NAR Cancer, 2026, **8**, zcag032
https://doi.org/10.1093/narcan/zcag032
Cancer Gene Regulation, Chromatin, and Epigenetics

OXFORD

Marker-based CRISPR screens identify POU2F1 as a regulator of DLL3 and neuroendocrine identity in small cell lung cancer

Patrick J. Cunniff[1,2,†], **Connor Fitzpatrick**[1,2,†], **Jack Bauer**[1,3,4], **Damianos Skopelitis**[1,5], **Sujay Pal**[1,5], **Liam Shanley**[1], **Kamil Taneja**[1], **Olaf Klingbeil**[1], **Toyoki Yoshimoto**[1,5], **Leemor Joshua-Tor**[1,3,6], **Christopher R. Vakoc**[1,5,*]

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, United States
[2]Cold Spring Harbor Laboratory School of Biological Sciences, Cold Spring Harbor, NY 11724, United States
[3]W. M. Keck Structural Biology Laboratory, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, United States
[4]Graduate Program in Molecular and Cellular Pharmacology, Stony Brook University, Stony Brook, NY 11794, United States
[5]Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN 38105, United States
[6]Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, United States

*To whom correspondence should be addressed. Email: Christopher.Vakoc@STJUDE.ORG
†The first two authors should be regarded as Joint First Authors.

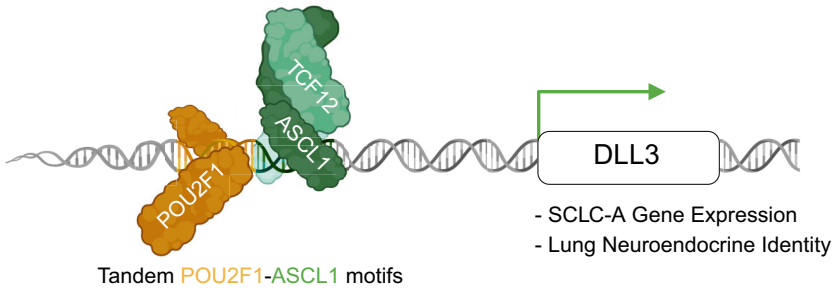
Abstract

Small cell lung cancers (SCLCs) often exhibit a neuroendocrine lineage identity marked by high expression of Delta-like ligand 3 (DLL3). Because DLL3 shows minimal expression in normal adult tissues, it serves as an SCLC-selective tumor antigen and is the basis for clinically efficacious targeted therapies. Understanding the mechanisms that regulate DLL3 expression is therefore critical for advancing therapeutic strategies in this disease. Here, we performed transcription factor-focused and genome-wide CRISPR screens to identify regulators of DLL3 expression in SCLC. Both approaches converged on POU2F1 as a top activator of *DLL3* in this tumor context. Despite its ubiquitous expression, we identify an SCLC-specific role for POU2F1 in activating *DLL3* and a broader set of neuroendocrine lineage genes. Epigenomic analyses reveal tandem POU2F1–ASCL1 motifs within the *DLL3* promoter that underlie the functional cooperativity between POU2F1 and the neuroendocrine master regulator ASCL1 for high-level DLL3 expression in SCLC. We provide evidence that tandem POU2F1–ASCL1 elements are part of a *cis*-regulatory code for the lung neuroendocrine cell fate. Together, these findings define a previously unrecognized transcriptional logic controlling DLL3 expression and establish POU2F1 as a context-specific regulator of neuroendocrine lineage in SCLC.

Impact statement

Small cell lung cancer (SCLC) is an aggressive disease increasingly treated with therapies targeting a protein called DLL3. Using systematic gene-disruption screens, this study identifies POU2F1, a gene regulator found throughout the body, as an unexpected controller of DLL3 expression in SCLC. POU2F1 works with ASCL1, a regulator of neuroendocrine cell identity, to activate DLL3 production. Cooperation between POU2F1 and ASCL1 appears to extend to many other genes that give these tumors their neuroendocrine characteristics. Findings provide a foundation for testing whether changes in the POU2F1–ASCL1 partnership contribute to loss of DLL3 expression and thus influence responses to DLL3-targeted therapies.

Graphical abstract



Introduction

Small cell lung cancer (SCLC) has long been recognized as a distinct clinicopathologic entity characterized by rapid cell proliferation, early dissemination, and a robust neuroendocrine phenotype [1]. From its earliest pathological descriptions, SCLC tumors were noted to express neuroendocrine

Received: April 18, 2026. **Revised:** August 27, 2026. **Accepted:** September 26, 2026
© The Author(s) 2026. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

markers such as synaptophysin and chromogranin A [2], linking this disease to neuroendocrine differentiation before the advent of molecular profiling [3]. Clinically, this lineage identity has shaped diagnostic criteria, informed therapeutic approaches, and distinguished SCLC from other lung cancer subtypes, despite shared etiologic factors such as tobacco exposure [1]. More recent genomic and transcriptomic studies have refined this view by revealing that SCLC comprises multiple molecular subtypes defined by dominant transcriptional regulators, including ASCL1, NEUROD1, and POU2F3 [4]. Among these, ASCL1-positive SCLC, hereafter referred to as SCLC-A, represents the most prevalent and canonical neuroendocrine subtype [4].

A prominent feature of the neuroendocrine state in SCLC is high expression of Delta-like ligand 3 (DLL3), an atypical member of the Notch ligand family [5]. In contrast to canonical Delta-like ligands that activate Notch signaling at the cell surface, DLL3 primarily localizes to intracellular compartments and functions as a negative regulator of Notch pathway activity during development [6, 7]. DLL3 expression is largely restricted to embryonic neuroendocrine contexts and is minimally expressed in most adult tissues [8]. In SCLC, however, DLL3 is highly and selectively expressed, particularly within tumors with neuroendocrine lineage identity [9]. This restricted expression pattern has positioned DLL3 as a clinically important tumor-associated antigen and has driven the development of multiple DLL3-directed therapeutic strategies [10–16]. Beyond its translational relevance, DLL3 has emerged as a robust marker of neuroendocrine differentiation in SCLC, linking suppression of Notch signaling to maintenance of lineage identity [17].

Despite its central role in SCLC biology and therapy, the transcriptional mechanisms that govern DLL3 expression in this disease remain incompletely defined. Prior work has largely inferred DLL3 regulation indirectly through its association with ASCL1-positive tumors and low Notch signaling, but the specific transcription factors and *cis*-regulatory elements responsible for DLL3 activation have not been systematically characterized [17]. This gap is notable given the growing clinical reliance on DLL3-directed therapies and the increasing recognition that SCLC exhibits substantial transcriptional plasticity [18, 19]. A clearer understanding of how DLL3 expression is encoded at the level of transcriptional regulation is therefore necessary to explain its subtype specificity, predict its stability under therapeutic pressure, and identify mechanisms by which DLL3 expression may be gained or lost during tumor evolution.

Functional genetic screening has emerged as a powerful and unbiased approach to dissect the transcriptional and lineage dependencies that underlie SCLC biology [20–23]. CRISPR-based loss-of-function screens have been particularly informative in this disease, revealing subtype-specific regulators, lineage vulnerabilities, and chromatin dependencies that are not readily inferred from gene expression or genomic data alone [23]. Using this strategy, our group previously identified POU2F3 as a master regulator of a non-neuroendocrine tuft cell-like SCLC subtype [22] (often referred to as SCLC-P) and uncovered a selective dependence of these tumors on specific transcriptional coactivators [21, 24], establishing a mechanistic framework for tuft cell-like identity in SCLC. Studies from other groups have similarly leveraged CRISPR screening to define core transcriptional regulators, chromatin modifiers, and signaling pathways that sustain distinct SCLC states and govern lineage plasticity [25–27]. These studies highlight the power of functional screening to identify previously unknown lineage regulators.

POU2F1, also known as OCT1, is a ubiquitously expressed POU domain transcription factor best known for its roles in basal promoter activity, metabolism, stress responses, and signal-dependent gene regulation across diverse cell types [28–32]. Unlike lineage-restricted POU transcription factors, POU2F1 has generally been viewed as a transcription factor that supports general cellular functions and integrates environmental cues [33]. Prior studies have implicated POU2F1 in processes such as immune activation [34], metabolic adaptation [35], and cellular stress tolerance [36], with context-dependent effects shaped by chromatin accessibility and cofactor availability. In a few contexts, POU2F1 has been found to regulate normal and malignant cell identity programs, often through cooperativity with other transcription factors [37–39]. However, POU2F1 has not been previously linked to neuroendocrine differentiation or to neuroendocrine lineage cancers. As a result, it remains unclear how a widely expressed transcription factor such as POU2F1 might acquire specialized functions within specific tumor lineages.

In this study, we set out to define the transcriptional mechanisms that drive DLL3 expression in SCLC. By combining two CRISPR screening approaches, we identify POU2F1 as a previously unappreciated regulator of DLL3 expression and of neuroendocrine transcriptional programs in this disease. We show that POU2F1 collaborates with ASCL1 through a defined *cis*-regulatory architecture at the *DLL3* promoter, providing a mechanistic explanation for how a ubiquitously expressed transcription factor can promote lineage-restricted gene expression. Extending beyond *DLL3*, our findings reveal a regulatory principle in which tandem POU2F1–ASCL1 motifs encode neuroendocrine transcriptional programs. Together, this work links lineage identity to tumor-selective antigen expression in SCLC and offers insight into how neuroendocrine cell states are maintained.

Materials and methods

Institutional approval

This study complies with all relevant ethical regulations, and all protocols were approved by the Cold Spring Harbor Laboratory Institutional Biosafety Committee.

Cancer cell lines and tissue culture

The following cell lines used in this study were obtained from ATCC: HEK293T (Cat# CRL-3216; RRID:CVCL_0063), NCI-H209 (male, Cat# HTB-172; RRID:CVCL_1525), NCI-H1836 (male, Cat# CRL-5898; RRID:CVCL_1498), and NCI-H1436 (male, Cat# CRL-5871; RRID:CVCL_1471).

HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS). NCI-H209, NCI-H1436, and NCI-H1836 cells were cultured in HITES media supplemented with 5% FBS. HITES media is composed of DMEM:F12 base medium supplemented with 0.005 mg/ml insulin, 0.01 mg/ml transferrin, 30 nM sodium selenite, 10 nM hydrocortisone, 10 nM β-estradiol, and 4.5 mM L-glutamine. One percent penicillin–streptomycin was added to all media. Cell lines were purchased from commercial vendors and their identity was validated by STR analysis. Cell lines were regularly tested

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

for *Mycoplasma* contamination. All antibiotic concentrations used to select gene cassettes were empirically titrated in each cell line to achieve maximum selection with minimum toxicity.

Lentiviral production and infection

Lentivirus was produced in HEK293T cells transfected with target plasmids and packaging plasmids (VSV-G and psPAX2) using polyethyleneimine. Transfection media was replaced with fresh HITES media supplemented with 5% FBS 6 h after transfection, and lentivirus-containing supernatant was collected 24, 48, and 72 h following transfection. All three collections were pooled and filtered using a 0.45 μm PES filter. For lentiviral infections, cell suspensions were exposed to lentiviral-containing supernatant supplemented with polybrene to a final concentration of 4 μg/ml and spun at 600 RCF for 30 min. Lentiviral media was replaced with fresh media after 24 h.

Intracellular FACS-based CRISPR screens

The lentiGuide-Puro backbone (Addgene #52963) containing the genome-wide Brunello [40] (Addgene #73178) single guide RNA (sgRNA) library was modified prior to library cloning to replace the puromycin-resistance cassette with a GFP-P2A-hygromycinR cassette. After empirical determination of the suitable lentiviral titer, $\sim 5 \times 10^7$ (transcription factor-focused screen) or 7.5×10^8 (genome-wide screen) Cas9-expressing cells were infected with human DNA Binding Domain-Focused [22] (Addgene #123334) or genome-wide Brunello sgRNA library-encoding lentiviral suspension for a 20%–30% infection efficiency. Media was replaced with HITES media at 48 h. For the genome-wide screen, hygromycin B was added for 72 h to select for transduced cells. Eight days post infection, cells were collected, treated with Accumax™ (Thermo Fisher) to dissociate cell clumps, washed in phosphate buffered saline (PBS), and then fixed in −20°C methanol at $\leq 10 \times 10^6$ cells/ml under gentle vortexing. Cells were stored in methanol at −20°C for at least 2 days and up to 1 month. One day before sorting, cells were pelleted, washed 1× in FACS buffer [1% (w/v) ultrapure bovine serum albumin, 0.05% (w/v) sodium azide, and 1 mM ethylenediaminetetraacetic acid (EDTA) in magnesium- and calcium-free PBS], and incubated overnight in 1:200 primary antibody (DLL3) in FACS buffer at 10×10^6 cells/ml rotating at 4°C. The next day, cells were pelleted, washed twice with FACS buffer, and incubated for 2 h in 1:500 secondary antibody (AlexaFluor647-conjugated anti-rabbit) in FACS buffer at 10×10^6 cells/ml rotating at 4°C protected from light. After washing 2× in FACS buffer, cells were resuspended in FACS buffer at 10×10^6 cells/ml and sorted. Stained cells were sorted using a BD FACS Aria II cell sorter. The total number of cells sorted per screen was a minimum of 2500× the size of the sgRNA library. Cells were sorted into three different populations, with ∼20% of the cells sorted into the $DLL3^{low}$ bin, 70% of cells sorted into the $DLL3^{bulk}$ bin, and 10% of the cells being sorted into the $DLL3^{high}$ bin. Cell pellets were then processed for DNA extraction and library preparation as described below. Custom sequencing primers were added for each respective cell population. All sequencing data from FACS-based screens were analyzed with MAGeCK [41] v0.5.9.3 using the MLE option. For the transcription factor-focused DLL3 marker-based screen, the $DLL3^{high}$ gate was insufficiently diverse for downstream analysis, so the $DLL3^{low}$ and $DLL3^{bulk}$ gates were used to analyze this screen.

