis, validation, investigation, visualization, methodology, writing–original draft, writing–review and editing. **H. Wollenzien:** Conceptualization, resources, data curation, formal analysis, validation, investigation, visualization, methodology, writing–original draft, writing–review and editing. **E. Thompson:** Conceptualization, formal analysis, investigation, methodology. **K. Kumar:** Formal analysis. **J. Feiner:** Formal analysis, investigation. **M. McNally:** Formal analysis, investigation. **H. Friesen:** Formal analysis, investigation. **M. Mukherjee:** Formal analysis, investigation. **Y. Afeworki:** Data curation, software, formal analysis. **M.S. Kareta:** Conceptualization, resources, data curation, software, formal analysis, supervision, funding acquisition, validation, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing.

Acknowledgments

We would like to acknowledge the NIH NIGMS Center for Pediatric Research 5P20GM103620 and the NIH NCI/NIGMS grant R01CA233661 for research support (to M.S. Kareta). M. McNally was supported by NIH NICHD grant R25HD097633. The Sanford Research Pathology and Flow Cytometry Cores are supported by the NIGMS Center for Cancer Research P20GM103548. H. Wollenzien is thankful to the University of South Dakota-Neuroscience, Nanotechnology and Networks (USD-N3) Program, supported by a grant from the National Science Foundation Research Traineeship program DGE-1633213. We are grateful to Ryan Askeland M.D. (Sanford Health Pathology Clinic) for the unbiased and blinded characterization of SCLC tissue samples. KP1, KP3, and NJH29 cells were a kind gift from Julien Sage. *Sox2* hairpin vectors were a kind gift from Alejandro Sweet-Cordero. pCMV-HA-hRb1-delta-CDK was a gift from Steven Dowdy (Addgene plasmid #58906), SOX2-t2A-GFP was a gift from Jennifer Mitchell (Addgene plasmid #127537), and TLCV2 was a gift from Adam Karpf (Addgene plasmid #87360).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received November 25, 2020; revised June 2, 2021; accepted September 24, 2021; published first September 30, 2021.

References

1. Gazdar AF, Bunn PA, Minna JD. Small-cell lung cancer: what we know, what we need to know and the path forward. Nat Rev Cancer 2017;17:725–37.
2. Byers LA, Rudin CM. Small cell lung cancer: where do we go from here? Cancer 2015;121:664–72.
3. Pietanza MC, Byers LA, Minna JD, Rudin CM. Small cell lung cancer: will recent progress lead to improved outcomes? Clin Cancer Res 2015;21:2244–55.
4. Sandler AB. Chemotherapy for small cell lung cancer. Semin Oncol 2003;30: 9–25.
5. Berns A. The therapy escapes of small-cell lung cancer. Nat Cancer 2020;1:374–5.
6. Wu Y, Liu Y, Sun C, Wang H, Zhao S, Li W, et al. Immunotherapy as a treatment for small cell lung cancer: a case report and brief review. Transl Lung Cancer Res 2020;9:393–400.
7. Rudin CM, Durinck S, Stawiski EW, Poirier JT, Modrusan Z, Shames DS, et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet 2012;44:1111–6.
8. Peifer M, Fernandez-Cuesta L, Sos ML, George J, Seidel D, Kasper LH, et al. Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet 2012;44:1104–10.
9. George J, Lim JS, Jang SJ, Cun Y, Ozretic L, Kong G, et al. Comprehensive genomic profiles of small cell lung cancer. Nature 2015;524:47–53.
10. Augustyn A, Borromeo M, Wang T, Fujimoto J, Shao C, Dospoy PD, et al. ASCL1 is a lineage oncogene providing therapeutic targets for high-grade neuroendocrine lung cancers. Proc Natl Acad Sci U S A 2014;111:14788–93.
11. Borromeo MD, Savage TK, Kollipara RK, He M, Augustyn A, Osborne JK, et al. ASCL1 and NEUROD1 reveal heterogeneity in pulmonary neuroendocrine tumors and regulate distinct genetic programs. Cell Rep 2016;16:1259–72.
