Cooperating lineage factors SOX2 and BRN2 govern the transcriptional program of the neural subtype of lung squamous cell carcinoma

Takashi Sato[1,2], Seungyeul Yoo[3], Ranran Kong[1,2,4], Maya Fridrikh[1,2], Abhilasha Sinha[1,2], Prashanth Chandramani-Shivalingappa[1], Ayushi Patel[1,2], Osamu Nagano[5], Takashi Masuko[6], Mary Beth Beasley[7], Charles A. Powell[1], Jun Zhu[2,3,8] & Hideo Watanabe[1,2,3,*]

[1]Division of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [2]Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [3]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [4]Department of Thoracic Surgery, The Second Affiliated Hospital of Medical School, Xi'an Jiaotong University, Xi'an, Shaanxi 710004, China. [5]Division of Gene Regulation, Institute for Advanced Medical Research, Keio University School of Medicine, Shinjuku-ku, Tokyo 160-8582, Japan. [6]Cell Biology Laboratory, School of Pharmacy, Kindai University, Higashiosaka, Osaka 577-8502, Japan. [7]Department of Pathology and Laboratory Medicine,

Icahn School of Medicine at Mount Sinai, NY 10029, USA. [8]Sema4, a Mount Sinai venture, Stamford, CT 06902, USA. *Correspondence: hideo.watanabe@mssm.edu

Abstract

Lung squamous cell carcinoma (LUSC) is one of the major subtypes of lung cancer, molecular characterization of which has not sufficiently improved its treatment strategies. Accumulating evidence suggests that lineage-specific transcriptional regulators control differentiation states not only during normal development but also during cancer evolution. By investigating the super-enhancer landscape of LUSC, we identified a unique 'neural' subtype defined by Sox2 and a neural lineage factor Brn2. Robust protein-protein interaction and genomic co-occupancy of these factors indicated their transcriptional cooperation in this 'neural' LUSC in contrast to the cooperation of Sox2 and p63 in the classical LUSC. Introduction of p63 expression in the 'neural' LUSC invoked the classical LUSC lineage accompanied by Brn2 downregulation and increased activities of ErbB/Akt and MAPK-ERK pathways. Collectively, our data demonstrate a unique LUSC lineage featured by Sox2 cooperation with Brn2 instead of p63, for which distinct therapeutic approaches may be warranted.

Introduction

Lung cancer is the leading cause of cancer-related death worldwide[1]. Despite recent progress in diagnosis and treatment including molecular-targeted therapeutics and immunotherapy which provide a considerable survival benefit for lung cancer patients, the overall 5-year survival remains less than adequate at 19%[2]. Lung squamous cell carcinoma (LUSC) is the second most common histological subtype of lung cancer and is strongly associated with cigarette smoking. Although achievements in tobacco control in developed countries have contributed to a decline in mortality rate of lung cancer, LUSC still remains a major cause of death worldwide and shows even higher incidence than lung adenocarcinoma in several countries[3,4]. Comprehensive genomic characterization of LUSC conducted by The Cancer Genome Atlas (TCGA) project revealed its heterogeneous features with complex genomic alterations[5]. Unfortunately, these findings have not resulted in the successful development of clinically approved targeted drugs for LUSC. Compared to patients with lung adenocarcinoma, where multiple genome-based targeted therapeutics have been approved, LUSC patients have very limited therapeutic options. More recently, immune checkpoint inhibitors have emerged as key therapeutic options for solid tumors including LUSC[6,7]; however, the response

rate for unselected LUSC population remains only approximately 20% and initially-responded tumors eventually progress in most cases. Therefore, innovative strategies to better characterize and classify LUSC for a better patient stratification for current and future therapeutic options are desperately needed.

We previously identified *SOX2* as the most commonly amplified oncogene in LUSC[8]. We proposed *SOX2* as a lineage-survival oncogene in squamous cell cancers for its essential role during the development in the specification of the squamous cell lineages by opposing the role of Nkx2-1 in the dividing foregut and its essentiality for LUSC cell survival[9]. In a following study, we identified another squamous lineage factor, p63 as an important cooperative partner of Sox2 in LUSC[10]. *SOX2* amplification on chromosome 3q in LUSC often extends to its telomeric side to include the locus of *TP63*, which encodes p63. While focal amplification and/or overexpression of both *SOX2* and *TP63* genes is found in only 7 % of LUSC tumors, broader copy number gains on 3q telomeric ends are observed in the vast majority of LUSCs[5]. Studies on expression profiles classified LUSCs into four expression subtypes (primitive, classical, secretory, and basal), suggesting the heterogeneity of transcriptional programs within LUSCs[5,11,12]. Based on this classification, co-amplification of *SOX2* and *TP63* is enriched in the

classical subtype of LUSC whereas *TP63* expression is relatively low in the primitive and secretory subtypes. However, it remains largely unknown what mechanisms are involved in controlling transcriptional programs in the heterogeneous group of LUSCs.

A series of recent genome-wide histone modification analyses have demonstrated the presence of large clusters of putative enhancers in close genomic proximity, coined super-enhancers[13,14]. Since these regions typically exhibit cell-lineage-specific patterns in health and in disease, super-enhancer profiling is becoming a powerful tool to identify novel cancer cell lineages governed by specific transcriptional regulators. For example, analysis of the super-enhancer landscape of neuroblastoma has demonstrated heterogeneity of lineage states governed by specific transcriptional programs that may underlie the cause of relapse after chemotherapy[15]. Similar analysis of the super-enhancer landscape in acute myeloid leukemia identified a novel epigenomic subtype, which have therapeutic implication for differentiation therapy[16]. In the present study, we investigated super-enhancer profiles in LUSC and identified a novel subtype in which a neural transcription factor Brn2 has a key role in determining its distinctive differentiation state. In addition, we show that Brn2 serves as an interacting partner for Sox2 in this novel subtype, instead of p63 in the classical subtype of LUSC and that forced

expression of p63 leads to classical squamous cell differentiation and suppression of Brn2 in this novel LUSC subtype.

Results

Super-enhancer profiling identifies a novel subtype of lung squamous cell carcinoma

To understand the inter-tumor heterogeneity of cell lineages in LUSC, we examined super-enhancer landscape in a panel of 13 LUSC cell lines. Unsupervised hierarchical clustering of these cell lines using H3K27 acetylation signals over super-enhancer regions near transcriptional regulators identified three subgroups of LUSC (Fig. 1a). Principal component analysis (PCA) using those signals supported this classification (Supplementary Fig. 1a). One of the three subgroups consists of five LUSC cell lines, in which common super-enhancers are observed at genetic loci of 9 transcriptional regulators including *SOX2* and *TP63* loci (Supplementary Fig. 1b and Supplementary Table 1), consistent with our previous findings that these genes play essential roles as lineage oncogenes in typical LUSCs[8,10]. This subgroup is enriched in *SOX2*-amplified LUSCs with high expression of both *SOX2* and *TP63* (Fig. 1b). In contrast, LK2 and NCI-H520 cells formed a small subgroup (Fig. 1a and Supplementary Fig. 1a),

in which the common super-enhancers with highest signals lie on only two loci at the *SOX2* gene and the *POU3F2* gene, which encodes Brn2 (Fig. 1c, d and Supplementary Table 2). Brn2 is a well-known lineage factor in neural progenitor cells, particularly in the hypothalamus, where it partners with Sox2 to exert its transcriptional functions [17-19]. We also noted that the third subgroup which consists of 5 cell lines shows relatively low expression of both *SOX2* and *TP63* (Fig. 1b) although human LUSC tumors with low expression (< mean–SD) of both genes accounts for only 6.6 % in the Cancer Genome Atlas Lung Squamous Cell Carcinoma (TCGA-LUSC) dataset. While the *SOX2* super-enhancers were shared by the former two subgroups (Fig. 1c, d and Supplementary Fig. 1b) and the *TP63* super-enhancers were found only in the first subgroup (Fig. 1d and Supplementary Fig. 1b), *POU3F2* super-enhancers were present only in the small subset of LK2 and NCI-H520 (Fig. 1d). We also found the *POU3F2* super-enhancers to be among the top differential super-enhancers in this subset compared to the 'classical' SOX2/TP63 subgroup (Supplementary Fig. 1c). Analyses using RNA-seq data on these cell lines obtained through Cancer Cell Line Encyclopedia (CCLE) dataset revealed that expression of *POU3F2* was indeed significantly higher in this small subset while squamous lineage factor *TP63* was expressed significantly higher in the 'classical' *SOX2*/*TP63* subgroup of

LUSC (Fig. 1b). By contrast, expression level of *SOX2* was found in both subgroups and highest in the *SOX2*/*POU3F2* subset (Fig. 1b). Immunoblotting showed the expression pattern to be consistent with protein level; the LUSC cell lines that harbor *POU3F2* super-enhancers do not express p63 at a detectable level but have high expression of Sox2 and Brn2 (Fig. 1e). These findings suggest that, while *SOX2* is highly expressed and likely serves as a lineage factor for both of these two subgroups, this small subgroup represents a unique subset of LUSCs signified by the neural lineage factor Brn2.

