

<https://doi.org/10.1038/s42003-021-01842-7>

OPEN

## Cell-autonomous immune gene expression is repressed in pulmonary neuroendocrine cells and small cell lung cancer

Ling Cai<sup>1,2,3✉</sup>, Hongyu Liu<sup>1,4</sup>, Fang Huang<sup>2</sup>, Junya Fujimoto<sup>5</sup>, Luc Girard<sup>3,6,7</sup>, Jun Chen<sup>4,8</sup>, Yongwen Li<sup>4</sup>, Yu-An Zhang<sup>6</sup>, Dhruba Deb<sup>6</sup>, Victor Stastny<sup>6</sup>, Karine Pozo<sup>9</sup>, Christin S. Kuo<sup>10</sup>, Gaoxiang Jia<sup>1</sup>, Chendong Yang<sup>2</sup>, Wei Zou<sup>11</sup>, Adeeb Alomar<sup>6</sup>, Kenneth Huffman<sup>6</sup>, Mahboubeh Papari-Zareei<sup>6</sup>, Lin Yang<sup>12</sup>, Benjamin Drapkin<sup>3,6,9</sup>, Esra A. Akbay<sup>13</sup>, David S. Shames<sup>11</sup>, Ignacio I. Wistuba<sup>5</sup>, Tao Wang<sup>1,3</sup>, Jane E. Johnson<sup>14</sup>, Guanghua Xiao<sup>1,3,15</sup>, Ralph J. DeBerardinis<sup>1,2,3,16</sup>, John D. Minna<sup>1,3,6,7,9,18✉</sup>, Yang Xie<sup>1,3,15,18✉</sup> & Adi F. Gazdar<sup>3,6,13,17,18</sup>

Small cell lung cancer (SCLC) is classified as a high-grade neuroendocrine (NE) tumor, but a subset of SCLC has been termed “variant” due to the loss of NE characteristics. In this study, we computed NE scores for patient-derived SCLC cell lines and xenografts, as well as human tumors. We aligned NE properties with transcription factor-defined molecular subtypes. Then we investigated the different immune phenotypes associated with high and low NE scores. We found repression of immune response genes as a shared feature between classic SCLC and pulmonary neuroendocrine cells of the healthy lung. With loss of NE fate, variant SCLC tumors regain cell-autonomous immune gene expression and exhibit higher tumor-immune interactions. Pan-cancer analysis revealed this NE lineage-specific immune phenotype in other cancers. Additionally, we observed MHC I re-expression in SCLC upon development of chemoresistance. These findings may help guide the design of treatment regimens in SCLC.

<sup>1</sup> Quantitative Biomedical Research Center, Department of Population and Data Sciences, UT Southwestern Medical Center, Dallas, TX, USA. <sup>2</sup> Children's Research Institute, UT Southwestern Medical Center, Dallas, TX, USA. <sup>3</sup> Simmons Comprehensive Cancer Center, UT Southwestern Medical Center, Dallas, TX, USA. <sup>4</sup> Tianjin Key Laboratory of Lung Cancer Metastasis and Tumor Microenvironment, Tianjin Lung Cancer Institute, Tianjin Medical University General Hospital, Tianjin, China. <sup>5</sup> Department of Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>6</sup> Hamon Center for Therapeutic Oncology Research, UT Southwestern Medical Center, Dallas, TX, USA. <sup>7</sup> Department of Pharmacology, UT Southwestern Medical Center, Dallas, TX, USA. <sup>8</sup> Department of Lung Cancer Surgery, Tianjin Lung Cancer Institute, Tianjin Medical University General Hospital, Tianjin, China. <sup>9</sup> Department of Internal Medicine, UT Southwestern Medical Center, Dallas, TX, USA. <sup>10</sup> Department of Pediatrics, Stanford University, Stanford, CA, USA. <sup>11</sup> Department of Oncology Biomarker Development, Genentech Inc., South San Francisco, CA, USA. <sup>12</sup> Department of Pathology, National Center/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. <sup>13</sup> Department of Pathology, UT Southwestern Medical Center, Dallas, TX, USA. <sup>14</sup> Department of Neuroscience, UT Southwestern Medical Center, Dallas, TX, USA. <sup>15</sup> Department of Bioinformatics, UT Southwestern Medical Center, Dallas, TX, USA. <sup>16</sup> Howard Hughes Medical Institute, UT Southwestern Medical Center, Dallas, TX, USA. <sup>17</sup> Deceased: Adi F. Gazdar. <sup>18</sup>These authors jointly supervised this work: John D. Minna, Yang Xie, Adi F. Gazdar. ✉email: [Ling.Cai@UTSouthwestern.edu](mailto:Ling.Cai@UTSouthwestern.edu); [John.Minna@UTSouthwestern.edu](mailto:John.Minna@UTSouthwestern.edu); [Yang.Xie@UTSouthwestern.edu](mailto:Yang.Xie@UTSouthwestern.edu)

**S**mall cell lung cancer (SCLC), accounting for 15% of lung cancer cases, with a 5-year survival of 6%, is designated by the US Congress as a “recalcitrant cancer”<sup>1,2</sup>. SCLC is classified as a high-grade neuroendocrine (NE) tumor<sup>3</sup>. A large fraction of SCLC tumors are driven by *ASCL1*, a lineage oncogene also important for pulmonary neuroendocrine cell (PNEC) fate determination<sup>4,5</sup>. In healthy lung tissue, PNECs are rare and dormant<sup>6</sup>, whereas upon lung injury, some act as stem cells to regenerate surrounding epithelial cells<sup>7</sup>. SCLC occurs primarily in heavy smokers, but despite the very high mutation burden<sup>8–10</sup> from SCLC genomes predicted to contribute an ample supply of neoantigens, SCLCs express low levels of major histocompatibility complex class I (MHC I) proteins to present tumor-specific antigens<sup>11,12</sup>. This could explain why, among various types of cancer, checkpoint-blockade immunotherapy underperforms in SCLC<sup>13,14</sup>.

Thirty-five years ago, it was observed that by contrast to classic SCLC cell lines (which grew in tissue culture as floating cell aggregates), a subset of patient-derived SCLC lines behaved differently—growing as adherent monolayers in culture, with morphologically larger cells, more prominent nucleoli, and expressed few or no NE markers<sup>15,16</sup>. These characteristics led such tumors to be termed “variant” or “non-NE” SCLC. Many of these variant SCLC lines were established from patients whose tumors had acquired resistance to chemotherapy and clinically relapsed, a context in which genomic *MYC* amplification was also noted to be more frequent<sup>17</sup>. Notch activation had been shown to mediate the transition from classic to variant subtypes and accounts for the intratumoral heterogeneity commonly seen in SCLC<sup>18</sup>.

Recently, extending the concepts of classic and variant SCLC, both intertumoral, and intratumoral heterogeneity in SCLC has been documented and has been associated with the expression of lineage-specific transcription factors (TFs) *ASCL1*, *NEUROD1*, *YAP1*, and *POU2F3*, and these various subtypes express different levels of NE markers<sup>19–21</sup>.

We have previously defined a 50-gene NE signature that helps us quantify the NE properties as a continuous NE score ranging from -1 to 1, with a more positive score indicating higher NE properties<sup>22</sup>. In the current study, we applied this NE scoring method to SCLC samples from preclinical models and patient tumors. We first assessed the relationship between NE scores and SCLC molecular subtypes. Then, we investigated the immune phenotypes associated with variable NE scores in SCLC and other cancer types.

## Results

**Relationship between NE scores and SCLC molecular subtypes.** Using the 50-gene NE signature updated with all available SCLC-related RNA-seq data (Supplementary Data 1), we computed NE scores for patient-derived SCLC lines and xenografts (PDXs) as well as four independent patient tumor data sets (including one newly generated for this study) (Table 1 and Supplementary Data 2). We examined the relationship between NE scores and expression of SCLC molecular subtype-specific TFs as proposed by Rudin et al. (Fig. 1a, b). Our findings are largely consistent with the previous proposal that assigns *ASCL1*<sup>+</sup> and *NEUROD1*<sup>+</sup> SCLCs to NE subtypes and *POU2F3*<sup>+</sup> and *YAP1*<sup>+</sup> SCLCs to non-NE subtypes. However, we note some discrepancies. First, we found that while expression of *ASCL1* and *NEUROD1* seems to be mutually exclusive in cell lines, they seem to co-express in many of the tumor samples; a small set of samples with low NE scores still express *ASCL1* or *NEUROD1*; in “George\_2015”, “Jiang\_2016” and our own data set, we have observed rare *POU2F3*<sup>+</sup> samples that have high NE scores.

With serially sectioned formalin-fixed paraffin-embedded (FFPE) slides from 9 out of the 18 tumors for which we had performed expression profiling, we examined the tumors with hematoxylin and eosin (H&E) staining as well as immunohistochemistry (IHC) staining of *ASCL1*, *NEUROD1*, and *POU2F3* (Fig. 1c–f). The high NE-score tumors exhibited predominantly classic SCLC morphology with dark nuclei, scant cytoplasm, and inconspicuous nucleoli. Notably, this was not only seen in *ASCL1*<sup>+</sup> tumors (for example, SCLC-04, NE score 0.4) but also in the *POU2F3*<sup>+</sup> tumor with a positive NE-score (SCLC-15, NE score 0.26) (Fig. 1c). On the other hand, while we observed variant morphology in tumors with low NE scores, we noticed intratumoral heterogeneity. In a tumor weakly positive for *ASCL1* (SCLC-20, NE score -0.05), the *ASCL1*-high regions were found to be more classic-like whereas the *ASCL1*-low regions were more variant-like (Fig. 1d). Our IHC-based quantifications largely agree with the microarray gene expression assessments (Fig. 1e). Tumors that were found to express both *ASCL1* and *NEUROD1* stained positive for both markers as well. In addition, intratumoral heterogeneity was commonly found within such tumors, which contain areas with high expression of both TFs but also areas with expression of only one TF (Fig. 1f).

**Immune gene repression is a NE lineage-specific property.** We performed a correlation between NE scores and SCLC transcriptomic data to identify gene expression changes associated with the NE program. Not surprisingly, gene ontology (GO) analyses revealed genes related to the neuronal system as highly expressed in high NE-score samples (Supplementary Fig. 1a, b). By contrast, genes negatively associated with NE score were enriched for GO terms related to immune response, and this was observed in both the cell line and human tumor data sets (Supplementary Fig. 1c, d). We also performed gene set enrichment analysis (GSEA)<sup>23</sup> with a variety of gene set libraries collected by Enrichr<sup>24</sup>. Consistent with the previous report that Notch signaling dependent REST (Neuron-Restrictive Silencer Factor) activation represses neuronal gene expression in variant SCLC<sup>18</sup>, we found REST targets (i.e., repressed by REST) are abundantly expressed in high NE-score SCLCs. On the other hand, interferon-stimulated genes (ISGs) are found to highly expressed in the low NE-score (variant) SCLC samples (Fig. 2a, b and Supplementary Data 3). As NF $\kappa$ B signaling mediates activation of ISGs, we examined reverse-phase protein array (RPPA) data from the cancer cell line encyclopedia (CCLE)<sup>25</sup> and found higher levels of activating serine 536 phosphorylation on p65<sup>26</sup> in low NE-score SCLC lines (Fig. 2c). This result aligns with the previous findings that innate immune genes are more highly expressed in an SCLC cell line H69 variant cell line with upregulation of mesenchymal genes—it was found that in this cell line, TBK1, IRF3, and STAT1 signaling is activated by a subclass of endogenous retroviruses (ERVs) through MAVS and STING to form a positive feedback loop and sustain innate immune gene expression<sup>27</sup>. As classic SCLC is driven by lineage factor *ASCL1* whereas variant SCLC with loss of NE genes is driven by *YAP1*, we examine the functional roles of *YAP1* and *ASCL1* in regulating the ISGs in SCLC cell lines. Horie et al. previously examined the consequences of silencing *YAP1* in *YAP1*<sup>+</sup> SCLC cell lines<sup>28</sup>. From their results, we observed a significant decrease of ISG expression in SCLC cell line SBC5 without upregulation of NE genes targeted by REST. *ASCL1* silencing from an *ASCL1*<sup>+</sup> SCLC cell line H2107 on the other hand<sup>29</sup>, did not increase the expression of ISGs (Supplementary Fig. 2).

