AD_____

AWARD NUMBER: W81XWH-17-1-0223

TITLE: Using hESC-Derived Lung Cells to Study the Initiation of Small Cell Lung Cancer

PRINCIPAL INVESTIGATOR: Harold Varmus, MD

RECIPIENT: Weill Medical College of Cornell University
New York, NY 10065

REPORT DATE: July 2018

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
July 2018	Annual	15 Jun 2017 – 14 June 2018

4. TITLE AND SUBTITLE
Using hESC-Derived Lung Cells to Study the Initiation of Small Cell Lung Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-17-1-0223

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Varmus, Harold; Chen, Huanhuan

E-Mail: varmus@med.cornell.edu; huc2003@med.cornell.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

AND ADDRESS(ES)
Joan and Sanford I. Weill
Medical College of Cornell
University
New York, NY 10065

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Cell culture models based on directed differentiation of human embryonic stem cells (hESCs) may reveal why certain constellations of genetic changes drive carcinogenesis in specialized human cell lineages. In research supported by this award, we have exploited our recent discovery that up to 10 percent of lung progenitor cells derived from hESCs can be induced by inhibition of NOTCH signaling to form pulmonary neuroendocrine cells (PNECs), the putative precursors to small cell lung cancers (SCLCs). By using small inhibitory RNAs in these cultures to reduce levels of retinoblastoma (RB) protein, the product of a gene commonly mutated in SCLCs, we can significantly expand the number of PNECs. However, similarly reducing levels of TP53 protein, the product of another tumor suppressor gene commonly mutated in SCLCs, or expressing mutant *KRAS* or *EGFR* genes, commonly implicated in lung adenocarcinomas, did not induce or expand PNECs, consistent with lineage-specific sensitivity to loss of *RB* function. Still, tumors resembling early stage SCLC grew in immunodeficient mice after subcutaneous injection of PNEC-containing cultures in which expression of both *RB and TP53* was blocked, implying distinct functional requirements for down-modulation of both tumor suppressor genes. Single-cell RNA profiles of PNECs are heterogeneous; when RB levels are reduced, the profiles show similarities to RNA profiles from early stage SCLC; when both RB and TP53 levels are reduced, the transcriptome is enriched with cell cycle-specific RNAs. Taken together, these findings demonstrate that genetic manipulation of hESC-derived pulmonary cells permits a detailed analysis of steps involved in the initiation, progression, and treatment of this recalcitrant cancer.

15. SUBJECT TERMS
Small cell lung cancer; human pluripotent stem cells; single cell transcriptomics; pulmonary neuroendocrine cells, tumor suppressor genes; carcinogenesis.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	15	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

1. **INTRODUCTION:** Narrative that briefly (one paragraph) describes the subject, purpose and

Cancers presumed to arise from different cell lineages display characteristic genotypes, but cells of origin are generally uncertain, and the relationships between lineage-specific attributes and genotypic differences among tumors are not understood. One of the main obstacles to greater knowledge about these relationships is the need for tractable systems that allow molecular changes observed in mature cancer cells to be evaluated for their contribution to hallmarks of neoplasia as they occur during the development of specific cell lineages. To study these issues, we have sought ways to assess functional changes that occur after specific genes are altered in human pulmonary cells at defined stages of tissue development. Recent advances in the induction, cultivation, and directed differentiation of human embryonic stem cells (hESCs) provide opportunities to study carcinogenesis in many human cell types derived from a variety of lineages, including cancers such as small cell lung cancer (SCLC), in which rapid onset and progression limit the availability of clinical samples, especially from early stage disease.

2. **KEYWORDS:** Provide a brief list of keywords (limit to 20 words).

Small cell lung cancer; human pluripotent stem cell; single cell transcriptomics; pulmonary neuroendocrine cell; tumor suppressor genes.

3. **ACCOMPLISHMENTS:** The PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency grants official whenever there are significant changes in the project or its direction.

What were the major goals of the project?
List the major goals of the project as stated in the approved SOW. If the application listed milestones/target dates for important activities or phases of the project, identify these dates and show actual completion dates or the percentage of completion.

The following text is organized in accord with the originally stated "specific aims" (in bold), followed by a brief summary of the progress achieved towards those aims. MOVE THE ACCOMPLISHMENTS TO THE NEXT SECTION

Specific Aim 1: To characterize CGRP+ cells in appearing in hESC cultures after inhibition of NOTCH signaling, with or without reduction of RB protein.

During the first year of this award, we have completed all of the objectives listed in our application and we have described the findings in a manuscript that has been posted on the bioRxiv preprint server (https://www.biorxiv.org/content/early/2018/07/02/261461.) and has been resubmitted to a leading journal after review. In summary, we have validated the induction of PNECs by inhibition of NOTCH signaling, using several experimental strategies; examined the characteristics of the PNECs by immunohistochemistry, flow sorting, and single cell RNA sequencing with both DropSeq and 10X sequencing methods; and compared the results with reported findings from human early stage SCLCs. REFER TO FIGURES IN SUPPLEMENT TOO (Additional characterization of these cells and tumors, using a wide variety of genetic changes and cells sorted by FACS methods is now planned, mainly under the terms of a recently awarded U01 grant from the National Cancer Institute.)

Specific Aim 2: To determine the mechanism that accounts for an increased proportion of PNEC like cells after reduction of RB expression

We have begun to examine this important question by measuring the effects of diminished RB mRNA and protein on DNA replication, entry into the cell cycle, and effects on differentiation of lung progenitors into PNECs. Since preliminary findings indicate that loss of RB function may influence all

three of these phenomena, we are refining our experimental plan to narrow the time interval in which the increases in PNECs occur and to separate subpopulations of DAPT-treated cells. ADD PRELIMINARY FINDINGS. We anticipate finishing this analysis in the next year of support as described below.

Specific Aim 3: To determine whether PNEC-like cells induced from hESCs can be transformed by known oncogenic events and can form SCLC-like tumors.

As described in our manuscript (https://www.biorxiv.org/content/early/2018/07/02/261461), we have shown that cultures containing DAPT-induced PNECs, in which both RB and TP53 tumor suppressor genes have been inhibited, can form tumors in subcutaneous injection sites in immune-deficient mice. These tumors appear similar histologically to SCLCs and express markers of SCLCs, but not the markers of teratomas; in addition, they are relatively indolent in growth properties, without evidence of invasion or metastasis. REFER TO FIGURES IN SUPPLEMENT In the next year, we will further explore the properties of these tumors, their genetic requirements, the influence of varied injection sites, and the consequences of cell sorting before injection.

Specific Aim 4: To test whether genotypic changes that convert PNEC-like cells to an oncogenic phenotype are specific for such cells or can be induced in other cell types in the lung lineage.

As documented in the submitted manuscript (https://www.biorxiv.org/content/early/2018/07/02/261461), we have shown that inhibition of NOTCH signaling induces PNECs and that suppression of expression of RB enhances the percentage of PNECs, but that other common genotypic changes in lung cancers (blockage of TP53 or mutation of EGFR or KRAS) do not affect the induction or expansion of PNECs. However, other effects of these and other lung cancer mutations on lung progenitors have not yet been fully explored.

What was accomplished under these goals?

For this reporting period describe: 1) major activities; 2) specific objectives; 3) significant results or key outcomes, including major findings, developments, or conclusions (both positive and negative); and/or 4) other achievements. Include a discussion of stated goals not met. Description shall include pertinent data and graphs in sufficient detail to explain any significant results achieved. A succinct description of the methodology used shall be provided. As the project progresses to completion, the emphasis in reporting in this section should shift from reporting activities to reporting accomplishments.

Significant results:

Generation of PNECs from cultured hESCs.

Methods have recently been described for generating most, but not all, of the cell types observed in adult lung tissues by using growth factors and chemicals to alter signaling pathways sequentially in cells derived from hESCs over several weeks (Fig. 1a). Using a protocol developed by Huang et al (Huang et al., 2015; Huang et al., 2014), we have confirmed that, by day 3, about 90% of hESCs (the RUES2 and ES02 lines) differentiate into definitive endoderm (DE), triple positive for the markers KIT, EPCAM, and CXCR4 (Supplementary Fig. 1a,b); anterior foregut endoderm (AFE) by day 6; increasing numbers of lung progenitors (LP),

SOX2+, NKX2.1+ and FOXA2+, between days 15 and 25 (Supplementary Fig. 1c, d, Supplementary Fig. 2a, b); and then a variety of airway and lung epithelial cells (basal progenitor cells, ciliated cells, goblet cells, club cells, and alveolar type 1 and type 2 cells (Treutlein et al., 2014; Warburton et al., 1998) [AT1 and AT2]) by day 55 (Supplementary Fig. 1e-g). However, this protocol and others produce few if any PNECs (<0.5%; Fig.1b, c; Supplementary Fig. 1g).

Studies of mouse development have suggested that inhibition of signaling via NOTCH receptors might influence cells to adopt a neuroendocrine fate (Ito et al., 2000; Linnoila, 2006; Morimoto et al., 2012; Shan et al., 2007). In addition, inactivation of *NOTCH* genes is found in about 25% of SCLCs (George et al., 2015). Based on these reports, we exposed LPs between day 25 and day 55 to N-[(3,5-Difluorophenyl)acetyl]-L-alanyl-2-phenyl]glycine-1,1-dimethylethyl ester (DAPT)(Geling et al., 2002)---a known inhibitor of γ-secretase, the protease that normally cleaves NOTCH receptors to yield a transcriptionally active, mobile, intracellular domain of NOTCH (NICD)(Schroeter et al., 1998)---to ask whether the loss of NOTCH signaling might promote the production of PNECs. (We define PNECs here as cells containing a general lung-specific marker, the transcription factor NKX2.1, and expressing one or more genes encoding well-recognized neuroendocrine markers, especially the cell surface-associated protein that includes the calcitonin gene-related peptide (CGRP) (Song et al., 2012) or the nuclear transcription factor ASCL1 (Borges et al., 1997; Borromeo et al., 2016).)

After exposure to 5 to 10uM DAPT for 30 days, a substantial number of differentiating LP cells (about 8.9 ± 1.9% in RUES2 cells, and 5.6 ± 1.4% in ES02 cells, versus 0.5 ± 0.20% or 0.4 ± 0.2% in control cultures) adopt properties of PNECs, as measured by counting CGRP+ cells with fluorescence-activated cell sorting (FACS) and confirmed qualitatively by immunofluorescence of cells in monolayer cultures with antibodies against NXK2.1 and CGRP (Fig.1b-d, Supplementary Fig. 2 c-e).

DAPT induces PNECs by blocking cleavage of NOTCH

We took several approaches to confirm the mechanism by which DAPT induced PNECs. First, we measured the abundance of NICD in extracts from cells treated with DAPT and observed the expected loss of the γ-secretase cleavage product (Fig. 1e). To confirm the reduction in NOTCH-mediated signaling, we measured the readout from two NOTCH target genes (Iso et al., 2003), *HES1* and *HEY1,* and found a marked decrease in levels of transcripts from *HEY1* and a slight but significant decrease of RNA from *HES1* (Supplementary Fig. 2f). We also tested another known inhibitor of γ-secretase, dibenzazepine (DBZ)(Milano et al., 2004), and produced percentages of PNECs at day 55 similar to those observed with DAPT (Fig. 1f). Finally, we reversed the effects of DAPT by providing NICD, the normal product of γ-secretase-mediated cleavage of NOTCH: expression of a tetracycline-inducible transgene encoding NICD from days 25 to 55 in differentiated RUES2 cells decreased the appearance of CGRP+ cells in cultures concurrently exposed to DAPT (Fig. 1 g, h). We also showed that expression of *HES1* and *HEY1*, which was down-regulated in DAPT-treated lung cells (Fig. 1e and Supplementary Fig. 2f), was increased by induction of NICD (Fig. 1g). In addition, DAPT induced expression of *ASCL1* after the RUES2 line was differentiated to form lung cells (Fig. 1e).

Single cell transcriptional profiling of induced PNECs.

To further characterize the presumptive PNECs generated by inhibition of NOTCH signaling, we used high-throughput single-cell RNA sequencing (scRNA-seq) applied to DAPT-treated and untreated cells at day 55 (Fig. 2a-e, Supplementary Fig. 1 f, g; see Methods). Clustering of the single-cell profiles revealed one cluster enriched with cells expressing genes that encode CGRP or ASCL1, thus identifying the presumed PNEC cells. In total, the presumed PNEC cells constituted 7.72% of the 9,824 high quality cells (pooled from two biological replicates). Analysis of differential gene expression in the various clusters revealed that cells in cluster 4 exhibit relatively high numbers of transcripts from a set of genes, including *CGRP, ASCL1*, *GRP*, *SYP*, and *UCHL1*, that encode canonical PNEC markers (Linnoila, 2006; Song et al., 2012) (Fig. 2c, e and Supplementary file 1). Figure 2e displays the expression of those and other relevant genes, including *MYC, MYCL* and *NEUROD1* (Mollaoglu et al., 2017), characteristically expressed in neuroendocrine cells or SCLC. In comparison, we detected *CGRP* or *ASCL1* RNA in only 1.3% percent of cells, and rarely together, from control cultures not treated with DAPT (Supplementary Fig. 1g).

Reduced expression of RB enlarges the proportion of PNECs

Since a central objective of this work is to assess the influence of known lung cancer genes on the behavior of cells in the lung lineage, we next examined the consequences of expressing or simulating known oncogenic mutations in hESC cultures undergoing differentiation, with or without inhibition of NOTCH signaling (Figs. 3 and 4). To that end, we equipped the RUES2 hESC line with Dox-inducible transgenes encoding small hairpin RNAs (shRNAs) that inhibit production of RNA from either of the two tumor suppressor genes most commonly inactivated by mutations in SCLC, the *RB* or *TP53* genes (Supplementary Fig. 3a, b, and Fig. 4 a). We also introduced into parallel cultures inducible transgenes encoding oncogenes commonly encountered in lung adenocarcinomas, mutant *EGFR* or mutant *KRAS* (Cancer Genome Atlas Research, 2014) (Fig. 4 b).

Induction of *RB*-specific shRNA in RUES2 cells differentiating to form LCs between days 25 and 50 markedly reduced the amount of RB protein (Supplementary Fig. 3a) but not the amounts of the closely related proteins p107 and p130 (Supplementary Fig. 3b). Reduced expression of the *RB* gene was associated with a significantly increased number of CGRP+ NKX2.1+ cells (putative PNECs) from 7.5 ± 2.0% to 37.8 ± 8.2% (Fig. 3a-d), as measured by FACS, but only when cells were also exposed to DAPT to inhibit processing of NOTCH (Fig.2d). Similarly, the proportions of cells expressing the PNEC transcription factor ASCL1 and the associated markers NCAM1 and CHGA were also significantly increased (Fig. 3e).

In contrast, induction of *TP53*-specific shRNA during the same interval had no effect on the number of CGRP+ cells, with or without DAPT and with or without induction of *RB*-shRNA (Fig. 4c-e). Similarly, induction of mutant forms of EGFR and KRAS proteins between days 25 and 55 (Fig. 4b) failed to increase the number of CGRP+ cells grown in the absence or presence of DAPT (Fig. 4f, g). These findings indicate that loss of TP53 or production of mutant EGFR or KRAS proteins do not induce or affect the abundance of PNECs.

The effects of reducing RB levels on PNEC transcriptomes

To examine the transcriptional phenotypes of cultures of differentiated (day 55) RUES2 cells in which both NOTCH and RB signaling were inhibited, we turned again to scRNA-seq. Cell clustering and differential expression analyses, similar to those used in Figure 1, indicate the

presence of multiple cell populations in our cultures, including an expanded PNEC-like cell compartment (11.7%), expressing markers similar to those observed in PNEC-like cells at day 55 with normal levels of RB protein (Supplementary Fig. 3c-g and Supplementary file 1). (We attribute the relatively modest increase in PNEC-like cells---as judged by scRNA-seq, compared to the increase measured by FACS---to differences in the sensitivity of detection methods that measure RNA levels as opposed to cell surface proteins.)

Transcriptomes of PNECs with reduced RB levels resemble transcriptomes from early SCLC.

To ask whether inhibition of NOTCH signaling, coupled with a reduction of RB protein, produces a transcriptional program that resembles the program in human SCLC, we compared the scRNA-seq profiles of PNECs and non-PNECs from day 55 RUES2 cells, with normal or reduced levels of *RB* gene expression, to the published RNA profiles from 29 early-stage (stage Ia or Ib) human SCLCs (George et al., 2015; Peifer et al., 2012) (Supplementary Fig.3 h). This analysis confirmed that PNEC expression profiles more closely resemble SCLC profiles than do the profiles from non-PNECs. In addition, the similarity to SCLC profiles is greater (P < 2.2e-16 by two sided Kolmogorov–Smirnov test) in PNECs in which RB levels were reduced than in PNECs in which RB levels were not perturbed.

To further characterize the effects of the reduction of RB protein on the transcriptional profile of the PNEC cells, we performed a differential expression analysis that compared PNECs appearing after DAPT with PNECs appearing after DAPT combined with reduced expression of *RB* (Supplementary Fig.4a). Subsequent gene function enrichment analysis (Chen et al., 2009) indicated that the most differentially expressed genes following reduction of *RB* gene expression are involved in regulation of nerve development, apoptosis, TP53 signal transduction, and other processes (Supplementary Fig.4b and Supplementary file 2). When investigating the heterogeneity within the DAPT-induced PNEC cell cluster, we detected three sub-populations (Supplementary Fig.4c) with substantially different expression profiles (Supplementary Fig.4c). A similar analysis of PNECs from DAPT-treated cultures in which RB levels were reduced also revealed three subpopulations, but with transcriptional profiles different from those in cultures in which RB expression was not altered (Supplementary Fig.4c). Some of these differentially expressed genes, including *HES1, SST, SCG3, STMN2, ELAVL3, IGFBP4*, as well as *NEUROD1*, have previously been reported to be responsible for the heterogeneity of pulmonary neuroendocrine tumors and implicated in the initiation and progression of human SCLC (Borromeo et al., 2016; Lim et al., 2017).

Reduction of both P53 and RB in DAPT-induced PNECs allows xenografts to form tumors

To assess the ability of cells in differentiated RUES2 cultures to form tumors, we performed subcutaneous injections of day 55 cells treated in various ways into immune-deficient NSG mice (NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ)(Shultz et al., 2005). The four tested cell populations all contained presumptive PNECs (measured as CGRP+ cells by FACS), ranging from about 10 to about 40 percent of the cultured cells. No growths greater than 250 mm^3 in volume were observed within 7 weeks after injection of parental cells exposed to DAPT alone or cells also carrying the *TP53*-shRNA or the *RB*-shRNA expression cassettes and treated with Dox (Fig 5 a). In contrast, cells with both shRNA cassettes and treated with DAPT and Dox formed visible tumors at 14 of 19 injection sites within 6 to 7 weeks. In general, these tumors were about 1 cm in diameter, formed of compact, darkly staining cells, with mitotic figures observed in some of them, morphologically resembling SCLC in mice and humans (Fig 5 b), and not locally invasive.

The origin of the tumor cells was confirmed by detection of GFP encoded by a component of the *RB*-shRNA cassette in nuclei of RUES2 cells (Supplementary Fig. 5 a, b); in addition, the nuclei were stained with antisera specific for human antigens (Supplementary Fig. 5 c). A PNEC-like phenotype was documented using IHC to display the neuroendocrine biomarkers CGRP, NCAM1, and ASCL1, as well as the lung marker NKX2.1(Figure 5b). In addition, we observed that some cells within the xenograft tumors produced C-MYC and MYCL (Mollaoglu et al., 2017), but not NEUROD1, another neuroendocrine transcription factor (Supplementary Fig. 5f-h).

