HHS Public Access

Author manuscript

*Cancer Cell*. Author manuscript; available in PMC 2021 July 13.

Published in final edited form as:

*Cancer Cell*. 2020 July 13; 38(1): 60–78.e12. doi:10.1016/j.ccell.2020.05.001.

# MYC drives temporal evolution of small cell lung cancer subtypes by reprogramming neuroendocrine fate

**Abbie S. Ireland**[1], **Alexi M. Micinski**[1], **David W. Kastner**[1], **Bingqian Guo**[1], **Sarah J. Wait**[1], **Kyle B. Spainhower**[1], **Christopher C. Conley**[2], **Opal S. Chen**[3], **Matthew R. Guthrie**[1], **Danny Soltero**[1], **Yi Qiao**[4], **Xiaomeng Huang**[4], **Szabolcs Tarapcsak**[4], **Siddhartha Devarakonda**[5], **Milind D. Chalishazar**[1], **Jason Gertz**[1], **Justin C. Moser**[6], **Gabor Marth**[4], **Sonam Puri**[7], **Benjamin L. Witt**[8,9], **Benjamin T. Spike**[1], **Trudy G. Oliver**[1,10,*]

[1]Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA

[2]Huntsman Cancer Institute Bioinformatic Analysis Shared Resource, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA

[3]Huntsman Cancer Institute High-Throughput Genomics Shared Resource, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA

[4]Utah Center for Genetic Discovery, Eccles Institute of Human Genetics, University of Utah, Salt Lake City, UT 84112, USA

[5]Division of Medical Oncology, Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA

[6]HonorHealth Research Institute, Scottsdale, AZ 85254, USA

[7]Department of Internal Medicine, University of Utah, Salt Lake City, UT 84112, USA

[8]Department of Pathology, University of Utah, Salt Lake City, UT 84112, USA

[9]ARUP Laboratories at University of Utah, Salt Lake City, UT 84108, USA

[10]Lead contact

## SUMMARY

[*]Correspondence: trudy.oliver@hci.utah.edu (T.G.O.).

AUTHOR CONTRIBUTIONS

Study design: A.S.I., T.G.O.

Data analysis and acquisition: A.S.I., A.M.M., D.W.K., O.S.C., S.J.W., K.B.S., M.R.G., D.S., B.G., M.D.C., J.G.

Bioinformatic analyses: A.S.I., C.C.C., Y.Q., X.H., S.T., S.D., G.M., B.T.S., J.G.

Pathology: B.L.W.

Clinical annotation: S.P., J.C.M.

Study supervision and funding: T.G.O.

Manuscript preparation: A.S.I., B.T.S., T.G.O.

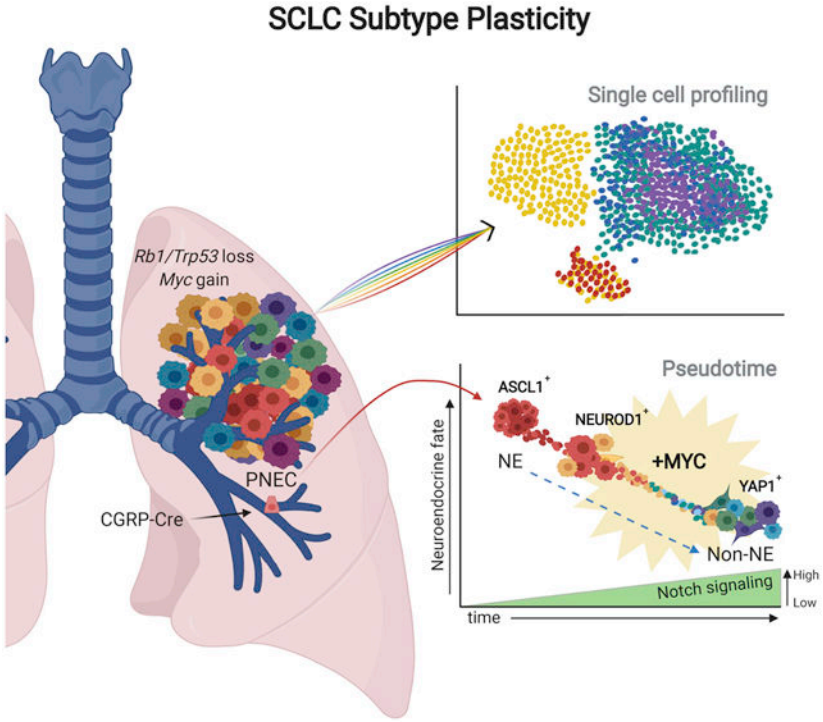


DECLARATION OF INTERESTS

TGO has pending patent applications related to subtype stratification of SCLC: US16/335368; JP2019522392, EP2017865057.

Small cell lung cancer (SCLC) is a neuroendocrine tumor treated clinically as a single disease with poor outcomes. Distinct SCLC molecular subtypes have been defined based on expression of *ASCL1, NEUROD1, POU2F3* or *YAP1*. Here, we use mouse and human models with a time-series single-cell transcriptome analysis to reveal that MYC drives dynamic evolution of SCLC subtypes. In neuroendocrine cells, MYC activates Notch to dedifferentiate tumor cells, promoting a temporal shift in SCLC from ASCL1$^+$ to NEUROD1$^+$ to YAP1$^+$ states. MYC alternatively promotes POU2F3$^+$ tumors from a distinct cell type. Human SCLC exhibits intratumoral subtype heterogeneity, suggesting this dynamic evolution occurs in patient tumors. These findings suggest that genetics, cell of origin, and tumor cell plasticity determine SCLC subtype.

## Graphical Abstract




## eTOC blurb

Ireland et al. show that MYC activates Notch signaling to dedifferentiate neuroendocrine small cell lung cancer (SCLC) in a conserved trajectory from ASCL1$^+$ to NEUROD1$^+$ to YAP1$^+$ non-neuroendocrine subtypes, suggesting these are not distinct subtypes but different stages of progressive evolution of SCLC.

## INTRODUCTION

Understanding molecular heterogeneity in cancer is critical for precision medicine to tailor cancer treatments to the specific features of a patient's disease. For example, lung adenocarcinoma comprises genetic subtypes with distinct, mutually-exclusive alterations in genes such as *EGFR, KRAS* or *ALK*, and targeting *EGFR*- or *ALK*-mutant tumors with targeted inhibitors prolongs survival and improves patient outcome (Collisson et al., 2014;

Lin et al., 2016; Remon et al., 2019). In mouse models of other tumor types such as prostate cancer, the same oncogenes can promote different tumor subtypes (i.e. luminal, basal) based on the initiating cell of origin (Park et al., 2016; Wang et al., 2013). The childhood cerebellar tumor medulloblastoma comprises genetic subtypes (i.e. WNT, SHH) and subtypes that appear to proceed along a developmental trajectory or continuum (i.e. Group C and D) (Hovestadt et al., 2019). For most cancers, the relative contribution of oncogenic pathway alterations, cell of origin, and biological plasticity to the overall tumor phenotype (i.e. subtype) is unknown.

Small cell lung cancer (SCLC) has historically been treated as a single disease without patient stratification. SCLC exhibits genetic loss of both tumor suppressors *RB1* and *TP53*, along with mutually exclusive expression of *MYCL, MYC* or *MYCN* (Bragelmann et al., 2017; Dammert et al., 2019; George et al., 2015; Peifer et al., 2012; Poirier et al., 2015; Rudin et al., 2012). Large-scale gene expression analyses of human tumors and cell lines suggest that SCLC comprises four distinct molecular subtypes based on expression of lineage-defining transcription factors: *ASCL1* (SCLC-A), *NEUROD1* (SCLC-N), *POU2F3* (SCLC-P) or *YAP1* (SCLC-Y) (Rudin et al., 2019). Studies suggest that SCLC subtypes have unique therapeutic vulnerabilities (Cardnell et al., 2017; Chalishazar et al., 2019; Dammert et al., 2019; Huang et al., 2018a; Huang et al., 2018b; Mollaoglu et al., 2017; Owonikoko et al., 2019), emphasizing the clinical importance of understanding these subtypes. The origins and relationships amongst SCLC molecular subtypes are currently unknown (Poirier et al., 2020; Rudin et al., 2019).

The SCLC-ASCL1 subtype comprises ~70% of human SCLC. ASCL1 is a master regulator of neuroendocrine (NE) fate that is required for pulmonary NE cell (PNEC) development, and labels adult PNECs (Ito et al., 2001). The adult PNEC is a cell of origin for $ASCL1^{+}$ tumors in multiple *Rb1/Trp53* (RP)-null genetically-engineered mouse models (GEMMs) of SCLC (Meuwissen et al., 2003; Park et al., 2011; Sutherland et al., 2011). In these GEMMs, ASCL1 is essential for tumor development (Borromeo et al., 2016; Kim et al., 2016). *MYCL* is amplified or highly expressed in the SCLC-A subtype and is necessary for SCLC-A development (Borromeo et al., 2016; Dooley et al., 2011; George et al., 2015; Huijbers et al., 2014; Kim et al., 2016; McFadden et al., 2014). In contrast, the other three SCLC subtypes representing ~30% of tumors (SCLC-N, SCLC-P, SCLC-Y) tend to exhibit amplification or overexpression of *MYC*, and exhibit a low or non-NE cell fate (George et al., 2015; Rudin et al., 2019). We showed that *Myc* expression drives a non-NE SCLC phenotype in RP GEMMs and tumors express NEUROD1 (Mollaoglu et al., 2017). However, the relationship between SCLC-A and SCLC-N, and whether MYC drives SCLC-P or SCLC-Y subtypes is unknown.

Notch signaling regulates PNEC fate and SCLC tumorigenesis with dichotomous functions. Notch acts as a tumor suppressor in SCLC with ~25% of tumors harboring loss-of-function alterations in Notch receptors (George et al., 2015). During lung injury, *Notch2* marks an NE-stem cell population ($NE^{stem}$) that undergoes self-renewal in the absence of Notch activation (Ouadah et al., 2019). Loss of Notch function in the context of *Rb1* and *Trp53* loss is postulated to lock cells in a self-renewing $NE^{stem}$-like state and thereby contribute to SCLC. However, Notch is active in a subset of human SCLC, and Notch activation can

promote non-NE SCLC fate (Lim et al., 2017). Moreover, an unknown signal activates Notch signaling during lung injury to promote transit amplification and deprogramming of $NE^{stem}$ cells to non-NE fates (Ouadah et al., 2019). While MYC and NOTCH have both been implicated in non-NE cell fate in SCLC, a functional relationship between MYC and NOTCH in this setting has not been defined.

Here, we investigate the function of MYC in the origins and relationships among SCLC molecular subtypes.

## RESULTS

### MYC drives multiple SCLC molecular subtypes in vivo

To determine which SCLC molecular subtypes are promoted by MYC, we analyzed tumors in *Myc*-driven (*Rb1*$^{fl/fl}$;*Trp53*$^{fl/fl}$;Lox-Stop-Lox (LSL)-*Myc*$^{T58A}$, RPM) and *Mycl*-driven (*Rb1*$^{fl/fl}$;*Trp53*$^{fl/fl}$;*Rbl2*$^{fl/fl}$, RPR2) SCLC GEMMs (Figure S1A) at time points with similar tumor burden (Kim et al., 2016; Schaffer et al., 2010). We used adenoviral-*Cgrp* promoter-*Cre* viruses (Ad-Cgrp-Cre) to specifically transform PNECs (Sutherland et al., 2011). Early in situ lesions in RPM and RPR2 models exhibit high expression of NE markers including ASCL1, SYP, CGRP, UCHL1, and DLL3 (Figures 1A, 1B, and S1B). Invasive RPR2 tumors retained a NE-high identity, whereas invasive RPM tumors displayed significantly reduced NE marker expression (Figures 1A, 1B, and S1B). Bulk gene expression data from human SCLC cell lines has shown that Myc, Notch/Rest, Hippo/Yap1, and epithelial-mesenchymal transition (EMT) pathways correlate with non-NE fate in SCLC (Zhang et al., 2018). We analyzed expression of SCLC subtype markers and a subset of the non-NE fate markers in RPM and RPR2 tumors. Invasive RPM tumors gained expression of non-NE markers including NEUROD1, YAP1, HES1, ZEB1, and CD44 (Figures 1C, 1D, and S1B). MYC expression was high in both in situ and invasive RPM tumors, and not detectable in RPR2 tumors (Figure S1B). Interestingly, POU2F3 was rarely detected in RPM tumors derived from PNECs (Figure 1E). However, tumors initiated in RPM mice with a general promoter (Ad-CMV-Cre), but not a club or alveolar type II (AT2) promoter, were enriched for POU2F3 expression (Figures 1E, 1F, and S1C). POU2F3 was not detected in *Mycl*-associated RPR2 tumors initiated from a general or NE promoter (Figure 1E). Analysis of serial sections of $POU2F3^+$ tumors revealed that ~44% lacked expression of other SCLC subtype markers, whereas the majority of the remaining tumors expressed POU2F3 and NEUROD1 (Figures 1G and 1H). The fraction of cells expressing other subtype markers within $POU2F3^+$ tumors was relatively minor (< 16% of cells) (Figures 1G and 1H). These data suggest that MYC-driven SCLC-P tumors predominantly arise from an unknown cell-of-origin that is not a PNEC, club, or AT2 cell, and that MYC (as opposed to MYCL) promotes SCLC-P development. Since POU2F3 is a master driver of the tuft cell lineage (Huang et al., 2018b), future studies will be required to determine whether MYC-driven SCLC-P tumors arise from tuft cells.

Together, these data suggest that in the context of *Rb1* and *Trp53* loss, MYC can promote SCLC-N and SCLC-Y molecular subtypes from an $ASCL1^+$ PNEC cell-of-origin. Further, MYC drives SCLC progression in PNECs from an NE-high to non-NE phenotype. These findings are consistent with the enrichment of *MYC* expression in human SCLC-N, -Y and

P subtypes (Rudin et al., 2019), and demonstrate that MYC is a driver of these SCLC molecular subtypes in vivo.

## MYC drives SCLC subtype evolution in vitro

Static observations of tumor histology provide limited insight into the potential temporal connections between SCLC subtypes. To address this point, we macro-dissected central portions of the lung of Ad-Cgrp-Cre-infected RPM mice prior to detection of macroscopic lesions at a time point with predominantly $ASCL1^{+}$ in situ tumors (Figure 2A). We cultured dissociated cells and within 3-4 days, large clusters of tumor cells grew in suspension (Figure 2B). The tumor cells initially form tight, round, spherical aggregates that resemble classic, NE-high human SCLC cell lines (Figures 2B and 2C), but eventually form amorphous clusters with a “chain-link” morphology that resembles variant, non-NE human SCLC cells (Figures 2B, 2C, and S2A). This striking transition occurred reproducibly over a period of ~20 days, consistent with the time frame from early-to-invasive SCLC progression in the RPM GEMM (Mollaoglu et al., 2017). There was a minor level of cell death during culture that was relatively consistent throughout the time course (Figure S2B). Importantly, early-stage RPR2 tumor cells maintained a classic morphology in culture (Figures 2B and S2A).

To characterize the RPM tumor cell transition, we harvested RPM tumor cells at multiple time points during culture and assessed markers of SCLC subtypes. Cells at day 5 displayed an NE-high identity, marked by ASCL1 and NE marker expression (Figure 2D). NEUROD1 was transiently expressed from day 5-12, followed by expression of the non-NE markers REST, YAP1, NOTCH2, NOTCH2’s active cleaved intracellular domain (N2ICD), and HES1 at day 10-24. In contrast, control RPR2 tumor cells expressed ASCL1, and did not induce non-NE markers (Figure S2C). Ectopic $Myc^{T58A}$-Ires-Gfp expression in RPR2 cells triggered a variant morphology of looser clusters or chains, and induced non-NE markers (Figures S2C and S2D). This finding indicates that the NE-high to non-NE fate transition is not simply a consequence of prolonged culture.

To test if MYC is sufficient to alter tumor cell fate in established human SCLC cells, we expressed $Myc^{T58A}$-Ires-Gfp in NE-high, *MYCL*-associated classic lines. Ectopic $MYC^{T58A}$ expression altered cell morphology from classic to variant-like, and induced non-NE marker expression (Figures 2E and 2F). The levels of ectopic MYC overexpression were consistently less than that of MYC-high SCLC cell lines and RPM tumors (Figure S2E), suggesting this phenotype is not due to supraphysiological MYC levels. NEUROD1 expression was not detected in these experiments (Figures 2E and S2C), suggesting that NEUROD1 expression is either early and transient (as in Figure 2D) or that YAP1 expression can be induced by MYC without NEUROD1 expression. Despite the gain of non-NE marker expression and variant morphology induced by MYC, mouse and human MYC-expressing cells retained expression of some NE-markers including ASCL1 (Figures 2E and S2C), suggesting either a heterogeneous population, and/or the presence of hybrid NE/non-NE cells. Overall, these data demonstrate MYC’s ability to promote non-NE tumor cell fate and subtype evolution from SCLC-A to SCLC-Y in tumor cells in vivo and in human SCLC cells.

## Human SCLC subtypes correspond with MYC-driven evolution

Next, we sought to investigate the transcriptional states of MYC-driven tumor evolution in human SCLC tumors. We performed bulk RNA-sequencing of RPM tumor cells from Ad-Cgrp-Cre infected mice at eight time points during the transition spanning 21 days in culture. Analysis of NE and non-NE pathway genes confirmed a temporal loss of NE identity, and a subsequent gain of non-NE signaling pathways including Notch/Rest, Hippo/Yap1, and EMT (Figures 3A and S3A) (Zhang et al., 2018). Gene set enrichment analyses (GSEA) of an established 50-gene signature comprised of 25 NE and 25 non-NE-related genes from human SCLC cell lines (Zhang et al., 2018) demonstrated that MYC promotes a shift from NE to non-NE transcriptional states (Figures 3B and S3A). We observed dynamic and sequential expression of the subtype-defining transcription factor genes *Ascl1, Neurod1* and *Yap1* during the transition (Figure 3C). Consistent with a lack of POU2F3 in RPM tumors derived from PNECs (Figure 1E), we observed extremely low counts of *Pou2f3* mRNA during the transition (Figure 3C).

To determine whether the expression of the subtype-defining transcription factors correlate with gene expression patterns in human tumors, we clustered ~81 human SCLC tumors and ~51 human SCLC cell lines according to molecular subtype using bulk RNA-sequencing data (Figure S3B). We then created human SCLC-subtype-specific gene signatures using the most highly expressed genes per subtype (Table S1), and applied GSEA to determine whether each signature was enriched or depleted during the RPM transition (Figure 3D). Consistent with the patterns of transcription factor gene expression (Figure 3C), Day 3-5 and Day 7-10 cells were enriched for human SCLC-A and SCLC-N signatures, and depleted for the SCLC-Y signature. In contrast, Day 14-21 cells were enriched for the SCLC-Y signature and depleted for SCLC-A and SCLC-N signatures, demonstrating that MYC-driven tumor cell evolution in the mouse corresponds with subtype-defining transcriptional signatures in human tumors. We confirmed that *Rb1* and *Trp53* were recombined in RPM tumor cells (Figures S3C and S3D). Whole-genome sequencing (WGS) of early and late time-point RPM tumor cells also confirmed complete loss of expected regions of *Rb1* and *Trp53* with no detectable copy number variations (CNVs) (Figures S3E and S3F). Together, MYC-driven tumor cell evolution in vitro reflects the temporal phenotypic changes observed in the RPM GEMM, and places three of four human SCLC subtypes along a defined MYC-driven trajectory from SCLC-A to -N to -Y.

## MYC-driven SCLC subtypes progress along a single evolutionary trajectory

To better understand the transcriptional changes during MYC-driven SCLC progression, we performed single cell RNA-sequencing (scRNA-seq). We isolated unsorted early-stage tumor cells from RPM-Rosa26-LSL-Cas9-Ires-Gfp (RPM-Cas9) mice as they were transitioning from Day 4-21 in culture at six distinct time points (Figure 4A). Approximately 4-8,000 cells were captured per time point (n = 31,519 total cells). Day 4 and -7 cells comprised both tumor and non-tumor cell populations and the non-tumor populations were depleted in culture by Day 11 (Figures 4A and S4A). For downstream analyses, non-tumor and low-quality cells were filtered out of the time-course (Figures 4B and S4A). Minor variation in gene expression due to cell cycle genes was regressed out for downstream analyses (Figure S4B). Unsupervised tSNE clustering of tumor cells based on top highly-

expressed genes revealed at least three distinct clusters corresponding with the vast majority of Days 4-7, 11, and 14-21 cells, respectively (Figure 4B). Cells in the Day 4-7 cluster expressed high levels of *Ascl1* and NE markers, which were largely absent in Day 11-21 cells that instead expressed high levels of non-NE markers (Figure 4C). Though we captured only a few *Neurod1*-expressing cells in the time-course, these cells clustered with NE-high cells (Figure S4C).

