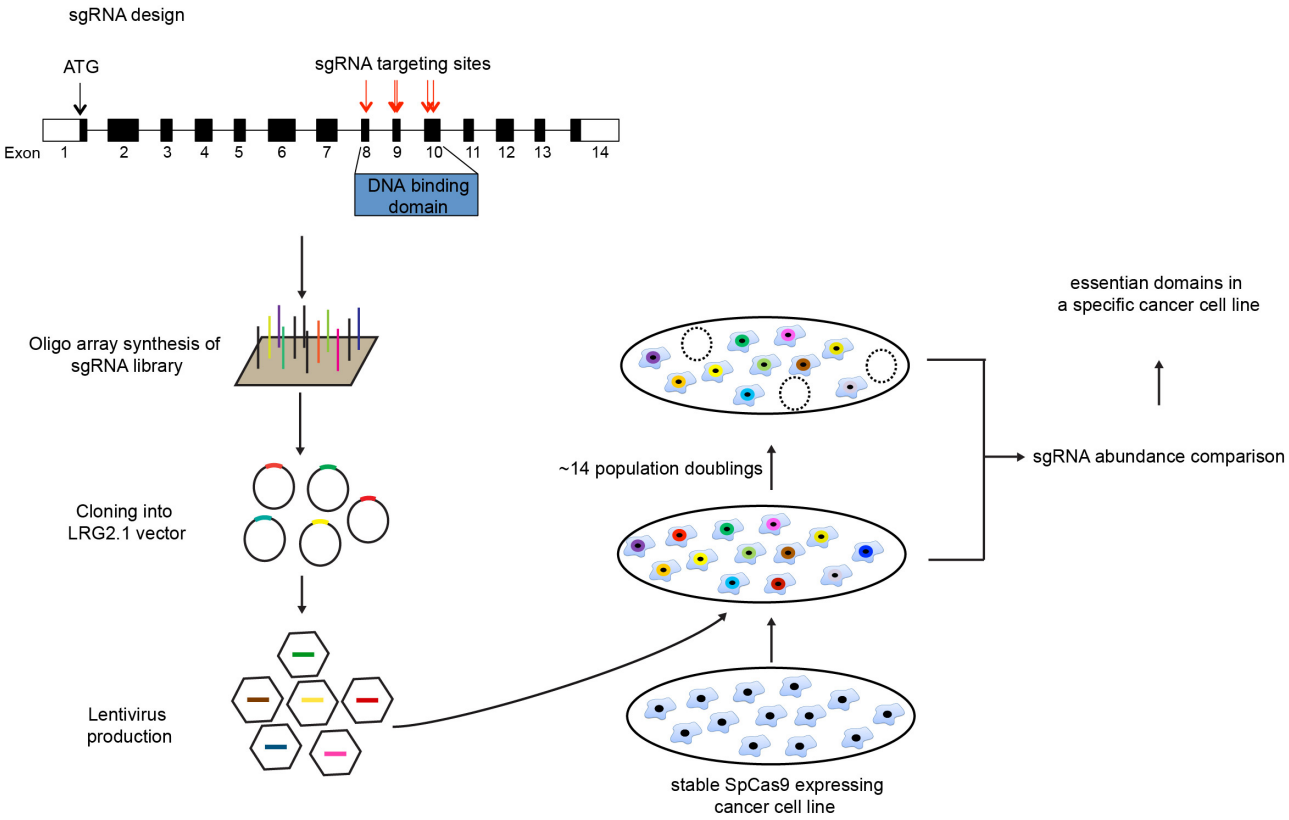
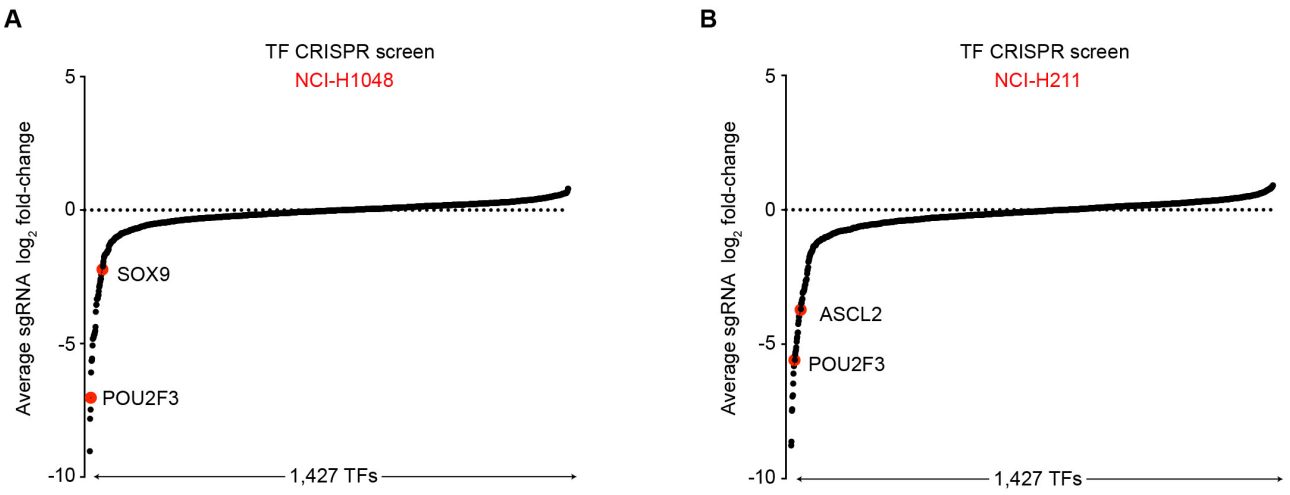
# Supplemental Information

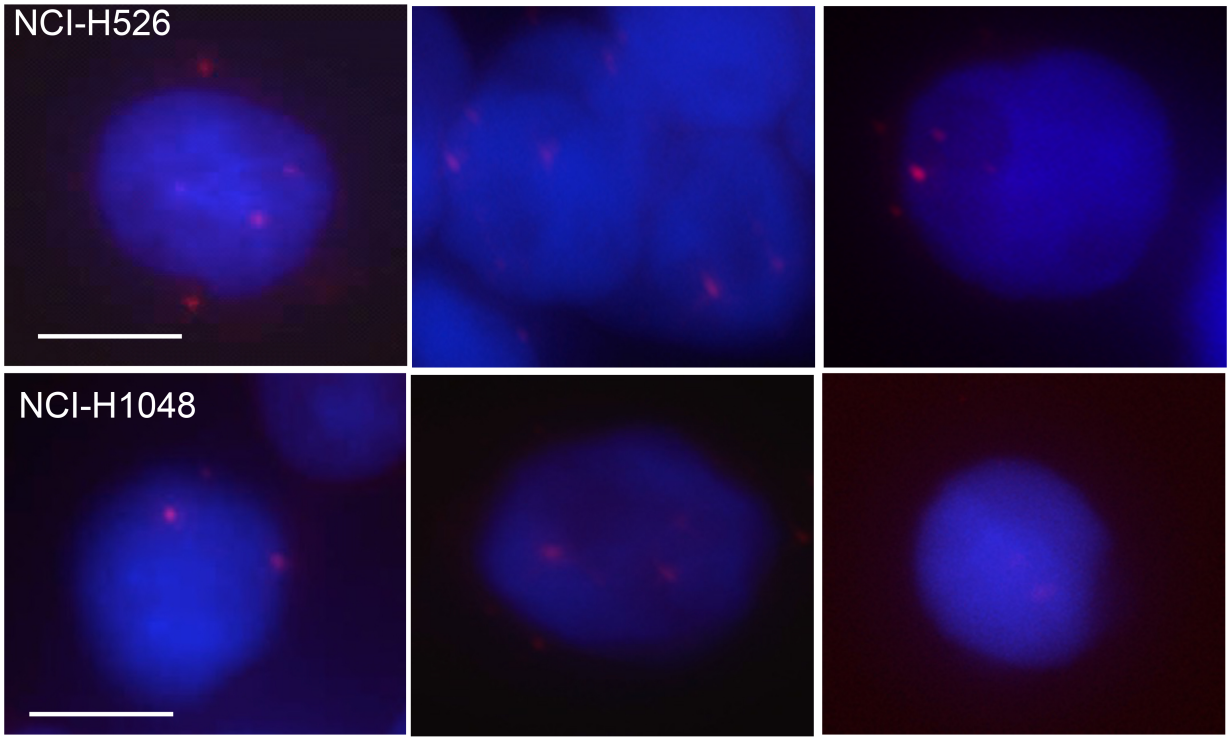



**Supplemental Figure S1. Transcription factor DNA-binding domain-focused CRISPR screening strategy.** A pooled library of 8,658 sgRNAs were designed to target the known (or putative) DNA-binding domain of 1,427 human transcription factors.




