

SOX2 and p63 colocalize at genetic loci in squamous cell carcinomas

Hideo Watanabe,^{1,2,3} Qiuping Ma,¹ Shouyong Peng,^{1,2} Guillaume Adelmant,^{4,5} Danielle Swain,¹ Wenyu Song,¹ Cameron Fox,¹ Joshua M. Francis,^{1,2,3} Chandra Sekhar Pedamallu,^{1,2} David S. DeLuca,² Angela N. Brooks,^{1,2} Su Wang,⁶ Jianwen Que,⁷ Anil K. Rustgi,⁸ Kwok-kin Wong,^{1,9} Keith L. Ligon,^{1,10} X. Shirley Liu,^{1,6,11} Jarrod A. Marto,^{4,5} Matthew Meyerson,^{1,2,3,10} and Adam J. Bass^{1,2,3,9}

Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. ²Cancer Program, The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. ³Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Boston, Massachusetts, USA.
 Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA.
 Department of Cancer Biology and Blais Proteomics Center, Dana-Farber Cancer Institute, Boston, Massachusetts, USA.
 Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, Massachusetts, USA.
 Department of Biomedical Genetics and Stem Cell and Regenerative Medicine Institute, University of Rochester, Rochester, New York, USA.
 Division of Gastroenterology, Department of Medicine, Abramson Cancer Center, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA.
 Department of Medicine and ¹⁰Department of Pathology, Harvard Medical School, Boston, Massachusetts, USA.
 Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, Massachusetts, USA.

The transcription factor SOX2 is an essential regulator of pluripotent stem cells and promotes development and maintenance of squamous epithelia. We previously reported that SOX2 is an oncogene and subject to highly recurrent genomic amplification in squamous cell carcinomas (SCCs). Here, we have further characterized the function of SOX2 in SCC. Using ChIP-seq analysis, we compared SOX2-regulated gene profiles in multiple SCC cell lines to ES cell profiles and determined that SOX2 binds to distinct genomic loci in SCCs. In SCCs, SOX2 preferentially interacts with the transcription factor p63, as opposed to the transcription factor OCT4, which is the preferred SOX2 binding partner in ES cells. SOX2 and p63 exhibited overlapping genomic occupancy at a large number of loci in SCCs; however, coordinate binding of SOX2 and p63 was absent in ES cells. We further demonstrated that SOX2 and p63 jointly regulate gene expression, including the oncogene ETV4, which was essential for SOX2-amplified SCC cell survival. Together, these findings demonstrate that the action of SOX2 in SCC differs substantially from its role in pluripotency. The identification of the SCC-associated interaction between SOX2 and p63 will enable deeper characterization the downstream targets of this interaction in SCC and normal squamous epithelial physiology.

Introduction

More than 1 million people worldwide die each year from squamous cell carcinomas (SCCs), and there is a paucity of targeted therapies for these diseases (1). While these tumors emerge from various epithelial tissues, certain features are commonly seen across these tumors regardless of origin, including highly recurrent amplifications involving chromosome 3q. We previously investigated 3q amplifications in lung and esophageal SCCs and found in both variants that the focus of amplification lies at the locus of the developmental transcription factor SOX2, which we further demonstrated to serve as an essential SCC oncogene (2). SOX2 amplification and oncogenicity have since been reported in a spectrum of SCCs (3-8) and, more recently, small-cell lung cancer (9). The most comprehensive SCC genomic characterization effort to date, The Cancer Genome Atlas (TCGA) lung SCC study, identified high-level amplification/overexpression of SOX2 in 21% of tumors, the third most frequent genomic alteration after inactivation of TP53 and CDKN2A (10).

Despite the strong genomic evidence, the functional rationale for recurrent *SOX2* amplifications in SCCs has not been estab-

Authorship note: Hideo Watanabe and Qiuping Ma contributed equally to this

Conflict of interest: The authors have declared that no conflict of interest exists. **Citation for this article:** *J Clin Invest.* doi:10.1172/JCI71545.

lished. SOX2 is largely studied in the context of pluripotency, as it is essential for ES cells and is able to cooperatively induce differentiated cells to become pluripotent stem cells (11). Because of this role in pluripotency, overexpression of SOX2 has been widely speculated to contribute to carcinogenesis by imparting upon cells stem-like properties, thus leading to the development of cancers characterized by aggressive clinical behavior and poor differentiation status (12–14). Indeed, we reported an expression signature of "ES cell-like" to be enriched in lung SCCs with higher SOX2 expression signature (2).

However, the hypothesis that oncogenic roles of SOX2 recapitulate its actions in pluripotency would not explain the preferential amplification of *SOX2* in SCCs as opposed to adenocarcinomas (15). The predilection for *SOX2* amplifications in SCC suggests that its contribution to SCC may reflect activities specific to the squamous epithelial lineage. Indeed, SOX2 has been recently noted to play essential roles in the development of squamous epithelial lineage and, in the adult, to mark precursor populations of both the esophagus and the large airways (16, 17). Therefore, it is plausible that SOX2's actions in SCC reflect this lineage-specific program.

While it may appear paradoxical that SOX2 is essential for pluripotency, yet also regulates the development and maintenance of a specific developmental lineage, these distinct SOX2 actions may follow its ability to act jointly with distinct cofactors. SOX2

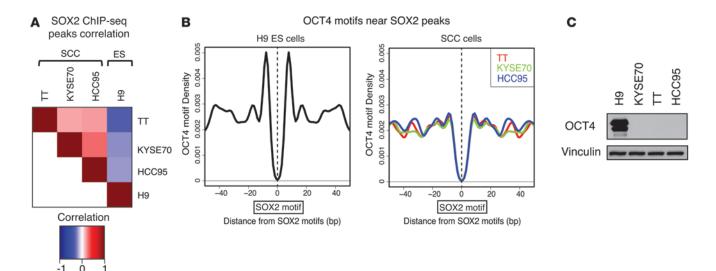


Figure 1
SOX2 genomic occupancy in SCC cells is distinct from that in ES cells. (A) Correlation matrix depicting pairwise comparisons of identified SOX2 binding peaks in the 3 SOX2-amplified SCC lines and in the ES cell line H9. Color scale represents degree of correlation (red, positive; blue, inverse). (B) Appearance of OCT4 DNA binding motif, plotted around the SOX2 motif near the summit of SOX2 binding peaks in the H9 ES cell line and in the 3 SCC lines. The OCT4 motif was highly enriched within 10 bp from the SOX2 motif in SOX2 peaks in H9 cells, but not in the SCC lines. (C) OCT4 protein expression in H9 ES cells and the 3 SCC lines. OCT4 expression was not detectable in any of the SCC lines by immunoblots.

belongs to a family of factors that largely bind to DNA as a heterodimer, typically with other transcription factors (18). Distinct SOX2 heterodimeric partners have been found in different lineages, such as SOX2-OCT4 pairing in ES cells (19, 20) and SOX2-BRN2 pairing in the neural lineage (21). However, unique SOX2 dimerization partners or protein complexes have not been identified in normal squamous epithelia or in SCC. We hypothesized that evaluating the genome-wide occupancy profile of SOX2 in SCCs compared with ES cells would enable us to identify the extent to which the actions of SOX2 in SCC recapitulates its roles in pluripotency. Furthermore, to the extent that SOX2's genomic localization differs from ES cells, we hypothesized that identification of novel SOX2 collaborating transcription factors in SCC may allow us to begin to characterize its mechanisms of action in these deadly cancers.