Next, we performed scRNA-seq on four invasive RPM-Cas9 tumors, one predominantly ASCL1-high (RPM1) and three that had varying levels of NEUROD1 and YAP1 expression (RPM2-4) (Figures S4D and S4E). Combining the RPM transition time-points and tumor samples, we constructed pseudotime trajectories to identify predicted transcriptional relationships among the tumor cells. Unsupervised pseudotime ordering of combined tumor cells predicted a single lineage trajectory that corresponds with early to late transition time points (Figures 4D-4F). Faceted pseudotime trajectories reveal that cells of the RPM tumors progress along the same trajectory as transitioning cells in vitro, with the bulk of tumor cells in the earliest stages of pseudotime and fewer cells progressing to later stages (Figure 4D). RPM2-4 tumors had more cells in the later time points than RPM1, consistent with their lower ASCL1 levels and higher YAP1 protein expression (Figures 4D and S4D).

Expression of *Ascl1, Neurod1* and *Yap1* were consistent with MYC-driven temporal evolution, whereas *Pou2f3* was rarely detected (Figure 4E). *Mycl* was associated with *Ascl1* expression in early time points and absent during tumor cell progression (Figure 4E), consistent with studies showing that *ASCL1* and *MYCL* are coordinately expressed in chemo-naive SCLC, and reduced in chemotherapy-relapsed SCLC (Wagner et al., 2018). Expression of the top-500 differentially-expressed genes across pseudotime demonstrate broad loss of NE genes followed by gain of non-NE genes (Figure 4F and Table S2).

We detected a small number of *Ascl1*-high cells at Day 4 (n = 8 cells) that distributed to the late end of pseudotime (Figures 4D and S4F); only one of these cells clustered with late-stage tumor cells in tSNE space (Figures S4F and S4G), suggesting these cells are transcriptionally dissimilar from both early and late-stage tumor cells. As an orthoganol approach to predicting cellular trajectories, we performed diffusion mapping (Angerer et al., 2016; Coifman et al., 2005; Giraddi et al., 2018). Diffusion mapping revealed the same cellular trajectory from early, NE-high to late, non-NE cell states in transitioning cells in culture and in tumors (Figure S4H). Diffusion components (DCs) that best distribute time points in culture were compounded to define a principle curve predicting diffusion pseudotime (i.e. cellular trajectory) (Figure S4I). Diffusion pseudotime placed a small number of Day 4 cells late in pseudotime (Figure S4J), but these cells were transcriptionally distinct from other Day 4 cells and the bulk of late cells from other timepoints and tumors (Figure S4K). Diffusion analysis also revealed some features distinguishing primary tumor cells and transitioning cultured cells (e.g. DC6, DC4) (Figure S4L). Monocle and diffusion map pseudotime coordinates were significantly, positively correlated (Figure S4M). Together, pseudotime analyses are consistent with our data suggesting that MYC drives the temporal evolution of SCLC fate from NE to non-NE states in culture and in tumors in vivo.

## MYC-driven murine tumors exhibit intratumoral molecular subtype heterogeneity

Pseudotime analyses suggest that individual RPM tumors are composed of cells representing multiple SCLC molecular subtypes (Figures 4D and 4E). To test this more comprehensively, we analyzed serial sections of individual RPM tumors at the in situ or invasive stage for subtype markers (Figures 5A and 5B). In situ tumors were predominantly ASCL1$^{+}$ and harbored a single molecular subtype marker (~ 87%). In contrast, invasive tumors predominantly harbored two or more subtypes, either ASCL1 and NEUROD1, or all three molecular subtypes (ASCL1, NEUROD1 and YAP1). CIBERSORT analyses from an additional 10 RPM tumors also predicted intratumoral heterogeneity with cells throughout the stages of MYC-driven progression (Figure 5C). Analyses of the RPM tumors by scRNA-seq suggest that there is minimal coexpression of these markers within individual cells, with the exception of *Ascl1* and *Neurod1* (Figure 5D), consistent with their close connection in pseudotime. The average abundance of *Ascl1, Neurod1*, and *Yap1* in these tumors is also relatively consistent with bulk RNA-seq and protein analyses (Figures 5B-5D). scRNA-seq analyses of the RPM transition in culture also shows little overlap between *Ascl1, Neurod1*, and *Yap1* in individual cells (Figure 5E), with Day 11-21 cells resembling later stages of tumor progression. Thus, independent methods suggest that MYC drives the evolution of multiple SCLC subtypes in vivo.

## MYC activates Notch signaling during NE dedifferentiation

To uncover transcriptional patterns associated with MYC-driven tumor cell evolution, we performed differential gene expression analysis using the scRNA-seq data (Figure 6A). ENRICHR analysis of top differentially-expressed genes between the major clusters of cells (Day 4-7 vs Day 11 vs Day 14-21) identified transcription factors predicted to be important for SCLC fate (Figure 6B) (Wooten et al., 2019), including REST, SUZ12, TCF3, NEUROD1, NHLH1, MYC and SOX2. As Notch/Rest signaling can promote non-NE fate in SCLC (Lim et al., 2017), and REST is a top predicted regulator of the earliest changes promoted by MYC, we focused on the NE phenotypic switch. Using the human-derived NE-score vectors from (Zhang et al., 2018) (Table S3), we assigned an NE-score to every cell in the RPM time-series and bulk tumors (Figure 6C). The NE-score accurately predicted the decrease in NE identity over time with a dramatic reduction in NE-score between Day 7 and 11 (Figure 6D).

To determine how MYC promotes this transition at a mechanistic level, we performed chromatin-immunoprecipitation-sequencing (ChIP-seq) for MYC in invasive RPM tumors (n = 4) representing a spectrum of tumor cell states (Figure S5A). The top-50 upregulated genes bound by MYC in RPM tumors whose expression was enriched in *MYC*-high vs -low human SCLC tumors and in mouse RPM vs RPR2 tumors were used to define a conserved MYC-ChIP score (Table S3). Application of the MYC-ChIP score to the time-series data revealed a small, but significant increase in MYC activity from Day 11 onward (Figure 6E), validating the ENRICHR predictions (Figure 6B). These analyses suggest that Day 7-11 represents a key transition state, characterized by loss of NE identity and high-MYC activity.

Because we observed Notch/Rest pathway induction during MYC-driven SCLC progression, we analyzed the expression of Notch-related machinery and target genes during the

transition. Expression of Notch-inhibitory factors were increased at early time points (Figures S5B and S5C), including the inhibitory Notch ligand gene *Dll3, Hes6* (HES6 is a repressor of HES1), *Fbxw7* (FBXW7 promotes NOTCH degradation), and *Ncor2* (NCOR2 functions in a Notch-corepressor complex) (Gratton et al., 2003; Matsumoto et al., 2011; Saunders et al., 2015). In contrast, pro-Notch signaling genes increased in expression over time, including the Notch target genes *Hes1* and *Rest*, and the Notch receptor gene, *Notch2* (Figures S5B and S5C). GSEA of Notch signaling and REST-transcriptional targets by bulk RNA-seq reveals increased Notch pathway activity and repressive REST activity in late compared to early transition time points (Figure S5D). Importantly, multiple Notch-signaling components were identified as MYC target genes including *Notch2, Hes1, Hes6* and *Jag2* (Figure 6F). Consistently, NOTCH2 and HES1 protein levels were induced by MYC in multiple cell types (Figures 2D, 2E and S2C). Pro-Notch signaling genes *NOTCH2, HES1*, and *REST* are also preferentially enriched in *MYC*-high human SCLC (Figure 6G), while Notch-inhibitory *HES6* is reduced, suggesting broad positive regulation of Notch signaling by MYC in SCLC. Together, these data suggest that MYC increases NOTCH/REST activity to destabilize NE identity during MYC-driven SCLC evolution.

## Notch activation is required for MYC-driven tumor evolution

To determine whether Notch signaling is required for MYC-driven SCLC evolution, early RPM tumor cells were cultured in vehicle control or the gamma-secretase inhibitor (GSI-IX) (e.g. DAPT) to block Notch signaling. In contrast to control cells that began to convert to variant morphology on Day 10, DAPT-treated RPM cells did not switch to variant morphology until ~Day 20 (Figure 7A). Notch blockade increased and prolonged ASCL1, INSM1, EPCAM, and NEUROD1 expression compared to control cells (Figures 7A, S6A and S6B). In contrast, DAPT-treatment delayed or blocked expression of non-NE markers (Figures 7B, S6A and S6B). Interestingly, when DAPT is added to Notch-active fully-progressed RPM tumor cells rather than cells in the classic/early state, we observe equivalent expression of non-NE markers and no reversion to NE marker expression, despite efficient Notch signaling blockade (Figure S6C). Therefore, it is possible that MYC-driven tumor evolution may be irreversible, but it remains to be determined whether blockade of other non-NE-related pathways could revert SCLC to an NE-high state. Next, we treated RPM mice with early stage tumors with DAPT to block Notch signaling in vivo and monitored tumor growth by microCT imaging for 10 days, since control animals live an average of ~12 days following tumor detection by microCT imaging. DAPT treatment led to a significant decrease in tumor burden in DAPT-treated vs control mice (Figure 7C). Quantification of H&E-stained tissue revealed that tumor burden was significantly reduced by DAPT treatment (Figure 7D). DAPT-treated RPM tumors exhibited an increase in classic morphology indicated by smaller cells with high nuclear:cytoplasmic ratios compared to control tumors and a significant increase in NE identity, measured by ASCL1 and DLL3 levels (Figures 7E and 7F). Moreover, DAPT-treated tumors exhibited a decrease in tumor progression compared to control tumors evident by reduced expression of NEUROD1 and YAP1. We did not detect differences in HES1 expression, likely because HES1 levels were already low in control tumors at these time points (Figures 7E and 7F). MYC levels were high in both treatment groups with > 90% of cells positive for MYC, suggesting that the impact of Notch inhibition on cell state is not due to reducing MYC levels. Treatment of

tumor-bearing RPM mice with a second Notch inhibitor (dibenzazepine, DBZ) also significantly reduced tumor burden compared to control animals, but was highly toxic limiting analysis to Day 7 (Figures S6D and S6E). Together, these findings suggest that Notch inhibition significantly inhibits MYC-driven tumor progression.

Loss-of-function *NOTCH* mutations occur in ~25% of human SCLC, with mutually exclusive alterations in *NOTCH1, −2, −3* and *−4* (George et al., 2015). We hypothesized that MYC requires intact NOTCH to drive SCLC subtype evolution. Using published functional classifications of *NOTCH* mutations, we grouped human SCLC tumors and cell lines as either *NOTCH*-wild-type (WT) or silent, non-damaging, or damaging. Consistent with our hypothesis, *MYC* expression is significantly increased in *NOTCH* WT or silent mutant compared to *NOTCH*-damaging mutant samples (Figure 7G). All *MYC*-high SCLC samples (n = 30) are predicted to have intact *NOTCH* (Figure 7G, left panel). Many *MYC*-expressing tumors with intact *NOTCH* exhibit low NE-scores, whereas all of the *NOTCH*-damaging mutant tumors are NE-high (Figure 7G, right panel). We mined a recent cell line genomics database (SCLC_CellMiner) (Tlemsani et al., 2020), which allowed a similar analysis with 50 additional human SCLC cell lines (Figure 7H). There was a significant difference in the quantity and proportion of *MYC*-high *NOTCH*-WT versus *MYC*-high *NOTCH*-mutant samples that were NE-low, consistent with our model that MYC promotes non-NE progression by activating NOTCH signaling. GSEA suggest that Notch signaling and MYC activity are enriched in *NOTCH* WT human SCLC compared to samples with damaging mutations (Figure 7I and Table S4). Together, these results suggest that MYC depends on NOTCH to promote NE dedifferentiation, and that Notch blockade can inhibit MYC-driven tumor progression.

## Human SCLC exhibits intratumoral molecular subtype heterogeneity

We sought to determine the abundance of RPM time-point gene signatures in bulk RNA-seq data from human SCLC tumors and cell lines (George et al., 2015; Newman et al., 2015). CIBERSORT analyses revealed that most human tumors are predicted to have cells resembling multiple stages of MYC-driven progression, regardless of their classified SCLC subtype (Figure 8A), reminiscent of a similar approach predicting varying proportions of SCLC phenotypes in human tumors (Wooten et al., 2019). The majority of tumors harbor cells that resemble the NE-high ASCL1$^{+}$ early time-point RPM cells (Figures 8A and 8B). The human SCLC-N subtype was enriched for the Day 7-10 time-point signature and the human SCLC-Y samples exhibit a significantly higher percentage of non-NE late-time-point signatures (Figures 8A and 8B). This suggests that individual human tumors have cells in multiple stages of tumor progression, with *MYC*-high tumors more likely to have cells at the latest stages of progression. Interestingly, some SCLC-A tumors were predicted to have a higher percentage of late-stage cells; these tumors had moderately higher *MYC* levels and significantly higher levels of *NOTCH2, HES1* and *REST*, consistent with a reduction in NE identity (Figures 8A and 8C).

To verify these predictions in human tissue, we obtained 21 chemo-naive human SCLC biopsies (n = 6 limited and 15 extensive stage), since surgical specimens in SCLC are rare and difficult to obtain due to metastatic disease. We performed IHC for ASCL1, NEUROD1,

YAP1, and MYC on serial sections. Many tumors harbored cells with more than one subtype, with the majority of tumors having some frequency of ASCL1 and/or NEUROD1 protein (Figures 8D and 8E), consistent with human gene expression data (Figure 8A). Approximately ~14% of samples harbored detectable MYC protein, and of these, none of them were in the ASCL1-only group. Finally, we obtained one fresh human SCLC biopsy from a patient who briefly responded to carboplatin and etoposide and then progressed on carboplatin-irinotecan for scRNA-seq analyses. The biopsy harbored distinct *ASCL1*$^{+}$ and *NEUROD1*$^{+}$ populations, and high *MYC* expression overlapped specifically with the *NEUROD1*-high population (Figures 8F, 8G, S7A and S7B). While these data warrant a more comprehensive analysis in human tissue, these findings are consistent with bioinformatic predictions suggesting that tumors frequently harbor cells representing multiple SCLC subtypes.

## DISCUSSION

While SCLC has historically been treated as a single disease, recent studies have converged on the concept that SCLC is composed of at least four molecular subsets with unique therapeutic vulnerabilities. Recent bioinformatic approaches suggest that SCLC subtypes may represent dynamic states of transition (Wooten et al., 2019). Our functional data here support that hypothesis and suggest that MYC has the capacity to shift SCLC molecular subtypes. We find in the context of *Rb1* and *Trp53* loss, MYC can promote three of the four molecular subsets from a PNEC cell of origin that proceed in a temporal evolution from SCLC-A, to SCLC-N to SCLC-Y in vivo. Studies in a limited number of human SCLC cell lines and patient-derived xenograft (PDX) models also suggest that MYC can convert ASCL1$^{+}$ SCLC to a variant morphology with NEUROD1 expression (Johnson et al., 1986; Patel et al., 2019; Simpson et al., 2020). Interestingly, we observed SCLC-P in RPM mice when an unknown cell type was targeted that we speculate could be the tuft cell (Huang et al., 2018b; Rudin et al., 2019). These data demonstrate that cell of origin, genetics, and tumor cell plasticity can determine SCLC subtype.

Our data suggest that MYC requires Notch pathway activity to promote tumor progression. Notch activity can promote non-NE fate in *Mycl*-associated SCLC models in the absence of MYC expression (George et al., 2015; Lim et al., 2017). In the RPR2 model, tumors with MYC-negative Notch-active cells do not develop variant morphology or express NEUROD1 or YAP1, in contrast to MYC-expressing Notch-active cells in our study. Together, these findings suggest that Notch activity alone is not sufficient to drive SCLC-N and SCLC-Y subtypes, and that MYC and NOTCH likely cooperate with one another to drive SCLC progression.

Recently identified NE$^{stem}$ cells are enriched for *Notch2* and *Hes1* expression (Ouadah et al., 2019). *Hes1* is particularly enriched in RPM time points where MYC-driven tumor cells are transitioning to a non-NE fate. *HES1* is also enriched in *ASCL1*$^{+}$ human SCLC tumors that exhibit a more non-NE fate in the absence of high *MYC*, consistent with a model whereby Notch can promote non-NE fate, but may be limited in its ability to promote progression without MYC. We identify *Notch2* and *Hes1* as MYC target genes, suggesting that MYC could be the unidentified signal that induces Notch signaling to deprogram

NE$^{stem}$ cells to other cell fates during lung injury (Ouadah et al., 2019). We propose a model whereby NE$^{stem}$-like tumors with defective Notch signaling are locked in a NE-high state, whereas tumors with intact Notch can be reprogrammed by MYC to non-NE fates—explaining the dichotomous nature of Notch signaling in SCLC (Figure 8H). While our pharmacological studies suggest that Notch activity is critical for MYC-driven tumor progression, genetic Notch knockout studies are warranted to fully test this model. These data warrant evaluation of pharmacological approaches that activate Notch in SCLC (Augert et al., 2019; Oser et al., 2019) to determine whether they can promote tumor progression particularly in collaboration with MYC.

The RPM primary cell culture model should be a valuable tool to better understand the role of additional signaling pathways in MYC-driven tumor evolution. Lineage-tracing approaches in this model in vitro and in vivo would be a powerful complement to the studies here. As some of our assays suggest that SCLC can progress from SCLC-A to SCLC-Y without evidence of the SCLC-N subtype, further studies are warranted to determine whether NEUROD1 is required or dispensable for SCLC progression. The functions of Hippo/Yap1 and EMT in SCLC (also induced by MYC) are less well understood than Notch signaling (Horie et al., 2016; Jia et al., 2018; McColl et al., 2017). Previous studies suggest that SCLC-Y human tumors are often *RB1* wild-type (McColl et al., 2017). We observe MYC-high YAP1$^{+}$ tumor cells that appear to lack *Rb1*, suggesting there may be more than one mechanism to induce YAP1 in SCLC.

Our work builds on an emerging concept that MYC and MYCL are not functionally redundant in SCLC. MYC and MYCL correlate with distinct gene expression and methylation profiles, and localize to distinct super enhancers (Borromeo et al., 2016; Christensen et al., 2014; Poirier et al., 2015). Functional studies modulating MYC reveal MYC's capacity to change cell fate, morphology, drug sensitivity, and molecular subtype (Dammert et al., 2019; Mollaoglu et al., 2017; Patel et al., 2019). In other tumor types like medulloblastoma, MYC and MYCN also appear to have distinct roles (Vo et al., 2016), illustrating divergent functions for MYC family members in cancer.

We find that many SCLC tumors are composed of cells representing multiple molecular subtypes at different frequencies, suggesting that bulk analysis methods may only identify the most abundant state in a given tumor. Studies are increasingly identifying unique therapeutic vulnerabilities for SCLC subtypes. For example, MYC-high tumors are preferentially sensitive to inhibition of AURKA/B, CHK1, IMPDH1/2, and arginine deprivation (Cardnell et al., 2017; Chalishazar et al., 2019; Dammert et al., 2019; Huang et al., 2018a; Mollaoglu et al., 2017; Sen et al., 2017). POU2F3$^{+}$ SCLC cells are sensitive to IGF1R inhibitors (Huang et al., 2018b), while ASCL1$^{+}$ tumors express more DLL3 and are sensitive to DLL3-targeting drugs (Cardnell et al., 2017; Saunders et al., 2015). Considering data that SCLC subtypes have distinct therapeutic vulnerabilities, this suggests that dynamically-evolving tumors represent a "moving therapeutic target", adopting unique therapeutic vulnerabilities as they progress. Given many failed clinical trials with targeted therapies in SCLC, we speculate that the reason why chemotherapy (a "blunt instrument") has remained the most effective therapeutic option is likely due to its non-specific cytotoxicity in the multiple subtypes of SCLC that evolve during progression.

As clinical trials begin to assess biomarkers of SCLC subtype and potentially enroll patients based on these subtypes (Owonikoko et al., 2019; Poirier et al., 2020), it will be critical to assess subtype heterogeneity and anticipate tumor evolution to other subtypes. Multiple studies have implicated an increase in MYC and a decrease in MYCL and ASCL1 in chemotherapy-resistant mouse and human SCLC (Chalishazar et al., 2019; Farago et al., 2019; Huang et al., 2018a; Mollaoglu et al., 2017; Stewart et al., 2020; Wagner et al., 2018) and have correlated high *MYC* with shorter patient survival and more aggressive, drug-resistant phenotypes (Carney et al., 1985; Gazdar et al., 1985; Johnson et al., 1986). Recent studies also demonstrate an increase in intratumoral transcriptional heterogeneity in chemo-resistant patient samples (Stewart et al., 2020). Together, these findings suggest the provocative notion that therapy selects for and/or promotes the latest stages of tumor cell progression identified here. We speculate that cancers in other tissues may also harbor cells in transcriptionally-dynamic states of progression, and would potentially benefit from either more blunt, combinatorial, or plasticity-directed therapies.

## STAR+METHODS

### RESOURCE AVAILABILITY

**Lead contact—**Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Trudy G. Oliver (Trudy.Oliver@hci.utah.edu).

**Materials availability—**The RPM mice used in this study are deposited at The Jackson Laboratory, JAX #029971. There are restrictions to the availability of the RPM-Cas9 mice due to further characterization of the Cas9 allele needed. Plasmids in this study will be deposited to Addgene.