BRN2 signifies the neural subtype of LUSC in human lung squamous cell carcinomas

Brn2 has been described to have an oncogenic role in neuroendocrine tumors such as small cell lung cancer (SCLC), neuroendocrine prostate cancer and glioblastoma [20-24]. To determine whether this unique 'neural' subset exists in human LUSC data and *POU3F2 (Brn2)* signifies this subgroup, we examined TCGA-LUSC dataset and identified a small subset of LUSC that expresses *POU3F2* (Fig. 2a) and their expression levels were comparable to that in SCLC [25] and hypothalamus in GTEx data[26] (Fig. 2b), suggesting functional relevancy of Brn2 in this unique subset of LUSC. Of note, this 'neural' subset of LUSC forms a distinct cluster from

super-enhancer profiles of SCLC cell lines (data not shown). In order to investigate how these tumors with Brn2 expression are represented in clinical specimens, we examined Brn2 expression pattern at protein level in human LUSC tissues by immunohistohemical staining. While all 10 specimens were positive for Sox2 as expected, we found Brn2 expression in 2 specimens (Fig. 2c and Supplementary Fig. 2a). Of note, the Brn2 expression was relatively confined to subpopulations of tumor cells in these two LUSC tissues. This suggests that some of the tumors with intermediate *POU3F2* expression level observed in bulk mRNA in the TCGA-LUSC dataset may reflect heterogeneous tumor populations with partial 'neural' differentiation and that, perhaps unlike cell lines, distinct lineage states can coexist within the same LUSC tumor. This further implies transdifferentiation during the evolution of LUSCs. We next examined the correlation of expression of *POU3F2* with *SOX2* and *TP63* expression in the TCGA-LUSC dataset. We found that expression of *POU3F2* was anti-correlated with that of *TP63* (Fig. 2b and Supplementary Fig. 2b), consistent with the data from LUSC cell lines in which high expression of *POU3F2* and *TP63* were mutually exclusive. We did not find significant correlation between *POU3F2* and *SOX2* expression (Supplementary Fig. 2c) in the entire dataset; however, when we focused on tumors with low *TP63* expression, we found a

significant correlation of *POU3F2* with *SOX2* (Fig. 2d) whereas there was a significant correlation of *TP63* with *SOX2* among LUSCs with high *TP63* expression (Supplementary Fig. 2d). These findings suggest that Brn2 and p63 have contrasting roles representing distinct lineages within LUSC by partnering with Sox2 to elicit distinct transcriptional outputs.

To investigate the difference in gene signatures between *POU3F2*-expressing LUSC tumors and *TP63*-expressing LUSC tumors, we next performed differential expression analysis in the TCGA-LUSC dataset comparing *POU3F2*-high/ *TP63*-low tumors vs. *TP63*-high/ *POU3F2*-low tumors (Fig. 2e). A functional enrichment analysis of differentially expressed genes identified genes which play roles in neural cell differentiation significantly enriched in *POU3F2*-high/ *TP63*-low tumors while those involved in epithelial/ epidermal cell differentiation were significantly enriched in *TP63*-high/ *POU3F2*-low tumors (Fig. 2f). Further, we evaluated the prognostic significance of *TP63*/ *POU3F2* expression status in the TCGA-LUSC dataset, finding that patients with *POU3F2*-high/ *TP63*-low tumors showed trends toward shorter survival, although the difference was not statistically significant (Supplemental Fig. 2e). These findings suggest the presence of a distinct LUSC subset with neural differentiation that has not been described previously.

BRN2 and SOX2 interact and co-localize at genetic loci in a variant subset of LUSC

Given the positive correlation of *POU3F2* and *SOX2* expressions in LUSC tumors with low *TP63* expression, and previously reported interaction of Brn2 and Sox2 in neural progenitor cells[18,19], we further investigated the relationship between Brn2 and Sox2 in LUSC. First, using immunofluorescence analysis we found that Sox2 and Brn2 co-localize in the nuclei of the Brn2-positive LUSC cell lines (Fig. 3a and Supplementary Fig. 3a). In contrast, Sox2 but not Brn2 was detected in nuclei of the 'classical' Sox2/p63-positive LUSC cells confirming the specificity of the antibody (Supplementary Fig. 3b). Second, we observed their robust protein-protein interaction by endogenous and reciprocal co-immunoprecipitations in the two 'neural' LUSC cell lines (Fig. 3b). Given their physical interaction, we hypothesized that Brn2 and Sox2 co-occupy genomic loci to coordinately regulate gene expression in these Brn2-positive LUSC cells. To test for genomic co-occupancy of Brn2 and Sox2, we explored genome-wide binding profiles of Brn2 and Sox2 in the 'neural' LUSC cells by chromatin immunoprecipitation sequencing (ChIP-seq) and found that the Brn2 binding peaks substantially overlap with the peaks for Sox2 in these cells (Fig. 3c, d). In contrast, the Sox2 binding peaks

from 'classical' LUSC cells showed distinct profiles with less overlaps with those for Brn2 in 'neural' LUSC cells (Fig. 3c, d), suggesting that in the absence of p63, Brn2 engages Sox2 to localize at genetic loci thereby imparting 'neural' features on LUSC cells.

p63 suppresses Brn2 expression in the LUSC cells

We previously reported that p63 and Sox2 physically interact and exhibit overlapping genomic occupancy, which suggest their cooperative transcriptional programs in 'classical' LUSCs[10]. The interaction of Brn2, in place of p63, with Sox2 that we observed in this 'neural' subset implies counteractive functions of Brn2 and p63 in lineage determination. To date, however, the role of the balance between these two factors in differentiation of squamous tissue, or in any lineages, has not been described. Therefore, to determine whether p63 has any effects on Brn2 or vice versa, we first overexpressed deltaNp63-alpha (DNp63), the predominant isoform in squamous lineage and LUSC, in the 'neural' LUSC cells (Fig. 4a and Supplementary Fig. 4a, b). Introduction of DNp63 substantially suppressed Brn2 expression, while it also suppressed expression of Sox2 to a lesser extent in these cells (Fig. 4a and Supplementary Fig. 4a, b). Notably, DNp63 overexpression led to morphological changes in these 'neural' LUSC

cells from small and/or elongated morphology to larger and cuboidal morphology in conventional 2-dimensional culture (Fig. 4b and Supplementary Fig. 4c) and decreased cell proliferation (Fig. 4c and Supplementary Fig. 4d). To confirm whether these observed changes *in vitro* reflect a shift in differentiation status, we next investigated the effects of DNp63 over-expression on the 'neural' LUSC *in vivo*. In the xenograft model, the tumor growth rates were significantly decreased in DNp63-overexpressed LUSC xenografts compared to control 'neural' LUSC (Fig. 4d). In addition, Brn2 expression was suppressed in p63-positive tumor cells (Fig. 4d and Supplementary Fig. 4e, f). Consistent with our finding *in vitro*, DNp63 expression downregulated expression of Sox2 (Supplementary Fig. 4d, f, g) presumably reflecting the squamous epithelial differentiation induced by p63[27,28]. Notably, patches of p63-negative tumor cells in DNp63-overexpressed xenograft tumor shows reduced expression of Brn2 and Sox2 (Supplementary Fig. 4e, f), likely due to imperfect selection or heterogeneous ectopic expression of DNp63, exhibiting a clear effect of DNp63 on suppressing Brn2. Morphologically, p63-positive tumor cells showed open chromatin, larger nucleoli, and on average, had a polygonal shape and more cytoplasm indicative of more classical squamous cell carcinoma histology while p63-negative/ Brn2-positive cells have a slightly oval shape, less

cytoplasm, denser chromatin and less frequent/smaller nucleoli (Fig. 4d and Supplementary Fig. 4f). These results suggest that DNp63 counteracts with Brn2 and induces a different tumor state characterized by more squamous/epithelial features of the cells.

On the other hand, overexpression of Brn2 in p63-positive 'classical' LUSC cells did not lead to suppression of p63 while modestly increased expression of Sox2, indicating non-reciprocal regulatory roles of Brn2 and DNp63, and regulation of Sox2 by Brn2 (Supplementary Fig. 5a, b). To further investigate the role of Brn2 in the 'neural' LUSC cells, we ablated expression of Brn2 in these cells by CRISPR-Cas9 mediated deletion of the *POU3F2* gene (Supplementary Fig. 5c). Brn2 ablation led to decreased cell proliferation in 'neural' LK2 cells (Supplementary Fig. 5d), suggesting its role in maintaining the lineage state; however, it did not lead to changes in p63 expression (Supplementary Fig. 5c), again suggesting that p63 expression is not under the regulation of Brn2. Although Brn2 overexpression led to increased Sox2 expression, Brn2 ablation did not result in suppression of Sox2 (Supplementary Fig. 5c). Instead, p63 ablation increased expression of Sox2 (Supplementary Fig. 5e), suggesting negative regulation of Sox2 by p63, consistent with the findings that p63-high LUSC cells have lower

expression of Sox2 than the Brn2-positive subset (Fig. 1e) and DNp63 induction suppressed SOX2 expression in the same subset (Fig. 4a and Supplementary Fig. 4a, b).