Results

Genomic occupancy of SOX2 in SCC cells is distinct from that in ES cells. To compare SOX2's genomic occupancy in SOX2-amplified SCCs and in ES cells, we performed ChIP-seq using an antibody against endogenous SOX2 in the esophageal SCC cell lines KYSE70 and TT and in the lung SCC line HCC95, all of which had genomic amplification at the SOX2 locus (Supplemental Figure 1A; supplemental material available online with this article; doi:10.1172/ JCI71545DS1), as well as in H9 human ES cells, in which SOX2 dimerizes with OCT4. Peaks of SOX2 binding were identified in each sample relative to input DNA using MACS algorithm (22). We confirmed strong enrichment of the presence of consensus SOX2 binding motifs in both SCC and ES cells (Supplemental Figure 1B) as well as a high degree of overlap (40.2%) between previously reported SOX2 occupancy in H1 ES cells (23) and our data from H9 ES cells (Supplemental Figure 1C). We then compared SOX2 binding peaks pairwise in these cells and found they were more similar across all 3 SCC cell lines - even between lines of lung and esophageal origin — than between the ES cells and any SCC cell line (Figure 1A). In addition, we found that overlaps of SOX2 occupancy in these 3 SCC cell lines with the published SOX2 occupancy in H1 ES cells were much less (4.9%–9.1%) than the H9 ES cells' overlap with the H1 line (Supplemental Figure 1C).

It has been hypothesized that SOX2 and OCT4 may collaborate in cancers, where SOX2 acts as an oncogene (24). Therefore, we evaluated the presence of OCT4 motifs around the SOX2-occupied regions in the data from SCC and ES cells. As expected, we observed strong enrichment of OCT4 motifs immediately adjacent to the SOX2 motif near the peak center in ES cells (Figure 1B). Conversely, OCT4 motifs were not enriched at SOX2-occupied loci of the SCC cell lines (Figure 1B), which suggests that one or more different SOX2 partners exist in SCC. In fact, we found no detectable expression of OCT4 protein in the SCC lines (Figure 1C), which suggests that factors other than OCT4 act with SOX2 in SCC.

SOX2 interactome in chromatin of SCC cells includes p63 protein. To identify potential novel SOX2-interacting proteins in SCC, we performed tandem affinity purification (TAP) followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) in the 3 SOX2-amplified SCC cell lines uniformly expressing FLAG-HAtagged SOX2 (FLAG-HA-SOX2) at near-physiological levels (Figure 2, A and B, and Supplemental Figure 2, B and C). To identify SOX2's interacting partners in the context of DNA binding, we performed TAP from the chromatin fraction solubilized by micrococcal nuclease digestion on the insoluble fraction of isolated nuclei (Figure 2A, Supplemental Figure 2C, and Supplemental Table 1). We identified a subset of 45 SOX2-interacting proteins common to all 3 cell lines (Supplemental Figure 2D), 19 of which were previously identified as pluripotency-associated factors (25-27). We also found that SOX2 in SCCs was associated with members of the NuRD complex similar to protein complexes purified from neural stem cells (Figure 2A and ref. 28). These observations suggest that certain SOX2 functions may be similar across cell lineages.



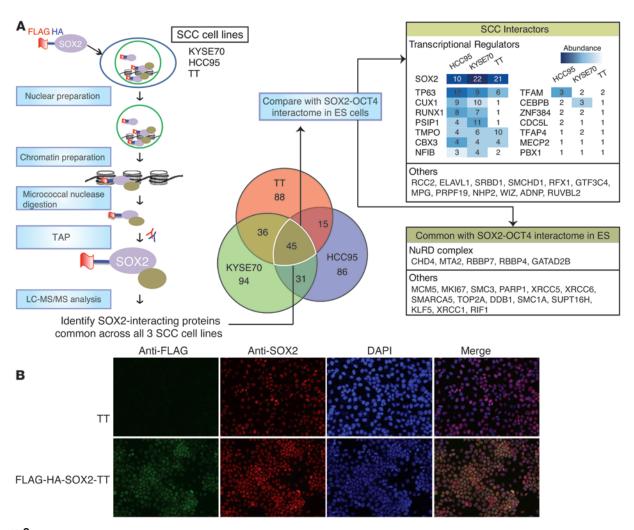


Figure 2
Identification of SOX2 interacting proteins in SCCs. (A) Strategy used to identify binding partners of SOX2 in SCC. Stable expression of FLAG-HA-SOX2 in SCCs was achieved by retroviral transduction. Chromatin prepared from the nuclear fraction was solubilized using micrococcal nuclease digestion. SOX2 complexes were purified from solubilized chromatin by TAP and identified by LC-MS/MS (see Supplemental Table 1 for the complete list). The Venn diagram shows the number of proteins identified in each of the 3 SCC cell lines and highlights the subset of 45 SOX2 binding partners found across all 3 cell lines. Proteins in this common subset were further classified based on their association with regulators of stem cell pluripotency. Among the 26 SCC-specific SOX2 partners, 14 transcriptional regulators were identified. The number of peptides for each of those factors, and their relative abundance with each purification, is indicated for each cell line. (B) Expression of ectopic FLAG-HA-SOX2 and endogenous SOX2 in parental and FLAG-HA-SOX2—expressing TT cells, determined by immunofluorescence with anti-FLAG (green) and anti-SOX2 (red) antibodies, respectively. DAPI staining (nuclei; blue) and merged images are also shown. Original magnification, ×200.

