Published in final edited form as:

*Cancer Res*. 2017 November 15; 77(22): 6267–6281. doi:10.1158/0008-5472.CAN-17-0570.

# ASXL3 is a Novel Pluripotency Factor in Human Respiratory Epithelial Cells and a Potential Therapeutic Target in Small Cell Lung Cancer

**Vivek Shukla**[1], **Mahadev Rao**[1], **Hongen Zhang**[2], **Jeanette Beers**[3], **Darawalee Wangsa**[2], **Danny Wangsa**[2], **Floryne O. Buishand**[2], **Yonghong Wang**[2], **Zhiya Yu**[5], **Holly S. Stevenson**[2], **Emily S. Reardon**[1], **Kaitlin C. McLoughlin**[1], **Andrew S. Kaufman**[1], **Eden C. Payabyab**[1], **Julie A. Hong**[1], **Mary Zhang**[1], **Sean Davis**[2], **Daniel Edelman**[2], **Guokai Chen**[3,*], **Markku M. Miettinen**[4], **Nicholas P. Restifo**[5], **Thomas Ried**[2], **Paul A. Meltzer**[2], and **David S. Schrump**[1,#]

[1]Thoracic Epigenetics Section, Thoracic and Gastrointestinal Oncology Branch, National Cancer Institute, Bethesda, MD

[2]Genetics Branch, National Cancer Institute, Bethesda, MD

[4]Laboratory of Pathology, National Cancer Institute, Bethesda, MD

[5]Surgery Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD

[3]NHLBI iPSC Core, National Institutes of Health, Bethesda, MD

## Abstract

In this study, we generated induced pluripotent stem cells (iPSC) from normal human small airway epithelial cells (SAEC) to investigate epigenetic mechanisms of stemness and pluripotency in lung cancers. We documented key hallmarks of reprogramming in lung iPSC (Lu-iPSC) that coincided with modulation of more than 15,000 genes relative to parental SAEC. Of particular novelty, we identified the PRC2-associated protein, ASXL3 which was markedly upregulated in Lu-iPSC and small cell lung cancer (SCLC) lines and clinical specimens. ASXL3 overexpression correlated with increased genomic copy number in SCLC lines. ASXL3 silencing inhibited proliferation, clonogenicity and teratoma formation by Lu-iPSC, and diminished clonogenicity and malignant growth of SCLC cells in-vivo. Collectively, our studies validate the utility of the Lu-iPSC model for elucidating epigenetic mechanisms contributing to pulmonary carcinogenesis, and highlight ASXL3 as a novel candidate target for SCLC therapy.

### Keywords

iPSC; ASXL3; lung cancer; small cell lung cancer; epigenetics

[#]Corresponding Author: David S. Schrump, MD, MBA, FACS, Thoracic and GI Oncology Branch, CCR/NCI, Building 10; 4-3942, 10 Center Drive, MSC 1201, Bethesda, MD 20892-1201, Tel: 301-496-2128, Fax: 301-451-6934, david.schrump@nih.gov.
[*]Present address: University of Macau, China

## Introduction

Lung cancer is the leading cause of cancer-related mortality worldwide, and presently accounts for ~160,000 deaths annually in the United States (1). Approximately 85% of pulmonary carcinomas are non-small cell lung cancers (NSCLC) comprised primarily of adeno-, squamous cell and large cell undifferentiated carcinomas with distinct histologic and molecular features (2). The remaining 15% are small cell lung cancers (SCLC) which exhibit varying degrees of neuroendocrine differentiation, and are typically widely metastatic at presentation with a high propensity for recurrence despite initial, often dramatic responses to chemotherapy (3). Whereas CT screening, targeted therapies, and immune checkpoint inhibitors have recently improved outcomes for some NSCLC patients (1), there have been no advances in detection or treatment of SCLC in over 30 years (3). As such, SCLC has been designated a NCI priority (http://www.lungcanceralliance.org/News/SCLC%20Congressional%20Response%206-30-14%20FINAL%20with%20appendices.pdf).

Although lung cancers may occasionally arise in never-smokers, cigarette smoking remains the predominant risk factor for NSCLC (4); virtually all SCLC arise in active or former smokers with substantial tobacco exposure (3). Despite extensive studies, the genetic and epigenetic mechanisms by which cigarette smoke mediates initiation and progression of lung cancers have not been fully elucidated (5). Previously, we have demonstrated that cigarette smoke condensate (CSC) mediates time and dose dependent epigenomic alterations in normal respiratory epithelial cells (NREC) including perturbations of the histone code, activation of oncomirs and cancer-germline genes, and epigenetic repression of tumor suppressor genes and microRNAs (6–8). These tobacco-induced alterations in NREC mirror those observed in lung cancer cells and primary tumors (9). A central theme emerging from these studies is that via complementary mechanisms including DNA methylation, polycomb repressive complexes (PRC) and noncoding RNAs, cigarette smoke induces stem-like phenotypes that coincide with progression to malignancy in NREC, as well as enhanced growth and metastatic potential of lung cancer cells (6–8, 10–12).

Reprogramming of somatic cells into induced pluripotent stem cells (iPSC) by induction of a defined set of transcriptional factors has opened a new era in regenerative medicine and cancer biology (13–15). iPSC exhibit numerous similarities to embryonic stem cells (ESCs) including morphology, gene expression profiles, in-vitro proliferation, and in-vivo teratoma formation (13, 16). As such, iPSC have emerged as important models with which to study epigenetic mechanisms contributing to pluripotency in normal cells, as well as stemness, chemo-resistance and metastatic phenotypes of cancer cells (17).

In an effort to further delineate epigenetic mechanisms contributing to the pathogenesis of lung cancers and possibly identify novel targets for lung cancer therapy, we sought to develop an iPSC model that might reflect epigenetic transitions from normal to malignant respiratory epithelia. Herein we describe reprogramming of normal human respiratory epithelial cells to pluripotency, and demonstrate that *Additional Sex Combs Like-3 (ASXL3)* is a novel factor critical for maintenance of pluripotent respiratory epithelial cells, and a potential therapeutic target in SCLC.

## Materials and Methods

### Cell Lines

All lung cancer lines were available in repositories at the NCI, or were purchased from ATCC, and cultured as recommended. cdk-4/h-TERT immortalized human bronchial epithelial cells (HBEC) were a generous gift from John Minna (UT-Southwestern, Dallas, TX), and were cultured as described (8). Cell lines were tested for mycoplasma regularly (tested the latest in July, 2017) using a kit from Sigma (Cat. no. MP0025), and were validated by HLA typing relative to original stocks.

### Primary cell culture

Normal human bronchial epithelial cells (NHBE), as well as SAEC derived from a fifty-seven year old Hispanic female non-smoker were purchased from Lonza, and cultured as recommended by the vendor. STEMCCA kit (Millipore, Cat. no. SCR545) was purchased from Millipore, and used as instructed. Irradiated mouse embryonic fibroblasts (MEFs) were obtained from NHLBI core facility, and Matrigel plates (Cat. no. 354230) were purchased from BD Biosciences. Normal foreskin fibroblast (CCD-1079Sk, ATCC Cat. no. CRL-2097) and induced pluripotent cells (ND1.2) derived from foreskin fibroblasts were obtained from the NHLBI core facility, and were grown in DMEM medium and stem cell medium, respectively. 8™ medium (Life technologies; Cat. no. A1517001) and Rho-associated kinase (Rock) inhibitor (Y-27632; Tocris; Cat. no. 1254) were used to culture the stem cells.

### Generation of iPSCs from SAECs

The STEMCCA vector (Supplementary Figure S1A) and protocol described by Beers at al. (18) were used to reprogram SAEC to pluripotency. Briefly, $2.5 \times 10^5$ SAEC were plated in each well of a 6-well plate. Once the cells were approximately 70% confluent, they were transduced with STEMCCA lentivirus using polybrene and left to incubate overnight. The transduced cells were then maintained in reprogramming medium with medium changed on alternate days. Six days after transduction, the cells were trypsinized and replated on irradiated mouse embryonic fibroblasts (MEFs; feeder cells). From this day forward, the cells were maintained in reprogramming medium supplemented with basic fibroblast growth factor. Medium was changed on alternate days and cell colonies grew in size. On day 25 after initial transduction, granular stem like cells were detached from the feeder layer using a P20 pipette, and transferred to Matrigel coated petri dishes for expansion and further analysis.

### Flow cytometry and alkaline phosphatase staining

Lu-iPSCs were trypsinized with TryPLE Express (Thermo Fisher Scientific), and the reaction was stopped by adding medium. The cells were centrifuged, and the pellets re-suspended in 4% paraformaldehyde for 10 minutes at room temperature to fix them. The cells were washed with PBS, and then permeabilized with 0.2% Tween-20 in PBS for 10 minutes at room temperature. Flow cytometry analysis (FACS) of pluripotency markers was performed using anti-OCT4- Alexa fluor 488 (Millipore #FCMAB113A4), anti-SSEA4-FITC (Bio legend #330410), anti-NANOG-Alexa Fluor 488 (Millipore #FCABS352A4) and

anti-Tra-1-60-FITC (Millipore #FCMAB115F). Nonimmune control (Millipore #MABC006F) was used at 0.5 μl per 50 μl reaction. All the FACS analyses were performed on a MACSQuant Flow Cytometer. The iPSC colonies were stained with alkaline phosphatase (BCIP/NBT alkaline phosphatase substrate kit IV, Vector Laboratories #SK-5400).

### Immunofluorescence staining

Expression of pluripotent marker proteins was assessed by immunofluorescence techniques using a Zeiss 780 confocal microscope, optimized for automated imaging. Briefly, cells were fixed in 4% paraformaldehyde and later permeabilized in PBS with 0.2% Triton-100X. After washing, the cells were blocked in 3% BSA. The cells were stained with primary antibodies (SSEA3, SSEA4, TRA-1-60, and TRA-1-81; Supplementary Table S1). Alexa 488 (mouse, rabbit), 555 (mouse, rabbit) were used as secondary antibodies at 1:1000 dilutions. Dapi was used as internal control for immunofluorescence.

### TaqMan quantitative RT-PCR

Total RNA was isolated using TRIzol (Invitrogen). cDNA synthesis was performed using iScript™ cDNA Synthesis (Bio-Rad #170-8891). Quantitative analysis of the genes was carried out in triplicates using an ABI PRISM7500 Sequence Detection System and primers listed in Supplementary Table S1. Copy numbers were calculated as described (10) using appropriate controls for specific genes.