DNA extraction and sgRNA sequencing for CRISPR screens

After pooling and pelleting of sorted cells, cells were resuspended in DNA extraction buffer (10 mM Tris–HCl, pH 8.0, 150 mM NaCl, and 10 mM EDTA) at a density of $\leq 5 \times 10^7$ cells/ml. Sodium dodecyl sulfate (SDS) and Proteinase K were added to final concentrations of 0.1% and 0.2 mg/ml, respectively. The mixture was incubated for 48 h at 56°C, after which DNA was purified by phenol extraction. Equilibrated phenol was added 1:1 to the lysis mixture, mixed well, and centrifuged for 10 min at 20 000 RCF. Supernatant was carefully removed, and another phenol purification round was performed. DNA was then precipitated by adding three volumes of isopropanol and NaOAc pH 5.2 to a final concentration of 75 mM and incubating overnight at −20°C. DNA was pelleted at 20 000 RCF for 1 h, washed in 70% ethanol, and air-dried until translucent. After resuspension in sterile, ultrapure water, DNA was assessed for quality by NanoDrop before proceeding to library prep. sgRNAs were directly amplified from genomic DNA by one step polymerase chain reaction (PCR) using NEBNext® Ultra™ II Q5® Master Mix (NEB). Each PCR reaction was performed with 10 μg of genomic DNA in 100 μl final volume. Titrations of amplification cycles were performed for LRG or LentiCRISPRv2 (95°C, 1 min; n cycles [95°C, 30s; 53°C, 30s; 72°C, 30s]; 72°C, 10 min) sgRNA cassettes. All PCR reactions for each sample were pooled and 400 μl were taken for double-sided AMPure bead cleanup [0.65× + 1× bead volume to preserve PCR amplicons (∼192 and ∼274 bp, respectively)]. Amplicons were sequenced using an Illumina NextSeq with 50% spike-in or pooled with high-diversity libraries (Cold Spring Harbor Genome Center, Woodbury, NY 11797).

RNA extraction and RNA sequencing

Total RNA was extracted using TRIzol reagent following manufacturer's instructions. For RNA-seq experiments following CRISPR-based targeting of *POU2F1*, *ASCL1*, or controls, cells stably expressing Cas9 were infected with control or target sgRNAs in an LRG2.1_Blast vector (Addgene #65656) to >90% GFP positivity. RNA was collected at the indicated timepoints (4 days or 6 days following lentiviral transduction).

RNA-sequencing libraries were constructed using the TruSeq Sample Prep Kit V2 (Illumina) following manufacturer's instructions. Briefly, 2 μg of extracted, purified RNA was poly-A selected and fragmented with fragmentation enzyme mix. Complementary DNA was synthesized with SuperScript™ II reverse transcriptase, followed by end repair, A-tailing, single-end indexed adaptor ligation, and PCR amplification. RNA-sequencing libraries were single-end sequenced for 76 bp using an Illumina NextSeq platform (Cold Spring Harbor Genome Center, Woodbury, NY 11797).

ChIP and ChIP-seq library construction

For each ChIP, cells were collected, washed 2× with room-temperature PBS, resuspended in Accumax™ (Thermo Fisher) at 10×10^6 cells/ml and rotated at room temperature for 2 h. Cells were then washed 2× with room tempera-

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

ture PBS and resuspended at 5–10 × 10^6 cells/ml in room-temperature PBS. Cell suspensions were crosslinked in 1% formaldehyde at room temperature for 15 min, followed by the addition of glycine to quench the reaction at a final concentration of 0.125 M. After two ice-cold PBS washes, cells were resuspended in cell lysis buffer (10 mM Tris, pH 8.0, 10 mM NaCl, 0.2% NP-40) at 10 × 10^6 cells/ml and incubated on ice for 15 min. After spinning down, supernatant was removed and nuclei were resuspended in nuclear lysis buffer (50 mM Tris, pH 8.0, 10 mM EDTA, pH 8.0, 1% SDS) at 20 × 10^6 cells/ml and sonicated in 15 ml tubes using a Bioruptor® Pico water-bath sonicator (13 × 30s on/off cycles, 1 ml per tube) to achieve an average chromatin size distribution of 200–500 base pairs. Each 1 ml of sonicated chromatin from 20 million cells was diluted with 7 ml of IP-Dilution buffer (20 mM Tris, pH 8.0, 2 mM EDTA, pH 8.0, 150 mM NaCl, 1% Triton X-100, 0.01% SDS), and 200 µl of the sample was saved for input. Chromatin from 20 million (H3K27ac) or 100 million (POU2F1, ASCL1) cells was incubated with 4 µg (H3K27ac) or 10 µg (POU2F1, ASCL1) of the appropriate antibody and 25–100 µl of magnetic protein-A (rabbit) or protein-G (mouse) beads at 4°C overnight. Beads were then pooled for each respective IP and washed once with 1 ml IP-wash buffer 1 (20 mM Tris, pH 8.0, 2 mM EDTA, pH 8.0, 50 mM NaCl, 1% Triton X-100, 0.1% SDS), twice with 1 ml High-salt buffer (20 mM Tris, pH 8.0, 2 mM EDTA, pH 8.0, 500 mM NaCl, 1% Triton X-100, 0.01% SDS), once with IP-wash buffer 2 (10 mM Tris, pH 8.0, 1 mM EDTA, pH 8.0, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate), and twice with 1 ml TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 8.0). Chromatin was eluted from beads and reverse-crosslinked in 200 µl nuclear lysis buffer supplemented with 12 µl NaCl and 1 µg/ml RNase A by shaking at 800 rpm for ≥4 h at 65°C. Supernatant was isolated from magnetic beads and protein digestion was performed by adding 4 µg/ml of Proteinase K and incubating the mixture for 2 h at 56°C. NaOAc pH 5.2 was added to a final concentration of 75 mM and the DNA to be used for library prep was purified using the QIAGEN PCR purification kit and following manufacturer's instructions.

Each ChIP-seq library was constructed using the Illumina TruSeq ChIP Sample Prep kit following manufacturer's instructions. Briefly, ChIP DNA was end repaired, A-tailed, and ligated to Illumina-compatible single-indexed adaptors. Fifteen PCR cycles were used for final library amplification. After amplification, the library was purified 2× with 1× AMPure XP beads and analyzed on a Bioanalyzer using a high-sensitivity DNA chip (Agilent). Library DNA concentrations were quantified using an Invitrogen Qubit 4 Fluorometer using the 1× High-Sensitivity double-stranded DNA (dsDNA) assay. ChIP-seq libraries were single-end sequenced for 76 or 100 bp at a sequencing depth of ≥50 million raw reads (H3K27ac or input) or ≥30 million raw reads (POU2F1, ASCL1) per sample using an Illumina NextSeq platform (Cold Spring Harbor Genome Center, Woodbury, NY 11797).

General computational, data, and statistical analyses

All sequencing data were analyzed using the CSHL High-Performance Computing system. Packages used to analyze next-generation sequencing data were installed in independent Anaconda environments to minimize dependency conflicts. Downstream analyses were performed using Python 3 in JupyterLab notebooks or in RStudio. Generalized Additive Mixed Model (GAMM) analysis, t-tests, Mann–Whitney U tests, two-way ANOVA with Šidák's test, and Fisher's Exact statistical tests were performed using R or GraphPad PRISM v10. Cancer Cell Line Encyclopedia data were downloaded from the Depmap portal (25Q3). Human SCLC gene expression data were reanalyzed from George *et al.* (2015).

RNA-sequencing analysis

Single-end 76 bp raw sequencing reads were pseudo-aligned to the hg38 genome using Kallisto [42] with bootstrap 100. Low-abundance transcripts were removed, and variance-stabilized normalized transcripts and differential expression were calculated using DESeq2 [43]. Transcripts per million (TPM) were calculated from aligned, mapped reads normalized for gene length using tximport.

Gene set enrichment analysis (GSEA) was conducted using GSEA_4.3.3 [44]. The DESCARTES Fetal Lung Neuroendocrine GSEA Signature was downloaded from the Molecular Signatures Database (MSigDB) [45]. GSEA was performed using rank-ordered $\log_2$(Fold Change) values and the GSEA "Preranked" option. Gene sets are included as Supplementary Table S7. Gene Ontology analysis was conducted using Enrichr [46] (Ma'ayan Lab) using the top 250 significantly downregulated genes (RNA-sequencing analysis), or all genes adjacent to each respective set of peaks (ChIP-sequencing analysis and motif analysis). For Supplementary Fig. S5B, genes associated with individual POU2F1 or ASCL1 motifs alone were randomly downsampled to 227 genes to match the size of the tandem-motif-associated gene set.

ChIP-sequencing analysis

Single-end 76 base pair sequencing reads were mapped to the hg38 genome using Bowtie2 [47] with default settings. MACS2 [48] v2.2.9.1 was used to call peaks using no input genomic DNA control. Annotation and motif analysis of ChIP-seq peaks was performed using HOMER [49] v5.11 with default settings. To visualize genomic tracks, bigWig files were generated from sorted, indexed BAM files using deepTools [50] v3.5.2 bamCoverage function. Reads from single-end sequencing were extended based on sonication fragment size for heatmap generation (300 base pairs).

To define BED files of peaks and peak overlaps, MACS2 output narrowPeak or broadPeak files were merged using bedtools [51] v2.30.0 intersect tools. Regions of high artifactual mapping to chromatin, "blacklisted" regions [52], were removed from BED files using bedtools intersect prior to each analysis. Heatmaps and average chromatin occupancy metaplots were generated using computeMatrix and plotHeatmap functions of deepTools, taking bigWig and BED files as input.

Motif analysis

Motif analysis of ChIP-seq peaks was performed using HOMER [49] v5.11 with default settings.

The tandem POU2F1–ASCL1 motif was generated manually by combining position weight matrix (PWM) values for the *Homo sapiens* POU2F1 (MA0785.1) and ASCL1 (MA1100.2) motifs. The tandem POU2F1–ASCL1 motif PWM graphic was generated using the Tomtom [53] motif comparison tool (v. 5.5.9) in the Multiple Expectation-Maximization for Motif Elicitation (MEME) suite.

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

The Find Individual Motif Occurrences (FIMO) [54] tool (v. 5.5.9) in the MEME suite was used to identify occurrences of ASCL1, POU2F1, and tandem POU2F1–ASCL1 motifs in the human genome (hg38). Significance thresholds of 1×10^{-7} (tandem POU2F1–ASCL1 motifs and "reversed" tandem POU2F1–ASCL1 motifs) and 1×10^{-5} (ASCL1 and POU2F1 motifs) were applied. Annotation of genomic loci containing each set of motifs was performed using HOMER [49] v5.11 with default settings.

AlphaFold 3 prediction analysis

AlphaFold 3 (using Google's AlphaFold 3 server) [55] was used to generate 20 predictions of the structure of the ASCL1–POU2F1–TCF12 complex from four separate starting seeds. Predictions were generated using either a tandem POU2F1–ASCL1 motif (5′-**ATTTGCATA***CAGCTG*T-3′ and complement), a "reversed" tandem motif that swaps the relative positions of the POU2F1- and ASCL1- binding motifs (5′-*CAGCTG*T**ATTTGCATA**-3′ and complement), or without any DNA. Generated predictions were then examined in ChimeraX [56] v1.11.1 and, using the "align" command, manually aligned to the structured domains of the highest-scoring model. For generation of the mean distance of residues/nucleotides from the mean aligned position, a Python script (written predominantly manually with some input by ChatGPT for the fixing of syntax errors and code annotation) was used to find the distance of each residue's Cα (for proteins) or C1' (for DNA) atom from the mean position of the atom across all 20 models, and repeated for all 20 models, to provide a measure of model consistency relative to the local environment. In addition, the corresponding pLDDT values for each Cα or C1' atom from the AlphaFold 3 output JSON files were extracted and exported to a CSV file along with the distance data and manually input into GraphPad Prism for further analysis. ChimeraX was used to find potential hydrogen bonds and van der Waals contacts, with hydrogen bonds split into "strict" matches to ChimeraX's hydrogen bonding criteria, and "relaxed" matches that allowed for a 1 Å increase to the distance cutoff and 35° to the angle cutoff, and van der Waals contacts were found using a VDW overlap of 0.4 Å in ChimeraX. A Python script was then used to extract the number of occurrences of "strict" hydrogen bonds, "relaxed" hydrogen bonds, and van der Waals contacts between each atom pair from the ChimeraX session, document the occurrence of each type of interaction between residue pairs across all 20 models (with only one allowed per residue, with "strict" hydrogen bonds taking priority, followed by "relaxed" hydrogen bonds, then van der Waals contacts), extract predicted aligned error (PAE) and estimated contact probability data from the AlphaFold 3 JSON outputs for each model, and output to a CSV file. Due to the large variability between predictions both within and across model seeds, especially with the predictions without DNA and the "reversed" motif predictions, all 20 outputs were used as individual model outputs. Welch's unpaired *t*-test was used to find significant differences in model confidence and consistency of the DNA bases, as changes to the DNA sequences used in "reversed" motif predictions made pairwise comparisons unreliable.