We excluded the possibility that the observed tumors were mischaracterized teratomas known to be formed in mice injected with hESCs. When we produced teratomas in NSG mice with undifferentiated RUES2 cells, they contain embryonic tissue markers (Liu et al., 2010), such as α-fetoprotein, Nanog, Oct4, and SSEA4 (Supplementary Fig. 5 d); exhibited morphological features different from tumors formed with our differentiated cultures in which NOTCH, RB, and P53 pathways were disrupted (Supplementary Fig. 5a); and did not contain PNEC markers CGRP, ASCL1 and NKX2.1 (Supplementary Fig. 5e). However, since the neuroendocrine tumors in Figure 5 arose from injections of heterogeneous cultures of RUES2-derived cells, we cannot completely exclude the highly unlikely possibility that the tumors originated from non-PNEC cells in the cultures (Huang et al., 2018).

The effects on PNEC transcriptomes of reduced expression of both RB and P53.

Since induction of *TP53*-specific shRNA, as well as *RB*-specific shRNA, was required to produce tumor-forming cells in DAPT-treated cells derived from RUES2 cells, we used scRNA-seq to seek evidence that might associate the tumorigenic phenotype with changes in PNEC transcriptomes. As indicated earlier (Fig 4c-e), inhibition of expression of *TP53* does not affect the proportion of PNECs in these cultures; as expected, scRNA-seq revealed that the PNEC-like compartment of cultures in which both *RB* and *TP53* were inhibited was similar in size (10.8%) to that of cultures in which only *RB* was inhibited (Supplementary Fig. 3c-g). Cell clustering and analysis of differentially expressed genes in these unfractionated cultures indicate the presence of multiple cell populations (Fig. 6 a-d), similar to cultures in which P53 was not reduced (Supplementary Fig. 3c-g).

In the analysis of cells in which P53 levels were reduced, however, we identified three well-defined clusters, one differing from the others because its cells show increased expression of genes associated with active cell cycling (Fig. 6 e-g). This is consistent with the idea that different proliferation rates exist within the PNEC compartment, suggesting that reduction of P53 levels, coupled with low levels of RB protein, enables a subset of PNECs to enter a proliferative mode---a phenomenon we did not observe in cells with normal levels of P53, regardless of RB status (Supplementary Fig. 4c). We also found that this subset of presumptively proliferating PNECs highly expresses genes associated with inhibition of apoptosis, as well as cell cycle genes, including *BIRC5, TOP2A, MKI67, CDK1, CDKN3, CDC20*. This indicates that the main effects of reducing P53 levels in PNECs are likely to be proliferative and anti-apoptotic (Fig. 6 e-g). These effects might account for the appearance of tumor-forming potential in xenografted mice.

A significant increase in the expression of RNA encoding the neuroendocrine transcription factor, NEUROD1,which has been associated with advanced stages of SCLC, was also observed in cultures in which P53 protein levels were reduced (Borromeo et al., 2016; Osborne et al., 2013a; Osborne et al., 2013b) (Fig. 6 e, f). The percentages of *NEUROD1* RNA-positive cells in

PNEC clusters produced by NOTCH inhibition alone, by NOTCH inhibition with reduced RB protein, and by NOTCH inhibition with reduction of both RB and P53, are 6.61±0.52%, 3.02±0.75% and 15.75±0.50%, respectively (mean ± sd, following repeated subsampling to account for sequencing depth).

(The figures, a brief method description and references are included in the file of supplementary materials.)

The summary in the preceding section provides a guide to our accomplishments in the first year of support under this grant, and the referenced manuscript offers a detailed account of the many experiments that have been completed and the goals that have already been achieved.

What opportunities for training and professional development has the project provided?

If the project was not intended to provide training and professional development opportunities or there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe opportunities for training and professional development provided to anyone who worked on the project or anyone who was involved in the activities supported by the project. "Training" activities are those in which individuals with advanced professional skills and experience assist others in attaining greater proficiency. Training activities may include, for example, courses or one-on-one work with a mentor. "Professional development" activities result in increased knowledge or skill in one's area of expertise and may include workshops, conferences, seminars, study groups, and individual study. Include participation in conferences, workshops, and seminars not listed under major activities.

Most of the work conducted under the terms of this award was performed by Joyce Chen, a post-doctoral fellow in the Varmus laboratory, and hence formed an integral part of her training. While her salary was paid by a post-doctoral fellowship from the Beckman Foundation, funds from this grant were used to support the extensive studies that she has undertaken. Since her background was in bioengineering, the DoD award allowed her to acquire new skills in cancer biology and molecular biology; these skills will be essential if she is to achieve her career aspirations as an independent investigator in her current field of work.

In addition, this award has supported work by a graduate student, Asaf Poran, in Olivier Elemento's laboratory. Asaf is a computer scientist with interests in the analysis of gene expression data from single cells. He has now assumed a senior position as a computational biologist at a biotechnology company, and the skills he has acquired during the conduct of this study have been essential to his success. (Note that he shares "first author" status with Dr. Chen on our posted manuscript.)

How were the results disseminated to communities of interest?

If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the results were disseminated to communities of interest. Include any outreach activities that were undertaken to reach members of communities who are not usually aware of these project activities, for the purpose of enhancing public understanding and increasing interest in learning and careers in science, technology, and the humanities.

The principal investigator (Harold Varmus) has had multiple opportunities to talk about this work before audiences that include scientists from a variety of biomedical fields, including students and fellows looking for new opportunities in cancer research. (The venues include several symposia, lectures at academic institutions, and meetings on cancer research.) Joyce Chen and Asaf Poran have also spoken on several occasions at meetings and at institutional seminars about this work.

What do you plan to do during the next reporting period to accomplish the goals?

If this is the final report, state "Nothing to Report."

Describe briefly what you plan to do during the next reporting period to accomplish the goals and objectives.

During the next funding period, we plan to accomplish the work remaining to accomplish all of the Specific Aims. We will place special emphasis on the following experiments:

(1) Under Aim 2, we will conduct a series of experiments to define more accurately the time during which RUES2 lines that have been differentiated to form lung progenitor cells must be exposed to DAPT to maximize induction of CGRP+ cells. More importantly, we will then determine the optimal time course for reducing RB levels to obtain maximal levels of GCGP+ cells; the kinetics of expansion of the CGRP+ cell population will be essential for determining the mechanism(s) by which loss of RB increases the proportion of PNECs through effects on differentiation, replication, or apoptosis. COULD SAY MORE ABOUT LINEAGE EXPERIMENTS

(2) Under Aim 3, we will test whether other oncogenic events (e.g. hyper-expression of a *MYC* gene, loss of *PTEN*, or mutation of *PIK3CA*) increases the tumorigenicity of RUES2-derived PNECs in which *TP53* and *RB* have been down-regulated. These experiments will be done by injecting cells (and appropriate control cells) into multiple sites in immunosuppressed mice, assessing rate of tumor growth, numbers of cells required for tumor formation, expression of lung cell markers, histology, and invasion and metastasis.

(3) Under Aim 4, we will test various types of lung cells (e.g. club cells, basal progenitor cells, alveolar type 1 and type 2 cells) derived from RUES2 hESCs for the acquisition of oncogenic properties after introducing changes in a variety of genes, including *RB*, *TP53*, PTEN, NOTCH, and genes known to serve as drivers of other kinds of lung cancer after characteristic mutations.

4. **IMPACT:** Describe distinctive contributions, major accomplishments, innovations, successes, or any change in practice or behavior that has come about as a result of the project relative to:

What was the impact on the development of the principal discipline(s) of the project?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how findings, results, techniques that were developed or extended, or other products from the project made an impact or are likely to make an impact on the base of knowledge, theory, and research in the principal disciplinary field(s) of the project. Summarize using language that an intelligent lay audience can understand (Scientific American style).

As described in the report, we have developed a novel experimental approach for studying the initiation and progression of small cell lung cancer (SCLC), by differentiating human embryonic stem cells (hESCs) into the lung lineage and perturbing three tumor suppressor genes frequently altered in the human disease. Of particular significance, using this experimental system we have (i) for the first time, generated pulmonary neuroendocrine cells (PNECs), the putative precursors of SCLC, by interfering with NOTCH signaling in committed lung progenitors; (ii) shown that reduced expression of the retinoblastoma tumor suppressor gene (*RB*), but not other alterations in genes commonly implicated in other forms of lung cancer, markedly expands the PNEC population, demonstrating the specificity of the role of RB signaling in the generation of SCLC; (iii) characterized the transcriptomes of PNECs generated by inhibition of NOTCH signaling, with or without interference with *RB* gene expression, using single cell RNA sequencing methods to show both informative heterogeneity and similarities to early stage human SCLC; and (iv) demonstrated that reduced production of TP53 *and* RB confers a tumorigenic phenotype on PNECs so that they form tumors when introduced into experimental animals.

These findings indicate the utility of this system for studying the initiation, progression, and treatment of SCLC in much greater depth than has been possible before, using human cells and the genes known to be altered in advanced human SCLC. This is especially critical because SCLC is a major public health concern---indeed the target of a major initiative at the National Cancer Institute---because of its high incidence, acquired resistance to standard chemotherapy, and essentially universal mortality. Furthermore, no significant improvements in diagnosis or therapy have occurred during the past few decades, despite major advances in the understanding, prevention, and treatment of many other forms of human cancer.

Finally, the general methods we have developed for studying the origins of SCLC, using human cells that can be characterized at the single cell level, are likely to be widely adopted to ask a wide range of questions about the origins and control of other kinds of cancer, as well as SCLC.

What was the impact on other disciplines?

If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the findings, results, or techniques that were developed or improved, or other products from the project made an impact or are likely to make an impact on other disciplines.

As indicated in the previous section, our work offers important precedents for work on other kinds of cancers. But we also foresee the application of our methods to the pathogenesis of other types of diseases, using hESCs or induced pluripotent stem cells with a variety of genotypes to understand the contribution of inherited mutations, somatic mutations, and environmental factors in the causation of disease and to envision new preventive and therapeutic strategies for controlling diseases of many kinds.

What was the impact on technology transfer?

If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe ways in which the project made an impact, or is likely to make an impact, on commercial technology or public use, including:

- *transfer of results to entities in government or industry;*
- *instances where the research has led to the initiation of a start-up company; or*
- *adoption of new practices.*

Thus far, we have not filed any claims for intellectual property protection based on our work under this award and we have no immediate plans to do so. Nevertheless, we are aware that many commercial firms seek to develop new means to prevent and treat cancers and that some companies are likely to consider our experimental strategies as the basis for the development of assays that could be used for testing new therapeutics.

What was the impact on society beyond science and technology?

If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how results from the project made an impact, or are likely to make an impact, beyond the bounds of science, engineering, and the academic world on areas such as:

- *improving public knowledge, attitudes, skills, and abilities;*
- *changing behavior, practices, decision making, policies (including regulatory policies), or social actions; or*
- *improving social, economic, civic, or environmental conditions.*

It would be premature and arrogant to claim that the work we have done thus far has materially changed our society in obvious ways. But it is fair to say that our work does demonstrate the value of using human embryonic or pluripotent cells for understanding human disease. At one time, there was significant skepticism about the use of such cells, based on ethical considerations, but the work we have done contributes to the emerging view that, when conducted with appropriate ethical oversight, work with hESCs can contribute in important and beneficial ways to our understanding of human development and human diseases. In that sense, the work reflects favorably on the scientific enterprise and helps to reassure the public about the value of science, including work that carries the potential to raise ethical concerns.

5. **CHANGES/PROBLEMS:** *The PD/PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency grants official whenever there are significant changes in the project or its direction. If not previously reported in writing, provide the following additional information or state, "Nothing to Report," if applicable:*

   Describe any changes in approach during the reporting period and reasons for these changes. Remember that significant changes in objectives and scope require prior approval of the agency.

   As indicated in the report, our work on this project has progressed well; only very minor adjustments in emphasis or methodology has been required.

   Actual or anticipated problems or delays and actions or plans to resolve them

   Describe problems or delays encountered during the reporting period and actions or plans to resolve them.

   Nothing to report

   Changes that had a significant impact on expenditures

   Describe changes during the reporting period that may have had a significant impact on expenditures, for example, delays in hiring staff or favorable developments that enable meeting objectives at less cost than anticipated.

   Nothing to report

   Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

   Describe significant deviations, unexpected outcomes, or changes in approved protocols for the use or care of human subjects, vertebrate animals, biohazards, and/or select agents during the reporting period. If required, were these changes approved by the applicable institution committee (or equivalent) and reported to the agency? Also specify the applicable Institutional Review Board/Institutional Animal Care and Use Committee approval dates.

   Nothing to report
   Significant changes in use or care of human subjects

   None.

   Significant changes in use or care of vertebrate animals

   None.

   Significant changes in use of biohazards and/or select agents

None.

6. PRODUCTS: List any products resulting from the project during the reporting period. If there is nothing to report under a particular item, state "Nothing to Report."

- **Publications, conference papers, and presentations**
  Report only the major publication(s) resulting from the work under this award.

  As noted earlier, work completed under the terms of this award is described in a manuscript now under review and posted on bioRxiv (https://www.biorxiv.org/content/early/2018/07/02/261461)

  In addition, Joyce Chen presented the work in Belfer Basic Research Working Group (BBRWG) seminar at Meyer Cancer Center at WCM in Febuary 2018 and at the International Society for Stem Cell Research (ISSCR) conference in June 2017. Harold Varmus presented results from these studies in multiple seminars at academic institutions in the US and abroad; at the National Cancer Institute's annual meeting of its small cell lung cancer research consortium in March 2018; and at several national and international meetings in 2017-2018.

  Journal publications. *List peer-reviewed articles or papers appearing in scientific, technical, or professional journals. Identify for each publication: Author(s); title; journal; volume: year; page numbers; status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).*

  MENTION STATUS OF MANUSCRIPT AND REITERATE POSTING ON BIORXIV

  Books or other non-periodical, one-time publications.

  Report any book, monograph, dissertation, abstract, or the like published as or in a separate publication, rather than a periodical or series. Include any significant publication in the proceedings of a one-time conference or in the report of a one-time study, commission, or the like. Identify for each one-time publication: author(s); title; editor; title of collection, if applicable; bibliographic information; year; type of publication (e.g., book, thesis or dissertation); status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).

Not applicable.

  Other publications, conference papers and presentations. *Identify any other publications, conference papers and/or presentations not reported above. Specify the status of the publication as noted above. List presentations made during the last year (international, national, local societies, military meetings, etc.). Use an asterisk (*) if presentation produced a manuscript.*

Not applicable.

- **Website(s) or other Internet site(s)**
  List the URL for any Internet site(s) that disseminates the results of the research activities. A short description of each site should be provided. It is not necessary to include the publications already specified above in this section.

We have noted above, multiple times, the URL that links to our manuscript.

- **Technologies or techniques**

  Identify technologies or techniques that resulted from the research activities. Describe the technologies or techniques were shared.

  Several protocols, both novel and adapted, have been used in our work and described in the manuscript. They are fully and freely available for others to use and are not appropriate for intellectual property protection.

- **Inventions, patent applications, and/or licenses**

  Identify inventions, patent applications with date, and/or licenses that have resulted from the research. Submission of this information as part of an interim research performance progress report is not a substitute for any other invention reporting required under the terms and conditions of an award.

  Not applicable

- **Other Products**

  Identify any other reportable outcomes that were developed under this project. Reportable outcomes are defined as a research result that is or relates to a product, scientific advance, or research tool that makes a meaningful contribution toward the understanding, prevention, diagnosis, prognosis, treatment and /or rehabilitation of a disease, injury or condition, or to improve the quality of life. Examples include:
  - *data or databases;*
  - *physical collections;*
  - *audio or video products;*
  - *software;*
  - *models;*
  - *educational aids or curricula;*
  - *instruments or equipment;*
  - *research material (e.g., Germplasm; cell lines, DNA probes, animal models);*
  - *clinical interventions;*
  - *new business creation; and*

- *other.*

Single cell RNA sequencing data is cited in our manuscript and available upon request. Any altered versions of the RUES2 cell line are available to qualified investigators upon request.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?
Provide the following information for: (1) PDs/PIs; and (2) each person who has worked at least one person month per year on the project during the reporting period, regardless of the source of compensation (a person month equals approximately 160 hours of effort). If information is unchanged from a previous submission, provide the name only and indicate "no change".

Example:

Name:	*Mary Smith*
Project Role:	*Graduate Student*
Researcher Identifier (e.g. ORCID ID):	*1234567*
Nearest person month worked:	*5*
Contribution to Project:	*Ms. Smith has performed work in the area of combined error-control and constrained coding.*
Funding Support:	*The Ford Foundation (Complete only if the funding support is provided from other than this award.)*

As indicated in the original grant application, multiple personnel are involved in this project. No significant changes have occurred:

Harold Varmus: no change
Joyce Chen: no change
Olivier Elemento: no change
Asaf Poran: no change

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

If there is nothing significant to report during this reporting period, state "Nothing to Report."

If the active support has changed for the PD/PI(s) or senior/key personnel, then describe what the change has been. Changes may occur, for example, if a previously active grant has closed and/or if a previously pending grant is now active. Annotate this information so it is clear what has changed from the previous submission. Submission of other support information is not necessary for pending changes or for changes in the level of effort for active support reported previously. The awarding agency may require prior written approval if a change in active other support significantly impacts the effort on the project that is the subject of the project report.

No changes in personnel involved in this project have occurred since the original submission. The Varmus laboratory has received additional support (a U01 cooperative agreement from the National Cancer Institute) for its studies of small cell lung cancer. The NCI is aware of the DoD grant and adjusted its funding of the U01 to avoid any redundancy of expense or effort.

What other organizations were involved as partners?

If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe partner organizations – academic institutions, other nonprofits, industrial or commercial firms, state or local governments, schools or school systems, or other organizations (foreign or domestic) – that were involved with the project. Partner organizations may have provided financial or in-kind support, supplied facilities or equipment, collaborated in the research, exchanged personnel, or otherwise contributed.

Provide the following information for each partnership:
Organization Name:
Location of Organization: (if foreign location list country)
Partner's contribution to the project (identify one or more)

- *Financial support;*
- *In-kind support (e.g., partner makes software, computers, equipment, etc., available to project staff);*
- *Facilities (e.g., project staff use the partner's facilities for project activities);*
- *Collaboration (e.g., partner's staff work with project staff on the project);*
- *Personnel exchanges (e.g., project staff and/or partner's staff use each other's facilities, work at each other's site); and*
- *Other.*

No organizations other than Weill Cornell Medicine are involved.

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS: For collaborative awards, independent reports are required from BOTH the Initiating Principal Investigator (PI) and the Collaborating/Partnering PI. A duplicative report is acceptable; however, tasks shall be clearly marked with the responsible PI and research site. A report shall be submitted to **https://ers.amedd.army.mil** for each unique award.

QUAD CHARTS: If applicable, the Quad Chart (available on https://www.usamraa.army.mil) should be updated and submitted with attachments.

Not applicable.

9. **APPENDICES:** Attach all appendices that contain information that supplements, clarifies or supports the text. Examples include original copies of journal articles, reprints of manuscripts and abstracts, a curriculum vitae, patent applications, study questionnaires, and surveys, etc.

   A PDF of the preprint posted on *bioRxiv* is attached.

Figure 1.