DNp63 changes the transcriptional program of 'neural' LUSC cells

To investigate how DNp63 changed transcriptional programs leading to the phenotypical changes in 'neural' LK2 cells, we performed transcriptomic profiling on these engineered cells (Fig. 5a). We found that genes which play roles in neuronal functions were enriched in down-regulated genes while genes involved in epithelial/ epidermal cell development were enriched in up-regulated genes (Fig. 5a). When we considered the levels of DNp63 expression depending on the experimental models, in addition to neuronal genes, those involved in cell cycle/ DNA replication were also enriched in down-regulated genes (Supplementary Fig. 6a). Consistent with this finding, higher proportion of cells in the S phase of cell cycle was found in control LK2 cells compared to the DNp63-overexpressed cells (Supplementary Fig. 6b, c). Altered cell-cycle dynamics might be a simply reflection of faster cell proliferation but it is also plausible that it is associated with neuronal differentiation in these 'neural' LUSC cells, which needs further investigation[29-31]. Of note, we found strong associations (OR=9.81 and 6.17 for up-

and down-regulated genes, respectively) between gene signatures in this model and those from human TCGA-LUSC tumors based on *POU3F2*/ *TP63* expression (Fig. 2e and Supplementary Table 3), suggesting biological relevancy of our model in the context of human LUSCs. To see whether introduction of DNp63 reconstitutes Sox2-binding profile toward classical LUSC, we profiled genomic occupancy of Sox2 in the DNp63-overexpressed cells and compared to the control 'neural' LK2 cells. We found that while average ChIP-seq signals of all the Sox2 peaks decreased in the DNp63-overexpressed cells compared to the control LK2 cells, the DNp63-overexpressed cells showed increased signals near the Sox2 peaks exclusive to the 'classical' LUSC lines (Fig. 5c), suggesting that presence of DNp63 re-engages Sox2 cistrome in the 'classical' LUSC cells specific Sox2-binding regions (Fig. 5d). Furthermore, we found p63 binding profile formed in the DNp63-overexpressed cells similar to that in 'classical' HCC95 cells including overlaps with Sox2 binding peaks from these cells (Fig. 5d and Supplementary Fig. 6d). These findings suggest that DNp63 overrides Sox2's transcriptional programs in the 'neural' state to redefine it into 'classical' squamous-cell state via superseding Brn2 as a Sox2 partner.

DNp63 changes activities of ErbB family signaling in 'neural' LUSC cells

Distinct patterns of cell signaling landscape including receptor tyrosine kinase (RTK) activity are associated with cell differentiation states. To investigate whether DNp63-overexpression induces a differential signaling activity profile in the 'neural' LUSC cells, we assayed for human phospho-RTK array on these cells (Fig. 6a). At the original state in the control cells, we detected ErbB4 phosphorylation along with activation of EGFR and IGF1R. In contrast, ErbB4 phosphorylation was diminished and instead higher signals of ErbB3 phosphorylation emerged in the DNp63-overexpressed cells.

We confirmed this switch in activity between ErbB family members by immunoblotting and further examined the difference in ErbB family signaling with the use of recombinant Neuregulin-1 (NRG1), a ligand for ErbB3 and ErbB4, as well as a neutralizing antibody directed against ErbB4 (Fig. 6b and Supplementary Fig. 7a). This line of experiments revealed that diminished phosphorylation of ErbB4 upon DNp63 overexpression was a consequence of suppressed ErbB4 protein level. In contrast, expression levels of ErbB3 are similar between DNp63 overexpressed and control cell lines while its activity was significantly increased upon DNp63 overexpression under the same growth condition (10% FBS)

(Supplementary Fig. 7a), suggesting that DNp63 expression led to altered expression levels of its ligands or expression/activity of its heterodimeric partners and/or interacting signaling molecules. Indeed, we observed increased protein level of EGFR upon DNp63 overexpression, and in response to NRG1 stimulation, phosphorylation of ErbB3 and EGFR increased to even higher level in the DNp63-overexpressed cells (Fig. 6b and Supplementary Fig. 7a), suggesting that increased EGFR protein level and their homo/hetero-dimerization led to increased phosphorylation of ErbB3 and EGFR. We also found that p63 binds to the *EGFR* locus in the DNp63 over-expressed cells suggesting a direct regulation of *EGFR* gene by p63 at the transcriptional level (Supplementary Fig. 7b).

To determine which major downstream signal pathway mediates ErbB3/EGFR activation, we examined phosphorylation levels of Akt and Erk1/2 in these cells. We observed increased Akt phosphorylation levels corresponding to phosphorylation levels of ErbB3 and EGFR in response to NRG1 stimulation (Fig. 6b and Supplementary Fig. 7a). In contrast, Erk1/2 phosphorylation levels were higher in the DNp63-overexpressed cells, but did not change after NRG1 stimulation, suggesting that the activation of MAPK-ERK pathway by DNp63 overexpression was independent from the phosphorylation of ErbB3 and EGFR. We next

examined response to NRG1 stimulation on cell growth in these cells and found significantly higher response in the DNp63-overexpressed cells compared to the control cells (Fig. 6c). These data suggest that the 'neural' LUSC cells are less dependent on ErbB/Akt and MAPK-ERK pathways compared to the more epithelial cell differentiation state induced by DNp63, and further imply different therapeutic strategies could be applied by targeting different molecules depending on the cell differentiation status.

Discussion

In this report, we identified a novel subtype of LUSC characterized by the neural transcription factor Brn2. We examined super-enhancer landscape of LUSC cell lines to reveal the heterogeneity of transcriptional programs in LUSC and identified three subgroups of LUSC based on the super-enhancer profiles. Notably, super-enhancers on the locus of the *SOX2* gene are shared by two subgroups. We previously reported that *SOX2* is the most commonly amplified gene in LUSC and its expression is essential for cell survival [8]. Following our study, it has been confirmed that Sox2 promotes lung squamous cancer lineage in mice model[32,33], supporting its important role as a lineage oncogene. However, in one subgroup of LUSC, *SOX2*

super-enhancers were not found and its expression was low, suggesting that Sox2 does not necessarily govern transcriptional programs among all LUSCs. While this subset is not represented as a significant proportion of primary human LUSC in TCGA cohort, it remains to be seen which lineage state this low Sox2 LUSC subset represents and which transcriptional program may be active. In addition, super-enhancers on the locus of the *TP63* gene, which we also identified as a collaborating lineage factor with Sox2 in LUSC, were only found in one 'classical' subgroup of LUSC, suggesting that in the other subgroups, other factors play roles in controlling their lineage states. In fact, our super-enhancer profiling found that super-enhancers on the loci of *SOX2* and *POU3F2* instead of *TP63* were commonly shared in a small subset of LUSC. Brn2, encoded by *POU3F2*, is a Class III POU transcription factor which play an essential role in neural cell differentiation[34-37] and is highly expressed in glioblastoma, neuroendocrine SCLC and neuroendocrine prostate cancer[21-23]. This subset exhibited gene expression patterns associated with neural development compared to the 'classical' subtype, suggesting the differentiation states of the subset is closer to neural lineages. This subset is not categorized as a neuroendocrine lung cancer which also commonly expresses Brn2 and is close to neural lineages. We speculate that this unique LUSC subset can have a potential to show

neuroendocrine phenotypes possibly with additional factors implicated for neuroendocrine lineages such as Rb1 and Ascl1, which warrants further investigation.

In this study, we highlighted the partnership of Brn2 and Sox2 in this unique 'neural' subset of LUSC for regulating their transcriptional programs. Sox2 has been described to have a role in determining cell differentiation states by cooperating with other lineage factors such as Oct4 in embryonic stem cells[38], Pax6 in eye lens development[39], p63 in squamous lineages[10] and Brn2 in neural progenitor cells[18,19]. Our data support distinct contributions of Sox2 across multiple cell types and confirmed that cooperation of Brn2 and Sox2 in neural progenitor cells is conserved in the 'neural' LUSC cells, which has not been described previously.

Prior studies including ours have reported a role of p63 as a lineage oncogene in defining squamous lineages[10,40,41]. To support and extend those findings, our study highlighted the critical role of p63 antagonizing Brn2 in defining more classical squamous cell states within the context of LUSC. Our findings suggest that switching the Sox2 collaborating partner from Brn2 to p63 contributes a shift in cell differentiation states in LUSC via dramatically changing Sox2 cistrome. This phenomenon also supports a role of Sox2 in determining cell differentiation states by having specific partners is maintained in malignant cells. As this switch has not been

described in normal differentiation or regeneration of lung cells, whether this transition from Brn2 to p63 is unique to LUSC remains to be investigated. Of note, p63 overexpression in the Brn2-positive 'neural' LUSC cells decreased cell growth both *in vitro* and *in vivo*. It suggests that rather than playing a role as a general oncogene, p63 defines more squamous differentiation states and the lineage program of this Brn2-positive 'neural' LUSCs could lead to more aggressive phenotypes as cancer cells compared to that of classical LUSCs defined by p63. Our study raised a hypothesis that a loss of p63 in classical LUSCs can unveil a latent neural state that is represented by the novel Brn2-high subtype. However, our classical LUSC cell lines were dependent on p63 for their survival and p63 knockout did not induce Brn2 expression in a short period of time. Further studies are needed to explore this hypothesis.

We also observed a change of RTK signaling activity profile through a shift in cell differentiation states induced by p63. It has been reported that p63 regulates a various signaling pathway molecules depending on cell contexts such as β-catenin, EGFR and Jagged1 in human airway epithelial basal cells[42], FGFR2 in murine squamous cell carcinoma model[43] and NRG1 in mammary basal cells[44]. In our study, we identified several RTKs whose expression and phosphorylation level are altered after p63 overexpression. While functions of positive ErbB4

signaling in the 'neural' LUSC cells are unclear, the p63-overexpressed LUSC cells showed more responses to ErbB3/EGFR pathway stimulation and high ERK activity even without stimulation. It is possible that the 'neural' LUSC cells are less dependent on AKT and MAPK pathways similarly to SCLC. Future studies are needed to investigate the relationships between lineage programs and networks of signaling pathways and this would help us making therapeutic strategies for each lineage cancer program.

Taken together, we identified a novel 'neural' lineage signified by Brn2 in LUSC by investigating its super-enhancer landscape. Brn2 cooperates with Sox2 in determining its transcriptional program, which is overwhelmed and reprogrammed by p63. Characterization of each cancer lineage to identify its unique vulnerabilities could lead to a novel approach to make therapeutic strategies for this heterogeneous disease.