Our 50-gene signature derived from lung cancer cell line mRNA data (Supplementary Data 1) contains several genes with immune-related functions that were found to highly express in

**Table 1** Data sets used for analyses.

Source	Name	Tissue source	Sample type	n	References
Human	SCLC cell lines/ NCI/Hamon Center	SCLC	Cell line	69	This study
Human	Horie_2016	SCLC	Cell line	4	28
Human	Pozo_2020	SCLC	Cell line	4	29
Human	Cañadas_2014	SCLC	Cell line	6	56
Human	Drapkin_2018	SCLC	PDX	19	57
Human	Rudin_2012	SCLC	Tumor	29	9
Human	George_2015	SCLC	Tumor	81	10
Human	Jiang_2016	SCLC	Tumor	79	80
Human	SCLC tumors (this study)	SCLC	Tumor	18	This study
Human	expO	Lung cancer	Tumor	109	82
Human	Rousseaux_2013	Lung cancer	Tumor	286	81,82
Human	CCLE	Pan-cancer	Cell line	_b	25
Human	TCGA	Pan-cancer	Tumor	10535	53
Human	TARGET	Pan-cancer	Tumor	734	53
Mouse	Lim_2017	SCLC	Pooled FACS-sorted tumor cells <sup>a</sup>	6	18
Human	Travagliini_2020	Healthy lung	Single cell	9384 <sup>c</sup>	42
Mouse	Ouadah_2019	Healthy lung	Single cell	46 <sup>c</sup>	7

PDX patient-derived xenografts; FACS fluorescence-activated cell sorting; scRNA-seq single-cell RNA sequencing.

<sup>a</sup>In Lim\_2017, Rb1<sup>flox/flox</sup>;p53<sup>flox/flox</sup>;p130<sup>flox/flox</sup>;Rosa26<sup>mTmG</sup>; Hes1<sup>GFP/+</sup> GEMM SCLC tumors were initiated by Ad-CMV-Cre, sorted by Tomato and GFP to obtain relatively pure tumor cells.

<sup>b</sup>CCLE data sets were used in multiple analyses with different numbers of cell lines; <sup>c</sup>Cells.

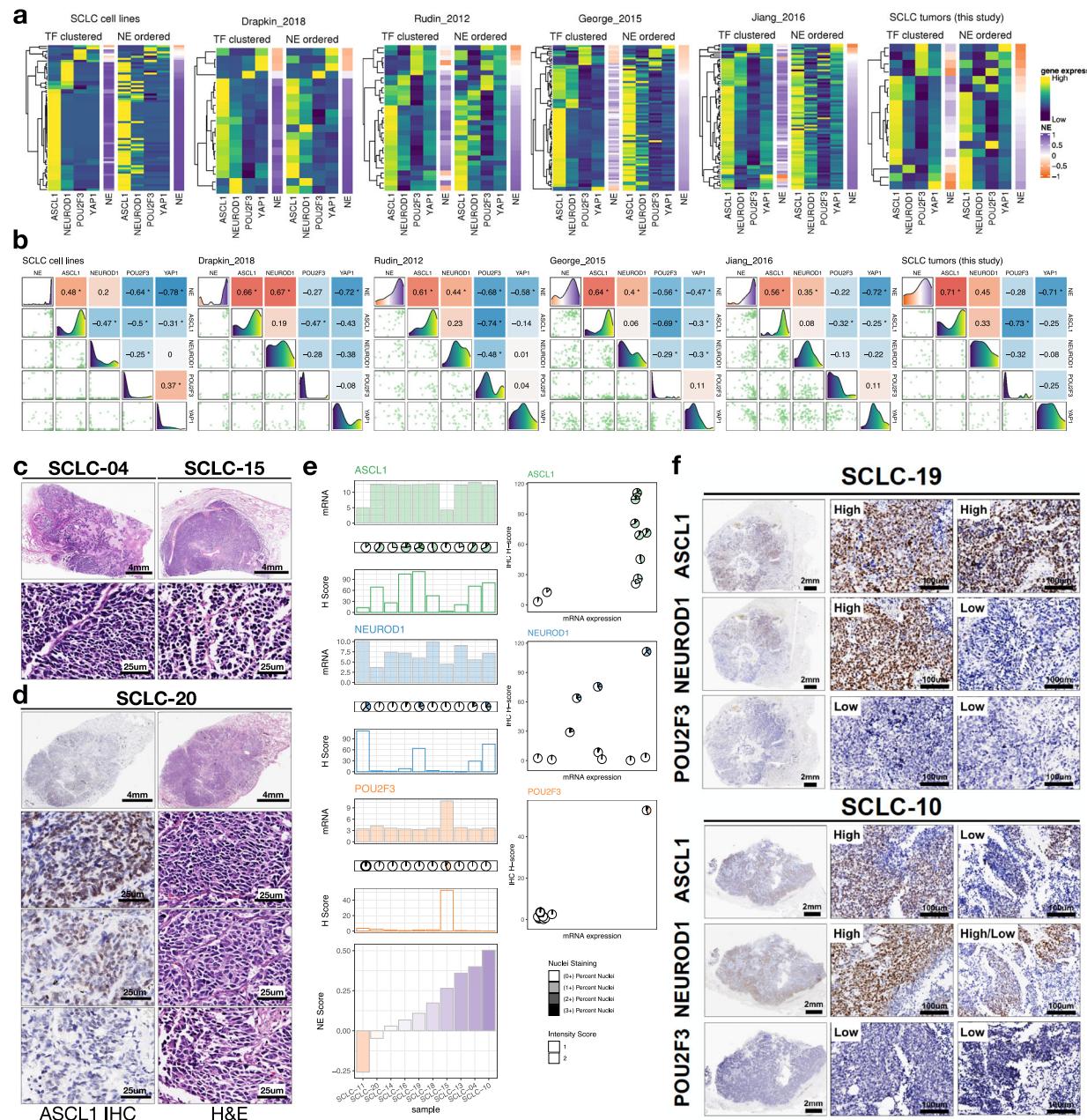
variant SCLC. Some are involved in cytokine signaling; for example, *IL18* encodes for a proinflammatory cytokine<sup>30</sup>, and *OSMR* encodes for a receptor for oncostatin M and IL-31<sup>31</sup>. Furthermore, many of these genes are involved with immunosuppressive processes, including *NT5E*<sup>32</sup>, *TGFBR2*<sup>33</sup>, *ANXA1*<sup>34</sup>, *EPHA2*<sup>35</sup>, *HFE*<sup>36</sup>, and *LGALS3*<sup>37</sup>. From our pathway analysis results, NE score negatively correlated genes are also enriched for genes upregulated in response to Tumor Necrosis Factor-alpha (TNFA) or interleukin-1 (IL1) (Supplementary Data 3). indicates that beyond these genes included in the NE expression signature, there is a broad immune program concordantly upregulated in low NE-score SCLC samples. We extended our analysis to a few immune gene sets that were previously identified to express cell-autonomously in cancer. These gene sets include the following: SPARCS genes (stimulated 3 prime antisense retroviral coding sequences) reported to express in mesenchymal tumors and mediate interferon-gamma signal amplification<sup>27</sup>; “parainflammation” genes in epithelial tumor cells<sup>38</sup>; and senescence-associated secretory phenotype (SASP) genes<sup>39</sup> that reinforce the senescence arrest, alter the microenvironment, and trigger immune surveillance of the senescent cells<sup>40</sup>. We observed that the expression of these genes also negatively correlate with NE scores in SCLC despite little overlap among genes in these various sets (Supplementary Fig. 2a, b). We also assessed about 1,000 innate immune genes cataloged by the InnateDB<sup>41</sup> and found a significantly higher proportion of these genes negatively associated with NE score from SCLC cell lines (Supplementary Fig. 3c and Supplementary Data 4).

While the expression of neuronal program genes in high NE-score SCLCs can be attributed to the NE lineage, we examined single-cell RNA-seq (scRNA-seq) data from the healthy human lung epithelial cells<sup>42</sup> to check whether the expression repression of ISGs is also a lineage-specific phenomenon that could be observed in PNECs rather than being cancer-specific. Consistent with the previous report that ASCL1 negatively regulates YAP1 during neuronal differentiation<sup>43</sup>, the highest expression of ASCL1 and lowest expression of YAP1 was observed in PNECs, relative to other cell types. We confirmed that while PNECs have increased expression of REST target genes, ISGs are indeed repressed as well (Fig. 2d and Supplementary Fig. 4a). Additionally, we specifically

examined interferon receptors in PNECs and found that they also have the lowest expression in PNECs (Fig. 2d). It has been estimated that 10% of the genes in the human genome have the potential to be regulated by IFN, many ISGs work in immune defense against viral infection, but some could be hijacked by viruses<sup>44</sup>. As some PNECs are rare stem cells, we reason that ISG repression might lower their risk from viral infection. In the context of the current COVID-19 pandemic, we examined scRNA-seq data from Ouadah et al., who performed lineage tracing with an *Ascl1*<sup>CreERT2</sup>; *Rosa26*<sup>LSL-ZsGreen</sup> mouse model to show that some PNECs can transdifferentiate into other cell types<sup>7</sup>. Supplementary Fig. 4b generated with their data shows that AT2 and ciliated cells originated from PNECs in this model have lost *Ascl1* but increased *Yap1* expression. *Ly6e* and *Tmprss2*, genes involved in coronavirus defense<sup>45</sup> and hijacked entry<sup>46</sup>, respectively, were also upregulated.

**Increased tumor-immune interaction in low NE-score SCLC tumor samples.** It has been long observed that the expression of MHC I is low in SCLC<sup>11</sup>. Using single sample gene set enrichment analysis (ssGSEA)<sup>47</sup>, we derived the MHC I scores for MHC I genes. From studies that had collected lung tumors of different histology, MHC I scores positively correlate with *PTPRC* (which encodes for pan-leukocyte marker CD45) levels (Fig. 3a and Supplementary Fig. 5a). The lowest MHC I and *PTPRC* gene expression were found in neuroendocrine tumors, including not only SCLC but also carcinoids (Fig. 3a), suggesting these NE tumors with decreased MHC I have fewer immune infiltrates. In SCLC data sets, low NE-score samples exhibited upregulation of MHC I genes (Fig. 3b) and were associated with higher *PTPRC* expression in patient tumor data sets (Supplementary Fig. 5b). We also estimated immune cell infiltration by deriving immune cell type-specific signature scores<sup>48</sup> and found that they negatively correlate with NE scores in SCLC patient tumors, suggesting increased tumor-immune interaction in low NE-score tumors (Fig. 3c).

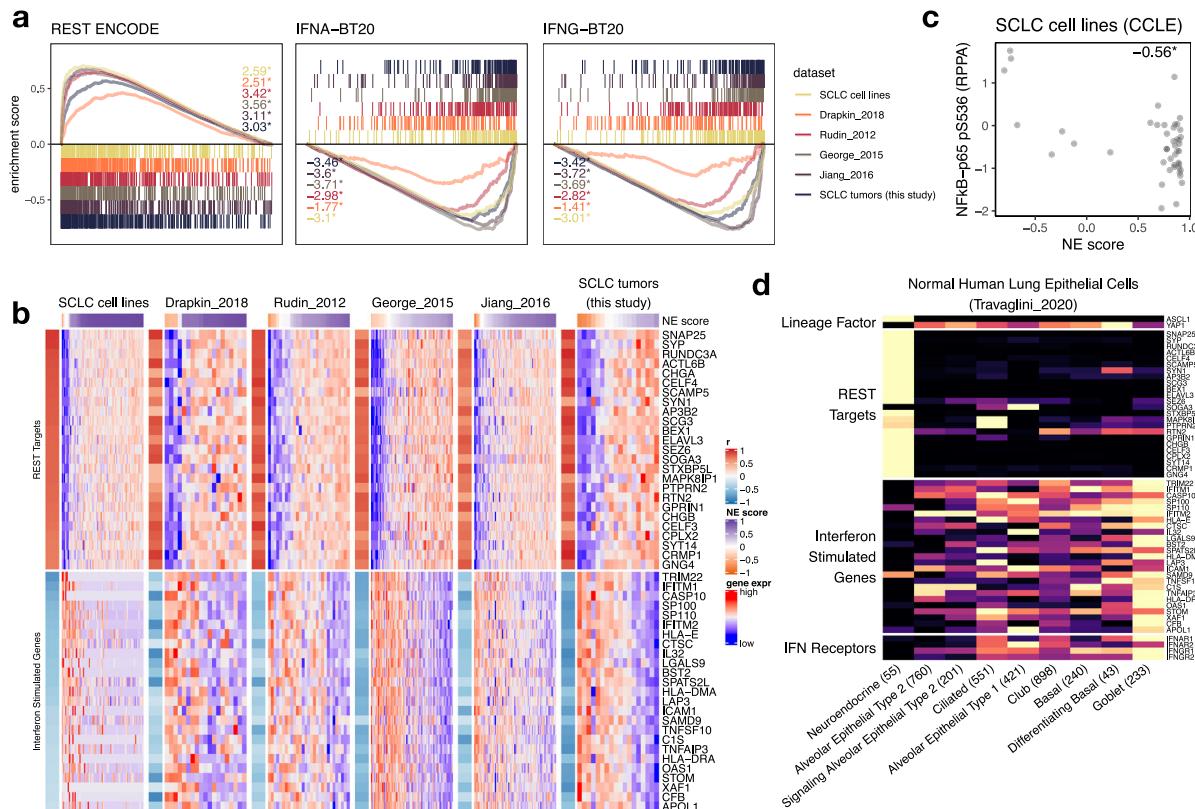
We saw a higher expression of PD-L1 (CD274) in low NE-score SCLCs for 3 out of 4 primary tumor data sets (Supplementary Fig. 6a). Furthermore, genes from an IFN-gamma related signature that has been shown to predict PD-1