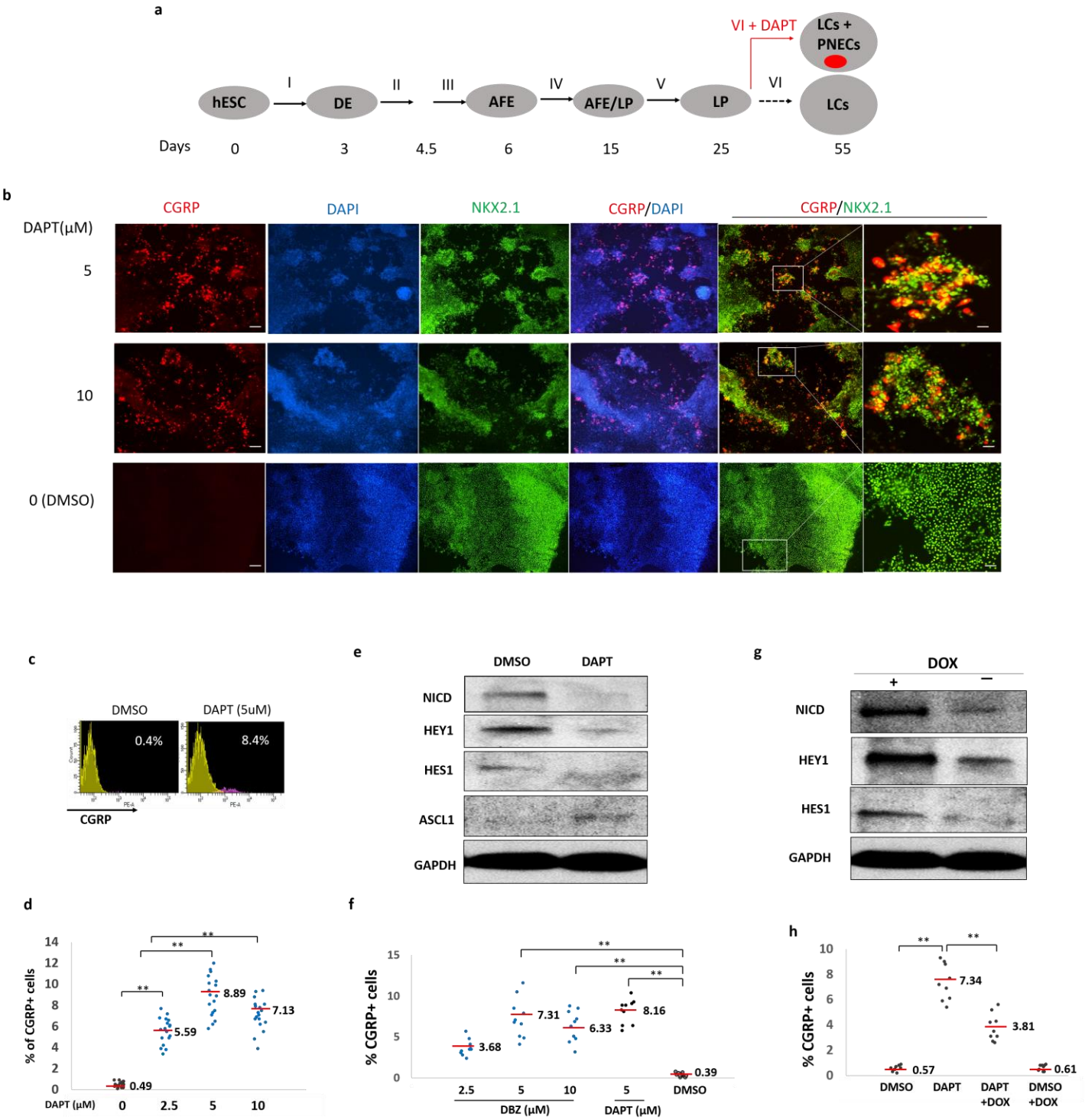

a
hESC
DE
AFE
AFE/LP
LP
LCs
LCs + PNECs
I
II
III
IV
V
VI
VI + DAPT
Days
0
3
4.5
6
15
25
55
b
CGRP
DAPI
NKX2.1
CGRP/DAPI
CGRP/NKX2.1
DAPT(μM)
5
10
0 (DMSO)
c
DMSO
DAPT (5uM)
0.4%
8.4%
CGRP
d
% of CGRP+ cells
DAPT (μM)
0
2.5
5
10
0.49
5.59
8.89
7.13
e
DMSO
DAPT
NICD
HEY1
HES1
ASCL1
GAPDH
f
% CGRP+ cells
3.68
7.31
6.33
8.16
0.39
DBZ (μM)
DAPT (μM)
DMSO
g
DOX
+
−
NICD
HEY1
HES1
GAPDH
h
% CGRP+ cells
0.57
7.34
3.81
0.61
DMSO
DAPT
DAPT +DOX
DMSO +DOX

Figure 1. Generating PNECs through directed differentiation of hESCs and suppression of NOTCH. a. Schematic of the protocol used to generate PNECs by stepwise differentiation of human embryonic stem cells (hESCs) to form definitive endoderm (DE) by day 3, anterior foregut endoderm (AFE) by day 6, and increasing numbers of lung progenitor cells (LPs) from days 15 to 25, using the differentiation mixtures I to V (defined in Methods; see Supplementary Figs 1 and 3). LPs were further differentiated in mixture VI from days 25 to 55 into the major types of lung cells (LCs) found in mature human lung parenchyma and airway epithelium (Treutlein, et.,al. 2014; Warburton, et.,al. 1998). Addition of DAPT to mixture VI induced formation of pulmonary neuroendocrine cells (PNECs; red dot), as described in the text. **b.** Detection of putative PNECs by IHC after treatment with DAPT. ESCs from the RUES2 line were differentiated according to the protocol in panel **a** to day 55 then stained to detect CGRP, NKX2.1, or both, with the indicated antisera; nuclei were detected by staining with DAPI. Scale bars: 100 μM (left) and 20 μM (right)**. c and d.** Percentages of CGRP+ cells were determined at day 55 by FACS and displayed as flow cytometry data (red, CGRP+; yellow, CGRP-) and a scatter graph (below). **e-g.** Confirmation of mechanism of action of DAPT as inhibitor of Y-secretase cleavage of NOTCH. **e.** DAPT (5uM) treatment from day 25 to day 55 decreased the level of the NOTCH intracytoplasmic domain (NICD) and protein products of the NOTCH target genes, HES1 and HEY1, while increasing levels of ASCL1 in day 55 lung cells, as detected by western blot. **f.** LPs treated with another Y-secretase inhibitor, DBZ, from day 25 to 55, also form CGRP+ cells at frequencies similar to those observed with DAPT (see panel d). **g and h.** Constitutive expression of NICD prevents the appearance of CGRP+ cells co-treated with DAPT. RUES2 cells carrying a doxycycline-inducible NICD were differentiated to form LPs and then treated with DOX, DAPT, or both for 30 days. Panel **g** demonstrates the induction of NICD by DOX, and expression of HES1 and HEY1 with Western blot; panel **h** shows by FACS that DOX (to induce expression of NICD) inhibits DAPT-mediated appearance of CGRP+ cells. ** $P < 0.01$, * $P < 0.05$ by one-way ANOVA test or (for panel **c**) by Student *t* test. Horizontal red lines denote average values; number of biological repeats (n) = 18 for panel **d**; n=10 for panel **f**; n=9 for panel **h**.

Figure 2.

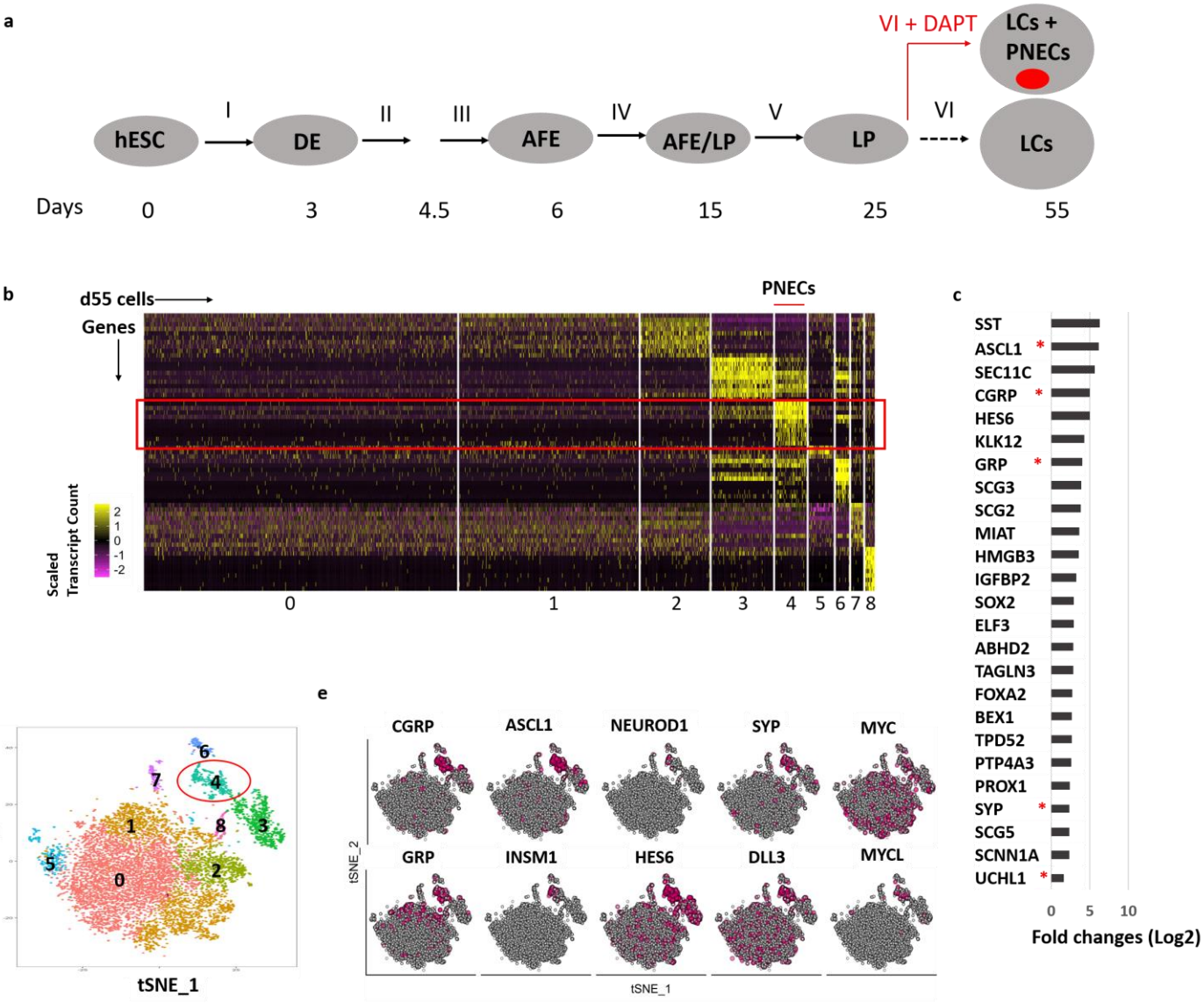
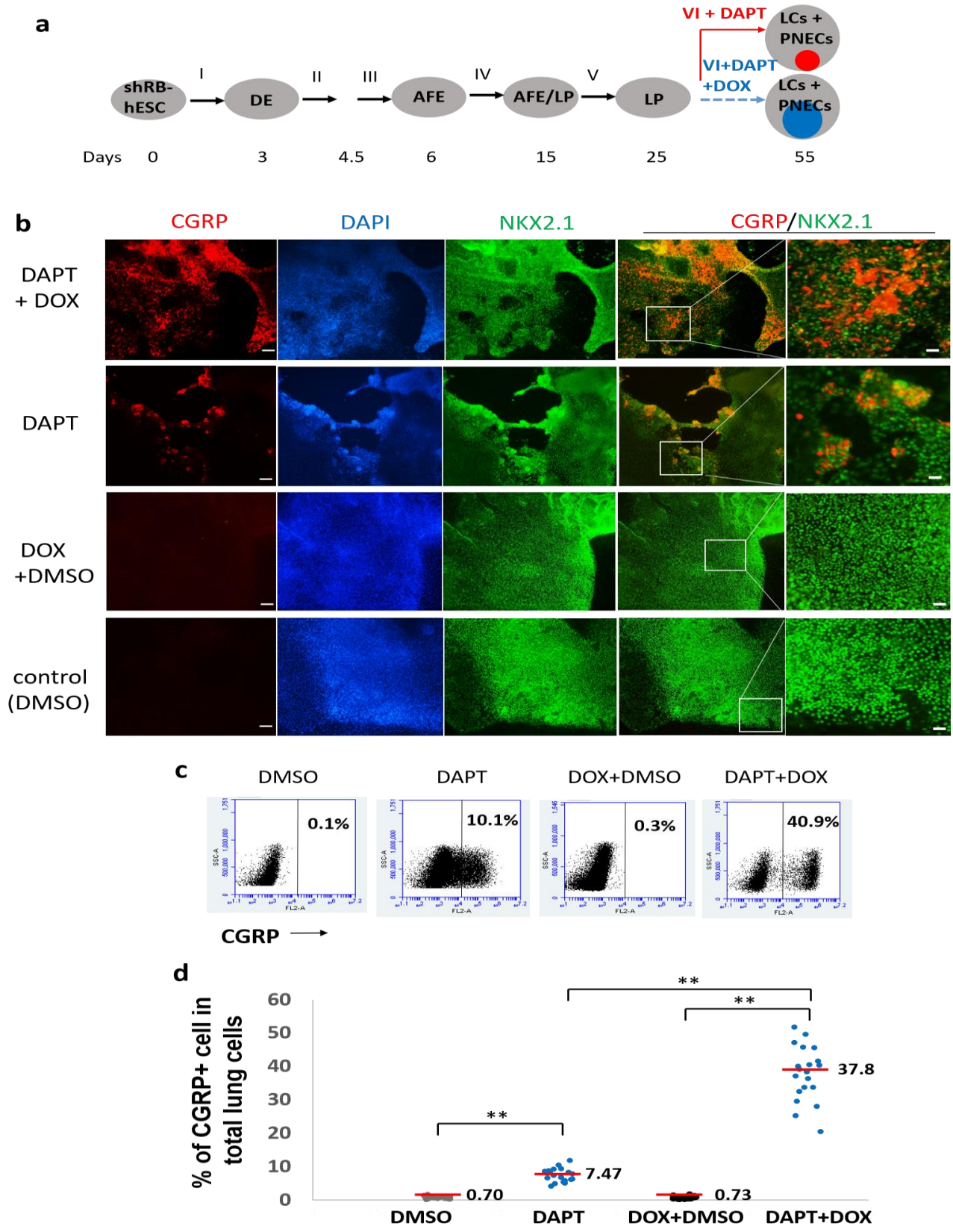

a
hESC
DE
AFE
AFE/LP
LP
LCs
LCs + PNECs
VI + DAPT
I
II
III
IV
V
VI
Days
0
3
4.5
6
15
25
55
b
d55 cells
Genes
PNECs
Scaled Transcript Count
2
1
0
-1
-2
0
1
2
3
4
5
6
7
8
c
SST
ASCL1 *
SEC11C
CGRP *
HES6
KLK12
GRP *
SCG3
SCG2
MIAT
HMGB3
IGFBP2
SOX2
ELF3
ABHD2
TAGLN3
FOXA2
BEX1
TPD52
PTP4A3
PROX1
SYP *
SCG5
SCNN1A
UCHL1 *
0
5
10
Fold changes (Log2)
d
tSNE_2
tSNE_1
e
CGRP
ASCL1
NEUROD1
SYP
MYC
GRP
INSM1
HES6
DLL3
MYCL
tSNE_2
tSNE_1

Figure 2. Single cell RNA profiling of RUES2-derived lung cells in which NOTCH signaling was blocked by DAPT. **a.** Schematic of the protocol used to generate PNECs by stepwise differentiation of human embryonic stem cells (hESCs) and DAPT treatment from Day 25 to Day 55. **b-e.** Single cell RNA profiling of day 55 lung cells derived from RUES2 cells treated with DAPT (5μM) from day 25 to day 55. **b**. Heatmap representing scaled expression of the most differentially expressed genes specific to different cell clusters. Rows represent genes and columns represent cells **c.** Putative PNEC markers differentially expressed in the PNEC-like cell cluster number 4 in panel **b**. Bars indicate log fold-change versus non-PNEC cells. Asterisks indicate canonical PNEC markers. (See full gene list in Supplementary File 1.) **d.** Reduced-dimensionality t-Distributed Stochastic Neighbor Embedding (tSNE) map colored by cluster assignment (see methods). **e.** Individual cells positive for PNECs markers and other genes associated with neuroendocrine differentiation are denoted by red dots.

Figure 3.


a
shRB-hESC
I
DE
II
III
AFE
IV
AFE/LP
V
LP
VI + DAPT
LCs + PNECs
VI+DAPT +DOX
LCs + PNECs
Days 0 3 4.5 6 15 25 55
b
CGRP
DAPI
NKX2.1
CGRP/NKX2.1
DAPT + DOX
DAPT
DOX +DMSO
control (DMSO)
c
DMSO
DAPT
DOX+DMSO
DAPT+DOX
0.1%
10.1%
0.3%
40.9%
CGRP
d
% of CGRP+ cell in total lung cells
60
50
40
30
20
10
0
**
**
**
0.70
7.47
0.73
37.8
DMSO
DAPT
DOX+DMSO
DAPT+DOX

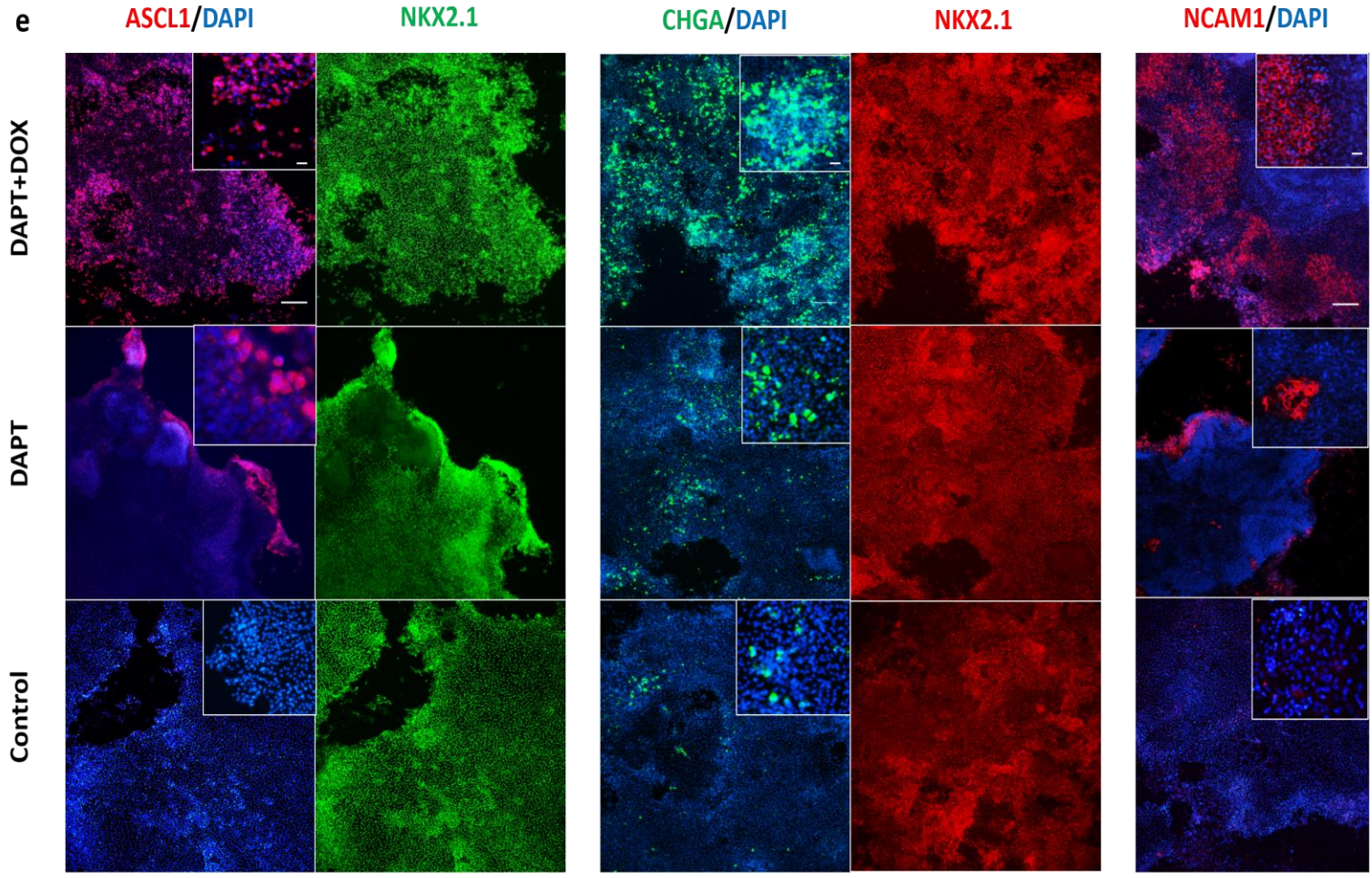


Figure 3. Inhibited expression of the RB tumor suppressor gene augments induction of PNECs by DAPT. a. Schematic of PNEC production from hESC cells carrying a tetracycline-inducible transgene that produces shRNA targeting RB1 (shRB). The format is similar to **Fig.1a,** except that mixture VI is supplemented with DAPT (5μM) with or without doxycycline (DOX). Internal colored circles at day 55 indicate PNECs induced by DAPT (red) or by DAPT and DOX (blue). **b-d.** Increased numbers of putative PNECs detected by co-staining for CGRP and NKX2.1 (panel **b**) or by FACS sorting with anti-human CGRP antibody (panels **c** and **d**) as described in the legend to Fig. 1 and Methods. ** $P < 0.01$ by one-way ANOVA test. Panel **d**, horizontal red lines denote average values; n= 20; Scale bars: 100 μM (left) and 20 μM (right)**. e.** Expression of shRNA-RB increases the percentages of ASCL1+, CHGA+ or NCAM+ cells. The indicated markers were detected by immunostaining as in panel **b.** Scale bars: 100 μM and 20 μM (in small window)

Figure 4.