**Supplemental Figure S2. POU2F3 is a top TF dependency in select SCLC cell lines.** Ranked average sgRNA $\log_2$ fold-change of 1,427 TFs in NCI-H1048 (A) and NCI-H211 (B).

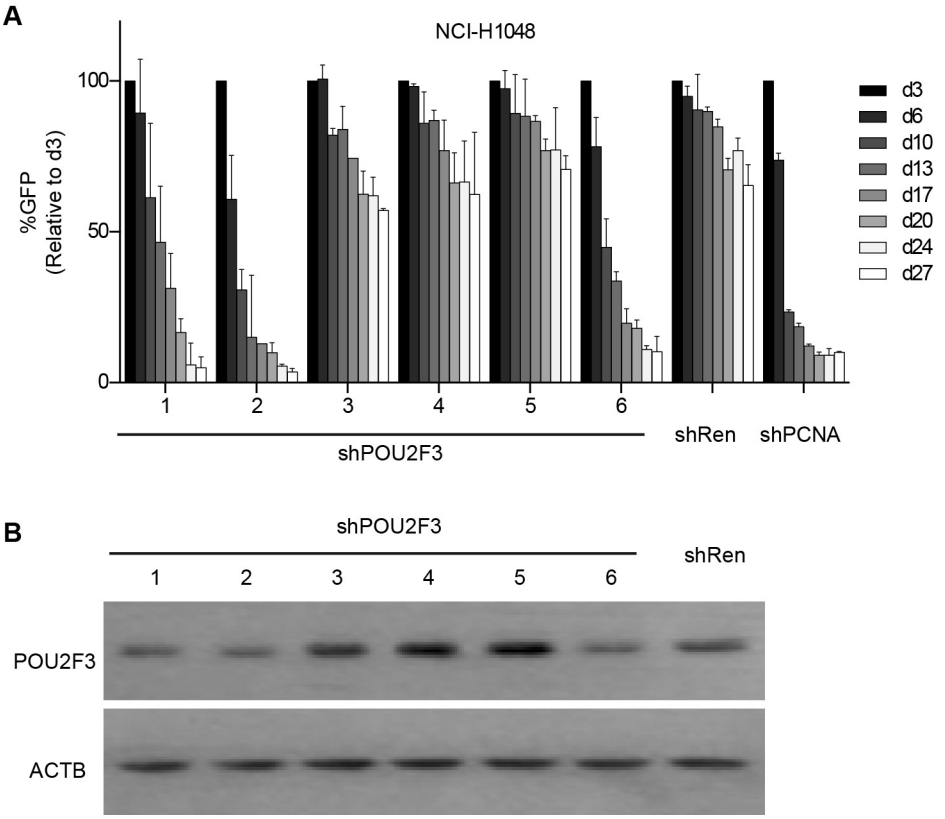
A



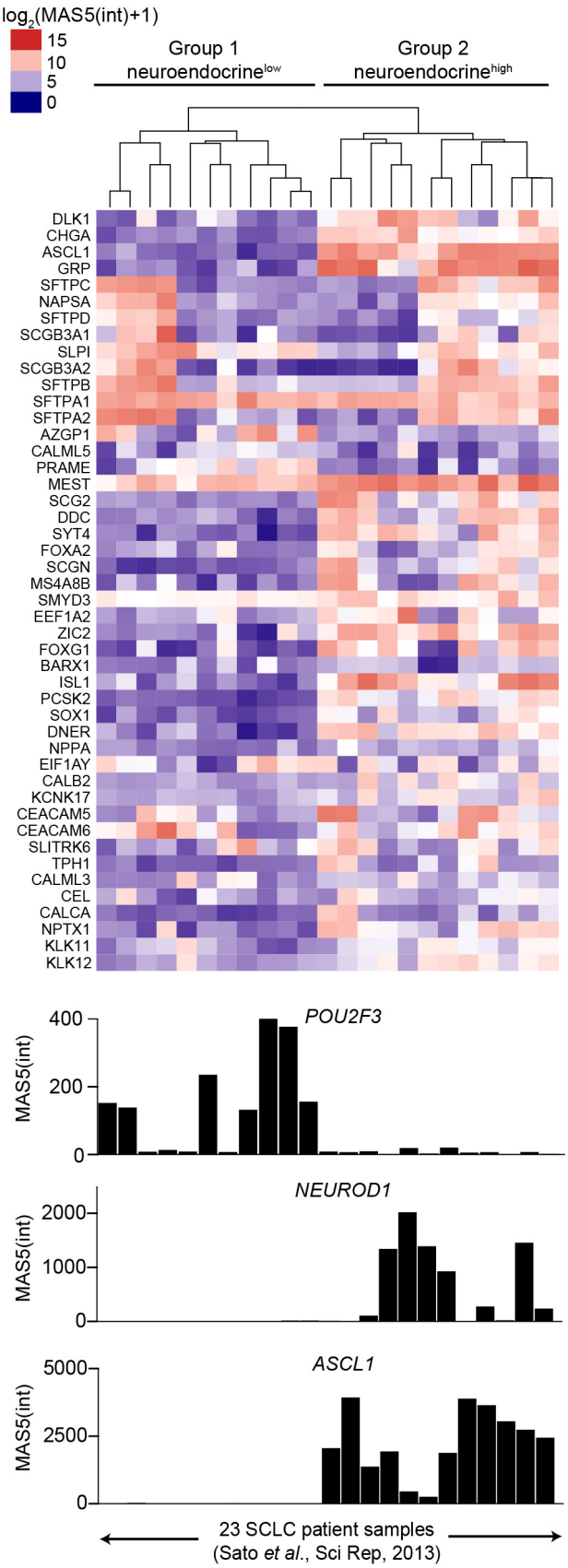

B

| Cell Line | Number of nuclei | Number of fields | 0 foci | 1 foci | 2 foci | 3 and >3 |
|---|---|---|---|---|---|---|
| NCI-H526 | 156 | 6 | 13 | 21 | 84 | 38 |
| NCI-H1048 | 91 | 5 | 7 | 11 | 65 | 8 |

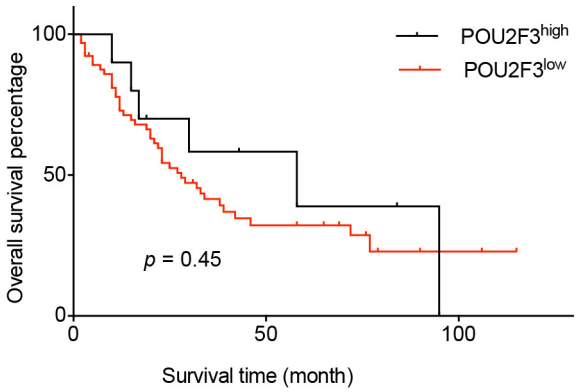
**Supplemental Figure S3. RNA-FISH of *POU2F3* gene transcription sites**.
(A) Representative *POU2F3* RNA-FISH images of NCI-H526 (top panels) and NCI-H1048 (bottom panels) cell lines. Nucleus is stained with DAPI. Red signal within the nucleus represents the transcription sites. Scale bar = 5μM.
(B) Quantitation of RNA-FISH signal from each of the cell lines. Total number of nuclear *POU2F3* foci per cell are counted and tabulated.




**Supplemental Figure S4. Validation of POU2F3 dependency using shRNA-based knockdown.** (A) Arrayed format competition-based proliferation assay in NCI-H1048 infected with indicated shRNAs. shRluc is a negative control shRNA targeting Renilla luciferase, which is not expressed in these cells. shPCNA is a positive control shRNA targeting the essential DNA replication protein PCNA. Plotted is the% GFP-positive relative to day 3 post-infection. Bar graphs represent the mean ± SEM (n = 3). (B) Western blotting of NCI-H1048 infected with control shRNA or shRNAs targeting POU2F3, blotting for POU2F3 or ACTB as a loading control. Note that the degree of POU2F3 knockdown correlates with the degree of growth arrest.



