Journal of Pathology
J Pathol September 2021; **255:** 16–29
Published online 19 June 2021 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/path.5733

ORIGINAL PAPER

SOX9 has distinct roles in the formation and progression of different non-small cell lung cancer histotypes

Jie Bao[1], Katja Närhi[1,2], Ana Teodòsio[3], Annabrita Hemmes[1], Nora M Linnavirta[1], Mikko I Mäyränpää[4,5], Kaisa Salmenkivi[5], John Le Quesne[3,6] and Emmy W Verschuren[1]*

[1] *Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland*
[2] *GlaxoSmithKline, Espoo, Finland*
[3] *MRC Toxicology Unit, University of Cambridge, Cambridge, UK*
[4] *HUSLAB, Division of Pathology, Helsinki University Hospital, Helsinki, Finland*
[5] *Department of Pathology, University of Helsinki, Helsinki, Finland*
[6] *Leicester Cancer Research Centre, University of Leicester, Leicester, UK*

**Correspondence to: EW Verschuren, Institute for Molecular Medicine Finland (FIMM), Helsinki Institute of Life Science (HiLIFE), University of Helsinki, Tukholmankatu 8, 00290 Helsinki, Finland. E-mail: emmy.verschuren@helsinki.fi*

Abstract

The transcription factor SOX9 is a key regulator of multiple developmental processes and is frequently re-expressed in non-small cell lung cancer (NSCLC). Its precise role in the progression of NSCLC histotypes has, however, remained elusive. We show that SOX9 expression relates to poor overall survival and invasive histopathology in human non-mucinous adenocarcinoma and is absent in murine early minimally invasive and low in human *in situ* adenocarcinoma. Interestingly, despite wide SOX9 expression across advanced NSCLC histotypes, its genetic deletion in the murine *Kras^{G12D};Lkb1$^{fl/fl}$* model selectively disrupted only the growth of papillary NSCLC, without affecting the initiation of precursor lesions or growth of mucinous or squamous tissue. Spatial tissue phenotyping indicated a requirement of SOX9 expression for the progression of surfactant protein C-expressing progenitor cells, which gave rise to papillary tumours. Intriguingly, while SOX9 expression was dispensable for squamous tissue formation, its loss in fact led to enhanced squamous tumour metastasis, which was associated with altered collagen IV deposition in the basement membrane. Our work therefore demonstrates histopathology-selective roles for SOX9 in NSCLC progression, namely as a promoter for papillary adenocarcinoma progression, but an opposing metastasis-suppressing role in squamous histotype tissue. This attests to a pleiotropic SOX9 function, linked to the cell of origin and microenvironmental tissue contexts.

Keywords: SOX9; non-small cell lung cancer; histopathology; cell of origin; metastasis; extracellular matrix; collagen IV; *in vivo* models

Received 8 November 2020; Revised 25 April 2021; Accepted 19 May 2021

No conflicts of interest were declared.

Introduction

Lung cancer remains the leading cause of cancer-related mortality, the majority of which is due to non-small cell lung cancer (NSCLC; 85%). NSCLC is subdivided into two major histotypes, adenocarcinoma (AC) and squamous cell carcinoma (SCC), of which the AC group is the most heterogeneous [1–4]. ACs are further subdivided based on the tissue's predominant histotype structure, particularly papillary AC (PAC), mucinous AC (MAC), and acinar AC (AAC) [5]. Certain driver mutations are selectively enriched in these AC subtypes. For example, *KEAP1* mutation or *NRF2* overexpression occurs more frequently in PACs compared with total NSCLCs, and *KRAS*, *EGFR* or *ALK* mutations have been associated with mucinous, lepidic growth, or signet ring structure, respectively [6,7]. Molecular and genetic functions therefore need to be interpreted within these histopathology-specific tissue contexts.