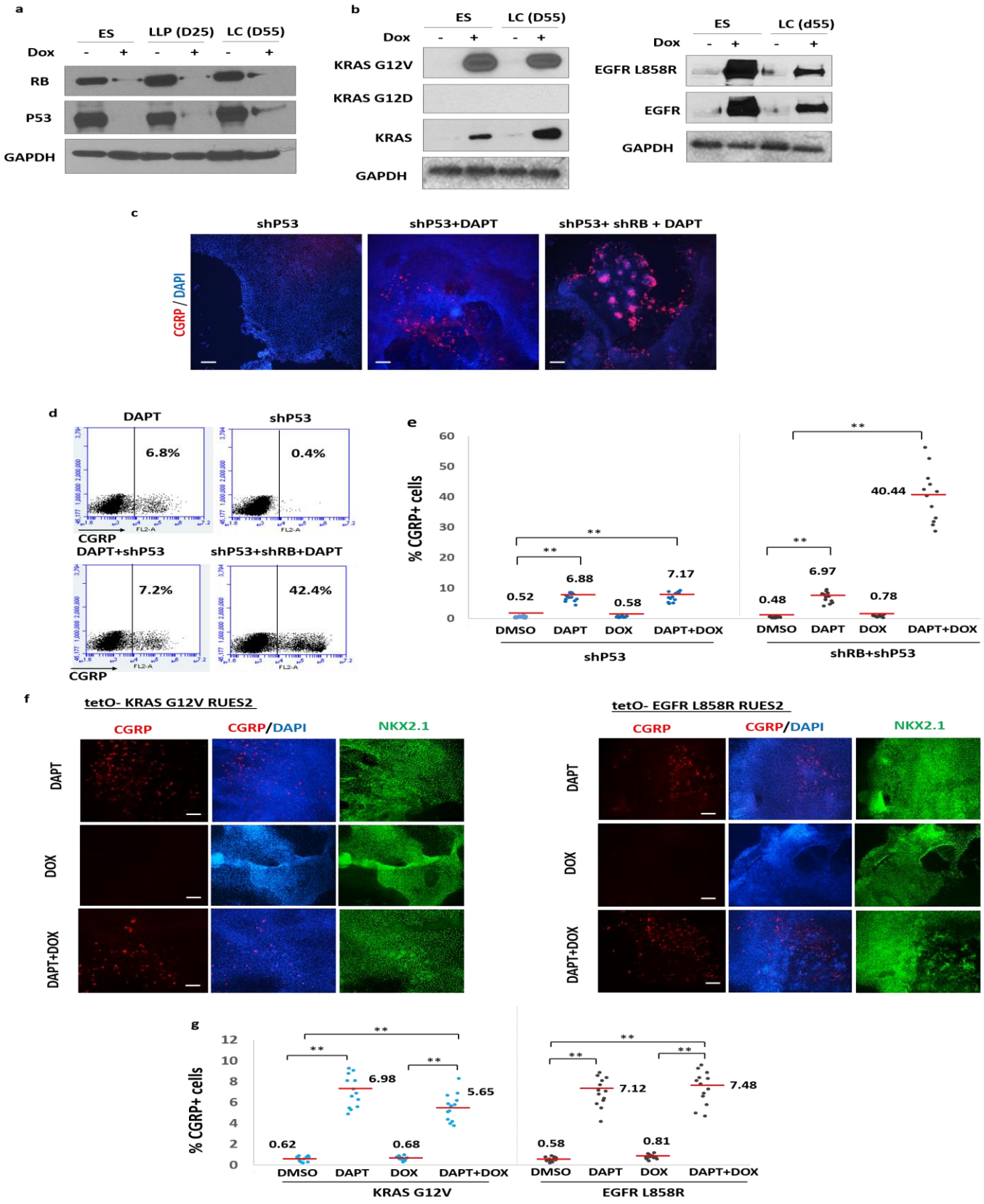
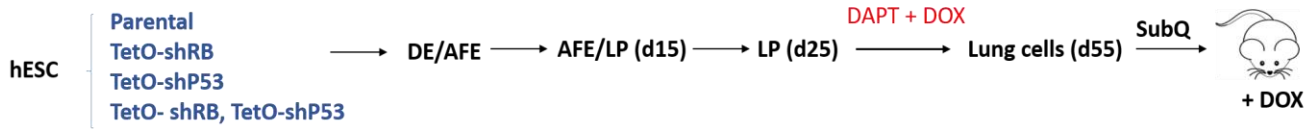

a
ES
LLP (D25)
LC (D55)
Dox
RB
P53
GAPDH
b
ES
LC (D55)
Dox
KRAS G12V
KRAS G12D
KRAS
GAPDH
ES
LC (d55)
Dox
EGFR L858R
EGFR
GAPDH
c
shP53
shP53+DAPT
shP53+ shRB + DAPT
CGRP / DAPI
d
DAPT
shP53
6.8%
0.4%
CGRP
DAPT+shP53
shP53+shRB+DAPT
7.2%
42.4%
CGRP
e
% CGRP+ cells
0.52
6.88
0.58
7.17
0.48
6.97
0.78
40.44
DMSO
DAPT
DOX
DAPT+DOX
shP53
shRB+shP53
f
tetO- KRAS G12V RUES2
tetO- EGFR L858R RUES2
CGRP
CGRP/DAPI
NKX2.1
DAPT
DOX
DAPT+DOX
g
% CGRP+ cells
0.62
6.98
0.68
5.65
0.58
7.12
0.81
7.48
DMSO
DAPT
DOX
DAPT+DOX
KRAS G12V
EGFR L858R

Figure 4. Inhibition of the p53 tumor suppressor gene or expression of two common lung cancer oncogenes do not augment production of PNECs in the presence and absence of DAPT. RUES2 cells carrying DOX-inducible transgenes that encode shRNAs targeting *P53* or *RB* or mutant alleles of *KRAS* (G12V) or *EGFR* (L858R) were used as in the experimental protocol shown in **Fig.2a** to measure the effects of DOX-induction of the indicated transgenes from day 25 to day 55 on production of PNECs. **a.** DOX-dependent expression of oncogenes or shRNA's in RUES2 cells. Western blots show RB and P53 tumor suppressor proteins after DOX-induction of shRNA cassettes. **b.** Production of proteins encoded by DOX-regulated transgenes encoding KRAS-G12V (left panel) and EGFR-L858R (right panel). Proteins were detected by the indicated antibodies (see Methods). **c-g.** Decreased P53 or production of mutant KRAS or EGFR proteins fail to increase the percentage of PNECs. The indicated RUES2 cell lines were tested for the appearance of PNECs at day 55 by immunofluorescence staining for CGRP (panels **c** and **f**) and by FACS (panels **d, e, g**). See text for interpretation. Scale bars: 100 μM in panels **c;** 200 μM in panels **f**, ** $P < 0.01$ by one-way ANOVA test; horizontal red lines denote average values; n= 12 for **e**, and 13 for **g**.

Figure 5.

a



Tumor formation in immunodeficient mice

Cells	% CGRP+ cells in total injected cells	Tumors / injection (≥ 250mm^3)
Parental (DAPT alone)	7.6 ± 1.0	0/12
DAPT+shRB	39.6 ± 4.4	0/14
DAPT+shP53	8.0 ± 1.3	0/11
DAPT+shRB+shP53	41.9 ± 4.6	14/19 **

b

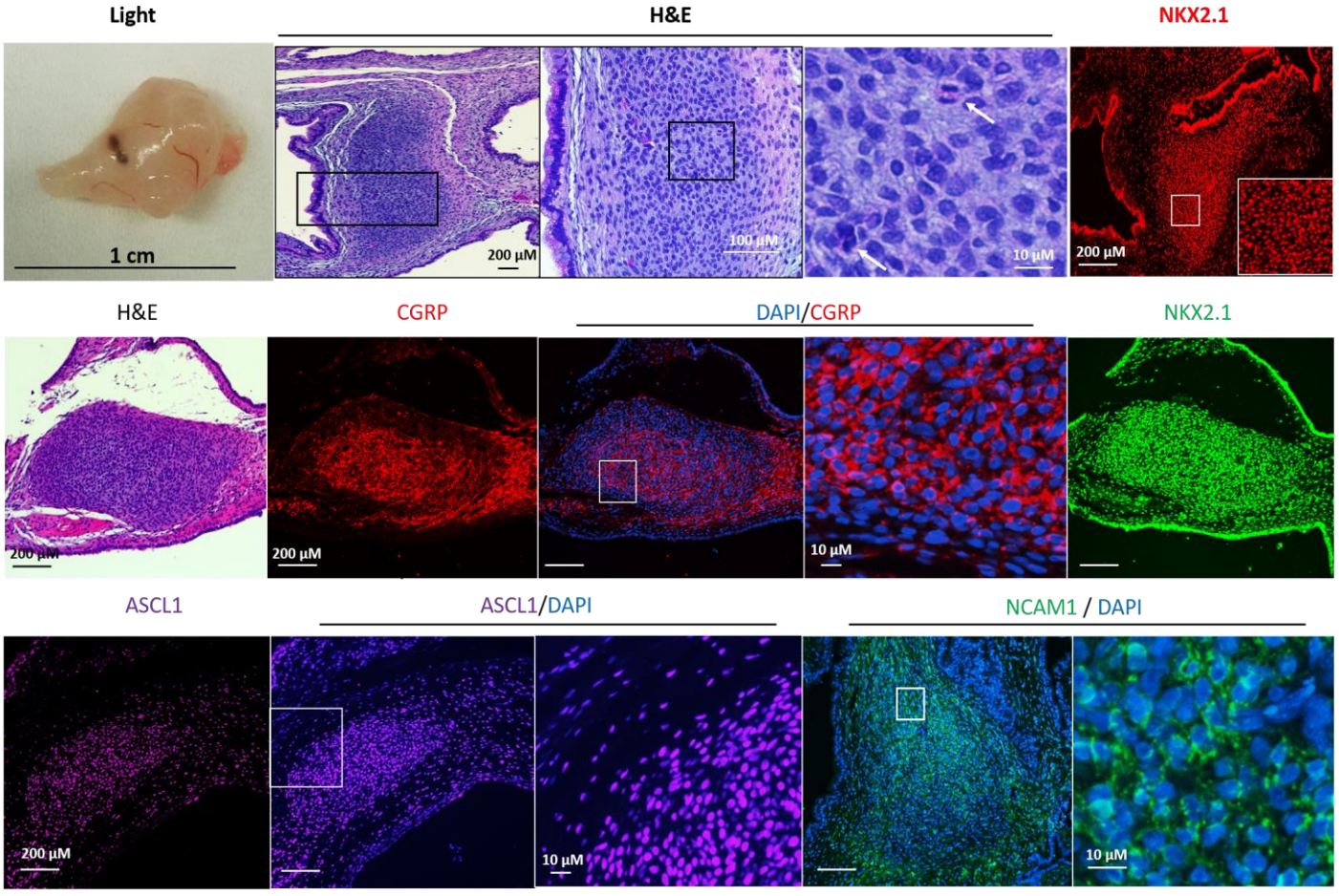
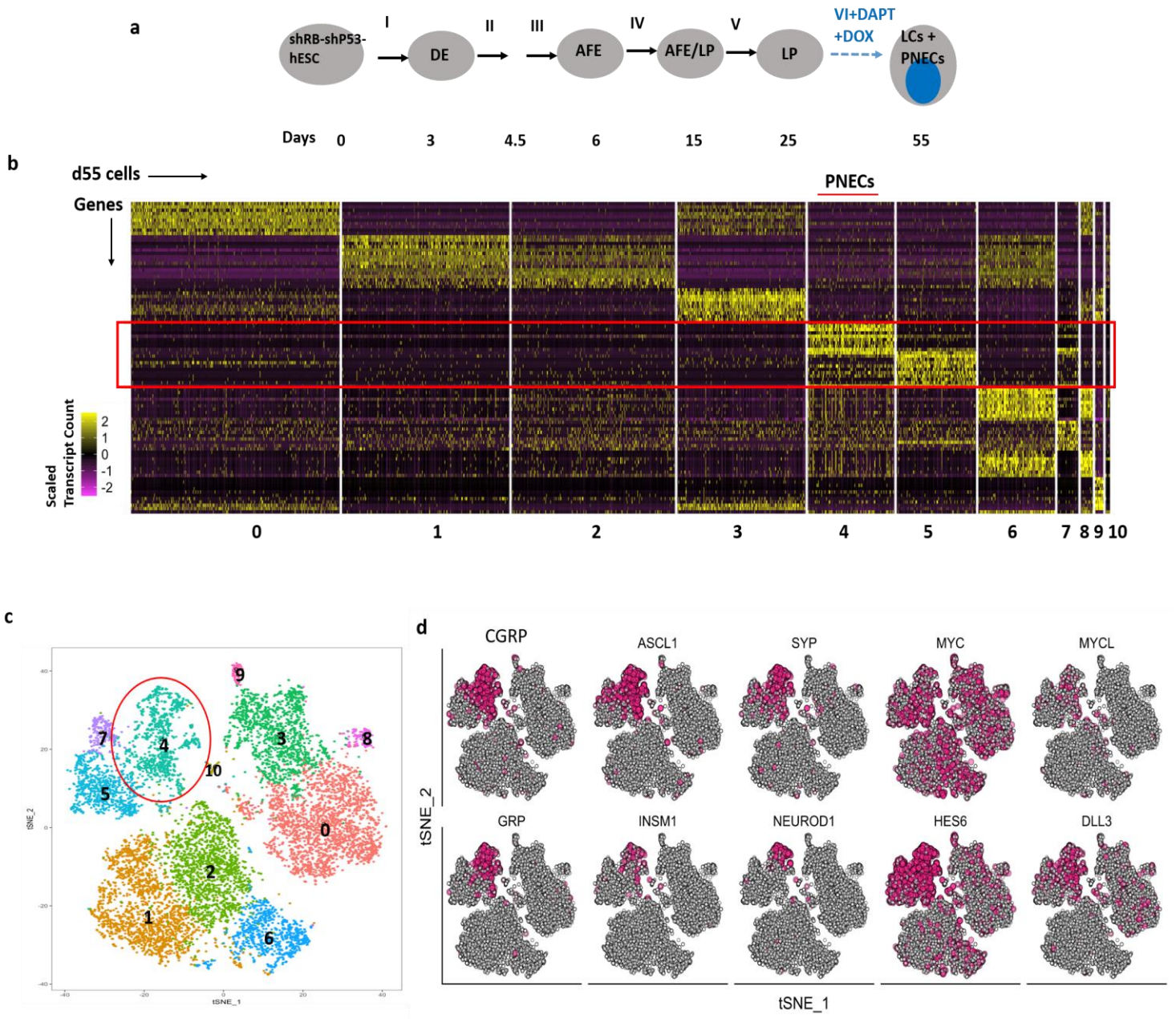


Figure 5. Characterization of xenografts formed with hESC – derived lung cells. a. Schematic representation of tumorigenesis experiment. The indicated transgenic and control lines of RUES2 hESCs were differentiated and grown in DAPT and DOX from day 25 to day 55. At day 55, total lung cells were injected subcutaneously into NSG mice. Xenografts grew into visible tumors after 6-7 weeks only from cells containing transgenes encoding shRNAs for both RB and P53. ** $P < 0.01$ by Fisher's test. **b.** Characterization of xenografts. **Upper panel:** left segment, a representative tumor by light microscopy; middle three segments, H & E staining of that tumor at different magnifications, white arrows designate the cells with mitotic figures; right segment, staining for NKX2.1; **Middle and Lower panels:** Additional samples stained with H & E, DAPI, or antibodies specific for the indicated marker proteins.