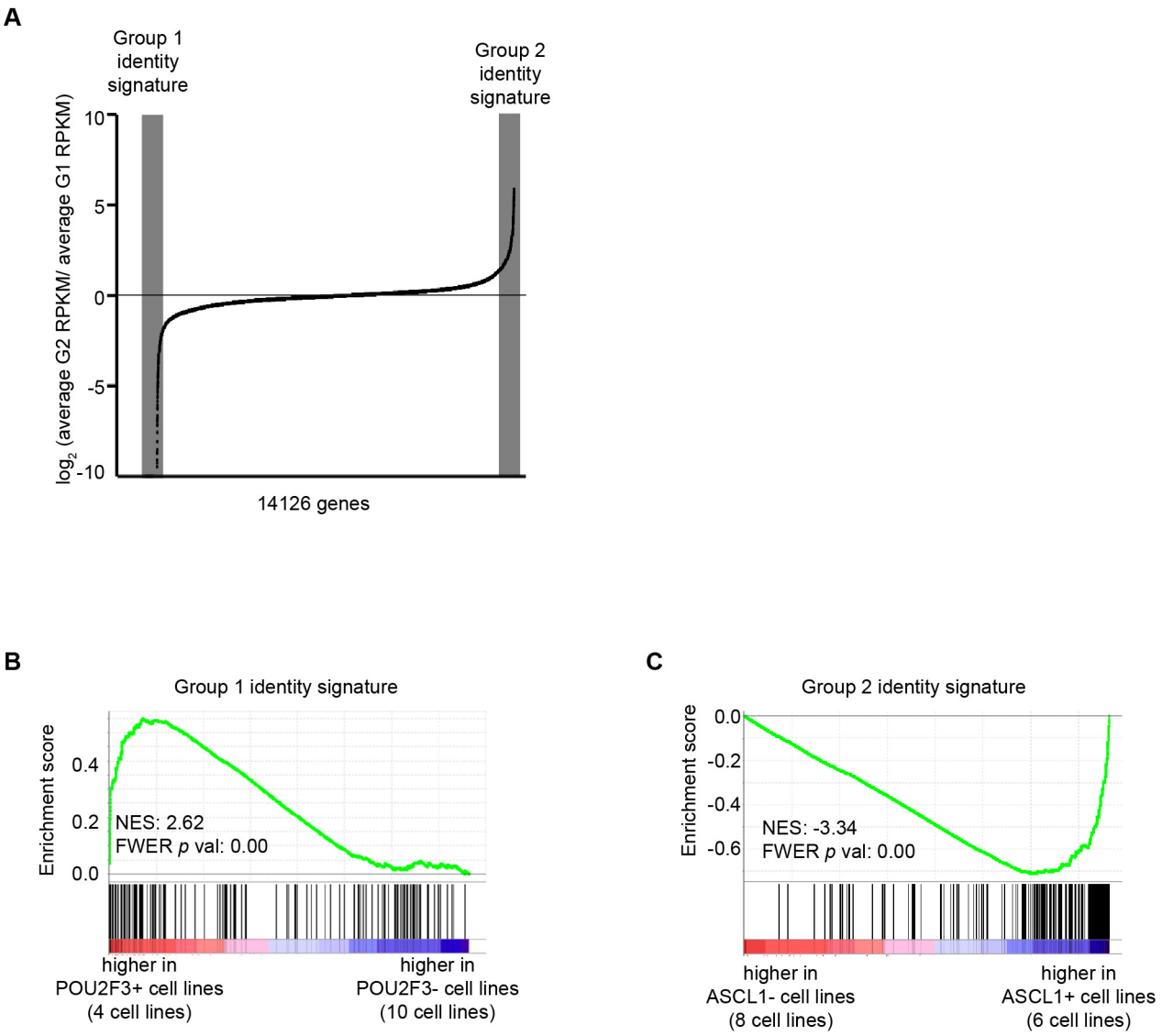

**Supplemental Figure S5. POU2F3 is expressed in Group 1 (neuroendocrine$^{low}$) SCLC.** We applied the method of unsupervised clustering analysis of RNA-seq data described in George *et al* 2015 (1) to the transcriptome dataset form 23 SCLC tumors Sato et al, 2013 (2). Tumor samples are arranged in columns, and genes are arranged in rows which selected based on differential expression in the two groups of samples. (bottom) Expression of POU2F3 across the 23 SCLC patient samples.



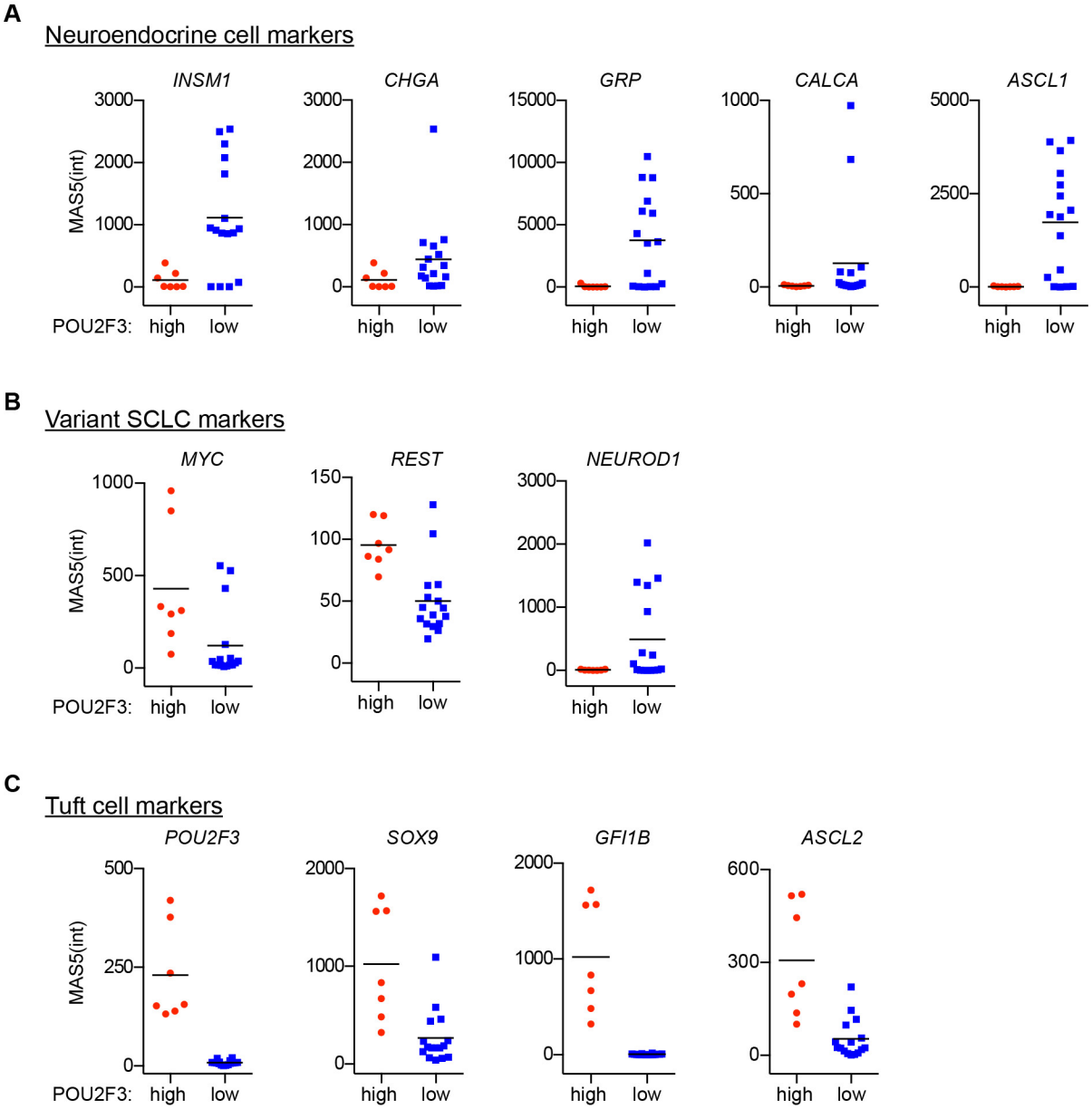

**Supplemental Figure S6. Overall survival analysis comparing POU2F3$^{high}$ and POU2F3$^{low}$ SCLC.** Kaplan-Meier survival curves comparing POU2F3 expressing (POU2F3$^{high}$) and non-expressing (POU2F3$^{low}$) SCLC patients using the dataset from George *et al.*, 2015 (1). A log-rank (Mantel-Cox) test failed to detect a significant survival difference between these two patient groups (p = 0.4507).