**Data and code availability—**All software is commercially available or cited in previous publications. Mouse RNA-seq, scRNA-seq and WGS data in this study are deposited in NCBI GEO: GSE149180. ChIP-seq data is deposited in GEO: GSE142496. Human scRNA-seq data is available at NCBI GSE149180. R scripts used to process single cell RNA-seq data are available upon request.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

**Human SCLC samples—**All patients provided informed consent for the collection of human specimens, and it was approved by the University of Utah Institutional Review Board (IRB_00010924) in accordance with the U.S. Common Rule. We queried for patients diagnosed with SCLC between 1st January 2011 and 1st April 2017 from the Cancer Clinical Research Database at Huntsman Cancer Institute (HCI). Chemotherapy-naive patients with availability and consent for utilization of archival tumor tissue for research purposes were isolated. We then retrospectively reviewed the available clinical data from the analyzed samples to obtain information about patient demographics, clinical staging, treatment response and survival. The original diagnosis of SCLC was made per standard-of-care by a board-certified pathologist at HCI, and a secondary confirmation of diagnosis for research purposes was obtained by board-certified pathologist Dr. Benjamin L. Witt. Single cell RNA

sequencing was performed on a tumor biopsy specimen collected from a patient with SCLC treated at the Siteman Cancer Center, Washington University School of Medicine in St. Louis. The biopsy specimen from this patient was banked for research purposes following consent through an IRB-approved protocol at the Washington University School of Medicine in St. Louis (HRPO 201305031).

**Clinical details—**21 unique patients with small cell lung cancer were identified between 1st January 2011 and 1st April 2017 with adequate archival tissue for analysis. Lung was the most common site of tissue biopsy (95%, 20/21). All patients were chemotherapy-naïve and the majority of patients had extensive disease at the time of initial diagnosis (71%, 15/21). The liver biopsy specimen for single cell sequencing was collected from a patient with extensive stage SCLC, at the time of progression, following initial therapy with carboplatin and etoposide - to which the disease briefly responded (platinum-resistant relapse). Disease at relapse failed to respond to subsequent therapy with carboplatin and irinotecan, and nivolumab and ipilimumab.

**Cell lines—**Human SCLC cell lines were obtained from ATCC, Dr. Minna (UTSW), Dr. Bunn (UC-Denver) or Martin Sos (Germany) and cultured in either RPMI (Fisher cat# MT10-040-CV) supplemented with 10% fetal bovine serum (FBS) (Sigma cat# 12303C), 1% L-glutamine (Invitrogen cat#2 5030-081), and 1% penicillin/streptomycin antibiotic cocktail (Invitrogen cat# 15140-122) (H889, GLC1, H1963, GLC8, H82); modified HITES medium DMEM/F12 (VWR cat#45000-344) supplemented with 0.005 mg/mL insulin (Fisher cat#12585-014), 0.01 mg/mL transferrin (Sigma cat# T2036), 30 nM sodium selenite (Sigma cat# S5261), 10 nM hydrocortisone (Sigma cat# H0888), 10 nM beta-estradiol (Sigma cat# E2758), 1% L-glutamine (Invitrogen cat#2 5030-081), and 1% penicillin/streptomycin cocktail (Invitrogen cat# 15140-122) (H1092); or DMEM (Fisher cat# MT-10-013-CV) with 10% FBS, 1% L-glutamine, and 0.1% penicillin/streptomycin antibiotic cocktail (HEK-293T). Cell line identity was confirmed by STR profiling in December, 2019.

**Mice—***Rb1*$^{fl/fl}$;*Trp53*$^{fl/fl}$;*Rbl2*$^{fl/fl}$ (RPR2) (Schaffer et al., 2010) (now deposited in MMRRC: 043692-UCD) and *Rb1*$^{fl/fl}$;*Trp53*$^{fl/fl}$;*MycT58A*$^{LSL/LSL}$ (RPM) (JAX #029971) (Mollaoglu et al., 2017) mice have been previously described. We crossed RPM mice with Rosa26-LSL-Cas9-Ires-Gfp mice (Platt et al., 2014) (JAX # 024857) to generate RPM-Cas9 mice with Cre-dependent expression of Cas9 and Egfp. Unfortunately, GFP was unreliable in marking tumor cells by flow cytometry and was highly variable by IHC in RPM-Cas9 mice, so we did not sort cells by GFP for these experiments. All RPM, RPM-Cas9, and RPR2 mice were housed and treated according to regulations set by the Institutional Animal Care and Use Committee of the University of Utah. Viral infections were performed in a Biosafety Level 2+ room following guidelines from the University of Utah Institutional Biosafety Committee. Male and female mice were distributed equally for all experiments.

## METHOD DETAILS

**Mouse lung tumor initiation—**Anesthetized RPM or RPR2 mice at 6-8 weeks of age were infected by intranasal instillation (DuPage et al., 2009) with $1x10^8$ plaque-forming

units of Ad5-CGRP-Cre, Ad5-CMV-Cre, Ad5-CCSP-Cre or Ad5-SPC-Cre adenovirus (University of Iowa). For harvesting in situ lung tumor tissue, RPM mice were sacrified at 3-6 weeks post-adenoviral infection, while RPR2 mice were sacrificed at 4-6 months post-infection due to differences in tumor latency (Figure S1C). For harvesting invasive lung tumor tissue, RPM mice were sacrified at 4-9 weeks post-adenoviral infection, while RPR2 mice were sacrificed at 4-8 months post-adenoviral infection in order to match stages of tumor development and burden.

**In vivo DAPT and DBZ—**For drug treatments, RPM mice were enrolled in treatment at the earliest signs of tumor burden determined by microCT imaging (typically 4-6 weeks post-Cre, 6-10% tumor burden). RPM mice (n = 8) were given freshly prepared DAPT (GSI-IX) γ-secretase inhibitor (20 mg/kg, Apexbio cat no. A8200) diluted in corn oil from a 20 mg/mL stock in 100% EtOH to a 4 mg/mL working concentration. Control mice (n = 6) received corn oil of equivalent volumes determined by body weight. Mice were treated via intraperitoneal (i.p.) injections every day for 10 days. Control and treatment cohorts were weighed daily to assess overall fitness and imaged 3-4 days/week to monitor tumor burden. At day 10, mice were imaged, then euthanized by $CO_2$ asphyxiation followed by necropsy. For DBZ (Apexbio cat no. A4018), RPM mice were given 14 mg/kg DBZ in vehicle (0.1% Tween-80, 0.5% hydroxypropylmethylcellulose, and 5% DMSO) or vehicle control i.p. for 7 days and imaged every 2-3 days.

**MicroCT analysis and quantification—**Beginning four weeks after tumor initiation by Cre-mediated recombination, RPM mice were imaged to monitor tumor development. Mice were anesthetized with isoflurane and imaged using a small animal Quantum GX2 microCT (Perkin Elmer). Quantum GX2 images were acquired with 18 s scans at 45 μm resolution, 90 kV, with 88 μA of current. RPM mice were sacrificed for transition studies at first signs of airway thickening and early tumor development.

To determine tumor burden, resulting images were processed with Analyze 11.0 software (Analyze Direct) as described previously (Mollaoglu et al., 2017). Scans were calibrated for Hounsfield Units (HU) by determining the mean value of “Bed” and “Air” for representative scans through the region of interest (ROI) tool and matching those values to their known HU (40 HU and −1000 HU, respectively) using the “Image Algebra” tool. Every image was then applied a 3x3x3 Median Filter from the “Spatial Filters” window. Thresholds of “Air” vs. “Dense Tissue” were established using the ROI and histogram tools. For tumor burden analysis, the object map was created using the previously established thresholds; adjustments were made manually using “Spline Edit”, “Draw”, “Trace” and “Nudge Edit” tools. With the “Morphology” tool, the object map was made binary by using the threshold morphing tool. Then, the map was dilated 3 times using 5x5x5 Jack-shaped structuring elements. The holes were then filled on every 2D-orientation. The map was finally brought back to its original size using the “Erode” tool 3 times using 5x5x5 Jack-shaped structuring elements. The volumetric analyses were then performed in the ROI window using the pre-established thresholds and non-airspace was calculated using the formula: Non- airspace = 1 – (VolAir/ROIVol).

**Immunohistochemistry—**Tissues were fixed in 10% neutral buffered formalin for 24 h at room temperature (RT), washed in PBS and transferred to 70% ethanol. Formalin-fixed paraffin embedded (FFPE) sections at 4-5 μm were dewaxed, rehydrated and subjected to high-temperature antigen retrieval by boiling 20 min in a pressure cooker in 0.01 M citrate buffer at pH 6.0. Slides were quenched of endogenous peroxide in 3% H2O2 for 15 min, then blocked in 5% goat serum in PBS/0.1% Tween-20 (PBS-T) for 1 h, and then stained overnight with primary antibodies in blocking buffer (5% goat serum or SignalStain antibody diluent, Cell Signaling Technology (CST) cat# 8112). For non-CST primary antibodies, an HRP-conjugated secondary antibody (Vector Laboratories) was used at 1:200 dilution in PBS-T, incubated for 45 min at RT followed by DAB staining (Vector Laboratories). Alternatively, CST primary antibodies were detected using 150 μL of SignalStain Boost IHC Detection Reagent (CST cat# 8114). All staining was performed with Sequenza cover plate technology. The primary antibodies include: ASCL1 (BD cat# BD556604) 1:200; UCHL1 (Sigma cat# HPA005993) 1:300; SYP (Thermo-fisher cat# RB1461P1) 1:200; CGRP (Sigma cat# C8198) 1:250; NEUROD1 (Abcam cat# 109224) 1:150; YAP1 (CST cat# 17074S) 1:400; ZEB1 (Abcam cat# ab133357) 1:2000; CD44 (CST cat# 37259) 1:250; POU2F3 (Sigma cat# HPA019652) 1:300; DLL3 (Abcam cat# 198505) 1:200; HES1 (CST cat# 11988) 1:500; MYC (mouse) (Santa Cruz cat#sc-764) and MYC (human) (Abcam cat# ab32072). For manual H-score quantification, images were acquired on a Nikon Ci-L LED Microscope with DS-Fi3 Camera. H-score was quantified on a scale of 0-300 taking into consideration percent positive cells and staining intensity as described (Flowers et al., 1986), where H Score = % of positive cells multiplied by intensity score of 0-3. For example, a tumor with 80% positive cells with high intensity of 3 = 240 H-Score.

For quantification of DAPT-treated vs control mice, H&E and IHC-stained slides were digitally scanned with the Zeiss Axio Scope.A1 microscope using AxioVision SE64 software. Whole slide images containing 4-5 lung lobes per animal were analyzed using CaseViewer software (3DHISTECH). Tumor regions were manually annotated and image analysis algorithms were applied only on tumor regions. The nuclear algorithm distinguishes cells as positive or negative based on the staining intensity per cell. Results are expressed as percent positive cells per tumor area (i.e. the number of positive cells divided by the number of total cells in a tumor area, multiplied by 100) (P; 0-100%).

**Primary tumor cell isolation—**Early tumor cells in RPM and RPR2 mice are identified prior to detection of overt tumors by microCT at a time point in which we observe only in situ tumors, as illustrated in Figure 2A. To generate mouse primary early-stage SCLC tumor cells for time-series transition experiments, whole lungs from RPM or RPR2 mice were digested at earliest stages of tumor development determined by thickening of airways in microCT images: ~3-6 weeks (RPM) or 4-6 months (RPR2) post-infection with $10^8$ plaque forming units of Ad5-CGRP-Cre virus (University of Iowa). Central portions of mouse lungs were mechanically dissociated with scissors, and digested to a single-cell suspension using an enzymatic digestion cocktail for 30 min at 37°C. Digestion media consists of 4200 mL HBSS-free (Thermo Fisher cat# 14175), 600 mL trypsin-EDTA (0.25%) (Thermo Fisher cat# 25200-072), 600 mL collagenase type 4 (Worthington Biochemical cat# LS004186) from 10 mg/mL stock prepared in HBSS with calcium and magnesium (Thermo Fisher cat#

14025), and 600 mL dispase (Worthington Biochemical cat# LS02104). 4 mL of digestion media was used per mouse lung. Enzymatic digestion was quenched on ice with 500 μL quench media containing 7.2 mL Leibovitz's L15 media (Thermo Fisher cat# 11415-064), 800 mL FBS (Sigma cat# 12303C), and 30 mL DNase (Sigma cat# D4527) at 5 mg/mL in HBSS-free media per mL of digestion media. Tissue was further dissociated by passing through a 16-gauge syringe to the point of no clogging. The tissue suspension was then passed through a 100 micron cell-strainer. Cells were spun at 2000 rpm for 5 min. Supernatant was removed and replaced with 3 mL ACK (Ammonium-Chloride-Potassium) lysis buffer per lung to remove whole blood cell contamination (Thermo Fisher cat#A10492, 3 min incubation at 37°C). Reaction was quenched with 10 mL cold 1X PBS. Cells were spun at 1500 rpm for 5 min and resuspended in culture media (RPMI) as described above. Following various times in culture as indicated, malignant cells in suspension were isolated from the adherent lung cells, pelleted by centrifugation, resuspended in fresh media, and processed for downstream analysis. Suspension culture images were taken using an EVOS XL Core Cell Imaging System (Invitrogen) at each collected time point of the transition study.

**Quantification of tumor cell morphology—**Tumor cell morphology was quantified using ImageJ area output as well as analysis plugins for the common shape factors: Circularity and Roundness. Area in pixels of each cluster was converted to square microns (based on image scale) to represent individual cluster size. ImageJ's Circularity function is calculated as 4*π*area/(perimeter^2). Roundness is calculated according to the equation 4*area/(π*major_axis^2), which represents the ratio of height to width. Both Circularity and Roundess values range from 0-1 where 1 would indicate a perfect circle and <1 indicates an increasingly elongated or linear/elliptical shape. Circularity is more complex than Roundness and captures perimeter smoothness (i.e. looser clusters with chains growing asymmetrically off sides would be low in circularity value) versus Roundness, which estimates how close a cluster is to a perfect circle (e.g. long chains would be much lower than classic clusters for roundness). Figure S2A represents ~15 clusters of cells per timepoint and n = 3 individual biological replicates. Figure 2F quantification represents 6-16 clusters per condition.

**Live/dead assays—**Live/dead assays were performed on n = 2-5 biological replicates of RPM transition experiments at indicated days in culture (Figure S2B). Approximately 500,000 suspension cells were collected from culture at indicated timepoints following initial digestion and washed 1x with PBS by centrifugation at 1000 rpm. Cells were resuspended in 500 μL Accutase dissociation reagent (Invitrogen cat#A1110501) and dissociated for 15 min at RT. Cells were washed with PBS, stained with Invitrogen's LIVE/DEAD Fixable Violet Dead Cell Stain Kit (Invitrogen cat#L34963) and fixed with formaldehyde according to the manufacturer's protocol. Human GLC8 cells were heat-treated at 56° C for 45 min and cooled to room temperature before staining and fixation for a dead, positive control. Cells were stored in the dark at 4° C until flow cytometry analysis. Data acquisition was performed with the BD Fortessa flow cytometer (BD Biosciences) running BD FACSDiva v8 software. Data were analyzed using FlowJo v10.6.2 and represented using Graphpad Prism.

**Immunoblotting—**Cell pellets were flash frozen and stored at −80°C until use. Total protein lysates were prepared as previously described, separated via SDS-PAGE and transferred to a PVDF membrane (Oliver et al., 2011). Membranes were blocked for 1 h in 5% milk followed by overnight incubation with primary antibodies at 4°C. Membranes were washed for 4 x 10 min at RT in TBS-T. Mouse and rabbit HRP-conjugated secondary antibodies (Jackson ImmunoResearch, 1:10,000) were incubated for 1 h in 5% milk at RT followed by washing 4 x 10 min at RT in TBS-T. Membranes were exposed to WesternBright HRP Quantum substrate (Advansta) and detected on Hyblot CL film (Denville Scientific Inc).

Primary antibodies include: ASCL1 (1:300, BD Pharmingen #BD556604), EPCAM (1:1000, Abcam #ab71916), INSM1 (1:300, Santa Cruz sc-271408), NEUROD1 (1:1000, Abcam ab109224), REST (1:1000, Millipore #17-641), YAP (1:1000, CST #14074), NOTCH2 XP (1:1000, CST #5732), HES1 (1:400, CST #11988), MYC (1:1000, CST #5605), NKX2-1 (1:2000, Abcam ab76013), ZEB1 (1:500, Bethyl Labs A301-922A), and HSP90 (1:1000, CST #4877) as loading control.

Quantification of immunoblots was performed using ImageJ across multiple biological replicates and normalized to the loading control for each replicate, HSP90.

**PCR for recombination efficiency—**Flash-frozen transition timepoints and established RPM cell lines were processed with the Qiagen DNeasy kit to isolate genomic DNA. Normal adult spleen DNA from an RPM mouse was isolated as an unrecombined control. DNA concentrations were measured on a BioTek Synergy HT plate reader. Equal quantities of tumor genomic DNA (100 ng) were amplified by PCR with GoTaq (Promega M7123) using primers to detect *Rb1* recombination include the following: D1 5'-GCA GGA GGC AAA AAT CCA CAT AAC-3', 1lox 5' 5'-CTC TAG ATC CTC TCA TTC TTC CC-3', and 3' lox 5'-CCT TGA CCA TAG CCC AGC AC-3'. PCR conditions used were 94 deg 3 min, 30 cycles of (94 deg 30 s, 55 deg 1 min, 72 deg 1.5 min), 72 deg 5 min, hold at 4 deg. Expected band sizes were ~500 bp for the recombined allele, and 310 bp for the floxed allele. Primers to detect *Trp53* recombination include the following: A 5'-CAC AAA AAC AGG TTA AAC CCA G-3', B 5'-AGC ACA TAG GAG GCA GAG AC-3', and D 5'-GAA GAC AGA AAA GGG GAG GG-3'. PCR conditions used were 94 deg 2 min, 30 cycles of (94 deg 30 s, 58 deg 30 s, 72 deg 50 s), 72 deg 5 min, hold at 4 deg. Expected band sizes were 612 bp for the recombined allele, and 370 bp for the floxed allele. PCR products were run on 1-2% agarose/TAE gels containing ethidium bromide and images were acquired using an Azure Biosystem C200 imager.

**MYC overexpression and virus production—**Mouse $Myc^{T58A}$ cDNA was cloned into MSCV-Puro-IRES-GFP (MSCV-PIG) (Addgene cat# 21654) (Mayr and Bartel, 2009) plasmid for retroviral overexpression of MYC in vitro. MSCV-PIG-$MYC^{T58A}$ plasmid was confirmed by direct sequencing. For generation of high-titer virus, HEK-293T cells were transfected with a three-plasmid system including: MSCV-PIG (Addgene cat# 21654) with an empty vector or *Myc* T58A insert, pCMV-VSVG (Stewart et al., 2003), (Addgene cat# 8454), and pCMV delta R8.2 (Stewart et al., 2003) (Addgene cat# 8455). Viruses were harvested at 48 and 72 h post-transfection, concentrated by ultracentrifugation (24,000 x g

for 1.45 h), and stored at −80°C until use. $MYC^{T58A}$ was overexpressed in human cell lines (H889, H1963) or RPR2 primary cells through retroviral spinoculation. Spinoculation was performed at 37°C, 900 x g, for 75 min. During spinoculation, 0.5-1 million cells per well of a 6-well plate were cultured with 2 mL RPMI, 25 μL HEPES buffer (Thermo Fisher cat# 15630080), 8 μg/mL polybrene (Santa Cruz cat# sc-134220), and 25 μL retroviral MSCV-PIG or MSCV-PIG-$MYC^{T58A}$ with titer $>10^6$ infectious units/mL. Cells were selected 48 h after spinoculation with puromycin at a concentration of 1 μg/mL (H889) or 0.5 μg/mL (H1963, RPR2) until uninfected control cells were dead.

**DAPT treatments in vitro—**RPM time-series transition cell lines were treated with 10 μM DAPT (GSI-IX) γ-secretase inhibitor (Apexbio cat# A8200) from a 10 mM stock in DMSO starting on day 3 following initial digest. Control cells were given equal volumes of DMSO only. DAPT was replaced every 3 days upon media change and pellet collection. Cells grew in DMSO or DAPT for 20-30 days. Images were taken using an EVOS XL Core Cell Imaging System (Invitrogen) at each collected time point of the DAPT transition study.

**ChIP-seq in RPM tumors—**Mouse lung tumor ChIP-seq data is previously described and published (Chalishazar et al., 2019) and deposited in the NCBI GEO (GEO: GSE142496). Due to limited size of individual tumors for ChIP, we micro-dissected pieces of individual tumors and assessed them for ASCL1 and NEUROD1 by immunoblot; the corresponding tumor fractions were then submitted for ChIP. Binding profiles were visualized using Integrated Genome Viewer (IGV version 2.6.3) aligned to mm10 genome build.

**Whole genome sequencing (WGS)—**30X WGS data was collected from Day 4 and Day 23 samples, as well as from a blood sample from RPM mice as the normal control. Genomic DNA was extracted from flash frozen cell pellets of Day 4 and −23 cells along with whole blood from the same RPM mouse using Qiagen's DNeasy Blood and Tissue kit (Qiagen cat#69504). Libraries were prepared using the Nextera DNA Flex Library Prep Kit (Illumina cat#20018705). Libraries were sequenced on a NovaSeq 6000 instrument targeting 300 million read-pairs on a 2 x 150 bp run (~30x coverage of whole genome). Sequencing reads were aligned to mouse genome mm10 by BWA 0.7.17-r1188 (Li and Durbin, 2009). *Rb1* and *Trp53* deletions were examined in the Integrated Genome Viewer (IGV) software v2.5.0. SNVs were jointly called by Freebayes 1.2.0 and somatic SNVs were filtered by the following criteria: DP >15, AO < 2 and AF <0.05 in the normal sample. Variants were annotated by SnpEff 4.3 (Cingolani et al., 2012). Somatic non-synonymous coding variants were manually reviewed in IGV. CNVs were called, and plots were generated by FACETS (Shen and Seshan, 2016).