Figure 6.


a
shRB-shP53-hESC
I
DE
II
III
AFE
IV
AFE/LP
V
LP
VI+DAPT +DOX
LCs + PNECs
Days 0 3 4.5 6 15 25 55
b
d55 cells
Genes
PNECs
Scaled Transcript Count
2
1
0
-1
-2
0 1 2 3 4 5 6 7 8 9 10
c
d
CGRP
ASCL1
SYP
MYC
MYCL
GRP
INSM1
NEUROD1
HES6
DLL3
tSNE_2
tSNE_1

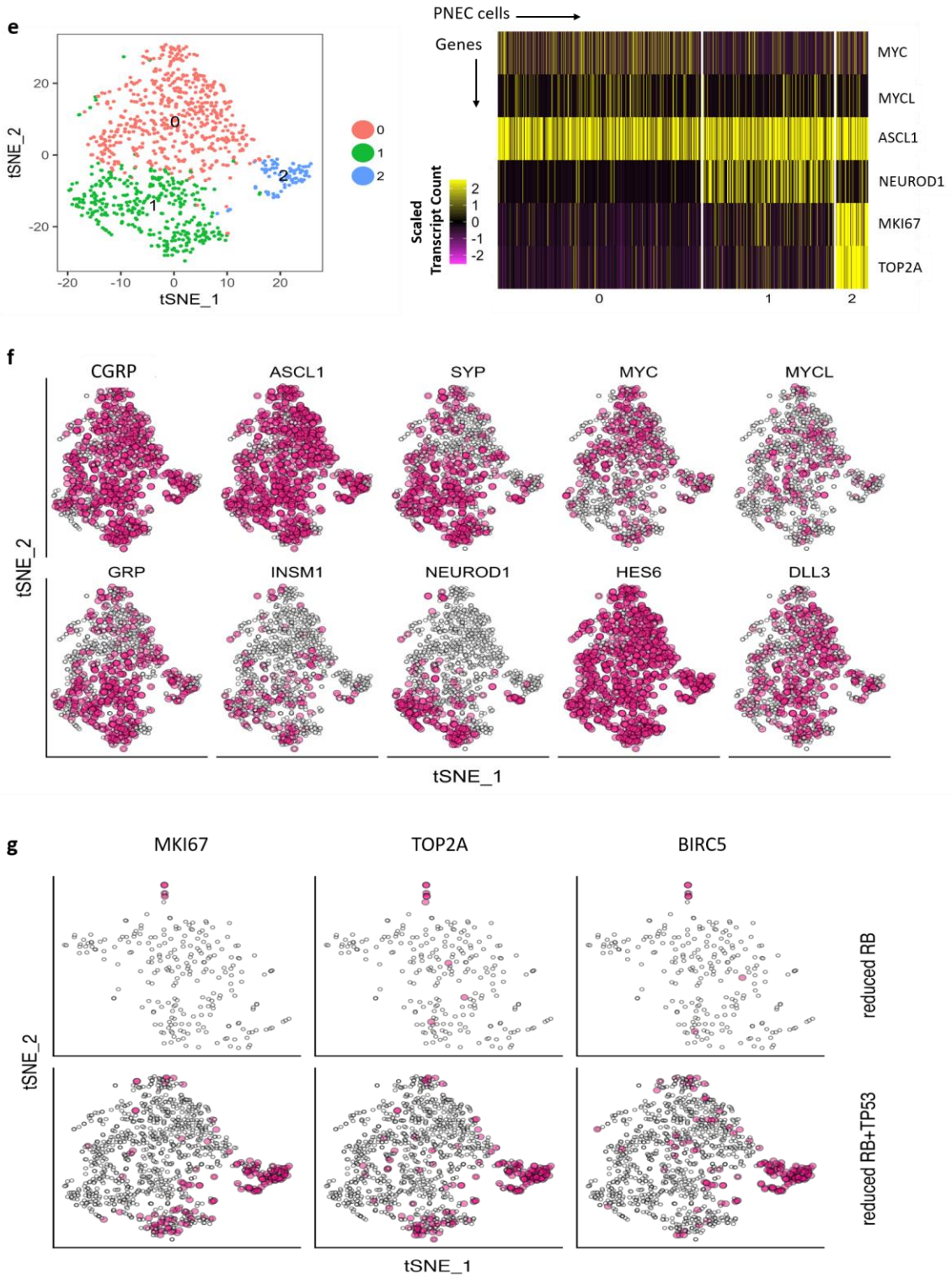
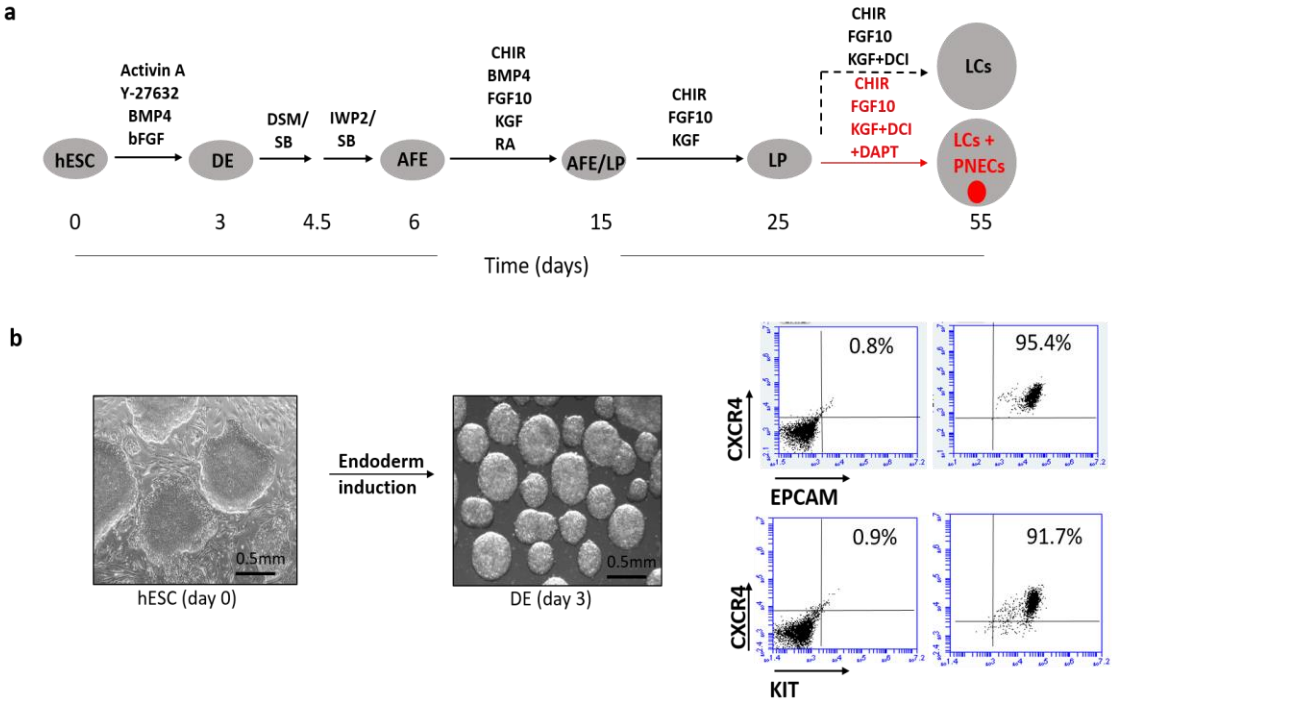
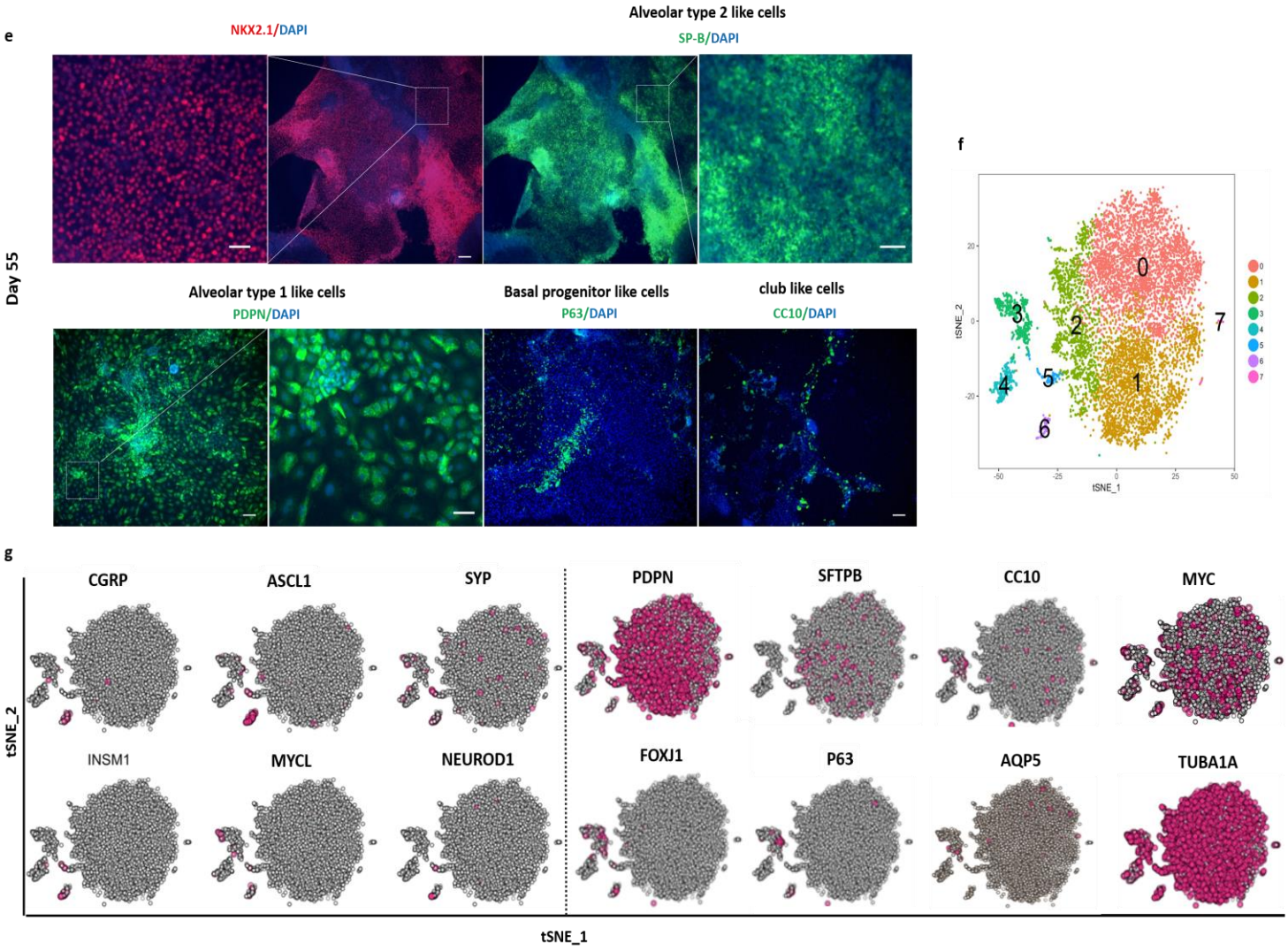

e
tSNE_2
tSNE_1
0
1
2
PNEC cells
Genes
MYC
MYCL
ASCL1
NEUROD1
MKI67
TOP2A
Scaled Transcript Count
f
CGRP
ASCL1
SYP
MYC
MYCL
GRP
INSM1
NEUROD1
HES6
DLL3
g
MKI67
TOP2A
BIRC5
reduced RB
reduced RB+TP53

Figure 6. Single cell RNA profiling of RUES2-derived lung cells in which NOTCH signaling was blocked by DAPT and RB and P53 protein levels were reduced by shRNA. a. Schematic of PNEC production from RUES2 cells carrying DOX-inducible transgenes encoding shRNAs against *RB* and *TP53* after exposure to DAPT and DOX (as in **Fig. 2a**). **b.** Single cell RNA profiling of day 55 lung cells treated with DAPT (5μM) and DOX from Day 25 to Day 55. Scaled expression of the most differentially expressed genes (horizontal rows) specific to different clusters of cells (vertical columns). **c.** PNEC-like cells are principally in cluster 4 but additional cells expressing some PNEC markers are found in other clusters. **c.** tSNE map colored by cluster assignment as in Figure 1, panel **j** . **d.** Individual cells positive for RNA encoding the indicated PNEC markers and for other proteins associated with neuroendocrine differentiation are denoted with red dots as in Fig. 1**k**. **e. (left)** tSNE map of re-clustering of the PNEC-enriched cluster only (cluster 4 in **c**.); **(right)** Scaled expression of the most differentially expressed key genes (horizontal rows) specific to cells in three PNEC sub-clusters called 0, 1 and 2 on the left (vertical columns). **f.** Individual cells positive for PNECs markers and other genes associated with neuroendocrine differentiation are denoted by red dots. g. Individual cells positive for cell-cycle related genes are denoted by red dots in PNECs in which RB levels were reduced (top row) or both RB and TP53 levels were reduced (bottom row).

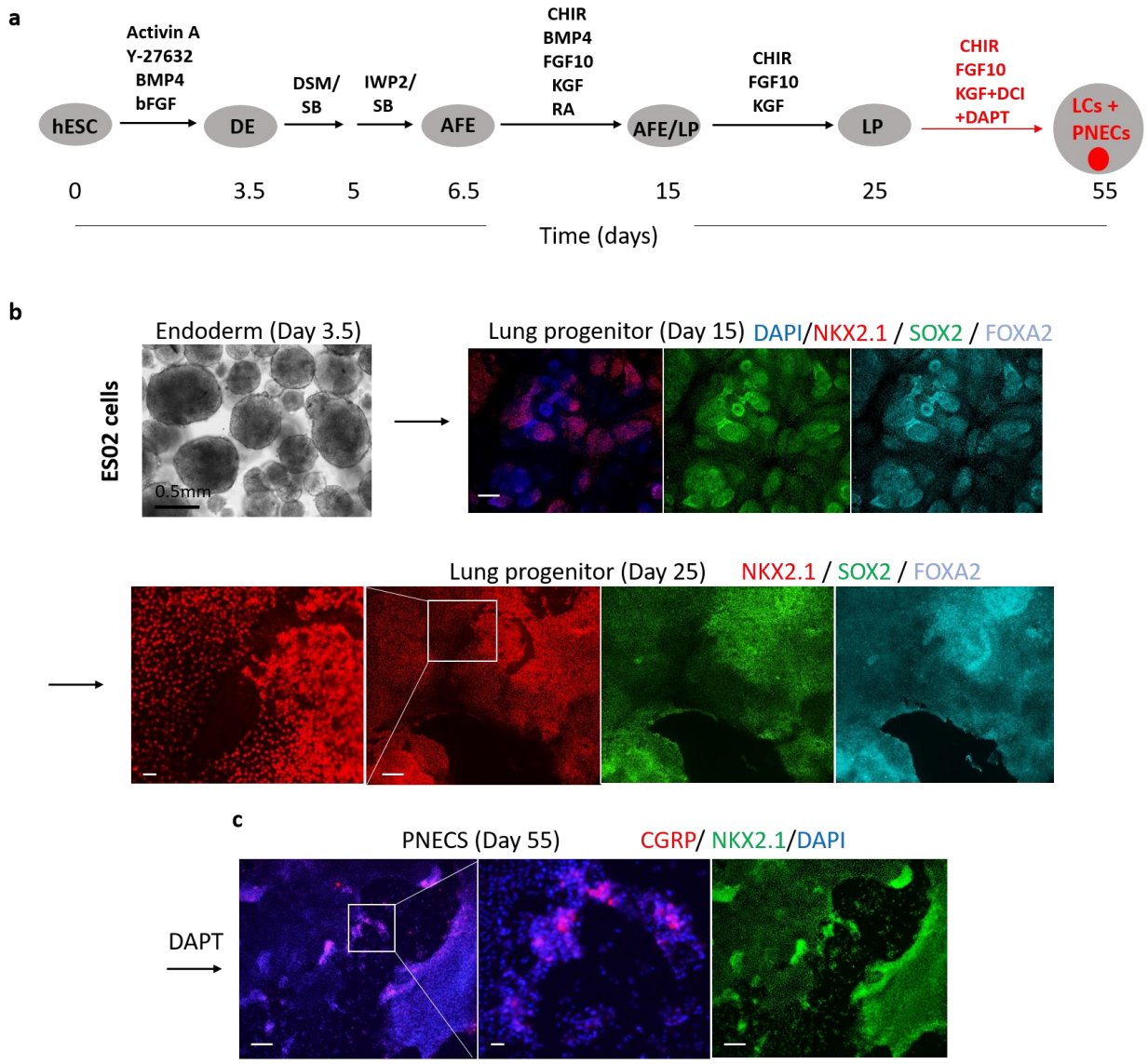
Supplementary Figure 1.

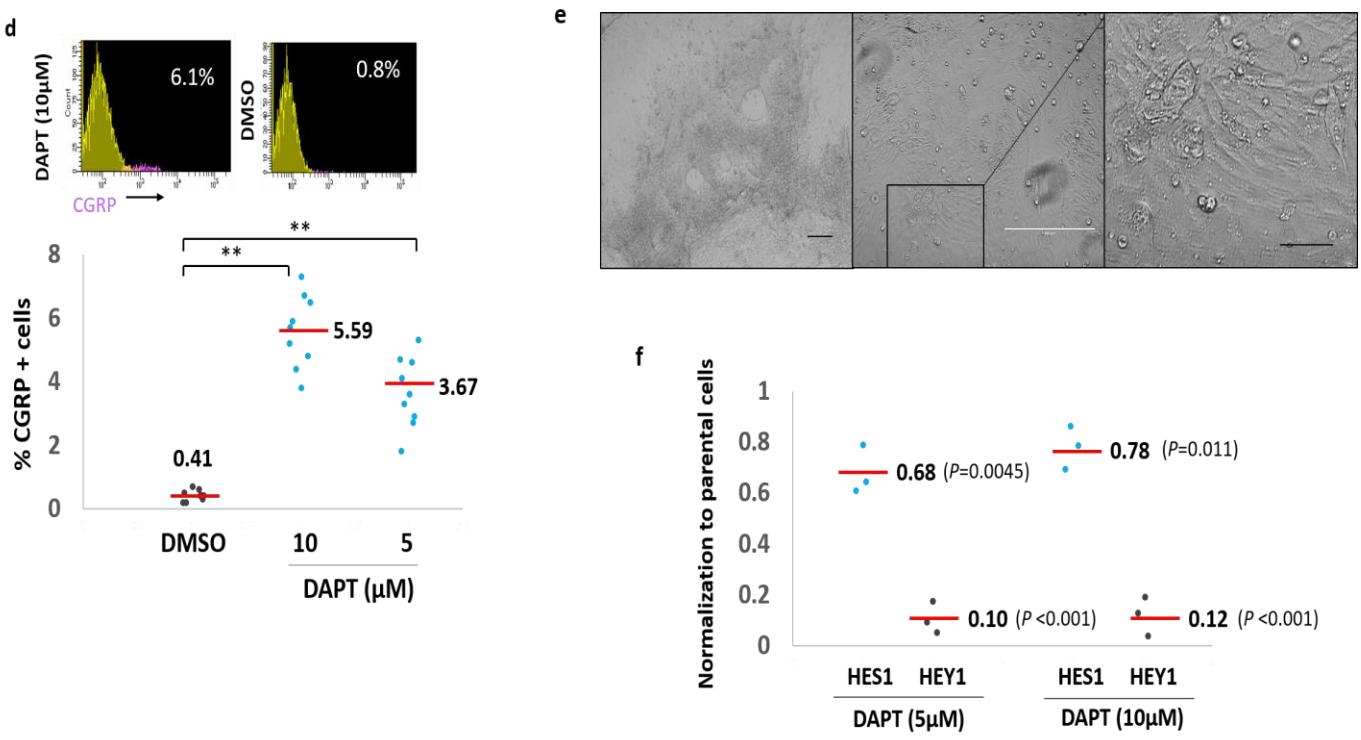



e
Day 55
NKX2.1/DAPI
Alveolar type 2 like cells
SP-B/DAPI
Alveolar type 1 like cells
PDPN/DAPI
Basal progenitor like cells
P63/DAPI
club like cells
CC10/DAPI
f
tSNE_2
tSNE_1
0
1
2
3
4
5
6
7
g
CGRP
ASCL1
SYP
PDPN
SFTPB
CC10
MYC
INSM1
MYCL
NEUROD1
FOXJ1
P63
AQP5
TUBA1A
tSNE_2
tSNE_1

Supplementary Figure 1. Differentiation of hESCs to form definitive endoderm and lung cells. a. Similar to **Fig. 1a**, but schematic of the protocol to generate lung cells from hESCs includes major components of differentiation mixtures I to VI. **b.** RUES2 cells were induced to definitive endoderm (DE) and DE cells were shown to be triply positive for CXCR4, EPCAM and KIT by FACS with the indicated antisera (right hand graphs). Samples stained with control IgG served as negative control (left hand graphs). Endoderm was differentiated to form lung progenitor cells (LPs) between day 15 **(c.)** and day 25 **(d.)** as shown by increasing proportions of cells positive for NKX2.1, SOX2 and FOXA2. Scale bars: 200 μM (upper panels) and 50 μM (lower panels)**. e.** Day 25 lung progenitor cells derived from hESCs (RUES2 line) were further differentiated into cells resembling major types of lung cells, including AT2 (SP-B+) and AT1 (PDPN+) lung epithelial cells; basal progenitor cells (P63+); and club cells (CC10+). All were detected at day 55 by immunostaining with the indicated antisera. Scale bars: 50 μM (long) and 100 μM (short). **f.** t-Distributed Stochastic Neighbor Embedding (tSNE) map of single cell RNA profiling of day 55 lung cells not exposed to DAPT, colored by cluster assignment; **g.** Individual cells positive (red dots) for RNA corresponding to PNEC-related genes **(left panel)** and canonical markers for other lung cell types **(Right panel)**. AT1 cells: PDPN&AQP5; AT2 cells: SFTPB; club cells: CC10; Ciliated cells: FOXJ1; and basal progenitor cells: P63. TUBA1A serves as internal control for gene expression during lung development**.**