CellTiter-Glo® cell viability analysis

Three Cas9 + SCLC-A cell lines (NCI-H209, NCI-H1436, and NCI-H1836) were hyper-infected with a guide targeting POU2F1, ROSA (negative control), or CDK1 (positive control). Two days after lentiviral transduction, cells were selected with blasticidin for 5 days. Following selection (day 7 post-transduction), cells were mechanically dissociated into single-cell suspensions and 2000 viable single cells in 100 µl of media were plated in each well (three wells per condition) in an opaque flat-bottom 96-well plate (Nunclon Delta Surface, Thermo Scientific, Cat. No. 136101) for the cell viability assay. Cell viability was checked one week (14 days post transduction) after plating. This experiment was performed in three independent biological replicates. For the cell viability assay, 100 µl of equilibrated CellTiter-Glo® 2.0 Cell Viability Assay reagent (Promega, Cat. No. G9241) was added to each well at a 1:1 ratio with the culture medium. Plates were shaken on an orbital shaker for 2 min and then incubated for 10 min at room temperature in the dark. Luminescence was measured using a SpectraMax i3 plate reader with standard luminescence settings.

Cloning and molecular biology

Oligonucleotides encoding primers and sgRNAs were ordered from Sigma–Aldrich. Pre-made vectors and gene cassettes were ordered from Addgene, and new DNA fragments were ordered as gBlocks from Integrated DNA Technologies. The Takara® In-Fusion HD or Snap Assembly cloning kit was used to clone new plasmids according to manufacturer's instructions. Plasmids were transformed into Stbl3™ competent *Escherichia coli* for antibiotic selection and plasmid amplification. All plasmid constructs were sequenced for validation prior to use.

Promoter activity assay

The genomic sequence upstream of the DLL3 transcriptional start site (939 bp, hg38 chr19: 39498008–39498947) was cloned upstream of the firefly *luc2* gene [in place of the minimal promoter] in the pGL4.23 vector (Promega). The tandem POU2F1–ASCL1 motif (5′-ATTTACATACAGCTG-3′) was mutated by either deleting the ASCL1 motif (5′-ATTTACATA——-3′), deleting the POU2F1 motif (5′———-CAGCTG-3′), or reversing the orientation of the motifs (5′-CAGCTGATTTACATA-3′) to generate mutant reporters. Two hundred nanograms of plasmid encoding GFP or 100 ng encoding ASCL1 and 100 ng encoding POU2F1; 100 ng of pGL4.23 reporter plasmid, 20 ng of *Renilla* luciferase-encoding plasmid (pGL4.74; Promega); 0.15 µl of Lipofectamine 3000 (Thermo Fisher); and 0.4 µl of P3000 reagent (Thermo Fisher) were diluted in Opti-MEM. The transfection mixture was added dropwise to HEK293T cells (10 000 cells per well in a 96-well plate) and incubated for 48 h. Dual-Glo® Luciferase Assay System (Promega) was used according to manufacturer's instructions to assess the relative luciferase activity. Luminescence per well was integrated over 1000 ms on a SpectraMax® i3 microplate reader (Molecular Devices). Experimental firefly luminescence within each well was normalized to *Renilla* luminescence (control). Fold change was calculated by dividing luminescence in the ASCL1 & POU2F1 co-transfection condition by luminescence in the GFP condition for a given promoter.

Western blots and protein analysis

Cells were collected, washed 2× with room temperature PBS, counted for each sample, and normalized by cell num-

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

ber. Equal numbers of cells were washed in PBS and resuspended in 1× Laemmli buffer, diluted in PBS, with 5% β-mercaptoethanol at a final concentration of ∼5 × 10^6 cells/ml. Samples were boiled at 98°C for 15 min and stored at 4°C for a maximum of 1 month, or at −20°C for long-term storage. Immediately prior to loading on an SDS–polyacrylamide gel electrophoresis gel, samples were boiled for 1 min and centrifuged at 20 000 RCF for 5 min. Antibodies are included as Supplementary Table S8.

HEK293T co-immunoprecipitation assays

Approximately 3 × 10^6 HEK293T cells were co-transfected with 3 μg each of HA-tagged POU2F1 and FLAG-tagged ASCL1 expression plasmids using polyethyleneimine. Forty hours after transfection, HEK293T cells were collected by scraping, washed once with ice-cold PBS, and lysed in buffer containing 20 mM Tris-Cl (pH 8.0), 20% glycerol, 2 mM EDTA, 150 mM KCl, 0.1% NP-40, protease inhibitor cocktail, 0.7 μl/ml β-mercaptoethanol, and 0.2 mM phenylmethylsulfonyl fluoride (PMSF). Cells were incubated in lysis buffer for 90 min at 4°C, and the lysates were subsequently clarified by centrifugation at 16 000 × *g* for 30 min at 4°C.

Five percent of each clarified lysate was saved as the input sample, with 0.8% loaded into the gel for western blotting). The remaining lysate was incubated with 20 μl of anti-HA magnetic beads (Thermo Scientific, Cat. No. 88836) for 12–14 h at 4°C with rotation. Following immunoprecipitation, the beads were washed three times with lysis buffer containing 0.1% NP-40 and resuspended in 45 μl of Laemmli sample buffer (Bio-Rad, Cat. No. 1610737) supplemented with 5% β-mercaptoethanol. Samples were heated at 98°C for 10 min and subsequently analyzed by western blotting.

Statistics and reproducibility

All statistical tests used to evaluate significance are detailed in the respective figure legends. All reported results were replicated across multiple experiments to generate reliable results, and replication is described in more detail in each figure legend. No statistical method was used to predetermine sample size. Sample sizes were selected based on previously published studies and established experimental practice. Figure legends indicate the sample sizes for each set of experiments. All sample sizes were determined to be sufficient given that the differences among groups were consistent. No data were excluded from any experiment. Randomization was not performed for any experiments. Investigators were not blinded to allocation during experiments and outcome assessment.

Results

A CRISPR screening strategy for identifying regulators of DLL3 expression in SCLC

Aberrant expression of the Notch ligand DLL3 is a defining feature of SCLC and underlies several therapeutic strategies, yet the mechanisms that drive DLL3 expression remain poorly understood. To identify upstream regulators of DLL3 in an unbiased manner, we developed a CRISPR screening strategy to systematically nominate genes required for DLL3 expression in SCLC. We selected the human SCLC cell line NCI-H209 for these screens based on its robust growth in culture and its high expression of both DLL3 and the neuroendocrine master regulator ASCL1 (Supplementary Fig. S1A and B; Supplementary Table S1). We identified an antibody capable of reliably and quantitatively detecting DLL3 protein and validated it by western blotting and flow cytometry, confirming its specificity using CRISPR-mediated inactivation of *DLL3* as a negative control (Fig. 1A and B).

For the CRISPR screens, Cas9-expressing NCI-H209 cells were transduced with two different lentiviral sgRNA libraries. The first library targets 1437 DNA-binding domains (DBDs) to enable efficient coverage of the human transcription factor repertoire (8739 sgRNAs total) [22]. The second was a genome-wide CRISPR sgRNA library targeting 19114 human genes (Brunello) [40]. On day 8 following transduction, 50 (DBD) or 750 (Brunello) million cells were harvested, dissociated to single cells, methanol-fixed, and stained with anti-DLL3 primary and fluorophore-conjugated secondary antibodies. Cells were then sorted by flow cytometry into populations with low, intermediate, or high DLL3 expression. Genomic DNA was isolated from each sorted population, and sgRNA cassettes were PCR-amplified and deep sequenced for quantification to identify genes required for DLL3 expression (Fig. 1C). The overall performance of both screens was supported by the behavior of sgRNAs targeting *ASCL1*, which is present in both libraries and is a known activator of DLL3 [57, 58], as well as sgRNAs targeting *DLL3* itself, which are present in the genome-wide library (Fig. 1D and E; Supplementary Table S2).

Two independent CRISPR screens identify POU2F1 as a potent activator of DLL3 expression in SCLC

Both CRISPR screens independently identified POU2F1 as one of the strongest positive regulators of DLL3 expression in NCI-H209 cells (Figs 1D and E; Supplementary Table S2). In addition to POU2F1 as a genetic outlier, the genome-wide screen also revealed enrichment for general messenger RNA (mRNA) processing factors as being needed for DLL3 expression (Fig. 1E and Supplementary Fig. S1C), consistent with a broader requirement for transcriptional and post-transcriptional machinery for DLL3 expression. To decouple DLL3-specific regulators from general transcriptional and post-transcriptional factors, we compared the genome-wide screening results to our previously published Serpin B5 marker-based screen performed in a pancreatic cancer cell line [59]. While many of the Mediator, RNA polymerase, and ribosomal subunits scored in both screens, this comparison validated ASCL1 and POU2F1 as specific regulators of DLL3 expression (Supplementary Fig. S1D). POU2F1 is a ubiquitously expressed member of the POU homeodomain transcription factor family, and its requirement for DLL3 expression was unexpected based on prior literature. Because POU2F1 has not been previously studied in neuroendocrine lineages or in SCLC, we focused subsequent analyses on validating its role in DLL3 regulation and assessing whether it contributes more broadly to neuroendocrine identity in SCLC.

Analysis of *POU2F1* expression across 1699 human cancer cell lines confirmed its widespread expression (Supplementary Fig. S1E and Supplementary Table S1). However, when cancer types were ranked by median *POU2F1* mRNA levels, SCLC lines exhibited the highest expression among all cancer types examined (Supplementary Fig. S1E). In both cancer cell lines and human tumors, *POU2F1* expression was higher in SCLC than in other lung cancer subtypes and was relatively uniform across the molecular subtypes of SCLC (Fig.