Conversely, we reasoned that the novel subset of 26 SOX2 interactors, including 14 transcriptional regulators unique to SCCs (Figure 2A and Supplemental Figure 2D), may dictate lineage-specific functions. Among them, p63 (encoded by TP63) was estimated to be the most abundant transcription factor associated with SOX2 (Figure 2A). Moreover, inspecting data from the lung SCC TCGA project, we found that of the 14 transcriptional regulators, mRNA expression of TP63 showed the most significant positive correlation with SOX2 expression (r = 0.629; Figure 3A and Supplemental Table 2). In fact, the TP63 locus is frequently coamplified with SOX2 (2) and was coamplified in the 3 SCC cell lines used in this study (Supplemental Figure 1A). Expression of both SOX2 and TP63 correlated with copy number status in lung TCGA dataset (Supplemental Figure 3A). p63 is well recognized as a master squamous regulator

and is the primary marker used to clinically diagnose SCCs (29). In normal squamous epithelia, SOX2 and p63 are coexpressed in the proliferative basal cell layer (6). In the large airways, SOX2 and p63 both mark a putative stem cell in normal physiology (30, 31). Based on the frequent coexpression of SOX2 and p63 and the biologic plausibility of this putative interaction, we pursued this candidate interaction in functional and mechanistic studies.

SOX2 interacts with $\Delta Np63\alpha$, and both are essential for SCC growth. The predominant p63 isoform expressed in the squamous basal layer lacks the full aminoterminal domain and is referred to as $\Delta Np63\alpha$. $\Delta Np63\alpha$ may help maintain the proliferative potential of basal cells and is necessary for epithelial stratification (32, 33). As in basal cells, we found that the predominant isoform of p63 in the 3 SCC cell lines was $\Delta Np63\alpha$ (Supplemental Figure 3B). Prelimi-

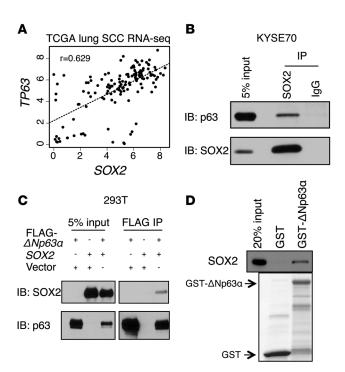


Figure 3

SOX2 associates with p63 in SCC. (**A**) Gene expression of SOX2 and TP63 in lung SCCs from TCGA. r=0.629; $P=1.02\times10^{-9}$. (**B**) SOX2-p63 interaction, shown by co-IP of p63 using an antibody against endogenous SOX2 in KYSE70 cells. (**C**) Co-IP of SOX2 with FLAG-tagged Δ Np63 α expressed in 293T cells. IPs by FLAG were immunoblotted with antibodies against SOX2 and p63. (**D**) Direct physical binding of SOX2 and GST- Δ Np63 α in GST pulldown assay. Top: Expression of SOX2 protein in 20% of input and GST- Δ Np63 α bound proteins, detected via immunoblotting with a SOX2 antibody. Bottom: GST and GST- Δ Np63 α expression, assessed by Coomassie blue staining.

nary evidence suggested a direct physical interaction of SOX2 and p63, obtained by co-IP of p63 with an antibody against endogenous SOX2 (Figure 3B); additional support for p63 as the immunoprecipitated protein was shown by co-IP after shRNA-mediated suppression of TP63 (shTP63; Supplemental Figure 3C). In addition, IPs with anti-FLAG antibody in 293T cells that were cotransfected with SOX2 and FLAG-tagged $\Delta Np63\alpha$ were found by immunoblotting to contain SOX2 (Figure 3C), further supporting a physical interaction between the proteins upon ectopic expression. An in vitro assay, pulldown with GST- Δ Np63 α , showed direct binding of SOX2 and Δ Np63 α under these conditions (Figure 3D), which suggests that the potential physical interaction between these factors may not be necessarily mediated by DNA or other factors.

To determine the requirement for $\Delta Np63\alpha$ expression for growth of SOX2-amplified SCC cells that predominantly express the $\Delta Np63\alpha$ isoform, we tested KYSE70 cells with 3 doxycycline-inducible (Doxinducible) TP63-directed shRNAs - 2 of which target all TP63 isoforms, including $\Delta Np63\alpha$ (shTP63), and 1 of which specifically targets only the $\Delta Np63\alpha$ isoform (sh $\Delta Np63$) - relative to shLacZ. Suppression of TP63 reduced growth of the KYSE70 cells (Figure 4A), similar to suppression of SOX2 with Dox-inducible SOX2-directed shRNAs (shSOX2; Figure 4B), indicating that both SOX2 and $\Delta Np63\alpha$ are essential for maintaining cell growth. Furthermore, combinato-

rial suppression of both SOX2 and $\Delta Np63\alpha$ led to a greater antiproliferative effect than either suppression individually (Figure 4C). However, we did not observe enhanced cell proliferation when SOX2 and p63 were ectopically overexpressed in AALE immortalized lung epithelial cells (data not shown).

SOX2 and p63 co-occupy the genome in SCC cells. Given their physical interaction and requirement for cell growth, we hypothesized that SOX2 and p63 would be colocalized on the genome in SCC and may coordinately regulate gene expression. Thus, we first used SOX2 occupancy data to query p63 consensus binding motifs (34) adjacent to the central SOX2 motif in SOX2-occupied peaks. p63 motifs were found to be enriched adjacent to the SOX2 motifs in SCC cells, with the centers of 2 motifs most commonly falling 25 bp apart (Figure 5A), partially reflecting the length of homodimeric p63 motif (19 bp). In contrast, enrichment of p63 motifs was absent in SOX2-occupied loci, and p63 expression was barely detectable, in H9 cells (Figure 5, A and B), which suggests that the SOX2-p63 collaboration is not a general feature of SOX2-expressing cells, but is possibly limited to squamous tissues.

To more directly test for genomic co-occupancy of SOX2 and p63, we profiled p63 occupancy in the 3 SCC cell lines by performing ChIP with an antibody against endogenous p63 and confirmed p63 motif enrichment in peaks (Supplemental Figure 4A). The overlaps of the peaks for SOX2 and p63 in the 3 cell lines were 13.3%–25.5%, and 33.4%–68.5% of the SOX2-p63 overlapping peaks were recurrently found in one or both of the other 2 lines (Supplemental Figure 4, B and C).

To identify high-confidence SOX2-p63 co-occupied regions, we first selected peaks present in 2 or more SCC SOX2 ChIP-seq data (5,083 regions). Within these recurrent SOX2 peaks, we distinguished peaks by evidence of p63 colocalization in the composite and recurrent p63 occupancy data from the 3 SCC lines (1,453 regions; 28.6%). These regions were supplemented with recurrent p63-occupied peaks (5,025 regions) similarly found co-occupied by SOX2 in composite and recurrent SOX2 occupancy data (1,906 regions; 37.9%). In all, we identified 2,426 high-confidence SOX2-p63-co-occupied regions as well as 3,549 SOX2-occupied regions without evidence of p63 binding (Figure 5C).