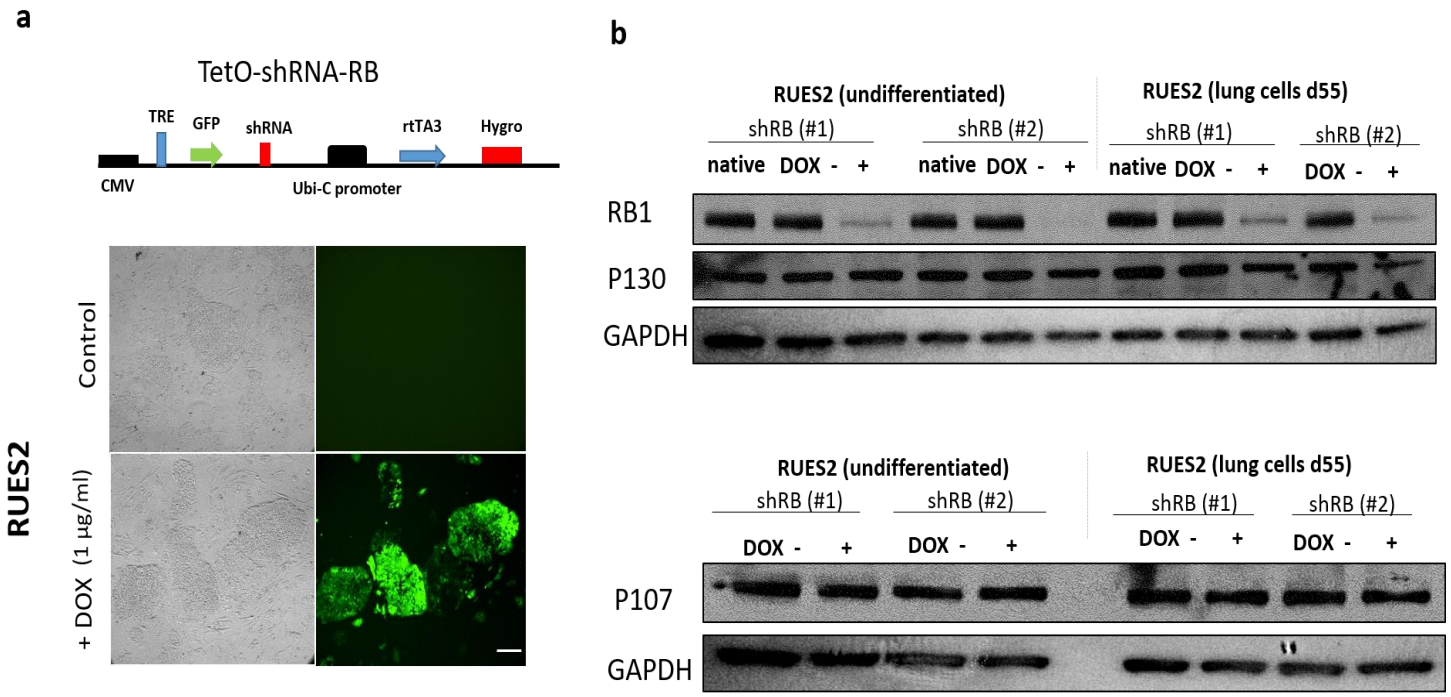
Supplementary Figure 2.

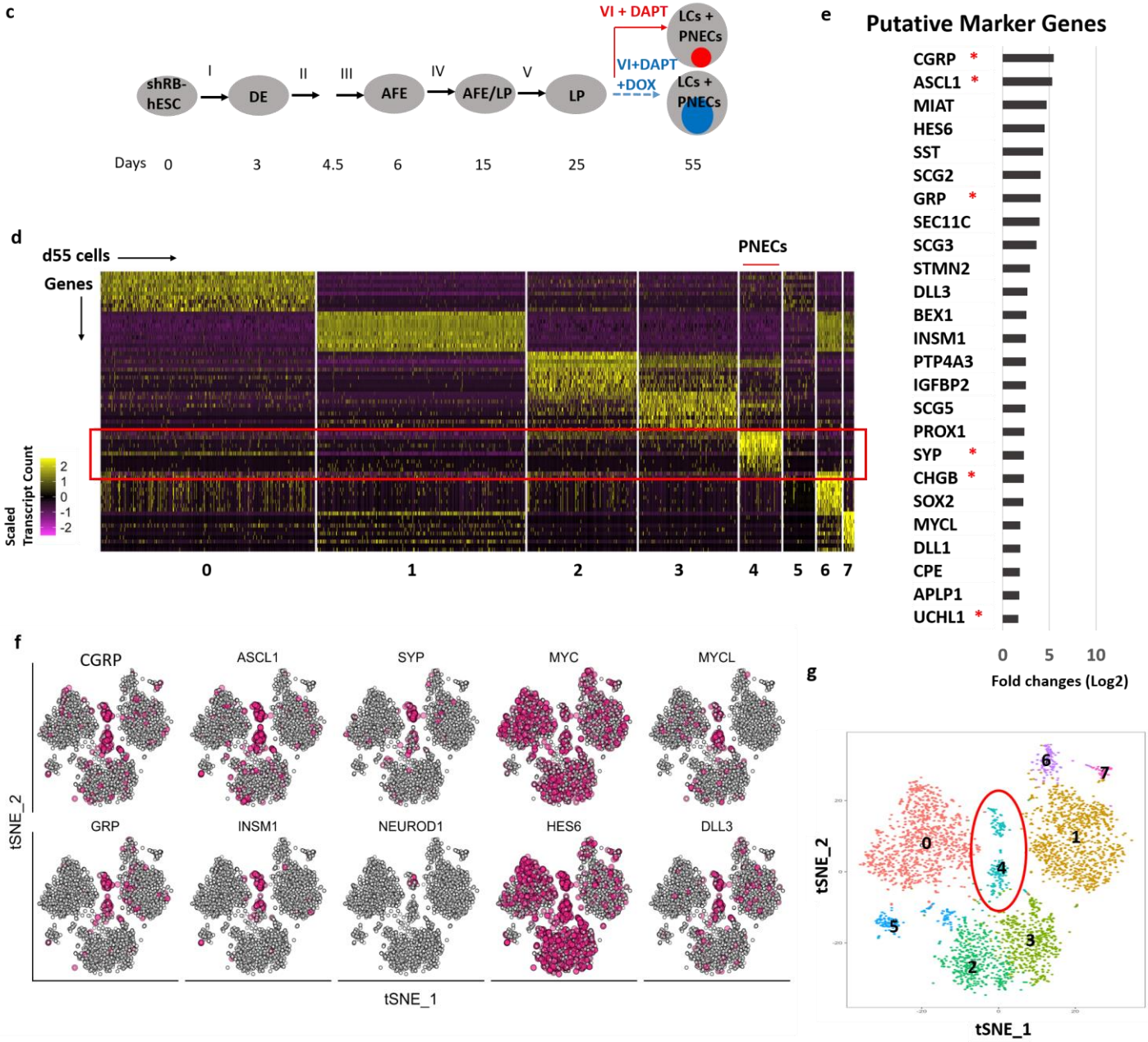


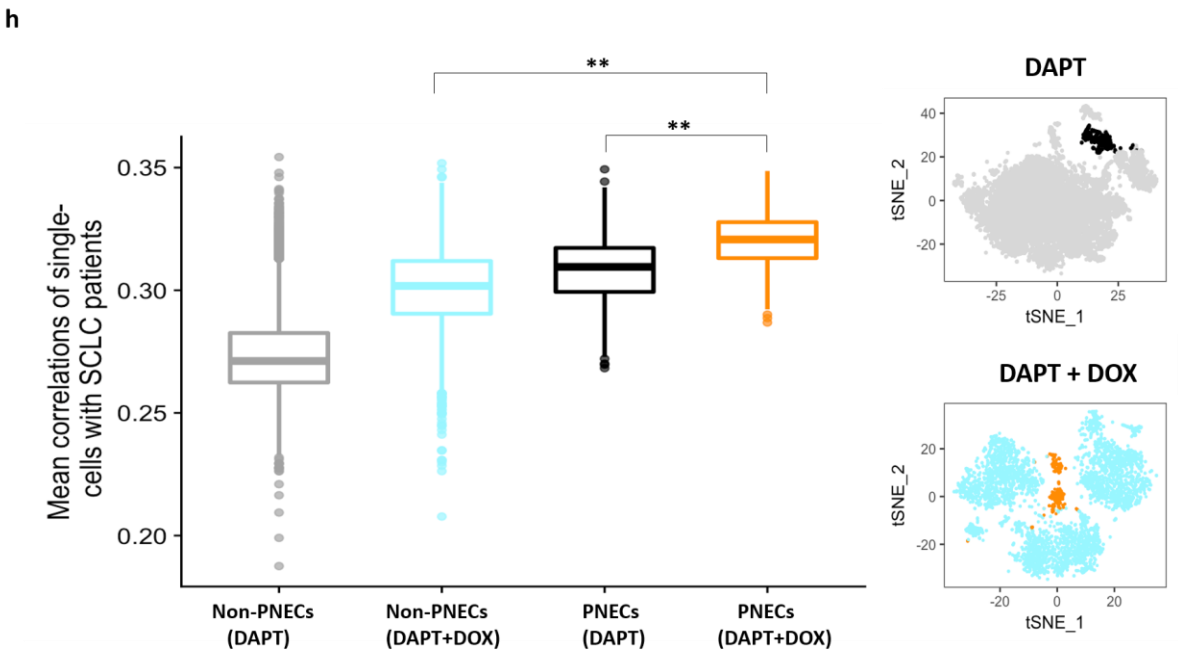


Supplementary Figure 2. DAPT induces PNECs during differentiation of the ES02 line of hESCs and gene expression after induction of PNECs with DAPT, an inhibitor of NOTCH signaling. a., Schematic of the protocol, as in **Fig.1a,** to generate PNECs with DAPT during differentiation of hESCs. **b.** Endoderm cells derived from the ES02 line of hESCs were further differentiated into lung progenitor cells (LPs), with increasing percentages of LPs from day 15 to day 25. **c.** PNECs, assessed by IHC staining for CGRP, were induced by DAPT from day 25 to day 55. Scale bars: 100 µM (long) and 20 µM (short). **d.** Percentages of CGRP+ cells in cultures grown with 0, 5, or 10 µM DAPT were determined at day 55 by FACS and displayed as flow cytometry data (10µM DAPT only; red, CGRP+; yellow, CGRP-) and by scatter graph (lower). ** $P < 0.01$ by student *t* test. Horizontal red lines denote average values; number of biological repeats (n) =9. **e**. Bright field images of the RUES2 hESC-derived lung cells treated with 5 µM DAPT from day 25 to day 55 (left), and the fractionated culture of CGRP+ cells after FACS sorting (middle and right). Scale bars: 400 µM in the left and middle panels; 50 µM in the right panel. **f.** Decreased expression of NOTCH target genes HES1 and HEY1 in day 55 lung cells after exposure to DAPT from day 25 during differentiation of the RUES2 line of hESCs. Results obtained by Quantitative-rtPCR assay; the relative levels were determined by comparison with parallel tests with day 55 lung cells not exposed to DAPT. P values by two-way ANOVA test; horizontal red lines denote average values; n= 3.

Supplementary Figure 3.

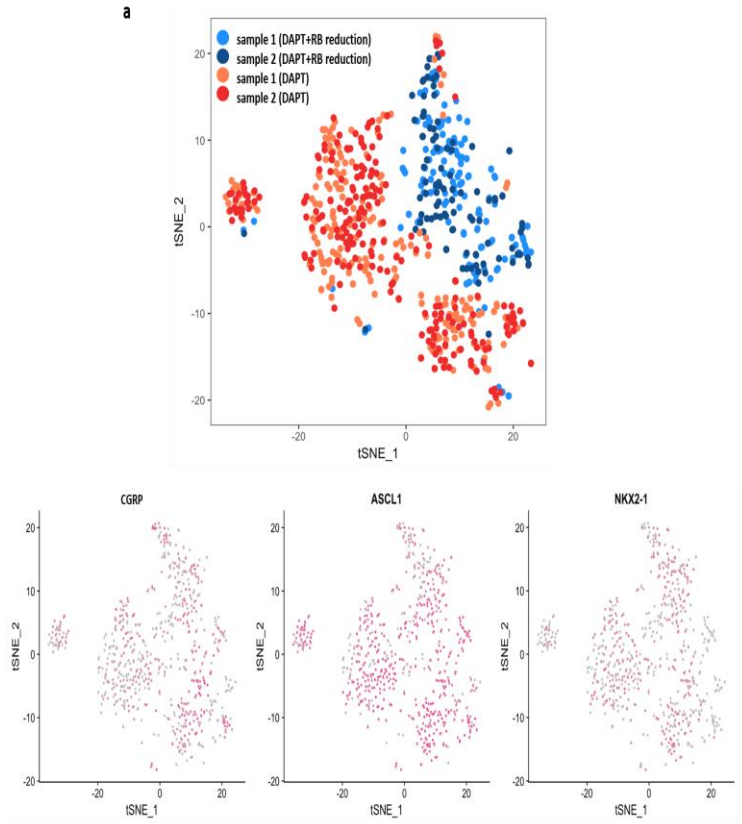



c
shRB-hESC
I
DE
II
III
AFE
IV
AFE/LP
V
LP
VI + DAPT
LCs + PNECs
VI+DAPT +DOX
LCs + PNECs
Days 0 3 4.5 6 15 25 55
d
d55 cells
Genes
PNECs
Scaled Transcript Count
2
1
0
-1
-2
0 1 2 3 4 5 6 7
e
Putative Marker Genes
CGRP *
ASCL1 *
MIAT
HES6
SST
SCG2
GRP *
SEC11C
SCG3
STMN2
DLL3
BEX1
INSM1
PTP4A3
IGFBP2
SCG5
PROX1
SYP *
CHGB *
SOX2
MYCL
DLL1
CPE
APLP1
UCHL1 *
0 5 10
Fold changes (Log2)
f
CGRP
ASCL1
SYP
MYC
MYCL
GRP
INSM1
NEUROD1
HES6
DLL3
tSNE_2
tSNE_1
g
0
1
2
3
4
5
6
7
tSNE_2
tSNE_1



Supplementary Figure 3. Generation and characterization of PNECs through directed differentiation of hESCs in which NOTCH signaling was blocked by DAPT and RB protein levels were reduced by shRNA. a. Schematic of vector constructed with shRNA targeting RB (see Methods) controlled by tetracycline operator (TetO) (upper). GFP-mediated fluorescence documents DOX-dependent expression in transgenic cells. **b**. Reduced production of RB1, but not of RB-related proteins P130 and P107 was documented by western blotting, using the indicated antisera and extracts from undifferentiated cells and day 55 LCs derived from the transgenic RUES1 line with and without doxycycline (DOX) treatment for 30 days. **c.** Schematic of PNEC production from RUES2 cells carrying a DOX-inducible transgene encoding RB-specific shRNA after exposure to DAPT and DOX (as in **Fig. 2a**). **d.** Single cell RNA profiling of day 55 lung cells treated with DAPT (5µM) and DOX from Day 25 to Day 55. Scaled expression of the most differentially expressed genes (rows) specific to different cell clusters in each cell (columns). PNEC like cells are principally in cluster 4 but additional cells expressing some PNEC markers are found in other clusters (0, 1, 2, 3). **e.** PNEC markers found to be differentially expressed in the PNEC-like cell cluster. Bars indicate log fold-change; asterisks indicate canonical PNEC marker. (See full gene list in Supplementary file 1). **g.** tSNE map colored by cluster assignment as in Figure 1 **j** . **f.** Individual cells positive for PNEC markers and other genes associated with neuroendocrine differentiation are denoted with red dots as in Fig. 1**k**. **g.** PNEC and non-PNEC single cell transcriptomes in cultures treated with DAPT alone or DAPT and DOX (to reduce RB expression) were correlated with 29 bulk RNA profiles of early stage human SCLCs (Stage 1a or 1b) (George, et.,al.2015). The distributions of the mean correlation for each cell are presented. See text for interpretation. (Left, Spearman's Correlation, ** $P < 2.2e\text{-}16$, by two sided Kolmogorov–Smirnov test). tSNE maps indicate PNEC and non-PNEC populations in cultures in which RB expression was not perturbed (top right) and cultures in which RB shRNA was induced by DOX (bottom right).

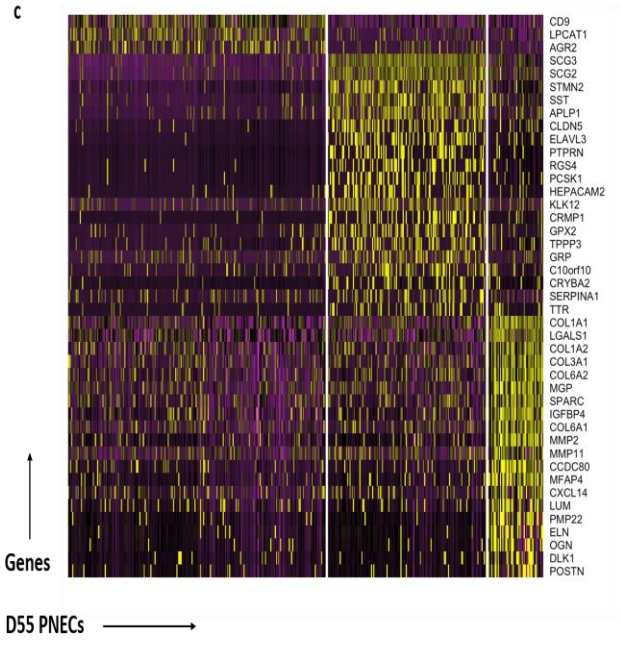
Supplementary Figure 4.