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

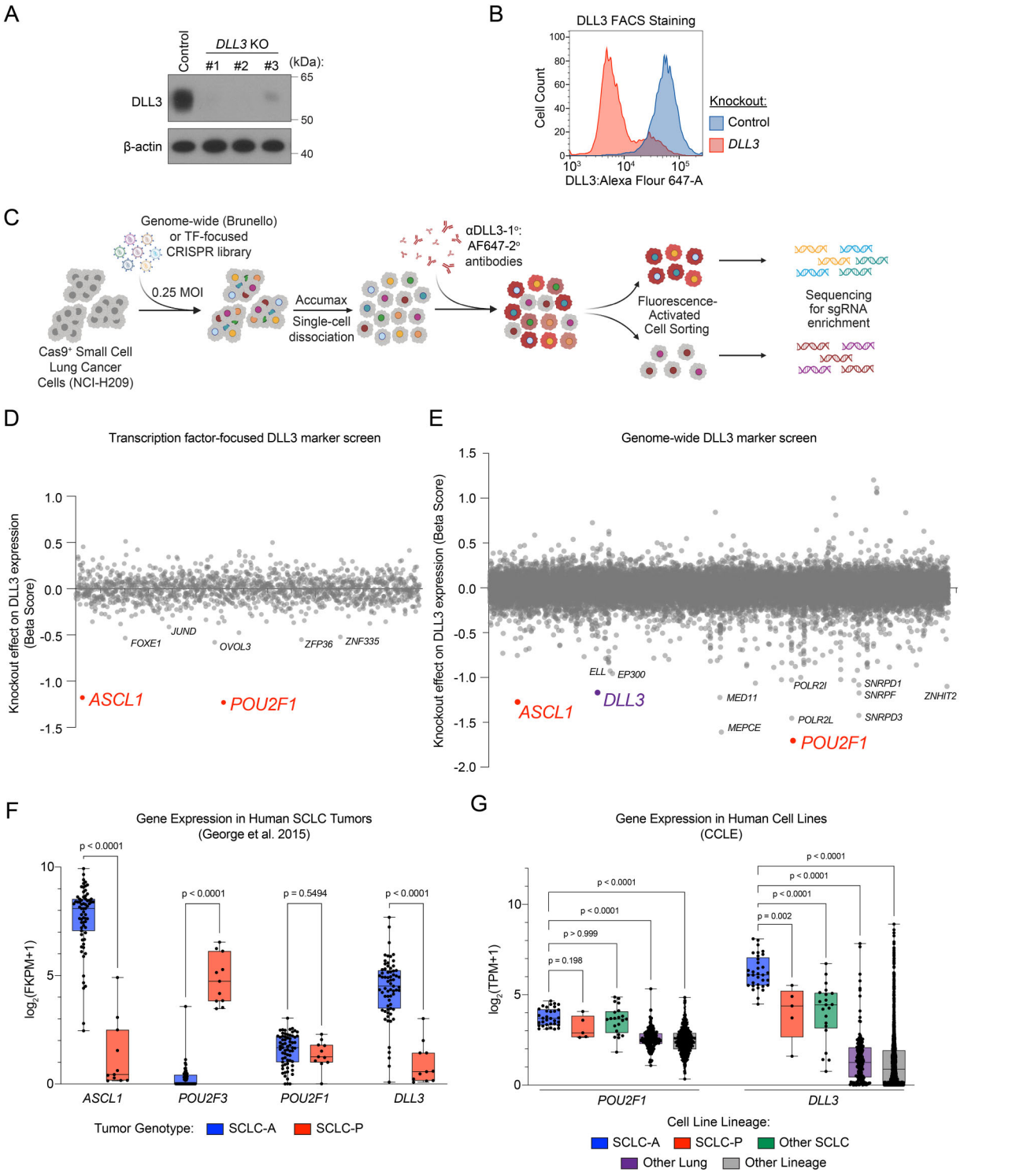


Figure 1. Parallel marker-based screens identify POU2F1 as a DLL3 regulator. (**A**) Western blot of DLL3 protein 7 days after CRISPR-Cas9 knockout (KO) of *DLL3* or *ROSA26* (control) in H209 cells. β-actin, loading control. Representative of two biological replicates. MOI, multiplicity of infection. (**B**) Flow cytometry analysis of H209 cells, methanol-fixed and stained with anti-DLL3 antibodies 8 days following KO of *DLL3* or control (*ROSA26*). Representative of two biological replicates. (**C**) Workflow of transcription factor-focused or genome-wide DLL3 reporter screen. Created in BioRender. Cunniff, P. (2026) https://BioRender.com/l9r3z2o. Transcription factor-focused (**D**) or Genome-wide (**E**) DLL3 reporter screen results in NCI-H209. Genes (dots) ordered alphabetically along the *x*-axis. Selected outlier genes are labeled. Beta scores and significance calculated using MAGeCK (maximum likelihood estimation). Negative beta scores indicate enrichment in the DLL3^{low} population. Gene expression in (**F**) resected human SCLC tumors – reanalyzed from George *et al.* (2015) or (**G**) human cell lines. Each point = 1 tumor or cell line. Boxes represent the median ± interquartile range (IQR). Error bars extend to the minimum and maximum values. Statistical significance was evaluated using two-sided Mann–Whitney tests with (**F**) no multiple comparison correction or (**G**) correction for four comparisons. FPKM = Fragments Per Kilobase of transcript per Million mapped reads. TPM = Transcripts Per Million.

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

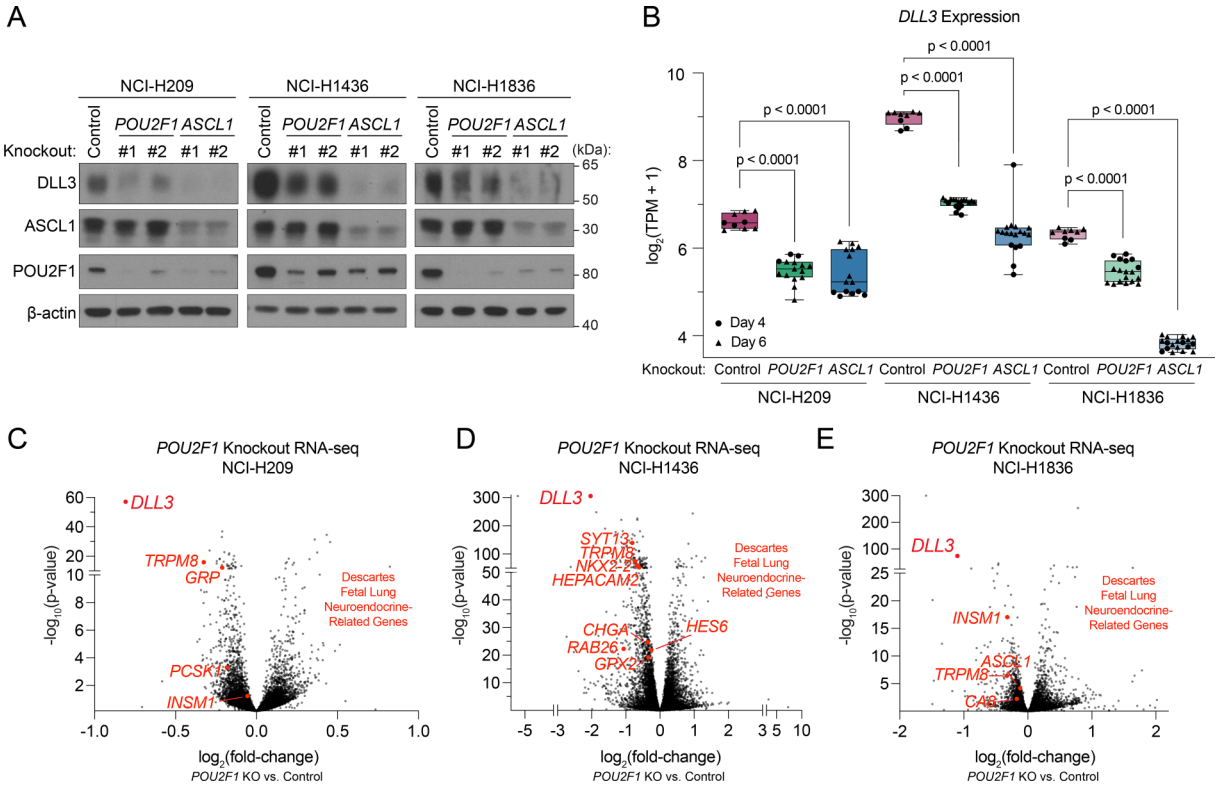


Figure 2. POU2F1 is required for DLL3 expression in SCLC. (**A**) Western blot of DLL3, ASCL1, and POU2F1. Protein lysates were collected from NCI-H209, NCI-H1436, and NCI-H1836 on day 7 following CRISPR-Cas9 KO of *ASCL1*, *POU2F1*, or *ROSA26* (control). Representative of two biological replicates. β-actin, loading control. (B–E) RNA sequencing in NCI-H209 (B–C), NCI-H1436 (B, D), and NCI-H1836 (B, E) on day 4 (B, C) or day 6 (B, D, E) following CRISPR-Cas9 KO of *POU2F1*, *ASCL1*, or *ROSA26* (control). (**B**) *DLL3* expression in TPM. Each point = 1 biological replicate. The day post-KO is indicated by the shape. Boxes represent the median ± IQR. Error bars extend to the minimum and maximum values. Statistical significance was evaluated using two-sided Mann–Whitney tests with no multiple comparison correction. (**C**–**E**) Volcano plots of differentially expressed genes following *POU2F1* KO. Fold change and significance calculated by DESeq2. Two independent experiments, each with six biological replicates. *DLL3* and additional genes related to fetal lung neuroendocrine identity (DESCARTES database; Cao *et al.* 2020 [45]) are labeled.

1F and G, Supplementary Fig. S1E, and Supplementary Tables S1 and S3) [60, 61]. This pattern contrasts with *DLL3* expression, which is selectively elevated in neuroendocrine SCLC-A relative to tuft cell-like SCLC-P subtypes (Fig. 1F and G, Supplementary Fig. S1F, and Supplementary Tables S1 and S3). Together, these findings suggest that while POU2F1 is highly expressed in SCLC, it is unlikely to be sufficient on its own to dictate DLL3 expression or neuroendocrine identity.

To validate the screening results, we performed western blot analysis following CRISPR-mediated KO of *POU2F1* in three independent neuroendocrine SCLC cell lines (NCI-H209, NCI-H1836, and NCI-H1436). Loss of POU2F1 resulted in a reduction of DLL3 protein levels in all three models (Fig. 2A). Consistent with a transcriptional mechanism, RNA-seq analysis demonstrated a significant decrease in *DLL3* mRNA expression following *POU2F1* KO (Fig. 2B and Supplementary Table S4). As a control, CRISPR-mediated inactivation of *ASCL1* similarly reduced *DLL3* mRNA and protein levels in these lines (Fig. 2A and B; Supplementary Table S4). Collectively, these data validate POU2F1 as a transcriptional activator required for DLL3 expression in multiple SCLC contexts.

To assess the functional consequences of POU2F1 loss, we measured the effects of *POU2F1* KO on SCLC cell proliferation. *POU2F1* KO produced a modest reduction in proliferation in all three SCLC-A cell lines (Supplementary Fig. S2A and B). To place these results in a broader context, we examined *POU2F1* dependency data from the Cancer Dependency Map (DepMap), which indicate that *POU2F1* is a weak dependency across all cancer types examined, with no lineage showing selective essentiality (Supplementary Fig. S2C and Supplementary Table S1). This contrasts with *ASCL1*, which is a strong dependency in SCLC-A (Supplementary Fig. S2B and D; Supplementary Table S1). These data led us to hypothesize that POU2F1 activates select neuroendocrine lineage identity genes in SCLC-A while having a minimal influence over genes that regulate cell proliferation.

POU2F1 supports the neuroendocrine transcriptional program in SCLC

We next examined the broader transcriptional consequences of POU2F1 inactivation in SCLC. Comparative analysis of RNA-seq datasets revealed that *DLL3* was among the most significantly and strongly downregulated genes following *POU2F1* KO across all three SCLC-A cell lines (Fig. 2B–E, Supplementary Fig. S3A–C, and Supplementary Table S4). Additional genes exhibiting strong POU2F1 dependence included canonical markers of neuroendocrine identity, such as *CHGA*, *INSM1*, *GRP*, and *TRPM8* (Fig. 2C–E, Supplementary Fig. S3A–C, and Supplementary Table S4) [62]. While the overall magnitude of transcriptional changes following POU2F1