**Mouse tumor and timepoint bulk RNA-seq—**RNA isolation from ~15 mg flash-frozen RPM (n = 10) and RPR2 (n = 5) primary tumors from mice infected with Ad5-CGRP-Cre or flash-frozen RPM transition cell pellets (n = 8: Days 3, 5, 7, 10, 12, 14, 19, 21) was performed using RNeasy Mini Kit (Qiagen) with the standard protocol. RNA from RPM tumors (n = 10) was subject to library construction with the Illumina TruSeq Stranded mRNA Sample Preparation Kit (cat# RS-122-2101, RS-122-2102) according to the manufacturer's protocol. RNA from RPR2 (n = 5) tumors and RPM transition timepoints (n

= 8) were subject to library construction with the Illumina TruSeq Stranded Total RNA Library Ribo-Zero Gold Prep kit (cat# RS-122-2301) according to the manufacturer's protocol. Chemically denatured sequencing libraries (25 pM) from RPM (n = 10) and RPR2 (n = 5) tumors and RPM transition timepoints (n = 8) were applied to an Illumina HiSeq v4 single read flow cell using an Illumina cBot. Hybridized molecules were clonally amplified and annealed to sequencing primers with reagents from an Illumina HiSeq SR Cluster Kit v4-cBot (GD-401-4001). Following transfer of the flowcell to an Illumina HiSeq 2500 instrument (HCSv2.2.38 and RTA v1.18.61), a 50-cycle single-end sequence run was performed using HiSeq SBS Kit v4 sequencing reagents (FC-401-4002).

**Bulk RNA-seq data analysis—**Fastq raw count files were aligned in the R statistical environment (version "3.6"). The mouse GRCm38 FASTA and GTF files were downloaded from Ensembl release 94 and the reference database was created using STAR version 2.6.1b (Dobin et al., 2013) with splice junctions optimized for 50 base pair reads. Optical duplicates were removed using clumpify v38.34 and adapters were trimmed using cutadapt 1.16 (Martin, 2011). The trimmed reads were aligned to the reference database using STAR in two pass mode to output a BAM file sorted by coordinates. Mapped reads were assigned to annotated genes in the GTF file using featureCounts version 1.6.3 (Liao et al., 2014). The output files from cutadapt, FastQC, Picard CollectRnaSeqMetrics, STAR and featureCounts were summarized using MultiQC (Ewels et al., 2016) to check for any sample outliers. To remove sources of unwanted variation from tumor RNA-seq sample preparation, all non-coding features, histones, and ribosomal RNAs were removed from count matrices for downstream analyses. The featureCount output files for RPM and RPR2 tumors were combined into a single raw count matrix. Differentially expressed genes (DEGs) between RPM and RPR2 tumors were identified using a 5% false discovery rate with DESeq2 version 1.24.0 (Love et al., 2014). PCA was performed on the first two principle components using the regularized log count (rlog) values of the top 500 variable genes. Log2(counts+1)-transformed, normalized intensity values were obtained and averaged across samples to obtain gene expression levels for RPM and RPR2 tumors. RPM time-series transition cells (n = 8) are annotated by timepoint and combined into a separate single raw count matrix, with counts as reads per kilobase million (RPKM).

**Human genomics data analysis—**We analyzed publically available transcriptome and whole exome sequencing (WES) data of human primary SCLC tumor samples (n = 81, (George et al., 2015)) and human SCLC cell lines (n = 52, CCLE). Transcript expression values in Figure S3B are represented as $\log_2$(TPM+1). Due to the relatively low prevalence of genomic *MYC* amplifications in the SCLC patient samples (George et al., 2015), but a considerable number of patients with high *MYC* TPM values, *MYC* expression was used for sample stratification by setting a log2-TPM+1 cut-off to >6.0, independent of copy number variation. Analyses were performed in R statistical environment using the pheatmap function for data presentation and unbiased hierarchical clustering. Figures 6G is analyses of only human tumor RNA-seq data (n = 70), excluding *POU2F3*$^+$ samples (George et al., 2015).

Figures 7G, 7I, and 8A are analyses of human tumor (n = 70; George et al., 2015) and cell line (n = 48; CCLE) bulk RNA-seq data with matching WES data, excluding *POU2F3*$^+$

samples. Predicted *NOTCH*-status in Figures 7G, 7I, and 8A are based on publicly available WES mutation data for human tumors (George et al., 2015) and for human cell lines on CCLE. Samples with wild-type *NOTCH1, −2, −3*, and/or *−4* or with silent mutations are grouped together. Predicted non-damaging mutations include those with missense or in-frame deletions in *NOTCH1, −2, −3*, and/or *−4*, and predicted damaging mutations include those with frame-shift deletions, nonsense, or splice variant mutations in *NOTCH1, −2, −3*, and/or *−4*. *MYC*-high samples in Figure 7G represent the upper quartile of *MYC* transcript levels. NE score for each sample was determined by calculating Spearman correlations of 50 genes from the NE and Non-NE expression vectors as described in Zhang et al. 2018.

Figure 7H represents expanded analysis of human cell lines (n = 98, excluding *POU2F3*$^{+}$ samples) that had publicly available expression and matching mutation data on the newly developed SCLC_Cellminer databse (https://discover.nci.nih.gov/SclcCellMinerCDB/), as well as the same human tumors (n = 70) used in Figure 7G. Cell line expression was downloaded from the Global SCLC dataset on the SCLC_Cellminer database (Tlemsani et al., 2020). In this dataset, global expression, including *MYC* expression, was determined based on average z-score intensity of gene expression from 5 sources (NCI SCLC, CCLE, CTRP, GDSC and UTSW). NE score values were calculated from the Global SCLC Dataset by Pearson correlation of 50 genes in the NE and Non-NE score vectors as described in Zhang et al, 2018. Mutation data was available from 3 of these 5 sources (CCLE, GDSC, and UTSW) on Cellminer—as values representing “Likelihood for Loss of Function (LOF) Mutation”. Mutation data for the 98 human cell lines that had matching expression data in the Global SCLC dataset were taken and averaged across the 3 sources. Cell lines with mutations in *NOTCH1, −2, −3*, and/or *−4* that were estimated to have >0 likelihood for LOF were categorized as “Mutant”, all other samples were “WT”. To combine the 70 tumors from George et al. with the cell line data, z-score intensity of *MYC* expression across tumors was calculated. NE score values were determined by calculating the Pearson correlation of 50 genes (in log2(TPM+1) counts) in the NE and Non-NE score vectors as described in Zhang et al, 2018. Human tumors in 7G with predicted non-damaging or damaging mutations were categorized as “Mutant” in 7H, all other tumors were “WT”. Cell line and tumor samples were marked as *MYC*-high with a z-score >0. The Fisher’s exact test, two-tailed p value in Figure 7H was calculated using Graphpad 2 x 2 contingency table.

**CIBERSORT—**CIBERSORT is a tool developed by (Newman et al., 2015) at Stanford University that uses expression data input to estimate abundances of specified cell types in a bulk cell population, such as a bulk tumor. We utilized the interactive user interface of CIBERSORT at https://cibersort.stanford.edu. Gene signatures were automatically determined by the software from a provided sample file with a matching phenotype class file (Table S5). For this sample file and class file, the bulk RNA-seq data from 8 RPM transition timepoints (adjacent groups of 2 representing 1 of 4 timepoint signatures) were input to derive a signature gene expression matrix with 4 “cell types” representing Day 3-5, Day 7-10, Day 12-14, and Day 19-21 timepoints. Mixture files of bulk sample data included a single normalized counts file from bulk RNA-seq of n = 10 RPM GEMM tumors infected with Ad-CGRP-Cre (Figure 5C), or a single file of human SCLC bulk RNA-seq data (n = 70 tumors, n = 48 SCLC cell lines from CCLE, not including *POU2F3*$^{+}$ samples) in TPM

counts (Figures 8A-8C). For each run, 100 permutations were performed. Relative and absolute modes were run together, with quantile normalization disabled for RNA-seq data, kappa = 999, q value cut-off = 0.3, and 50-150 barcode genes considered when building the signature matrix.

**Single cell RNA-seq sample and library prep—**For single-cell RNA seq (scRNA-seq) of bulk RPM-Cas9 tumors, ~100-200 mg live or viably frozen tumor (in 95% FBS, 5% DMSO) was digested into a single-cell suspension following the same methods described above, in “Primary tumor cell isolation”, for generation of RPM and RPR2 time-series transition cells. Single cell populations from bulk tumors were immediately processed for sequencing following digestion. Live RPM time-series transition cells underwent library preparation directly from cell culture at days 4 (n = 2 of the same sample), 7, 11, 14, 17, and 21 post-digestion for a total of 6 distinct transition timepoints. RPM time-series transition cells were made into single cell suspensions through gentle enzymatic dissociation of cell clusters with Accutase (Innovative Cell Tech Inc, cat#AT104-500) for 10-20 min with constant shaking at RT. The liver biopsy specimen collected from a patient with relapsed SCLC was viably frozen before being made into a single cell suspension through gentle enzymatic digestion with type 1A collagenase (Sigma, cat#C9891) for 10-20 min with shaking at RT. The cell suspension was filtered through a 40 micron cell filter and the resulting single cell suspension was flow sorted to isolate viable cells using the Zombie Green viability fluorescent dye (Biolegend, cat#423111) on a Sony Synergy SY3200 cell sorter. All single-cell suspensions were further prepared for sequencing according to 10x Chromium platform protocols found on: https://support.10xgenomics.com/single-cell-gene-expression.

In brief, the Chromium Single Cell Gene Expression Solution with 3' chemistry, version 3 (PN-1000075) was used to barcode individual cells with 16 bp 10x barcodes and to tag cell specific transcript molecules with 10 bp Unique Molecular Identifiers (UMIs) according to the manufacturer’s instructions. The following protocol was performed at High-Throughput Genomics Shared Resources at Huntsman Cancer Institute, University of Utah. Cells were suspended in PBS with 1.0% bovine serum albumin. The cell suspension was filtered through a 40 micron cell strainer. Viability and cell count were assessed on a Countess I (Thermo Scientific). Suspensions were equilibrated to targeted cell recovery of 8,000 cells. 10x Gel Beads and reverse transcription reagents were added and cell suspensions were loaded to the Chromium Single Cell Controller (PN-120263) to form Gel-Bead-In EMulsions (GEMs)—the micro-droplets. Within individual GEMs, cDNA generated from captured and barcoded mRNA was synthesized by reverse transcription at the setting of 53°C for 45 min followed by 85°C for 5 min. Subsequent A tailing, end repair, adaptor ligation and sample indexing were performed in bulk according to the manufacturer’s instructions. The resulting barcoding libraries were assessed on Agilent D1000 ScreenTape on an Agilent Technology 2200 TapeStation system and quantified by quantitative PCR using KAPA Biosystems Library Quantification Kit for Illumina Platforms (KK4842). Multiple libraries were then normalized and sequenced on NovaSeq 6000 with 2x150 paired-end (PE) mode. Single cell 5' PE chemistry was utilized for the SCLC liver biopsy

specimen, which was sequenced according to manufacturer's instructions at the McDonnell Genome Institute, Washington University School of Medicine in St. Louis.

### 10X single cell RNA-seq data processing

**Demultiplexing and data alignment:** Single cell RNA-seq data from the RPM time-series and bulk tumors were demultiplexed with 10x cellranger mkfastq version 3.1.0 to create fastq files with the I1 sample index, R1 cell barcode+UMI and R2 sequence. Reads were aligned to the mouse genome (mm10 custom EGFP+Cas9 reference v3.0.0) and count barcodes and UMIs were generated using cellranger count 3.1.0 with expected-cells set to 8000 per library. Single-cell RNA-seq data from the human biopsy were demultiplexed with 10x cellranger mkfastq version 3.0.1 and reads were aligned to the human genome (GRCh38). QC reporting, clustering, dimension reduction, and differential gene expression analysis using default parameter inputs (https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/using/reanalyze) were performed for primary data evaluation in 10x Genomics' Cell Loupe Browser (v3.1.1). Primary QC reporting on aggregate time-series cells revealed capture of 31,519 cells total with 36,854 post-normalization mean reads per cell and 3,856 median genes per cell—averaging 1,161,920,045 total reads or ~165 million reads per sample timepoint submission. Primary QC reporting on RPM1 tumor revealed capture of 5,367 cells total with 41,503 mean reads per cell and 1,810 median genes per cell or ~222 million reads total. Primary QC reporting on RPM2 tumor revealed capture of 2,961 cells total with 90,733 mean reads per cell and 995 median genes per cell or ~269 million reads total. Primary QC reporting on RPM3 tumor revealed capture of 998 total cells with 222,206 mean reads per cell and 1,046 median genes per cell or ~222 million reads total. Primary QC reporting on RPM4 tumor revealed capture of 2,129 total cells with 101,198 mean reads per cell and 1,208 median genes per cell or ~215 million reads total. Lastly, primary QC reporting on the human liver biopsy revealed capture of 1,081 total cells with 637,671 mean reads per cell and 1,859 median genes per cell or ~689 million reads total. For further details of the primary Cell Ranger data processing, see: https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/algorithms/overview.

**Filtering of non-tumor cell populations:** Data from RPM time-series and bulk tumors were visualized by t-distributed stochastic neighbor embedding (tSNE) in 10x Genomics' Cell Loupe Browser. Cell types were identified based on established cell type signatures derived from published single-cell RNA sequencing data (Treutlein et al., 2014; Xie et al., 2018; Zilionis et al., 2019). Low quality cell clusters were identified based on Loupe differential gene expression analysis, with clusters predominantly marked by mitochondrial genes tagged as low quality (Table S2). High to low gene expression in Figures S4A and S4E represent summed expression values of all cell-type-specific genes (Table S2). Tumor cells were identified and respective cell barcodes were exported from the Loupe browser to perform all tumor-cell-specific downstream analyses. Tumor cell barcodes from the time-series and bulk tumors can be found in Table S2. Cell type assignment was validated by visualization of cell types via Seurat dot plots to confirm enrichment of cell-type specific gene signatures per population (Figures S4A, S4E).

All cells from the human SCLC liver biopsy were imported directly into Seurat for cell-type assignment. Data dimensionality of the human biopsy was determined based on the ElbowPlot function to visualize data variance. To represent the majority of total data variance, n = 10 principal components were used for initial tSNE clustering to determine cell types. Unbiased tSNE clustering was performed and revealed 12 clusters (Figure S7A). Expression of cell-type-specific gene signatures (for tumor markers, endothelial cells, hematopoietic cells, and immune cells) were visualized per cluster using Seurat dot plot. From this analysis, clusters 0, 1, 4, and 6 were defined as tumor cells, clusters 2, 3, 5, 8, and 11 were identified as potential immune cells, clusters 9 and 10 as potential endothelial, and cluster 7 as potential hematopoietic. All tumor cells (n = 1021) were subsetted as a separate Seurat object for downstream analysis.

**Initial quality control:** Basic quality control was performed on all identified mouse tumor cells with the following parameters. For each 10X library, aligned matrices were converted to Single Cell Experiment objects. All cells with high mitochondrial content, low library size (summed counts), or low gene content (number of genes>0) per cell (exceeding 5 absolute deviations from median of each metric per library) were flagged as outlying cells. Outlying cells were visualized per 10X library using the Scater package (McCarthy et al., 2017). Cells tagged as low quality were dropped from analyses, and remaining cells were processed further following the Simple Single Cell workflow. In brief, to remove cell-specific biases, cells were clustered based on genes with average expression in the top 50% and size factors per cluster were determined using the deconvolution method (Lun et al., 2016). Normalized log expression values were then determined for all cells. Cell-cycle phase assignment was then performed using a robust and commonly-used cell-cycle assignment method specific to single-cell transcriptomic data (Scialdone et al., 2015). Finally, doublet distribution was simulated by adding random cells together and comparing doublet density estimates to original cell density estimates. Cells with doublet scores exceeding 20 median absolute deviations per library were marked as putative doublets and additionally excluded from downstream analyses. Following tumor cell identification and initial QC, transition timepoints included 15,434 high-quality tumor cells of 31,519 total cells, RPM1 included 2,107 high-quality tumor cells of 5,367 total cells, RPM2 included 809 high-quality tumor cells of 2,961 total cells, RPM3 included 248 high-quality tumor cells of 998 total cells, and RPM4 included included 772 high-quality tumor cells of 2,129 total cells. Tumor cell barcodes passing QC can be found in Table S2.

Human SCLC liver biopsy cells were subject to intial QC in Seurat. Cells passing QC included those with >100 and <10,000 genes detected per cell and with less than 25% mitochondrial content. Following tumor cell identification and initial QC, the biopsy included 1021 high-quality tumor cells of 2,201 total cells.

**Calculation of a NE and MYC-ChIP scores:** Following initial QC, pre-processed matrices were called into the R statistical computing environment (v3.5.3) for cell signature scoring (NE, MYC ChIP, and Cell Cycling/Prolif), tSNE, pseudotime, and differential gene expression (DGE) analyses using Monocle2 (tSNE and pseudotime) (Qiu et al., 2017; Trapnell et al., 2014) and Seurat (DGE) (Butler et al., 2018; Stuart et al., 2019) algorithms,

along with the Scater package (McCarthy et al., 2017) for NE and MYC-ChIP Score data visualization. NE score for cells of each single-cell RNA-seq library were determined based on Spearman correlation with an established 50-gene NE vs non-NE expression vector derived from (Zhang et al., 2018), where ~41 NE and ~87 non-NE human cell lines were used to identify a core 50-gene signature comprised of 25 NE genes and 25 non-NE genes that robustly predict NE phenotype (Table S3). NE score was added as metadata to Single Cell Experiment objects and visualized using the Scater plotColData function before conversion to CellDataSet for Monocle2 analyses.

MYC ChIP score was derived as follows. From all established MYC ChIP target genes according to ChIP-seq of four bulk RPM tumors, genes at least 3-fold up in RPM vs RPR2 tumors by bulk RNA-sequencing were analyzed in human MYC-high vs MYC-low tumors (George et al., 2015). Of those target genes, only the top-50 differentially expressed genes in MYC-high vs MYC-low human tumors were used in calculating a conserved MYC ChIP score (Table S3). MYC ChIP scores were computed for each cell per single-cell RNA-seq library by converting expression data to a Seurat object, and utilizing the AddModuleScore function. Seurat objects were converted back to Single Cell Experiment objects and MYC ChIP score was visualized per library using the Scater plotColData function before proceeding to Monocle2 analyses.

**Regression of cycling/proliferation effects:** Variation due to expression of cell cycle/proliferation genes was regressed out for all clustering and pseudotime analyses of RPM time-series and bulk tumor cells. A cell cycling/proliferation score was generated and added as metadata to Seurat objects using AddModuleScore based on the following cell cycle or proliferation related gene signatures from MSigDB: “Ben-Porath Proliferation”, “Ben-Porath Cycling Genes”, “KEGG cell cycle” (Table S2). Cycling/proliferation was regressed out in Monocle2 reduceDimension analyses using the residualModelFormulaStr input option and tSNE clustering and cell ordering was not based on cycling/proliferation effects. Proper regression of cell cycling/proliferation effects was validated based on cell cycle phase assignment according to established single-cell-RNA seq-specific methods (cyclone) (Scialdone et al., 2015). Following cell cycle regression based on generic cycling/proliferation gene sets above, there was no detected enrichment for distinct cell cycle phases in clusters of tSNE (Figure S4B) and no significant correlation of cyclone-assigned cell-cycle phase with cell ordering in pseudotime (as determined by Pearson correlation of pseudotime coordinate vs likelihood of G1, S, G2M score vectors for all cells in the trajectory analysis)—confirming appropriate regression of cycling/proliferation effects.

**tSNE Clustering:** tSNE dimension reduction and visualization on timepoint data was performed using the Monocle2 package (Qiu et al., 2017; Trapnell et al., 2014). Cells were grouped with clusterCells according to global expression profiles and clustered in an unsupervised manner with setOrderingFilter based on genes with mean_expression >= 0.1. Dimensionality reduction was performed using Monocle2’s reduceDimension function with max_components set to 2, num_dim set to 10, and cycling/prolif regressed out using residualModelFormulaStr. 10 principle components (PCs) were sufficient to describe the biological variation of interest and led to coherent cell groupings as shown. Using more PCs

may encode higher-order biological complexities, but also minor variations and biological and technical noise. Interpretation of tSNE embeddings based on a larger number of PCs were thus confounded.

For tSNE clustering of the human SCLC liver biopsy tumor cells (identified as explained above using Seurat), cells were grouped with clusterCells according to global expression profiles and clustered in an unsupervised manner with setOrderingFilter based on genes with mean_expression >= 0.1. Dimensionality reduction was performed using Monocle2's reduceDimension function with max_components set to 2 and num_dim set to 5. 5 PCs were sufficient to describe the majority of human tumor cell variation, as validated by the ElbowPlot function used to visualize data variability in Seurat (Figure 8F).