As the majority of SOX2 peaks lacked evidence for p63 localization, we searched for additional potential collaborating factors independent of p63 by performing motif analysis among those SOX2 peaks without evidence for p63 colocalization in SCC and found enrichment of motifs for AP-1 factors ($P=1\times10^{-83}$; Supplemental Figure 5, A and C), complexes of transcription factors such as FOS and JUN. These motifs were not found to be similarly enriched in p63–co-occupied SOX2 loci in SCCs, nor in SOX2 peaks in H9 ES cells (Supplemental Figure 5, B and C). We confirmed with ChIP–quantitative PCR (ChIP-qPCR) that JunD occupied several SOX2 binding sites with an AP-1 motif nearby (Supplemental Figure 5D), which suggests that AP-1 and other factors may also collaborate with SOX2 in SCC.

SOX2 collaborates with p63 in gene regulation. Focusing on those peaks showing SOX2-p63 colocalization, we sought to connect SOX2-p63-co-occupied loci to gene targets. First, we evaluated the effect of RNAi-induced suppression of SOX2 or TP63 upon gene expression in KYSE70 cells by mRNA sequencing (RNA-seq). At the global level, we found a positive correlation of the effects of SOX2 and TP63 suppression on gene expression (r = 0.492; Figure 6A); i.e., those genes whose expression was affected by SOX2 suppression also tended to be affected by TP63 suppression in the



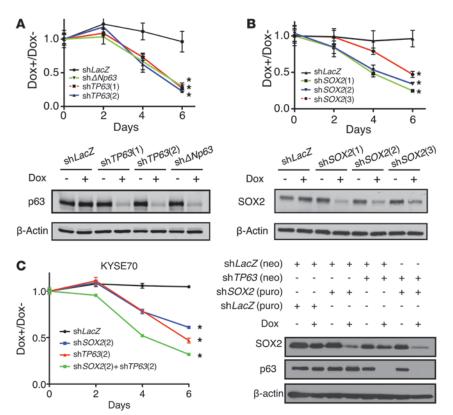


Figure 4

SOX2 and p63 are essential for SOX2-amplified SCC cell growth. (A) Effect of TP63-specific shRNA on viable cell numbers over time. Growth of KYSE70 cells after suppression of TP63 or LacZ control by Dox-inducible shRNAs — shTP63(1) and shTP63(2) target all TP63 isoforms, whereas sh∆Np63 specifically targets $\Delta Np63\alpha$ — is plotted as the ratio of Dox-treated to non-Dox-treated cell viability over time after seeding. Immunoblots show expression of p63 and β -actin as loading controls before and after Dox treatment. (B) Effect of shSOX2 on viable cell numbers over time. Growth of KYSE70 cells after suppression of SOX2 or LacZ by Dox-inducible shRNAs is plotted as in A. Immunoblots for SOX2 and β -actin are shown. (C) Effect of SOX2 and TP63 double suppression compared with suppression of each gene on viable cell numbers over time. Growth of KYSE70 cells and immunoblots after suppression of both SOX2 and TP63 by shSOX2 construct with puromycin resistance cassette and shTP63 construct with neomycin resistance as well as suppression of either SOX2 or TP63 alone are plotted as in A. Values are mean \pm SD of cells plated in 6 wells. *P < 0.0001, sum-of-squares F test of curve fitting.

same direction. By integrating these data into ChIP data, we found that genes whose expression was downregulated by suppression of either SOX2 or TP63 in KYSE70 cells were significantly enriched with the genes near our high-confidence SOX2-p63-co-occupied regions ($P = 6.35 \times 10^{-6}$ or $P = 7.66 \times 10^{-7}$, respectively), and genes whose expression was downregulated by suppression of both SOX2 and TP63 were also significantly enriched with the presence of SOX2 and p63 co-occupancy ($P = 2.04 \times 10^{-4}$) (Supplemental Table 3). We next identified genes whose expression changed >1.5-fold relative to control after TP63 suppression as well as after SOX2 suppression. We then overlapped this gene list with 2,523 genes with transcriptional start sites (TSSs) within 50 kb of one of the 2,426 high-confidence SOX2-p63-co-occupied peaks (Supplemental Table 4). We identified 93 genes that were commonly downregulated, and another 95 that were commonly upregulated, after SOX2 and TP63 suppression (Figure 6B, Supplemental Figure 6A, and Supplemental Tables 5 and 6).

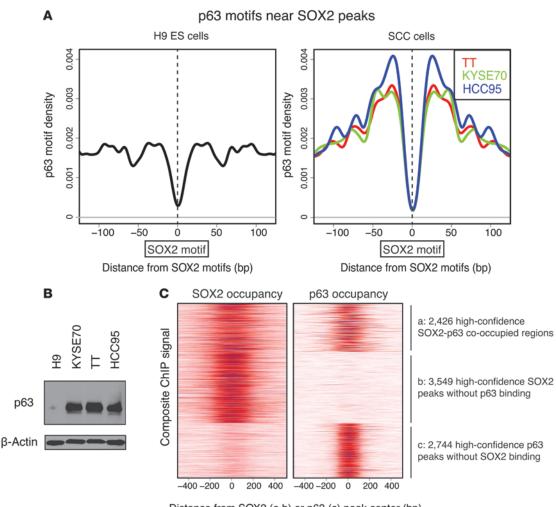
For those 93 genes with evidence for positive regulation by p63 and SOX2, we next evaluated their mRNA expression across primary SCCs from the lung SCC TCGA (10) and their correlation with SOX2 mRNA expression, with a focus on 5 genes present in the Cancer Gene Census (http://cancer.sanger.ac.uk/cancergenome/projects/census/#cl_analysis) (Figure 6C). Among them, ETV4, a member of the Ets family of transcription factors, was notable for both its correlation with SOX2 in primary SCCs and its previously reported role in SCC pathogenesis (35). ETV4 has also been reported to be overexpressed in airway basal cells (31), thus linking it to a SOX2-p63-positive cell population. We confirmed that ETV4 transcription was dependent on expression of both p63 and SOX2 by quantitative RT-PCR after treatment with Dox to induce shRNAs for the respective genes (Figure 6D).

We identified that 2 loci near the *ETV4* TSS were co-occupied by SOX2 and p63 (Figure 7A), at positions approximately 3 kb upstream and approximately 1 kb downstream of the TSS. Suppression of p63 by sh*TP63* led to decreased occupancy of SOX2 at the above-described positions in the *ETV4* locus, as measured by ChIP-PCR (Figure 7B), which suggests that SOX2 binding to these regions depends on simultaneous binding of p63. Furthermore, consistent with previous reports (35), proliferation and anchorage-independent growth of SCC cell lines were dependent on *ETV4*, as they were diminished by shRNA-mediated suppression of *ETV4* (sh*ETV4*; Figure 7, C and D, and Supplemental Figure 6B). However, ectopic expression of *ETV4* was not itself sufficient to rescue the effect of *SOX2* suppression upon proliferation of KYSE70 SCC cells (data not shown). These data suggest that *ETV4* and additional targets contribute to SOX2-p63's role in SCC.