### HLA typing

HLA typing was performed in the NIH Clinical Center HLA laboratory.

### Spectral Karyotyping (SKY)

SKY was performed as previously described (19). Briefly, metaphases were prepared in culture by incubating for approximately one hour in 0.02 mg/ml Colcemid (Invitrogen). Cells were incubated in hypotonic solution (0.075 M KCl) for 20 minutes before being fixed in methanol and acetic acid (3:1). SKY probes were generated in-house, and consisted of a specific combination of labeled chromosome painting probes that were hybridized simultaneously onto metaphase chromosomes. Protocols used for SKY can be accessed at https://ccr.cancer.gov/Genetics-Branch/thomas-ried?qt-staff_profile_tabs=9#qt-staff_profile_tabs. SKY imaging was performed on a Leica DMRXE microscope equipped with a Xenon lamp and Spectracube (Applied Spectral Imaging). For each cell line, 15–20 metaphases were imaged, and analyzed using the SkyView software. Each metaphase was scored for numerical and structural aberrations according to human chromosome nomenclature rules from ISCN (International Standing Committee on Human Cytogenetic Nomenclature 2009).

### Array-comparative genomic hybridization

Isolated DNA was labeled using the Genomic DNAULS Labeling Kit (Agilent) and subsequently hybridized on Agilent SurePrint G3 Human CGH Microarrays 4 × 180 K (Agilent) according to the manufacturer’s protocol version 3.5. Briefly, 1 μg DNA of each

cell line and 1 μg normal genetic reference DNA (Promega) were differentially labeled with ULS-Cy5 and ULS-Cy3 (both Agilent), respectively. After hybridizing and washing according to the manufacturer's instructions, slides were scanned with microarray scanner G256BA (Agilent), and images were analyzed by Feature Extraction software version 10.7.1.1 (Agilent). The aCGH data were visualized and analyzed using Nexus Copy Number software version 7.5 (BioDiscovery). The correlation between average CGH copy number and average gene expression was performed using Pearson's correlation with aCGH $Log_2$ ratio probe mean values as the *X*-axis versus *ASXL3* expression $Log_2$ ratios as the *Y*-axis. The significance threshold was $P < 0.05$ (two-sided t test).

### Teratoma assay

Human iPSCs were grown on BD matrigel-plates. The cells were collected using EDTA-method and resuspended in cold iPSC medium containing matrigel and 10 μM ROCK inhibitor (Tocris). Approximately $3–4\times10^6$ cells from each iPSC colony were injected intramuscularly in the hind leg quadriceps along its long axis of SCID mice (NOD.Cg-*Prkdc*$^{scid}$ Il2$^{rgtm1Wjl}$/SzJ, Stock number-00557, Jackson laboratory) as described previously (18). At the same time, $3\times10^6$ to $5\times10^6$ of parental SAEC cells were injected intramuscularly into other SCID mice. Approximately 15 weeks post injection of cells, teratomas were excised and fixed in 10% buffered formalin. Teratomas were embedded in paraffin and sectioned. Hematoxylin and eosin (H&E) staining as well as immunostaining was performed to identify three germ layers. α-1-fetoprotein (endoderm, A000829, Dako), α-smooth muscle actin (SMA) (mesoderm, Sigma #A5228) and Neuronal Class III β-Tubulin (TUJ1) (ectoderm; MMS-435P Covance) were analyzed.

### In-vitro matrigel clonogenic assay

Clonogenic assay was performed as described (8) with minor modifications. Briefly, 6-well plates were prepared by resuspending Matrigel (BD Biosciences #354230) in cell culture medium without serum and antibiotics. The medium containing Matrigel was evenly distributed in the wells and kept at room temperature in the biosafety cabinet for 30 minutes. Single cell suspension was prepared and plated on the plate. After 7 days, colonies were stained with crystal violet and colonies were counted.

### In vivo tumorigenicity assay

Tumor formation assay was performed as described (10).

### Western blot

Cell pellets were lysed with RIPA buffer containing protease inhibitors. 30 μg of proteins were loaded on 4–14% SDS-PAGE and electroblotted on PVDF membrane. Membranes were blocked in 3% BSA/PBS followed by incubation with primary antibodies in 3% BSA/PBS overnight at 4°C. Membranes were probed using primary antibodies and horseradish peroxidase-conjugated anti-rabbit or mouse IgG (Millipore) secondary antibodies as listed in Supplementary Table S1.

### DNA methylation array

DNA was isolated using the QIAamp DNA Mini Kit (Qiagen #51304), and treated with bisulfite using the EpiTect Bisulfite Kit (Qiagen). DNA methylation profiles were assessed using Infinium HumanMethylation450 BeadChip techniques (Illumina, San Diego, CA. USA) as described (8). The raw data file "FinalReport.txt" generated from Illumina "GenomeStudio" (Illumina) was normalized with "Simple Scaling Normalization" (SSN) method implemented in the "lumi" R package following the color-bias adjustment. Two files were produced, one with beta value for individual target and another one with corresponding "M" values for the beta values. The file with "M" values was used for statistical analysis with FDR<0.05 (adjusted based on Benjamini-Hochberg procedure) as significant cutoff unless otherwise indicated, and also the absolute beta value difference greater than 0.2 from the same contrast comparison was applied. Partek Genomics Suite (Partek Inc.), R packages of lumi, methlumi and other related R packages were used for data processing, analyzing and data presentation. Data have been deposited to: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE102726.

### Pyrosequencing and MSP

Genomic DNA was isolated from SAECs, iPSCs and from other control cell lines using QIAamp DNA Mini Kit (Qiagen #51304). The DNA was treated with sodium bisulfite using EpiTect Bisulfite kit (Qiagen #59104). Pyrosequencing analysis of *LINE-1*, *NBL2*, *D4Z4*, *NANOG*, *OCT4,* and *ASXL3* was performed as previously described (8) using primers listed in Supplementary Table S1. MSP analysis of the *ASXL3* promoter was performed using primers listed in Supplementary Table S1.

### RNA-Seq

Total RNA was isolated using the RNeasy Mini Kit (Qiagen #74104) followed by removal of ribosomal RNA (rRNA) using biotinylated, target-specific oligos combined with Ribo-Zero rRNA removal beads. The RNA was fragmented into small pieces and the cleaved RNA fragments were copied into first strand cDNA using reverse transcriptase and random primers, followed by second strand cDNA synthesis using DNA Polymerase I and RNase H. The resulting double-strand cDNA was used as the input to a standard Illumina library prep with end-repair, adapter ligation and PCR amplification being performed to yield a library that would go to the sequencer. The HiSeq Real Time Analysis software (RTA 1.12) was used for processing image files, the Illumina CASAVA_v1.8.2 was used for demultiplex and converting binary base calls and qualities to fastq format. The sequencing reads were trimmed adapters and low quality bases using Trimmomatic (version 0.3), the trimmed reads were aligned to human hg19 reference genome (GRCh37/UCSC hg19) and Ensembl annotation version 70 using TopHat_v2.0.8 software. RNA-Seq reads mapping and feature count: bam files were generated from fastq files by alignment short reads to hg19 genome (http://hgdownload.soe.ucsc.edu/goldenPath/hg19/) with STAR version 2.4.0a (20). RNA expression was summarized from bam files with featureCounts tools provided by subread software (21) at gene level defined by gene annotation (gencode.v19.annotation.gtf) from GENCODE Project (http://www.gencodegenes.org/releases/current.html). Differential gene expression analysis was performed with R limma package (22). RNA-Seq feature counts of

all samples were quantile normalized first to log2-counts per million (logCPM) and differentially expressed gene list was selected with adjusted *P* value (fdr) from linear-model fit between different groups. All graphics were generated with R. Scatter plot: mean logCPM of different cell lines were plotted with R ggplot2 package and coefficients between two cell lines were calculated with fitting linear model in R stat package. Heatmaps of differentially expressed genes were generated with R gplots package and scaled by gene. Coverage of a single gene: raw short read counts covering a specific gene in different sample bam files were plotted with R Gviz package. Circos plot was generated with R RCircos package. Gene lists were selected from linear model fitting between Lu-iPSC cell lines and SAEC cell lines (11629 genes, adjusted *P* value < 0.05) and heatmap was generated with mean logCPM of each cell line. Data have been deposited to: http://www.ncbi.nlm.nih.gov/bioproject/398652.

### Quantitative ChIP and ChIP-Seq

Cell lines were treated with paraformaldehyde and sonicated to generate DNA fragments. Immune complexes were formed with either nonspecific IgG or chromatin grade antibodies recognizing H3K4me3 and H3K27me3. DNA was eluted and purified from complexes. Quantitative ChIP assay was performed as described (10) using primers listed in Supplementary Table S1.

ChIP-Seq was performed as described (23). Briefly, purified DNA after chromatin immunoprecipitation was sent to Frederick National Laboratory for Cancer Research. For sequencing, the fragments were blunt-end ligated to the Illumina adaptors, amplified and processed by using Hi-Seq 2000 sequencer (Illumina). Sequencing files in Bam format (bam file) were provided by Frederick National Laboratory for Cancer Research and all bam files were mapped to hg19 human genome with software bowtie2. For peak enrichment in each sample, original bam files were directly used. For finding overlapping and differentially bound peaks, bam files from triplicates were merged first with samtools software. Peaks enrichment and annotations were all processed with software HOMER, version 4.7.2, and hg19 genome. R gviz package was used to visualize peaks around specific genes. Data have been deposited to: http://www.ncbi.nlm.nih.gov/bioproject/398652.

## Results

### Generation and Characterization of Lu-iPSCs from SAECs

The overall method of reprogramming SAEC to pluripotency is depicted in Supplementary Figure S1A. Figure 1A images a-f depict parental SAEC and various stages of reprogramming following lentiviral transduction of Yamanaka factors. No iPSC-like colonies were observed in non-transduced SAEC after 30 days of culture in reprogramming medium (Figure 1A; image g). Many lung-iPSC (Lu-iPSC) clones were generated, of which Lu-iPSC6 and Lu-iPSC7 were randomly selected for further expansion and analysis.