**Fig. 1 NE score and SCLC molecular subtypes.** **a** Heatmaps visualizing relative expression of molecular subtype-specific TFs and NE scores. Two heatmaps were generated for each study, with one ordered by complete linkage hierarchical clustering of TFs and the other ordered by NE scores. Gene expression was z-score standardized for each sample. **b** Pairs plot for NE scores and molecular subtype-specific TFs. Lower left panels are scatter plots showing a pairwise relationship between variables, diagonal panels are density plots showing the distribution of each variable and upper right panels are correlation coefficients from pairwise Pearson correlation. Refer to the top or right of each  $5 \times 5$  matrix for subplot x or y axis variable labels. For example, for each data set, the first subplot on the top row shows the distribution of NE scores from that data set, the scatterplot below it shows the relationship between NE score (x axis value) and ASCL1 expression (y axis value), and the Pearson correlation coefficient between NE score and ASCL1 expression is provided in the second cell of the top row. \* $p$ -value  $< 0.05$ . Note that in all studies: most samples have positive NE score; under TF-based classification, ASCL1<sup>+</sup> subtype dominates; NE scores positively correlate with ASCL1 and NEUROD1 but negatively correlate with POU2F3 and YAP1 expression. **c** H&E staining of two high NE-score SCLC tumor samples showing classic SCLC morphology with dark nuclei, scant cytoplasm, and inconspicuous nucleoli. **d** ASCL1 IHC staining and H&E staining of a low NE-score SCLC tumor, showing variable morphology at different selected areas, where ASCL1-low areas appear to be more variant-like. **e** Quantifications of TF expression from IHC staining or microarray profiling, samples are ordered by increasing NE scores. **f** IHC of ASCL1, NEUROD1, and POU2F3 in two tumors that express both ASCL1 and NEUROD1. Two areas per tumor were selected for showing intratumoral heterogeneity in ASCL1 and NEUROD1 expression patterns.

blockade response in multiple cancer types<sup>49</sup> are highly expressed in low NE-score SCLC tumors across multiple data sets (Fig. 3c). We also examined a list of 21 immune checkpoint genes<sup>50</sup>, immune-suppressive cytokines (IL-10 and TGF-beta), and their

receptors<sup>51</sup>, for their association with NE scores. We found that these genes also have higher expression in low NE-score SCLC tumors (Fig. 3c). Finally, the expression of 995 immunosuppressive genes from the Human Immunosuppression Gene Atlas<sup>50</sup>



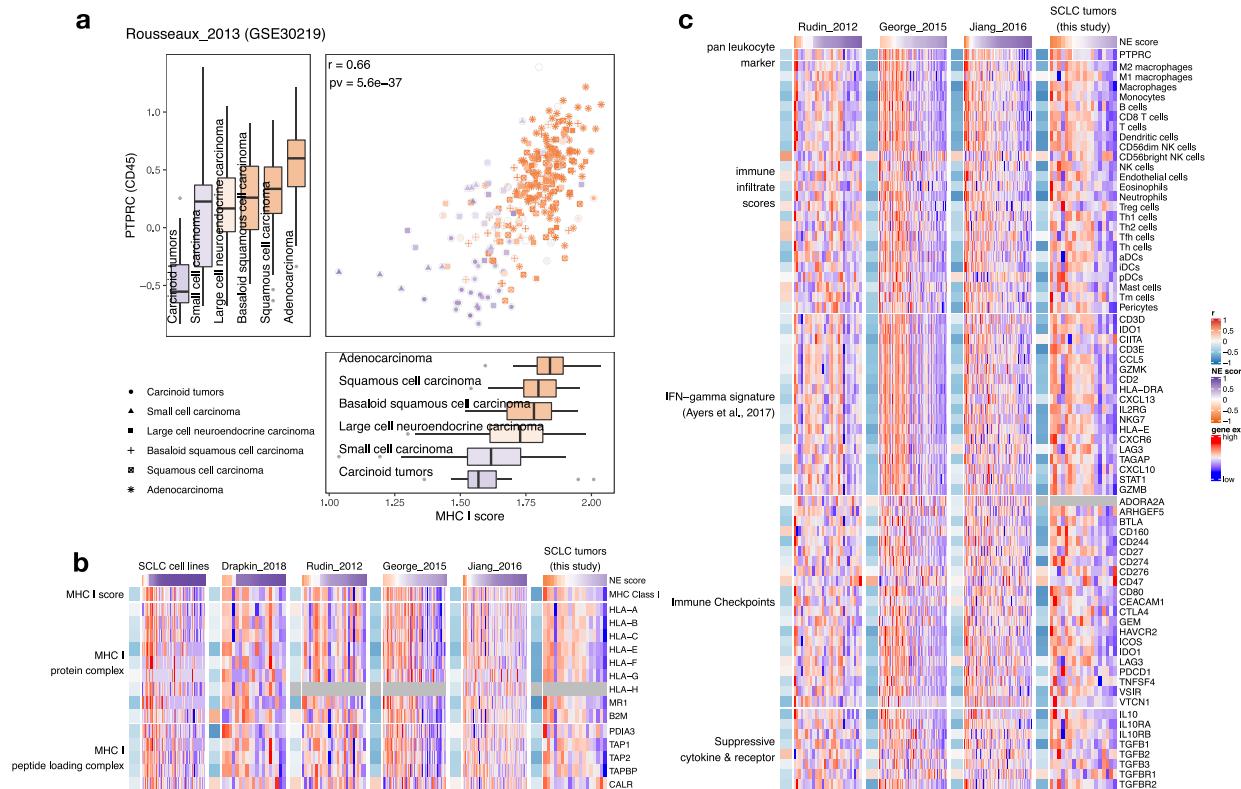
**Fig. 2** Repression of ISGs in high NE-score SCLC and PNECs. **a** GSEA enrichment plots for selected gene sets. Results from SCLC cell lines, PDxs (Drapkin\_2018), and patient tumor data sets were superimposed. Normalized enrichment score (NES) was provided. \* multiple comparison-adjusted  $p$ -value  $< 0.05$ . **b** Heatmaps for top 25 leading-edge genes selected from gene sets in **a**. Gene expression matrix of each data set was annotated with color-coded Pearson correlation coefficient (from correlating NE score with gene expression) as a left-side column, and a top bar indicating NE scores. For each data set, gene expression was z-score standardized across samples. **c** Scatter plot showing negative correlation between NE score and Ser536 phosphorylation on NFκB-p65 in SCLC cell lines. Pearson correlation coefficient was provided. \* $p$ -value  $< 0.05$ . **d** Heatmap showing relative expression of selected lineage factors (ASCL1 and YAP1), REST targets, and ISGs (same genes as used in **b**, determined from **a**) as well as interferon receptor genes in healthy human lung epithelial cells based on scRNA-seq experiments. Gene expression from all cells in each cell-type cluster was averaged and then min-max scaled across cell types to a scale between 0 and 1.

was assessed, and again, the majority of these genes exhibit a negative correlation between mRNA expression and NE scores across different SCLC tumor data sets (Supplementary Fig. 6b and Supplementary Data 5).

Besides gene expression-based analyses, we also performed immunohistochemistry (IHC) with our 9 SCLC tumor samples to quantify tumor-infiltrating CD8 $^{+}$  and CD4 $^{+}$  T cells (Supplementary Data 2). Of importance, both intratumoral and intertumoral heterogeneity were observed in T cell infiltration. Within the same tumor, areas with low tumor ASCL1 levels exhibited more CD8 $^{+}$  and CD4 $^{+}$  T cell infiltration, whereas areas with high tumor ASCL1 levels showed fewer CD8 $^{+}$  or CD4 $^{+}$  T cells (Fig. 4a). Across all the SCLC tumor specimens assessed, CD8 $^{+}$  and CD4 $^{+}$  T cell per area cell count positively correlated with the T cell score computed from gene expression data, and both IHC-based T cell counts and gene expression-based T cell scores negatively correlated with NE scores (Fig. 4b).

**Pan-cancer analyses for NE score expression and immune response genes.** These findings had prompted us to examine other cancer types to see whether immune gene repression is seen in other NE tumors and whether variant subtype from NE lineage loss could also be observed (Fig. 4c). A recent study identified SCLC-like epithelial tumors in pan-cancer samples using a principal component analysis-based approach. They found that tumors across many lineages with a higher SCLC-like score had lower immune gene

expression<sup>52</sup>. We applied our NE scoring method across all cancer lineages (pan-cancer analysis) to compute NE scores and assess their relationship with immune phenotypes. In pediatric (Therapeutically Applicable Research to Generate Effective Treatments—TARGET) and adult (The Cancer Genome Atlas—TCGA) pan-cancer studies<sup>53</sup>, neuroendocrine tumor neuroblastoma (NBL), as well as pheochromocytoma & paraganglioma (PCPG) were identified as containing the highest NE scores (Fig. 5a, b). Tumors of glial origin, including Low-Grade Glioma (LGG) and Glioblastoma Multiforme (GBM) also have high NE scores. Besides these NE/glial tumors, a small number of high NE-score samples were observed for bladder urothelial carcinoma (BLCA), breast invasive carcinoma (BRCA), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), pancreatic adenocarcinoma (PDAC) and stomach adenocarcinoma (STAD), for which it is also known that neuroendocrine tumors, while uncommon, still comprise a small proportion of the cases (Fig. 5a, b). In a previous immunogenomic analysis that had classified pan-cancer TCGA samples into six immune subtypes<sup>54</sup>, we found samples from the “immunologically quiet” subclass have the highest NE score, followed by the “lymphocyte depleted” subclass (Fig. 5c). We further assessed the relationship between NE scores and the tumor-infiltrating lymphocytes and leukocyte regional fractions previously reported for the pan-cancer samples<sup>54</sup>, these immune metrics negatively correlate with NE scores across all samples (Fig. 5d) and also within specific tumor types (Supplementary Fig. 7).