In summary, our present demonstration of cooperative regulation by SOX2 and p63 of a factor essential for SCC proliferation and anchorage-independent growth supports our hypothesis that joint function of SOX2 and p63 contribute to squamous epithelial neoplasia.

Discussion

SOX2 is a highly recurrent amplified oncogene in multiple forms of SCC and has been implicated in breast and prostate carcinomas and glioblastoma (12–14). Many have hypothesized that SOX2 and OCT4 act coordinately in cancer to induce "stem-like" tumors that are poorly differentiated and demonstrate poor survival (36–38). In SCCs, in contrast, SOX2 amplification is largely a positive prognostic marker (2, 15). Together with the predilection for SOX2 amplification in cancers of squamous cell origin, these associations suggest that SOX2's oncogenic role in SCC may



Distance from SOX2 (a,b) or p63 (c) peak center (bp)

Figure 5
SOX2 genomic occupancy in SCC overlaps with p63 occupancy. (A) Appearance of p63 DNA binding motifs, plotted around the SOX2 motif near the summit of SOX2 peaks in the H9 ES cell line and the 3 SCC lines. The p63 motif was highly enriched within 50 bp from the SOX2 motif in SOX2 peaks in all 3 SCC lines, but not in H9 cells. (B) p63 protein expression in H9 cells and the 3 SCC lines. p63 expression was barely detectable in H9 cells by immunoblots. (C) Heatmap showing composite signals from SOX2 and p63 ChIP-seq data in the 3 SCC lines. ChIP signal intensity is shown by red shading. Shown are SOX2-p63—co-occupied loci, high-confidence SOX2 peaks without evidence of p63 binding, and high-confidence p63 peaks without evidence of SOX2 binding.

reflect its physiologic function in squamous epithelium more than its role in pluripotency.

We demonstrated that SOX2 genome-wide occupancy in SCCs of esophageal and lung origin was distinct from that in ES cells, where SOX2 functioned in cooperation with the pluripotency factor OCT4. Instead, we discovered p63 as a novel collaborative interacting partner protein of SOX2 in SCCs. Our data further suggested that the differences in SOX2 occupancy between SCC and pluripotent ES cells likely largely reflect the presence of different SOX2 partners (i.e., p63 and OCT4, respectively). SOX2 and p63 coregulated scores of genes in SCC, including the previously described squamous cell oncogene ETV4 (35), which we confirmed to be essential for SCC proliferation. In light of our demonstration that amplification and overexpression of SOX2 and TP63 could be readily used as clinical markers, it would be interesting to identify additional target genes of functional col-

laboration between these 2 factors that may be more relevant to translational research.

Expression of p63 is accompanied by SOX2 expression in precursor populations of both the esophagus and the large airways, and expression of both factors is downregulated during cellular maturation, which suggests they may act to maintain an immature cell state (16, 17, 39, 40). Indeed, loss of function of Δ Np63 enhances terminal maturation in stratified squamous epithelium (41). On one hand, Δ Np63\$\alpha\$ overexpression has been shown in a rat cell line to accelerate tumor growth in nude mice (42), and transgenic mice overexpressing Δ Np63\$\alpha\$ develop epidermal hyperplasia (43). Tp63 was recently shown to be required for maintenance of tumors, as inducible genetic deletion of all its isoforms led to blockade of chemically induced SCC formation in mice (44). In vivo models have shown that ectopic overexpression of Δ Np63 cooperates with oncogenic Hras to induce



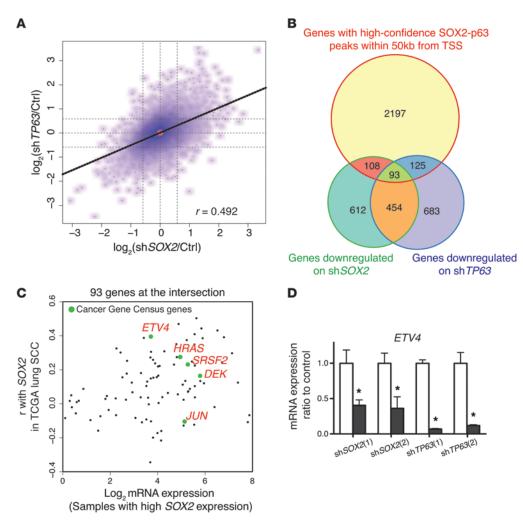


Figure 6
Identification of target genes coregulated by SOX2 and p63. (**A**) Gene expression changes by suppression of SOX2 and of TP63 in KYSE70 cells. Cloud color represents plot density. Changes were significantly positively correlated (trend line; r = 0.492). (**B**) Overlap among genes with high-confidence SOX2-p63—co-occupied loci within 50 kb from their TSS, and genes whose expression was downregulated (>1.5 fold) after suppression of SOX2 or TP63, in KYSE70 cells. See Supplemental Table 5 for the 93 genes at the intersection. (**C**) Correlation of expression of the 93 genes co-occupied by SOX2 and p63 and downregulated upon SOX2 and TP63 suppression with SOX2 mRNA expression from the lung SCC TCGA. Of these 93 genes within TCGA lung SCC samples with the top quartile of SOX2 expression, 5 genes were on the list of Cancer Gene Census (green circles). (**D**) ETV4 mRNA, determined by quantitative RT-PCR, before and after induction of shSOX2 or shTP63 in KYSE70 cells (the stable cell lines used in Figure 4). Mean ratio \pm SD of triplicates is shown. *P < 0.05, 2-way ANOVA with Bonferroni post-test.

SCCs from keratinocytes by promoting progenitor cell expansion (45). On the other hand, ectopic *SOX2* in the esophageal/ forestomach epithelium has been demonstrated to induce basal cell hyperplasia with loss of maturation and, in cooperation with inflammation and *STAT3* activation, lead to SCCs (8). These 2 squamous factors, however, have not been previously shown to functionally collaborate. Although we have not studied normal SOX2-p63-positive precursor cells, we speculate that these factors likely interact and may coregulate gene expression in this cell population as well.

While SOX2 plays pivotal roles in ES cell and adult stem cell maintenance across a diverse array of tissues, including eye lens, neurons, gastrointestinal glands, and squamous epithelia (16, 39, 46), the roles of p63 in development appear to be confined more to stratified squamous epithelia (33). Our findings

of overlapping SOX2 and p63 localization and function within squamous epithelium may explain the distinct contributions of SOX2 across multiple tissue types, indicating that SOX2 cooperates with p63 in a manner analogous to its actions with OCT4 in ES cells, BRN2 in neural stem cells, and PAX6 in the eye lens.