### Characterization of Lu-iPSCs

A variety of methods were used to characterize Lu-iPSCs. Initially, histochemical techniques were used to examine alkaline phosphatase (AP) expression, which is typically high in iPSC (24). Lu-iPSC had positive staining for AP within 10 minutes of exposure of cells to alkaline

phosphatase substrate, whereas parental SAEC showed no staining at 24 hours (Supplementary Figure S1B). Immunofluorescence experiments demonstrated increased surface expression of stage specific embryonic antigen (SSEA)-3 and SSEA-4, as well as TRA-1-60 and TRA-1-81 in Lu-iPSC relative to parental SAEC (Figure 1B, and Supplementary Figure S1C); these antigens are known to be expressed at high levels in undifferentiated human ESCs, iPSCs, embryonal carcinoma and embryonic germ cells (13, 24). Similar findings were observed in ND1.2 (control iPSCs established from dermal fibroblasts; Supplementary Figure S1D). Quantitative RT-PCR (qRT-PCR) as well as immunoblot experiments confirmed up-regulation of pluripotency-related genes in Lu-iPSC as well as ND1.2 relative to SAEC (Figures 1C and 1D). Up-regulation of the stemness genes coincided with repression of numerous differentiation-associated genes including *KRT5A, KRT6A*, *MT2A*, *CXCL1*, *HOXA1* and *HOXB2* in Lu-iPSC (Supplementary Figure S2).

Additional experiments were performed to ascertain if the phenotypic changes observed in Lu-iPSC were due to persistent expression of the reprogramming genes. qRT-PCR analysis demonstrated expression of the transduced genes in SAEC six days following lentiviral transduction; however, all four transgenes were silenced in Lu-iPSC6 and Lu-iPSC7 (Figure 1E).

The initial stages of reprogramming of SAEC required the use of irradiated MEF feeder layers. These feeder cells did not proliferate, and remained adherent to the culture dish while the Lu-iPSC were harvested. Despite the low likelihood of hybridoma formation or simultaneous outgrowth of human and murine cells, additional experiments were performed to confirm that the Lu-iPSC were pure derivatives of SAEC, and ascertain if the Lu-iPSC exhibited any genomic aberrations. PCR based HLA typing (assays performed in the Department of Transfusion Medicine, Clinical Center, NIH) demonstrated identical results for Lu-iPSC clones 6 and 7 and parenteral SAECs (Supplementary Figure S3A), indicating no contamination with murine DNA. Spectral karyotyping (SKY) experiments (Supplementary Figure S3B) revealed no chromosomal abnormalities in Lu-iPSC relative to parental SAEC. In contrast, numerous chromosomal aberrations were observed in A549 and H1299 lung cancer cells. Collectively, these experiments confirmed that the Lu-iPSC were derived only from SAEC, and that the reprogramming had not induced major chromosomal alterations.

## Lu-iPSC form teratomas in SCID mice

Additional experiments were performed to ascertain if the reprogrammed respiratory epithelial cells formed teratomas in immunodeficient SCID mice, the “gold standard” for pluripotency (13, 25). Briefly, Lu-iPSC clones 6 and 7 or parental SAEC were implanted with matrigel bilaterally into the superficial layers of the gastrocnemius muscles (five mice per group). All mice implanted with Lu-iPSC had tumors evident approximately 12 weeks after inoculation which were collected for analysis approximately three weeks later. In contrast, no tumors were detected in mice inoculated with parental SAEC following 20 weeks of observation (results summarized in Figure 2A; upper panel). Histologic examination of Lu-iPSC-derived tumors demonstrated immature teratomas, containing

neural elements (ectoderm), cartilage and skeletal muscle (mesoderm), and kidney and intestinal tissue (endoderm) (Figure 2A; lower panel, and Supplementary Figure S4A). Subsequent immunofluorescence experiments demonstrated expression of βIII tubulin (ectodermal origin), α-smooth muscle actin (mesodermal origin) and α–fetoprotein (endodermal origin) in all of the harvested teratomas (Figure 2B and Supplementary Figure S4B).

## Epigenetic Alterations in Lu-iPSC

Previously we have demonstrated genomic hypomethylation as evidenced by DNA demethylation of LINE-1, NBL2 and D4Z4 repetitive elements in association with up-regulation of cancer-testis (cancer germline) genes in lung cancer cells, as well as immortalized respiratory epithelial cells following cigarette smoke exposure (8). As such, pyrosequencing and qRT-PCR experiments were performed to examine DNA methylation of LINE-1, NBL2 and D4Z4 elements, as well as expression status of several cancer-germline genes typically de-repressed in lung cancer cells. Pyrosequencing did not demonstrate demethylation of these repetitive elements in Lu-iPSCs compared to SAECs, possibly due to the limited number of CpG sites interrogated with these assays (Supplementary Figure S4C). qRT-PCR experiments demonstrated very modest increases in *NY-ESO-1*, *MAGE-A1* and *MAGE-A3* expression in Lu-iPSC relative to parental SAEC (Figure 2C; left panel). Consistent with these findings, quantitative chromatin immunoprecipitation (qChIP) experiments (Figure 2C; right panel) demonstrated decreased levels of the repressive histone mark H3K27me3 without corresponding increases in H3K4me3 (histone activation mark) in the *NY-ESO-1*, *MAGE-A1* and *MAGE-A3* promoters, indicative of incomplete de-repression of these genes following reprogramming. In contrast, pyrosequencing and ChIP-seq experiments demonstrated that the *NANOG* and *OCT4* promoters were significantly hypomethylated, with decreased H3K27me3 and increased H3K4me3 levels in Lu-iPSC relative to primary SAEC (Figures 2D and 2E). These findings were consistent with results of aforementioned qRT-PCR and immunoblot experiments.

To more comprehensively examine DNA methylation changes associated with reprogramming of respiratory epithelial cells, micro-array experiments were performed using Illumina Infinium 450 arrays (Infinium Human Methylation 450 Bead Chip), which can interrogate 480,000 CpG sites throughout the genome (99% of RefSeq genes and 96% of CpG islands). This analysis demonstrated significant differences in DNA methylation profiles in Lu-iPSC compared to SAEC with a shift toward a hypermethylated genome in the reprogrammed cells (Figure 3A). Additional analysis was performed to ascertain if this hypermethylation was related to location of CpG sites (i.e., gene body, promoter, CpG island vs non-CpG island). CpG targets in non-promoter regions were hypermethylated in Lu-iPSC relative to SAEC, with a significant percent appearing to be fully methylated (Figure 3B). Targets in promoter regions were also hypermethylated in Lu-iPSC relative to SAEC, with some sites appearing to be hemimethylated, whereas others exhibited bi-allelic methylation (Figure 3C). A similar phenomenon was observed for CpG sites in CpG islands (Figure 3D). Lastly, CpG sites in CpG islands within promoters were also hypermethylated in Lu-iPSC compared to SAEC, with many of these sites being hemimethylated (Figure 3E).

In order to ascertain if the hypermethylated phenotype was unique to Lu-iPSC, unsupervised analysis was performed using 20,769 CpG targets with variance greater than 0.01 in several iPSC and ESC relative to normal somatic cells in the GSE31848 database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31848). This analysis revealed a shift to a hypermethylated phenotype in iPSC and ESC relative to normal somatic cells. A similar phenomenon was observed for the same 20,769 targets in Lu-iPSC relative to SAEC (Figure 3F). Subsequent examination of non-CPG methylation (CHH) sites demonstrated a similar shift to a more methylated phenotype in Lu-iPSCs relative to parental SAEC (Figure 3G), consistent with what was observed following analysis of pluripotent cell line data in the public database.

Because PRC2 has been shown to be a critical mediator of pluripotency in normal as well as cancer stem cells (26, 27), additional experiments were performed to examine expression of core components of this complex. qRT-PCR and immunoblot experiments (Figures 4A and 4B) demonstrated significant increases in *EZH2*, *EED* and *SUZ12* expression in Lu-iPSC relative to SAEC. Consistent with these findings, qRT-PCR and immunoblot experiments (Figures 4C and 4D) demonstrated down-regulation of *DKK1*, *CDKN1A* and *CDKN2A*, which are well established PRC2 targets that are known to be repressed in iPSC as well as ESC, and are frequently downregulated in lung cancer cells. DNA methylation array and RNA-seq analysis demonstrated that up-regulation of PRC2 components coincided with marked alterations in DNA methylation and expression of numerous PRC2 target genes in Lu-iPSC (representative results are depicted in Figure 4E; full DNA heat map and corresponding list of PRC2 targets are provided in Supplementary Figures S4D and S4E, and Supplementary Table S2, respectively). DNA methylation tended to be associated with repression of the respective PRC2 targets genes, although this was not observed in all cases (Figure 4E).

## ASXL3 is up-regulated and contributes to pluripotency in Lu-iPSC

Additional RNA-seq analysis of more than 33,000 factors revealed that 18,983 genes were significantly altered in Lu-iPSC clone 6, whereas 18,068 genes were altered in Lu-iPSC clone 7 relative to parental SAEC (Figure 5A). RNA-seq data showed that 15,453 genes were commonly altered in Lu-iPSC6 and Lu-iPSC7 (significance adj. $P < 0.05$; Figure 5A); 1,851 genes were commonly altered more than 4-fold in Lu-iPSCs relative to SAECs (significance adj. $P < 0.05$; Figure 5B). Differences in gene expression between the two Lu-iPSC clones which were derived from the same SAEC stock may be related to sites of transgene integration and plasticity of the reprogrammed cells.

CIRCOS and scatter plot analysis of whole genome data demonstrated that *Additional Sex Combs Like 3 (ASXL3)*, a mammalian homologue of drosophila *Additional Sex Combs (Asx)* (28, 29), was highly up-regulated in Lu-iPSC relative to SAEC (Figure 5C). RNA-seq and ChIP-seq data pertaining to ASXL3 in Lu-iPSC and SAEC are depicted in Figure 5D and 5E. qRT-PCR and immunoblot experiments demonstrated over-expression of ASXL3 in Lu-iPSCs and their respective teratomas relative to SAECs (Figure 5F). DNA methylation arrays, as well as pyrosequencing and methylation specific PCR (MSP) experiments which interrogated three regions of a potential CpG island within the ASXL3 promoter revealed no

significant differences in DNA methylation in Lu-iPSC compared to SAEC (Supplementary Figure S5A); consistent with these findings, deoxyazacytidine exposure was insufficient to up-regulate ASXL3 in SAEC or HBEC (Supplementary Figure S5B).