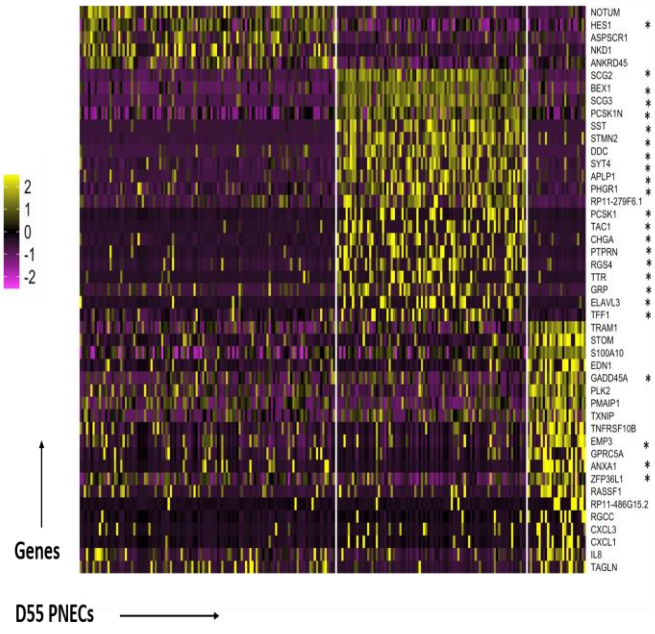


b

Differential Gene Analysis

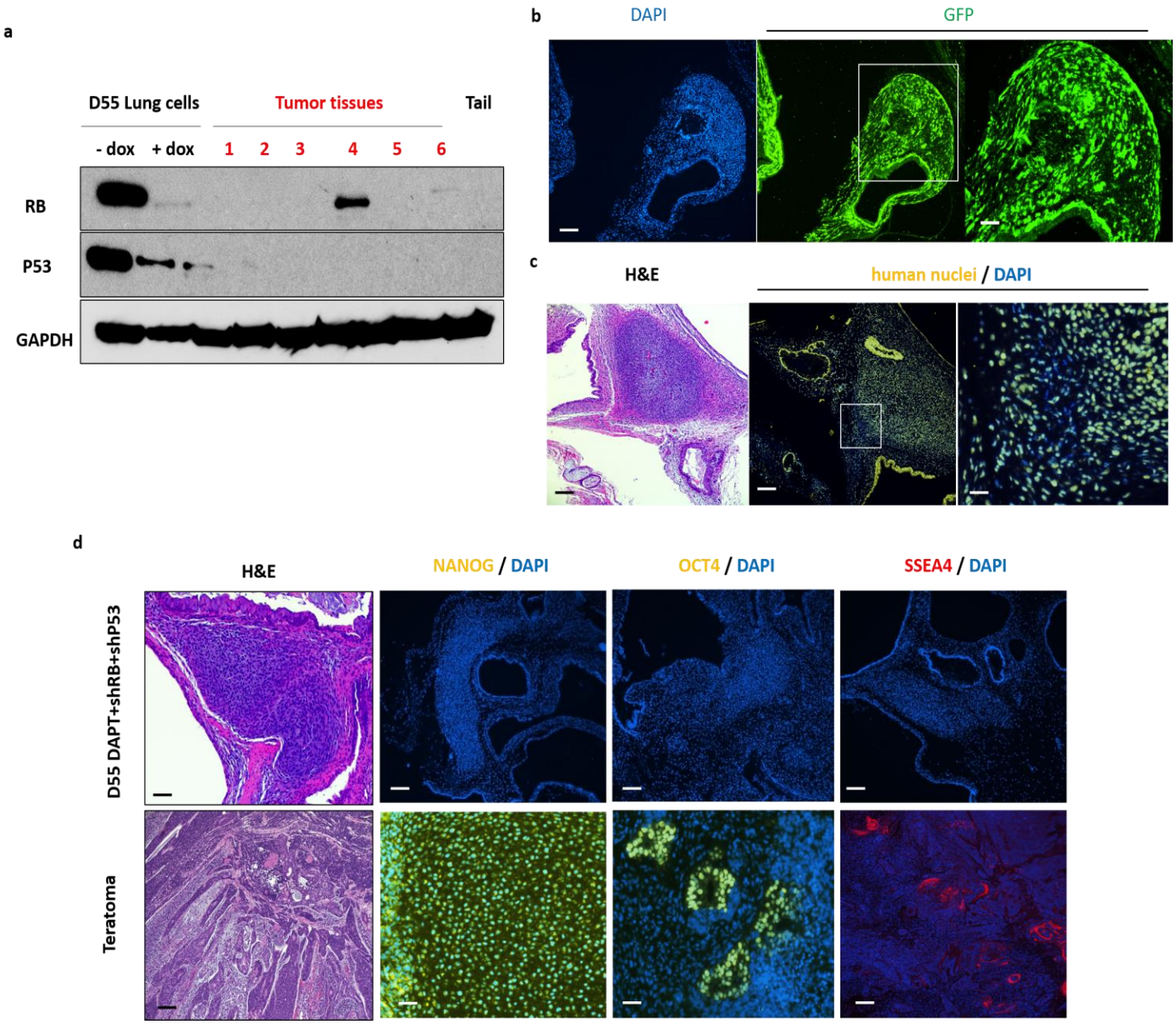
Biological Processes	Genes	*P*-value
• Regulation of apoptotic process	TNFRSF12A, RHOB, ATF3, SFPQ, PMAIP1, ANKRD1, KLF6, etc	3.820E-07
• regulation of signal transduction by P53 class mediator	ATF3, PMAIP1	1.021E-05
• Regulation of nerve system differentiation	ID4, DUSP10, SOX4, HES1, Catenin Beta 1, etc	1.27E-05
• Growth factor binding	COL3A1, COL5A1, COL6A1, IGFBP4, COL1A1	1.612E-05
• extracellular matrix organization	MMP2,11, POSTN, COL5A1, 6A1, MFAP2, FBLN1, SPARC, LUM, etc	5.467E-13
• Impaired lung alveolus development	FBLN1, HIF3A, NEUROD1, MMP2	4.411E-06

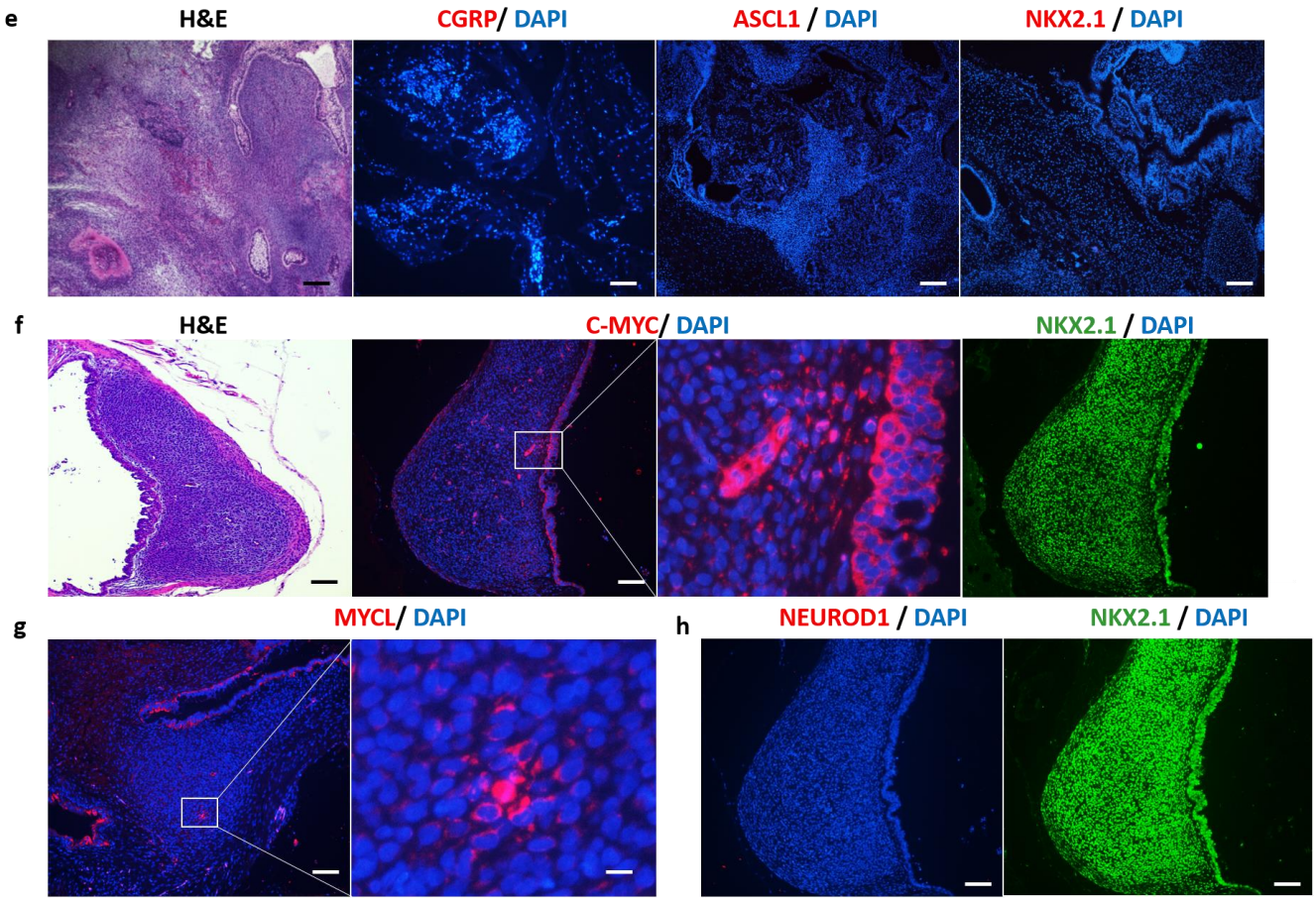




Supplementary Figure 4. Comparisons of single-cell RNA profiles from cultures of RUES2 cells in which RB levels were and were not reduced by induction of shRNA. a. Reduced-dimensionality tSNE map of PNEC cells from day 55 lung cells treated with DAPT (5μM) and DOX (to reduce RB expression, blue) or with DAPT alone (red). **Upper:** Colors denote biological replicates as listed in the key. **Lower:** Individual cells positive for PNECs markers (CGRP or ASCL1) and lung lineage marker NKX2.1 are denoted by red dots. **b.** Pathways and genes identified by Topp Gene software[27] based on differential gene expression from comparison of PNEC cells identified by scRNA profiling of cultures in which RB levels were and were not reduced (Supplementary file 2.) **c.** Scaled expression of the most differentially expressed genes (rows) specific to different PNEC cell sub-clusters in each cell (columns). The left-hand heatmap shows results with PNECs from cultures exposed to DAPT; the right hand map shows results with PNECs from cultures exposed to DAPT, with DOX used to reduce RB levels by induction of RB shRNA. Asterisks denote the most differential genes expressed in the subsets of PNECs treated with DAPT and DOX are also expressed accordingly in the subsets of PNECs treated with DAPT alone. The heterogeneity within the PNEC-like cell cluster is discussed in the text.

Supplementary Figure 5.





Supplementary Figure 5. Characterization of xenografts formed with hESC-derived lung cells. Panels a-c: DAPT-treated day55 lung cells, induced with DOX to express shRNA specific for RB1 or P53 or both, were injected subcutaneously into NSG mice. Xenografts developed after 6-7 weeks only with DAPT-treated lung cells in which RB and P53 were reduced by DOX-induction of shRNAs. **a.** P53 and RB proteins were measured in extracts of day55 lung cells (the first two lanes, without and with DOX treatment; six xenografts (1-6); and a mouse tail in right hand lane) by western blot using antibodies against human RB or P53; GAPDH served as in internal loading control. **b.** Origin of the tumors formed with day55 lung cells was validated by fluorescence-based detection of GFP encoded in the RB-shRNA vector. **c.** Cells in tumor nodules formed afer injection of day55 lung cells were confirmed to be of human origin by immunostaining with antibody against human nuclear specific protein. Scale bars: 100µM (long) & 20µM (short). Panels **d-h.** Comparison of hESC-derived lung cell xenografts with undifferentiated hESC-derived teratomas. **d.** Morphology of xenograft tumors formed with Day 55 lung cells (upper, left) compared to teratomas (lower, left) formed with undifferentiated RUES2 cells (H & E staining); tumors from day55 lung cells did not stain with antisera against embryonic tissue markers NANOG, OCT4, and SSEA4 (upper) that were detected in tumors (teratomas; lower) grown from undifferentiated RUES2 cells**.** e. CGRP, ASCL1 and NKX2.1 --- detected in the xenograft tumors

from differentiated lung cells (Figure 4) ---were not detected in teratomas. Scale bars: 100μM (long) & 20μM (short). **f-h**. Some cells in the xenograft tumors formed with differentiated lung calls expressed C-MYC **(f)** and MYCL **(g)**, but did not express detectable amounts of NEUROD1 **(h)**. Scale bars: 100μM (long) & 20μM (short).

References

Borges, M., R.I. Linnoila, H.J. van de Velde, H. Chen, B.D. Nelkin, M. Mabry, S.B. Baylin, and D.W. Ball. 1997. An achaete-scute homologue essential for neuroendocrine differentiation in the lung. *Nature* 386:852-855.

Borromeo, M.D., T.K. Savage, R.K. Kollipara, M. He, A. Augustyn, J.K. Osborne, L. Girard, J.D. Minna, A.F. Gazdar, M.H. Cobb, and J.E. Johnson. 2016. ASCL1 and NEUROD1 Reveal Heterogeneity in Pulmonary Neuroendocrine Tumors and Regulate Distinct Genetic Programs. *Cell reports* 16:1259-1272.

Cancer Genome Atlas Research, N. 2014. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 511:543-550.

Cancer Genome Atlas Research, N., J.N. Weinstein, E.A. Collisson, G.B. Mills, K.R. Shaw, B.A. Ozenberger, K. Ellrott, I. Shmulevich, C. Sander, and J.M. Stuart. 2013. The Cancer Genome Atlas Pan-Cancer analysis project. *Nature genetics* 45:1113-1120.

Chen, J., E.E. Bardes, B.J. Aronow, and A.G. Jegga. 2009. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic acids research* 37:W305-311.

Funato, K., T. Major, P.W. Lewis, C.D. Allis, and V. Tabar. 2014. Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. *Science* 346:1529-1533.

Garraway, L.A., and E.S. Lander. 2013. Lessons from the cancer genome. *Cell* 153:17-37.

Gazdar, A.F., P.A. Bunn, and J.D. Minna. 2017. Small-cell lung cancer: what we know, what we need to know and the path forward. *Nature reviews. Cancer* 17:725-737.

Geling, A., H. Steiner, M. Willem, L. Bally-Cuif, and C. Haass. 2002. A gamma-secretase inhibitor blocks Notch signaling in vivo and causes a severe neurogenic phenotype in zebrafish. *EMBO reports* 3:688-694.

George, J., J.S. Lim, S.J. Jang, Y. Cun, L. Ozretic, G. Kong, F. Leenders, X. Lu, L. Fernandez-Cuesta, G. Bosco, C. Muller, I. Dahmen, N.S. Jahchan, K.S. Park, D. Yang, A.N. Karnezis, D. Vaka, A. Torres, M.S. Wang, J.O. Korbel, R. Menon, S.M. Chun, D. Kim, M. Wilkerson, N. Hayes, D. Engelmann, B. Putzer, M. Bos, S. Michels, I. Vlasic, D. Seidel, B. Pinther, P. Schaub, C. Becker, J. Altmuller, J. Yokota, T. Kohno, R. Iwakawa, K. Tsuta, M. Noguchi, T. Muley, H. Hoffmann, P.A. Schnabel, I. Petersen, Y. Chen, A. Soltermann, V. Tischler, C.M. Choi, Y.H. Kim, P.P. Massion, Y. Zou, D. Jovanovic, M. Kontic, G.M. Wright, P.A. Russell, B. Solomon, I. Koch, M. Lindner, L.A. Muscarella, A. la Torre, J.K. Field, M. Jakopovic, J. Knezevic, E. Castanos-Velez, L. Roz, U. Pastorino, O.T. Brustugun, M. Lund-Iversen, E. Thunnissen, J. Kohler, M. Schuler, J. Botling, M. Sandelin, M. Sanchez-Cespedes, H.B. Salvesen, V. Achter, U. Lang, M. Bogus, P.M. Schneider, T. Zander, S.

Ansen, M. Hallek, J. Wolf, M. Vingron, Y. Yatabe, W.D. Travis, P. Nurnberg, C. Reinhardt, S. Perner, L. Heukamp, R. Buttner, S.A. Haas, E. Brambilla, M. Peifer, J. Sage, and R.K. Thomas. 2015. Comprehensive genomic profiles of small cell lung cancer. *Nature* 524:47-53.

Huang, S.X., M.D. Green, A.T. de Carvalho, M. Mumau, Y.W. Chen, S.L. D'Souza, and H.W. Snoeck. 2015. The in vitro generation of lung and airway progenitor cells from human pluripotent stem cells. *Nature protocols* 10:413-425.

Huang, S.X., M.N. Islam, J. O'Neill, Z. Hu, Y.G. Yang, Y.W. Chen, M. Mumau, M.D. Green, G. Vunjak-Novakovic, J. Bhattacharya, and H.W. Snoeck. 2014. Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. *Nature biotechnology* 32:84-91.

Huang, Y.H., O. Klingbeil, X.Y. He, X.S. Wu, G. Arun, B. Lu, T.D.D. Somerville, J.P. Milazzo, J.E. Wilkinson, O.E. Demerdash, D.L. Spector, M. Egeblad, J. Shi, and C.R. Vakoc. 2018. POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer. *Genes & development* 32:915-928.

Iso, T., L. Kedes, and Y. Hamamori. 2003. HES and HERP families: multiple effectors of the Notch signaling pathway. *Journal of cellular physiology* 194:237-255.

Ito, T., N. Udaka, T. Yazawa, K. Okudela, H. Hayashi, T. Sudo, F. Guillemot, R. Kageyama, and H. Kitamura. 2000. Basic helix-loop-helix transcription factors regulate the neuroendocrine differentiation of fetal mouse pulmonary epithelium. *Development* 127:3913-3921.

Lim, J.S., A. Ibaseta, M.M. Fischer, B. Cancilla, G. O'Young, S. Cristea, V.C. Luca, D. Yang, N.S. Jahchan, C. Hamard, M. Antoine, M. Wislez, C. Kong, J. Cain, Y.W. Liu, A.M. Kapoun, K.C. Garcia, T. Hoey, C.L. Murriel, and J. Sage. 2017. Intratumoural heterogeneity generated by Notch signalling promotes small-cell lung cancer. *Nature* 545:360-364.

Linnoila, R.I. 2006. Functional facets of the pulmonary neuroendocrine system. *Laboratory investigation; a journal of technical methods and pathology* 86:425-444.

Liu, A., L. Cheng, J. Du, Y. Peng, R.W. Allan, L. Wei, J. Li, and D. Cao. 2010. Diagnostic utility of novel stem cell markers SALL4, OCT4, NANOG, SOX2, UTF1, and TCL1 in primary mediastinal germ cell tumors. *The American journal of surgical pathology* 34:697-706.

Milano, J., J. McKay, C. Dagenais, L. Foster-Brown, F. Pognan, R. Gadient, R.T. Jacobs, A. Zacco, B. Greenberg, and P.J. Ciaccio. 2004. Modulation of notch processing by gamma-secretase inhibitors causes intestinal goblet cell metaplasia and induction of genes known to specify gut secretory lineage differentiation. *Toxicological sciences : an official journal of the Society of Toxicology* 82:341-358.

Mollaoglu, G., M.R. Guthrie, S. Bohm, J. Bragelmann, I. Can, P.M. Ballieu, A. Marx, J. George, C. Heinen, M.D. Chalishazar, H. Cheng, A.S. Ireland, K.E. Denning, A. Mukhopadhyay, J.M. Vahrenkamp, K.C. Berrett, T.L. Mosbruger, J. Wang, J.L. Kohan, M.E. Salama, B.L. Witt, M. Peifer, R.K. Thomas, J. Gertz, J.E. Johnson, A.F. Gazdar, R.J. Wechsler-Reya, M.L. Sos, and T.G. Oliver. 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.

Morimoto, M., R. Nishinakamura, Y. Saga, and R. Kopan. 2012. Different assemblies of Notch receptors coordinate the distribution of the major bronchial Clara, ciliated and neuroendocrine cells. *Development* 139:4365-4373.

Osborne, J.K., J.E. Larsen, J.X. Gonzales, D.S. Shames, M. Sato, Wistuba, II, L. Girard, J.D. Minna, and M.H. Cobb. 2013a. NeuroD1 regulation of migration accompanies the differential sensitivity of neuroendocrine carcinomas to TrkB inhibition. *Oncogenesis* 2:e63.

Osborne, J.K., J.E. Larsen, M.D. Shields, J.X. Gonzales, D.S. Shames, M. Sato, A. Kulkarni, Wistuba, II, L. Girard, J.D. Minna, and M.H. Cobb. 2013b. NeuroD1 regulates survival and migration of neuroendocrine lung carcinomas via signaling molecules TrkB and NCAM. *Proc Natl Acad Sci U S A* 110:6524-6529.