| | POU2F3$^{high}$ (9 cases) | POU2F3$^{low}$ (60 cases) |
|---|---|---|
| TP53 | 88.9% (8) | 93.3% (56) |
| RB1 | 55.6% (5) | 81.7% (49) |
| KIAA1211 | 11.1% (1) | 18.3% (11) |
| COL22A1 | 0.0% (0) | 18.3% (11) |
| RGS7 | 33.3% (3) | 8.3% (5) |
| FPR1 | 11.1% (1) | 3.3% (2) |
| EP300 | 33.3% (3) | 10.0% (6) |
| CREBBP | 11.1% (1) | 8.3% (5) |
| ASPM | 11.1% (1) | 10.0% (6) |
| ALMS1 | 0.0% (0) | 11.7% (7) |
| PDE4DIP | 11.1% (1) | 11.7% (7) |
| XRN1 | 0.0% (0) | 11.7% (7) |
| PTGFRN | 11.1% (1) | 3.3% (2) |
| TP73 | 11.1% (1) | 5.0% (3) |
| FMN2 | 22.2% (2) | 23.3% (14) |
| NOTCH1 | 0.0% (0) | 21.7% (13) |
| NOTCH2 | 11.1% (1) | 3.3% (2) |
| NOTCH3 | 11.1% (1) | 6.7% (4) |
| NOTCH4 | 0.0% (0) | 6.7% (4) |
| PIK3CA | 0.0% (0) | 1.7% (1) |

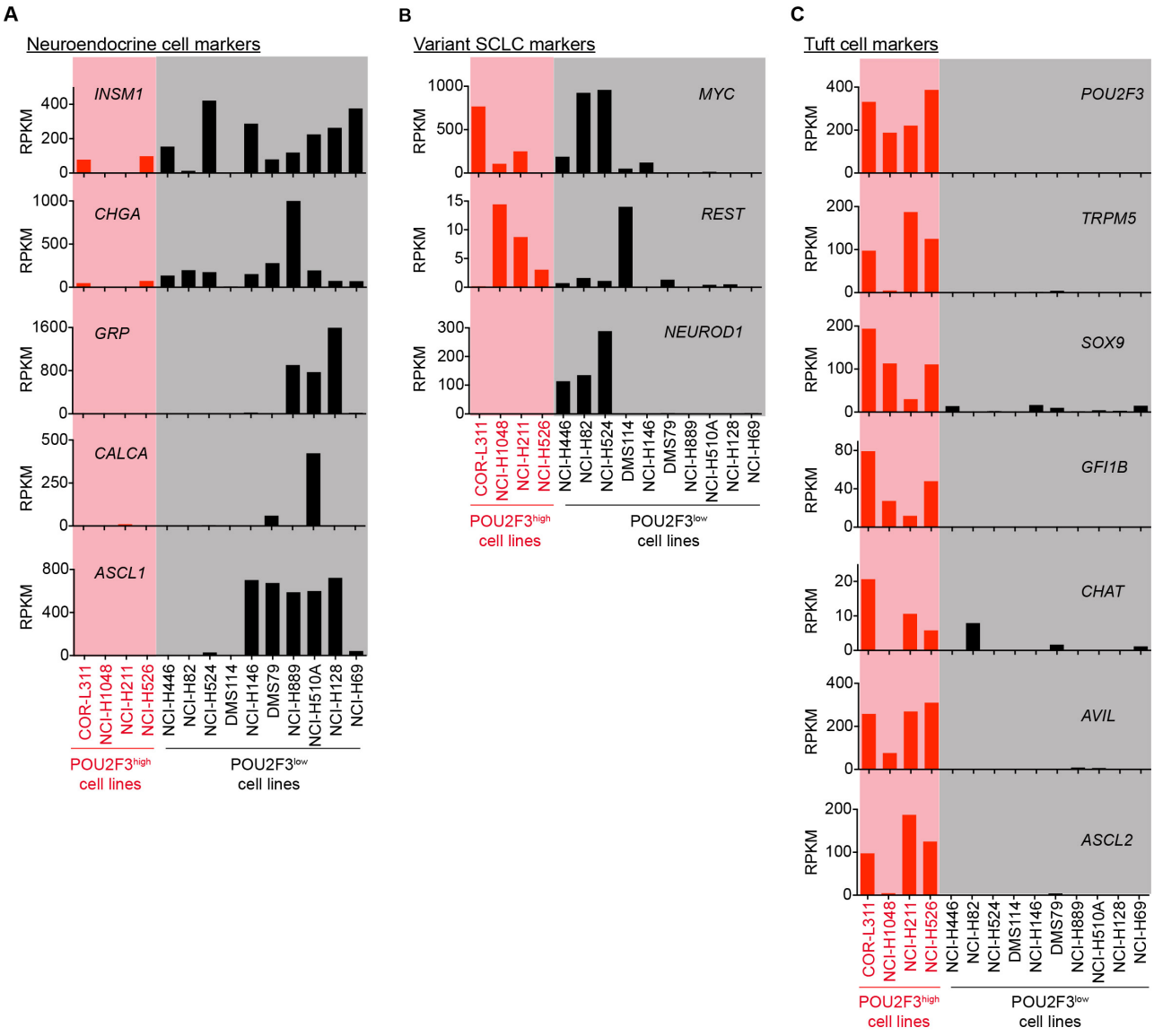
**Supplemental Figure S7**. Comparison of gene mutation frequency in POU2F3$^{high}$ and POU2F3$^{low}$ patient samples using the dataset from George *et al*, 2015 (1). Only samples with matched RNA-seq and whole geneome sequencing data are considered for the analysis here.