12. McColl K, Wildey G, Sakre N, Lipka MB, Behtaj M, Kresak A, et al. Reciprocal expression of INSM1 and YAP1 defines subgroups in small cell lung cancer. Oncotarget 2017;8:73745–56.
13. Huang YH, Klingbeil O, He XY, Wu XS, Arun G, Lu B, et al. POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer. Genes Dev 2018;32: 915–28.
14. Rudin CM, Poirier JT, Byers LA, Dive C, Dowlati A, George J, et al. Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. Nat Rev Cancer 2019;19:289–97.
15. Mollaoglu G, Guthrie MR, Bohm S, Bragelmann J, Can I, Ballieu PM, et al. MYC drives progression of small cell lung cancer to a variant neuroendocrine subtype with vulnerability to Aurora kinase inhibition. Cancer Cell 2017;31:270–85.
16. Kim DW, Wu N, Kim YC, Cheng PF, Basom R, Kim D, et al. Genetic requirement for Mycl and efficacy of RNA Pol I inhibition in mouse models of small cell lung cancer. Genes Dev 2016;30:1289–99.
17. Semenova EA, Kwon MC, Monkhorst K, Song JY, Bhaskaran R, Krijgsman O, et al. Transcription factor NFIB is a driver of small cell lung cancer progression in mice and marks metastatic disease in patients. Cell Rep 2016;16:631–43.
18. Ireland AS, Micinski AM, Kastner DW, Guo B, Wait SJ, Spainhower KB, et al. MYC drives temporal evolution of small cell lung cancer subtypes by reprogramming neuroendocrine fate. Cancer Cell 2020;38:60–78.
19. Shue YT, Lim JS, Sage J. Tumor heterogeneity in small cell lung cancer defined and investigated in pre-clinical mouse models. Transl Lung Cancer Res 2018;7: 21–31.
20. Burkhart DL, Sage J. Cellular mechanisms of tumour suppression by the retinoblastoma gene. Nat Rev Cancer 2008;8:671–82.
21. Chinnam M, Goodrich DW. RB1, development, and cancer. Curr Top Dev Biol 2011;94:129–69.
22. Dyson NJ. RB1: a prototype tumor suppressor and an enigma. Genes Dev 2016; 30:1492–502.
23. Kareta MS, Gorges LL, Hafeez S, Benayoun BA, Marro S, Zmoos AF, et al. Inhibition of pluripotency networks by the rb tumor suppressor restricts reprogramming and tumorigenesis. Cell Stem Cell 2015;16:39–50.
24. Abdelalim EM, Emara MM, Kolatkar PR. The SOX transcription factors as key players in pluripotent stem cells. Stem Cells Dev 2014;23:2687–99.
25. Arnold K, Sarkar A, Yram MA, Polo JM, Bronson R, Sengupta S, et al. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell Stem Cell 2011;9:317–29.
26. Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev 2003;17:126–40.
27. Driessens G, Blanpain C. Long live sox2: sox2 lasts a lifetime. Cell Stem Cell 2011; 9:283–4.
28. Ellis P, Fagan BM, Magness ST, Hutton S, Taranova O, Hayashi S, et al. SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. Dev Neurosci 2004;26:148–65.
29. Sutherland KD, Proost N, Brouns I, Adriaensen D, Song JY, Berns A. Cell of origin of small cell lung cancer: inactivation of Trp53 and Rb1 in distinct cell types of adult mouse lung. Cancer Cell 2011;19:754–64.
30. Schaffer BE, Park KS, Yiu G, Conklin JF, Lin C, Burkhart DL, et al. Loss of p130 accelerates tumor development in a mouse model for human small-cell lung carcinoma. Cancer Res 2010;70:3877–83.
31. Shaham O, Smith AN, Robinson ML, Taketo MM, Lang RA, Ashery-Padan R. Pax6 is essential for lens fiber cell differentiation. Development 2009;136: 2567–78.
32. DuPage M, Dooley AL, Jacks T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. Nat Protoc 2009;4: 1064–72.
33. Cardona A, Arganda-Carreras I, Saalfeld A. Register virtual stack slices. Available from: https://imagej.net/Register_Virtual_Stack_Slices.