Additional experiments were performed to examine if ASXL3 contributes to pluripotency in Lu-iPSC. Briefly, Lu-iPSC7 cells were transduced with lentiviral shRNAs targeting *ASXL3*. qRT-PCR and immunoblot experiments demonstrated that two of five lentiviral constructs mediated 50–70% reductions of *ASXL3* expression (Figure 5G). The phenotypic effects of *ASXL3* knock-down in Lu-iPSCs were evident within 72 hours following lentiviral transduction; relative to Lu-iPSC transduced with control lentivirus, *ASXL3* knock-down cells grew more slowly and lost pluripotent morphology (Figure 5H). Knock-down of *ASXL3* significantly decreased soft agar colony formation by Lu-iPSC (Figure 5I). Furthermore, knock-down of *ASXL3* decreased the number and size of teratomas in immunodeficient mice (Figure 5J). Notably, ASXL3 expression levels in teratomas from ASXL3 knock-down Lu-iPSC were similar to control teratomas (Supplementary Figure S5C), suggesting that ASXL3 expression is indispensable for maintenance of pluripotency in respiratory epithelial cells.

## ASXL3 Functions as an Oncogene in Lung Cancer Cells

Examination of the Oncomine (https://www.oncomine.org), GENT (http://medicalgenome.kribb.re.kr/GENT/search/g570_gsymbol_ASXL3_20161224_2OQ275_tissue.html), and Cancer Cell Line Encyclopedia (CCLE; http://software.broadinstitute.org/software/cprg/?q=node/11) databases revealed that *ASXL3* is expressed in some human lung cancers. As such, qRT-PCR and immunoblot experiments were performed to examine expression of ASXL3 in a panel of NSCLC and SCLC cell lines relative to normal or immortalized human bronchial epithelial cells. As shown in Figure 6A, *ASXL3* was not expressed in SAEC, short-term normal human bronchial epithelial cells (NHBE), or SV-40 or cdk4/h-TERT-immortalized respiratory epithelial cells (BEAS and HBEC, respectively). Furthermore, ASXL3 expression was either absent or very low in NSCLC lines. In contrast, ASXL3 was expressed in virtually all of the SCLC lines examined, with some lines exhibiting very high ASXL3 mRNA as well as protein levels. Consistent with these findings, immunohistochemistry (IHC) analysis demonstrated significantly higher ASXL3 expression in SCLC relative to NSCLC or normal lung (Figure 6B and 6C); ASXL3 immunoreactivity in SCLC was predominantly nuclear (Figure 6B).

Additional experiments were performed to ascertain the potential relevance of *ASXL3* expression in lung cancer cells. Briefly, H69 SCLC cells were transduced with either lentiviral shRNA targeting *ASXL3* or control lentiviral vector. qRT-PCR and immunoblot analysis demonstrated that lentiviral shRNA mediated 40–60% knockdown of *ASXL3* in H69 cells (Figure 6D). Knock-down of *ASXL3* significantly decreased soft agar colony formation (Figure 6E), and markedly diminished growth of H69 xenografts in athymic nude mice (Figure 6F and 6G), suggesting that *ASXL3* functions as an oncogene in SCLC.

### Influence of genomic imbalances on ASXL3 expression in SCLC Lines

Previously, it has been demonstrated that gene expression levels are directly dependent on chromosomal aneuploidies in colorectal and other carcinomas (30–32). Whereas the strongest correlations have been found between genomic copy number and average chromosome-wide expression levels; the expression of individual genes has also correlated with genomic copy numbers (33). Because SCLC exhibit a variety of amplifications and chromosomal fusions which directly influence expression levels of genes such as *MYC* family members and *SOX2* (34, 35), CGH array (aCGH) experiments were performed to examine if *ASXL3* mRNA expression levels in SCLC were associated with alterations in gene copy number. As shown in Figure 6H, a significant correlation was observed between genomic copy number and *ASXL3* expression levels in SCLC lines (R = 0.78420, $P = 0.007$). Similar analysis revealed no increase in *ASXL3* copy number in Lu-iPSC6 and Lu-iPSC7 relative to parental SAEC. Collectively, these data suggest that genomic amplification contributes to *ASXL3* over-expression in SCLC.

## Discussion

Elucidation of epigenetic mechanisms mediating aberrant gene expression during initiation and progression of lung cancers may hasten the development of novel therapies for these neoplasms. In the present study, we developed iPSCs from short term normal human small airway epithelial cells using well established techniques (18). The Lu-iPSC exhibited similarities to other pluripotent cell lines reported in the literature (13) including surface antigen and stem cell gene expression profiles, in-vitro growth, and teratoma formation. Although previously reported to be activated in stem cells (36, 37), cancer-germline genes commonly up-regulated in lung cancers remained transcriptionally repressed in Lu-iPSC. These unexpected findings are consistent with a report by Loriot *et al.* (38) demonstrating no up-regulation of 18 cancer-germline genes in human ESC, mesenchymal stem cells or adipose derived stem cells. Additionally, Lu-iPSC exhibited a hypermethylated genome relative to parental SAEC. Whereas these unanticipated findings could be indicative of incomplete reprogramming, DNA hypermethylation profiles in Lu-iPSC were also observed following analysis of databases pertaining to previously published ESC and iPSC. Using a micro-array platform identical to what was used in our studies, He *et al.* (39) observed significant increases in DNA methylation in two human ESCs, as well as two episomally derived iPSCs and four virally derived iPSCs relative to two human fibroblast lines used to generate the iPSC (~12% vs 1.1% of all CpG sites, respectively). Our analysis, which appears more extensive than that performed by He *et al.* (39) demonstrated that DNA hypermethylation in Lu-iPSC is not restricted to CpG islands in promoter regions but instead occurs throughout the genome. As yet, we have not directly compared methylation signatures in our Lu-iPSC with other iPSCs reported in the literature. Additional studies are required to more comprehensively characterize epigenomic alterations during reprogramming of human respiratory epithelial cells, and ascertain their relevance with respect to lung cancer biology.

In addition to complex alterations in DNA methylation, our experiments demonstrated that more than 15,000 genes were commonly regulated in Lu-iPSC relative to parental SAEC.

Notably, *ASXL3*, one of three mammalian homologues of drosophila *Additional Sex Combs (Asx)* (28, 29), was markedly up-regulated in Lu-iPSC. Analysis of previously published RNA-seq data summarized by Liu et al (40) revealed almost no expression of *ASXL3* in iPSC generated from different fibroblasts, suggesting that *ASXL3* activation may be context dependent, and possibly reflect transcriptional memory of the somatic cells undergoing reprogramming (41, 42). Subsequent experiments demonstrated marked up-regulation of *ASXL3* in SCLC; *ASXL3* did not appear to be expressed in NSCLC lines. Interestingly, SCLC is the highest expresser of *ASXL3* among a large panel of cancer lines of diverse histologies in the CCLE database. Knock-down of *ASXL3* decreased teratoma formation by Lu-iPSC, and inhibited growth of SCLC *in vitro* and *in vivo*. To the best of our knowledge, these experiments are the first description of reprogramming of human airway epithelial cells to pluripotency, and the first demonstration that *ASXL3* is a potential therapeutic target in SCLC.

The translational relevance of these findings may be inferred from observations that ASXL proteins physically interact with EZH2 and SUZ12 (core components of PRC2) to target this “initiation” complex to polycomb response elements throughout the genome (28, 29). ASXL proteins also interact with BRCA-1 Associated Protein-1 (BAP1), a ubiquitin hydrolase that mediates de-ubiquitination of H2AK119Ub, the repressive mark placed by the “maintenance” PRC1 (43). BAP1 does not interact with PRC2 proteins (29). As such, ASXL-BAP1 complexes may function to mitigate repressive activities of ASXL-PRC2. Particularly germane to SCLC, ASXL3 interacts with LSD1 (KDM1A), a histone demethylase which co-localizes with OCT4 and NANOG to regulate expression of pluripotency genes, and which together with PRC2 is required for maintenance of bivalent chromatin within promoters of lineage specific genes in undifferentiated stem cells (44).

Among a large panel of normal tissues in the CCLE database, *ASXL3* expression is highest in testis and to a lesser extent, ovary; *ASXL3* is expressed at much lower levels in brain, but does not appear to be expressed in normal lung or many other organs. This highly restricted expression suggests that *ASXL3* expression contributes to germ cell/stem cell homeostasis and neural development. With the exception of one report indicating that an evolutionarily conserved retrotransposon modulates *ASXL3* expression during neural development by encoding an alternatively spliced exon leading to nonsense mediated decay of the transcript (45), there are no published data pertaining to mechanisms regulating *ASXL3* expression in normal cells. Furthermore, whereas truncating mutations of *ASXL3* have been associated with autism and Bainbridge-Ropers syndrome (46), as well as a low percentage of melanomas (29, 47), there are no published data regarding mechanisms and clinical implications of *ASXL3* over-expression in cancers. *ASXL3* is located on 18q11, in a region not previously associated with genomic abnormalities by cytogenetic analysis in SCLC (48). Our preliminary experiments have demonstrated no induction of *ASXL3* in lung cancer cells following exposure to Decitabine, suggesting that *ASXL3* up-regulation in SCLC is not simply a manifestation of genomic demethylation. Instead, our aCGH experiments strongly suggest that *ASXL3* over-expression in SCLC is attributable, at least in some cases, to genomic amplification. This mechanism does not account for induction of *ASXL3* in Lu-iPSC. Our array, as well as pyrosequencing and MSP experiments did not demonstrate any changes in DNA methylation within the *ASXL3* promoter in Lu-iPSC relative to parental

SAEC; these findings as well as results of our ChIP-seq experiments suggest that *ASXL3* is repressed in normal respiratory epithelia by polycomb-mediated mechanisms (as evidenced by H3K27me3 levels within the ASXL3 promoter in SAEC relative to Lu-iPSC). These findings are consistent with our previously published observations pertaining to repression of *DKK1* in lung cancer cells by PRC independent of DNA hypermethylation (7). On the other hand, our data do not rule out the possibility of focal DNA demethylation coinciding with the observed histone alterations that facilitate up-regulation of *ASXL3* during reprogramming. Conceivably, up-regulation of *ASXL3* in Lu-iPSC is due to activation of one or more highly restricted stem cell transcription factors during reprogramming of parental SAEC; such mechanisms might also contribute to over-expression of *ASXL3* in SCLC. Experiments are underway using DNase hypersensitivity-seq and ChIP-seq techniques to comprehensively examine mechanisms of *ASXL3* up-regulation in Lu-iPSC and SCLC.