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

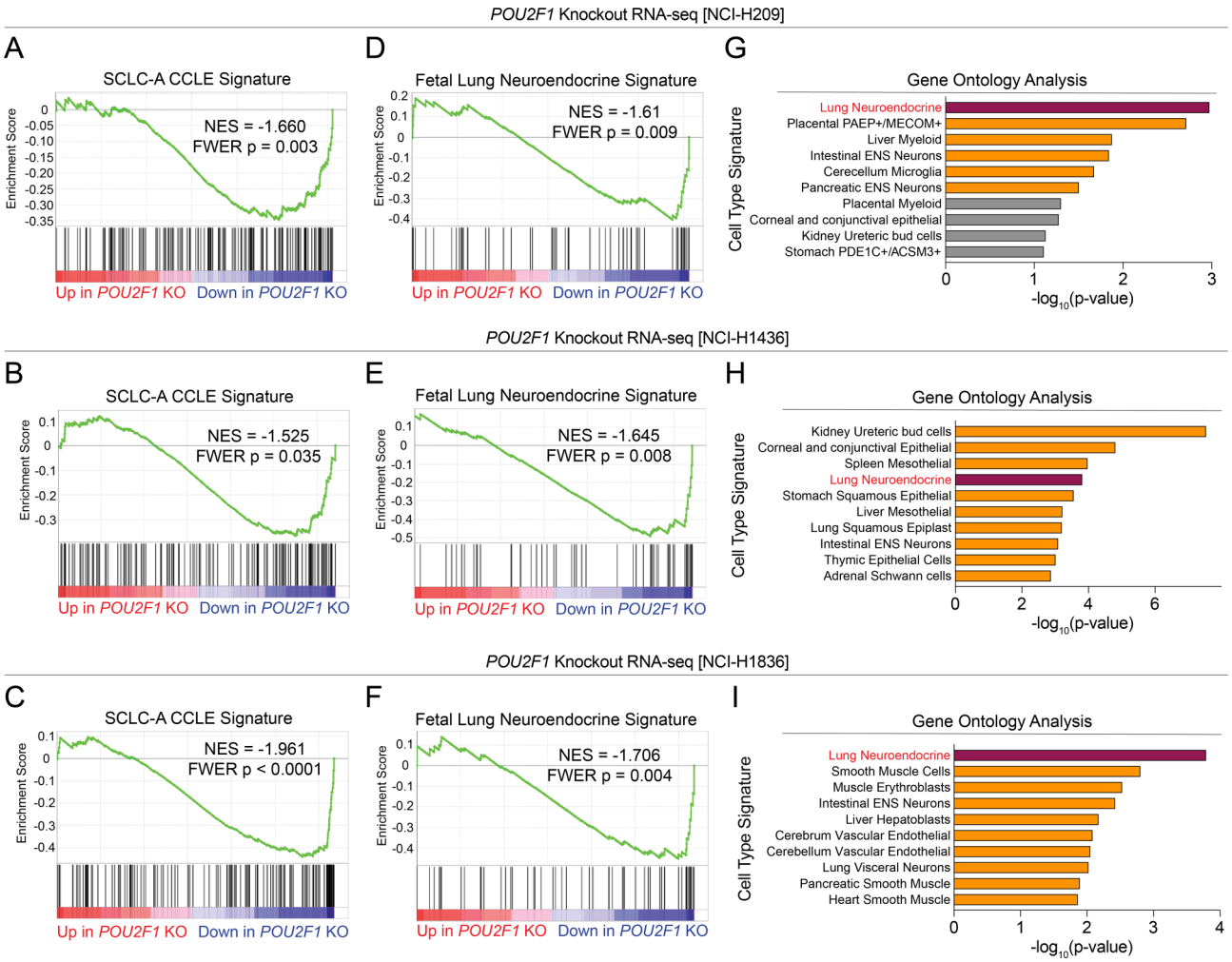


Figure 3. POU2F1 activates a neuroendocrine identity gene expression program in three independent SCLC-A cell lines. (**A–I**) RNA sequencing in NCI-H209 (A, D, G), NCI-H1436 (B, E, H), and NCI-H1836 (C, F, I) on day 4 (A, D, G) or day 6 (B, C, E, F, H, I) following CRISPR-Cas9 KO of *POU2F1* or *ROSA26* (control). Fold change and significance calculated by DESeq2. Two independent experiments performed in each cell line and time point, each with six biological replicates. (A–F) GSEA of differentially expressed genes following *POU2F1* KO. Significance calculated by GSEA. NES = Normalized Enrichment score. FWER = Family-wise Error Rate. (G–I) Enrichr Gene Ontology analysis of top 250 downregulated genes following *POU2F1* KO. Top 10 DESCARTES Cell Type signatures are rank ordered by significance (*P*-value), and the most significant terms ($P < .05$) are highlighted.

KO varied across cell lines, these effects were consistent across biological replicates and supported that POU2F1 is required for expression of neuroendocrine lineage markers (Supplementary Fig. S3A–C). To assess whether these changes reflected a coordinated loss of neuroendocrine identity, we performed GSEA [44] using gene signatures associated with ASCL1-positive neuroendocrine SCLC (SCLC-A) or with normal human fetal lung neuroendocrine cells derived from the Developmental Single Cell Atlas of Gene Regulation and Expression (DESCARTES) database [45]. Both signatures showed a significant reduction following POU2F1 KO (Fig. 3A–F). Unbiased Gene Ontology analysis further identified "Lung Neuroendocrine genes" as the most significantly enriched category among genes downregulated upon POU2F1 KO (Fig. 3G–I) [46]. A parallel analysis of *ASCL1*-KO transcriptomes in these same cell line models yielded similar results (Supplementary Fig. S3D–O and Supplementary Table S4). Together, these findings indicate that POU2F1 and ASCL1 each contribute to sustaining the neuroendocrine transcriptional program of SCLC cells.

POU2F1 and ASCL1 occupy *cis*-regulatory elements at neuroendocrine genes

We next asked whether POU2F1 directly regulates *DLL3* and other neuroendocrine genes and whether its activity is integrated with ASCL1 at the level of chromatin occupancy. Using two independent antibodies validated for chromatin immunoprecipitation, we performed POU2F1 ChIP-seq in NCI-H209 cells. Although the total number of high-confidence POU2F1 peaks was relatively modest (fewer than 700), the data sets showed concordance between antibodies and robust enrichment of the known POU2F1 DNA recognition motif (Fig. 4A–C; Supplementary Table S5). Motif enrichment analysis [49] further identified the ASCL1 recognition sequence as significantly enriched within POU2F1-occupied regions (Fig. 4C and Supplementary Table S5). This observation prompted us to perform ASCL1 ChIP-seq in the same cell line, which confirmed overlapping occupancy of POU2F1 and ASCL1 at a subset of *cis*-regulatory elements (Figs 4B and D; Supplementary Fig. S4A–C). Notably, both transcription factors occupied the *DLL3* promoter (Fig. 4D). Recip-

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

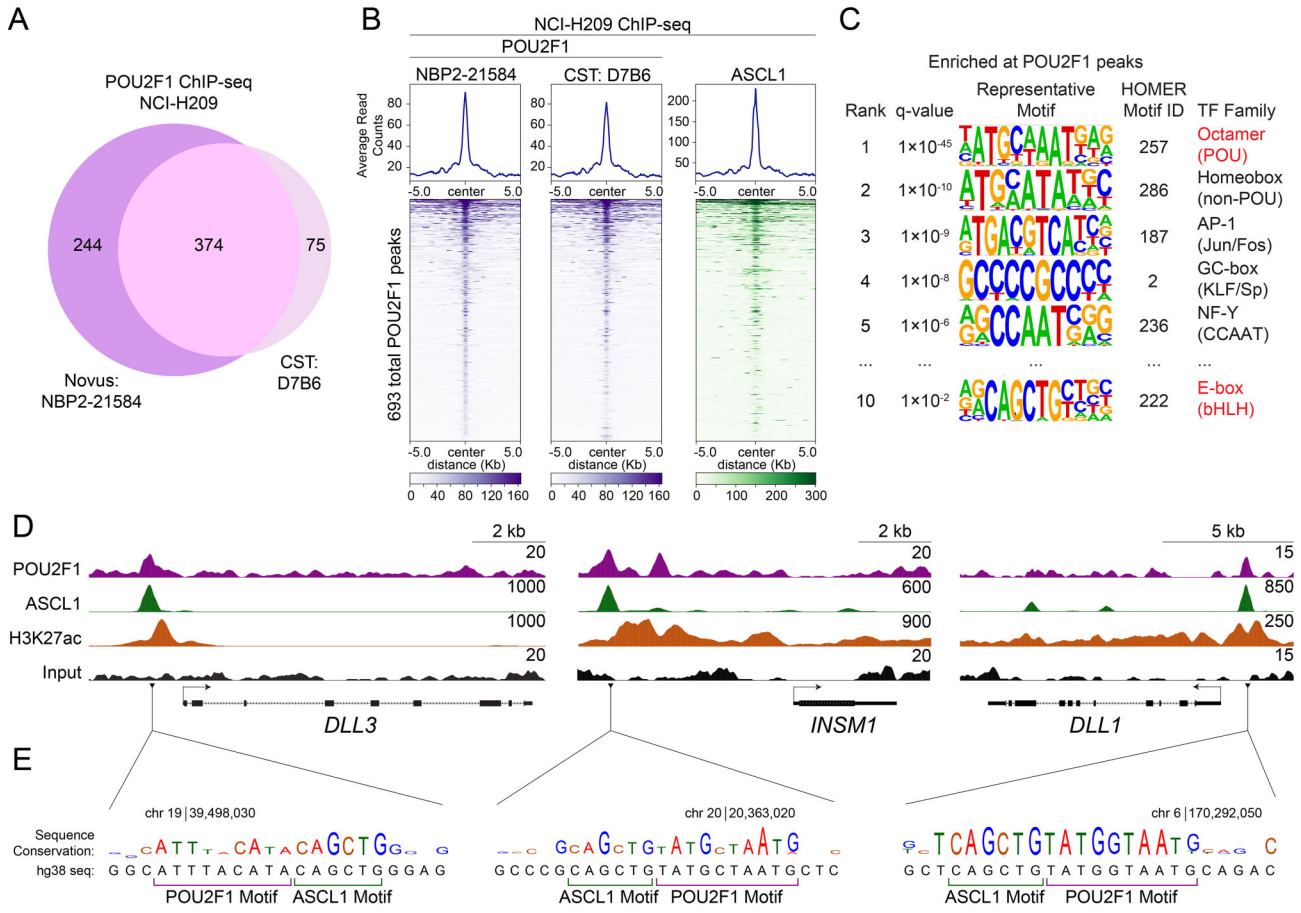


Figure 4. POU2F1 and ASCL1 occupy chromatin at tandem POU2F1–ASCL1 motifs. (**A**) Venn diagram of 693 total POU2F1 peaks (MACS2 $q < 0.01$) in NCI-H209 cells, including IP replicates using two independent antibodies (denoted). (**B**) Heatmaps for POU2F1 and ASCL1 ChIP-seq at all POU2F1 peaks. Rows = 10 kb genomic regions centered on a POU2F1 peak summit. Rows are ordered by POU2F1 signal, and this ordering is applied to both heatmaps. Metagene plots show the average signal for each factor across all peaks, plotted above each heatmap. (**C**) HOMER motif enrichment analysis of 693 called POU2F1 peaks (MACS2 $q < 0.01$). The top five transcription factor family motifs were selected. (**D**) ChIP-seq tracks in NCI-H209 showing ASCL1, POU2F1, H3K27ac, and input enrichment at representative loci, visualized in the UCSC genome browser. Matched, scaled track heights are listed (right). (**E**) hg38 sequence (bottom) and sequence conservation across 100 vertebrates (top) at the locus of ASCL1/POU2F1 peak summits adjacent to *DLL3*, *INSM1*, and *DLL1*. Base height indicates conservation across 100 vertebrates in UCSC genome browser. Genomic location (above) and locations of POU2F1 and ASCL1 motifs (below) are labeled.

rocal motif analysis of ASCL1 peaks revealed enrichment of the POU2F1 recognition motif (Supplementary Fig. S4D and Supplementary Table S5). Consistent with this observation, the *DLL3* promoter contains tandem POU2F1 and ASCL1 motifs (Fig. 4E). A similar pattern of occupancy was observed at a regulatory element located ∼6 kilobases upstream of the *INSM1* gene, again in association with tandem POU2F1 and ASCL1 motifs (Fig. 4D and E).

Tandem POU2F1–ASCL1 motifs contribute to a neuroendocrine identity *cis*-regulatory code

These findings led us to hypothesize that tandem POU2F1–ASCL1 motifs represent a feature of *cis*-regulatory elements that control neuroendocrine identity. To test this idea, we constructed a custom PWM for tandem POU2F1–ASCL1 motifs and scanned the human genome (hg38) for occurrences of this composite element using the FIMO algorithm (Fig. 5A) [54]. This analysis identified 882 instances of tandem motifs in the human genome (Supplementary Table S6). Most tandem motifs were located at distal intergenic or intragenic regions rather than promoters, with *DLL3* representing a notable exception (Supplementary Fig. S5A). Using HOMER, we annotated genes located near each tandem motif, identifying 227 genes within 20 kilobases of a tandem POU2F1–ASCL1 motif [49]. Remarkably, Gene Ontology analysis of these 227 genes revealed "Lung Neuroendocrine" as the top enriched category associated with the presence of the nearby tandem motif (Fig. 5B) [46]. Importantly, this enrichment was substantially stronger for the tandem motif than for either the POU2F1 or ASCL1 motif alone, and this effect was orientation-dependent, indicating a unique regulatory association of the composite element (Fig. 5B–D, Supplementary Fig. S5B, and Supplementary Table S6).