34. Kamentsky L, Jones TR, Fraser A, Bray MA, Logan DJ, Madden KL, et al. Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. Bioinformatics 2011; 27:1179–80.
35. Jahchan NS, Dudley JT, Mazur PK, Flores N, Yang D, Palmerton A, et al. A drug repositioning approach identifies tricyclic antidepressants as inhibitors of small cell lung cancer and other neuroendocrine tumors. Cancer Discov 2013;3:1364–77.
36. Tan YS, Sansanaphongpricha K, Xie Y, Donnelly CR, Luo X, Heath BR, et al. Mitigating SOX2-potentiated immune escape of head and neck squamous cell carcinoma with a STING-inducing nanosatellite vaccine. Clin Cancer Res 2018; 24:4242–55.
37. Burkhart DL, Wirt SE, Zmoos AF, Kareta MS, Sage J. Tandem E2F binding sites in the promoter of the p107 cell cycle regulator control p107 expression and its cellular functions. PLoS Genet 2010;6:e1001003.
38. Janssens D, Henikoff S. CUT&RUN: targeted in situ genome-wide profiling with high efficiency for low cell numbers. protocols.io 2019. Available from: https://dx.doi.org/10.17504/protocols.io.zcpf2vn.
39. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife 2017;6:e21856.
40. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–9.
41. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9:R137.
42. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 2010;38:576–89.
43. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 2012;481:389–93.
44. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559.
45. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
46. Gouyer V, Gazzeri S, Bolon I, Drevet C, Brambilla C, Brambilla E. Mechanism of retinoblastoma gene inactivation in the spectrum of neuroendocrine lung tumors. Am J Respir Cell Mol Biol 1998;18:188–96.
47. Meuwissen R, Linn SC, Linnoila RI, Zevenhoven J, Mooi WJ, Berns A. Induction of small cell lung cancer by somatic inactivation of both Trp53 and Rb1 in a conditional mouse model. Cancer Cell 2003;4:181–9.
48. Park KS, Liang MC, Raiser DM, Zamponi R, Roach RR, Curtis SJ, et al. Characterization of the cell of origin for small cell lung cancer. Cell Cycle 2011;10:2806–15.
49. Gierut JJ, Jacks TE, Haigis KM. In vivo delivery of lenti-Cre or adeno-Cre into mice using intranasal instillation. Cold Spring Harb Protoc 2014;2014: 307–9.
50. Drivsholm L, Paloheimo LI, Osterlind K. Chromogranin A, a significant prognostic factor in small cell lung cancer. Br J Cancer 1999;81:667–71.
51. Narasimha AM, Kaulich M, Shapiro GS, Choi YJ, Sicinski P, Dowdy SF. Cyclin D activates the Rb tumor suppressor by mono-phosphorylation. Elife 2014;3: e02872.

52. Nikitin AY, Juarez-Perez MI, Li S, Huang L, Lee WH. RB-mediated suppression of spontaneous multiple neuroendocrine neoplasia and lung metastases in Rb+/- mice. Proc Natl Acad Sci U S A 1999;96:3916–21.
53. Tomioka M, Nishimoto M, Miyagi S, Katayanagi T, Fukui N, Niwa H, et al. Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. Nucleic Acids Res 2002;30:3202–13.
54. Kamachi Y, Kondoh H. Sox proteins: regulators of cell fate specification and differentiation. Development 2013;140:4129–44.
55. Suh H, Consiglio A, Ray J, Sawai T, D'Amour KA, Gage FH. In vivo fate analysis reveals the multipotent and self-renewal capacities of Sox2+ neural stem cells in the adult hippocampus. Cell Stem Cell 2007;1:515–28.
56. Bennett L, Yang M, Enikolopov G, Iacovitti L. Circumventricular organs: a novel site of neural stem cells in the adult brain. Mol Cell Neurosci 2009;41: 337–47.
57. Zhou C, Yang X, Sun Y, Yu H, Zhang Y, Jin Y. Comprehensive profiling reveals mechanisms of SOX2-mediated cell fate specification in human ESCs and NPCs. Cell Res 2016;26:171–89.