In light of the fact that over-expression of *EZH2* (the catalytic core component of PRC2), as well as up-regulation of *LSD-1* have been implicated in SCLC (49, 50), our findings pertaining to *ASXL3* over-expression in these neoplasms suggest that perturbations of *ASXL3*, *PRC2* and *LSD1* expression/activities are central themes of small cell lung carcinogenesis; as such, epigenomic landscapes and clinical phenotypes of these malignancies may be established by specific and potentially opposing mechanisms dysregulating expression, stoichiometry and targeting of these epigenetic modifiers. Experiments using inducible over-expression/knock-down techniques are in progress to examine these issues. Collectively, our findings support further analysis of the mechanisms and clinical implications of *ASXL3* over-expression in SCLC, and the evaluation of novel pharmacologic regimens targeting ASXL3 for SCLC therapy.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

**Financial Support:** This work was supported by NCI Intramural grants ZIA BC 011122 (to D.S. Schrump) and ZIA BC 011418 (to D.S. Schrump), and the Stephen J. Solarz Memorial Fund (to D.S. Schrump).

The authors wish to thank Drs. Beverly Teicher and Yves Pommier for assistance with access to cell lines and the CCLE database, Ms. Madeline Kim for assistance with editing the manuscript, and Ms. Jan Pappas as well as Ms. Joanna Lamot for administrative assistance.

## References

1. Rajan A, Schrump DS. Precision therapy for lung cancer: tyrosine kinase inhibitors and beyond. Semin Thorac Cardiovasc Surg. 2015; 27:36–48. [PubMed: 26074108]
2. Zheng M. Classification and pathology of lung cancer. Surg Oncol Clin N Am. 2016; 25:447–68. [PubMed: 27261908]
3. Kalemkerian GP, Schneider BJ. Advances in small cell lung cancer. Hematol Oncol Clin N Am. 2017; 31:143–56.
4. Schrump, DS., Carter, D., Kelsey, CR., Marks, LB., Giaccone, G. Cancer Principles and Practice of Oncology. DeVita, V.Lawrence, TS., Rosenberg, SA., editors. Philadelphia: Lippincott Williams Wilkins; 2011. p. 799-847.

5. Dong N, Shi L, Wang DC, Chen C, Wang X. Role of epigenetics in lung cancer heterogeneity and clinical implicatio. Semin Cell Devel Biol. 2017; 64:18–25. [PubMed: 27575638]
6. Hussain M, Rao M, Humphries AE, Hong JA, Liu F, Yang M, et al. Tobacco smoke induces polycomb-mediated repression of Dickkopf-1 in lung cancer cells. Cancer Res. 2009; 69:3570–8. [PubMed: 19351856]
7. Xi S, Yang M, Tao Y, Xu H, Shan J, Inchauste S, et al. Cigarette smoke induces C/EBP-beta-mediated activation of miR-31 in normal human respiratory epithelia and lung cancer cells. PLoS One. 2010; 5:e13764. [PubMed: 21048943]
8. Liu F, Killian JK, Yang M, Walker RL, Hong JA, Zhang M, et al. Epigenomic alterations and gene expression profiles in respiratory epithelia exposed to cigarette smoke condensate. Oncogene. 2010; 29:3650–64. [PubMed: 20440268]
9. Schrump DS. Targeting epigenetic mediators of gene expression in thoracic malignancies. Biochim Biophys Acta. 2012; 1819:836–45. [PubMed: 22507242]
10. Xi S, Xu H, Shan J, Tao Y, Hong JA, Inchauste S, et al. Cigarette smoke mediates epigenetic repression of miR-487b during pulmonary carcinogenesis. J Clin Invest. 2013; 123:1241–61. [PubMed: 23426183]
11. Zhang M, Mathur A, Zhang Y, Xi S, Atay S, Hong JA, et al. Mithramycin represses basal and cigarette smoke-induced expression of ABCG2 and inhibits stem cell signaling in lung and esophageal cancer cells. Cancer Res. 2012; 72:4178–92. [PubMed: 22751465]
12. Portal-Nunez S, Shankavaram UT, Rao M, Datrice N, Atay S, Aparicio M, et al. Aryl hydrocarbon receptor-induced adrenomedullin mediates cigarette smoke carcinogenicity in humans and mice. Cancer Res. 2012; 72:5790–800. [PubMed: 22993405]
13. Takahashi K, Yamanaka S. A decade of transcription factor-mediated reprogramming to pluripotency. Nature Rev Mol Cell Biol. 2016; 17:183–93. [PubMed: 26883003]
14. Xie N, Tang B. The Application of human iPSCs in neurological diseases: from bench to bedside. Stem Cells Int. 2016:6484713. [PubMed: 26880979]
15. Lim KL, Teoh HK, Choong PF, Teh HX, Cheong SK, Kamarul T. Reprogramming cancer cells: overview & current progress. Expert Opin Biol Ther. 2016; 16:941–51. [PubMed: 27070264]
16. Fan J, Robert C, Jang YY, Liu H, Sharkis S, Baylin SB, et al. Human induced pluripotent cells resemble embryonic stem cells demonstrating enhanced levels of DNA repair and efficacy of nonhomologous end-joining. Mutat Research. 2011; 713:8–17.
17. Papapetrou EP. Patient-derived induced pluripotent stem cells in cancer research and precision oncology. Nat Med. 2016; 22:1392–401. [PubMed: 27923030]
18. Beers J, Gulbranson DR, George N, Siniscalchi LI, Jones J, Thomson JA, et al. Passaging and colony expansion of human pluripotent stem cells by enzyme-free dissociation in chemically defined culture conditions. Nat Protoc. 2012; 7:2029–40. [PubMed: 23099485]
19. Schrock E, du Manoir S, Veldman T, Schoell B, Wienberg J, Ferguson-Smith MA, et al. Multicolor spectral karyotyping of human chromosomes. Science. 1996; 273:494–7. [PubMed: 8662537]
20. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29:15–21. [PubMed: 23104886]
21. Liao Y, Smyth GK, Shi W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014; 30:923–30. [PubMed: 24227677]
22. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43:e47. [PubMed: 25605792]
23. Roychoudhuri R, Hirahara K, Mousavi K, Clever D, Klebanoff CA, Bonelli M, et al. BACH2 represses effector programs to stabilize T(reg)-mediated immune homeostasis. Nature. 2013; 498:506–10. [PubMed: 23728300]
24. Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H, et al. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell. 2008; 2:151–9. [PubMed: 18371436]
25. Smith KP, Luong MX, Stein GS. Pluripotency: toward a gold standard for human ES and iPS cells. J Cellular Physiol. 2009; 220:21–9. [PubMed: 19326392]

26. Collinson A, Collier AJ, Morgan NP, Sienerth AR, Chandra T, Andrews S, et al. Deletion of the polycomb-group protein EZH2 leads to compromised self-renewal and differentiation defects in human embryonic stem cells. Cell Rep. 2016; 17:2700–14. [PubMed: 27926872]
27. Kadoch C, Copeland RA, Keilhack H. PRC2 and SWI/SNF chromatin remodeling complexes in health and disease. Biochemistry. 2016; 55:1600–14. [PubMed: 26836503]
28. Katoh M. Functional and cancer genomics of ASXL family members. Br J Cancer. 2013; 109:299–306. [PubMed: 23736028]
29. Katoh M. Functional proteomics of the epigenetic regulators ASXL1, ASXL2 and ASXL3: a convergence of proteomics and epigenetics for translational medicine. Expert Rev Proteomics. 2015; 12:317–28. [PubMed: 25835095]
30. Grade M, Hormann P, Becker S, Hummon AB, Wangsa D, Varma S, et al. Gene expression profiling reveals a massive, aneuploidy-dependent transcriptional deregulation and distinct differences between lymph node-negative and lymph node-positive colon carcinomas. Cancer Res. 2007; 67:41–56. [PubMed: 17210682]
31. Habermann JK, Paulsen U, Roblick UJ, Upender MB, McShane LM, Korn EL, et al. Stage-specific alterations of the genome, transcriptome, and proteome during colorectal carcinogenesis. Genes Chromosomes Cancer. 2007; 46:10–26. [PubMed: 17044061]
32. Camps J, Grade M, Nguyen QT, Hormann P, Becker S, Hummon AB, et al. Chromosomal breakpoints in primary colon cancer cluster at sites of structural variants in the genome. Cancer Res. 2008; 68:1284–95. [PubMed: 18316590]
33. Ried T, Hu Y, Difilippantonio MJ, Ghadimi BM, Grade M, Camps J. The consequences of chromosomal aneuploidy on the transcriptome of cancer cells. Biochim Biophys Acta. 2012; 1819:784–93. [PubMed: 22426433]
34. Hwang DH, Sun H, Rodig SJ, Hornick JL, Sholl LM. Myc protein expression correlates with MYC amplification in small-cell lung carcinoma. Histopathology. 2015; 67:81–9. [PubMed: 25407018]
35. 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. [PubMed: 22941189]
36. Cronwright G, Le BK, Gotherstrom C, Darcy P, Ehnman M, Brodin B. Cancer/testis antigen expression in human mesenchymal stem cells: down-regulation of SSX impairs cell migration and matrix metalloproteinase 2 expression. Cancer Res. 2005; 65:2207–15. [PubMed: 15781633]
37. Lifantseva N, Koltsova A, Krylova T, Yakovleva T, Poljanskaya G, Gordeeva O. Expression patterns of cancer-testis antigens in human embryonic stem cells and their cell derivatives indicate lineage tracks. Stem Cells Int. 2011:795239. [PubMed: 21785609]
38. Loriot A, Reister S, Parvizi GK, Lysy PA, De Smet C. DNA methylation-associated repression of cancer-germline genes in human embryonic and adult stem cells. Stem Cells. 2009; 27:822–4. [PubMed: 19350682]
39. He WY, Kang XJ, Du HZ, Song B, Lu ZY, Huang YL, et al. Defining differentially methylated regions specific for the acquisition of pluripotency and maintenance in human pluripotent stem cells via microarray. Plos One. 2014; 9:e108350. [PubMed: 25250679]
40. Liu Y, Cheng D, Li Z, Gao X, Wang H. The gene expression profiles of induced pluripotent stem cells (iPSCs) generated by a non-integrating method are more similar to embryonic stem cells than those of iPSCs generated by an integrating method. Genet Mol Biol. 2012; 35:693–700. [PubMed: 23055811]
41. Kim K, Zhao R, Doi A, Ng K, Unternaehrer J, Cahan P, et al. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. Nature Biotechnol. 2011; 29:1117–9. [PubMed: 22119740]
42. Ohi Y, Qin H, Hong C, Blouin L, Polo JM, Guo T, et al. Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. Nat Cell Biol. 2011; 13:541–9. [PubMed: 21499256]
43. Daou S, Hammond-Martel I, Mashtalir N, Barbour H, Gagnon J, Iannantuono NV, et al. The BAP1/ASXL2 histone H2A deubiquitinase complex regulates cell proliferation and is disrupted in cancer. J Biol Chem. 2015; 290:28643–63. [PubMed: 26416890]