**Differential gene expression analyses:** Differentially-expressed genes (DEGs) per transition timepoint were determined using the FindAllMarkers function in Seurat (Butler et al., 2018; Stuart et al., 2019) with MAST test, which identifies DEGs between two groups of cells using a hurdle model tailored to scRNA-seq data. With this method, only genes with at least 0.25-log fold increase between a group and its comparators was taken as a marker gene of that group (Figure 6A and Table S3). DEGs per timepoints were visualized using Seurat's DoHeatmap function. Notch machinery expression per timepoint was also visualized using Seurat's DoHeatmap function with maximum and minimum thresholds set to 1. ENRICHR analyses (Kuleshov et al., 2016) were performed on all genes at least 0.25 log fold change increased between named groups shown in key of Figure 6B, as determined by Seurat's FindMarkers function (Table S3).

**Pseudotime analyses:** To analyze the trajectory development of RPM transition and tumor cells (Figure 4D, 4E, 4F, 5D, 6C, S4B, S4F, S5C), an unsupervised pseudotemporal analysis was performed using Monocle2 (Qiu et al., 2017; Trapnell et al., 2014). Timepoint data (n = 6) and bulk tumor libraries (n = 4) were combined into a single Seurat object using the merge function, following initial QC of each library as described. MYC-ChIP, cell cycling/prolif, and NE scores were added to the merged Seurat object before conversion to a Monocle2 CellDataSet. To predict biological trajectories, Monocle2 uses DDRTree, a reversed graph embedding algorithm, to reduce the high-dimensional scRNA-seq data space, and predict how cells progress through a given biological process based on global gene expression levels. Monocle2 offers ideal unsupervised pseudotime analysis for this study as it predicts branch points and trajectory states without cell fate input information. Monocle2's accuracy has been tested in multiple biological contexts and has yielded more robust pseudotimes compared to algorithms predicting pseudotime based on pairwise geodesic distances between cells rather than explicit tree construction (Qiu et al., 2017). Following size factor and dispersion estimates, trajectory ordering genes were called by mean_expression >=0.1 and dispersion_empirical >=1. Data dimensionality was reduced using the reduceDimension function with max_components set to 2, reduction_method set to DDRTree, and cycling/prolif regressed out with residualModelFormulaStr input. DEGs across pseudotime were determined using the differentialGeneTest function on all genes detected by the detectGenes function with min_expr set to 1 and expressed in >100 cells.

Resultant genes were ordered by q value and the top-500 genes changing in pseudotime were visualized using the plot_pseudotime_heatmap function (Table S2).

**Diffusion mapping:** Diffusion mapping utilized destiny, princurves and rgl packages all in R (v 3.5.2) similar to (Giraddi et al., 2018). For diffusion maps of RPM time-series and bulk tumors, genes present in more than 5 cells at more than 5 counts were included, and upper and lower bounds on cellular complexity (i.e. # of unique genes expressed) were imposed at 8000 and 2000, respectively, which excluded an upper 0.7% of samples and a lower 7.6% of samples from analysis. As in Giraddi et al., high local variance genes were used for diffusion mapping of cell-cell relationships and lambda values of a fit principal curve were considered as pseudotime coordinates for cell ordering. Briefly, the diffusion trajectory was determined by first transforming the DC1,6 and 9 components to a 'flattened' coordinate graph as these DCs showed the greatest sequential separation of sequential culture stages (among DC's 1:10). The coordinates of the flattened graph are compound DCs designated DC1' and DC2', that were derived as follows to track with stage specific diffusion mapping across the DM: DC1' = 0.5xlog(DC1+.1)+1.5+DC2'; DC2'=0.15+5xDC1+2xDC2+1.5xDC3. They represent uniform arithmetic modulation of DCs to form a square coordinate map. A principle curve was generated through the compound DC graph using the R package "PrinCurve" and positions along this curve were mapped as the diffusion trajectory. Diffusion map pseudotime assignment was directly compared to Monocle2 pseudotime assignment on RPM time-series and bulk tumors via Pearson correlation analysis in Rstudio (cor.test R function). Results were plotted using the ggscatter function from the "ggpubr" library (Figure S4M).

**GSEA—**Gene set enrichment analysis (GSEA) was performed using GSEA version 4.03 software with default parameters, classic enrichment, inclusion gene set size between 15 and 5000, and the phenotype permutation at 1,000 times. Normalized enrichment scores (NES) and p values are shown below each respective GSEA plot in the figures. For Figure 3D, GSEA was performed on pre-ranked gene lists representing log2 fold change expression of day 3-5 bulk RPM transition timepoints vs all later timepoints (Day 3-5 signature, Figure 3D), day 7-10 bulk RPM transition timepoints vs all other timepoints (Day 7-10 signature, Figure 3D), and Day 14-21 bulk RPM transition timepoints vs all earlier timepoints (Day 14-21 signature, Figure 3D) (Table S1). Gene signatures for hSCLC-A, hSCLC-N, and hSCLC-Y were determined by taking all genes increased per subtype by at least log2 fold change >1.5 vs all other subtypes (Table S1). Cell line and tumor samples were assigned by subtype according to unbiased grouping established in Figure S3A. Remaining GSEA included ranked gene expression files based on three late RPM transition timepoints (day 14, 19, 21) vs three early timepoints (day 3, 5, 7) (Figure S5D) (Tables S1 and S3) or on *NOTCH*-damaging-mutant vs *NOTCH*-wild-type (WT) or silent mutant human SCLC tumors (George et al., 2015) and CCLE cell lines (Figure 7I) (Table S4). A catalog of functional gene sets from Molecular Signature Database (MSigDB, version 6.2, July 2018, www.broad.mit.edu/gsea/msigdb/msigdb_index.html) was used for the "Kegg Notch signaling" (Figure S5D), "Rest targets NRSF_01" (Figure S5D), or "Notch signaling" (PID_Notch_Pathway) (Figure 7I) gene sets. "MYC ChIP score" gene set consists of the 50-gene MYC ChIP signature described in Methods section "10X single cell RNA-seq data

processing: *Calculation of a NE and MYC-ChIP scores"* (Table S3). All gene sets can be viewed in Tables S1, S3, and/or S4.

**Quantification and statistical analysis—**Remaining statistical analysis was performed using GraphPad Prism. Error bars show mean ± SEM unless otherwise specified. Significance was determined by Student's two-tailed unpaired t tests with 95% confidence intervals and p values <0.05 considered statistically significant, unless otherwise indicated. All statistical details are further described in respective figure legends. Additional statistical methods related to bioinformatic analyses can be found in Bioinformatic Analyses under Method Details. No statistical methods were used to predetermine sample sizes. Please see details of IHC quantifications in Immunohistochemistry section of Method Details.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

We thank members of the Oliver Lab for technical assistance and mouse colony management, R. Olsen, R. Dahlgren, and L. Houston for administrative support, A. Andersen for editorial guidance, and B. Dalley, K. Rondem, and C. Stubben for bioinformatics support. We are grateful to our late colleague Dr. Adi Gazdar for insightful discussions and inspiration for this work. We acknowledge support from the National Cancer Institute (NCI) of the National Institutes of Health (NIH) under award P30CA042014 to Huntsman Cancer Institute for the use of core facilities and shared resources, including Biorepository and Molecular Pathology, High-Throughput Genomics and Bioinformatic Analysis, and Flow Cytometry Shared Resources; we acknowledge support and resources from the Center for High Performance Computing at the University of Utah. T.G.O. was supported in part by NIH NCI (U01-CA231844, U24-CA213274 and R21-CA216504-01A1).

## REFERENCES

Angerer P, Haghverdi L, Buttner M, Theis FJ, Marr C, and Buettner F (2016). destiny: diffusion maps for large-scale single-cell data in R. Bioinformatics 32, 1241–1243. [PubMed: 26668002]

Augert A, Eastwood E, Ibrahim AH, Wu N, Grunblatt E, Basom R, Liggitt D, Eaton KD, Martins R, Poirier JT et al. (2019). Targeting NOTCH activation in small cell lung cancer through LSD1 inhibition. Sci Signal 12.

Borromeo MD, Savage TK, Kollipara RK, He M, Augustyn A, Osborne JK, Girard L, Minna JD, Gazdar AF, Cobb MH, and Johnson JE (2016). ASCL1 and NEUROD1 Reveal Heterogeneity in Pulmonary Neuroendocrine Tumors and Regulate Distinct Genetic Programs. Cell Rep 16, 1259–1272. [PubMed: 27452466]

Bragelmann J, Bohm S, Guthrie MR, Mollaoglu G, Oliver TG, and Sos ML (2017). Family matters: How MYC family oncogenes impact small cell lung cancer. Cell Cycle 16, 1489–1498. [PubMed: 28737478]

Butler A, Hoffman P, Smibert P, Papalexi E, and Satija R (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36, 411–420. [PubMed: 29608179]

Cardnell RJ, Li L, Sen T, Bara R, Tong P, Fujimoto J, Ireland AS, Guthrie MR, Bheddah S, Banerjee U, et al. (2017). Protein expression of TTF1 and cMYC define distinct molecular subgroups of small cell lung cancer with unique vulnerabilities to aurora kinase inhibition, DLL3 targeting, and other targeted therapies. Oncotarget 8, 73419–73432. [PubMed: 29088717]

Carney DN, Gazdar AF, Bepler G, Guccion JG, Marangos PJ, Moody TW, Zweig MH, and Minna JD (1985). Establishment and identification of small cell lung cancer cell lines having classic and variant features. Cancer Res 45, 2913–2923. [PubMed: 2985257]

Chalishazar MD, Wait SJ, Huang F, Ireland AS, Mukhopadhyay A, Lee Y, Schuman SS, Guthrie MR, Berrett KC, Vahrenkamp JM, et al. (2019). MYC-Driven Small-Cell Lung Cancer is Metabolically Distinct and Vulnerable to Arginine Depletion. Clin Cancer Res 25, 5107–5121. [PubMed: 31164374]

Christensen CL, Kwiatkowski N, Abraham BJ, Carretero J, Al-Shahrour F, Zhang T, Chipumuro E, Herter-Sprie GS, Akbay EA, Altabef A, et al. (2014). Targeting transcriptional addictions in small cell lung cancer with a covalent CDK7 inhibitor. Cancer Cell 26, 909–922. [PubMed: 25490451]

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, and Ruden DM (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6, 80–92. [PubMed: 22728672]

Coifman RR, Lafon S, Lee AB, Maggioni M, Nadler B, Warner F, and Zucker SW (2005). Geometric diffusions as a tool for harmonic analysis and structure definition of data: multiscale methods. Proc Natl Acad Sci U S A 102, 7432–7437. [PubMed: 15899969]

Collisson EA, Campbell JD, Brooks AN, Berger AH, Lee W, Chmielecki J, Beer DG, Cope L, Creighton CJ, Danilova L, et al. (2014). Comprehensive molecular profiling of lung adenocarcinoma. Nature 511, 543–550. [PubMed: 25079552]

Dammert MA, Bragelmann J, Olsen RR, Bohm S, Monhasery N, Whitney CP, Chalishazar MD, Tumbrink HL, Guthrie MR, Klein S, et al. (2019). MYC paralog-dependent apoptotic priming orchestrates a spectrum of vulnerabilities in small cell lung cancer. Nat Commun 10, 3485. [PubMed: 31375684]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Dooley AL, Winslow MM, Chiang DY, Banerji S, Stransky N, Dayton TL, Snyder EL, Senna S, Whittaker CA, Bronson RT, et al. (2011). Nuclear factor I/B is an oncogene in small cell lung cancer. Genes Dev 25, 1470–1475. [PubMed: 21764851]

DuPage M, Dooley AL, and Jacks T (2009). Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. Nat Protoc 4, 1064–1072. [PubMed: 19561589]

Ewels P, Magnusson M, Lundin S, and Kaller M (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048. [PubMed: 27312411]

Farago AF, Yeap BY, Stanzione M, Hung YP, Heist RS, Marcoux JP, Zhong J, Rangachari D, Barbie DA, Phat S, et al. (2019). Combination Olaparib and Temozolomide in Relapsed Small-Cell Lung Cancer. Cancer Discov 9, 1372–1387. [PubMed: 31416802]

Flowers JL, Burton GV, Cox EB, McCarty KS Sr., Dent GA, Geisinger KR, and McCarty KS Jr. (1986). Use of monoclonal antiestrogen receptor antibody to evaluate estrogen receptor content in fine needle aspiration breast biopsies. Ann Surg 203, 250–254. [PubMed: 3954477]

Gazdar AF, Carney DN, Nau MM, and Minna JD (1985). Characterization of variant subclasses of cell lines derived from small cell lung cancer having distinctive biochemical, morphological, and growth properties. Cancer Res 45, 2924–2930. [PubMed: 2985258]

George J, Lim JS, Jang SJ, Cun Y, Ozretic L, Kong G, Leenders F, Lu X, Fernandez-Cuesta L, Bosco G, et al. (2015). Comprehensive genomic profiles of small cell lung cancer. Nature 524, 47–53. [PubMed: 26168399]

Giraddi RR, Chung C-Y, Heinz RE, Balcioglu O, Novotny M, Trejo CL, Dravis C, Hagos BM, Mehrabad EM, Rodewald LW, et al. (2018). Single-Cell Transcriptomes Distinguish Stem Cell State Changes and Lineage Specification Programs in Early Mammary Gland Development. Cell Rep 24, 1653–1666.e1657. [PubMed: 30089273]

Gratton MO, Torban E, Jasmin SB, Theriault FM, German MS, and Stifani S (2003). Hes6 promotes cortical neurogenesis and inhibits Hes1 transcription repression activity by multiple mechanisms. Mol Cell Biol 23, 6922–6935. [PubMed: 12972610]

Horie M, Saito A, Ohshima M, Suzuki HI, and Nagase T (2016). YAP and TAZ modulate cell phenotype in a subset of small cell lung cancer. Cancer Sci 107, 1755–1766. [PubMed: 27627196]

Hovestadt V, Smith KS, Bihannic L, Filbin MG, Shaw ML, Baumgartner A, DeWitt JC, Groves A, Mayr L, Weisman HR, et al. (2019). Resolving medulloblastoma cellular architecture by single-cell genomics. Nature 572, 74–79. [PubMed: 31341285]

Huang F, Ni M, Chalishazar MD, Huffman KE, Kim J, Cai L, Shi X, Cai F, Zacharias LG, Ireland AS, et al. (2018a). Inosine Monophosphate Dehydrogenase Dependence in a Subset of Small Cell Lung Cancers. Cell Metab 28, 369–382.e365. [PubMed: 30043754]

Huang YH, Klingbeil O, He XY, Wu XS, Arun G, Lu B, Somerville TDD, Milazzo JP, Wilkinson JE, Demerdash OE, et al. (2018b). POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer. Genes Dev 32, 915–928. [PubMed: 29945888]

Huijbers IJ, Bin Ali R, Pritchard C, Cozijnsen M, Kwon MC, Proost N, Song JY, de Vries H, Badhai J, Sutherland K, et al. (2014). Rapid target gene validation in complex cancer mouse models using re-derived embryonic stem cells. EMBO Mol Med 6, 212–225. [PubMed: 24401838]

Ito T, Udaka N, Ikeda M, Yazawa T, Kageyama R, and Kitamura H (2001). Significance of proneural basic helix-loop-helix transcription factors in neuroendocrine differentiation of fetal lung epithelial cells and lung carcinoma cells. Histol Histopathol 16, 335–343. [PubMed: 11193209]

Jia D, Augert A, Kim DW, Eastwood E, Wu N, Ibrahim AH, Kim KB, Dunn CT, Pillai SPS, Gazdar AF, et al. (2018). Crebbp Loss Drives Small Cell Lung Cancer and Increases Sensitivity to HDAC Inhibition. Cancer Discov 8, 1422–1437. [PubMed: 30181244]

Johnson BE, Battey J, Linnoila I, Becker KL, Makuch RW, Snider RH, Carney DN, and Minna JD (1986). Changes in the phenotype of human small cell lung cancer cell lines after transfection and expression of the c-myc proto-oncogene. J Clin Invest 78, 525–532. [PubMed: 3016030]

Kim DW, Wu N, Kim YC, Cheng PF, Basom R, Kim D, Dunn CT, Lee AY, Kim K, Lee CS, et al. (2016). Genetic requirement for Mycl and efficacy of RNA Pol I inhibition in mouse models of small cell lung cancer. Genes Dev 30, 1289–1299. [PubMed: 27298335]

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 44, W90–97. [PubMed: 27141961]

Li H, and Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. [PubMed: 19451168]

Liao Y, Smyth GK, and Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. [PubMed: 24227677]

Lim JS, Ibaseta A, Fischer MM, Cancilla B, O'Young G, Cristea S, Luca VC, Yang D, Jahchan NS, Hamard C, et al. (2017). Intratumoural heterogeneity generated by Notch signalling promotes small-cell lung cancer. Nature 545, 360–364. [PubMed: 28489825]

Lin JJ, Cardarella S, Lydon CA, Dahlberg SE, Jackman DM, Janne PA, and Johnson BE (2016). Five-Year Survival in EGFR-Mutant Metastatic Lung Adenocarcinoma Treated with EGFR-TKIs. J Thorac Oncol 11, 556–565. [PubMed: 26724471]

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

Lun AT, Bach K, and Marioni JC (2016). Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. Genome Biol 17, 75. [PubMed: 27122128]

Martin M (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17.

Matsumoto A, Onoyama I, Sunabori T, Kageyama R, Okano H, and Nakayama KI (2011). Fbxw7-dependent degradation of Notch is required for control of "stemness" and neuronal-glial differentiation in neural stem cells. J Biol Chem 286, 13754–13764. [PubMed: 21349854]

Mayr C, and Bartel DP (2009). Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell 138, 673–684. [PubMed: 19703394]

McCarthy DJ, Campbell KR, Lun AT, and Wills QF (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 33, 1179–1186. [PubMed: 28088763]

McColl K, Wildey G, Sakre N, Lipka MB, Behtaj M, Kresak A, Chen Y, Yang M, Velcheti V, Fu P, and Dowlati A (2017). Reciprocal expression of INSM1 and YAP1 defines subgroups in small cell lung cancer. Oncotarget 8, 73745–73756. [PubMed: 29088741]

McFadden DG, Papagiannakopoulos T, Taylor-Weiner A, Stewart C, Carter SL, Cibulskis K, Bhutkar A, McKenna A, Dooley A, Vernon A, et al. (2014). Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. Cell 156, 1298–1311. [PubMed: 24630729]

Meuwissen R, Linn SC, Linnoila RI, Zevenhoven J, Mooi WJ, and Berns A (2003). Induction of small cell lung cancer by somatic inactivation of both Trp53 and Rb1 in a conditional mouse model. Cancer Cell 4, 181–189. [PubMed: 14522252]

Mollaoglu G, Guthrie MR, Bohm S, Bragelmann J, Can I, Ballieu PM, Marx A, George J, Heinen C, Chalishazar MD, et al. (2017). MYC Drives Progression of Small Cell Lung Cancer to a Variant Neuroendocrine Subtype with Vulnerability to Aurora Kinase Inhibition. Cancer Cell 31, 270–285. [PubMed: 28089889]

Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, and Alizadeh AA (2015). Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 12, 453–457. [PubMed: 25822800]

Oliver TG, Meylan E, Chang GP, Xue W, Burke JR, Humpton TJ, Hubbard D, Bhutkar A, and Jacks T (2011). Caspase-2-mediated cleavage of Mdm2 creates a p53-induced positive feedback loop. Mol Cell 43, 57–71. [PubMed: 21726810]

Oser MG, Sabet AH, Gao W, Chakraborty AA, Schinzel AC, Jennings RB, Fonseca R, Bonal DM, Booker MA, Flaifel A, et al. (2019). The KDM5A/RBP2 histone demethylase represses NOTCH signaling to sustain neuroendocrine differentiation and promote small cell lung cancer tumorigenesis. Genes Dev 33, 1718–1738. [PubMed: 31727771]

Ouadah Y, Rojas ER, Riordan DP, Capostagno S, Kuo CS, and Krasnow MA (2019). Rare Pulmonary Neuroendocrine Cells Are Stem Cells Regulated by Rb, p53, and Notch. Cell 179, 403–416.e423. [PubMed: 31585080]

Owonikoko TK, Niu H, Nackaerts K, Csoszi T, Ostoros G, Mark Z, Baik C, Joy AA, Chouaid C, Jaime JC, et al. (2019). Randomized Phase II Study of Paclitaxel plus Alisertib versus Paclitaxel plus Placebo as Second-Line Therapy for Small-Cell Lung Cancer: Primary and Correlative Biomarker Analyses. J Thorac Oncol 19, 33572–33575.

Park JW, Lee JK, Phillips JW, Huang P, Cheng D, Huang J, and Witte ON (2016). Prostate epithelial cell of origin determines cancer differentiation state in an organoid transformation assay. Proc Natl Acad Sci U S A 113, 4482–4487. [PubMed: 27044116]

Park KS, Liang MC, Raiser DM, Zamponi R, Roach RR, Curtis SJ, Walton Z, Schaffer BE, Roake CM, Zmoos AF, et al. (2011). Characterization of the cell of origin for small cell lung cancer. Cell Cycle 10, 2806–2815. [PubMed: 21822053]

Patel AS, Yoo S, Kongaelin RJ, Sato T, Fridrikh M, Sinha A, Nudelman G, Powell C, Beasley M, Zhu J, and Watanabe H (2019). Integrative genomic and epigenomic analyses identify a distinct role of c-Myc and L-Myc for lineage determination in small cell lung cancer. bioRxiv, doi: 10.1101/852939.