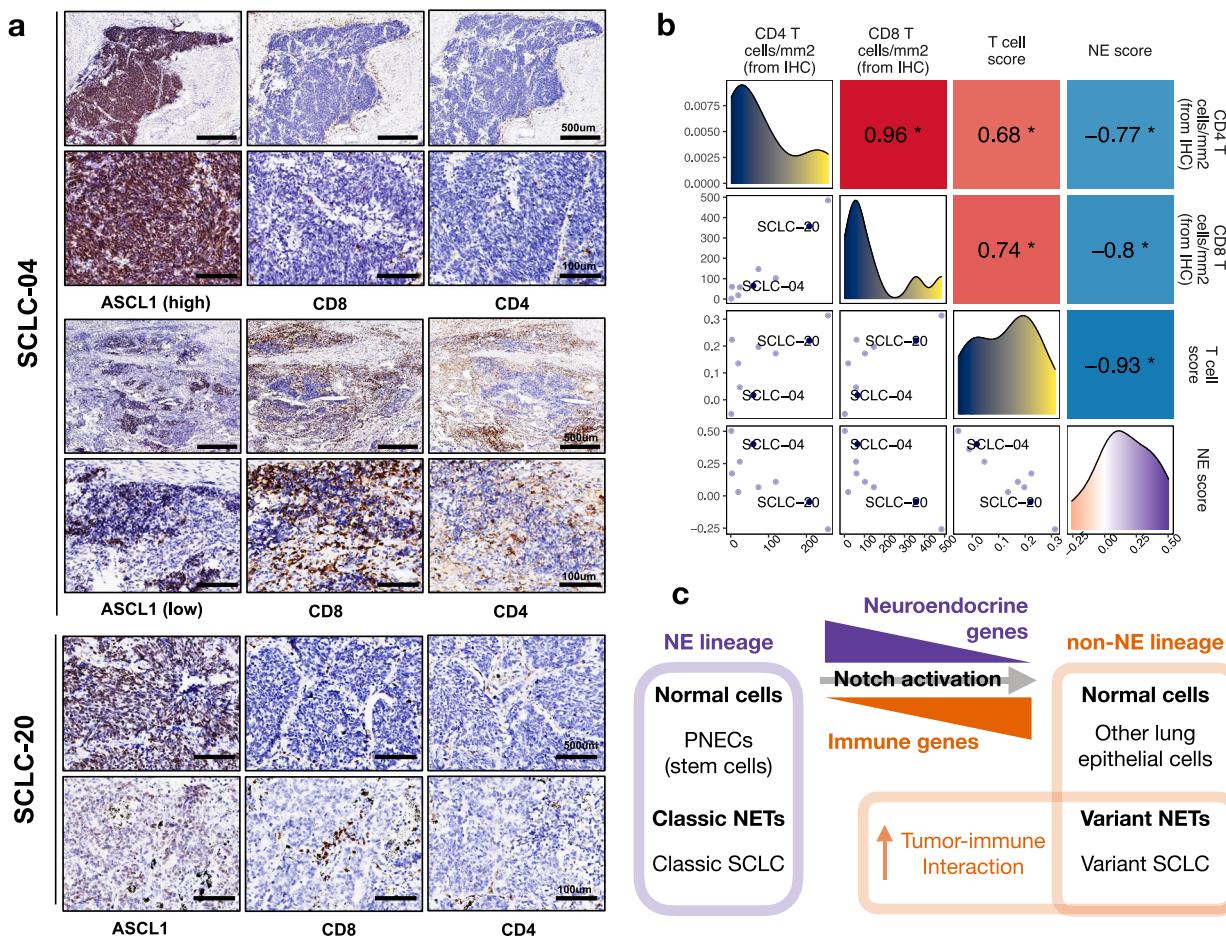


**Fig. 3 Low NE-score variant SCLC has increased tumor-immune interaction.** **a** Expression of MHC I genes and pan-leukocyte marker *PTPRC* in lung tumors from the “Rousseaux\_2013” data set. Box whisker plots are filled with color reflecting the median NE score in different histological subtypes. The centerline in the boxplot represents the median, the lower and upper hinges correspond to the first and third quartiles, the whiskers extend from the hinge to the value no further than  $1.5 \times \text{IQR}$  away. Color for scatterplot symbols reflects NE score for different samples. **b** Heatmaps visualizing expression of MHC I genes across multiple SCLC data sets. **c** Heatmaps visualizing expression of *PTPRC*, immune-cell-type-specific signature scores, PD-1 blockade response-predicting IFN-gamma-related signature genes<sup>49</sup>, immune checkpoint genes, and suppressive cytokines and receptors in SCLC tumor data sets. MHC I score and immune infiltrate scores were computed using ssGSEA methodology. Sample-wise z-score standardized values were used for the heatmaps.

We took a close examination of NBL using cell line expression data from CCLE<sup>25</sup> along with patient tumor data from TARGET<sup>55</sup> for lineage factors *ASCL1* and *YAP1*, REST targets, ISGs, MHC I, immune cell-specific signature scores<sup>48</sup>, Ayer et al.’s PD-1 blockade response signature<sup>49</sup>, immune checkpoints<sup>50</sup>, and suppressive cytokines and receptors<sup>51</sup>. The pattern for NBL (Fig. 5e) highly resembles that of SCLC (Figs. 2b and 3b, c) suggesting the existence of a variant NBL subset with decreased neuroendocrine features, increased cell-autonomous expression of immune genes as well as increased tumor–immune interaction. Like SCLC, we also found higher levels of NFkB-p65 phosphorylation in the low NE-score variant NBL cell lines (Fig. 5f).

**MHC I re-expression in chemoresistant SCLC.** As it was previously observed that variant SCLC cell lines were frequently derived from patients whose tumors had relapsed on chemotherapy<sup>17</sup>, we wondered if the development of chemoresistance in tumors was associated with the altered expression of immune genes, especially MHC I. Five sets of data with origin-matched chemosensitive and chemoresistant samples were examined to address this question. In 2017, using a genetically engineered mouse model (GEMM), Lim et al. showed that Notch-active SCLC cells were more chemoresistant<sup>18</sup>. Using their data we found the Notch-active SCLC cells had switched from *ASCL1*<sup>+</sup> to *YAP1*<sup>+</sup>, have reduced NE scores, and increased expression of ISGs and MHC I genes (Fig. 6a). We next examined a series of preclinical models we and others have developed for human SCLC. Classic,

high NE-score SCLC cell lines predominantly grow as floating aggregates in culture, but contain a small proportion of cells growing adherently in a monolayer. By selecting for adherent growth, we generated an adherent subline H69-AD/(NCI-H69-AD) from parental, chemosensitive H69/(NCI-H69) cells (Fig. 6b). Increased resistance to Cisplatin (~10 fold) and Etoposide (~6 fold) was observed in H69-AD compared to the parental H69 cells (Fig. 6c). We found H69-AD had transitioned to become a low NE-score (-0.02) *YAP1*<sup>+</sup> variant line compared to the parental high NE-score (0.91) classic *ASCL1*<sup>+</sup> line. Both ISGs and MHC I genes were found to have increased expression in H69-AD (Fig. 6d). In a previous study, Cañadas et al. also derived sublines from H69. Hepatocyte growth factor treatment was used to induce the mesenchymal transition of H69 cells, resulting in H69-M lines that were found to be chemoresistant both *in vitro* and *in vivo*<sup>56</sup>. From their data set, we also found MHC I expression increased in H69-M compared to parental H69 cells. There was some increase in ISGs too, but less prominent compared to H69-AD from us. Notably, although *YAP1* expression increased in H69-M, *ASCL1* levels did not change (Fig. 6d). As the fourth set of data, PDX models established sequentially from SCLC tumors (Drapkin\_2018) collected before and after chemotherapy from the same patient<sup>57</sup> were examined. In PDXs from patient MGH1518 for which chemoresistance had developed in the relapsed sample, we found upregulation of MHC I, but not ISGs (Fig. 6d). Of note, this relapsed sample maintained a high NE score but expressed higher levels of *MYC*, consistent with previous findings that *MYC*



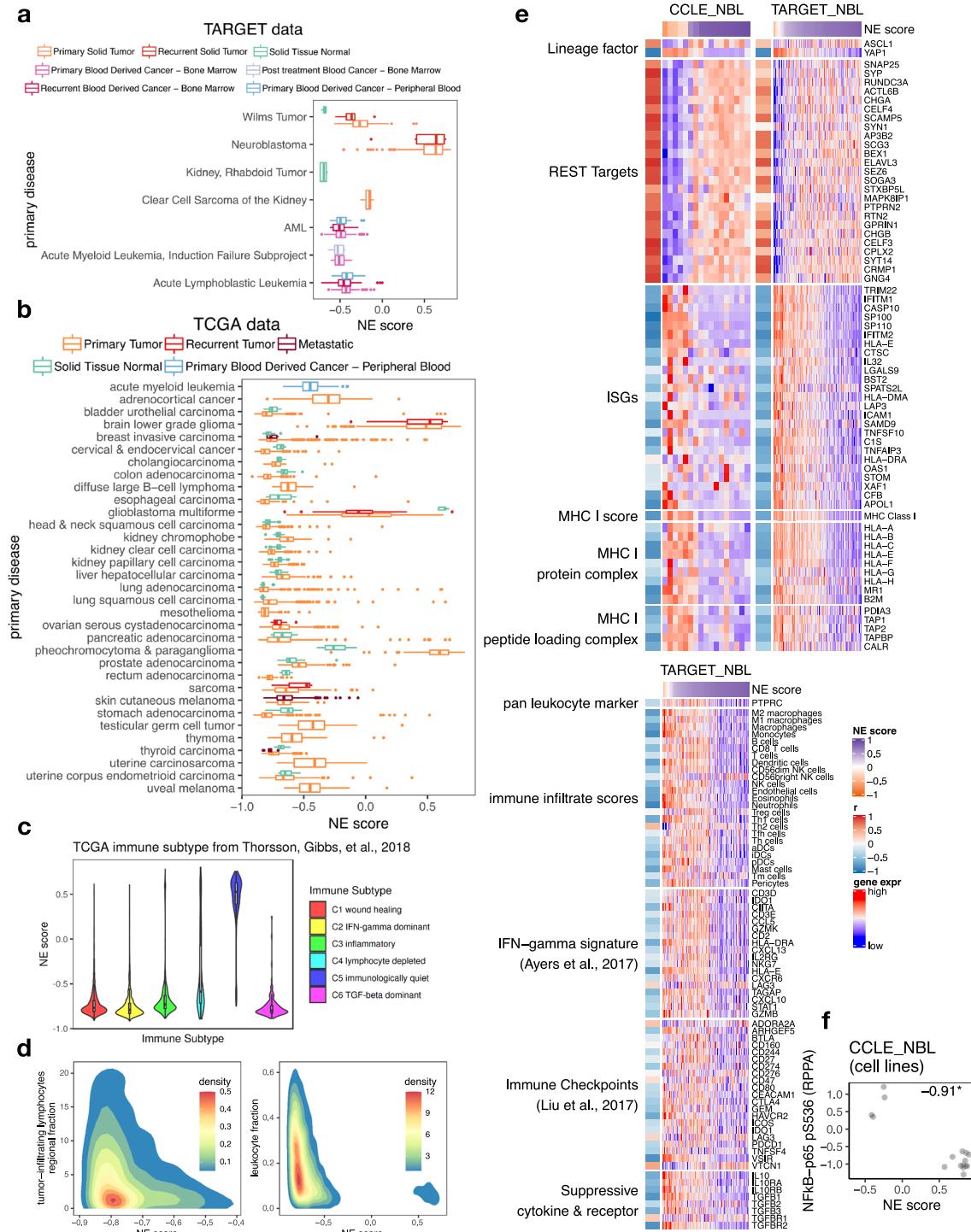
**Fig. 4 Intertumoral and intratumoral heterogeneity in T cell infiltration from SCLC tumors with variable NE features.** **a** IHC of ASCL1, CD4, and CD8 in selected tumors. SCLC-04 is an SCLC tumor with an NE score of 0.4. CD8 or CD4 T cells were few in the ASCL1-high regions but abundant in the ASCL1-low regions; SCLC-20 is a tumor with NE score of -0.05, similar reciprocal relationship of ASCL1 staining and T cell infiltration was observed. Representative regions with high or low ASCL1 staining were shown. **b** Relationship between IHC-determined per area CD4 and CD8 T cell count, gene expression-based T cell score, and NE score in all nine tumors assessed. Lower left panels are scatter plots showing a pairwise relationship between variables, diagonal panels are density plots showing the distribution of each variable and upper right panels are correlation coefficients from pairwise Pearson correlation. \*p-value < 0.05. T cell score was calculated by the ssgSEA method. **c** Schematic diagram showing a relationship between neuroendocrine and immune gene expression in normal cells and neuroendocrine tumors (NETs).

mediates chemoresistance<sup>57</sup>. Lastly, we generated a set of subcutaneous xenograft models from a high NE-score human SCLC cell line NCI-H1436 with or without selection for resistance to Cisplatin and Etoposide in mice (Fig. 6e). Compared to the parental xenograft, the drug-resistant xenografts maintained ASCL1 expression but exhibited increased *B2M* (MHC I complex subunit), *PSMB8* (immunoproteasome subunit) (Fig. 6e), and *MYC*<sup>58</sup>. Collectively, these findings suggest MHC I can re-express upon the development of chemoresistance—in some cases, with lineage transition; and in some other cases, accompanied by an increase in MYC expression.