58. Narayan S, Bryant G, Shah S, Berrozpe G, Ptashne M. OCT4 and SOX2 work as transcriptional activators in reprogramming human fibroblasts. Cell Rep 2017; 20:1585–96.
59. Ng SY, Bogu GK, Soh BS, Stanton LW. The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. Mol Cell 2013;51:349–59.
60. Fang X, Yoon JG, Li L, Yu W, Shao J, Hua D, et al. The SOX2 response program in glioblastoma multiforme: an integrated ChIP-seq, expression microarray, and microRNA analysis. BMC Genomics 2011;12:11.
61. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 2005; 122:947–56.
62. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 2006;441:349–53.
63. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 2005;4:Article17.
64. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483:603–7.
65. Coles GL, Cristea S, Webber JT, Levin RS, Moss SM, He A, et al. Unbiased proteomic profiling uncovers a targetable GNAS/PKA/PP2A axis in small cell lung cancer stem cells. Cancer Cell 2020;38:129–43.
66. Novak D, Huser L, Elton JJ, Umansky V, Altevogt P, Utikal J. SOX2 in development and cancer biology. Semin Cancer Biol 2020;67:74–82.
67. Gontan C, de Munck A, Vermeij M, Grosveld F, Tibboel D, Rottier R. Sox2 is important for two crucial processes in lung development: branching morphogenesis and epithelial cell differentiation. Dev Biol 2008;317:296–309.
68. Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell 2012;151:994–1004.
69. Yang D, Denny SK, Greenside PG, Chaikovsky AC, Brady JJ, Ouadah Y, et al. Intertumoral heterogeneity in SCLC is influenced by the cell type of origin. Cancer Discov 2018;8:1316–31.
70. Christensen CL, Kwiatkowski N, Abraham BJ, Carretero J, Al-Shahrour F, Zhang T, et al. Targeting transcriptional addictions in small cell lung cancer with a covalent CDK7 inhibitor. Cancer Cell 2014;26:909–22.
71. Amador-Arjona A, Cimadamore F, Huang CT, Wright R, Lewis S, Gage FH, et al. SOX2 primes the epigenetic landscape in neural precursors enabling proper gene activation during hippocampal neurogenesis. Proc Natl Acad Sci U S A 2015;112: E1936–45.
72. Kuwabara T, Hsieh J, Muotri A, Yeo G, Warashina M, Lie DC, et al. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. Nat Neurosci 2009;12:1097–105.
73. Ferone G, Song JY, Sutherland KD, Bhaskaran R, Monkhorst K, Lambooij JP, et al. SOX2 is the determining oncogenic switch in promoting lung squamous cell carcinoma from different cells of origin. Cancer Cell 2016;30:519–32.
74. Lu Y, Futtner C, Rock JR, Xu X, Whitworth W, Hogan BL, et al. Evidence that SOX2 overexpression is oncogenic in the lung. PLoS One 2010;5: e11022.
75. Metz EP, Wuebben EL, Wilder PJ, Cox JL, Datta K, Coulter D, et al. Tumor quiescence: elevating SOX2 in diverse tumor cell types downregulates a broad spectrum of the cell cycle machinery and inhibits tumor growth. BMC Cancer 2020;20:941.
76. Jahchan NS, Lim JS, Bola B, Morris K, Seitz G, Tran KQ, et al. Identification and targeting of long-term tumor-propagating cells in small cell lung cancer. Cell Rep 2016;16:644–56.
77. Stewart CA, Gay CM, Xi Y, Sivajothi S, Sivakamasundari V, Fujimoto J, et al. Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. Nat Cancer 2020;1:423–36.
78. Simpson KL, Stoney R, Frese KK, Simms N, Rowe W, Pearce SP, et al. A biobank of small cell lung cancer CDX models elucidates inter- and intratumoral phenotypic heterogeneity. Nat Cancer 2020;1:437–51.
79. Lim JS, Ibaseta A, Fischer MM, Cancilla B, O'Young G, Cristea S, et al. Intratumoural heterogeneity generated by Notch signalling promotes small-cell lung cancer. Nature 2017;545:360–4.