While our data suggest novel direct collaborative functions of SOX2 and p63 in SCCs, we also demonstrated that SOX2 may additionally cooperate with AP-1 transcription complex in SCCs independent of p63 function. AP-1 complexes have been shown to function in normal squamous epithelial development to help prevent maturation (47). The degree to which SOX2 and AP-1 may jointly act in SCC and in normal squamous epithelial tissues will need to be addressed in future studies. Similarly, SOX2 may act with other factors in SCC and squamous development.

7

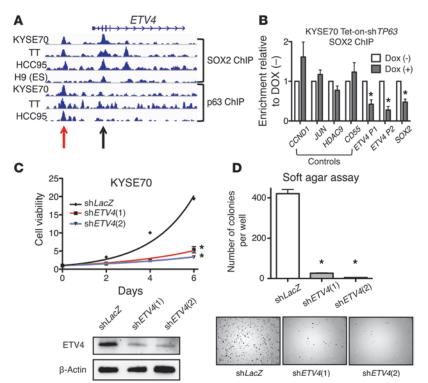


Figure 7

SOX2 and p63 regulate ETV4 expression that is essential for cell growth in SCCs. (A) ChIP signals of SOX2 or p63 for the indicated cell lines near ETV4. 2 loci near the TSS of ETV4 were co-occupied by SOX2 and p63 in all 3 SCC lines (arrows); the locus upstream of TSS was not enriched for SOX2 ChIP in H9 ES cells (red arrow). (B) ChIP enrichment of SOX2, determined by qPCR, was attenuated after shTP63 induction in KYSE70 cells at the locus 3 kb upstream of ETV4 (2 distinct sets of primers were used [P1 and P2]) and the SOX2 locus. However, other SOX2-occupied regions (CCND1, JUN, HDAC9, and CD55 promoters) were not affected. Percent recovery of input for ChIP was calculated based on 10% non-IP DNA sample for each experiment. Mean ± SD of triplicates are shown. *P < 0.05, t test. (C) Growth of KYSE70 cells after seeding, following infection with 2 independent shETV4s or with shLacZ control. Data are mean ± SD of cell plated in 6 wells. *P < 0.0001 vs. shLacZ, sum-ofsquares F test of exponential growth model. Immunoblots for ETV4 protein as well as β -actin in the lysates of the respective cells are also shown. (D) Anchorageindependent growth of the same KYSE70 cells as in C. Mean colony numbers ± SD of triplicate wells 3 weeks after seeding are shown. *P < 0.001, t test with Dunnett multiple-comparison test. Representative images are also shown. Original magnification, ×6.3.

Identification of SOX2-p63 collaboration in SCC may further explain why SOX2 amplifications are so enriched in SCCs; given that TP63 is located approximately 7 Mb from the SOX2 locus, these genes were often coamplified. The squamous-specific joint actions of SOX2 and p63, each of which helps maintain the immature precursor population of squamous epithelia, may thus be frequently co-opted during the process of squamous carcinogenesis. However, it remains to be determined what differences may exist between SOX2 function in SCC and in squamous epithelial precursor cells. As SOX2 expression increases in the process of amplification/overexpression and new chromatin sites become accessible during transformation to cancer, it is likely that new transcriptional targets emerge. The extent to which SOX2 binding to these predicted novel sites contributes to its oncogenic function, and the relative importance of p63 collaboration with SOX2 in distinct classes of genomic targets, are important future questions. Here, by demonstrating the protein-protein interaction and genomic co-occupancy of SOX2 and p63, we established a new foundation by which to study the function of SOX2 in SCC and in normal squamous epithelia, and our data suggest that disruption of SOX2-p63 interaction may be therapeutically valuable for SCC.

Methods

Further information can be found in Supplemental Methods.

Cell lines. SCC cells were maintained in RPMI 1640 (HCC95 and KYSE70) or DMEM (TT) with 10% FBS. H9 ES cells were provided by T. Schlaeger (Boston Children's Hospital, Boston, Massachusetts, USA).

TAP. TAP of FLAG-HA-SOX2 with nuclear pellets prepared from SCC cells was performed as recently described (48). Briefly, chromatin pellet separated from solubilized nuclear fraction (NE) was digested with Micrococcal nuclease. The chromatin extract (CE) was incubated with anti-FLAG-agarose (Sigma-Aldrich) followed by elution with FLAG peptide

(Sigma-Aldrich). The eluate was filtered and further incubated with anti-HA-agarose conjugate (Santa Cruz) followed by elution with HA peptide (4 mg/ml). The final HA eluate from CE fractions was processed for LC-MS/MS sequencing and data analysis.

Protein digestion, LC-MS/MS data acquisition, and database searching. Purified tryptic peptides of CE fractions were analyzed on an LTQ-Orbitrap-XL mass spectrometer equipped with a Digital PicoView electrospray source platform, modified from a recently described method (49–52). MS spectra were searched against 3 appended databases and processed to remove peptide spectral matches (PSMs) to the forward database with FDR > 1.0%. Protein abundance was estimated by calculating the sum of the extracted ion chromatograms area of the 3 most intense peptides (53). Proteins identified in >1% of 108 negative TAP controls were removed from the sets of SOX2 interactors. See Supplemental Methods for details.

Immunofluorescence. TT cells stably expressing FLAG-HA-SOX2 and parental TT cells were seeded on cover slips. Cells were fixed for 20 minutes with 4% paraformaldehyde, permeabilized for 20 minutes with PBS containing 0.02% Triton X-100, and blocked with 10% goat serum in PBS for 3 hours. Cells were then stained with either mouse anti-FLAG or rabbit anti-SOX2 antibody, along with rabbit IgG antibody as control. Expression of endogenous or ectopic SOX2 was detected by Alexa Fluor 594 (red) goat anti-rabbit IgG (Invitrogen, A11037) or Alexa Fluor 488 (green) goat anti-mouse IgG (Invitrogen, A11029), respectively. Nuclei were counterstained with DAPI (blue) with VECTASHIELD mounting medium (Vector Laboratories, H-1200). Fluorescent microscopic images were obtained with a Nikon Diaphot microscope using a Photometrix PXL cooled CCD camera. The microscope was equipped with the appropriate filters for 3-color imaging of cells and with a motorized stage for obtaining z series images.