We also checked whether expression levels of MHC I and *MYC* differ by tumor source and anatomical site based on the “NCI/Hamon Center” patient-derived SCLC lines data set (Supplementary Fig. 8). Interestingly, the lowest MHC I and *MYC* levels were both observed for cell lines derived from primary site lung tumor specimens and they are all high-NE tumors, whereas higher MHC I levels were observed in SCLC lines isolated from metastatic tumor samples especially those from lymph node and bone marrow. These observations remain to be validated with primary and metastatic samples from the same patients.

## Discussion

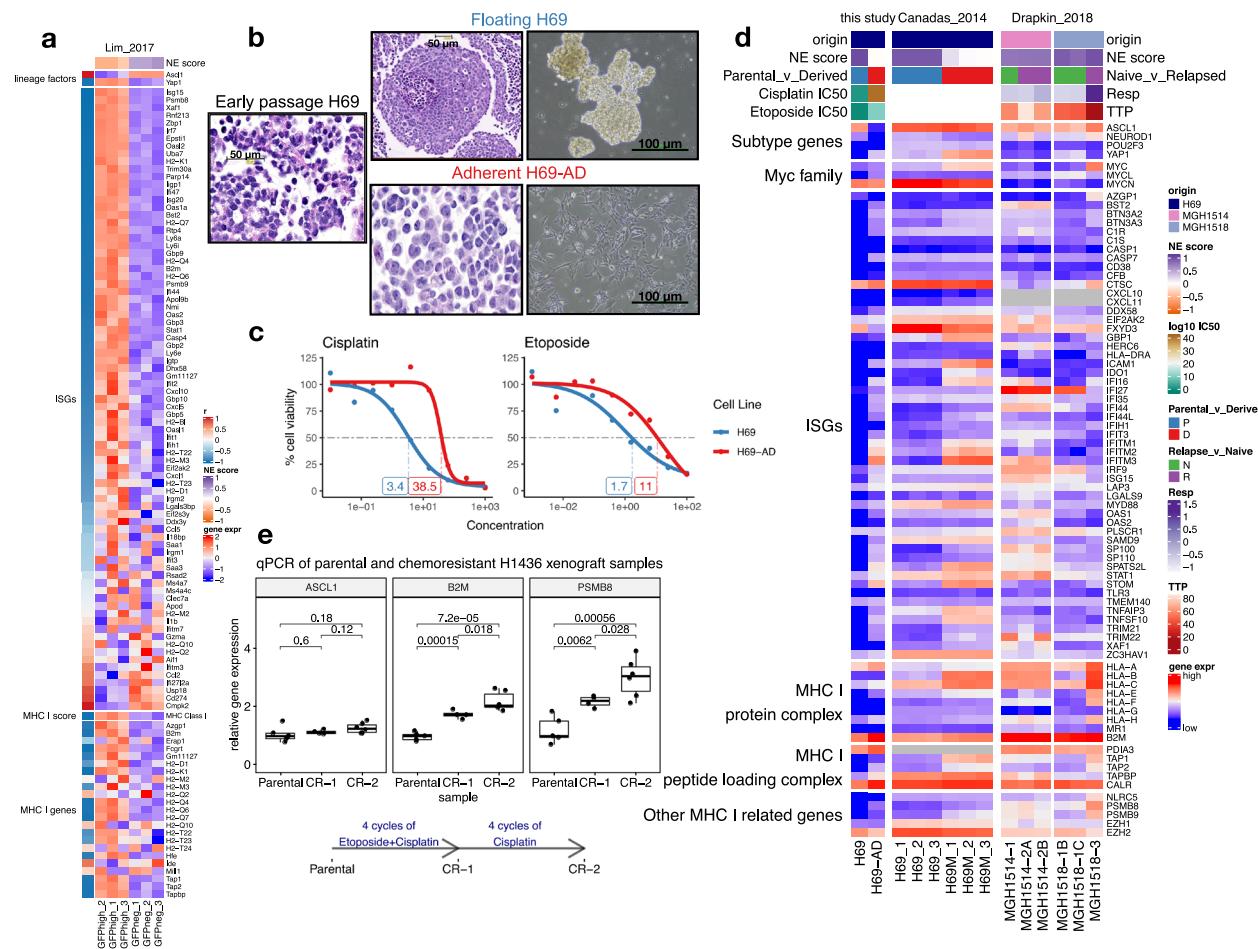
In this study, we examined NE properties of patient-derived SCLC cell lines, PDXs, and human tumors based on NE scores estimated from a gene expression signature. Currently, it is believed that Notch activation drives the lineage transition from ASCL1<sup>+</sup> to NEUROD1<sup>+</sup> to YAP1<sup>+</sup> subtype<sup>59</sup>, whereas POU2F3 + SCLC is a standalone subtype that originated from tuft cells<sup>60</sup>. While we observed mutually exclusive patterns of ASCL1 and NEUROD1 expression in cell lines, their co-expression was identified in many patient tumors. Our IHC results further revealed intratumoral heterogeneity in such tumors, suggesting ongoing lineage transition in primary treatment-naïve tumors. From the alignment of NE scores and molecular subtype-specific TF expression, we observed rare high NE-score POU2F3<sup>+</sup> tumors in three independent data sets, raising the possibility that POU2F3<sup>+</sup> tumors could also arise from NE lineage. Does this represent the fact that all of the respiratory epithelium arises from a common stem cell that then differentiates into sub-niches giving rise to the other differentiated cells? Such concept would also explain the observations in GEMMs<sup>61,62</sup> that SCLC may occasionally arise from other non-NE cells of the respiratory epithelium.



**Fig. 5 Relationship between NE scores and immune phenotypes in pan-cancer samples.** **a, b** NE scores of pan-cancer samples in the TARGET pediatric cancer cohorts (**a**) and TCGA adult cancer cohorts (**b**). **c** NE scores by immune subtype in TCGA pan-cancer samples. **d** 2D density plots visualizing the relationship between NE scores and tumor-infiltrating lymphocytes regional fraction (Pearson correlation coefficient  $-0.15$ ,  $p\text{-value} < 2.2\text{e-}16$ ) or leukocyte fraction in TCGA pan-cancer samples (Pearson correlation coefficient  $-0.31$ ,  $p\text{-value} < 2.2\text{e-}16$ ). **e** Heatmap visualizing expression of various genes and summary scores previously assessed for SCLC and now in NBL with cell line and tumor data sets side-by-side. For each data set, sample-wise z-score standardized values were used. **f** Scatter plots showing a negative correlation between NE score and Ser536 phosphorylation on NFkB-p65 in NBL cell lines. Pearson correlation coefficient was provided. \* $p\text{-value} < 0.05$ . For all box whisker plots, the centerline represents the median, the lower, and upper hinges correspond to the first and third quartiles, the whiskers extend from the hinge to the value no further than  $1.5 \times \text{IQR}$  away.

Our investigation of immune phenotypes associated with variable NE scores had identified repression of ISGs in classic high NE-score SCLC. While it remains to be determined what other pulmonary epithelial cells besides PNECs can function as cells of origin for SCLC<sup>61–65</sup>, the

gene expression similarities between PNECs and SCLC suggests many of the SCLC properties could be tied to PNEC characteristics. We confirmed ISG repression in PNECs relative to other lung epithelial cells through the examination of scRNA-seq data from



**Fig. 6 MHC I upregulation in chemoresistant SCLC.** **a** Heatmap visualizing increased expression of ISGs, MHC I genes as cells switch from Ascl1<sup>+</sup> to Yap1<sup>+</sup> in SCLC GEMM tumors from the “Lim\_2017”. GFP was expressed from an endogenous promoter of a Notch target gene *Hes1* in *Rb*<sup>-/-</sup>/*p53*<sup>-/-</sup>/*p130*<sup>-/-</sup> (TKO) background. Using flow cytometry, the authors first sorted out SCLC tumor cells and then further sorted by GFP to obtain relatively pure tumor cells with different Notch activation status. Three biological replicates were each provided for GFP negative (Notch inactive, classic high-NE) cells and GFP-positive cells (Notch-active, low-NE). **b** Different morphology and culture characteristics of adherent H69-AD and the parental H69. **c** Dose-response curves for Cisplatin and Etoposide in the H69 cell line pair. Note that H69-AD, the adherent line, is more resistant with higher IC50s. **d** Expression changes of selected genes in H69 cell line pair from this study, H69 and derived mesenchymal H69-M cell lines from “Canadas\_2014” and autologous PDX samples before and after chemotherapy from “Drapkin\_2018”. PDX parameters: TTP, time to progression, defined by time to 2× initial tumor volume; RESP, change in tumor volume between initial tumor volume and a minimum of days 14–28. Relapsed sample from MGH-1514 did not show increased chemoresistance based on the RESP and TTP parameters. Note that unlike other heatmaps, due to the small number of samples in each data set, the expression is not scaled across samples in this heatmap. **e** qPCR measurement of normalized ASCL1, B2M, and PSMB8 expression in naive parental and chemoresistant H1436 xenograft tumors. *p*-values are based on a two-sided *t*-test. The centerline in the boxplot represents the median, the lower and upper hinges correspond to the first and third quartiles, the whiskers extend from the hinge to the value no further than 1.5 \* IQR away.

healthy human lungs. As a limitation of our current study, the biological significance and mechanism of ISG repression in healthy and cancer cells of NE lineage remains to be determined with additional functional studies. ISGs provide viral defense for cells but some can be hijacked by viruses<sup>44</sup>. Since PNECs assume stem cell roles for tissue regeneration after injury<sup>7</sup>, lowering expression of ISGs and other genes involved in viral entry presumably have some role in self-protection. While we observed repression of basal ISG expression in PNECs, it remains to be determined in the presence of interferon stimulation whether PNECs would also be less primed for further activation of ISGs. Like SCLC, PNECs can switch from ASCL1<sup>+</sup> to YAP1<sup>+</sup> through Notch activation, but in the context of tissue repair<sup>7</sup>. Our findings suggest that the increase of cell-autonomous immune gene expression as high NE-score SCLCs transition to low NE-score SCLCs are mirroring the changes that normally take place during the transdifferentiation of PNECs to other lung epithelial cell types, and in SCLC, this had, in turn, led to

increased tumor-immune interaction. Interestingly, our pan-cancer analysis had extended this finding of a reciprocal relationship between neuroendocrine and immune gene expression to other cancer types. For NBL, a cancer with drastically distinct etiology compared to SCLC, we were also able to identify more inflammatory variant cancer cell line and tumors with loss of NE lineage gene expression. It would be interesting to explore more of such molecular similarities between SCLC and other tumors with NE/neuronal lineage.

The full repertoire of immune evasion strategies employed by SCLC remains to be elucidated. However, our results tying together with current clinical treatment findings raise several important questions and paradoxes. The first paradox is that we found a depletion of immune infiltrates in high NE-score SCLC tumors and other neuroendocrine tumors (NETs), associated with the down-regulation of MHC I expression. While this MHC I expression explains the presence of very few T cells, it raises the question of how

high NE-score NETs evade natural killer cells that normally would recognize the missing self that such MHC I expression loss conveys<sup>66</sup>. Thus, we feel the high NE-score low-MHC I expression pairing indicates we need to understand how natural killer cell mechanisms are avoided in NET pathogenesis. The second paradox is that low NE-score variant SCLCs appear to be associated with expression of MHC I and a more immune infiltrated tumor microenvironment, yet clinical trials of immune checkpoint-blockade do not clearly show these are the tumors responding to such therapy. Since we found these tumors also express many immunosuppressive genes it will be important to know which of these immunosuppressive gene functions need to be targeted to achieve anti-tumor immune responses. Finally, we observed the expression of MHC I in selected SCLC samples with chemoresistance and increased MYC expression even without changes in ASCL1 expression. We need to know whether immunosuppressive mechanisms are the same or different in the high vs. low NE-score SCLC resistant to chemotherapy. We conclude, that some 30 years after the first description of classic (high NE-score) and variant (low NE-score) SCLCs there are important links between these NE phenotypes and the expression of immune phenotypes, and between similar gene expression profiles of SCLC and pulmonary neuroendocrine cells. Importantly, these correlations identify important problems to be solved of clinical therapeutic translational relevance.