44. Adamo A, Sese B, Boue S, Castano J, Paramonov I, Barrero MJ, et al. LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. Nat Cell Biol. 2011; 13:652–9. [PubMed: 21602794]

45. Lowe CB, Bejerano G, Salama SR, Haussler D. Endangered species hold clues to human evolution. Journal Hered. 2010; 101:437–47.

46. Hori I, Miya F, Ohashi K, Negishi Y, Hattori A, Ando N, et al. Novel splicing mutation in the ASXL3 gene causing Bainbridge-Ropers syndrome. Am J Med Genet A. 2016; 170:1863–7. [PubMed: 27075689]

47. Micol JB, Abdel-Wahab O. The role of additional sex combs-like proteins in cancer. Cold Spring Harb Perspect Med. 2016; 6 pii:a026526.

48. Ashman JN, Brigham J, Cowen ME, Bahia H, Greenman J, Lind M, et al. Chromosomal alterations in small cell lung cancer revealed by multicolour fluorescence in situ hybridization. Int J Cancer. 2002; 102:230–6. [PubMed: 12397641]

49. Poirier JT, Gardner EE, Connis N, Moreira AL, de Stanchina E, Hann CL, et al. DNA methylation in small cell lung cancer defines distinct disease subtypes and correlates with high expression of EZH2. Oncogene. 2015; 34:5869–78. [PubMed: 25746006]

50. Mohammad HP, Smitheman KN, Kamat CD, Soong D, Federowicz KE, Van Aller GS, et al. A DNA hypomethylation signature predicts antitumor activity of LSD1 inhibitors in SCLC. Cancer Cell. 2015; 28:57–69. [PubMed: 26175415]

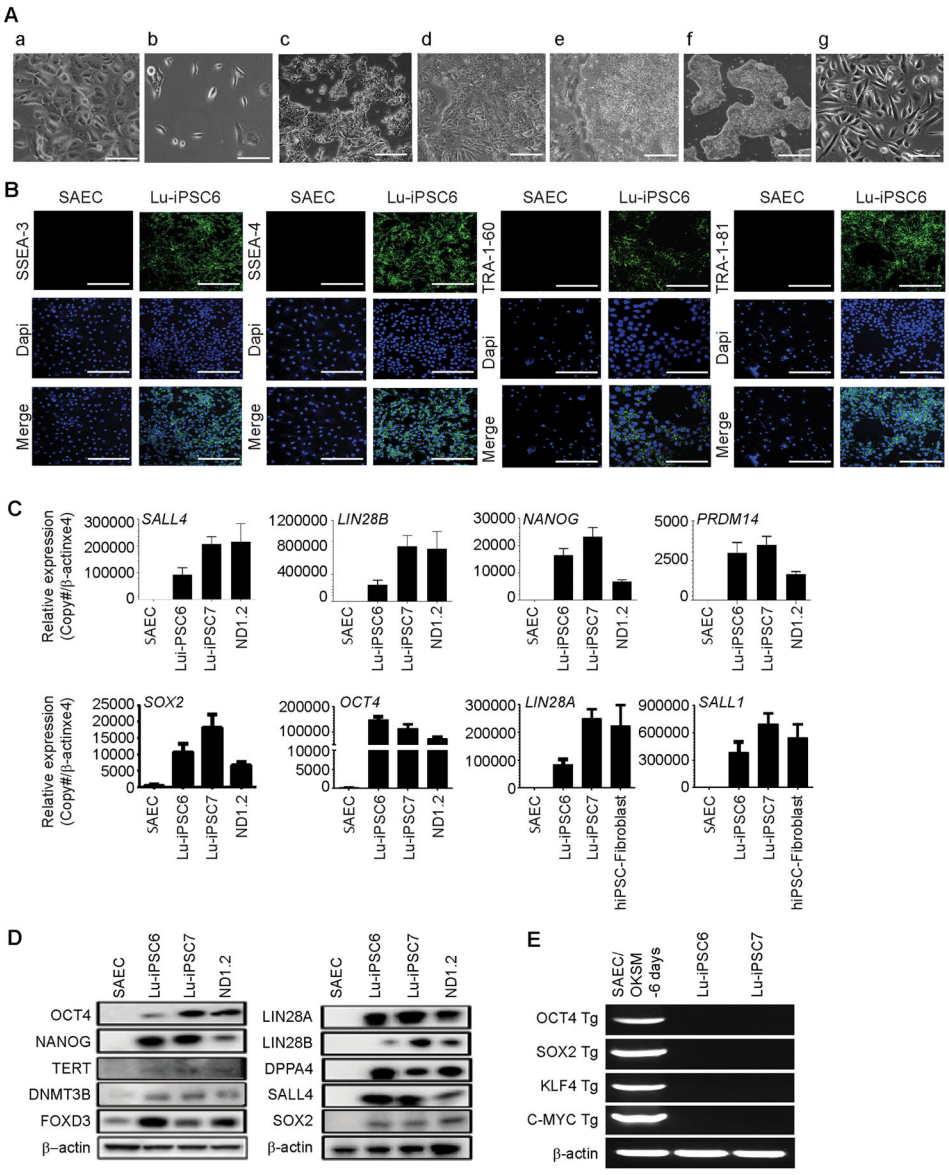


**Figure 1.**

Generation of Lu-iPSC from SAEC. (A-a) SAEC before reprogramming. (A-b) SAEC 6 days after transduction with STEMCCA. (A-c) Emergence of Lu-iPSC colonies after 10–15 days of reprogramming. (A-d) Lu-iPSC in ES cell medium after 20 days of reprogramming. (A-e) Lu-iPSC colonies before dissociation after 25 days of reprogramming. (A-f) Established Lu-iPSC. (A-g) Non-transduced SAEC cultured for 30 days in reprogramming medium. (B) Immunofluorescence analysis of SSEA-3, SSEA-4, TRA-1-60, and TRA-1-81 expression in Lu-iPSC6 relative to SAEC. Scale bars, 100μm. (C) qRT-PCR analysis of stemness-associated genes in Lu-iPSC6, Lu-iPSC7 and SAEC. ND1.2 was used as a positive control. (D) Immunoblot analysis of pluripotency-related proteins in Lu-iPSCs, SAEC and ND1.2. β-actin was used as a control. (E) RT-PCR analysis of *OCT4, SOX2, KLF4* and *C-MYC* transgenes in SAEC 6 days after transduction with STEMCCA, and Lu-iPSC.