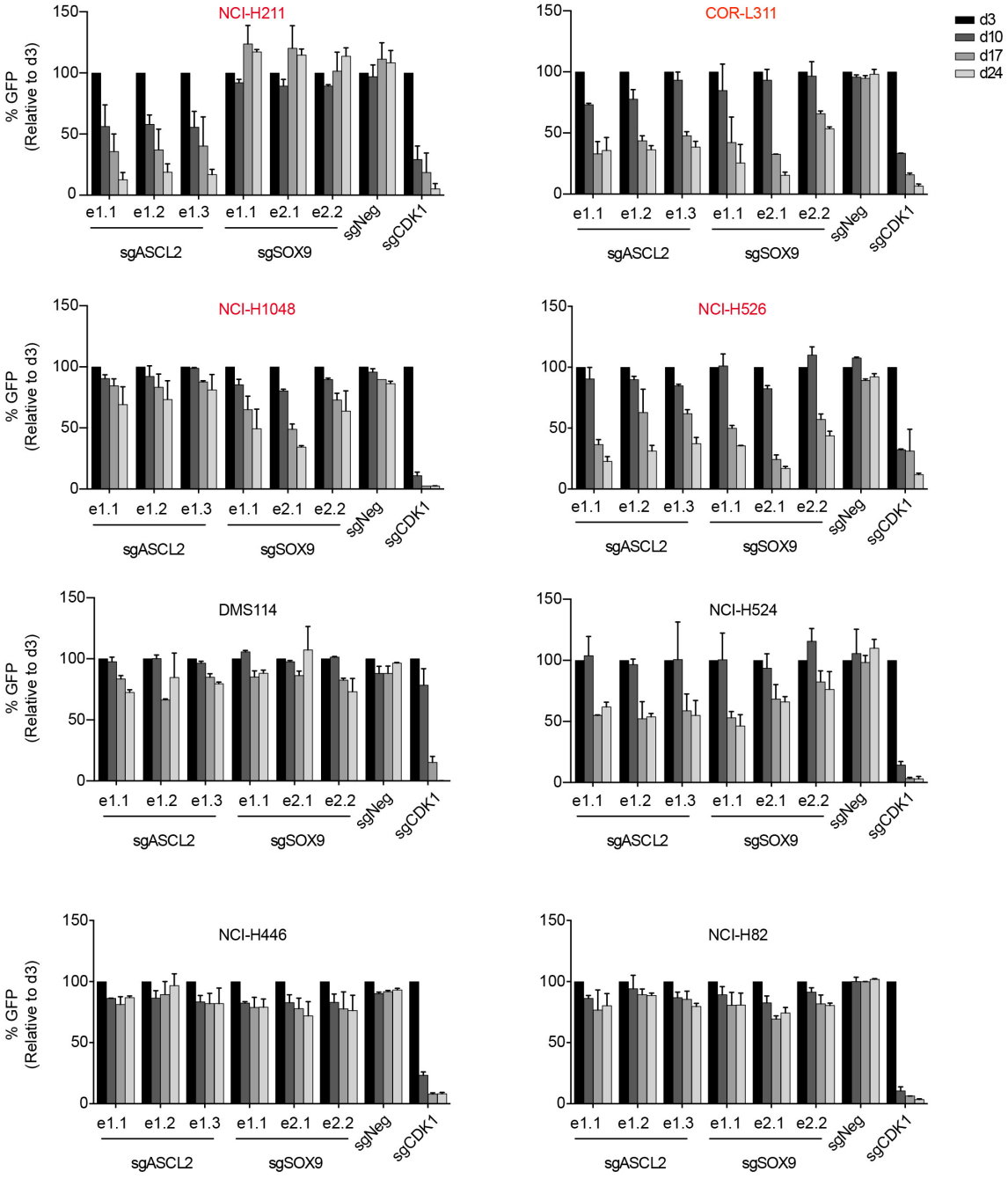

**Supplemental Figure S8. Definition of Group 1 signature in SCLC patient samples.** (A) Method for defining the 'Group 1 identity signature' and 'Group 2 identity signature'. All expressed genes (14,126 in total) in the SCLC patient sample cohort obtained from George *et al*, 2015 (1) were ranked based on the average RPKM in Group 2 samples divided by the average RPKM of Group 1 samples. Genes with greater than 5-fold bias to Group 1 samples were selected as the 'Group 1 identity signature', and genes with greater than 5-fold bias to Group 2 samples were selected as "Group 2 identity signature. Group 1 identity signature and Group 2 identity signature gene sets are provided in Supplemental Table 4. (B) Gene set enrichment analsysis (GSEA) comparing four POU2F3$^{high}$ cell lines to ten POU2F3$^{low}$ cell lines using the Group 1 identity signature identified in (a). (C) GSEA comparing six ASCL1$^{high}$ cell lines to eight ASCL1$^{low}$ cell lines using the Group 2 identity signature identified in (A). NES: normalized enrichment score. FWER *p* val: family-wise error rate *p* value.



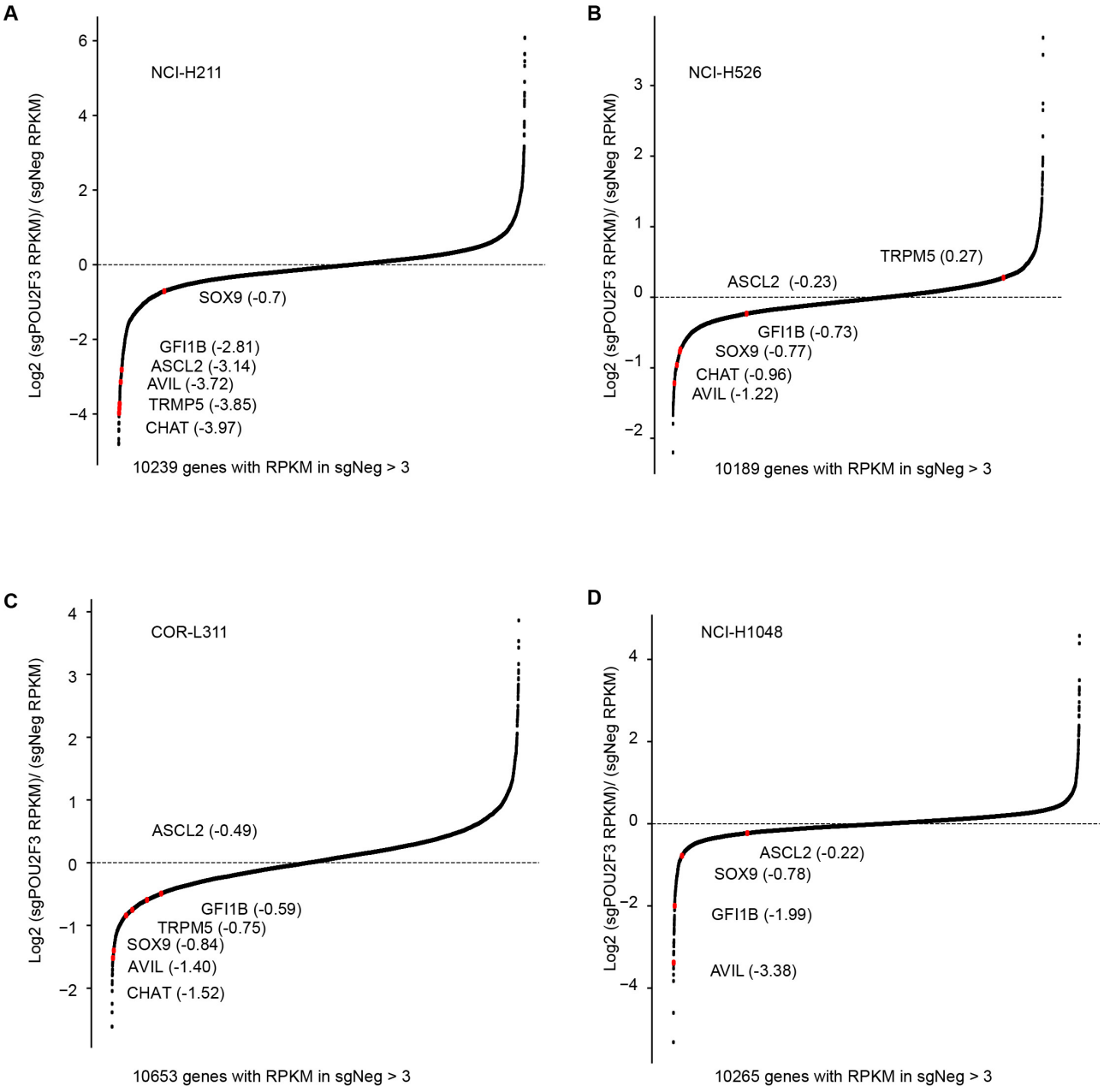

**Supplemental Figure S9. Expression of lineage markers in POU2F3$^{high}$ and POU2F3$^{low}$ SCLC cell patient samples from Sato et al. cohort (2).** (A) Expression of neuroendocrine cell markers in POU2F3$^{high}$ and POU2F3$^{low}$ SCLC patient samples. Two-tailed t test: *INSM1*: $p = 0.008$, *CHGA*: $p = 0.1782$, *GRP*: $p = 0.0165$, *CALCA*: $p = 0.2742$, *ASCL1*: $p = 0.006$. (B) Expression of variant SCLC markers in POU2F3$^{high}$ and POU2F3$^{low}$ SCLC patient samples. Two-tailed t test: *MYC*: $p < 0.0001$; *REST*: $p = 0.0011$ , *NEUROD1*: $p = 0.0825$. (C) Expression of tuft cell markers in POU2F3$^{high}$ and POU2F3$^{low}$ SCLC patient samples. Two-tailed t test: *POU2F3*: $p < 0.0001$, *SOX9*: $p = 0.0003$, *GFI1B*: $p < 0.0001$, *ASCL2*: $p < 0.0001$. Horizontal line represents the mean. Only genes with MAS5(int) value > 200 in at least one sample were analyzed here.