## Methods

**Computation of NE score.** The construction of the original NE signature has been described by Zhang et al.<sup>22</sup>. In this study, this signature has been updated with expression data from RNA-seq experiments. A quantitative NE score can be generated from this signature using the formula: NE score = (correl NE – correl non-NE)/2, where correl NE (or non-NE) is the Pearson correlation between expression of the 50 genes in the test sample and expression of these genes in the NE (or non-NE) cell line group. This score has a range of -1 to +1, where a positive score predicts for NE while a negative score predicts for non-NE cell types. The higher the score in absolute value, the better the prediction.

**Pathway enrichment analysis with GO terms.** Gene Ontology enRICHment analYsis and visuaLizAtion tool (GOrilla<sup>67</sup>, <http://cbl-gorilla.cs.technion.ac.il/>) was used to identify enriched GO terms<sup>68</sup> related to biological processes (BP) from gene lists ranked by increasing or decreasing Pearson correlation with NE scores in cell line data sets or “George\_2015” tumor data set. The *p*-value threshold was set at  $10^{-3}$  for resulting GO terms. The output was visualized by Treemap R scripts generated from “reduce + visualize gene ontology” (REVIGO<sup>69</sup>, <http://revigo.irb.hr/>) and further customized with a modified color scheme.

**Gene set enrichment analysis (GSEA).** Gene set libraries were downloaded from Enrichr<sup>24</sup> (<https://amp.pharm.mssm.edu/Enrichr/>). Fast GSEA based on gene label permutation from R package “fgsea”<sup>70</sup> was first used for a fast screening across a large number of gene set libraries. After reviewing the results for SCLC cell lines, sample label permutation-based GSEA<sup>23</sup> was run for selected gene set libraries to obtain normalized enrichment scores and multiple comparison-adjusted *p*-values. Pearson correlation was used as the ranking metric from correlating gene expression with NE scores.

**Visualization.** All heatmaps were generated by R package “ComplexHeatmap”<sup>71</sup>. Other R packages used for visualization include “ggplot2”<sup>72</sup>, “ggridges”<sup>73</sup>, “ggrepel”<sup>74</sup>, “ggpubr”<sup>75</sup>, “treemap”<sup>76</sup>, “RColorBrewer”<sup>77</sup>, “jcolors”<sup>78</sup> and “patchwork”<sup>79</sup>.

**Expression data.** Horie\_2016<sup>28</sup> is a set of microarray data that examines the effect of YAP1 KD in SCLC cell lines, the author-processed data were downloaded from GEO with accession id GSE93400. Pozo\_2020<sup>29</sup> is a set of RNA-seq data that examines the effect of ASCL1 KD in SCLC cell lines, FPKM data was obtained from the author. The original data have been deposited to GSE151002. Drapkin\_2018<sup>57</sup> is a set of RNA-seq data for SCLC PDX. Transcript per million (TPM) data processed by the original authors were downloaded from GEO with accession id GSE110853; Rudin\_2019<sup>9</sup> is a set of RNA-seq data for human SCLC tumors, it was obtained from the authors; George\_2015<sup>10</sup> is a set of RNA-seq data for human SCLC tumors, FPKM (fragments per kilobase of exon per million fragments mapped) data processed by the original authors were obtained from a supplementary table of the original publication; Jiang\_2016 is a set of RNA-seq data for human SCLC tumors<sup>80</sup>; DESeq normalized read count data was downloaded from GEO with accession id GSE60052; IGC’s Expression Project for Oncology—expO (GSE2109) and Rousseaux\_2013 (GSE30219)<sup>81</sup> are microarray data for lung cancer

samples with different histological subtypes, processed previously for the lung cancer explorer (LCE)<sup>82</sup>. Pan-cancer RNA-seq data from TCGA and TARGET processed by Expectation-Maximization (RSEM) algorithm was downloaded from Toil xena hub<sup>53</sup>. CCLE cell line RPPA data, as well as TPM RNA-seq data, were downloaded from the DepMap portal (19Q1)<sup>25</sup>. Travaglini\_2020 is a set of scRNA-seq data from healthy human lung<sup>42</sup>. Author-processed count data were downloaded from Synapse with accession id syn21041850. FACS-sorted SmartSeq2 data was used. Cell types with less than 10 cells were removed from analyses. Oudah\_2019 is a set of scRNA-seq data from healthy mouse lung<sup>7</sup>. TPM data were downloaded from GEO with accession id GSE136580. For data from GEO, R package GEOquery<sup>83</sup> was used for extracting the expression and phenotype data. Quantile normalization was performed for bulk expression data by running the “normalize.quantiles” function from R package “preprocessCore”<sup>84</sup>. Log 2 transformation of gene expression data was performed as necessary. Library size normalization was performed for author-processed scRNA-seq data by running the “library.size.normalize” function from R package “phateR”<sup>85</sup>.

**Gene signatures.** SPARCS gene set is from a study by Cañadas et al. (Fig. 1S in original article)<sup>27</sup>. Parainflammation gene set is from a study by Aran et al. (Fig. 1C in original article)<sup>38</sup>. SASP gene set is from a study by Ruscetti et al. (Fig. 2C in original article)<sup>86</sup>. InnateDB genes were downloaded from InnateDB<sup>41</sup>, non-human genes were filtered out.

Gene set “REST\_ENCODE” is from the “ENCODE\_and\_ChEA\_Consensus\_TFs\_from\_ChIP-X” library, and “IFNA-BT2” and “IFNG-BT2” are from the “LINCS\_L1000\_Ligand\_Perturbations\_up” library. Both libraries were downloaded from Enrichr<sup>24</sup>. The top 25 genes from the leading edge and are common to all SCLC data sets were selected for heatmap visualization. For ISGs, the leading edge genes from “IFNA-BT2” and “IFNG-BT2” were first combined and then the top 25 genes were selected.

Mouse ISGs, from Cilloniz et al.<sup>87</sup>, were identified from interferome<sup>88</sup> by specifying “mouse” as the species of interest and “lung” as the organ of interest. An unfiltered ISG set was used for Fig. 6A.

Human MHC I gene set is a combination of genes under GO terms “GO\_MHC\_CLASS\_I\_PROTEIN\_COMPLEX” and “GO\_MHC\_CLASS\_I\_PEPTIDE\_LOADING\_COMPLEX” from Molecular Signatures Database (MSigDB)<sup>23,89</sup>. Mouse MHC I genes were selected from GO: 0019885, “antigen processing and presentation of endogenous peptide antigen via MHC class I” based on the Mouse Genome Informatics (MGI) database<sup>90</sup>. Immune-cell-specific gene sets in human are from DisHET<sup>48</sup>. Interferon-gamma signature that predicts response to PD-1 blockade is from Ayers et al.<sup>49</sup>. The 21-gene immune checkpoint set and 995-gene immunosuppressive set are from HisgAtlas, a human immunosuppression gene database<sup>50</sup>.

**MHC I and Immune infiltrate scores.** R package GSVA<sup>91</sup> was used to compute immune infiltrate scores by single sample GSEA (ssGSEA) method<sup>23,47</sup>.

**Patients and tissue specimens.** Study participants included 18 patients who were diagnosed with SCLC and underwent surgical resection of lung cancer between 2006 and 2010 at the Department of Lung Cancer Surgery, Tianjin Medical University General Hospital. Written informed consent was obtained, and the institutional ethics committee of Tianjin Medical University General Hospital approved the study. The cases were selected based on the following criteria: (1) diagnosis of primary lung cancer clinical stage I to IV (pTNM); (2) undergoing surgical resection. Pathologic diagnosis was based on WHO criteria. Lung cancer staging for each patient has performed according to the AJCC Cancer Staging Manual, 8th edition, and was based on findings from physical examination, surgical resection, and computed tomography of the chest, abdomen, pelvis, and brain. The following information was collected from the patients’ medical records: age, gender, clinical stage, pathologic diagnosis, differentiation, lymph node status, metastasis, smoking status, and overall survival time. Resected lung and lymph node tissues were immediately immersed in liquid nitrogen until RNA extraction.

**Immunohistochemistry—histology and immunohistochemistry.** Tissue blocks, once collected, were reviewed by a staff thoracic pathologist to confirm SCLC histology. Consecutive four-micrometer-thick tissue sections were cut for immunohistochemistry. IHC staining was performed with a Bond Max automated staining system (Leica Microsystems Inc., Vista, CA) using IHC parameters optimized previously. Antibodies used in this study included ASCL1 (dilution 1:25; Clone 24B7D11.1, BD Biosciences, Cat# 556604), NEUROD1 (dilution 1:100; Clone EPR20766, Abcam, ab213725), POU2F3 (dilution 1:200; polyclonal, Novus Biologicals, NBP1-83966), CD4 (dilution 1:80; Leica Biosystems, CD4-368-L-CE-H) and CD8 (dilution 1:25; Thermo Scientific, MS-457s) in a Leica Bond Max automated stainer (Leica Biosystems Nussloch GmbH). The expression of proteins was detected using the Bond Polymer Refine Detection kit (Leica Biosystems, Cat# DS9800) with diaminobenzidine as chromogen<sup>92</sup>. The slides were counterstained with hematoxylin, dehydrated and cover-slipped. FFPE cell lines pellets with known expression of ASCL1, NEUROD1, and POU2F3 were used to establish and optimize IHC conditions and assess sensitivity and specificity for each antibody.

**Immunohistochemistry—image analysis.** The stained slides were digitally scanned using the Aperio ScanScope Turbo slide scanner (Leica Microsystems Inc.) under  $\times 200$  magnification. The images were visualized by ImageScope software (Leica Microsystems, Inc.) and analyzed using the Aperio Image Toolbox (Leica Microsystems Inc.). Different intensity levels of ASCL1, NEUROD1, or POU2F3 nuclear expression were quantified using a 4-value intensity score (0, none; 1, weak; 2, moderate; and 3, strong) and the percentage (0–100%) of the extent of reactivity. A final expression score (*H*-score) was obtained by multiplying the intensity and reactivity extension values (range, 0–300) as previously described<sup>93</sup>. For example, for SCLC-04, (3+)% nuclei is 1.22, (2+)% nuclei is 9.97, (1+)% nuclei is 47.81, the resulting *H*-score is  $3^*1.22 + 2^*9.97 + 1^*47.81 = 71.41$ .

The lymphocyte cells expressing CD4<sup>+</sup> and CD8<sup>+</sup> were counted by a pathologist using Aperio Image Toolbox analysis software (Aperio, Leica Biosystems) and expressed as cell density (CD4<sup>+</sup> and CD8<sup>+</sup> cells/mm<sup>2</sup> of analyzed tissue)<sup>92,94</sup>.

**Microarray assay.** The Human Genome U133 Plus 2.0 microarray with 54,000 probe sets was purchased from the Affymetrix (Lot#: 4032359). Total RNA was extracted with the Trizol reagent (Invitrogen) from the tissue samples. The extracted RNA was purified using the Oligotex mRNA Midi Kit (Qiagen). Then double-strand cDNA synthesis was made using a one-cycle cDNA synthesis kit (Affymetrix) and purified again by column followed by the synthesis of complementary RNA (cRNA) with in vitro transcription (IVT) kit (Affymetrix). The cRNA was fragmented after purification by column and the quality was verified by ultraviolet spectrophotometer and 1.2% denaturing agarose gel. After the test gene-chip (Lot#: 4020852, Affymetrix) was affirmed satisfactory, the real chip hybridization of cRNA fragmentation was performed and then stained and washed. Finally, the real chip was scanned in an Affymetrix scanner and the data were collected by GCOS (gene-chip operation software). CEL files were read into an AffyBatch object by “AffyBatch” function under the “affy”<sup>95</sup> R package. Alternative cdf package<sup>96</sup> “hgu133plus2hsentrezg” was downloaded from “[http://mbni.org/customcdf/22.0.0/entrezg\\_download/hgu133plus2hsentrezg.db\\_22.0.0.zip](http://mbni.org/customcdf/22.0.0/entrezg_download/hgu133plus2hsentrezg.db_22.0.0.zip)” and was specified in the function so that the resulting expression data were processed to gene level rather than the original probe level. (Probe name follows format concatenating Entrez ID for the gene and “\_at”. For example “3939\_at” corresponds to gene LDHA). The AffyBatch object was then converted to an expression set using robust multi-array average (RMA) expression measure by running the “rma” function under R package “affy”. Quantile normalization was performed by running “normalize.quantiles” function from the R package “preprocessCore”<sup>84</sup>.