RNAi. The shΔNp63 sequence was published previously (54). Other shRNA sequences were obtained from the RNAi Consortium (TRC) (http://www.broadinstitute.org/rnai/trc; Supplemental Table 7). Oligonucleotides were cloned into pLKO.1, Dox-inducible pLKO-Tet-On, or pLKO-Tet-On-Neo lentiviral vector as indicated.



Cell growth assays and anchorage-independent growth assays. shRNA viral transfer, induction with Dox, cell growth assays, and anchorage-independent growth assays were performed as previously described (2, 55).

IP and GST pulldown assays. Whole-cell lysate of KYSE70 cells or 293T cells transfected with pcDNA3- Δ Np63 α -FLAG and pcDNA3-SOX2 were incubated with SOX2 antibody or goat IgG antibody and Dynabeads Protein G (Invitrogen) or anti-FLAG-agarose. GST pulldown assay was performed as previously described (56).

ChIP, qPCR, and Illumina sequencing. SOX2, p63 ChIP-seq, and JunD ChIP-qPCR were performed as previously described with modifications (55). See Supplemental Table 8 for primers used for qPCR. DNA libraries for Illumina cluster generation and sequencing with Illumina HiSeq 2000 were performed according to the manufacturer's protocol.

ChIP-seq data analysis and motif analysis. The binding sites for SOX2 or p63 were detected using MACS (22) after aligning to hg18 and normalization for copy number variation. Correlation matrix of ChIP-seq datasets was constructed from Pearson correlations between peak occurrence profiles. SOX2 and p63 binding peaks present in 2 or more SOX2 ChIP-seq data from SCC cells with and without evidence for significant p63 occupancy from composite p63 ChIP-seq analysis were distinguished, and vice versa for p63 peaks. Target genes were defined as genes whose TSSs were within 50 kb from SOX2 or p63 binding sites. Motifs for SOX2, p63, OCT4, or AP1 consensus binding closest to the center of SOX2 occupancy were profiled to the relative distance from SOX2 motif. See Supplemental Table 9 for details for threshold and number of occurrences along the genome for each motif.

cDNA library construction for RNA-seq. KYSE70 cells expressing Doxinducible shSOX2 or shTP63 were treated with or without Dox for 4 days. cDNA library prepared from extracted RNA samples were sequenced with HiSeq 2000 (Illumina).

RNA-seq analysis. RNA-seq reads were aligned to hg19 and exon-exon junctions (ensembl v64) with PRADA pipeline (57). Transcriptome was collapsed to gene level using GENCODE version 12. Gene expression was represented by the mean of duplicates, with log₂ fold changes >1.5 considered to be differentially expressed upon suppression via shRNA. Gene

- Jemal A, et al. Annual Report to the Nation on the Status of Cancer, 1975-2009, featuring the burden and trends in human papillomavirus(HPV)-associated cancers and HPV vaccination coverage levels. *J Natl Cancer Inst*. 2013;105(3):175-201.
- Bass AJ, et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. Nat Genet. 2009;41(11):1238–1242.
- Gen Y, et al. SOX2 identified as a target gene for the amplification at 3q26 that is frequently detected in esophageal squamous cell carcinoma. *Cancer Genet Cytogenet*. 2010;202(2):82–93.
- 4. Hussenet T, et al. SOX2 is an oncogene activated by recurrent 3q26.3 amplifications in human lung squamous cell carcinomas. *PLoS One.* 2010; 5(1):e8960.
- Maier S, et al. SOX2 amplification is a common event in squamous cell carcinomas of different organ sites. Hum Pathol. 2011;42(8):1078–1088.
- 6. Chen X, et al. Multilayered epithelium in a rat model and human Barrett's esophagus: similar expression patterns of transcription factors and differentiation markers. BMC Gastroenterol. 2008:8:1.
- 7. Lu Y, et al. Evidence that SOX2 overexpression is oncogenic in the lung. *PLoS One*. 2010;5(6):e11022.
- 8. Liu K, et al. Sox2 cooperates with inflammation-mediated Stat3 activation in the malignant transformation of foregut basal progenitor cells. *Cell Stem Cell*. 2013;12(3):304–315.
- 9. Rudin CM, et al. Comprehensive genomic analy-

expression values in RPKM for TCGA lung SCC samples were obtained via TCGA (https://tcga-data.nci.nih.gov/docs/publications/lusc_2012/). SOX2 high-expressing lung SCC samples at the top quartile were used for average expression of each annotated gene.

Statistics. Effects of RNAi on cellular proliferation were analyzed using F tests of curve fitting. Correlation of gene expression to SOX2 expression was determined by Pearson r. Other data were examined using 2-way ANOVA with Bonferroni post-test or using 2-tailed t test with or without Dunnett multiple-comparison test, as indicated in the figure legends. A P value less than 0.05 was considered significant.

Accession numbers. ChIP-seq (accession no. GSE46837) and RNA-seq (accession no. GSE47058) data were deposited in GEO.

Acknowledgments

We thank Ramesh Shivdasani, Myles Brown, and Orit Rozenblatt-Rosen for helpful discussions; Thorsten Schlaeger for providing H9 ES cells; and Fujiko Duke for technical assistance. This work was supported by the National Cancer Institute (K08 CA134931 to A.J. Bass), a Dana-Farber/Harvard Cancer Center Lung Cancer SPORE Developmental Project Award (NIH P50 CA090578 to A.J. Bass and K. Wong), the Lung Cancer Research Foundation (to A.J. Bass), and a Clinical Scientist Development Award from the Doris Duke Charitable Foundation (to A.J. Bass). J. Que is supported by grant R00 DX082650; A.K. Rustgi is supported by grants P01 CA098101, R01 DK060694, R01 DK056645, and P30 DK050306; and M. Meyerson is supported by grants R01 CA109038, P01 CA154303, P20 CA90578, and U24 CA126546.

Received for publication June 13, 2013, and accepted in revised form December 30, 2013.

Address correspondence to: Adam J. Bass, Dana-Farber Cancer Institute, 450 Brookline Ave, Boston, Massachusetts 02215, USA. Phone: 617.632.2477; Fax: 617.582.9830; E-mail: Adam_Bass@dfci.harvard.edu.