We next hypothesized that a tandem motif architecture might permit DNA-dependent interactions between POU2F1 and ASCL1. Because ASCL1 functions as a heterodimer with E proteins such as TCF12 [63], we used multiple AlphaFold 3 predictions to generate structural models of the ASCL1–TCF12 heterodimer in complex with POU2F1 and a 16 bp dsDNA fragment containing a consensus tandem POU2F1 and ASCL1 motifs (5′-**ATTTGCATA***CAGCTGT*-3′ and complement, Fig. 6A) [55]. Across multiple predictions,

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

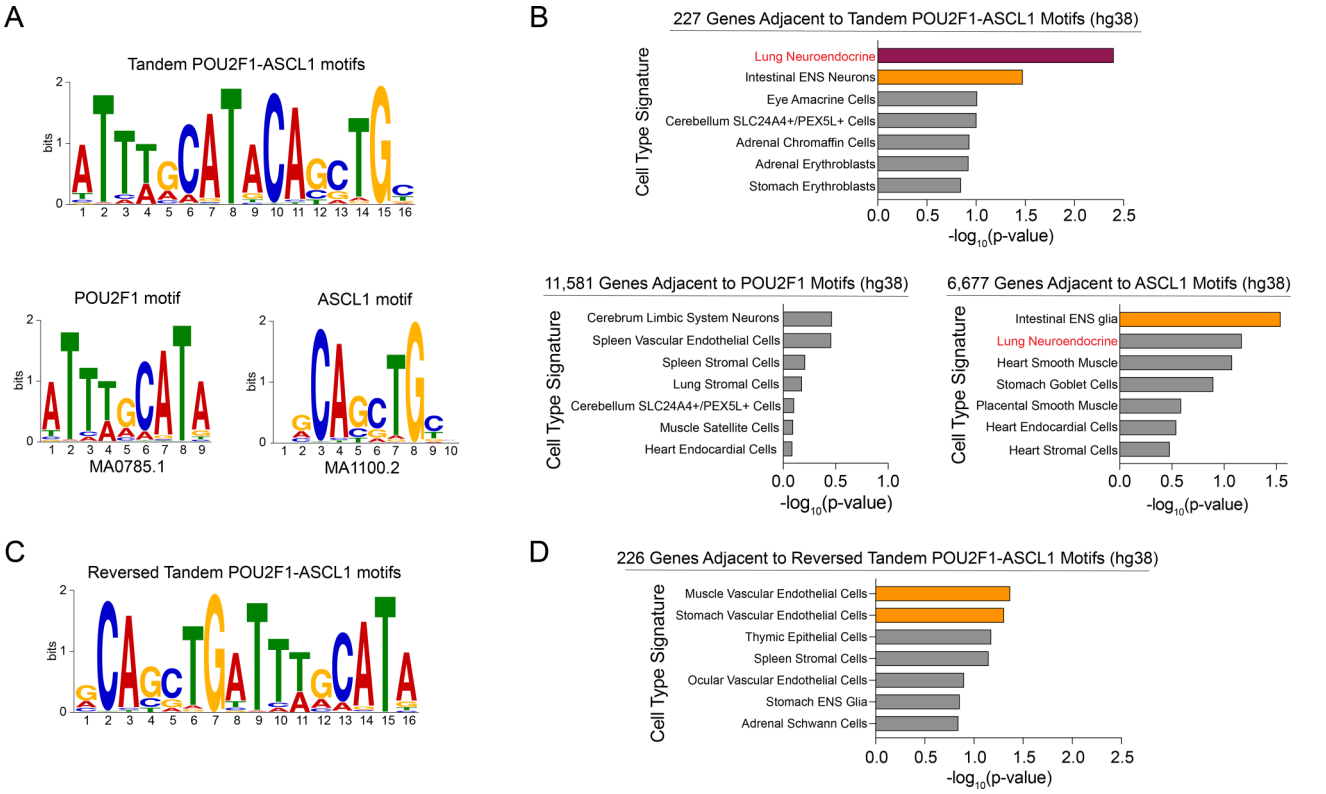


Figure 5. Tandem POU2F1–ASCL1 motifs represent a lung neuroendocrine identity-defining *cis*-regulatory grammar. (**A**, **C**) Schematic of the ASCL1, POU2F1, tandem POU2F1–ASCL1, and reversed tandem POU2F1–ASCL1 motifs. Position weight matrices generated using Tomtom (MEME suite). Height (bits) represents the conservation of the annotated base pair at that position. JASPAR Motif IDs are labeled below. (**B**, **D**) Enrichr Gene Ontology analysis of annotated genes within 20 kb of a tandem POU2F1–ASCL1 ($P < 1 \times 10^{-7}$), reversed tandem POU2F1–ASCL1 ($P < 1 \times 10^{-7}$), POU2F1 ($P < 1 \times 10^{-5}$), or ASCL1 ($P < 1 \times 10^{-5}$) motif (hg38). Top 7 DESCARTES Cell Type signatures are rank ordered by their significance (*P*-value), and the significant terms ($P < .05$) are highlighted.

the structured regions of all three proteins were positioned with high confidence relative to the DNA, whereas the relative placement of the proteins with respect to one another showed greater variability (Figs 6A–C and Supplementary Fig. S6A–C). One recurrent higher-confidence feature was a predicted interaction between a charged cluster at the N-terminus of the TCF12 bHLH domain and a basic patch in POU2F1 encompassing residues 377–382, located adjacent to the POU2F1 DNA motif (Fig. 6C and D; Supplementary Fig. S6A and B). In addition, a short region of ASCL1 immediately N-terminal to its bHLH domain was predicted to contact POU2F1 in two distinct conformations, albeit with lower confidence than the charged interaction (Fig. 6C and E; Supplementary Fig. S6A and B). Both the overall model confidence and the relative confidence in protein positioning were markedly reduced when the tandem motif orientation was reversed (5′-*CAGCTG*T**ATTTGCATA**-3′ and complement) or when DNA was omitted from the prediction, consistent with a DNA-stabilized interaction between POU2F1 and the ASCL1–TCF12 complex (Fig. 6B and C; Supplementary Fig. S6A–D). Together, these findings are consistent with a model in which tandem POU2F1 and ASCL1 motifs facilitate DNA-dependent interactions between POU2F1 and the ASCL1–TCF12 complex.

We next evaluated whether the tandem POU2F1–ASCL1 motifs are relevant for transcriptional activation. Using transient co-transfection in HEK293T cells, we found that epitope-tagged POU2F1 and ASCL1 co-immunoprecipitate in cellular lysates, supporting an association between the proteins and consistent with the association modeled in the AlphaFold 3 predictions (Fig. 7A). To evaluate the *cis*-regulatory requirements for this cooperativity, we cloned the *DLL3* promoter sequence upstream of a firefly luciferase reporter and assessed its induction by co-expressed POU2F1 and ASCL1. As expected, co-expression of both factors led to robust induction of luciferase activity (Fig. 7B). Targeted disruption of either the POU2F1 or the ASCL1 motif within the tandem POU2F1–ASCL1 motif or reversal of the POU2F1–ASCL1 orientation abrogated the inducible activation of this promoter (Fig. 7B). These findings suggest that the *cis*-regulatory function of the POU2F1–ASCL1 element is both motif- and orientation-dependent.

Finally, we assessed whether genes linked to tandem POU2F1–ASCL1 motifs were functionally dependent on these transcription factors. GSEA demonstrated significant down-regulation of tandem-associated genes following *POU2F1* KO across all three SCLC cell lines (Fig. 7C). This pattern was not recapitulated when examining genes associated with POU2F1 motifs alone (Supplementary Fig. S7A). Similar results were observed following *ASCL1* inactivation in the same cell lines (Supplementary Fig. S7B and C). Taken together, these data support a model in which tandem POU2F1–ASCL1 motifs mark critical *cis*-regulatory elements that integrate the activities of both transcription factors to maintain neuroendocrine identity in SCLC, and potentially in normal neuroendocrine cellular contexts.

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

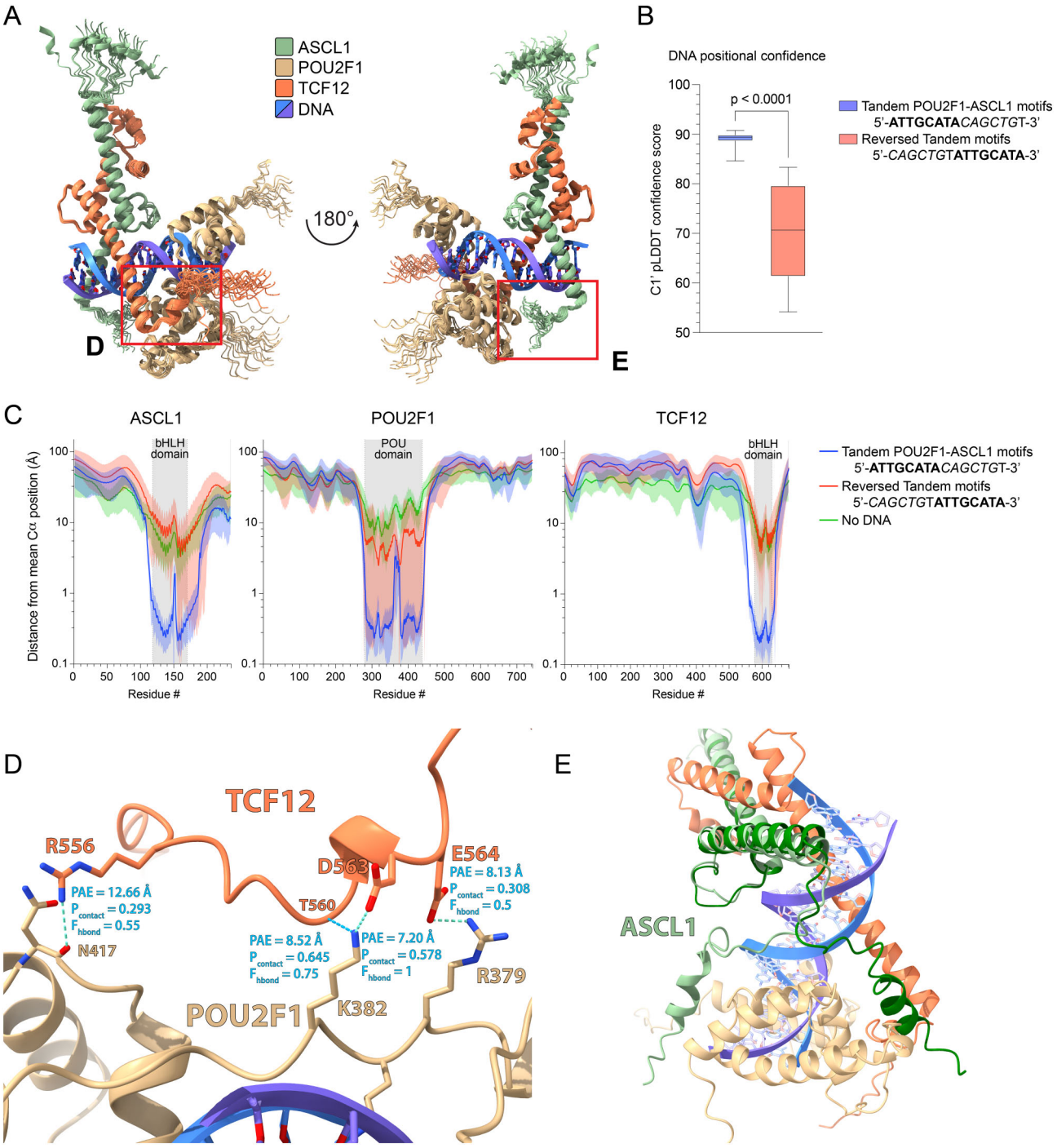


Figure 6. AlphaFold 3 supports an orientation-specific tandem POU2F1–ASCL1 motif. (**A**) Cartoon schematic of 20 (four starting seeds, five models each) aligned AlphaFold 3 predictions of the tandem motif bound to the ASCL1–POU2F1–TCF12 complex. (**B**) The C1′ carbon is the carbon atom within the deoxyribose sugar of DNA that forms the covalent bond between the nitrogenous base and the sugar backbone. Higher pLDDT scores at DNA nucleotide positions indicate greater confidence in the modeled DNA position within the complex. Confidence scores (pLDDT) in the placement of the DNA nucleotides ($N = 32$ each) in AlphaFold 3 models containing the ASCL1–POU2F1 tandem motif or a "reversed" motif are plotted. Each data point represents the mean pLDDT score at a given nucleotide position in the DNA sequence. Boxes represent the median ± IQR. Error bars extend to the minimum and maximum values. Statistical significance was evaluated using Welch's unpaired *t*-test. Bold letters were used to represent the POU2F1 binding motif, while *italics* indicate the ASCL1 binding motif. (**C**) The Cα carbon is the central backbone carbon of each amino acid, bonded to the amino, carboxyl, and side chain groups. Smaller distances from the mean aligned position indicate greater model-to-model consistency across AlphaFold 3 predictions. The distance of each residue's Cα atom from the mean aligned position of predictions containing the tandem motif sequence, a "reversed" tandem motif, or models that did not include DNA are plotted. Central line represents the mean distance from the aligned Cα, with the standard deviation of the mean residue distance surrounding it. (**D**) The predicted POU2F1–TCF12 interface as seen in a representative AlphaFold 3 prediction. PAE and contact probability values are represented by the mean of those values over all twenty models in both directions. F_{hbond} represents the fraction of AlphaFold 3 predictions in which a hydrogen bond was projected to occur using relaxed constraints in ChimeraX. (**E**) The N-terminal region of ASCL1, as shown in two representative AlphaFold 3 predictions, fell into one of two distinct conformations in most predictions.