Methods

Further information can be found in Supplemental Methods.

Cell lines

Lung cancer cell lines (HCC95, KNS62, HCC2814, LK2, NCI-H520, SQ-1, EBC-1, HCC2279, NCI-H2887, HCC2450, Calu-1, HARA, LC-1/SQSF) were maintained in RPMI-1640 (Gibco) with 10% FBS and 1% penicillin–streptomycin (Gibco). HEK293T cells were maintained in DMEM with 10% FBS and 1% penicillin–streptomycin (Gibco). Cultured cells were regularly tested for mycoplasma using the mycoAlert Detection Kit (Lonza).

ChIP-seq

H3K27ac, Sox2 and p63 ChIP were performed as previously described with modifications[10]. Brn2 ChIP was performed using chromatin digested by micrococcal nuclease (NEB, M0247S). DNA libraries for Illumina cluster generation and sequencing with NextSeq500 (Illumina) were performed according to the manufacturer's protocol.

ChIP-seq data analysis

For H3K27ac ChIP, peaks were identified by MACS[45] after aligning to hg19/GRCh37. To define 'super-enhancers', we used the ROSE2 pipeline[46]. Unsupervised hierarchical clustering was performed using signals near transcriptional regulators. Principle component

analysis was performed based on the same signals per region, and top 3 components were depicted in three dimensional plots. Differentially enriched super-enhancers were determined by comparing the 'neural' subgroup against the 'classical' subgroup using 'samr' R package[47].

For Sox2, Brn2 and p63 ChIP, peaks were identified by MACS[45] after aligning to hg19/GRCh37. ChIP signals on Sox2, Brn2 and p63 peaks detected on any samples were visualized as a heatmap generated by 'heatmap' function in Cistrome analysis pipeline (http://cistrome.org/) or plotHeatmap function of deepTools[48].

RNA-seq analyses

RNA-seq gene expression data for *SOX2*, *TP63* and *POU3F2* as well as SNP-array copy number data for *SOX2* were obtained from Cancer Cell Line Encyclopedia (CCLE) (http://www.broadinstitute.org/ccle/home). RNA-seq data of 501 tumor tissues were obtained from TCGA-LUSC dataset[5]. RNA-seq data from the brain hypothalamus region were obtained from GTEx data portal[26]. Two RNA-seq data (EGAD00001001244 and GSE60052) were obtained for SCLC[25,49].

Hierarchical clustering of the LUSC tumors was performed using the differentially expressed genes between *POU3F2*-high/ *TP63*-low and *POU3F2*-low/ *TP63*-high tumors with complete-linkage clustering. Kaplan-Meier curves were plotted to compare overall survival for patients with *POU3F2*-high/ *TP63*-low tumors versus patients with *POU3F2*-low/ *TP63*-high tumors.

To perform RNA-seq analyses for the engineered LK2 cell lines, sequencing reads were aligned to hg19/GRCh37 and read counts per gene symbol were estimated using SAMtools[50]. Then DEseq2[51] was used to identify differentially expressed genes between control and DNp63-overexpressed LK2 cells and to identify genes significantly associated with DNp63 overexpression levels (control -> inducible -> stable).

To identify potentially enriched functions of selected gene sets of interest, we compared these gene sets with the genes annotated by the same Gene Ontology (GO) terms curated in the Molecular Signature Database (MSigDB)[52]. Each of 5917 GO terms included in "C5" collection (version 6.0 downloaded in 2017) was compared with query gene sets.

Immunofluorescence

Cells were grown on glass coverslips coated with poly D-lysine (Neuvitro), fixed in 4% paraformaldehyde for 10 min and permeabilized with 1x PBS containing 0.25% Triton X-100 for 10 min. Primary antibodies were used at the following dilutions: anti-Brn2 at 1:200 (Cell Signaling, #12137), and anti-Sox2 at 1:400 (Abcam, ab171380). Fluorescent signal was detected with secondary antibodies conjugated with Alexa Fluor (ThermoFisher #A-11034, #A-21236) diluted at 1:2000, and coverslips were mounted (ProLong Diamond Antifade Mountant with DAPI: ThermoFisher). Images were obtained with a Leica DM5500 B fluorescence microscope at 20x objective magnification. Images were processed using the Fiji distribution of ImageJ (https://fiji.sc/).

Western blotting and co-immunoprecipitation

For western blotting to profile ErbB family signaling, engineered LK2 cells were serum-starved for four hours, and then incubated with anti-ErbB4 monoclonal antibody P6-1 (a kind gift from the Hideyuki Saya laboratory, Keio University) for 15 minutes, followed by stimulation with recombinant human NRG1-β1 EGF domain (R&D) for 30 minutes.

For co-immunoprecipitation, whole cell lysate of cells was incubated with anti-BRN2 antibody, anti-SOX2 antibody, or normal IgG antibody and Dynabeads Protein G (ThemoFisher). After washing, the beads were boiled in 1x SDS sample buffer and the immunoprecipitates were then processed for immunoblotting.

Lentiviral introduction of genes

POU3F2, *ΔNp63α* or *GFP* open reading frame (ORF) was cloned into pLEX_306 (a gift from David Root, Addgene #41391), pLEX_307 (a gift from David Root, Addgene #41392) or pLIX_403 (a gift from David Root, Addgene #41395). Cells were infected with lentiviral vectors produced in HEK293T cells and then selected with puromycin. Ectopic protein expression was confirmed via immunoblotting.

CRISPR-Cas9 genome editing

sgRNAs were cloned downstream of the human U6 promoter in a lentiviral vector containing eGFP downstream of the human PGK promoter (a kind gift from the Brian Brown laboratory, Icahn School of Medicine at Mount Sinai). Cells were first transduced with the

lentiCas9-Blast (a gift from Feng Zhang, Addgene, # 52962) lentivirus, and selected with blasticidin. Cells were then transduced with the pLenti-GFP-sgRNA lentivirus. sgRNA target sequences are listed in Supplemental Table 4.

Immunohistochemistry

Immunohistochemical analyses were performed for xenograft tumor specimens and human primary LUSC tumor specimens. Total 10 FFPE specimens of human primary LUSCs resected from patients at Icahn School of Medicine at Mount Sinai (ISMMS, USA) were obtained. The sections were incubated with anti-Brn2 antibody, anti-Sox2 antibody or anti-p63 antibody. All slides were counterstained with hematoxylin before being mounting.

Cell proliferation assay

Cell proliferation was assayed in sextuplicate at 24, 72 and 120 h after plating with alamarBlue Cell Viability Reagent (ThermoFisher) according to the manufacturer's protocol.

Cell viability at 72 and 120 h were corrected for the ratio to control cells from the 24-h reading to account for plating unevenness.

Cell cycle analysis

Cells were fixed with 70 % ethanol at least overnight at 4°C. After washing with PBS, cells were incubated in PBS containing 100 μg/ml RNase and 50 μg/ml propidium iodide overnight at 4°C overnight. DNA content was analyzed by FACS Canto II (BD Bioscience), and quantitative analyses for the proportions of cells in cell cycle were performed using FlowJo software. Differences between two groups were examined using two-tailed *t*-test. Bonferroni correction was performed for multiple comparisons.

Xenograft model

Engineered LK2 cells (4 x 10^6 cells) were injected with a 1:1 mixture of 50 μl cell suspension and 50 μl Matrigel (Corning) subcutaneously into both flank regions of 4- to 6-week-old male NOD-scid or NOD-scid gamma mice (Jackson Laboratory). Tumor volume (length x width^2 /2) was measured twice a week. Before tumor size reached at 1000 mm^3, mice

were sacrificed and tumors were immersed in formalin for immunohistological analysis. Difference of the tumor size was examined using two-tailed *t*-test. All animal procedures and studies were approved by the Mount Sinai IACUC (protocol number, IACUC-2018-0021).

Phospho-RTK array analysis

The Human Phospho-RTK Array Kit (R&D Systems) was used to determine the relative levels of tyrosine phosphorylation of 42 distinct RTKs, according to the manufacturer's protocol. 300 μg of total protein was used for each membrane. Chemiluminescent signals were captured with a Chemidoc MP Imaging System (Bio-Rad laboratories) and images were processed using Image Lab software (Bio-Rad laboratories).

References

1. Bray, F. *et al*. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **68**, 394-424 (2018).

2. Siegel, R.L., Miller, K.D. & Jemal, A. Cancer statistics, 2019. *CA Cancer J Clin* **69**, 7-34 (2019).

3. Cheng, T.Y. *et al*. The International Epidemiology of Lung Cancer: Latest Trends, Disparities, and Tumor Characteristics. *J Thorac Oncol* **11**, 1653-71 (2016).

4. Islami, F., Torre, L.A. & Jemal, A. Global trends of lung cancer mortality and smoking prevalence. *Transl Lung Cancer Res* **4**, 327-38 (2015).

5. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519-25 (2012).

6. Brahmer, J. *et al*. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. *N Engl J Med* **373**, 123-35 (2015).

7. Herbst, R.S. *et al*. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet* **387**, 1540-50 (2016).

8. Bass, A.J. *et al*. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat Genet* **41**, 1238-42 (2009).

9. Que, J. *et al*. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. *Development* **134**, 2521-31 (2007).

10. Watanabe, H. *et al*. SOX2 and p63 colocalize at genetic loci in squamous cell carcinomas. *J Clin Invest* **124**, 1636-45 (2014).

11. Wilkerson, M.D. *et al*. Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types. *Clin Cancer Res* **16**, 4864-75 (2010).

12. Wu, D. *et al*. Gene-expression data integration to squamous cell lung cancer subtypes reveals drug sensitivity. *Br J Cancer* **109**, 1599-608 (2013).