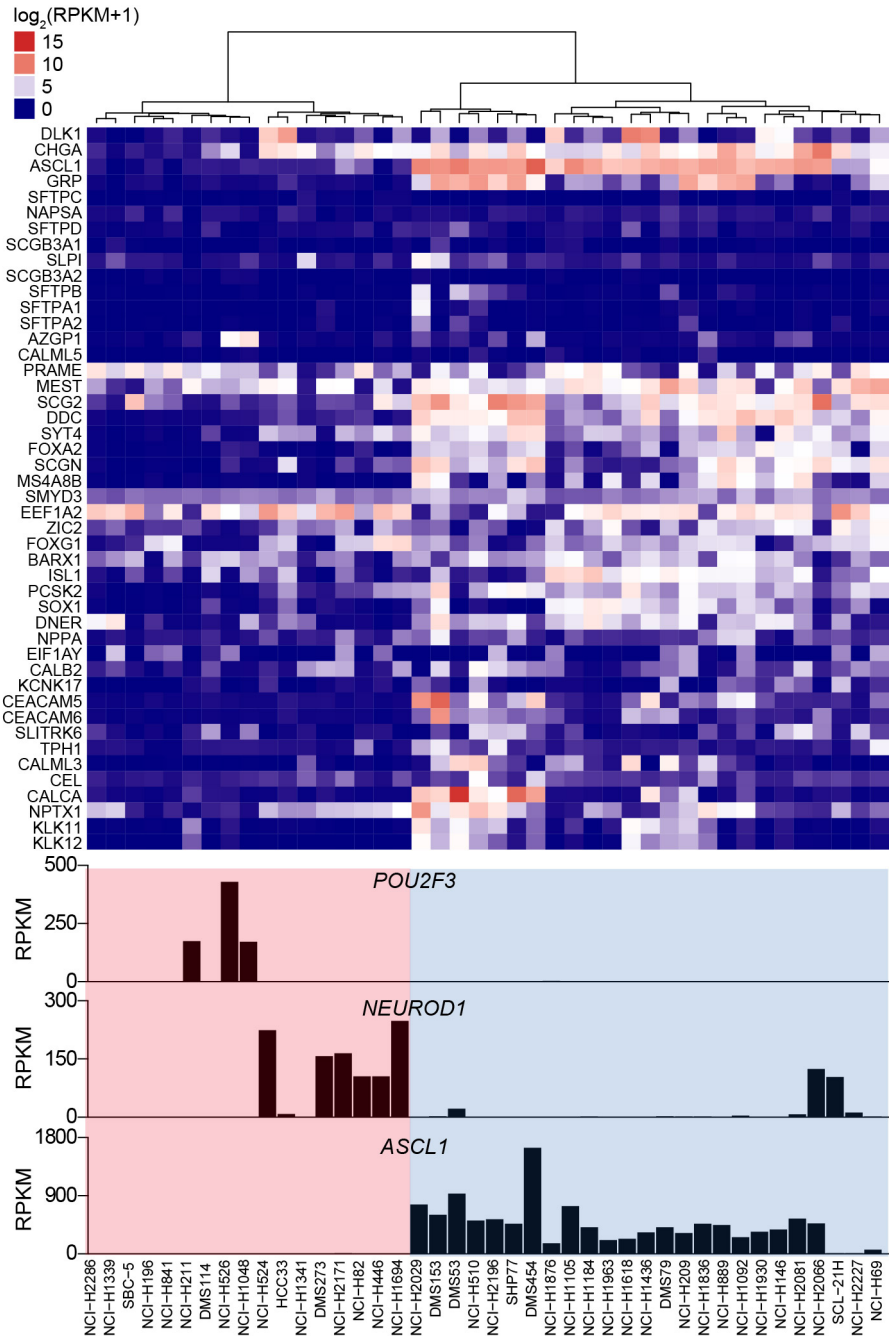

**Supplemental Figure S10. RNA-seq analysis of lineage markers in 14 human SCLC lines.** Expression of indicated neuroendocrine cell markers (A), variant SCLC markers (B), and tuft cell markers (C) in POU2F3$^{high}$ and POU2F3$^{low}$ SCLC cell lines.



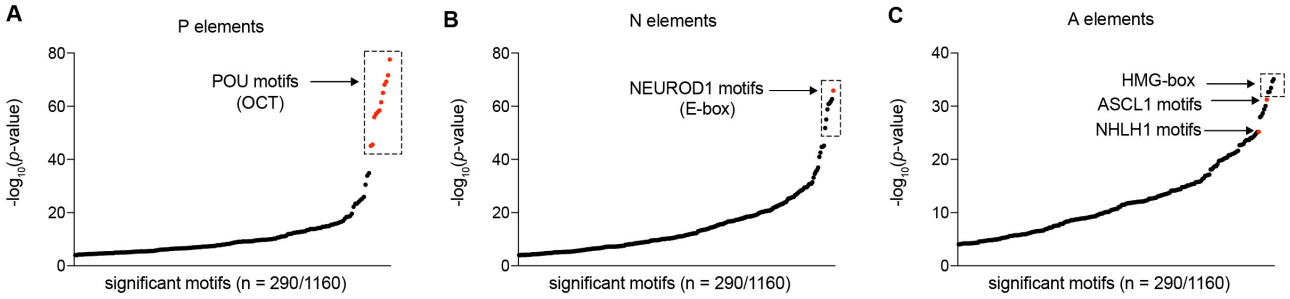

**Supplemental Figure S11. ASCL2 and SOX9 TFs are unique dependencies in POU2F3$^{high}$ SCLC lines.** Arrayed format competition-based assays evaluating effects of the indicated sgRNAs on cell proliferation. sgRNA expression in the LRG 2.1T vector is linked to GFP, and hence GFP loss over time in culture reflects the sgRNA-induced fitness disadvantage relative to non-transduced cells in the same well. As shown in Fig. S10, ASCL2 is expressed at low levels in NCI-H1048 and SOX9 expressed at low levels in NCI-H211.



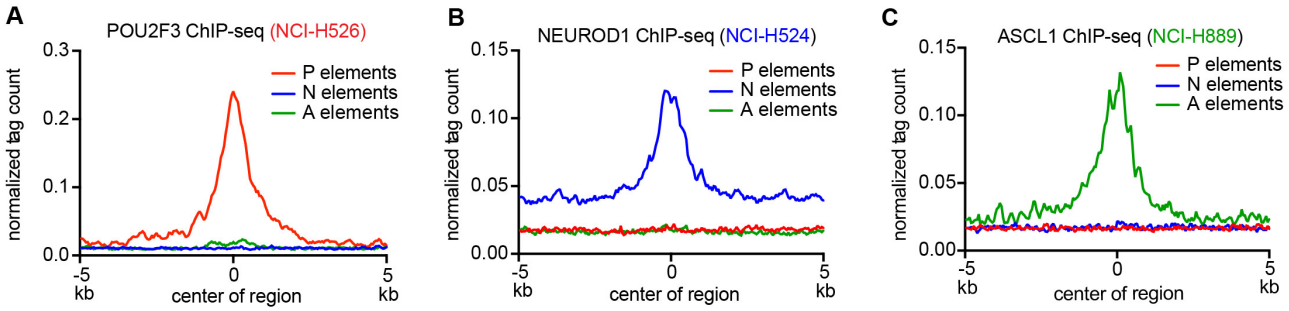

**Supplemental Figure S12. Decreased expression of tuft cell marker genes upon the knock out of POU2F3 in POU2F3$^{high}$ SCLC cell lines. POU2F3 and control** sgRNAs were infected into NCI-H211 (A), NCI-H526 (B), COR-L311 (C), or NCI-H1048 (D) cells and cell were collected for RNA-seq analysis. Genes with the RPKM value > 3 in sgNeg in each cell line are shown, and genes are ranked by the $\log_2$ sgPOU2F3 RPKM / sgNeg RPKM. Tuft cell marker genes are labled in red in each cell line, with the value of $\log_2$ sgPOU2F3 RPKM / sgNeg RPKM.