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

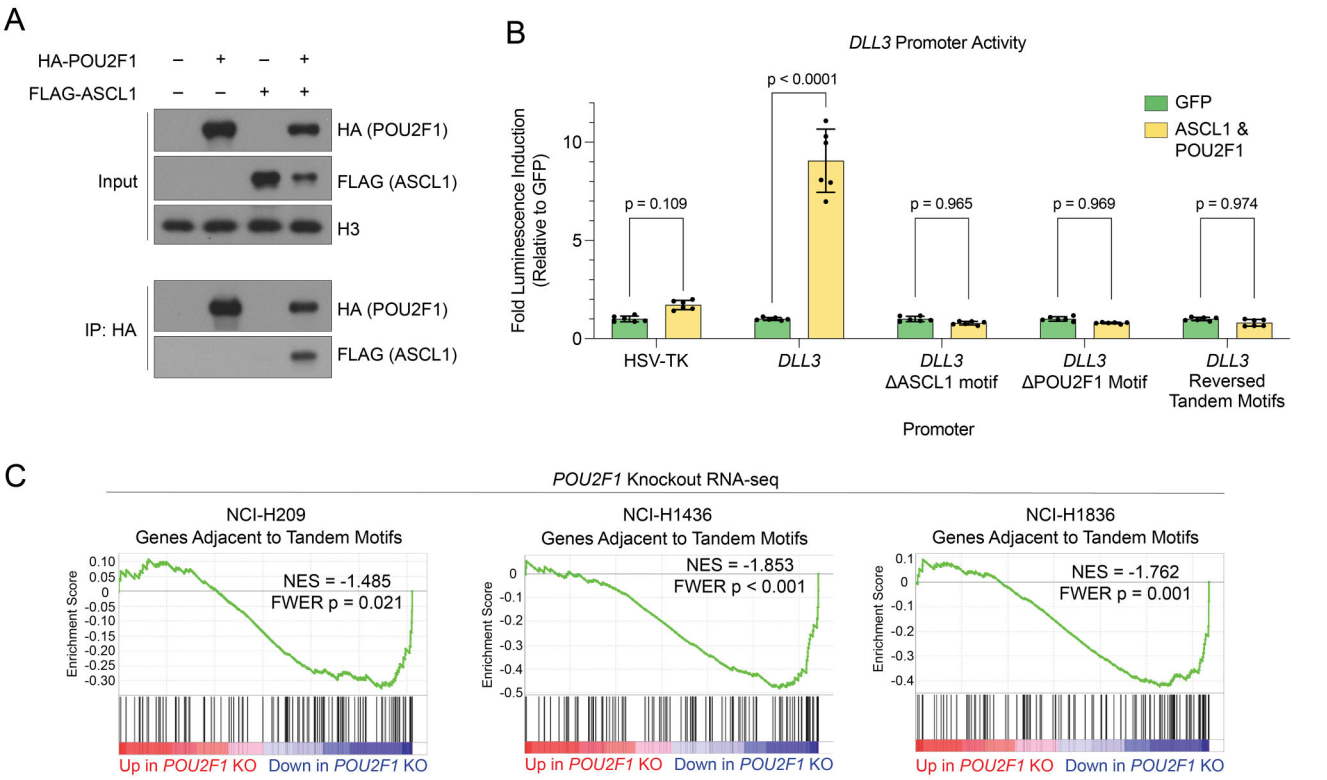


Figure 7. POU2F1 and ASCL1 cooperate to activate DLL3 and the neuroendocrine transcriptional program. (**A**) Anti-HA co-immunoprecipitation–western blot. 0.8% input loaded. H3 = loading control. Representative of two biological replicates. (**B**) DLL3 promoter luciferase activity measured in the presence of reporter plasmids with wild-type or mutant POU2F1–ASCL1 tandem motif and in the presence of ASCL1 and POU2F1 or GFP (control) in HEK293T cells. Luciferase activity is reported as fold change in the ASCL1 and POU2F1 condition relative to control. Two independent experiments, each containing three replicate transfections. Data represent mean ± SD. Significance calculated by two-way analysis of variance with Šidák's multiple-comparisons test. HSV-TK = control promoter. (**C**) RNA sequencing in NCI-H209, NCI-H1436, and NCI-H1836 following CRISPR-Cas9 KO of *POU2F1* or *ROSA26* (control). Fold change and significance calculated by DESeq2. Two independent experiments, each with six biological replicates. GSEA of differentially expressed genes following *POU2F1* KO. NES = Normalized Enrichment score. FWER = Family-wise Error Rate.

Discussion

This study indicates that DLL3 expression in SCLC is governed by a defined *cis*-regulatory logic rather than serving as a passive marker of neuroendocrine identity. While DLL3 has been widely used as a biomarker and therapeutic target, its regulation has largely been inferred indirectly through associations with ASCL1-positive tumors and suppressed Notch signaling [9, 57, 64]. By applying unbiased CRISPR screening approaches, we directly map the transcriptional requirements that sustain DLL3 expression and demonstrate that high-level DLL3 transcription requires coordinated activity of multiple transcription factors. These findings position DLL3 as an encoded output of a lineage-specific transcriptional program, with implications for understanding both its biological role and its behavior under therapeutic pressure.

A key insight from this work is that a ubiquitously expressed transcription factor, POU2F1, acquires a lineage-restricted function in SCLC through cooperative interactions with ASCL1 within a specific *cis*-regulatory architecture. POU2F1 has traditionally been viewed as a broadly acting transcriptional regulator rather than as a determinant of cell identity [30]. Our findings add to a growing body of evidence challenging this view by showing that POU2F1 contributes to neuroendocrine gene expression in SCLC when paired with ASCL1 at tandem DNA motifs [37–39]. One potential explanation for this context specificity is the pioneer factor activity of ASCL1. In multiple developmental and reprogramming systems, ASCL1 can bind nucleosome-occluded chromatin and induce chromatin remodeling, subsequently recruiting additional transcriptional regulators [65–67]. It is possible that, in SCLC, ASCL1 functions as the primary lineage-defining factor that establishes accessible chromatin at neuroendocrine *cis*-regulatory elements, in turn enabling binding of POU2F1 to tandem POU2F1–ASCL1 motifs, which leads to functional cooperativity via a DNA-dependent interaction between these two transcription factors. Consistent with this model, ASCL1 chromatin binding during induced fibroblast-to-neuronal reprogramming precedes recruitment of BRN2, a POU family transcription factor, which cannot productively access chromatin alone [68]. Following ASCL1 binding, BRN2 effectively binds neuronal *cis*-regulatory elements and activates transcription to support reprogramming. This observation supports a similar hierarchical model of chromatin engagement during SCLC progression in which ASCL1 initiates the global regulatory landscape, while POU2F1 reinforces transcriptional activation at a defined subset of neuroendocrine-specific loci. Consistent with a privileged relationship between ASCL1 and POU2F1, POU2F1 expression is markedly elevated in lung neuroendocrine cancer cell lines, relative to other cancer cell line lineages. POU2F1 protein levels are also reduced following ASCL1 loss despite relatively unchanged *POU2F1* mRNA expression, suggesting that ASCL1 may influence POU2F1 abundance through a post-transcriptional mechanism. Defining the

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

mechanism underlying this relationship will require further investigation.

These findings also have implications for understanding lineage plasticity in SCLC and how tumors might transition between neuroendocrine and non-neuroendocrine states. Prior work has shown that tuft cell-like SCLC is driven by a distinct transcriptional circuitry centered on POU2F3, often in conjunction with ASCL2, a paralog of ASCL1 [22, 24]. The parallels between POU2F1–ASCL1 cooperation in neuroendocrine SCLC and POU2F3–ASCL2 function in tuft cell-like carcinomas suggest a modular transcriptional system in which different POU and ASCL family members pair to specify alternative lineage programs. Such pairing could provide a mechanism for lineage switching, whereby changes in expression or activity of one component enable engagement of an alternative *cis*-regulatory code. Future investigations will define the molecular determinants that govern selective pairings between POU and ASCL family members and will determine whether lineage transitions reflect modified transcription factor paralog pairings, competition for access to shared *cis*-regulatory elements, or the activation of distinct enhancer networks. A more comprehensive characterization of how alternative POU–ASCL modules are assembled and reprogrammed during cancer progression and upon therapeutic pressure may reveal vulnerabilities that can be exploited to constrain lineage plasticity or to redirect tumor identity.

These findings also suggest potential, albeit still speculative, strategies for pharmacologically manipulating the POU2F1–ASCL1 interaction. Directly disrupting this interaction may offer limited therapeutic value: the broad expression of POU2F1 raises uncertainty regarding a therapeutic window, and destabilizing neuroendocrine identity could reduce DLL3 expression and thereby potentially promote resistance to DLL3-directed therapies. An alternative strategy might be to stabilize the POU2F1–ASCL1–E-protein complex, potentially through a molecular-glue approach, to reinforce DLL3 expression and reduce antigen heterogeneity. Such an approach could conceivably restore or maintain DLL3 expression in tumors that have become less responsive to DLL3-directed therapy. However, the therapeutic utility of this approach remains uncertain, as the relationship between DLL3 transcription and cell-surface antigen abundance has not been established, and pharmacologically reinforcing a neuroendocrine program could have unintended biological consequences. More broadly, these considerations underscore how lineage-specific transcriptional circuitry may influence the stability of tumor-associated antigens and the durability of antigen-directed therapies.

Acknowledgements

L.J. is an investigator of the Howard Hughes Medical Institute.

Author contributions: P.J.C. performed the formal analysis and wrote the initial draft of the manuscript. C.F. designed the methodology and carried out the investigation and validation studies, with assistance from D.S. in performing the marker-based CRISPR screens. J.B. performed the AlphaFold 3 analysis, and L.J. provided supervision and advice. S.P. performed the coimmunoprecipitation assays. L.S. performed the promoter reporter assay. K.T. performed the CellTiter-Glo analysis. O.K. and T.Y. subcloned the Brunello genome-wide sgRNA library. C.R.V. conceived the project, developed the methodology, supervised the research, acquired funding, and contributed to writing, editing, and review of the manuscript.

Supplementary data

Supplementary data is available at NAR Cancer online.

Conflict of interest

C.R.V. has received consulting fees from Flare Therapeutics, Roivant Sciences, and C4 Therapeutics; has served on the advisory boards of KSQ Therapeutics, Syros Pharmaceuticals, and Treeline Biosciences; has received research funding from Boehringer-Ingelheim and Treeline Biosciences; and owns stock in Treeline Biosciences. The remaining authors declare no competing interests.

Funding

This study was supported by the Cold Spring Harbor Laboratory NCI Cancer Center Support Grant (P30 CA045508). C.R.V. was supported by the Pershing Square Sohn Cancer Research Alliance and NCI grant CA290004. Additional funding was provided by Cold Spring Harbor Laboratory, Northwell Health Affiliation, and Treeline Biosciences. P.J.C. was supported by NCI grant CA278591. D.S. was supported by NCI R50 grant CA305054.

Data availability

All genomic datasets are available at the GEO database under accession codes GSE317998 (CRISPR screens), GSE317997 (RNA-seq), and GSE317996 (ChIP-seq). Raw uncropped western blots are published alongside this paper. ChIP-seq tracks are accessible on the UCSC genome browser at https://genome.ucsc.edu/s/pcunniff/DLL3_POU2F1_ASCL1_H209_hg38. Source data are available on Figshare at https://figshare.com/projects/Marker-based-CRISPR-screens-identify-POU2F1-as-a-regulator-of-DLL3-and-neuroendocrine-identity-in-small-cell-lung-cancer/280107.

References

1. Rudin CM, Brambilla E, Faivre-Finn C *et al.* Small-cell lung cancer. *Nat Rev Dis Primers* 2021;7:3. https://doi.org/10.1038/s41572-020-00235-0
2. Leube RE, Wiedenmann B, Franke WW. Differentiation markers as an aid in the histological diagnosis of small-cell carcinoma of the lung: synopsis of intermediate filament protein and synaptophysin expression. *Klin Wochenschr* 1988;66:80–6.
3. Linnoila RI, Mulshine JL, Steinberg SM *et al.* Neuroendocrine differentiation in endocrine and nonendocrine lung carcinomas. *Am J Clin Pathol* 1988;90:641–52. https://doi.org/10.1093/ajcp/90.6.641
4. Rudin CM, Poirier JT, Byers LA *et al.* Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. *Nat Rev Cancer* 2019;19:289–97. https://doi.org/10.1038/s41568-019-0133-9
5. Yao J, Bergsland E, Aggarwal R *et al.* DLL3 as an emerging target for the treatment of neuroendocrine neoplasms. *Oncologist* 2022;27:940–51. https://doi.org/10.1093/oncolo/oyac161
6. Geffers I, Serth K, Chapman G *et al.* Divergent functions and distinct localization of the Notch ligands DLL1 and DLL3 *in vivo*.