- sis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nat Genet.* 2012; 44(10):1111–1116.
- Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. 2012;489(7417):519–525.
- Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663–739.
- Ben-Porath I, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet.* 2008; 40(5):499–507.
- 13. Jia X, et al. SOX2 promotes tumorigenesis and increases the anti-apoptotic property of human prostate cancer cell. *J Mol Cell Biol*. 2011;3(4):230–238.
- 14. Alonso MM, et al. Genetic and epigenetic modifications of Sox2 contribute to the invasive phenotype of malignant gliomas. *PLoS One*. 2011; 6(11):e26740.
- Wilbertz T, et al. SOX2 gene amplification and protein overexpression are associated with better outcome in squamous cell lung cancer. *Mod Pathol*. 2011;24(7):944–953.
- 16. Que J, et al. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. *Development*. 2007;134(13):2521-2531.
- 17. Yuan P, et al. Sex determining region Y-Box 2 (SOX2) is a potential cell-lineage gene highly

- expressed in the pathogenesis of squamous cell carcinomas of the lung. *PLoS One*. 2010;5(2):e9112.
- Kamachi Y, Uchikawa M, Kondoh H. Pairing SOX off: with partners in the regulation of embryonic development. Trends Genet. 2000;16(4):182–187.
- Chew J-L, et al. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol*. 2005; 25(14):6031–6046.
- Boyer LA, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell.* 2005; 122(6):947–956.
- Tanaka S, Kamachi Y, Tanouchi A, Hamada H, Jing N, Kondoh H. Interplay of SOX and POU factors in regulation of the Nestin gene in neural primordial cells. *Mol Cell Biol*. 2004;24(20):8834–8846.
- Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.
- 23. Lister R, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009;462(7271):315–322.
- Sarkar A, Hochedlinger K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. Cell Stem Cell. 2013;12(1):15–30.
- Wang J, et al. A protein interaction network for pluripotency of embryonic stem cells. *Nature*. 2006; 444(7117):364–368.
- Pardo M, et al. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell. 2010;6(4):382–395.
- 27. van den Berg DL, et al. An Oct4-centered protein

- interaction network in embryonic stem cells. Cell Stem Cell. 2010;6(4):369–381.
- Engelen E, et al. Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. Nat Genet. 2011;43(6):607–611.
- Ring BZ, et al. A novel five-antibody immunohistochemical test for subclassification of lung carcinoma. Mod Pathol. 2009;22(8):1032–1043.
- 30. Que J, Luo X, Schwartz RJ, Hogan BL. Multiple roles for Sox2 in the developing and adult mouse trachea. *Development*. 2009;136(11):1899–1907.
- Rock JR, et al. Basal cells as stem cells of the mouse trachea and human airway epithelium. Proc Natl Acad Sci U S A. 2009;106(31):12771–12775.
- Parsa R, Yang A, McKeon F, Green H. Association of p63 with proliferative potential in normal and neoplastic human keratinocytes. *J Invest Dermatol*. 1999;113(6):1099–1105.
- Yang A, et al. p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. *Nature*. 1999;398(6729):714–718.
- 34. Kouwenhoven E, et al. Genome-wide profiling of p63 DNA-binding sites identifies an element that regulates gene expression during limb development in the 7q21 SHFM1 locus. *PLoS Genet.* 2010;6(8):e1001065.
- 35. Yuen HF, et al. The role of Pea3 group transcription factors in esophageal squamous cell carcinoma. *Am J Pathol.* 2011;179(2):992–1003.
- Li X, et al. Expression of sox2 and oct4 and their clinical significance in human non-small-cell lung cancer. *Int J Mol Sci.* 2012;13(6):7663–7675.
- Huang P, et al. Role of Sox2 and Oct4 in predicting survival of hepatocellular carcinoma patients after hepatectomy. Clin Biochem. 2011;44(8–9):582–589.

- Saigusa S, et al. Correlation of CD133, OCT4, and SOX2 in rectal cancer and their association with distant recurrence after chemoradiotherapy. *Ann Surg Oncol.* 2009;16(12):3488–3498.
- Arnold K, et al. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell Stem Cell. 2011;9(4):317–329.
- 40. Kumar PA, et al. Distal airway stem cells yield alveoli in vitro and during lung regeneration following H1N1 influenza infection. Cell. 2011;147(3):525–538.
- 41. Romano RA, Ortt K, Birkaya B, Smalley K, Sinha S. An active role of the DeltaN isoform of p63 in regulating basal keratin genes K5 and K14 and directing epidermal cell fate. PLoS One. 2009;4(5):e5623.
- Hibi K, et al. AIS is an oncogene amplified in squamous cell carcinoma. Proc Natl Acad Sci U S A. 2000;97(10):5462–5467.
- 43. Yang X, et al. ΔNp63 versatilely regulates a Broad NF-κB gene program and promotes squamous epithelial proliferation, migration, and inflammation. *Cancer Res.* 2011;71(10):3688–3700.
- 44. Ramsey MR, et al. FGFR2 signaling underlies p63 oncogenic function in squamous cell carcinoma. *J Clin Invest.* 2013;123(8):3525–3538.
- 45. Keyes WM, et al. ΔNp63α is an oncogene that targets chromatin remodeler Lsh to drive skin stem cell proliferation and tumorigenesis. *Cell Stem Cell*. 2011;8(2):164–176.
- Pevny L, Placzek M. SOX genes and neural progenitor identity. Curr Opin Neurobiol. 2005;15(1):7–13.
- Ding X, et al. Epigenetic activation of AP1 promotes squamous cell carcinoma metastasis. Sci Signal. 2013;6(273):ra28.
- 48. Adelmant G, et al. DNA ends alter the molecular composition and localization of Ku multi-

- component complexes. *Mol Cell Proteomics*. 2012; 11(8):411–421.
- Parikh JR, et al. multiplierz: an extensible API based desktop environment for proteomics data analysis. BMC Bioinformatics. 2009;10:364.
- Ficarro SB, et al. Improved electrospray ionization efficiency compensates for diminished chromatographic resolution and enables proteomics analysis of tyrosine signaling in embryonic stem cells. *Anal Chem.* 2009;81(9):3440–3447.
- 51. Wang Q, Moore MJ, Adelmant G, Marto JA, Silver PA. PQBP1, a factor linked to intellectual disability, affects alternative splicing associated with neurite outgrowth. *Genes Dev.* 2013;27(6):615–626.
- Rozenblatt-Rosen O, et al. Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature*. 2012; 487(7408):491-495.
- Silva JC, Gorenstein MV, Li G-Z, Vissers JPC, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition.
 Mol Cell Proteomics. 2006;5(1):144–156.
- 54. Sabbisetti V, et al. p63 promotes cell survival through fatty acid synthase. PLoS One. 2009; 4(6):e5877.
- Watanabe H, et al. Integrated cistromic and expression analysis of amplified NKX2-1 in lung adenocarcinoma identifies LMO3 as a functional transcriptional target. Genes Dev. 2013;27(2):197–210.
- 56. Ma Q, et al. FoxO1 mediates PTEN suppression of androgen receptor N- and C-terminal interactions and coactivator recruitment. *Mol Endocrinol*. 2009;23(2):213–225.
- 57. Berger MF, et al. Integrative analysis of the melanoma transcriptome. *Genome Res.* 2010;20(4):413–427.