Peifer M, Fernandez-Cuesta L, Sos ML, George J, Seidel D, Kasper LH, Plenker D, Leenders F, Sun R, Zander T, et al. (2012). Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. Nat Genet 44, 1104–1110. [PubMed: 22941188]

Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, Dahlman JE, Parnas O, Eisenhaure TM, Jovanovic M, et al. (2014). CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell 159, 440–455. [PubMed: 25263330]

Poirier JT, Gardner EE, Connis N, Moreira AL, de Stanchina E, Hann CL, and Rudin CM (2015). DNA methylation in small cell lung cancer defines distinct disease subtypes and correlates with high expression of EZH2. Oncogene 34, 5869–5878. [PubMed: 25746006]

Poirier JT, George J, Owonikoko TK, Berns A, Brambilla E, Byers LA, Carbone D, Chen HJ, Christensen CL, Dive C, et al. (2020). New Approaches to SCLC Therapy: From the Laboratory to the Clinic. J Thorac Oncol.

Qiu X, Hill A, Packer J, Lin D, Ma YA, and Trapnell C (2017). Single-cell mRNA quantification and differential analysis with Census. Nat Methods 14, 309–315. [PubMed: 28114287]

Remon J, Ahn MJ, Girard N, Johnson M, Kim DW, Lopes G, Pillai RN, Solomon B, Villacampa G, and Zhou Q (2019). Advanced-Stage Non-Small Cell Lung Cancer: Advances in Thoracic Oncology 2018. J Thorac Oncol 14, 1134–1155. [PubMed: 31002952]

Rudin CM, Durinck S, Stawiski EW, Poirier JT, Modrusan Z, Shames DS, Bergbower EA, Guan Y, Shin J, Guillory J, et al. (2012). Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. Nat Genet 44, 1111–1116. [PubMed: 22941189]

Rudin CM, Poirier JT, Byers LA, Dive C, Dowlati A, George J, Heymach JV, Johnson JE, Lehman JM, MacPherson D, et al. (2019). Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. Nat Rev Cancer 19, 289–297. [PubMed: 30926931]

Saunders LR, Bankovich AJ, Anderson WC, Aujay MA, Bheddah S, Black K, Desai R, Escarpe PA, Hampl J, Laysang A, et al. (2015). A DLL3-targeted antibody-drug conjugate eradicates high-grade pulmonary neuroendocrine tumor-initiating cells in vivo. Sci Transl Med 7, 302ra136.

Schaffer BE, Park KS, Yiu G, Conklin JF, Lin C, Burkhart DL, Karnezis AN, Sweet-Cordero EA, and Sage J (2010). Loss of p130 accelerates tumor development in a mouse model for human small-cell lung carcinoma. Cancer Res 70, 3877–3883. [PubMed: 20406986]

Scialdone A, Natarajan KN, Saraiva LR, Proserpio V, Teichmann SA, Stegle O, Marioni JC, and Buettner F (2015). Computational assignment of cell-cycle stage from single-cell transcriptome data. Methods 85, 54–61. [PubMed: 26142758]

Sen T, Tong P, Stewart CA, Cristea S, Valliani A, Shames DS, Redwood AB, Fan YH, Li L, Glisson BS, et al. (2017). CHK1 Inhibition in Small-Cell Lung Cancer Produces Single-Agent Activity in Biomarker-Defined Disease Subsets and Combination Activity with Cisplatin or Olaparib. Cancer Res 77, 3870–3884. [PubMed: 28490518]

Shen R, and Seshan VE (2016). FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. Nucleic Acids Res 44, e131. [PubMed: 27270079]

Simpson KL, Stoney R, Frese KK, Simms N, Rowe W, Pearce SP, Humphrey S, Booth L, Morgan D, Dynowski M, et al. (2020). A biobank of small cell lung cancer CDX models elucidates inter- and intratumoral phenotypic heterogeneity. Nature Cancer 1, 437–451.

Stewart CA, Gay CM, Xi Y, Sivajothi S, Sivakamasundari V, Fujimoto J, Bolisetty M, Hartsfield PM, Balasubramaniyan V, Chalishazar MD, et al. (2020). Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. Nature Cancer 1, 423–436.

Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, Sabatini DM, Chen IS, Hahn WC, Sharp PA, et al. (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. Rna 9, 493–501. [PubMed: 12649500]

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, and Satija R (2019). Comprehensive Integration of Single-Cell Data. Cell 177, 1888–1902.e1821. [PubMed: 31178118]

Sutherland KD, Proost N, Brouns I, Adriaensen D, Song JY, and Berns A (2011). Cell of origin of small cell lung cancer: inactivation of Trp53 and Rb1 in distinct cell types of adult mouse lung. Cancer Cell 19, 754–764. [PubMed: 21665149]

Tlemsani C, Pongor L, Girard L, Roper N, Elloumi F, Varma S, Luna A, Rajapakse VN, Sebastian R, Kohn KW, et al. (2020). SCLC_CellMiner: Integrated Genomics and Therapeutics Predictors of Small Cell Lung Cancer Cell Lines based on their genomic signatures. bioRxiv, doi: 10.1101/2020.03.09.980623.

Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, and Rinn JL (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 32, 381–386. [PubMed: 24658644]

Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH, Desai TJ, Krasnow MA, and Quake SR (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509, 371–375. [PubMed: 24739965]

Vo BT, Wolf E, Kawauchi D, Gebhardt A, Rehg JE, Finkelstein D, Walz S, Murphy BL, Youn YH, Han YG, et al. (2016). The Interaction of Myc with Miz1 Defines Medulloblastoma Subgroup Identity. Cancer Cell 29, 5–16. [PubMed: 26766587]

Wagner AH, Devarakonda S, Skidmore ZL, Krysiak K, Ramu A, Trani L, Kunisaki J, Masood A, Waqar SN, Spies NC, et al. (2018). Recurrent WNT pathway alterations are frequent in relapsed small cell lung cancer. Nat Commun 9, 3787. [PubMed: 30224629]

Wang ZA, Mitrofanova A, Bergren SK, Abate-Shen C, Cardiff RD, Califano A, and Shen MM (2013). Lineage analysis of basal epithelial cells reveals their unexpected plasticity and supports a cell-of-origin model for prostate cancer heterogeneity. Nat Cell Biol 15, 274–283. [PubMed: 23434823]

Wooten DJ, Groves SM, Tyson DR, Liu Q, Lim JS, Albert R, Lopez CF, Sage J, and Quaranta V (2019). Systems-level network modeling of Small Cell Lung Cancer subtypes identifies master regulators and destabilizers. PLoS Comput Biol 15, e1007343. [PubMed: 31671086]

Xie T, Wang Y, Deng N, Huang G, Taghavifar F, Geng Y, Liu N, Kulur V, Yao C, Chen P, et al. (2018). Single-Cell Deconvolution of Fibroblast Heterogeneity in Mouse Pulmonary Fibrosis. Cell Rep 22, 3625–3640. [PubMed: 29590628]

Zhang W, Girard L, Zhang YA, Haruki T, Papari-Zareei M, Stastny V, Ghayee HK, Pacak K, Oliver TG, Minna JD, and Gazdar AF (2018). Small cell lung cancer tumors and preclinical models display heterogeneity of neuroendocrine phenotypes. Transl Lung Cancer Res 7, 32–49. [PubMed: 29535911]

Zilionis R, Engblom C, Pfirschke C, Savova V, Zemmour D, Saatcioglu HD, Krishnan I, Maroni G, Meyerovitz CV, Kerwin CM, et al. (2019). Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. Immunity 50, 1317–1334.e1310. [PubMed: 30979687]

## SIGNIFICANCE

SCLC has historically been treated as a single disease, but is now recognized to comprise multiple molecular subtypes. We find that MYC directly activates NOTCH signaling to reprogram neuroendocrine SCLC from ASCL1$^{+}$ to NEUROD1$^{+}$ to YAP1$^{+}$ non-neuroendocrine states. Therefore, SCLC molecular subtypes are not distinct, but rather represent dynamic stages of MYC-driven tumor evolution. We find that individual human SCLC tumors are composed of multiple molecular subtypes. Given the reported unique therapeutic vulnerabilities of each subtype, we postulate that SCLC tumors represent a "moving therapeutic target" that may require more general, combinatorial, or plasticity-directed therapeutic approaches to combat this transcriptional flexibility. We speculate that molecular subsets of other cancer types may also represent dynamic stages of tumor evolution.

## Highlights

- Multiple SCLC molecular subtypes arise from a neuroendocrine cell of origin
- MYC drives the NEUROD1$^{+}$ and YAP1$^{+}$ subtypes of SCLC in a temporal evolution
- MYC directly activates NOTCH signaling to reprogram neuroendocrine fate
- Multiple SCLC molecular subtypes are present within individual human tumors

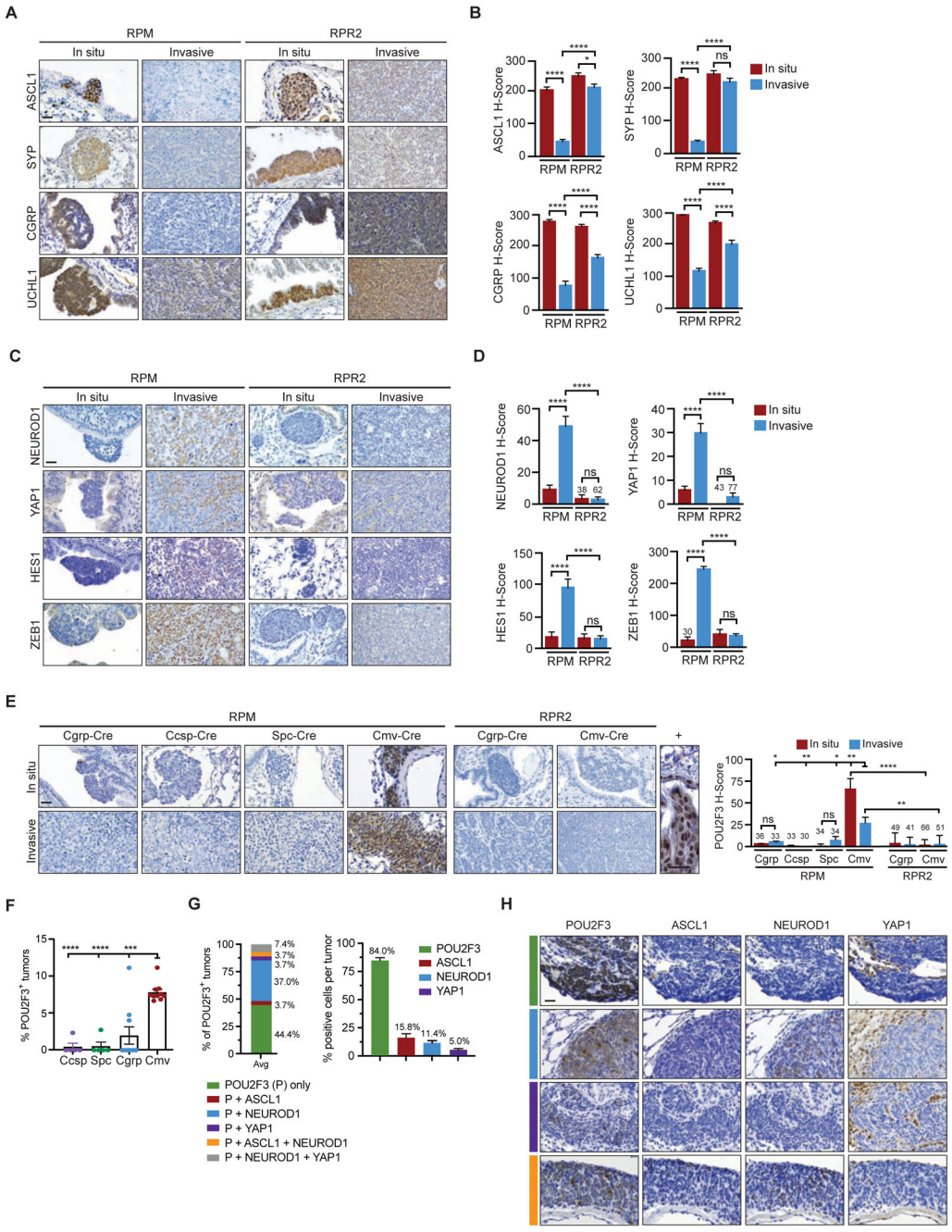



**Figure 1. MYC drives multiple SCLC molecular subtypes in vivo**

(A) Representative immunohistochemistry (IHC) for NE markers in early-stage (in situ) or invasive tumors in indicated GEMMs infected with Ad-Cgrp-Cre.

(B) IHC quantification from panel A.

(C) Representative IHC for non-NE markers in in situ or invasive tumors in indicated GEMMs infected with Ad-Cgrp-Cre.

(D) IHC quantification from panel C.

(E) Representative IHC for POU2F3 in in situ or invasive tumors in indicated GEMMs infected with cell-type-specific Cre viruses. Positive control (+) is adult mouse skin. IHC quantification in right panel.
(F) Percent RPM tumors per animal that are $POU2F3^+$ by IHC following cell-type-specific Cre viral infection. Indicated p values relative to Cmv.
(G) Left panel: Percent of $POU2F3^+$ (P) tumors expressing subtype markers analyzed by IHC from serial sections (n = 27 total tumors). Right panel: Average percentage of positive cells for each marker within individual $POU2F3^+$ tumors.
(H) Representative IHC from serial sections of $POU2F3^+$ tumors analyzed for SCLC subtype markers. Left color label indicates classification as in panel G.
Data from n = 6-9 mice per genotype, except Ccsp- and Spc-Cre mice are n = 5. Number of tumors scored by manual H-Score method is indicated within bar graphs. Scale bars, 25 μm. Mean +/− standard error of the mean (SEM), Student's two-tailed unpaired t test, * $p < 0.045$, ** $p < 0.009$, *** $p = 0.002$, **** $p < 0.0001$, ns = not significant. See also Figure S1.

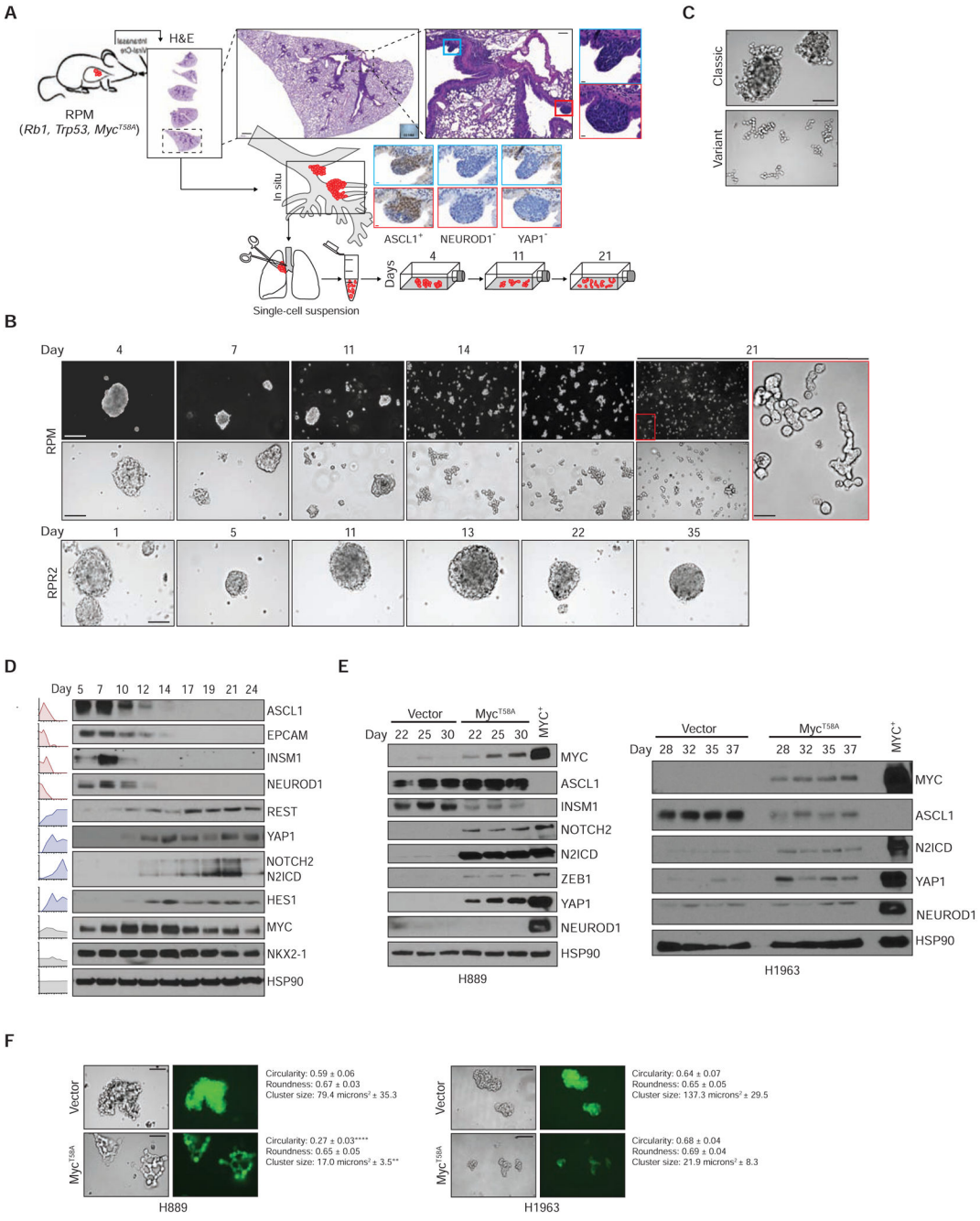



**Figure 2. MYC drives SCLC subtype evolution in vitro**

(A) Schematic of early RPM tumor cell isolation and culture. From left to right: whole slide H&E with bottom lung lobe (dashed box) provided in higher magnification inset (scale bar, 500 μm), with successive magnification of airway lesions (scale bar, 100 μm) with in situ tumors (red and blue boxes) labeled in right top panels (scale bars, 10 μm). IHC of serial sections of the same tumors (scale bars, 10 μm). Lungs were dissociated and single cell suspensions placed in culture.

(B) Representative brightfield images of tumor cells in culture at indicated days following plating. Scale bars, 200 (top row), 100 (middle and bottom rows) and 50 (red inset) μm. Representative of > 60 independent assays for RPM cells and 5 assays for RPR2 cells.
(C) Representative brightfield images of human classic (H1092) and variant (H82) cell lines. Scale bar, 100 μm.
(D) Representative immunoblot on specified days following culture of early stage RPM tumor cells. Representative of n = 5 independent assays. Bands normalized to HSP90 values, with fold-change relative to Day 5 (red and black graphs) or Day 24 (blue graphs) and averaged across 3-5 experiments.
(E) Representative immunoblot on specified days following puromycin selection from H889 (left) or H1963 (right) cells infected with vector control or $MYC^{T58A}$ constructs. MYC-expressing human SCLC cells (GLC1) are used for positive controls ($MYC^{+}$).
(F) Representative brightfield and GFP fluorescent images of H889 (left) and H1963 (right) human SCLC cells infected with retroviral control or $Myc^{T58A}$-Ires-Gfp viruses. Scale bars, 100 μm, except H889 $Myc^{T58A}$ is 50 μm. Average circularity, roundness and cluster size indicated +/− SEM. Student's unpaired two-tailed t test for MYC vs vector control, **** $p < 0.0001$, ** $p < 0.009$.
HSP90 serves as loading control. See also Figure S2.