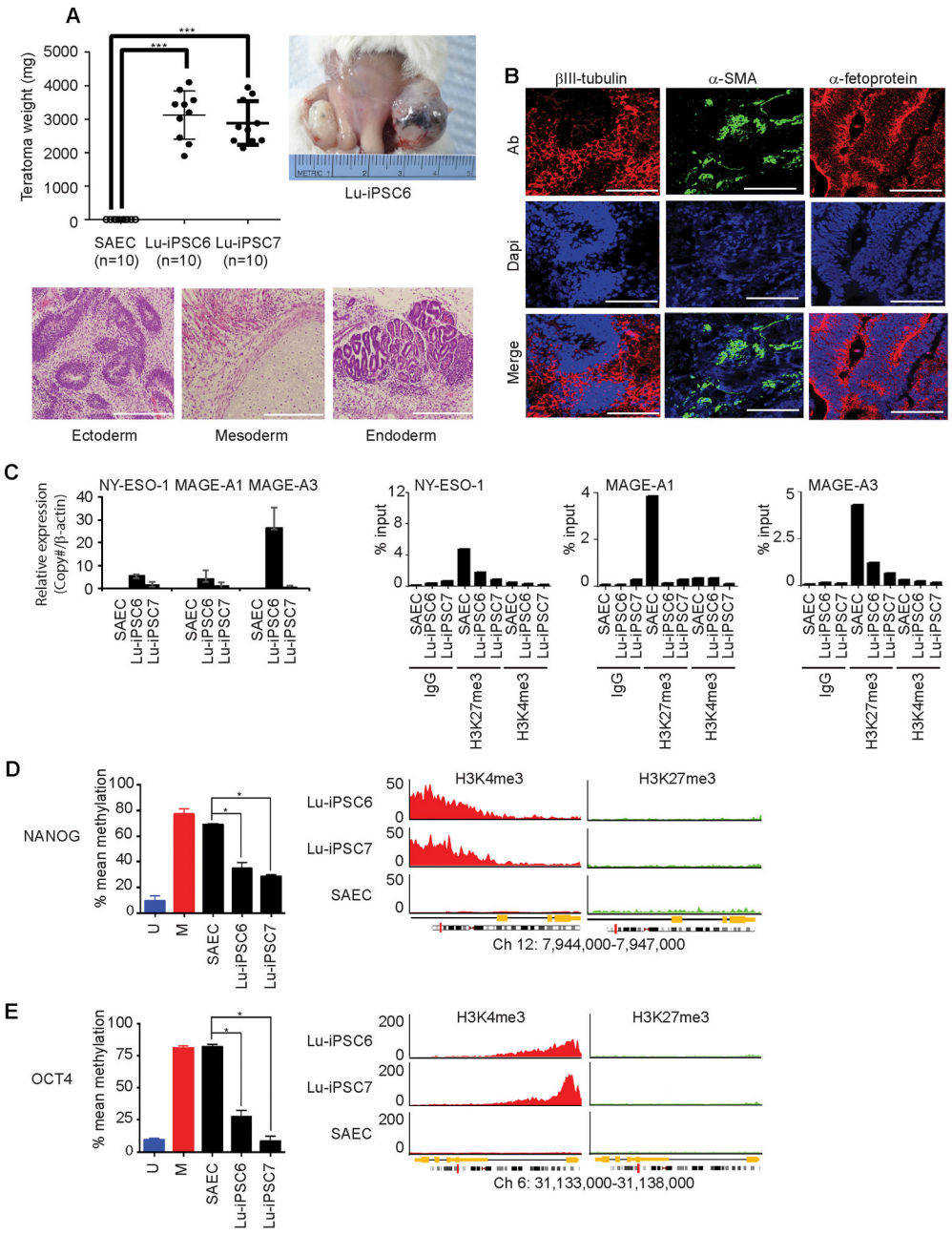


**Figure 2.**

Characterization of Lu-iPSC. (A; left upper panel) Summary of xenograft experiments. Tumors were observed following inoculation with Lu-iPSC but not parental SAEC. (A; right upper panel) Representative photograph of teratoma formation by Lu-iPSC6. A, lower panel) Hematoxylin and eosin staining of teratomas derived from Lu-iPSC6 demonstrating ectoderm (neural tissues), mesoderm (cartilage and skeletal tissues) and endoderm (kidney and intestinal tissues). (B) Immunofluorescence staining for different germ layer markers in Lu-iPSC6-derived teratomas (βIII tubulin for ectoderm, α-smooth muscle actin (α-SMA) for mesoderm, and α-fetoprotein for endoderm). Scale bars, 100μm. (C) qRT-PCR analysis

of *NY-ESO-1*, *MAGE-A1*, and *MAGE-A3* expression in Lu-iPSC and SAEC (left panel). qChIP analysis of H3K27me3 and H3K4me3 levels within the *NY-ESO-1*, *MAGE-A1* and *MAGE-A3* promoters (right panel). (D/E) Pyrosequencing (left panel) and ChIP-seq analysis (right panel) demonstrating that *NANOG* (D) and *OCT4* (E) promoters are hypomethylated with increased H3K4me3 levels in Lu-iPSC. U: unmethylated, M: methylated.

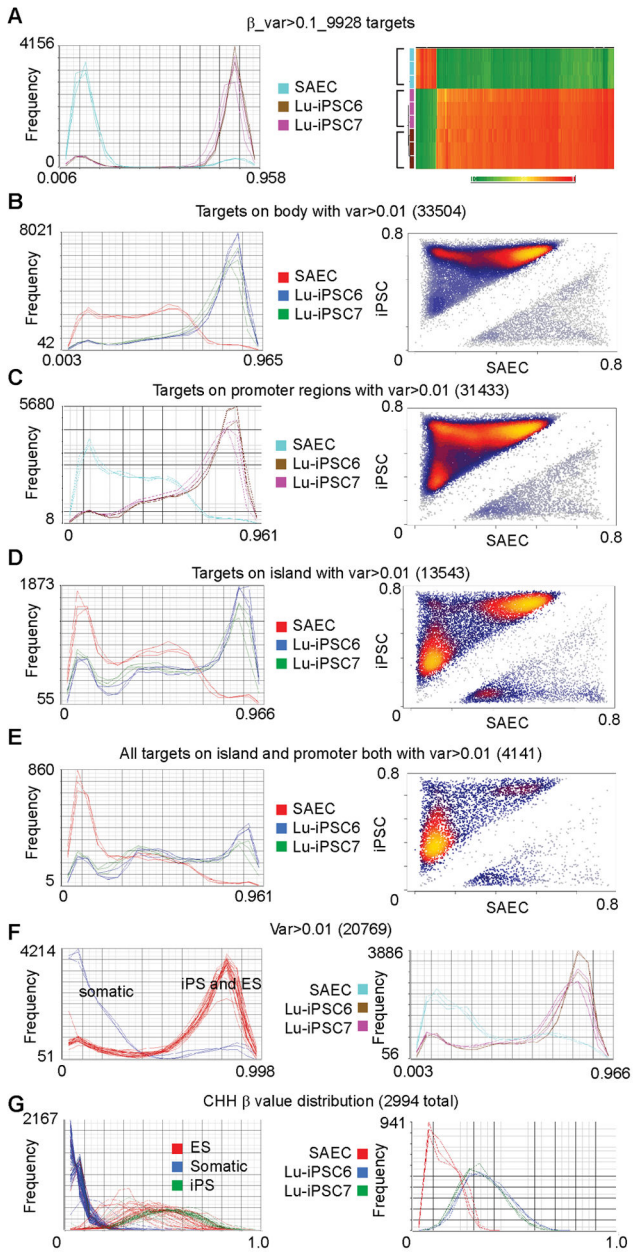


**Figure 3.**

DNA methylation array analysis of Lu-iPSC and SAEC. (A) Unsupervised analysis of beta value distribution for targets with variance greater than 0.1 as specified. DNA methylation profiling for targets on body (B), on promoters (C), in islands (D) and collectively in islands and promoters (E) demonstrating hypermethylation in Lu-iPSCs relative to parental SAEC. DNA methylation changes in Lu-iPSC are similar to those observed following analysis of a public database. (F) Unsupervised analysis of beta value distribution for CpG targets with variance greater than 0.01 from downloaded dataset of GSE31848 (left panel) and the corresponding beta value distribution of the same targets in our dataset (right panel). (G)

Beta value distributions of all 2994 CHH targets designed on the H450K methylation array from downloaded data set (GSE31848, left panel) and our dataset (right panel).

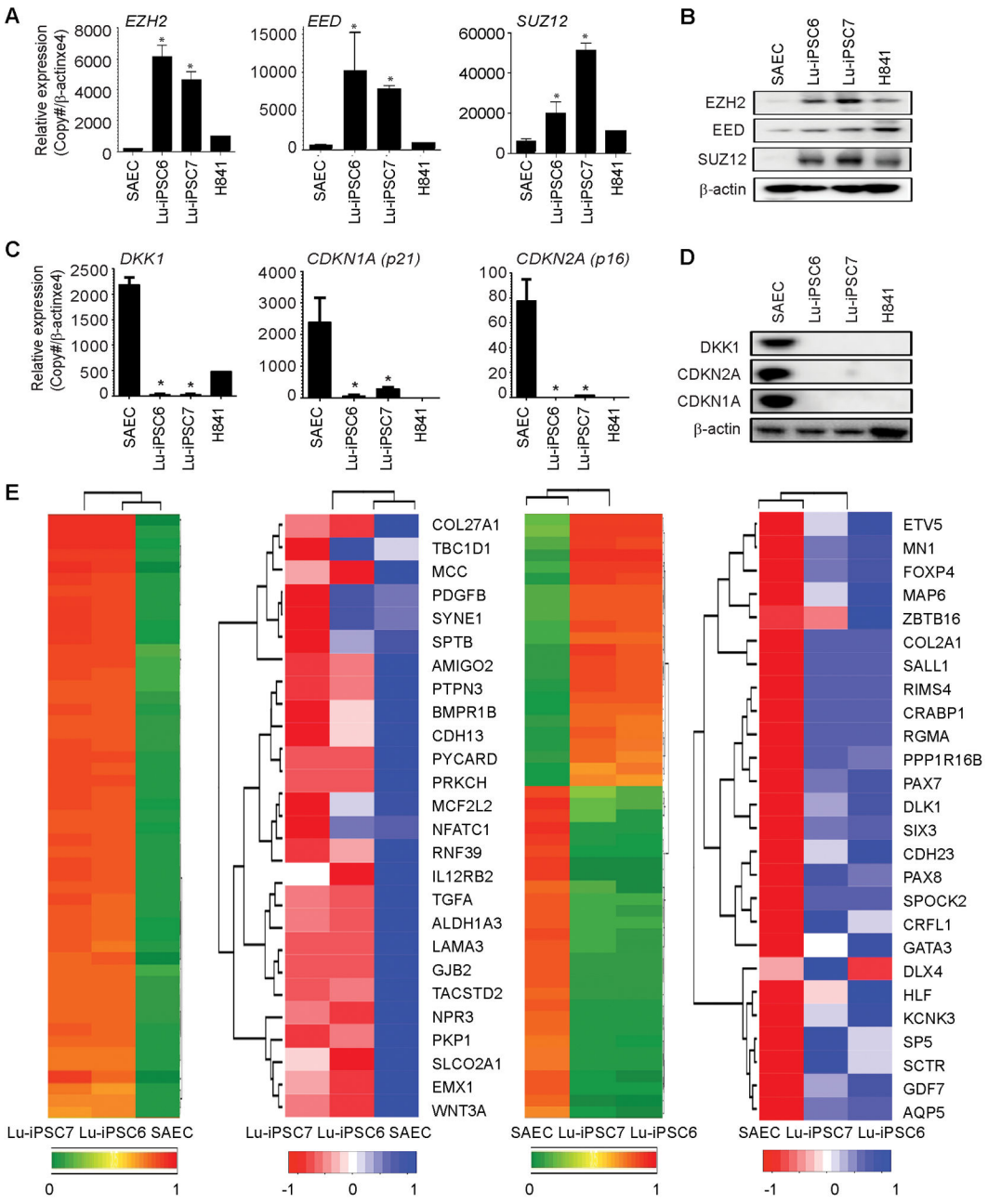


**Figure 4.**

PRC2 genes are upregulated and tumor suppressors are downregulated in Lu-iPSC relative to SAEC. *P<0.05. (A) qRT-PCR analysis demonstrating upregulation of *EZH2*, *EED* and *SUZ12* in Lu-iPSC relative to SAEC. (B) Immunoblot analysis of EZH2, EED and SUZ12 in Lu-iPSC and SAEC. (C) qRT-PCR analysis of *DKK1*, *CDKN1A (p21)*, and *CDKN2A (p16)* in Lu-iPSC and SAEC. (D) Immunoblot analysis of DKK1, CDKN1A and CDKN2A expression in Lu-iPSC and SAEC. (E) Representative comparison of DNA methylation array and RNA-seq analysis of PRC2 targets in Lu-iPSC compared to parental SAEC.