13. Hnisz, D. *et al*. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-47 (2013).

14. Pott, S. & Lieb, J.D. What are super-enhancers? *Nat Genet* **47**, 8-12 (2015).

15. Boeva, V. *et al*. Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. *Nat Genet* **49**, 1408-1413 (2017).

16. McKeown, M.R. *et al*. Superenhancer Analysis Defines Novel Epigenomic Subtypes of Non-APL AML, Including an RARalpha Dependency Targetable by SY-1425, a Potent and Selective RARalpha Agonist. *Cancer Discov* **7**, 1136-1153 (2017).

17. Schonemann, M.D. *et al*. Development and survival of the endocrine hypothalamus and posterior pituitary gland requires the neuronal POU domain factor Brn-2. *Genes Dev* **9**, 3122-35 (1995).

18. Tanaka, S. *et al*. Interplay of SOX and POU factors in regulation of the Nestin gene in neural primordial cells. *Mol Cell Biol* **24**, 8834-46 (2004).

19. Lodato, M.A. *et al*. SOX2 co-occupies distal enhancer elements with distinct POU factors in ESCs and NPCs to specify cell state. *PLoS Genet* **9**, e1003288 (2013).

20. Ishii, J. *et al*. POU domain transcription factor BRN2 is crucial for expression of ASCL1, ND1 and neuroendocrine marker molecules and cell growth in small cell lung cancer. *Pathology International* **63**, 158-168 (2013).

21. Ishii, J. *et al*. Class III/IV POU transcription factors expressed in small cell lung cancer cells are involved in proneural/neuroendocrine differentiation. *Pathology International* **64**, 415-422 (2014).

22. Bishop, J.L. *et al*. The Master Neural Transcription Factor BRN2 Is an Androgen Receptor-Suppressed Driver of Neuroendocrine Differentiation in Prostate Cancer. *Cancer Discovery* **7**, 54-71 (2017).

23. Schreiber, E. *et al*. Astrocytes and glioblastoma cells express novel octamer-DNA binding proteins distinct from the ubiquitous Oct-1 and B cell type Oct-2 proteins. *Nucleic Acids Res* **18**, 5495-503 (1990).

24. Suva, M.L. *et al*. Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. *Cell* **157**, 580-94 (2014).

25. George, J. *et al*. Comprehensive genomic profiles of small cell lung cancer. *Nature* **524**, 47-53 (2015).

26. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-5 (2013).

27. Truong, A.B. & Khavari, P.A. Control of keratinocyte proliferation and differentiation by p63. *Cell Cycle* **6**, 295-9 (2007).

28. Arnold, K. *et al*. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. *Cell Stem Cell* **9**, 317-29 (2011).

29. Ferguson, K.L. & Slack, R.S. The Rb pathway in neurogenesis. *Neuroreport* **12**, A55-62 (2001).

30. Gobert, R.P. *et al*. Convergent functional genomics of oligodendrocyte differentiation identifies multiple autoinhibitory signaling circuits. *Mol Cell Biol* **29**, 1538-53 (2009).

31. Magri, L. *et al*. E2F1 coregulates cell cycle genes and chromatin components during the transition of oligodendrocyte progenitors from proliferation to differentiation. *J Neurosci* **34**, 1481-93 (2014).

32. Ferone, G. *et al*. SOX2 Is the Determining Oncogenic Switch in Promoting Lung Squamous Cell Carcinoma from Different Cells of Origin. *Cancer Cell* **30**, 519-532 (2016).

33. Tata, P.R. *et al*. Developmental History Provides a Roadmap for the Emergence of Tumor Plasticity. *Dev Cell* **44**, 679-693 e5 (2018).

34. Sugitani, Y. *et al*. Brn-1 and Brn-2 share crucial roles in the production and positioning of mouse neocortical neurons. *Genes Dev* **16**, 1760-5 (2002).

35. Jin, Z. *et al*. Different transcription factors regulate nestin gene expression during P19 cell neural differentiation and central nervous system development. *J Biol Chem* **284**, 8160-73 (2009).

36. Iwafuchi-Doi, M. *et al*. The Pou5f1/Pou3f-dependent but SoxB-independent regulation of conserved enhancer N2 initiates Sox2 expression during epiblast to neural plate stages in vertebrates. *Dev Biol* **352**, 354-66 (2011).

37. Dominguez, M.H., Ayoub, A.E. & Rakic, P. POU-III transcription factors (Brn1, Brn2, and Oct6) influence neurogenesis, molecular identity, and migratory destination of upper-layer cells of the cerebral cortex. *Cereb Cortex* **23**, 2632-43 (2013).

38. Boyer, L.A. *et al*. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-56 (2005).

39. Kamachi, Y., Uchikawa, M., Tanouchi, A., Sekido, R. & Kondoh, H. Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. *Genes Dev* **15**, 1272-86 (2001).

40. Devos, M. *et al*. Elevated DeltaNp63alpha Levels Facilitate Epidermal and Biliary Oncogenic Transformation. *J Invest Dermatol* **137**, 494-505 (2017).

41. Somerville, T.D.D. *et al*. TP63-Mediated Enhancer Reprogramming Drives the Squamous Subtype of Pancreatic Ductal Adenocarcinoma. *Cell Rep* **25**, 1741-1755 e7 (2018).

42. Warner, S.M. *et al*. Transcription factor p63 regulates key genes and wound repair in human airway epithelial basal cells. *Am J Respir Cell Mol Biol* **49**, 978-88 (2013).

43. Ramsey, M.R. *et al*. FGFR2 signaling underlies p63 oncogenic function in squamous cell carcinoma. *J Clin Invest* **123**, 3525-38 (2013).

44. Forster, N. *et al*. Basal cell signaling by p63 controls luminal progenitor function and lactation via NRG1. *Dev Cell* **28**, 147-60 (2014).

45. Zhang, Y. *et al*. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

46. Lin, C.Y. *et al*. Active medulloblastoma enhancers reveal subgroup-specific cellular origins. *Nature* **530**, 57-62 (2016).

47. Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**, 5116-21 (2001).

48. Liu, T. *et al*. Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol* **12**, R83 (2011).

49. Jiang, L. *et al*. Genomic Landscape Survey Identifies SRSF1 as a Key Oncodriver in Small Cell Lung Cancer. *PLoS Genet* **12**, e1005895 (2016).

50. Li, H. *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-9 (2009).

51. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

52. Subramanian, A. *et al*. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-50 (2005).

Acknowledgements

We thank Chiara Vardabasso, Dan Hasson, Katsutoshi Sato, Aleksandra Wroblewska, Almuneda Bosch and Iman Tavassoly for helpful discussions; Aleksandra Wroblewska and Brian D. Brown for providing a lentiviral vector for cloning sgRNAs; NextSeq Sequencing Facility of the Department of Oncological Sciences at Icahn School of Medicine at Mount Sinai (ISMMS), Saboor Hekmaty, Irene Salib, Gayatri Panda and Ravi Sachidanandam for sequencing assistance; and Stephanie Tuminello and Michael William for technical assistance. The authors also thank the Flow Cytometry Core facility and the Biorepository and Pathology Core Facility at ISMMS. This work was supported in part through Tisch Cancer Institute at ISMMS and the computational resources and staff expertise provided by Scientific Computing at ISMMS. This study was supported by grants from the 2017 ATS Foundation Unrestricted Grant: Pulmonary (H. Watanabe) and the American Lung Association of the Northeast Lung Cancer Discovery Award (H. Watanabe). T. Sato is supported by the Japanese Respiratory Society the 6th Lilly Fellowship Program and the Uehara Memorial Foundation. R. Kong is supported by Shaanxi Provincial Natural Science Foundation, China (No. 2017JM8046).

Author contributions

T.S. and H.W. conceived and designed the study. T.S., S.Y., R.K., M.F., A.S., P.C., A.P., and M.B.B. conducted experiments and/or acquired data. T.S., S.Y., M.B.B., and H.W. analyzed data. O.N., T.M., M.B.B., and C.A.P. provided administrative, technical and/or material support. C.A.P., J.Z., and H.W. supervised the study. T.S., S.Y., and H.W. wrote, reviewed, and revised the manuscript. All authors reviewed and commented on the manuscript.

Competing interests: The authors declare no competing interests.