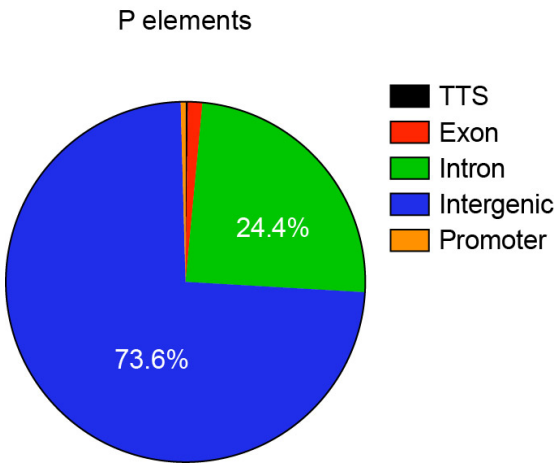

**Supplemental Figure S13. Mutually exclusive expression of POU2F3, ASCL1, and NEUROD1 in SCLC lines in the cancer cell line encyclopedia (CCLE).**



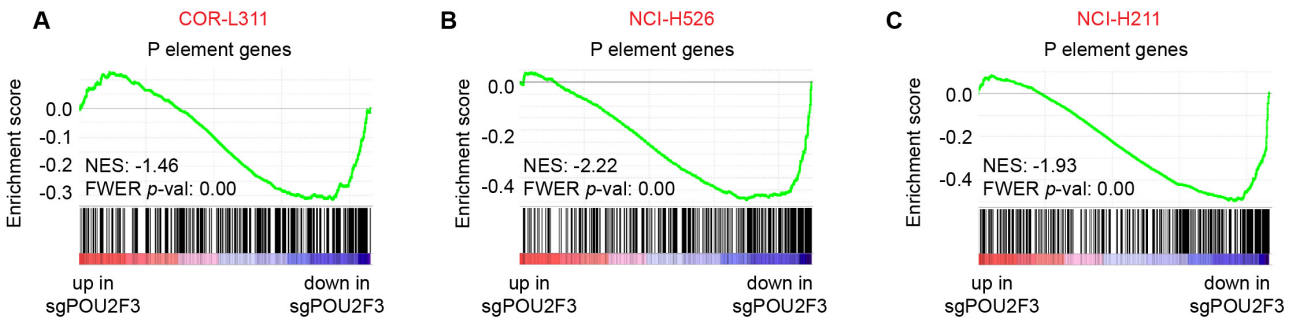

**Supplemental Figure S14. Motif analysis of P, N, and A elements. A** 500 bp region surround the summit of H3K27ac enrichment of indicated elements were evaluated by using motif enrichment analysis, using the HOCOMO and JASPAR databases plus additional NEUROD1 and ASCL1 motifs.



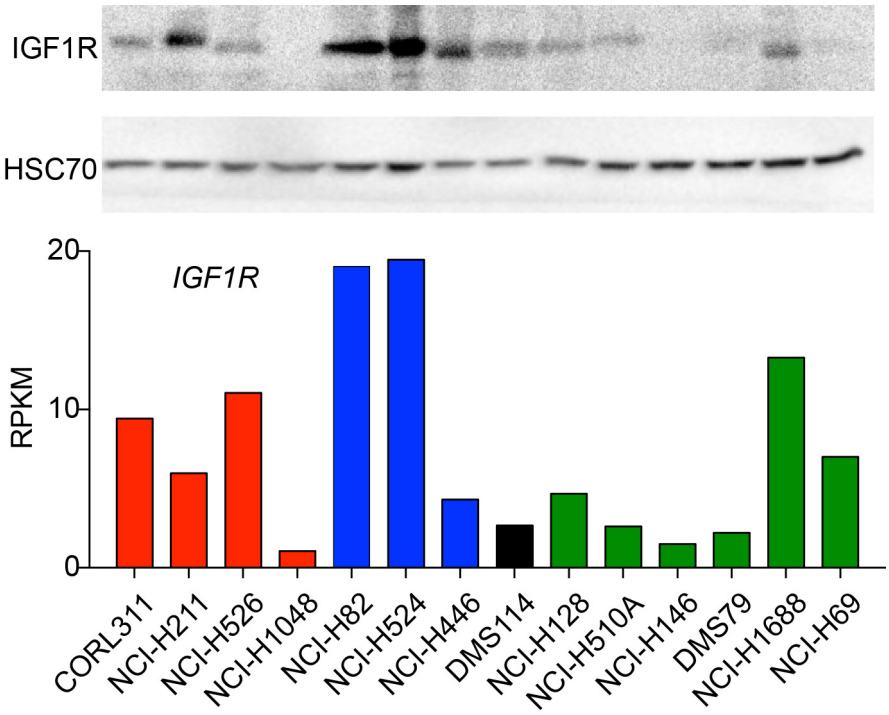

**Supplemental Figure S15. Metaprofile analysis of POU2F3, NEUROD1, and ASCL1 ChIP-seq data at P, N, and A elements.** ChIP-seq meta-profiles of POU2F3 (A), NEUROD1 (B), ASCL1 (C) occupancy at P (red), N (blue), and A (green) elements in the indicated cell lines. NEUROD1 and ASCL1 ChIP-seq datasets were obtained from Borromeo *et al.* (3).



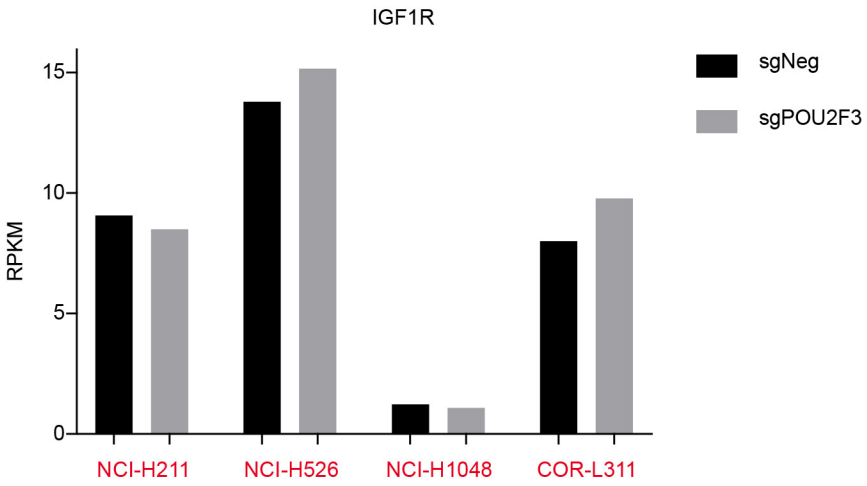

**Supplemental Figure S16. P elements tend to be located a distal locations from promoters**. Pie chart showing genomic location of 463 P elements. TTS: Transcription termination site.