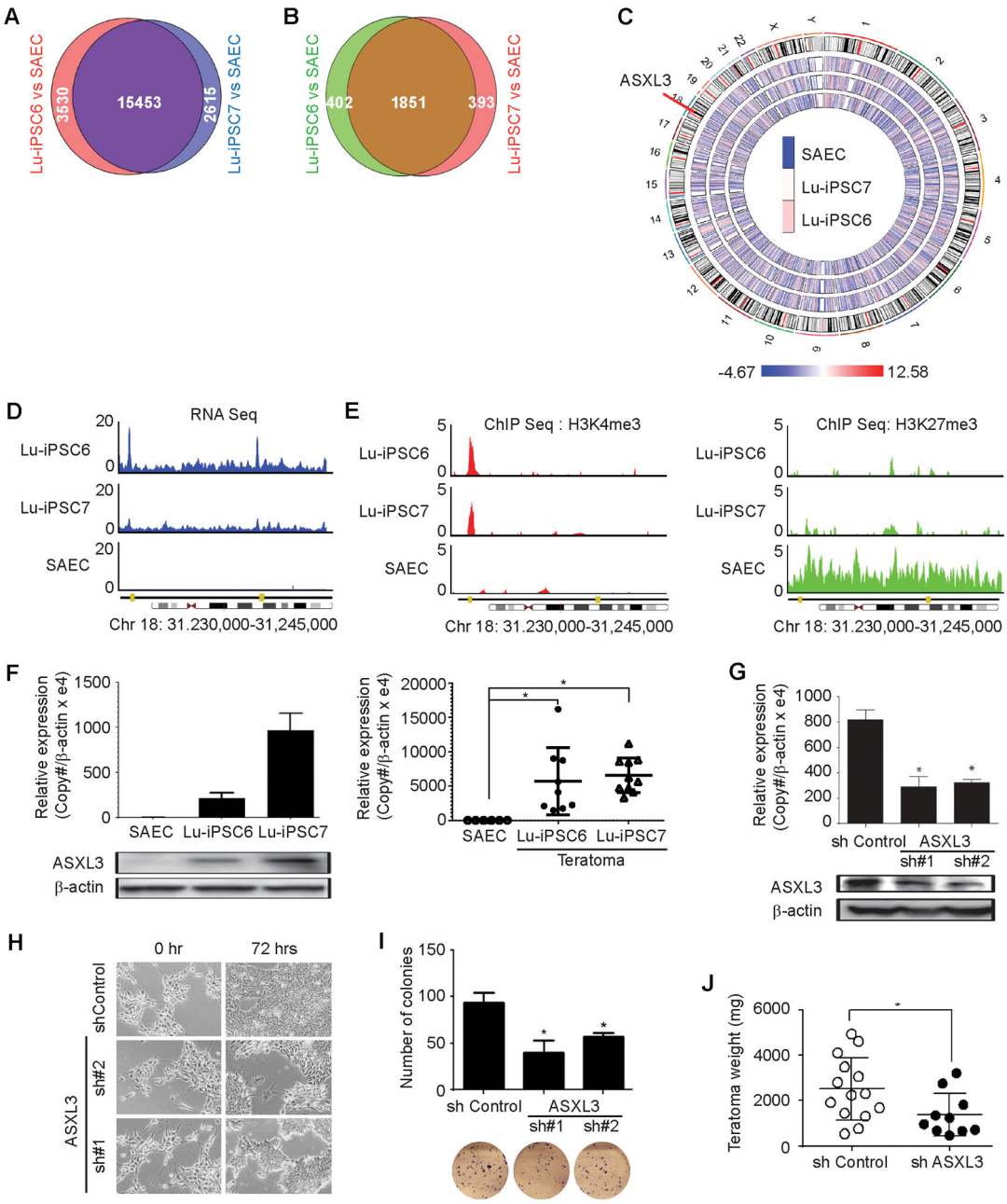


**Figure 5.**

*Additional Sex Combs Like 3 (ASXL3)* is upregulated in Lu-iPSC. (A) Venn diagram showing overlap of genes between Lu-iPSC6 and 7 relative to SAEC (RNA-seq after normalizing read counts (log2/million)). (B) Upon more in-depth analysis, 1851 genes were altered in Lu-iPSC compared to SAEC. (C) Circos plot showing over-expression of *ASXL3* in Lu-iPSC, which is represented by red line. (D) Normalized RNA read counts for *ASXL3* in Lu-iPSC and SAEC. (E) Histone modifications peak read counts of for H3K4me3 (left) and H3K27me3 (right) within the *ASXL3* promoter. (F) qRT-PCR and immunoblot showing over-expression of *ASXL3* in Lu-iPSC (left) and their teratomas (right). *P<0.05. (G)

Validation of *ASXL3*-shRNA lentiviral particles showing reduction of *ASXL3* expression at transcriptional and protein levels. (H) Knockdown of *ASXL3* reduces growth and alters morphology of Lu-iPSC7. (I) Effects of *ASXL3* knockdown on soft agar colony formation by Lu-iPSC7. (J) Results of tumorigenicity assays demonstrating that knockdown of *ASXL3* decreases number and size of teratomas induced by Lu-iPSC7 in SCID mice. *P<0.05.

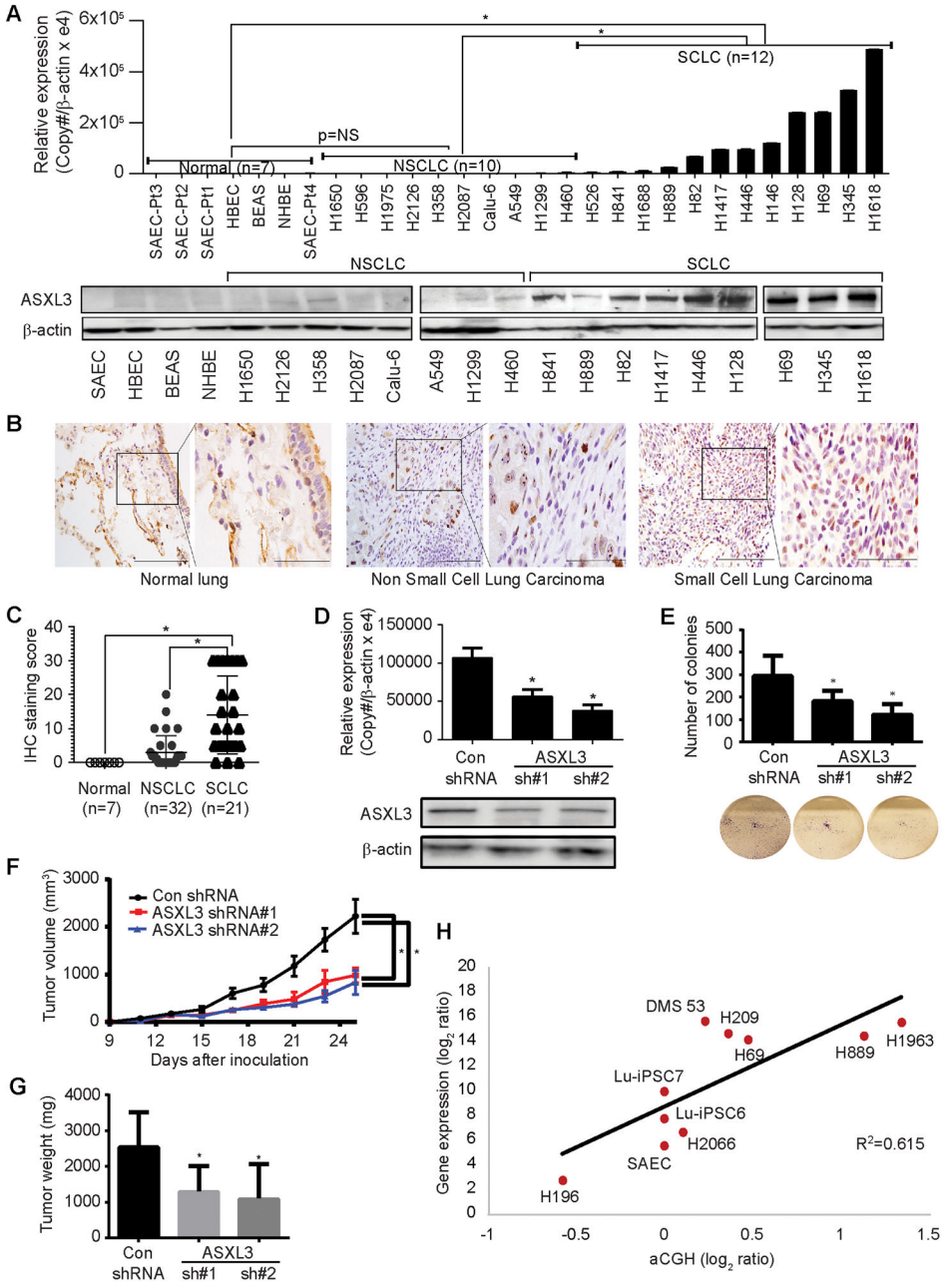


**Figure 6.**

*ASXL3* is an oncogene in SCLC. *P<0.05. (A) qRT-PCR and immunoblot analyses of *ASXL3* expression in lung cancer lines and primary or immortalized respiratory epithelial cells. (B/C) IHC analysis demonstrating higher nuclear expression of ASXL3 in SCLC relative to NSCLC or normal lung. (D) Effects of lentiviral knockdown of *ASXL3* on *ASXL3* mRNA and protein levels in H69 cells. (E) Effects of *ASXL3* knockdown on soft agar clonogenicity of H69 cells. (F/G) Effects of *ASXL3* knockdown on growth (F) and weight (G) of H69 xenografts in athymic nude mice. Scale bars, 100μm. (H) Correlation of genomic copy numbers of *ASXL3* by CGH array with *ASXL3* mRNA expression levels by

qRT-PCR. A significant correlation was observed for genomic copy number and *ASXL3* expression levels in SCLC lines, suggesting a direct effect of gene copy on relative message levels (R = 0.78420, P = 0.007). Upregulation of *ASXL3* in Lu-iPSC was not associated with an increase in genomic copy number.