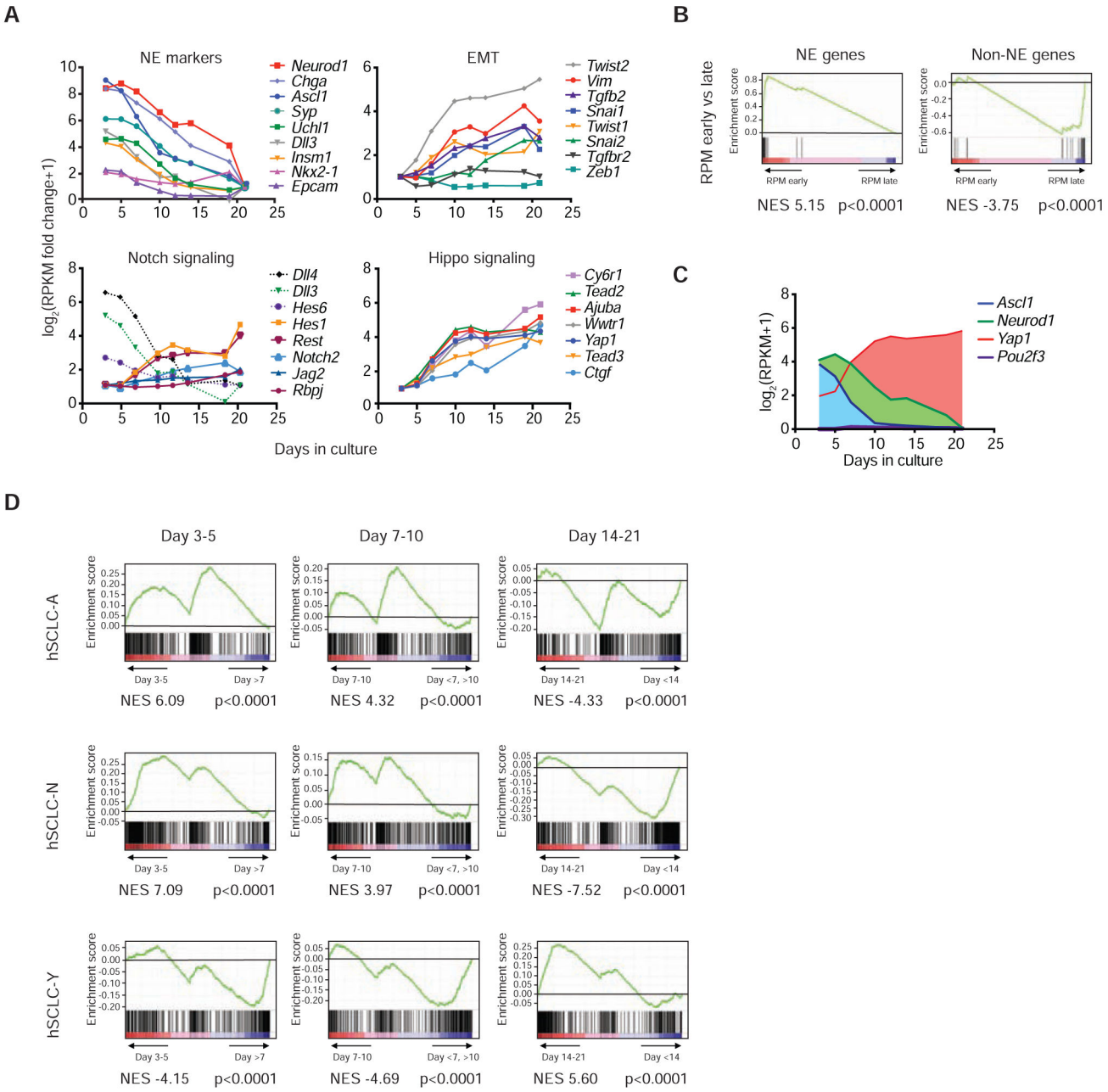



**Figure 3. Human SCLC subtypes correspond with MYC-driven evolution**

(A) Log2-fold change of indicated NE (relative to the last time point) and non-NE pathway genes (relative to the first time point) from bulk RNA-seq of primary RPM tumor cells at specific days in culture. Dashed lines in Notch signaling panel indicate genes predicted to be Notch-inhibitory.

(B) GSEA of 50-gene NE and non-NE gene signature from (Zhang et al., 2018) applied to early (day 3-7) vs late (day 14-21) time points of RPM transition.

(C) Log2 expression of SCLC-subtype defining transcription factor genes at indicated time points from bulk RNA-seq data of RPM transition.

(D) GSEA for human SCLC-ASCL1 (SCLC-A), SCLC-NEUROD1 (SCLC-N) and SCLC-YAP1 (SCLC-Y) gene signatures applied to bulk RNA-seq data grouped in day increments. See also Figure S3 and Table S1.

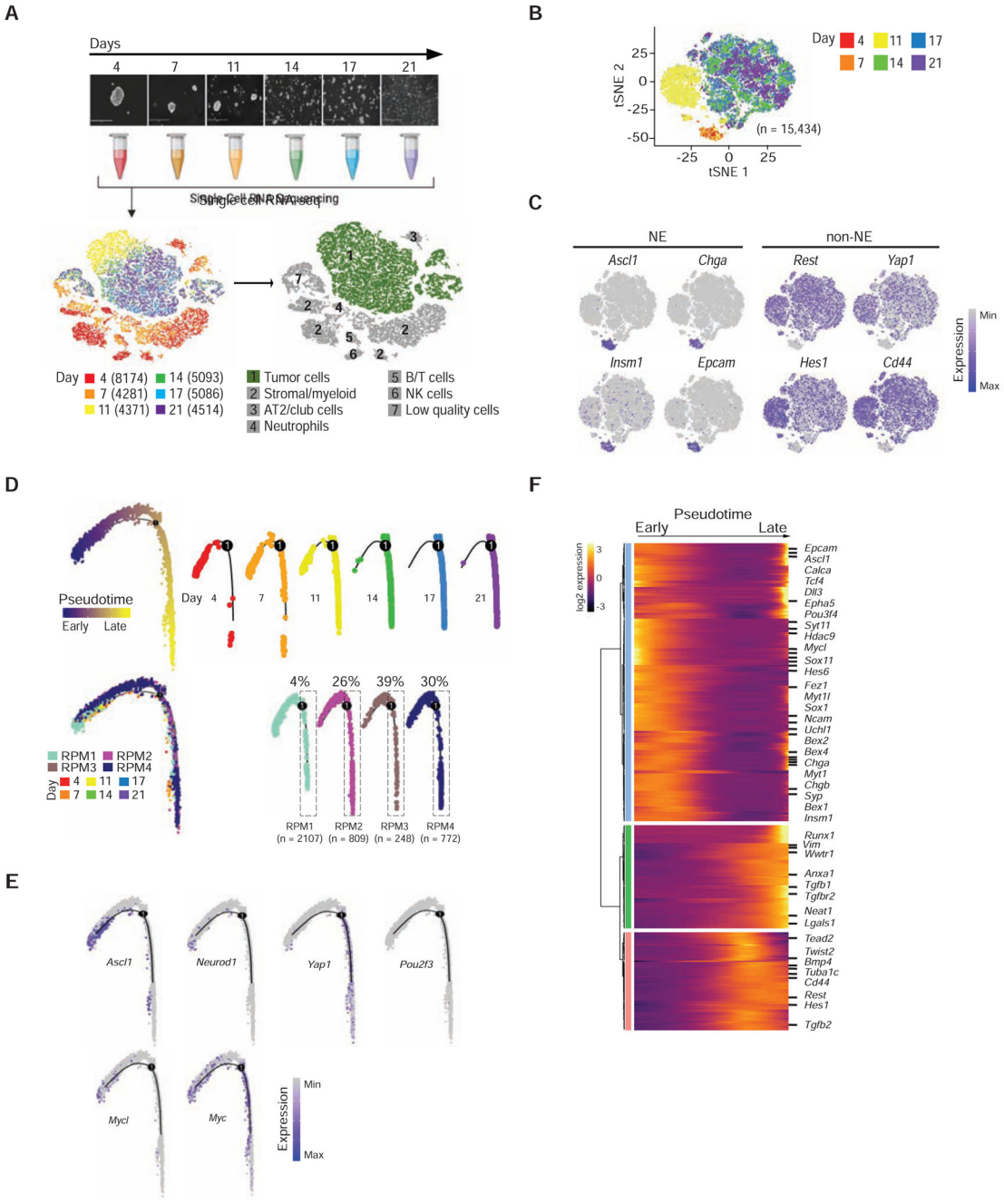



**Figure 4. MYC-driven SCLC subtypes progress along a single evolutionary trajectory**

(A) Schematic of primary RPM tumor cell transition analyzed by scRNA-seq with tumor cell populations colored by day, and number of cells analyzed per time point in the legend (bottom left). Predicted cell types based on gene expression labeled on the same cells by number in the bottom right panel.

(B) Following removal of non-tumor and low-quality cells, RPM tumor cells labeled by day in tSNE space using Monocle 2.

(C) Expression of individual NE and non-NE marker genes in tSNE space from cells in panel B.

(D) Pseudotime trajectory by Monocle 2 from early to late time points of primary RPM transition and 4 RPM-Cas9 tumors from Ad-Cgrp-Cre-infected mice. Faceted pseudotime plots indicated on top right and bottom right panels. Dashed insets in RPM tumors highlight percent of cells in late-stages of progression; total post-QC number of tumor cells indicated below RPM1-4 labels.

(E) Expression of indicated genes projected onto pseudotime space as in panel D.

(F) Heatmap of top-500 differentially expressed genes over pseudotime from the primary RPM tumor cell transition and 4 RPM tumors.

See also Figure S4 and Table S2.

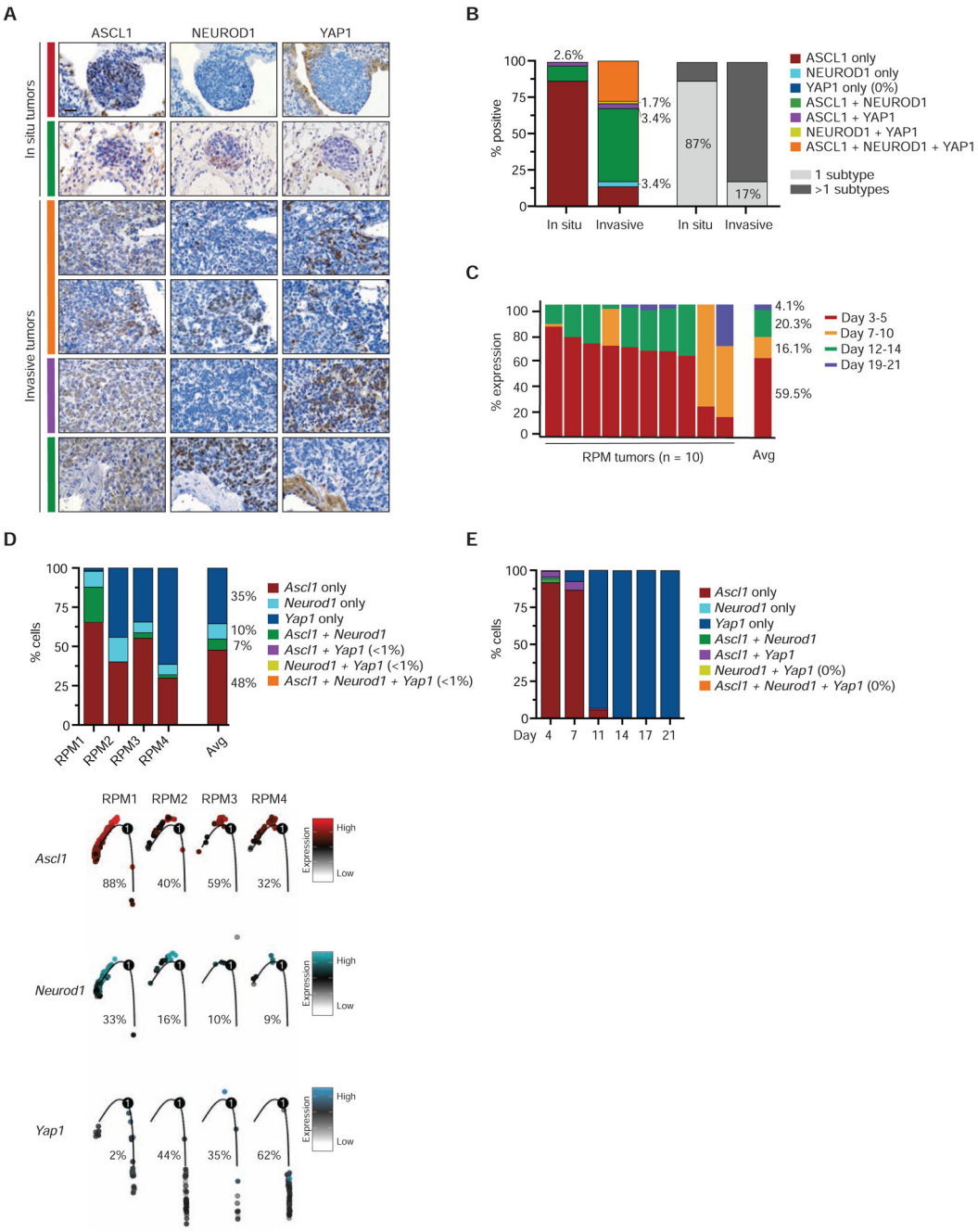



**Figure 5. MYC-driven murine tumors exhibit intratumoral SCLC subtype heterogeneity**

(A) Representative IHC from in situ (n = 38) or invasive (n = 59) RPM tumors analyzed in serial sections. Left color panel indicates subtype classification matching panel B. Scale bar, 25 μm.

(B) IHC quantification from serial sections in panel A where individual tumors have detection of 1 or more than 1 (> 1) subtypes.

(C) CIBERSORT analyses of bulk RNA-seq data from RPM tumors with average (Avg) percent similarity to gene expression signatures of RPM transition cells in culture.

(D) Top panel: Percentage of cells per tumor expressing subtype-defining genes with average (Avg) across tumors. Bottom panel: Individual RPM 1-4 trajectories with localization of positive cells in pseudotime. Percentage of total tumor cells expressing indicated genes shown below trajectories.
(E) Percentage of cells expressing subtype-defining genes in the RPM transition experiment from Figure 4D.

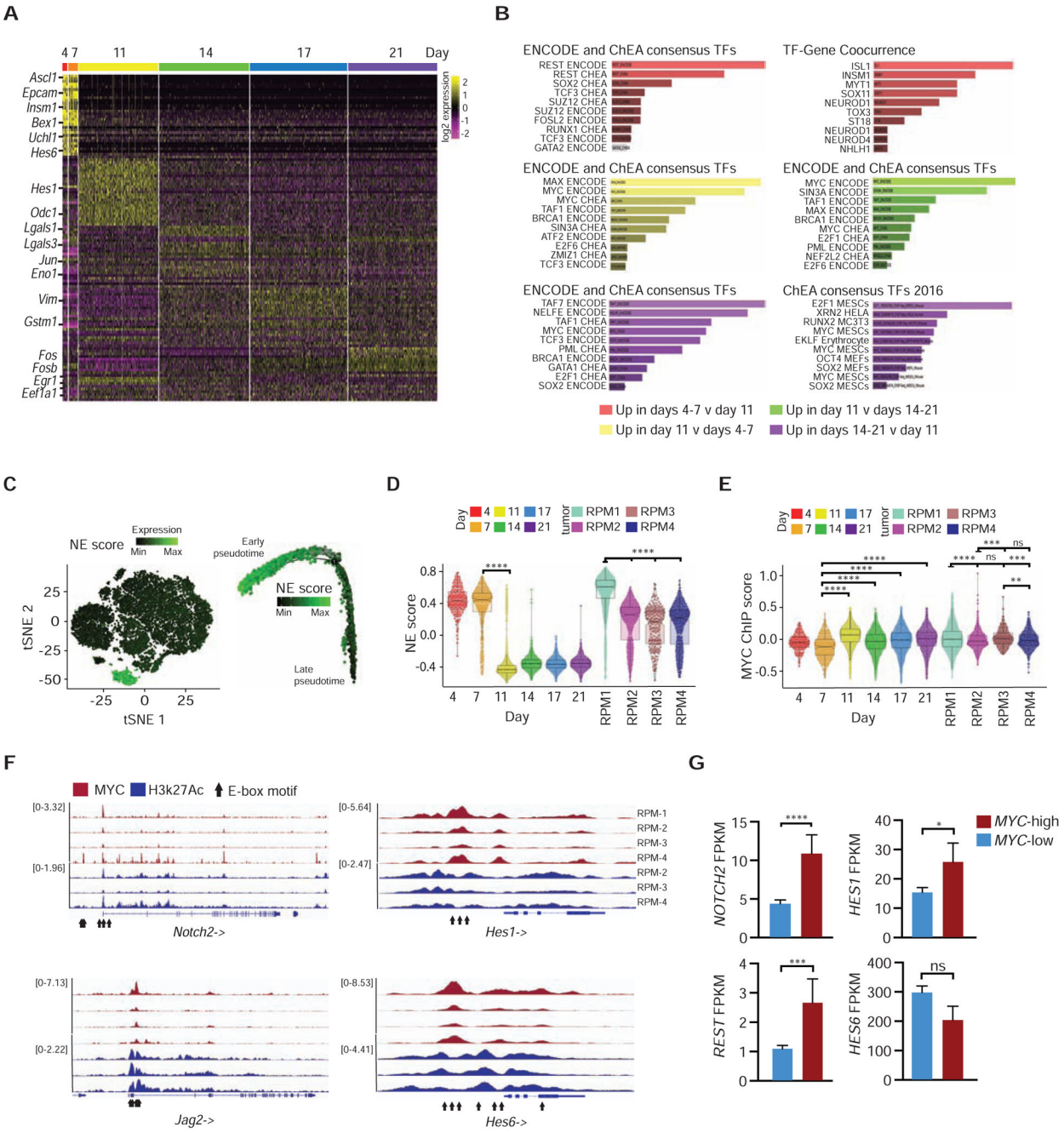



**Figure 6. MYC activates Notch signaling during NE reprogramming**

(A) Heatmap of top 30 (or fewer, if < 30 significant) differentially-expressed genes for each time point of RPM transition using Seurat.

(B) ENRICHR analysis of top differentially-expressed genes (0.25 log-fold change) in RPM tumor cell transition compared by day as color-coded in the bottom legend.

(C) 50-gene NE score from (Zhang et al., 2018) applied to RPM tumor cell transition in tSNE (left) as in Figure 4B and in pseudotime space (right) as in Figure 4D.

(D) Violin plots of NE-score from panel C applied to every cell of the RPM transition time points and individual RPM tumors. Student's two-tailed unpaired t test, **** p < 2.22e-16.

(E) Violin plots of MYC ChIP score applied to every cell of the RPM transition time points and individual RPM tumors. Student's two-tailed unpaired t test, **** p < 2.22e-16, *** p < 0.0004, ** p < 0.002, ns = not significant.

(F) ChIP-seq analysis of MYC (red) and H3K27Ac (blue) genomic binding at indicated gene loci from n = 3-4 independent RPM tumor samples. Blue rectangles below plots indicate gene exons with directionality of gene (->) near gene name. Black up-arrows indicate canonical E-Box 5'-CACGTG-3' or non-canonical '5-CANNTG-3'; motifs ("E-box motif") selected in the vicinity of observed MYC binding.

(G) Gene expression from (George et al., 2015) analyzed according to *MYC* status (n = 59 *MYC*-low and n = 11 *MYC*-high tumors). Mean +/− SEM, Student's two-tailed unpaired t test, **** p < 0.0001, *** p < 0.0007, * p < 0.03, ns = not significant.

For box plots, median and interquartile range are shown (lower bar is 25$^{th}$ percentile, upper bar is 75$^{th}$ percentile, and end points indicate minimum and maximum values). See also Figure S5 and Table S3.

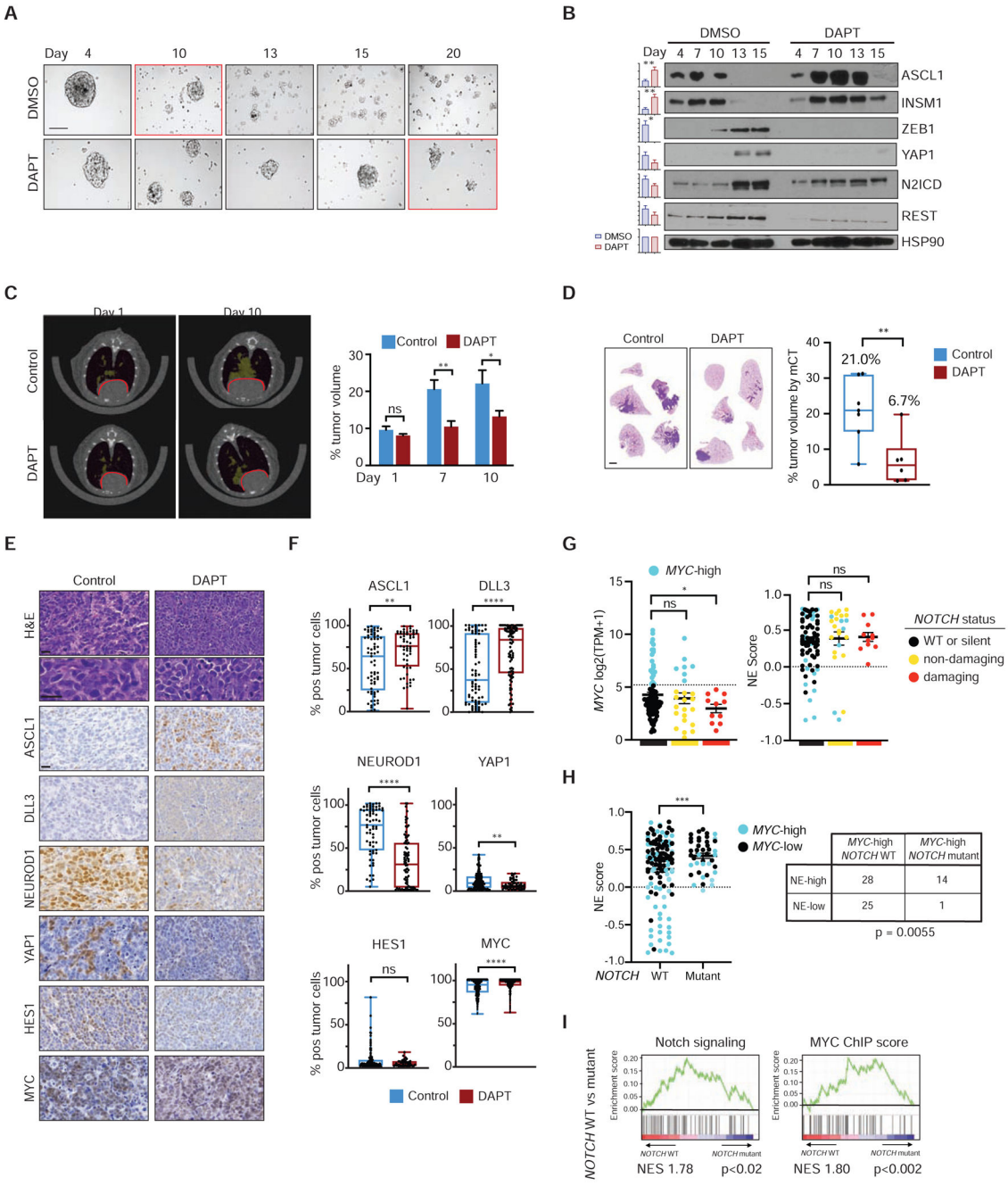



**Figure 7. Notch activation is required for MYC-driven tumor evolution**

(A) Representative brightfield images of primary RPM tumor cells cultured in DMSO or 10 μM DAPT treated every three days and visualized at indicated days. Red box on each row indicates the day that variant morphology was first observed. Scale bar, 100 μm. n = 4 biological experiments.