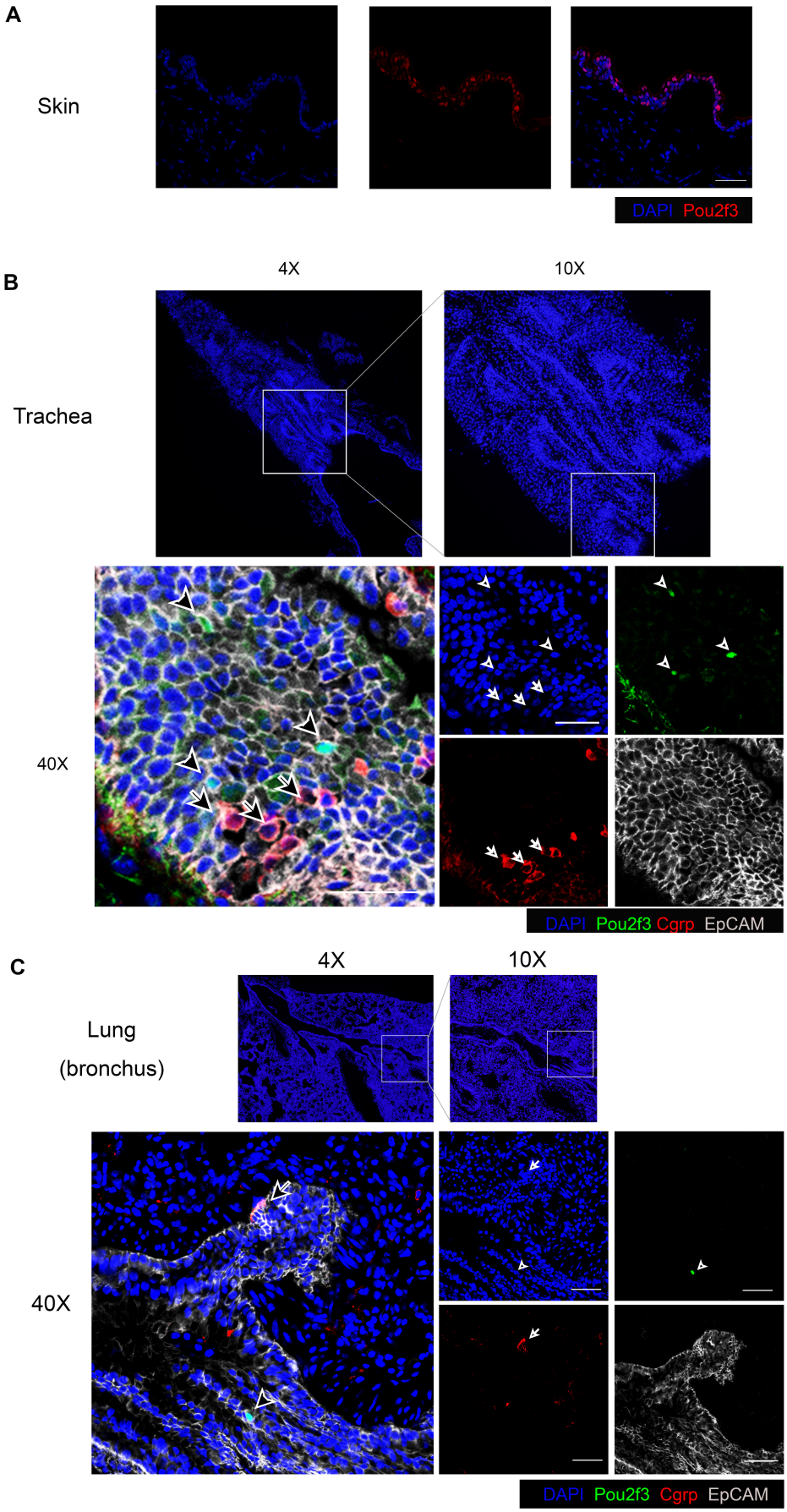

**Supplemental Figure S17. Inactivating POU2F3 leads to reduced expression of genes located near P elements.** sgRNA (sgPOU2F3 or sgNeg) infected POU2F3$^{high}$ cell line COR-L311 (A), NCI-H526 (B), and NCI-H211 (C) cells were collected for RNA-seq analysis. The P element genes defined as the nearest expressed genes to each P element. NES: normalized enrichment score. FWER p-val: family-wise error rate p value. The P element gene set is provided in Supplemental Table 5.




**Supplemental Figure S18. IGF1R protein and mRNA levels across a panel of SCLC cell lines.** (top) IGF1R protein levels detected by western blotting. (Bottom) IGF1R mRNA levels (reads per kilobase of transcript per million mapped reads from RNA-seq analysis) in the indicated SCLC cell lines.

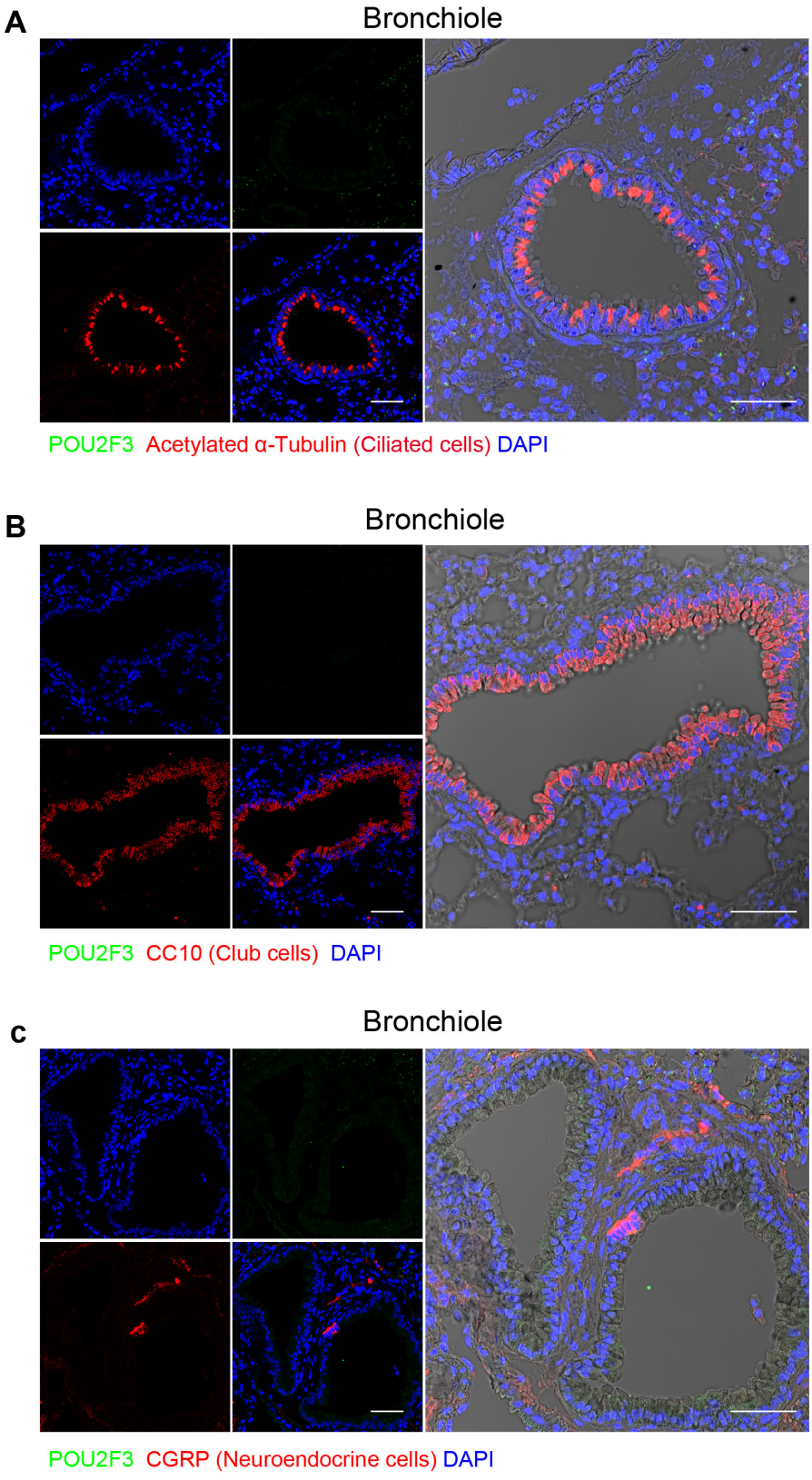



**Supplemental Figure S19. IGF1R expression level is not affected by perturbation of POU2F3.** sgRNA (sgPOU2F3 or sgNeg) infected POU2F3$^{high}$ cell NCI-H211, NCI0H526, NCI-H1048 and COR-L311 cells were collected for RNA-seq analysis. IGF1R mRNA expression levels were compared between control (sgNeg) and the knock out of POU2F3 (sgPOU2F3e10.1 and sgPOU2F3e10.2) in indicated cell lines.




**Supplemental Figure S20. POU2F3 expressing cells and neuroendocrine cells are distinct populations in the mouse respiratory tract.** A) Immunofluorescence staining of POU2F3 in the mouse epidermis as a positive control. B) Representative images of immunofluorescence staining

for POU2F3 (green), CGRP (red), and EpCAM(white) in mouse trachea (A) and lung (B) tissues. Arrow heads: POU2F3 expressing cells; arrows: CGRP expressing cells.




**Supplemental Figure S21. A lack of POU2F3 expressing cells in the epithelial layer of mouse bronchioles.** Representative immunofluorescence images staining for POU2F3, CGRP, CC10, or DAPI.

**References**


1. 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(7563):47-53 doi 10.1038/nature14664.

2. Sato T, Kaneda A, Tsuji S, Isagawa T, Yamamoto S, Fujita T, *et al*. PRC2 overexpression and PRC2-target gene repression relating to poorer prognosis in small cell lung cancer. Sci Rep **2013**;3:1911 doi 10.1038/srep01911.

3. 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 reports **2016**;16(5):1259-72 doi 10.1016/j.celrep.2016.06.081.