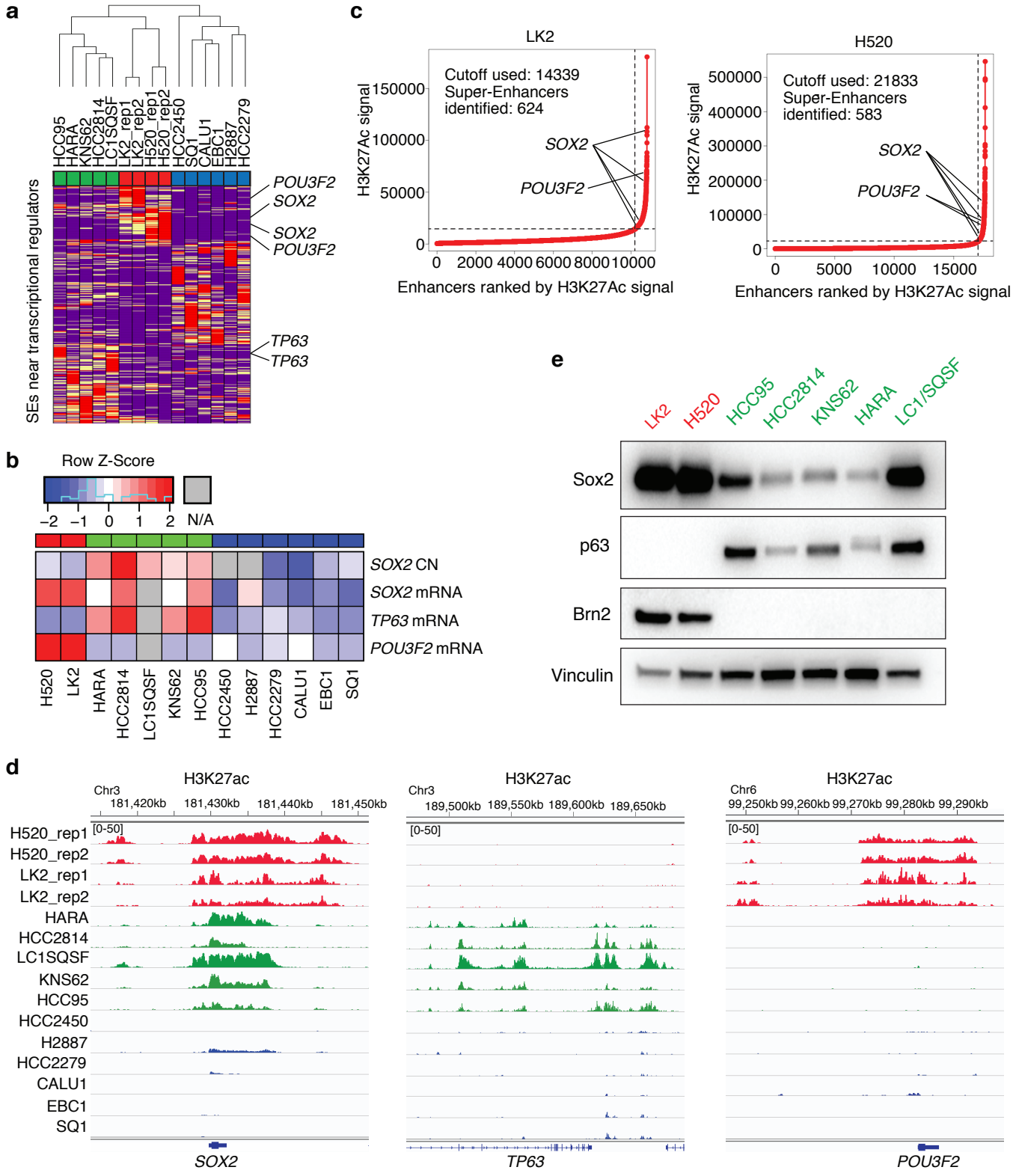

Fig. 1 Super-enhancer profiling identifies a novel subtype of LUSC. **a** Unsupervised hierarchical clustering of 13 LUSC cell lines using super-enhancer scores near transcriptional regulator genes. **b** Copy number of *SOX2* and mRNA expression of *SOX2*, *TP63* and *POU3F2* in LUSC cell lines from CCLE. **c** Super-enhancer plots using H3K27ac scores in the small subset of LUSC cell lines; LK2 cells (*left*) and NCI-H520 cells (*right*). **d** Genome view tracks of H3K27ac signal at loci of *SOX2* (*left*), *TP63* (*middle*) and *POU3F2* (*right*) in LUSC cell lines. **e** Protein expression of Brn2, p63, Sox2 and vinculin as a loading control in LUSC cell lines.

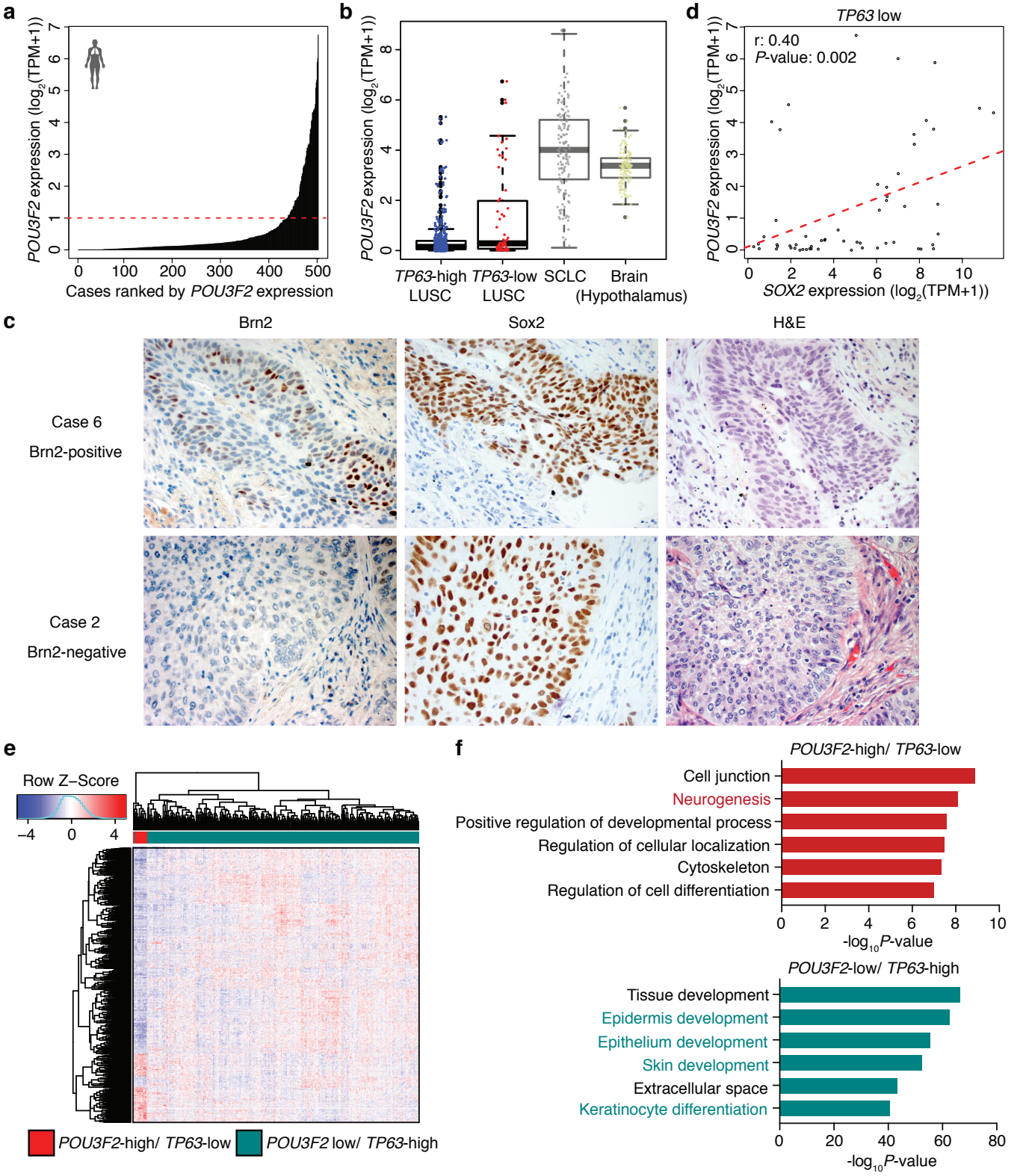

Fig. 2 Brn2 signifies the neural subtype of LUSC in primary LUSC tumors. **a** Expression of *POU3F2* in 501 TCGA LUSC tumor tissues. The red dashed line (TPM=1) shows the cutoff to separate samples into *POU3F2* high and low samples. **b** Box plots of *POU3F2* expression in *TP63*-high and low LUSC tumors from TCGA, SCLC tumors from publicly available datasets and normal brain tissues (hypothalamus) from GTEx. **c** Immunohistochemical staining of Brn2 and Sox2 and H&E staining in Brn2-positive and negative human LUSC tumors. Representative images are shown (original images, ×400). **d** Scatter plots of expression of *SOX2* and *POU3F2* in *TP63*-low LUSC tumors from TCGA. **e** Heatmap showing hierarchical clustering of LUSC tumors from TCGA using differentially expressed genes between *POU3F2*-high/ *TP63*-low (n=20) and *POU3F2*-low/ *TP63*-high tumors (n=396). With cutoffs of fold change>2 and FDR<0.01, 196 genes are differentially up-regulated and 735 genes are differentially down-regulated in the *POU3F2*-high/ *TP63*-low tumors. Color scheme represents Z-score distribution. **f** Gene ontology analyses for the differentially up-regulated (*top*) and down-regulated (*bottom*) genes. Enriched functions for these genes are identified based on Fisher's exact test against GO terms curated in MSigDB.