(B) Immunoblot from cells in panel A. HSP90 serves as loading control. Bar graphs on left represent fold change in expression, summed across all timepoints where protein was detected, relative to HSP90. Error bars represent mean +/− SEM for n = 4 biological replicates. Student's two-tailed unpaired t test, ** $p < 0.004$, * $p < 0.03$, ns = not significant.

(C) Representative microCT (mCT) imaging from vehicle control (corn oil) (n = 6) and 20 mg/kg DAPT-treated (n = 8) RPM mice. Lung tumors pseudocolored in yellow and heart outlined in red. Quantification of microCT imaging data for total tumor burden (% lung tumor volume/total lung volume) in graph on right at indicated days. Mean +/− SEM, Student's two-tailed unpaired t test, ** $p < 0.004$, * $p < 0.03$, ns = not significant.
(D) Representative H&E from vehicle control (n = 7) or DAPT-treated (n = 6) RPM mice at Day 10 following treatment. Scale bar, 2,000 μm. Quantification of average tumor burden (% tumor area/total lung area) in right panel with box plots where each dot represents one animal, Student's two-tailed unpaired t test, ** $p < 0.009$.
(E) Representative H&E and IHC in vehicle control (n = 7) or DAPT-treated (n = 6) RPM mice at Day 10 following treatment. Scale bars, 20 μm.
(F) Digital IHC quantification from lung tumor tissue (% positive tumor cells) in panel E where each dot represents a tumor. Student's two-tailed unpaired t test, **** $p < 0.0001$, ** $p < 0.03$, ns = not significant.
(G) Left panel: *MYC* expression in normalized transcripts per million (TPM) + 1 grouped by *NOTCH* status with *MYC*-high samples in blue; RNA-seq data from n = 70 human tumors (George et al., 2015) and n = 48 human cell lines (CCLE), excluding *POU2F3*$^{+}$ samples. Right panel: NE score applied to the same samples. "Non-damaging" indicates missense or in-frame deletions, and "damaging" indicates frame-shift deletions, nonsense, or splice variant mutations. Mean +/− SEM, Student's two-tailed unpaired t test with Welch's correction, * $p < 0.01$, ns = not significant.
(H) NE score applied to RNA-seq data from n = 70 human tumors (George et al., 2015) and n = 98 human cell lines (SCLC_CellMiner), excluding *POU2F3*$^{+}$ samples; Student's two-tailed unpaired t test with Welch's correction, *** p value = 0.0009. Contingency table analyzed by Fisher's exact test.
(I) GSEA comparing hSCLCs with *NOTCH* WT or silent ("WT") vs "damaging" mutations from samples in panel G. Normalized enrichment scores (NES) indicated.
For box plots, median and interquartile range is shown (lower bar is 25$^{th}$ percentile, upper bar is 75$^{th}$ percentile, and end points indicate minimum and maximum values). See also Figure S6 and Table S4.

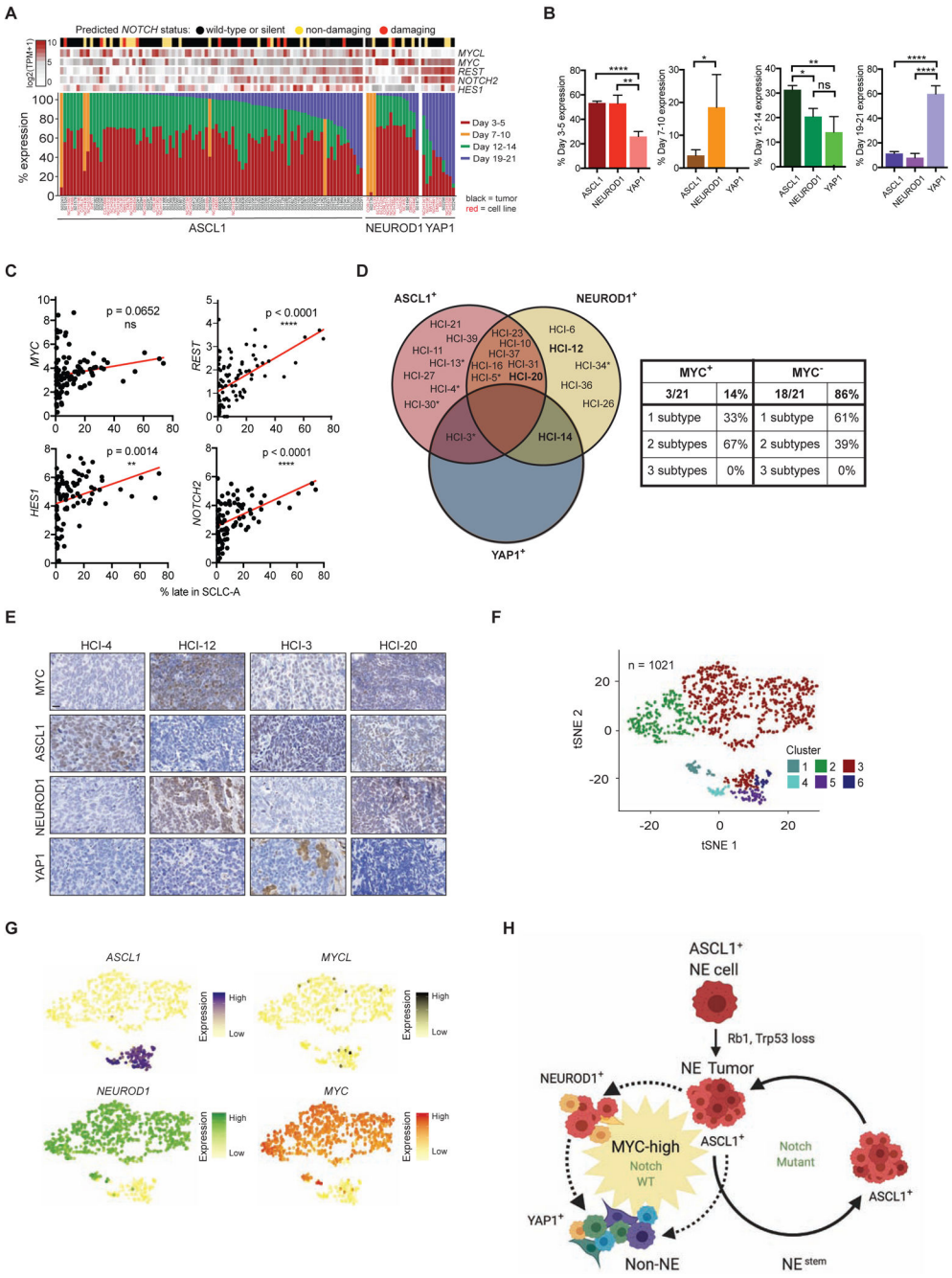



**Figure 8. Human SCLC exhibits intratumoral molecular subtype heterogeneity**

(A) CIBERSORT analysis of RPM transition signatures in 70 human SCLC tumors (George et al., 2015) (black ID) and 48 human SCLC cell lines from CCLE (red ID), excluding *POU2F3*$^{+}$ samples, grouped according to SCLC molecular subtype. Gene expression values for indicated *MYC* family member or Notch-related genes overlaid above the stacked bar graphs. Predicted *NOTCH*-status marked by top color bar.

(B) Percent of RPM time-point signatures (Y-axis) within hSCLC molecular subtypes derived from panel A. Mean +/− SEM, Student's two-tailed unpaired t test, **** $p < 0.0001$, ** $p < 0.007$, * $p < 0.02$, ns = not significant.

(C) Correlation of indicated genes (Y-axis) with percent late (Day 19-21) signature determined by CIBERSORT within the human $ASCL1^{+}$ subset only, with Pearson correlation p values indicated.

(D) Venn diagram of human SCLC tissue IHC results with deidentified HCI-patient# for positive samples for ASCL1 (red circle), NEUROD1 (yellow circle) or YAP1 (blue circle) (n = 21 total). Samples in bold text harbor at least some cells with $MYC^{+}$ IHC (HCI-12, −20, −14). Table on the right summarizes % of samples with indicated number of subtypes present at any frequency organized by MYC expression. * indicates tissue from limited stage as opposed to extensive stage.

(E) Representative IHC in indicated patient biopsies. Serial sections were stained, but multiple tumor regions are shown to illustrate tumor heterogeneity. Scale bar, 20 μm.

(F) tSNE unbiased clustering of tumor cell populations derived from chemotherapy-relapsed human SCLC liver biopsy analyzed by scRNA-seq. Number of post-QC tumor cells indicated.

(G) Relative expression of indicated genes in tSNE as in panel F.

(H) We propose a hypothetical model whereby *NOTCH*-deficient SCLC tumors are locked in an $NE^{stem}$-like ($ASCL1^{+}$) state (right circle) similar to Ouadah et al., 2019. In contrast, MYC reprograms *NOTCH*-WT tumors to non-NE SCLC fates that proceed either from NEUROD1 to YAP1 (left circle, top dashed line), or directly from ASCL1 to YAP1 (left circle, bottom dashed line).

See also Figure S7.

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Peroxidase AffiniPure Donkey Anti-Rabbit IgG (H+L) | Jackson ImmunoResearch | Cat#711-035-152; RRID: AB_10015282 |
| Peroxidase AffiniPure Goat Anti-Mouse IgG1, Fcg Subclass 2b Specific | Jackson ImmunoResearch | Cat#115-035-205; RRID: AB_2338513 |
| HSP90 | Cell Signaling Technology | Cat#4877 |
| ASCL1 | BD Pharmingen | Cat#BD556604 |
| UCHL1 | Sigma | Cat#HPA005993 |
| SYP | Thermo Fisher Scientific | Cat#RB1461P1 |
| CGRP | Sigma | Cat# C8198 |
| CD44 | Cell Signaling Technology | Cat#37259 |
| POU2F3 | Sigma | Cat#HPA019652 |
| EPCAM | Abcam | Cat#ab71916 |
| INSM1 | Santa Cruz Biotechnology | Cat#sc-271408 |
| NEUROD1 [EPR4008] | Abcam | Cat#ab109224 |
| REST | Millipore | Cat#17-641 |
| YAP XP | Cell Signaling Technology | Cat#14074 |
| NOTCH2 XP | Cell Signaling Technology | Cat#5732 |
| HES1 | Cell Signaling Technology | Cat#11988 |
| MYC, for immunoblot | Cell Signaling Technology | Cat#5605 |
| MYC, for human IHC | Abcam | Cat#ab32072 |
| MYC, for mouse IHC | Santa Cruz Biotechnology | Cat#sc-764 |
| NKX2-1 | Abcam | Cat#ab76013 |
| ZEB1 | Bethyl Laboratories | Cat#A301-922A |
| DLL3 | Thermo Fisher Scientific | Cat#PA5-23448 |
| MYC, for ChIP-seq | Cell Signaling Technology | Cat#13987 |
| H3K27Ac | Active Motif | Cat#39133 |
| Bacterial and Virus Strains | | |
| Ad5-CMV-Cre | University of Iowa Viral Vector Core Facility | Cat#VVC-U of Iowa-5 |
| Ad5-CGRP-Cre | University of Iowa Viral Vector Core Facility | Cat#VVC-Berns-1160 |
| Ad5-CCSP-Cre | University of Iowa Viral Vector Core Facility | Cat#VVC-Berns-1166 |
| Ad5-SPC-Cre | University of Iowa Viral Vector Core Facility | Cat#VVC-Berns-1168 |
| Biological Samples | | |
| Mouse tissues | This paper | N/A |
| Human SCLC tissue | Huntsman Cancer Institute | N/A |
| Chemicals, Peptides, and Recombinant Proteins | | |

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| DAPT | Apexbio | Cat#A8200 |
| DBZ | Apexbio | Cat#A4018 |
| Normal Goat Serum | Jackson Immunoresearch | Cat#005-000-121 |
| ACK Lysing Buffer | Thermo Fisher Scientific | Cat#A10492 |
| Collagenase Type 1A | Sigma | Cat#C9891 |
| Collagenase, Type 4 | Worthington Biochemical | Cat#LS004186 |
| Dispase | Worthington Biochemical | Cat#LS02104 |
| StemPro Accutase Cell Dissociation | Invitrogen | Cat#A1110501 |
| DMSO | Fisher Scientific | Cat#BP231-100 |
| Puromycin | Fisher Scientific | Cat#BP2956-100 |
| Gotaq G2 Colorless Master Mix | Promega | Cat#M7832 |
| Formaldehyde (37% by weight) | Thermo Fisher Scientific | Cat#BP531-500 |
| Polybrene | Santa Cruz Biotechnology | Cat#sc-134220 |
| Hydroxypropylmethylcellulose (HPMC) | Sigma | Cat#H7509 |
| Tween-80 | Fisher Scientific | Cat#BP338 |
| Corn oil | Sigma | Cat#C8267 |
| Critical Commercial Assays | | |
| LIVE/DEAD Fixable Violet Dead Cell Stain Kit | Invitrogen | Cat#L34963 |
| Zombie Green Viability Fluorescent Dye | Biolegend | Cat#423111 |
| DAB Peroxidase (HRP) Substrate Kit (with Nickel), 3,3'-diaminobenzidine | Vector Laboratories | Cat#SK-4100 |
| VECTASTAIN ABC Kit (Rabbit IgG) | Vector Laboratories | Cat#PK-4001 |
| SignalStain Boost IHC Detection Reagent (HRP, Rabbit) | Cell Signaling Technology | Cat#8114 |
| SignalStain Antibody Diluent | Cell Signaling Technology | Cat#8112 |
| Advansta WesternBright ECL HRP Substrate Kit | VWR | Cat# 490005-020 |
| Mouse on Mouse (M.O.M) Basic Kit | Vector Laboratories | Cat#BMK-2202 |
| Qiagen DNeasy Blood and Tissue Kit | Qiagen | Cat#69504 |
| Nextera DNA Flex Library Prep Kit | Illumina | Cat#20018705 |
| TruSeq Stranded mRNA Library Prep Kit | Illumina | Cat#RS-122-2101 |
| TruSeq Stranded Total RNA Library Ribo-Zero Gold Prep Kit | Illumina | Cat#RS-122-2301 |
| HiSeq SR Cluster Kit v4-cBot | Illumina | Cat#GD-401-4001 |
| HiSeq SBS Kit v4 | Illumina | Cat#FC-401-4002 |
| Chromium Single Cell 3' Library & Gel Bead Kit v3 | 10X Genomics | Cat#PN-1000075 |
| Chromium Single Cell 5' Library & Gel Bead Kit | 10X Genomics | Cat#PN-1000020 |
| Chromium Single Cell Controller | 10X Genomics | Cat#PN-120263 |
| D1000 Screen Tape | Agilent | Cat#5067-5582 |
| D1000 Reagents | Agilent | Cat#5067-5583 |
| KAPA Library Quantification Kit | Roche | Cat#KK4842 |
| RNeasy Mini Kit | Qiagen | Cat#74106 |
| Deposited Data | | |

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Bulk RNA Seq RPM Transition Timepoints | This paper | GEO: GSE149180 |
| Bulk RNA Seq RPM and RPR2 Tumors | This paper | GEO: GSE149180 |
| Single Cell RNA-Seq RPM Transition Timepoints | This paper | GEO: GSE149180 |
| Single Cell RNA-Seq Bulk RPM Tumors | This paper | GEO: GSE149180 |
| MYC and H3K27Ac ChIP-Seq from RPM Tumors | Chalishazar et al., 2019 | GEO: GSE142496 |
| Single Cell RNA-Seq Human SCLC Liver Biopsy | This paper | GEO: GSE149180 |
| Experimental Models: Cell Lines | | |
| Mouse: RPM Transition Primary Cells | This paper | N/A |
| Mouse: RPR2 Transition Primary Cells | This paper | N/A |
| Human: NCI-H889 | Gift of John Minna, UTSW | CRL-5817; CVCL_1598 |
| Human: GLC1 | Gift of Martin Sos, Germany | CVCL_8200 |
| Human: H1963 | Gift of Paul Bunn, UC-Denver | CVCL_1510 |
| Human: NCI-H1092 | Gift of David MacPherson, Fred Hutch | CRL-5855 |
| Human: NCI-H446 | Gift of John Minna, UTSW | HTB-171 |
| Human: NCI-H82 | Gift of John Minna, UTSW | HTB-175 |
| Human: GLC8 | Gift of Martin Sos, Germany | CVCL_8218 |
| Human: HEK-293T | ATCC | CRL-3216; CVCL_0063 |
| Experimental Models: Organisms/Strains | | |
| Mouse: *Rb1*$^{fl/fl}$;*Trp53*$^{fl/fl}$;*Rbl2*$^{fl/fl}$ (RPR2) | Julien Sage, Stanford, Schaffer et al, 2010 | MMRRC 043692-UCD |
| Mouse: *Rb1*$^{fl/fl}$;*Trp53*$^{fl/fl}$;*MycT58A*$^{LSL/LSL}$ (RPM) | Trudy G. Oliver, University of Utah, Mollaoglu et al, 2017 | JAX#029971 |
| Mouse: *Rosa26-LSL-Cas9-Ires-GFP* | Feng Zhang, MIT, Platt et al, 2014. | JAX#024857 |
| Mouse: *Rb1*$^{fl/fl}$;*Trp53*$^{fl/fl}$;*MycT58A*$^{LSL/LSL}$-*Cas9* | This paper | N/A |
| Oligonucleotides | | |
| *Rb1* recombination primers | Tyler Jacks, MIT | https://jackslab.mit.edu/protocols/genotyping/rb1lox |
| *Trp53* recombination primers | Tyler Jacks, MIT | https://jackslab.mit.edu/protocols/genotyping/p53_cond_recomb |
| Recombinant DNA | | |
| Puro-Ires-GFP Empty | Mayr and Bartel., 2009 | Addgene Plasmid #21654 |
| Puro-Ires-GFP MYC T58A | Mollaoglu et al., 2017 | Will deposit to Addgene |
| pCMV-VSVG | Stewart et al., 2003 | Addgene Plasmid #8454 |
| pCMV-delta-R8.2 | Stewart et al., 2003 | Addgene Plasmid #8455 |
| Software and Algorithms | | |
| Graphpad Prism 8 | Graphpad Software | www.graphpad.com/scientific-software/prism/ |
| Image Studio Lite 5 | LI-COR | www.licor.com/bio/products/software/image_studio_lite/ |
| Quantum GX2 mCT Software | PerkinElmer | N/A |
| Analyze 11.0 | AnalyzeDirect | https://analyzedirect.com/ |

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| FlowJo v10 | FlowJo LLc | www.flowjo.com |
| BD FACSDiva 8 | BD Biosciences | http://www.bdbiosciences.com/us/instruments/research/software/flow-cytometry-acquisition/bd-facsdiva-software/m/111112/overview |
| ImageJ: Image processing and analysis in Java | ImageJ | https://imagej.nih.gov/ij/ |
| Cell Ranger Analysis Pipeline | 10X Genomics | https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger |
| Loupe Cell Browser 3.1.1 | 10X Genomics | https://support.10xgenomics.com/single-cell-gene-expression/software/visualization/latest/what-is-loupe-cell-browser |
| Integrative Genomics Viewer (IGV) | Broad Institute, and the Regents of the University of California | http://software.broadinstitute.org/software/igv/ |
| Gene Set Enrichment Analysis (GSEA) | Broad Institute, and the Regents of the University of California | http://software.broadinstitute.org/gsea/index.jsp |
| Enrichr | Kuleshov et al., 2016 | http://amp.pharm.mssm.edu/Enrichr/ |
| Cibersort | Newman et al., 2015 | https://cibersort.stanford.edu |
| Monocle2 | Trapnell et al., 2014; Qiu et al., 2017 | http://cole-trapnell-lab.github.io/monocle-release/docs/ |
| Seurat | Butler et al., 2018; Stuart and Butler et al., 2019 | https://satijalab.org/seurat/ |
| Scater | McCarthy et al., 2017 | http://bioconductor.riken.jp/packages/3.9/bioc/vignettes/scater/inst/doc/vignette-intro.html |
| R Statistical Programming | The R Foundation | www.r-project.org |
| BioRender | BioRender 201 | https://www.getbiorender.com |
| SCLC_Cellminer Database | Tlemsani, et al., BioRxiv preprint, 2020 | https://discover.nci.nih.gov/SclcCellMinerCDB/ |