Peifer, M., L. Fernandez-Cuesta, M.L. Sos, J. George, D. Seidel, L.H. Kasper, D. Plenker, F. Leenders, R. Sun, T. Zander, R. Menon, M. Koker, I. Dahmen, C. Muller, V. Di Cerbo, H.U. Schildhaus, J. Altmuller, I. Baessmann, C. Becker, B. de Wilde, J. Vandesompele, D. Bohm, S. Ansen, F. Gabler, I. Wilkening, S. Heynck, J.M. Heuckmann, X. Lu, S.L. Carter, K. Cibulskis, S. Banerji, G. Getz, K.S. Park, D. Rauh, C. Grutter, M. Fischer, L. Pasqualucci, G. Wright, Z. Wainer, P. Russell, I. Petersen, Y. Chen, E. Stoelben, C. Ludwig, P. Schnabel, H. Hoffmann, T. Muley, M. Brockmann, W. Engel-Riedel, L.A. Muscarella, V.M. Fazio, H. Groen, W. Timens, H. Sietsma, E. Thunnissen, E. Smit, D.A. Heideman, P.J. Snijders, F. Cappuzzo, C. Ligorio, S. Damiani, J. Field, S. Solberg, O.T. Brustugun, M. Lund-Iversen, J. Sanger, J.H. Clement, A. Soltermann, H. Moch, W. Weder, B. Solomon, J.C. Soria, P. Validire, B. Besse, E. Brambilla, C. Brambilla, S. Lantuejoul, P. Lorimier, P.M. Schneider, M. Hallek, W. Pao, M. Meyerson, J. Sage, J. Shendure, R. Schneider, R. Buttner, J. Wolf, P. Nurnberg, S. Perner, L.C. Heukamp, P.K. Brindle, S. Haas, and R.K. Thomas. 2012. Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nature genetics* 44:1104-1110.

Pietanza, M.C., L.A. Byers, J.D. Minna, and C.M. Rudin. 2015. Small cell lung cancer: will recent progress lead to improved outcomes? *Clinical cancer research : an official journal of the American Association for Cancer Research* 21:2244-2255.

Schroeter, E.H., J.A. Kisslinger, and R. Kopan. 1998. Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain. *Nature* 393:382-386.

Semenova, E.A., R. Nagel, and A. Berns. 2015. Origins, genetic landscape, and emerging therapies of small cell lung cancer. *Genes & development* 29:1447-1462.

Shan, L., J.C. Aster, J. Sklar, and M.E. Sunday. 2007. Notch-1 regulates pulmonary neuroendocrine cell differentiation in cell lines and in transgenic mice. *American journal of physiology. Lung cellular and molecular physiology* 292:L500-509.

Shultz, L.D., B.L. Lyons, L.M. Burzenski, B. Gott, X. Chen, S. Chaleff, M. Kotb, S.D. Gillies, M. King, J. Mangada, D.L. Greiner, and R. Handgretinger. 2005. Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *Journal of immunology* 174:6477-6489.

Song, H., E. Yao, C. Lin, R. Gacayan, M.H. Chen, and P.T. Chuang. 2012. Functional characterization of pulmonary neuroendocrine cells in lung development, injury, and tumorigenesis. *Proc Natl Acad Sci U S A* 109:17531-17536.

Sutherland, K.D., N. Proost, I. Brouns, D. Adriaensen, J.Y. Song, and A. Berns. 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.

Treutlein, B., D.G. Brownfield, A.R. Wu, N.F. Neff, G.L. Mantalas, F.H. Espinoza, T.J. Desai, M.A. Krasnow, and S.R. Quake. 2014. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* 509:371-375.

Warburton, D., C. Wuenschell, G. Flores-Delgado, and K. Anderson. 1998. Commitment and differentiation of lung cell lineages. *Biochemistry and cell biology = Biochimie et biologie cellulaire* 76:971-995.

(References for Supplementary Figures)

Chen, H.J. et al. Chemokine 25-induced signaling suppresses colon cancer invasion and metastasis. The Journal of clinical investigation 122, 3184-3196 (2012).

Conklin, J.F., Baker, J. & Sage, J. The RB family is required for the self-renewal and survival of human embryonic stem cells. Nature communications 3, 1244 (2012).

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F. & Regev, A. Spatial reconstruction of single-cell gene expression data. Nature biotechnology 33, 495-502 (2015).

Unni, A.M., Lockwood, W.W., Zejnullahu, K., Lee-Lin, S.Q. & Varmus, H. Evidence that synthetic lethality underlies the mutual exclusivity of oncogenic KRAS and EGFR mutations in lung adenocarcinoma. eLife 4, e06907 (2015).

METHODS

Generation of lung cells, including PNECs, from hESCs. Protocols for maintenance of hESCs and generation of lung cells were slightly modified from previous studies (Huang, et.,al 2014; Huang, et.,al 2015). Two hESC lines---RUES2 (Rockefeller University Embryonic Stem Cell Line 2, NIH approval number NIHhESC-09-0013, Registration number 0013; passage 7-10) and ES02 (HES2, NIH registry, WiCell Research Insititute. INC. passage 3-7)---- were cultured on irradiated mouse embryonic fibroblasts (Global Stem, cat. no. GSC-6001G) at a density of 20,000- 25,000 cells/cm2 in a medium of DMEM/F12, 20% knockout serum replacement (Life Technologies, Grand Island, NY), 0.1 mM β-mercaptoethanol (Sigma-Aldrich, St. Louis, MO) and 20 ng/ml FGF-2 (R&D Systems, Minneapolis, MN), and medium was changed daily. hESC cultures were maintained in an undifferentiated state at 37 °C in a 5% CO2/air environment until stem cells reached about 90% confluence.

hESC differentiation into endoderm was performed in serum-free differentiation (SFD) media of DMEM/F12 (3:1) (Life Technologies) supplemented with N2 (Life Technologies), B27 (Life Technologies), ascorbic acid (50 μg/ml, Sigma), Glutamax (2 mM, Life Technologies), monothioglycerol (0.4 μM, Sigma), 0.05% bovine serum albumin (BSA) (Life Technologies) at

37 °C in a 5% CO2/5% O2/95% N2 environment. hESCs were treated with Accutase (Stemcell technology) and plated onto low attachment 6-well plates (Corning Incorporated, Tewksbury MA), resuspended in endoderm induction media containing Y-27632, 10 μM (R&D Systems), human BMP4, 0.5 ng/ml (R&D Systems); human bFGF, 2.5 ng/ml (R&D Systems); human Activin A, 100 ng/ml (R&D Systems) for 72 – 84 hours dependent on the formation rates of endoderm cells. On day 3 or 3.5, the embryoid bodies were dissociated into single cells using 0.05% Trypsin/0.02% EDTA and plated onto fibronectin (Sigma)-coated, 24-well tissue culture plates (~100,000–150,000 cells/well). For induction of anterior foregut endoderm, the endoderm cells were cultured in SFD medium supplemented with 1.5 μM Dorsomorphin dihydrochloride (R&D Systems) and 10 μM SB431542 (R&D Systems) for 36-48 h, and then switched to 36-48 h of 10 μM SB431542 and 1 μM IWP2 (R&D Systems) treatment.

For induction of early stage lung progenitor cells (day 6–15), the resulting anterior foregut endoderm was treated with CHIR99021, 3 μM (WNT signaling agonist), human FGF10, 10 ng/ml; human FGF7, 10 ng/ml; human BMP4, 10 ng/ml; and all-*trans* retinoic acid (ATRA), 50-60nM (all from R&D system and ATRA from Sigma) in SFD medium for 8–10 d. The day 10–15 cultures were maintained in a 5% CO2/air environment. On days 15 and 16, the lung field progenitor cells were replated after brief one minute trypsinization onto fibronectin-coated plates, in the presence of SFD containing either a combination of five factors (CHIR99021, 3 μM; human FGF10, 10 ng/ml; human FGF7, 10 ng/ml; human BMP4, 10 ng/ml; and ATRA, 50 nM), or three factors (CHIR99021, 3 μM, human FGF10, 10 ng/ml; human FGF7, 10 ng/ml) for day 14-16. Day 16–25 cultures of late stage lung progenitor cells were maintained in SFD media containing CHIR99021, 3 μM, human FGF10, 10 ng/ml; human FGF7, 10 ng/ml in a 5% CO2/air environment.

For differentiation of mature lung cells (day 25 to 55), cultures were re-plated after brief trypsinization onto 3.3% matrigel coated 24-well plates in SFD media containing maturation components containing 3 μM CHIR99021, 10 ng/ml human FGF10; 10 ng/ml human FGF7, and DCI (50 nM Dexamethasone, 0.1 mM 8-bromo-cAMP (Sigma) and 0.1 mM IBMX (3,7-dihydro-1-methyl-3-(2-methylpropyl)-1*H*-purine-2,6-dione) (Sigma)). DAPT or DBZ (5-10 μM, Sigma) was added to the maturation media for induction of pulmonary neuroendocrine cells (PNECs).

In Fig. 1a, Fig. 2a, Supplementary Figs.1a, 2a, 3c, I-VI denote the following mixtures of factors:

I	II	III	IV	V	VI
Activin A, Y-27632, BMP4, bFGF	DSM, SB	IWP2, SB	CHIR, BMP4, FGF10, KGF, RA	CHIR, FGF10, KGF	CHIR, FGF10, KGF, DCI

Single cell sequencing and transcriptomics. Single-cell capture, reverse transcription, cell lysis, and library preparation were performed using the Single Cell 3′ version 2 kit and chip according to the manufacturer's protocol (10x Genomics, USA). Single-cell suspensions were generated by dissociating the cultured RUES2 cells with 0.05% Trypsin/0.02% EDTA for 10-15 min, followed with passing through 40μM strainer. The single cell suspension was achieved through sorting the dissociated cells in flow cytometry singlets. Cell count was adjusted to 6000-9000 cells to achieve an estimated capture of 4000-5000 cells. Six input wells were used.

Sequencing was performed on a HiSeq 2500 (Illumina, paired-end protocol with 26 base pairs for read 1 and 98 for read 2). Alignment of the raw reads to the human reference genome (hg19), removing duplicated transcripts using the unique molecular identifiers (UMIs) and assignment to single cells was performed using CellRanger (10x Genomics, USA).

Single-cell analyses, including quality filtering, clustering, differential gene expression and reduced-dimensionality visualization, were performed using the Seurat package as described in the package tutorial (Version 2.1.0) (Satija, et.,al.2015). Briefly, cells to be included in the analysis were required to have at least 2000 and at most 40000 unique molecular identifiers (UMIs). In addition, cells were excluded if more than 7.5% of the determined RNA sequences mapped to mitochondrial genes. In total, 8,716 cells from the sample without DAPT treatment, 9,824 cells from the DAPT samples, 4,148 cells from the samples treated with DAPT and with DOX to induce RB-shRNA, and 11,361 from the samples treated with DAPT and with DOX to induce RB and P53-shRNA passed these filters for quality. Following the package suggestions, a linear model was used to mitigate the variation stemming from the number of detected unique molecules per cell. Principal Component Analysis was performed on a subset of the gene expression matrix using 1,500-2,000 genes with the most variable expression level. Using the top 50 principal components, clustering and tSNE visualization were performed. Clustering resolution, which affects the number of clusters was set at 0.6, 0.4, 0.6, 0.4 for the four samples, respectively (for more information, see package tutorial at http://satijalab.org/seurat/). Differential gene expression across clusters was performed using the Wilcoxon rank sum test implemented in the Seurat package.

Correlation with SCLC patient samples. Each of the single cell data sets from DAPT-treated, RUES2-derived lung cells and from cells treated with DAPT and with DOX to induce shRNA-RB were divided to two groups classified as PNECs and non-PNECs. To correct for the effects of UMI/cell, the single cell transcriptomes from either PNECS or non-PNECs in each treatment group (the samples treated with DAPT only or samples treated with DAPT and doxycycline) were subsampled to form transcriptomes with an equal number of UMIs, and correlated (Spearman's Correlation) with the previously published bulk transcriptomes of 29 early stage (stage Ia or Ib) SCLC patients (George, et.,al 2015). The 29 correlation values for each cell were averaged, and the distributions of the mean correlations were compared using the Kolmogorov-Smirnov test. Subsampling to 2000, 4000 or 10000 UMIs resulted in a similar pattern.

Gene list enrichment analysis. Genes that showed at least 2.5-fold changes in expression in comparisons of PNECs derived by DAPT treatment alone and PNECs derived by DAPT treatment plus reduction of RB RNA were used for ToppGene analysis (Chen, et.,al 2009) (https://toppgene.cchmc.org/).

Intra-PNEC heterogeneity. Following the clustering procedure described above, the PNEC single-cell transcriptomes from DAPT-treated cells, DAPT-treated cells in which RB RNA was reduced, and DAPT-treated cells in which RB and P53 RNA was reduced were assigned separately to sub-clusters, using up to 14 top principal components and a clustering resolution of 0.6, 0.6, 0.1. Each of the two datasets was assigned to three clusters, and a differential expression analysis was performed as described above. Percent of PNECs positive for NEUROD1 was calculated following

repetitive subsampling (50 repeats) of the transcriptomes to 7,500 UMIs per cell, to account for sequencing depth.

Lentivirus transduction of hESCs. The lentiviral vectors expressing tetracycline inducible (TetO) shRNAs against human *RB* constructed in the pSLIK vector system were obtained as a gift from Julien Sage's lab at Stanford University (Conklin, et., al 2012). The lentiviral vector expressing tetracycline inducible shRNA against human P53 was purchased from Gentarget. Inc. (cat# LVP-343-RB-PBS). The lentiviral vectors expressing tetracycline inducible (TetO) *KRAS* G12V or *EGFR* L858R were constructed in the pInducer vector system by the Varmus lab (Unni, et., al 2015). The lentiviral vector with a tetracycline-inducible cassette encoding the human NOTCH1-intracytoplasmic domain (NICD) in TetO-FUW vector system was obtained from Addgene (plasmid# 61540). ShRNA target sequences are as follows: RB #1: 50-GGACATGTGAACTTATATA-30, RB #2: 50-GAACGATTATCCATTCAAA-30, p53: 50-CACCATCCACTACAACTACAT-30.

To generate the lentiviral particles, the above plasmids were transfected into HEK293T cells with the PC-Pack2 lentiviral packaging mix (Cellecta, Inc.), according to the manufacturer's protocol. High titer viral particles were used to transduce hESCs in serum-free conditions and the antibiotic selection of transduced hESCs was performed without MEF feeder cells, using mTeSR1 stem cell media (Stemcell Tech. Inc.) The efficiency of *RB* or *P53* knockdown and production of KRAS G12V, EGFR L858R or NICD were verified by western blotting after antibiotic selection using the following antibodies: anti-RB, Cell Signaling, clone 4H1, Cat# 9309; ant-P107, Santa Cruz, clone C-18, Cat# Sc-318; anti-P-130, Santa Cruz, clone C-20, Cat# Sc-317; anti-P53, Santa Cruz, clone DO-1, Cat# Sc-126; anti-KRAS G12V, Cell Signaling, clone D2H12, Cat# 14412; anti-KRAS G12D, Cell Signaling, clone D8H7, Cat# 14429; anti-KRAS, Cell Signaling, clone D2C1, Cat# 8955; anti-EGFR L858R, Cell Signaling, clone 43B2, Cat# 3197; anti-EGFR, Cell Signaling, clone D38B1, Cat# 4267;or anti-NICD, Cell Signaling, clone D3B8, Cat# 4147.

Immunohistochemistry. Living cells in culture were directly fixed in 4% paraformaldehyde for 25 min, followed with 15 min permeabilization in 1% triton X-100. Histology on tissues from mice was performed on paraffin-embedded or frozen sections from xenografted tumors and corresponding normal tissues as previously described (Chen, et., al 2012). Tissues were either fixed overnight in 10% buffered formalin and transferred to 70% ethanol, followed by paraffin embedding, or snap frozen in O.C.T (Fisher Scientific, Pittsburgh, PA) and fixed in 10% buffered formalin, followed by paraffin embedding. For immunofluorescence, cells or tissue sections were immunostained with antibodies and counterstained with 4,6-diamidino-2-phenylindole (DAPI). Adjacent sections stained with H and E were used for comparison.

The antibodies used for immunostaining or western blot experiments are as follows: anti-c-Kit, Invitrogen, Cat# CD11705; anti- CXCR4, Biolegend, Clone 12G5, Cat# 306506; anti-EpCAM, Invitrogen, Clone G8.8, Cat# 17-5791-80; anti-NKX2.1, SEVEN HILLS, Cat# WRAB-1231; anti-NKX2.1, Invitrogen, Clone 8G7G3/1, Cat# 18-0221; anti-SOX2, Santa Cruz, Clone Y-17, Cat# sc-17320; anti-FOXA2, Santa Cruz, clone M-20, Cat# sc-6554; anti-SP-B, SEVEN HILLS, Cat# WRAB-48604; anti-SP-C, SEVEN HILLS, Cat# WRAB-76694; anti-pro-SP-C, SEVEN HILLS, Cat# WRAB-9337; anti-PDPN, Santa Cruz, Clone FL-162, Cat# sc-134482; anti-CC10, Santa

Cruz, clone T-18, Cat# sc-9772; anti-P63, Biolegend, Cat# 619001; anti-CGRP, Sigma, Clone CD8, Cat# c9487; anti-ASCL1, Sigma, clone 2D9, Cat# SAB1403577; anti-NCAM1, R&D systems, cat# AF2408; anti-CHGA, Sigma, Cat# HPA017369; anti-NOTCH1 or Cleaved NOTCH1, Cell Signaling, clone D3B8, Cat# 4147; anti-NOTCH1 or NICD, , clone 2D9, R&D systems, Cat# AF3647; anti-NOTCH2, Cell Signaling, clone 8a1, Cat# 2420; anti-Ki67, Cell Signaling, clone D2H10, Cat# 9027; anti-human nuclei, EMD Millipore, Cat# MAB1281; anti-OCT4, Abcam, Cat# ab19857; anti-Nanog, Abcam, Cat# ab21624; anti-SSEA4, Abcam, clone MC813, Cat# ab16287; anti-β-Actin, Cell Signaling, clone D6A8, Cat# 8457; anti-GAPDH, Cell Signaling, clone 14C10, Cat# 2118.

Fluorescent activated cell sorting (FACS). FACS with anti-c-Kit, Invitrogen, Cat# CD11705; anti- CXCR4, Biolegend, Clone 12G5, Cat# 306506; anti-EpCAM, Invitrogen, Clone G8.8, Cat# 17-5791-80 was used to detect definitive endoderm cells. Basically, cells were incubated with antibodies for 30 minutes at 4°C, followed with washed and suspended in 0.1% BSA/PBS buffer. PE and APC filters were then used to detect cells double positive for Kit and CXCR4, or EpCAM and CXCR4 by signal intensity gating. FACS with anti-CGRP antibody (Abcam, clone 4902, Cat# ab81887) was used to detect CGRP+ cells. Cells were first incubated with anti-human CGRP antibody for 30 minutes at room temperature followed with incubation of secondary antibody conjugated with R-phycoerythrin (PE) for 30 min at room temperature. Then cells were washed and suspended in 0.1% BSA/PBS buffer. PE filter was then used to separate cells into CGRP + and CGRP - sub-groups by signal intensity gating. Negative controls stained with control IgG instead of primary antibodies were always performed with sample measurements. Flowcytometry machine of BD FACSAria II and software of Flowjo were mainly used to collect and analyze the flowcytometry data.

Xenograft formation. 1-2 x 10^6 undifferentiated hESCs or hESC-derived lung cells (at day 55 with or without 30 days of prior exposure to DAPT alone or to DAPT and doxycycline to reduce P53 or RB RNA or both RNAs) were subcutaneously injected into 6-8 weeks old NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) mice (Jackson Laboratory, Bar Harbor, Maine). Doxycycline was added to mouse food beginning the day after injection; tumor incidence was monitored at least 2-3 times weekly. When mice became moribund or tumor size reached the IACUC allowable burden, they were sacrificed immediately, necropsy performed, and tumors harvested for further histological or molecular study.

Study approval

All embryonic stem cell studies were approved by the Tri-Institutional ESCRO committee (Weill Cornell Medicine, Memorial Sloan Kettering Cancer Center, and Rockefeller University). All animal protocols in this study were approved by the IACUC committee of Weill Cornell Medicine.

Data Availability and Accession

Single cell sequence data have been deposited in NCBI SRA (SRP136659). Source data for all figures are available from the corresponding authors on reasonable request.