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

J Cell Biol 2007;178:465–76. https://doi.org/10.1083/jcb.200702009
7. Ladi E, Nichols JT, Ge W *et al.* The divergent DSL ligand Dll3 does not activate Notch signaling but cell autonomously attenuates signaling induced by other DSL ligands. *J Cell Biol* 2005;170:983–92. https://doi.org/10.1083/jcb.200503113
8. He P, Lim K, Sun D *et al.* A human fetal lung cell atlas uncovers proximal-distal gradients of differentiation and key regulators of epithelial fates. *Cell* 2022;185:4841–60. https://doi.org/10.1016/j.cell.2022.11.005
9. Saunders LR, Bankovich AJ, Anderson WC *et al.* A DLL3-targeted antibody-drug conjugate eradicates high-grade pulmonary neuroendocrine tumor-initiating cells *in vivo*. *Sci Transl Med* 2015;7:302ra136. https://doi.org/10.1126/scitranslmed.aac9459
10. Jaspers JE, Khan JF, Godfrey WD *et al.* IL-18-secreting CAR T cells targeting DLL3 are highly effective in small cell lung cancer models. *J Clin Invest* 2023;133:e166028. https://doi.org/10.1172/JCI166028
11. Morgensztern D, Johnson M, Rudin CM *et al.* SC-002 in patients with relapsed or refractory small cell lung cancer and large cell neuroendocrine carcinoma: phase 1 study. *Lung Cancer* 2020;145:126–31. https://doi.org/10.1016/j.lungcan.2020.04.017
12. Paz-Ares L, Champiat S, Lai WV *et al.* Tarlatamab, a first-in-class DLL3-targeted bispecific T-cell engager, in recurrent small-cell lung cancer: an open-label, phase I study. *J Clin Oncol* 2023;41:2893–903. https://doi.org/10.1200/JCO.22.02823
13. Rudin CM, Pietanza MC, Bauer TM *et al.* Rovalpituzumab tesirine, a DLL3-targeted antibody-drug conjugate, in recurrent small-cell lung cancer: a first-in-human, first-in-class, open-label, phase 1 study. *Lancet Oncol* 2017;18:42–51. https://doi.org/10.1016/S1470-2045(16)30565-4
14. Tully KM, Tendler S, Carter LM *et al.* Radioimmunotherapy targeting delta-like ligand 3 in small cell lung cancer exhibits antitumor efficacy with low toxicity. *Clin Cancer Res* 2022;28:1391–401. https://doi.org/10.1158/1078-0432.CCR-21-1533
15. Wermke M, Felip E, Gambardella V *et al.* Phase I trial of the DLL3/CD3 bispecific T-cell engager BI 764532 in DLL3-positive small-cell lung cancer and neuroendocrine carcinomas. *Future Oncol* 2022;18:2639–49. https://doi.org/10.2217/fon-2022-0196
16. Zhang Y, Tacheva-Grigorova SK, Sutton J *et al.* Allogeneic CAR T cells targeting DLL3 are efficacious and safe in preclinical models of small cell lung cancer. *Clin Cancer Res* 2023;29:971–85. https://doi.org/10.1158/1078-0432.CCR-22-2293
17. Kim JW, Ko JH, Sage J. DLL3 regulates Notch signaling in small cell lung cancer. *iScience* 2022;25:105603. https://doi.org/10.1016/j.isci.2022.105603
18. Chan JM, Quintanal-Villalonga Á, Gao VR *et al.* Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. *Cancer Cell* 2021;39:1479–96. https://doi.org/10.1016/j.ccell.2021.09.008
19. Groves SM, Ildefonso GV, McAtee CO *et al.* Archetype tasks link intratumoral heterogeneity to plasticity and cancer hallmarks in small cell lung cancer. *Cell Syst* 2022;13:690–710. https://doi.org/10.1016/j.cels.2022.07.006
20. Bunn PA, Minna JD, Augustyn A *et al.* Small cell lung cancer: can recent advances in biology and molecular biology be translated into improved outcomes? *J Thorac Oncol* 2016;11:453–74. https://doi.org/10.1016/j.jtho.2016.01.012
21. Duplaquet L, So K, Ying AW *et al.* Mammalian SWI/SNF complex activity regulates POU2F3 and constitutes a targetable dependency in small cell lung cancer. *Cancer Cell* 2024;42:1352–69. https://doi.org/10.1016/j.ccell.2024.06.012
22. Huang YH, Klingbeil O, He XY *et al.* POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer. *Genes Dev* 2018;32:915–28. https://doi.org/10.1101/gad.314815.118
23. Tsherniak A, Vazquez F, Montgomery PG *et al.* Defining a cancer dependency map. *Cell* 2017;170:564–76. https://doi.org/10.1016/j.cell.2017.06.010
24. Wu XS, He XY, Ipsaro JJ *et al.* OCA-T1 and OCA-T2 are coactivators of POU2F3 in the tuft cell lineage. *Nature* 2022;607:169–75. https://doi.org/10.1038/s41586-022-04842-7
25. Brumage L, Best S, Hippe DS *et al.* *In vivo* functional screens reveal *KEAP1* loss as a driver of chemoresistance in small cell lung cancer. *Sci Adv* 2025;11:eadq7084. https://doi.org/10.1126/sciadv.adq7084
26. Duplaquet L, Li Y, Booker MA *et al.* KDM6A epigenetically regulates subtype plasticity in small cell lung cancer. *Nat Cell Biol* 2023;25:1346–58. https://doi.org/10.1038/s41556-023-01210-z
27. Li L, Ng SR, Colón CI *et al.* Identification of DHODH as a therapeutic target in small cell lung cancer. *Sci Transl Med* 2019;11:eaaw7852. https://doi.org/10.1126/scitranslmed.aaw7852
28. Préfontaine GG, Lemieux ME, Giffin W *et al.* Recruitment of octamer transcription factors to DNA by glucocorticoid receptor. *Mol Cell Biol* 1998;18:3416–30. https://doi.org/10.1128/MCB.18.6.3416
29. Sturm RA, Das G, Herr W. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes Dev* 1988;2:1582–99. https://doi.org/10.1101/gad.2.12a.1582
30. Zwilling S, Annweiler A, Wirth T. The POU domains of the Oct1 and Oct2 transcription factors mediate specific interaction with TBP. *Nucl Acids Res* 1994;22:1655–62. https://doi.org/10.1093/nar/22.9.1655
31. Shakya A, Cooksey R, Cox JE *et al.* Oct1 loss of function induces a coordinate metabolic shift that opposes tumorigenicity. *Nat Cell Biol* 2009;11:320–7. https://doi.org/10.1038/ncb1840
32. Tantin D, Schild-Poulter C, Wang V *et al.* The octamer binding transcription factor Oct-1 is a stress sensor. *Cancer Res* 2005;65:10750–8. https://doi.org/10.1158/0008-5472.CAN-05-2399
33. Kang J, Gemberling M, Nakamura M *et al.* A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative stress. *Genes Dev* 2009;23:208–22. https://doi.org/10.1101/gad.1750709
34. Shakya A, Goren A, Shalek A *et al.* Oct1 and OCA-B are selectively required for CD4 memory T cell function. *J Exp Med* 2015;212:2115–31. https://doi.org/10.1084/jem.20150363
35. Lin J, Xia L, Oyang L *et al.* The POU2F1-ALDOA axis promotes the proliferation and chemoresistance of colon cancer cells by enhancing glycolysis and the pentose phosphate pathway activity. *Oncogene* 2022;41:1024–39. https://doi.org/10.1038/s41388-021-02148-y
36. Stepchenko AG, Portseva TN, Glukhov IA *et al.* Primate-specific stress-induced transcription factor POU2F1Z protects human neuronal cells from stress. *Sci Rep* 2021;11:18808. https://doi.org/10.1038/s41598-021-98323-y
37. Maddox J, Shakya A, South S *et al.* Transcription factor Oct1 is a somatic and cancer stem cell determinant. *PLoS Genet* 2012;8:e1003048. https://doi.org/10.1371/journal.pgen.1003048
38. Perovanovic J, Wu Y, Abewe H *et al.* Oct1 cooperates with the Smad family of transcription factors to promote mesodermal lineage specification. *Sci Signal* 2023;16:eadd5750. https://doi.org/10.1126/scisignal.add5750
39. Shen Z, Kang J, Shakya A *et al.* Enforcement of developmental lineage specificity by transcription factor Oct1. *eLife* 2017;6:e20937. https://doi.org/10.7554/eLife.20937
40. Doench JG, Fusi N, Sullender M *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 2016;34:184–91. https://doi.org/10.1038/nbt.3437
41. Li W, Xu H, Xiao T *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

screens. *Genome Biol* 2014;15:554. https://doi.org/10.1186/s13059-014-0554-4
42. Bray NL, Pimentel H, Melsted P *et al.* Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016;34:525–7. https://doi.org/10.1038/nbt.3519
43. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550. https://doi.org/10.1186/s13059-014-0550-8
44. Subramanian A, Tamayo P, Mootha VK *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;102:15545–50. https://doi.org/10.1073/pnas.0506580102
45. Cao J, O'Day DR, Pliner HA *et al.* A human cell atlas of fetal gene expression. *Science* 2020;370:eaba7721. https://doi.org/10.1126/science.aba7721
46. Chen EY, Tan CM, Kou Y *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 2013;14:128. https://doi.org/10.1186/1471-2105-14-128
47. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–9. https://doi.org/10.1038/nmeth.1923
48. Zhang Y, Liu T, Meyer CA *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;9:R137. https://doi.org/10.1186/gb-2008-9-9-r137
49. Heinz S, Benner C, Spann N *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. https://doi.org/10.1016/j.molcel.2010.05.004
50. Ramírez F, Dündar F, Diehl S *et al.* deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 2014;42:W187–91. https://doi.org/10.1093/nar/gku365
51. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841–2. https://doi.org/10.1093/bioinformatics/btq033
52. Amemiya HM, Kundaje A, Boyle AP. The ENCODE Blacklist: identification of problematic regions of the genome. *Sci Rep* 2019;9:9354. https://doi.org/10.1038/s41598-019-45839-z
53. Gupta S, Stamatoyannopoulos JA, Bailey TL *et al.* Quantifying similarity between motifs. *Genome Biol* 2007;8:R24. https://doi.org/10.1186/gb-2007-8-2-r24
54. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. *Bioinformatics* 2011;27:1017–8. https://doi.org/10.1093/bioinformatics/btr064
55. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w
56. Meng EC, Goddard TD, Pettersen EF *et al.* UCSF ChimeraX: tools for structure building and analysis. *Protein Sci* 2023;32:e4792. https://doi.org/10.1002/pro.4792
57. Borromeo MD, Savage TK, Kollipara RK *et al.* ASCL1 and NEUROD1 reveal heterogeneity in pulmonary neuroendocrine tumors and regulate distinct genetic programs. *Cell Rep* 2016;16:1259–72. https://doi.org/10.1016/j.celrep.2016.06.081
58. Henke RM, Meredith DM, Borromeo MD *et al.* Ascl1 and Neurog2 form novel complexes and regulate Delta-like3 (Dll3) expression in the neural tube. *Dev Biol* 2009;328:529–40. https://doi.org/10.1016/j.ydbio.2009.01.007
59. Cunniff PJ, Sivetz N, Skopelitis D *et al.* KLF5 enables dichotomous lineage programs in pancreatic cancer via the AAA+ ATPase coactivators RUVBL1 and RUVBL2. *Nat Commun* 2025;16:9996. https://doi.org/10.1038/s41467-025-66007-0
60. Barretina J, Caponigro G, Stransky N *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012;483:603–7. https://doi.org/10.1038/nature11003
61. George J, Lim JS, Jang SJ *et al.* Comprehensive genomic profiles of small cell lung cancer. *Nature* 2015;524:47–53. https://doi.org/10.1038/nature14664
62. Rooper LM, Sharma R, Li QK *et al.* INSM1 demonstrates superior performance to the individual and combined use of synaptophysin, chromogranin and CD56 for diagnosing neuroendocrine tumors of the thoracic cavity. *Am J Surg Pathol* 2017;41:1561–9. https://doi.org/10.1097/PAS.0000000000000916
63. Johnson JE, Birren SJ, Saito T *et al.* DNA binding and transcriptional regulatory activity of mammalian achaete-scute homologous (MASH) proteins revealed by interaction with a muscle-specific enhancer. *Proc Natl Acad Sci USA* 1992;89:3596–600. https://doi.org/10.1073/pnas.89.8.3596
64. Lim JS, Ibaseta A, Fischer MM *et al.* Intratumoural heterogeneity generated by Notch signalling promotes small-cell lung cancer. *Nature* 2017;545:360–4. https://doi.org/10.1038/nature22323
65. Marro S, Pang ZP, Yang N *et al.* Direct lineage conversion of terminally differentiated hepatocytes to functional neurons. *Cell Stem Cell* 2011;9:374–82. https://doi.org/10.1016/j.stem.2011.09.002
66. Păun O, Tan YX, Patel H *et al.* Pioneer factor ASCL1 cooperates with the mSWI/SNF complex at distal regulatory elements to regulate human neural differentiation. *Genes Dev* 2023;37:218–42. https://doi.org/10.1101/gad.350269.122
67. Vierbuchen T, Ostermeier A, Pang ZP *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 2010;463:1035–41. https://doi.org/10.1038/nature08797
68. Wapinski OL, Vierbuchen T, Qu K *et al.* Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* 2013;155:621–35. https://doi.org/10.1016/j.cell.2013.09.028

Downloaded from academic.oup.com/narcancer/article/8/4/zcag032/9089570 by Cold Spring Harbor Laboratory user on 20 November 2026

Received: April 18, 2026. **Revised:** August 27, 2026. **Accepted:** September 26, 2026
© The Author(s) 2026. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.