**RNA-seq.** RNA samples from SCLC cell lines ( $n = 69$ ) were prepared at UT Southwestern (Dallas TX) and sent to Baylor College of Medicine (David Wheeler, Houston TX) for paired-end RNA sequencing. The analysis was then performed at UT Southwestern: Reads were aligned to the human reference genome GRCh38 using STAR-2.7<sup>97</sup> (<https://github.com/alexdobin/STAR>) and FPKM values were generated with cufflinks-2.2.1<sup>98</sup> (<http://cole-trapnell-lab.github.io/cufflinks/>). All data were then pooled, upper-quartile normalized<sup>99</sup>, and log-transformed.

**Cell culture.** All SCLC cell lines used in these studies were originally established in the John D. Minna and Adi F. Gazdar laboratories. The cultured SCLC cell lines were obtained from both the National Cancer Institute (NCI) and Hamon Cancer Center (HCC) libraries. Cells were cultured in RPMI-1640 media (Sigma Life Science, St. Louis, MO) supplemented with 5% Fetal Bovine Serum (FBS). RPMI-1640 supplemented with 5% FBS will be referred to as R5. All cells were incubated in NuAire (NuAire, Plymouth, MN) humidified incubators at 37 °C at 5% CO<sub>2</sub>. All cell lines were regularly tested for mycoplasma contamination (Bulldog Bio, Portsmouth, NH) and fingerprinted using a PowerPlex 1.2 Kit (Promega, Madison, WI) to confirm the cell line identity.

**Establishing adherent H69 (H69-AD).** The early passage of the parental H69 cell line grew as a mixture of floating and adherent cells. To enrich for adherent cells, the floating population of H69 was washed off during growth media replacement and fresh media was provided for the expansion of the remaining adherent cells. This was repeated until every passage grew as adherent cells with few to no suspension cells. This derived adherent subline was designated as H69-AD.

**Drug-response assay.** Cisplatin and Etoposide were obtained from Selleck Chemicals LLC, USA. 5000 cells of H69 and H69-AD were cultured in 100 µl R5 growth media per well in ultra-low adherent, clear, round bottom, 96 well plates (BD Biosciences, USA) for 48 h. An additional 100 µl R5 plus either a control (DMSO) or drug was added to the plate. 96 h after drug treatment, each cell line was assayed using the Cell-Titer-Glo reagent (Promega, Inc.). The fluorescence intensity was recorded at 570 nM. A standard 4-parameter log-logistic fit between the survival rate and the dosage was generated by the “drm” function from the R package “drc”<sup>100</sup>.

**Xenograft models for parental and chemoresistant SCLC tumors.** Subcutaneous xenograft in NSG mice was derived from direct implantation of

untreated H1436 cells or re-implantation of chemoresistant tumors after 4 cycles of Cisplatin and Etoposide (EC), or plus 4 cycles of Cisplatin (reduced from EC due to toxicity). Specifically, a million H1436 cells were resuspended in 100 µl mixture of serum-free RPMI-1640 and Matrigel (BD Bioscience #356237) at 1:1 ratio and immediately injected in the flank of 6–8-week-old female NSG mouse (Jackson Laboratory #005557). Mice were randomized after tumor cell injection. Treatment starts after a week when the tumor becomes palpable. 5 mg/kg/w Cisplatin (Sigma P4394) in saline, 10 mg/kg/w Etoposide (Sigma E1383) in 30% PEG 300 (Sigma 202371) were freshly prepared and administered by intraperitoneal injection, for 4 cycles in total to obtain the first group of chemoresistant tumors. An additional 4 cycles of Cisplatin were administered in the second group of mice to obtain tumors with further potentiated chemoresistance. To harvest the tumor, 10 ml digestion media was used per mouse. This was prepared freshly by supplementing 9 ml HBSS with 1 ml type IV collagenase, 50 µl DNase II and 50 µl 1 M CaCl<sub>2</sub>. Tumors were collected and placed in HBSS immediately following dissection. A fraction of the tumor was cut into a few pieces and flash-frozen in liquid nitrogen to be saved in aliquots for molecular assays. The remaining chunk was finely minced with a sterile scalpel blade. For re-implantation, the minced tissue was resuspended in digestion media, rotated at 37 °C for 20 min, filtered through a 40 µm filter, centrifuged at 300 × g for 5 min.

**Quantitative reverse transcription PCR.** ~20 mg flash-frozen tumor fragments were weighed out and homogenized in 1 ml TRIzol (Invitrogen #15596-026) in Precellys tissue homogenizing mixed beads kit (Cayman Chemical #10409). 0.2 ml chloroform (Fisher #S25248) was added to the TRIzol lysate and the mixture was vortexed for 10 s and centrifuged at 12,000 × g for 15 min at 4 °C for phase separation. 450 µl aqueous phase was collected, mixed well with 0.5 ml isopropanol (Fisher #A451-1), and precipitated RNA was collected by centrifugation at 12,000 × g for 10 min at 4 °C. The RNA pellet was rinsed in 1 ml 75% ethanol, then dissolved in 100 µl deionized water by incubating at 55 °C for 5 min. 500 ng total RNA was reverse-transcribed to cDNA in a 20 µl reaction with 4 µl iScript reverse transcription supermix (Bio-Rad #1708841) at 25 °C for 5 min, 46 °C for 20 min, and 95 °C for 1 min. The mixture was then 1:5 diluted with deionized water. Target sequences in cDNA library were amplified in 10 µl qPCR reaction (5 µl SYBR Green supermix (Bio-Rad #1725121), 0.675 µl 2.5 µM primer mix and 0.45 µl diluted cDNA) at 95 °C 10 s, 60 °C 30 s, for 40 cycles. All procedures were performed under RNase-free conditions unless specified. For data analysis, the median was taken from triplicates, normalized by Ct values of control gene *PPIA*, exponentiated with base 2 then divided by the median of parental samples. Primer sequences are provided in Supplementary Table 1.

**Study approval.** The protocol of collecting human SCLC tumor tissue for research was approved by the Ethics Committee of Tianjin Medical University General Hospital. Written informed consent was received from participants prior to inclusion in the study. Specimen collection did not interfere with standard diagnostic and therapeutic procedures. All mouse procedures were performed with the approval of the University of Texas Southwestern Medical Center Institutional Animal Care and Use Committee.

**Statistics and reproducibility.** All statistical analyses were performed with R<sup>101</sup>. Pearson correlation was used to assess the association between continuous variables. t-test was used for group comparison. Pearson’s chi-squared test was used to test for association between categorical variables. Statistical significance was set at  $p \leq 0.05$ . Benjamini–Hochberg procedures were used for generating *p*-values adjusted for multiple comparisons.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The RNA-seq gene expression data from UTSW SCLC has been added to dbGaP (accession phs001823.v1.p1)<sup>102</sup>. SCLC tumor microarray data used in this study have been deposited to GEO with accession id GSE149507. Source data used for the figures presented in this manuscript have been deposited in the Dryad Digital Repository ([https://datadryad.org/stash/share/BkmPdMrwhae1VxhkSLIG\\_532FLCqcYiMFUpY1yKmGA](https://datadryad.org/stash/share/BkmPdMrwhae1VxhkSLIG_532FLCqcYiMFUpY1yKmGA)). All other data are presented in Supplementary Data 1–5 or available from the corresponding authors upon reasonable request.

## Code availability

Scripts and input data used for this manuscript have been deposited in the Dryad Digital Repository ([https://datadryad.org/stash/share/BkmPdMrwhae1VxhkSLIG\\_532FLCqcYiMFUpY1yKmGA](https://datadryad.org/stash/share/BkmPdMrwhae1VxhkSLIG_532FLCqcYiMFUpY1yKmGA)).

Received: 7 December 2020; Accepted: 9 February 2021;