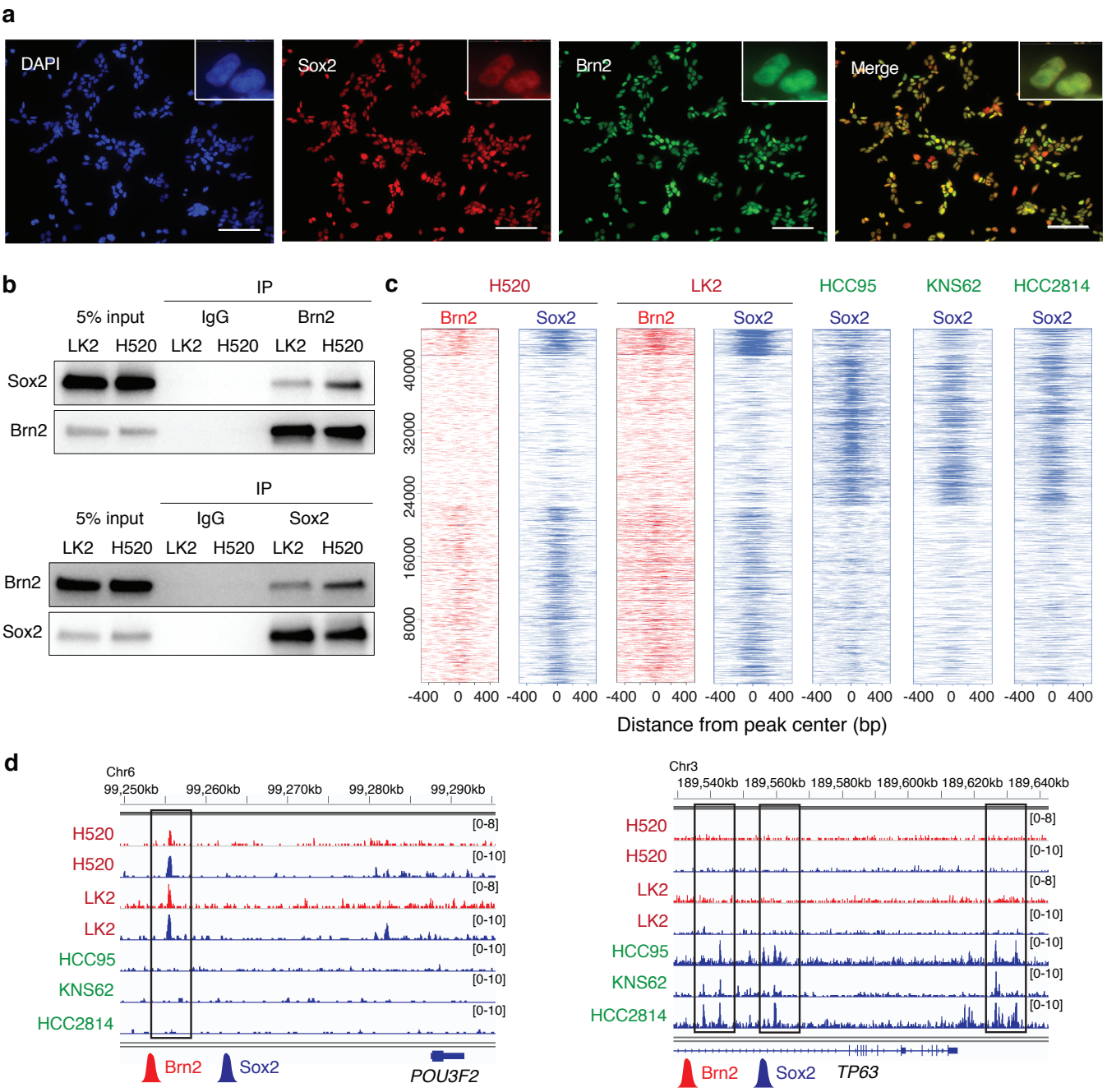

Fig. 3 Brn2 and Sox2 interact and co-localize at genetic loci in the 'neural' subset of LUSC. **a** Expression of endogenous Brn2 and Sox2 in LK2 cells, determined by immunofluorescence with anti-Sox2 (*green*) and anti-Brn2 (*red*) antibodies, respectively. DAPI staining (nuclei; *blue*) and merged images are also shown. Original magnification, ×200. Scale bar, 100 μm. **b** Sox2-Brn2 interaction, shown by co-immunoprecipitation of Sox2 using an antibody against endogenous Brn2 (*top*) and co-immunoprecipitation of Brn2 using an antibody against endogenous Sox2 (*bottom*) in LK2 and NCI-H520 cells. **c** Heatmap depicting global analysis of ChIP-seq signals for Brn2 and Sox2 in 'neural' NCI-H520 and LK2 cells and those for Sox2 in 'classical' HCC95, KNS62 and HCC2814 cells at all the peak loci. ChIP-seq signal intensity is shown by color shading. **d** Genome view tracks of Brn2 and Sox2 ChIP-seq signals in NCI-H520 and LK2 cells and Sox2 ChIP signals in HCC95, KNS62 and HCC2814 cells at loci of *POU3F2* (*left*) and *TP63* (*right*).

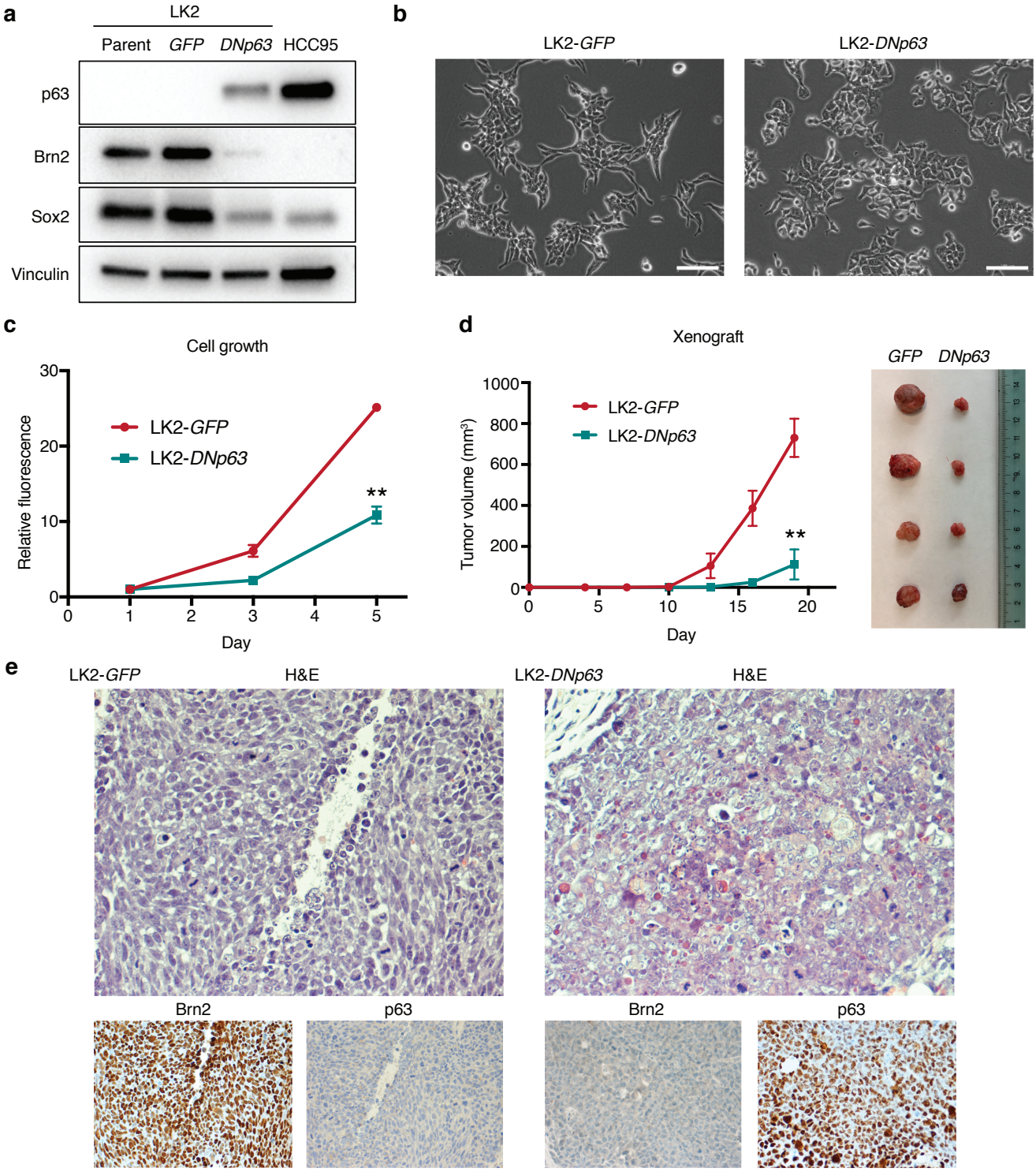

Fig. 4 DNp63 overexpression in the 'neural' LUSC cells suppresses Brn2 expression and induces phenotypic changes. **a** Protein expression of p63, Brn2, Sox2 and vinculin as a loading control in parental, GFP-overexpressed and DNp63-overexpressed LK2 cells. **b** Phase-contrast microphotographs of GFP-overexpressed and DNp63-overexpressed LK2 cells. Bar = 100 μm. **c** Cell growth of GFP-overexpressed and DNp63-overexpressed LK2 cells. Mean ± SD of sextuplicates are shown. **, $P<0.001$ vs. GFP-overexpressed LK2 cells, *t*-test. **d** Tumor growth of GFP-overexpressed and DNp63-overexpressed LK2 cells *in vivo*. Mean ± SD of tetraplicates are shown. **, $P<0.001$ vs. GFP-overexpressed xenografts, *t*-test. Xenograft tumors resected 19 days after inoculation are shown in the right picture. **e** H&E staining and immunohistochemical staining of Brn2 and p63 in GFP-overexpressed or DNp63-overexpressed LK2 xenograft. Original Images, ×400.