Published online: 09 March 2021

## References

- Gazdar, A. F., Bunn, P. A. & Minna, J. D. Small-cell lung cancer: what we know, what we need to know and the path forward. *Nat. Rev. Cancer* **17**, 765 (2017).
- H.R.733, U. C. Recalcitrant Cancer Research Act of 2012. <https://www.congress.gov/bill/112th-congress/house-bill/733> (2012).
- Travis, W. D. et al. The 2015 World Health Organization Classification of Lung Tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J. Thorac. Oncol.* **10**, 1243–1260 (2015).
- Augustyn, A. et al. ASCL1 is a lineage oncogene providing therapeutic targets for high-grade neuroendocrine lung cancers. *Proc. Natl Acad. Sci. USA* **111**, 14788–14793 (2014).
- Borges, M. et al. An achaete-scute homologue essential for neuroendocrine differentiation in the lung. *Nature* **386**, 852–855 (1997).
- Boers, J. E., den Brok, J. L., Koudstaal, J., Arends, J. W. & Thunnissen, F. B. Number and proliferation of neuroendocrine cells in normal human airway epithelium. *Am. J. Respir. Crit. Care Med.* **154**, 758–763 (1996).
- Ouadah, Y. et al. Rare pulmonary neuroendocrine cells are stem cells regulated by Rb, p53, and Notch. *Cell* **179**, 403–416.e423 (2019).
- Pleasance, E. D. et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463**, 184–190 (2010).
- Rudin, C. M. et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nat. Genet.* **44**, 1111–1116 (2012).
- George, J. et al. Comprehensive genomic profiles of small cell lung cancer. *Nature* **524**, 47–53 (2015).
- Doyle, A. et al. Markedly decreased expression of class I histocompatibility antigens, protein, and mRNA in human small-cell lung cancer. *J. Exp. Med.* **161**, 1135–1151 (1985).
- Burr, M. L. et al. An evolutionarily conserved function of polycomb silences the MHC class I antigen presentation pathway and enables immune evasion in cancer. *Cancer Cell* **36**, 385–401.e388 (2019).
- Yarchoan, M., Hopkins, A. & Jaffee, E. M. Tumor mutational burden and response rate to PD-1 inhibition. *N. Engl. J. Med.* **377**, 2500–2501 (2017).
- Iams, W. T., Porter, J. & Horn, L. Immunotherapeutic approaches for small-cell lung cancer. *Nat. Rev. Clin. Oncol.* <https://doi.org/10.1038/s41571-019-0316-z> (2020).
- Gazdar, A. F., Carney, D. N., Nau, M. M. & Minna, J. D. Characterization of variant subclasses of cell lines derived from small cell lung cancer having distinctive biochemical, morphological, and growth properties. *Cancer Res.* **45**, 2924–2930 (1985).
- Carney, D. N. et al. Establishment and identification of small cell lung cancer cell lines having classic and variant features. *Cancer Res.* **45**, 2913–2923 (1985).
- Johnson, B. E. et al. MYC family DNA amplification in 126 tumor cell lines from patients with small cell lung cancer. *J. Cell Biochem. Suppl.* **24**, 210–217 (1996).
- Lim, J. S. et al. Intratumoural heterogeneity generated by Notch signalling promotes small-cell lung cancer. *Nature* **545**, 360–364 (2017).
- Gazdar, A. F., Bunn, P. A. & Minna, J. D. Small-cell lung cancer: what we know, what we need to know and the path forward. *Nat. Rev. Cancer* **17**, 725–737 (2017).
- Rudin, C. M. et al. Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. *Nat. Rev. Cancer* **19**, 289–297 (2019).
- Gazdar, A. F. Morphologic and other forms of heterogeneity in small cell lung cancer: what can we learn from them? *J. Thorac. Oncol.* **13**, 148–150 (2018).
- Zhang, W. et al. Small cell lung cancer tumors and preclinical models display heterogeneity of neuroendocrine phenotypes. *Transl. Lung Cancer Res.* **7**, 32–49 (2018).
- Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
- Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
- Ghandi, M. et al. Next-generation characterization of the cancer cell line encyclopedia. *Nature* **569**, 503–508 (2019).
- Sakurai, H., Chiba, H., Miyoshi, H., Sugita, T. & Toriumi, W. IkappaB kinases phosphorylate NF-kappaB p65 subunit on serine 536 in the transactivation domain. *J. Biol. Chem.* **274**, 30353–30356 (1999).
- Canadas, I. et al. Tumor innate immunity primed by specific interferon-stimulated endogenous retroviruses. *Nat. Med.* **24**, 1143–1150 (2018).
- Horie, M., Saito, A., Ohshima, M., Suzuki, H. I. & Nagase, T. YAP and TAZ modulate cell phenotype in a subset of small cell lung cancer. *Cancer Sci.* **107**, 1755–1766 (2016).
- Pozo, K. et al. Lineage transcription factors co-regulate subtype-specific genes providing a roadmap for systematic identification of small cell lung cancer vulnerabilities. Preprint at <https://doi.org/10.1101/2020.08.13.249029> (2020).
- Dinarello, C. A. Interleukin-18, a proinflammatory cytokine. *Eur. Cytokine Netw.* **11**, 483–486 (2000).
- Cornelissen, C., Luscher-Firzlaff, J., Baron, J. M. & Luscher, B. Signaling by IL-31 and functional consequences. *Eur. J. Cell Biol.* **91**, 552–566 (2012).
- Kordass, T., Osen, W. & Eichmuller, S. B. Controlling the immune suppressor: transcription factors and microRNAs regulating CD73/NT5E. *Front. Immunol.* **9**, 813 (2018).
- Yang, L., Pang, Y. & Moses, H. L. TGF-beta and immune cells: an important regulatory axis in the tumor microenvironment and progression. *Trends Immunol.* **31**, 220–227 (2010).
- D'Acquisto, F., Perretti, M. & Flower, R. J. Annexin-A1: a pivotal regulator of the innate and adaptive immune systems. *Br. J. Pharm.* **155**, 152–169 (2008).
- Markosyan, N. et al. Tumor cell-intrinsic EPHA2 suppresses anti-tumor immunity by regulating PTGS2 (COX-2). *J. Clin. Invest.* **130**, 3594–3609 (2019).
- Reuben, A., Chung, J. W., Lapointe, R. & Santos, M. M. The hemochromatosis protein HFE 20 years later: an emerging role in antigen presentation and in the immune system. *Immun. Inflamm. Dis.* **5**, 218–232 (2017).
- Kouo, T. et al. Galectin-3 shapes antitumor immune responses by suppressing CD8+ T cells via LAG-3 and inhibiting expansion of plasmacytoid dendritic cells. *Cancer Immunol. Res.* **3**, 412–423 (2015).
- Aran, D. et al. Widespread parainflammation in human cancer. *Genome Biol.* **17**, 145 (2016).
- Coppe, J. P. et al. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.* **6**, 2853–2868 (2008).
- Chien, Y. et al. Control of the senescence-associated secretory phenotype by NF-kappaB promotes senescence and enhances chemosensitivity. *Genes Dev.* **25**, 2125–2136 (2011).
- Breuer, K. et al. InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Res.* **41**, D1228–D1233 (2013).
- Travaglini, K. J. et al. A molecular cell atlas of the human lung from single cell RNA sequencing. *Nature* **587**, 619–625 (2019).
- Zhang, H., Deo, M., Thompson, R. C., Uhler, M. D. & Turner, D. L. Negative regulation of Yap during neuronal differentiation. *Dev. Biol.* **361**, 103–115 (2012).
- Schoggins, J. W. Interferon-stimulated genes: what do they all do? *Annu. Rev. Virol.* **6**, 567–584 (2019).
- Pfaender, S. et al. LY6E impairs coronavirus fusion and confers immune control of viral disease. *Nat. Microbiol.* **5**, 1330–1339 (2020).
- Shen, L. W., Mao, H. J., Wu, Y. L., Tanaka, Y. & Zhang, W. TMPRSS2: A potential target for treatment of influenza virus and coronavirus infections. *Biochimia* **142**, 1–10 (2017).
- Barbie, D. A. et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108–112 (2009).
- Wang, T. et al. An empirical approach leveraging tumorgrafts to dissect the tumor microenvironment in renal cell carcinoma identifies missing link to prognostic inflammatory factors. *Cancer Discov.* **8**, 1142–1155 (2018).
- Ayers, M. et al. IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* **127**, 2930–2940 (2017).
- Liu, Y. et al. HigAtlas 1.0: a human immunosuppression gene database. *Database* **2017**, <https://doi.org/10.1093/database/bax094> (2017).
- Lee, S. & Margolin, K. Cytokines in cancer immunotherapy. *Cancers* **3**, 3856–3893 (2011).
- Balanis, N. G. et al. Pan-cancer convergence to a small-cell neuroendocrine phenotype that shares susceptibilities with hematological malignancies. *Cancer Cell* **36**, 17–34.e17 (2019).
- Vivian, J. et al. Toil enables reproducible, open source, big biomedical data analyses. *Nat. Biotechnol.* **35**, 314–316 (2017).
- Thorsson, V. et al. The immune landscape of cancer. *Immunity* **48**, 812–830.e814 (2018).
- Pugh, T. J. et al. The genetic landscape of high-risk neuroblastoma. *Nat. Genet.* **45**, 279–284 (2013).
- Canadas, I. et al. Targeting epithelial-to-mesenchymal transition with Met inhibitors reverts chemoresistance in small cell lung cancer. *Clin. Cancer Res.* **20**, 938–950 (2014).
- Drapkin, B. J. et al. Genomic and functional fidelity of small cell lung cancer patient-derived xenografts. *Cancer Discov.* **8**, 600–615 (2018).
- Huang, F. et al. Guanosine triphosphate links MYC-dependent metabolic and ribosome programs in small cell lung cancer. *J. Clin. Invest.* <https://doi.org/10.1172/JCI139929> (2020).
- Ireland, A. S. et al. MYC drives temporal evolution of small cell lung cancer subtypes by reprogramming neuroendocrine fate. *Cancer Cell* <https://doi.org/10.1016/j.ccr.2020.05.001> (2020).
- Huang, Y. H. et al. POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer. *Genes Dev.* **32**, 915–928 (2018).

61. Yang, D. et al. Intertumoral heterogeneity in SCLC is influenced by the cell type of origin. *Cancer Discov.* **8**, 1316–1331 (2018).
62. Sutherland, K.D., Proost, N., Brouns, I., Adriaensen, D., Song, J.-Y., Berns, A. 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 (2011).
63. Park, K. S. et al. Characterization of the cell of origin for small cell lung cancer. *Cell Cycle* **10**, 2806–2815 (2011).
64. Semenova, E. A., Nagel, R. & Berns, A. Origins, genetic landscape, and emerging therapies of small cell lung cancer. *Genes Dev.* **29**, 1447–1462 (2015).
65. Gazdar, A. F. et al. The comparative pathology of genetically engineered mouse models for neuroendocrine carcinomas of the lung. *J. Thorac. Oncol.* **10**, 553–564 (2015).
66. Vivier, E. et al. Innate or adaptive immunity? The example of natural killer cells. *Science* **331**, 44–49 (2011).
67. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOriella: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
68. The Gene Ontology Consortium The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019).
69. Supek, F., Bosnjak, M., Skunca, N. & Smuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **6**, e21800 (2011).
70. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M. N., Sergushichev, A., Fast gene set enrichment analysis. Preprint at <https://www.biorxiv.org/content/10.1101/060012v3.article-info> (2016).
71. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
72. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (2016).
73. Wilke, C. O., *ggridges: Ridgeline Plots in ggplot2* (2020).
74. Slowikowski K., *ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'* (2019).
75. Kassambara, A., *ggpubr: 'ggplot2' Based Publication Ready Plots* (2020).
76. Tennekes, M., *treemap: Treemap Visualization* (2017).
77. Neuwirth, E., *RCColorBrewer: ColorBrewer Palettes* (2014).
78. Huling, J., *jcolor: Colors Palettes for R and 'ggplot2', Additional Themes for 'ggplot2'* (2020).
79. Pedersen, T. L., *patchwork: The Composer of Plots* (2019).
80. Jiang, L. et al. Genomic landscape survey identifies SRSF1 as a key oncodriver in small cell lung cancer. *PLoS Genet.* **12**, e1005895 (2016).
81. Rousseaux, S. et al. Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci. Transl. Med.* **5**, 186ra166 (2013).
82. Cai, L. et al. LCE: an open web portal to explore gene expression and clinical associations in lung cancer. *Oncogene* **38**, 2551–2564 (2019).
83. Davis, S. & Meltzer, P. S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* **23**, 1846–1847 (2007).
84. Bolstad, B. *preprocessCore: A Collection of Pre-Processing Functions* (2019).
85. Moon, K. R. et al. Visualizing structure and transitions in high-dimensional biological data. *Nat. Biotechnol.* **37**, 1482–1492 (2019).
86. Ruscetti, M. et al. NK cell-mediated cytotoxicity contributes to tumor control by a cytostatic drug combination. *Science* **362**, 1416–1422 (2018).
87. Cilloniz, C. et al. Molecular signatures associated with Mx1-mediated resistance to highly pathogenic influenza virus infection: mechanisms of survival. *J. Virol.* **86**, 2437–2446 (2012).
88. Rusinova, I. et al. Interferome v2.0: an updated database of annotated interferon-regulated genes. *Nucleic Acids Res.* **41**, D1040–D1046 (2013).
89. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
90. Eppig, J. T. Mouse Genome Informatics (MGI) Resource: genetic, genomic, and biological knowledgebase for the laboratory mouse. *ILAR J.* **58**, 17–41 (2017).
91. Hanzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).
92. Parra, E. R. et al. Image analysis-based assessment of PD-L1 and tumor-associated immune cells density supports distinct intratumoral microenvironment groups in non-small cell lung carcinoma patients. *Clin. Cancer Res.* **22**, 6278–6289 (2016).
93. Fujimoto, J. et al. G-protein coupled receptor family C, group 5, member A (GPRC5A) expression is decreased in the adjacent field and normal bronchial epithelia of patients with chronic obstructive pulmonary disease and non-small-cell lung cancer. *J. Thorac. Oncol.* **7**, 1747–1754 (2012).
94. Tang, C. et al. Development of an immune-pathology informed radiomics model for non-small cell lung cancer. *Sci. Rep.* **8**, 1922 (2018).
95. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
96. Dai, M. et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* **33**, e175 (2005).
97. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
98. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
99. Bullard, J. H., Purdom, E., Hansen, K. D. & Dudoit, S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* **11**, 94 (2010).
100. Ritz, C., Baty, F., Streibig, J. C. & Gerhard, D. Dose-response analysis using R. *PLoS ONE* **10**, e0146021 (2015).
101. R Core Team. *R: A Language and Environment for Statistical Computing* (2019).
102. McMillan, E. A. et al. Chemistry-first approach for nomination of personalized treatment in lung cancer. *Cell* **173**, 864–878.e829 (2018).

## Acknowledgements

Support for this work comes from the National Institutes of Health [R135GM136375, R35CA22044901, 5P30CA142543, 5U01CA213338-04, 3P50CA070907, and R01GM115473], and the Cancer Prevention Research Institute of Texas [RP180805]. We dedicate this work to Dr. Adi F. Gazdar, who had initially conceptualized and led this project until his passing in December 2018. We thank Ms. Jessie Norris for proofreading the manuscript.

## Author contributions

Conception and design, A.F.G. and L.C.; development of methodology, L.C., L.G., and T.W.; acquisition of data, L.C., H.L., F.H., J.F., L.G., J.C., Y.L., Y.-A.Z., D.D., V.S., K.P., C.S.K., C.Y., A.A., K.H., M.P.-Z., and B.D.; analysis and interpretation of data, L.C., J.F., G.J., L.Y., and W.Z.; writing, review and/or revision of the manuscript, L.C., H.L., E.A., R.J.D., and J.D.M.; study supervision, D.S.S., I.I.W., J.E.J., G.X., J.D.M., Y.X., and A.F.G.

## Competing interests

J.D.M. receives licensing fees from the NCI and UT Southwestern to distribute cell lines. R.J.D. is on the advisory board for Agios Pharmaceuticals. D.S.S. and W.Z. are currently employed by Genentech Inc. and own stock in Roche Holdings. I.I.W. is a speaker at Medscape, MSD, Genentech/Roche, PlatformQ Health, Pfizer, AstraZeneca, Merck; receives research support from Genentech, Oncoplex, HTG Molecular, DepArray, Merck, Bristol-Myers Squibb, MedImmune, Adaptive, Adaptimmune, EMD Serono, Pfizer, Takeda, Amgen, Karus, Johnson & Johnson, Bayer, Iovance, 4D, Novartis, and Akoya; and is on the advisory boards for Genentech/Roche, Bayer, Bristol-Myers Squibb, AstraZeneca/MedImmune, Pfizer, HTG Molecular, Asuragen, Merck, GlaxoSmithKline, Guardant Health, Oncocyte, and MSD. The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-021-01842-7>.

**Correspondence** and requests for materials should be addressed to L.C., J.D.M. or Y.X.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.