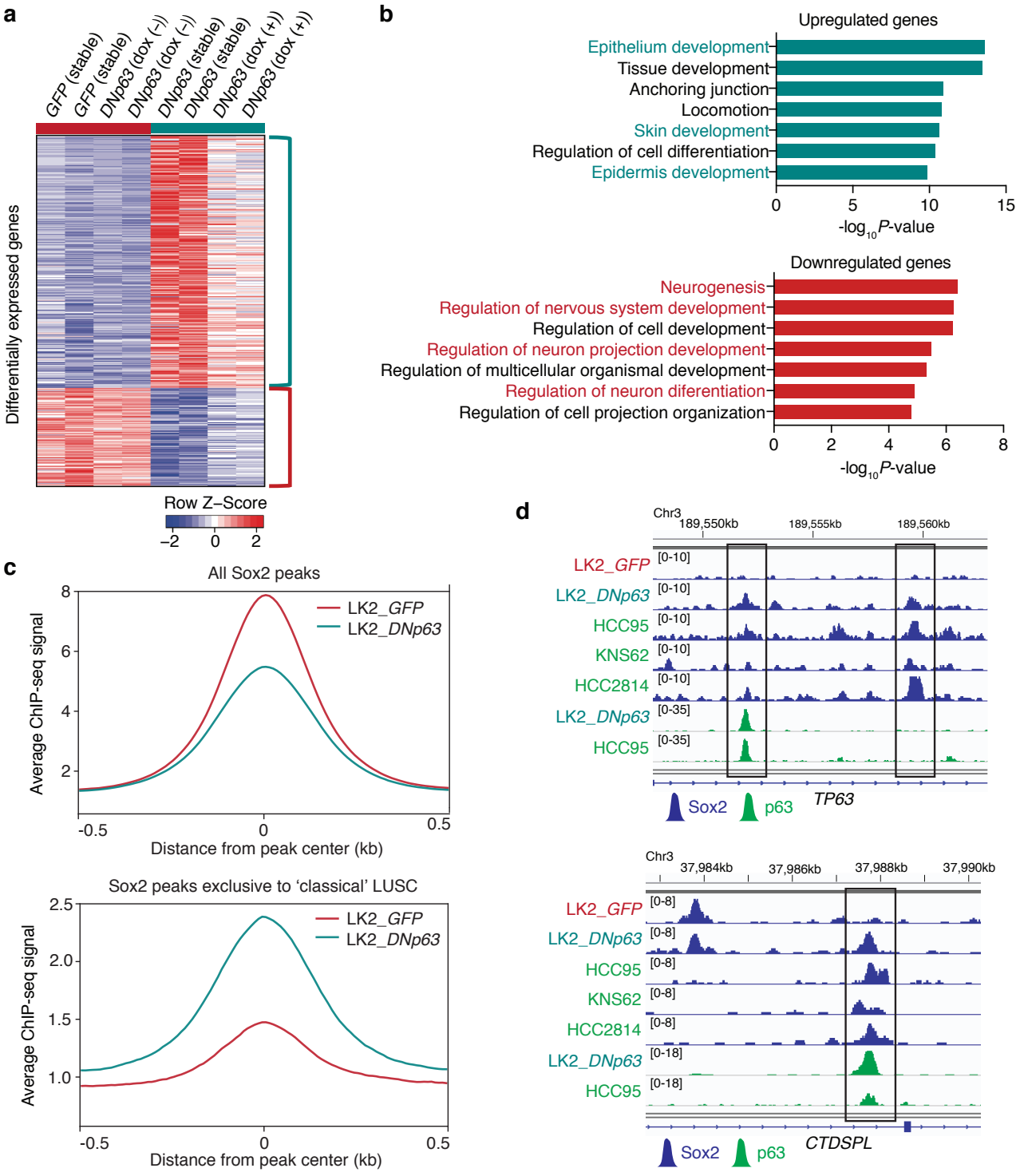

Fig. 5 DNp63 induces a classical squamous-cell transcriptional program in the 'neural' LUSC cells. **a** Heatmap showing 399 differentially expressed genes (287 upregulated and 112 downregulated genes) between control LK2 cells (stable GFP-overexpressed cells and doxycycline-inducible DNp63-overexpressing cells without doxycycline in duplicate) and DNp63-overexpressed LK2 cells (stable DNp63-overexpressed cells and doxycycline-inducible DNp63-overexpressing cells with 2 μg/ml doxycycline in duplicate), sorted by fold change. Color scheme represents Z-score distribution. **b** Gene ontology analyses for the differentially up-regulated (*top*) and down-regulated (*bottom*) genes. Enriched functions for these genes are identified based on Fisher's exact test against GO terms curated in MSigDB. **c** Average ChIP-seq signals for all Sox2 peaks (*top*) and Sox2 peaks exclusively found in 'classical' HCC95, KNS62 and HCC2814 cells (*bottom*) in GFP-overexpressed and DNp63-overexpressed LK2 cells. **d** Genome view tracks of Sox2 ChIP-seq signals in GFP-overexpressed and DNp63-overexpressed LK2 cells, HCC95, KNS62 and HCC2814 cells and p63 ChIP signals in DNp63-overexpressed LK2 cells and HCC95 cells at loci of *TP63* (*top*) and *CTDSPL* (*bottom*).

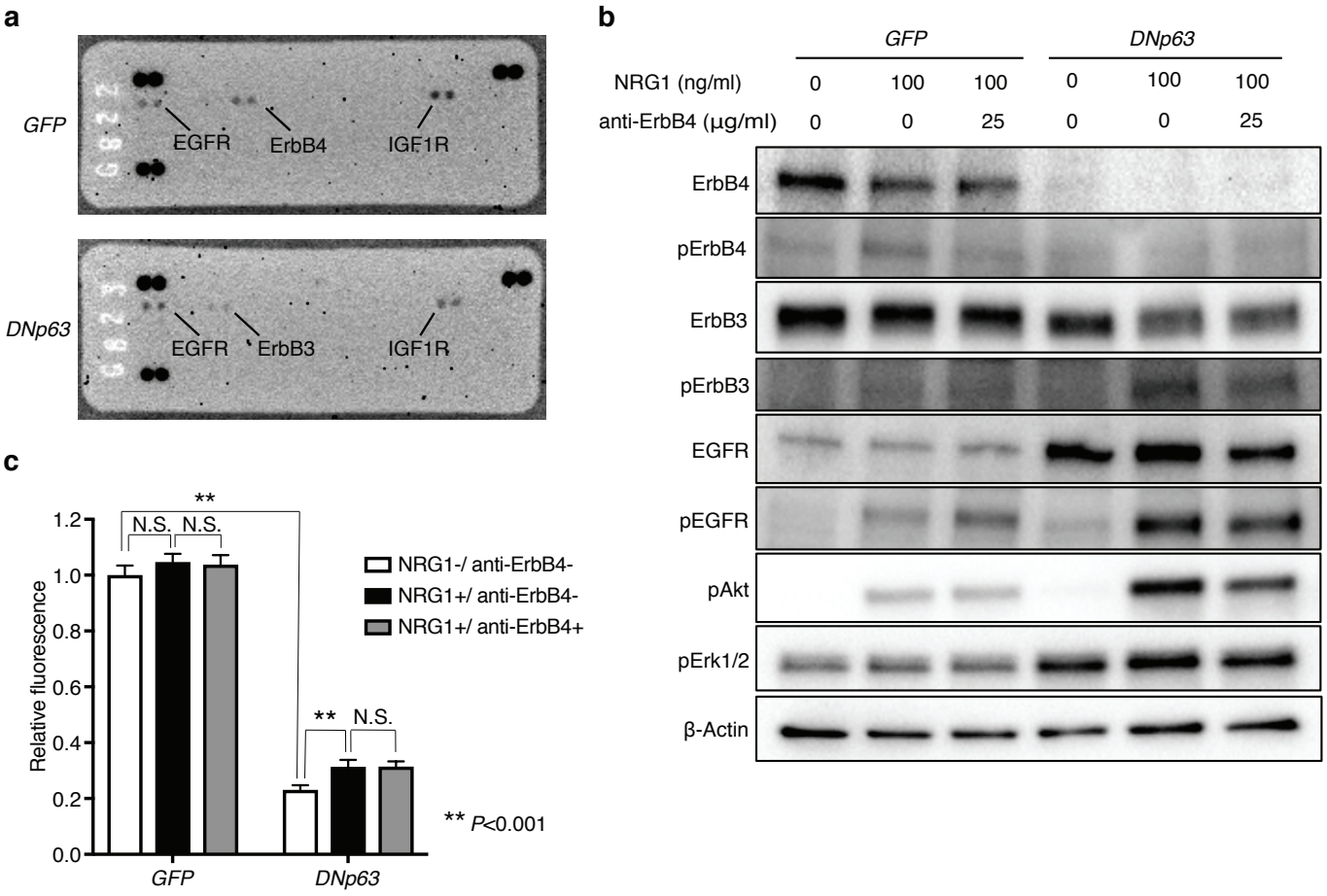

Fig. 6 DNp63 alters ErbB family signaling profile in the 'neural' LK2 cells. **a** Human phospho-receptor tyrosine kinase array was performed for GFP-overexpressed and DNp63-overexpressed LK2 cells. **b** Protein expression of ErbB4, phospho-ErbB4, ErbB3, phospho-rbB3, EGFR, phospho-EGFR, phospho-Akt and phospho-Erk1/2 and β-Actin as a loading control in GFP-overexpressed and DNp63-overexpressed LK2 cells. Cells were serum-starved for 4 hours, and then incubated with anti-ErbB4 monoclonal antibody at the indicated concentrations for 15 minutes, followed by stimulation with NRG1 for 30 minutes at the indicated concentrations. **c** Cell proliferation of GFP-overexpressed and DNp63-overexpressed LK2 cells. Cells were cultured with 5% FBS and with or without NRG-1 (100 ng/ml) and anti-ErbB4 monoclonal antibody (25 μg/ml) for 4 days. Mean ± SD of sextuplicates are shown. **, $P<0.001$ and N.S., not significant ($P>0.05$), *t*-test with Bonferroni correction..