The transcription factor SOX9 is a member of the SOX (sex-determining region Y-related) family [8] and it regulates a variety of developmental processes [9–11]. During lung development in the mouse, SOX9 is expressed in a timed manner to coordinate lung branching morphogenesis and alveolar differentiation [12,13]. SOX9 is silenced in adult tissues but frequently re-expressed in cancer, where it has been associated with increased cell proliferation, transcriptional reprogramming, and self-renewal of stem cells, and increased cell invasiveness and drug resistance [14–17]. The *SOX9* gene resides on the long arm of chromosome 17 (17q24.3), which is frequently gained in non-smokers with lung AC [18]. In patients, elevated SOX9 expression has been associated

with poor survival and advanced clinical NSCLC stage, and its expression was shown to promote the proliferation of AC cells *in vitro* through affecting the CDKN1A and CDK4 cell cycle regulators [19,20]. However, while many studies have found oncogenic roles for SOX9, supporting its therapeutic targeting, the downregulation of SOX9 has also been linked to relapse of prostate cancer [21] and stage II colon cancer [22], indicating that SOX9 has pleiotropic functions. Despite NSCLC constituting a heterogeneous set of diseases, the potential histopathology-selective role of SOX9 in NSCLC progression has not been studied.

Distinct sets of developmental signals and transcription factors are expressed in different lung progenitor cells and thereby exert lineage-specific functions [23]. In recent years, the ways in which different cells of origin influence NSCLC progression have been addressed by applying genetically engineered mouse models [24]. Using adenovirus-mediated Cre recombination, we previously showed that conditional $Kras^{G12D}$ expression combined with loss of *Lkb1* (the official HUGO symbol for the human gene is *STK11*) in CC10-expressing club cells (CC10; the HUGO symbol for the gene is *SCGB1A1*) gives rise to a wide NSCLC histotype spectrum encompassing minimally invasive AC (IAC), MAC, AAC, and PAC, as well as adenosquamous carcinoma (ASC) [25], a rare NSCLC subtype exhibiting both SCC and AC features [26]. In contrast, in alveolar type II (AT2) progenitors expressing surfactant protein C (SPC), the same genetic alterations mainly lead to IAC and PAC formation [25]. Among murine NSCLC models, these $Kras^{G12D}$;$Lkb1^{fl/fl}$ mice are unique in carrying a wide spectrum of NSCLC histotypes, including the more aggressive ASC subtype, permitting comparative study of the putative histopathology-specific roles of SOX9 expression in NSCLC progression. We therefore established heterozygous and homozygous conditional *Sox9* null $Kras^{G12D}$;$Lkb1^{fl/fl}$ cohorts and examined how *Sox9* loss influenced the progression of lesions of distinct histotypes. We describe opposing histopathology-selective roles of SOX9 in NSCLC progression and metastasis, and link these to the cell of origin and tumour microenvironment contexts, with ramifications for future directions in therapeutic intervention.

Materials and methods

Human and mouse tumour tissues

Breeding of $Kras^{G12D/+}$;$Lkb1^{flox/flox}$ ($Kras^{G12D}$;$Lkb1^{fl/fl}$) mice, genotyping, and lung tumour initiation by intranasal infection with progenitor cell-directed Ad5-Cre viruses were performed as described previously [23]. *Sox9*-defficient cohorts were generated by crossing $Kras^{G12D}$;$Lkb1^{fl/fl}$ and $Sox9^{flox}$ mice (B6.129S7-$Sox9^{tm2Crm}$/J; The Jackson Laboratory, Bar Harbor, ME, USA). *Sox9* genotypes were determined using the supplied protocol 29713 (The Jackson Laboratory). Animal work was approved by the Experimental Animal Committee of the University of Helsinki and the State Provincial Office of Southern Finland (ESAVI/6365/2019). A tissue microarray (TMA-1) was constructed previously [27] that included human NSCLC SCC, ASC, and PAC lesions collected from the Helsinki Biobank's pathology sample archive (Helsinki Biobank project number HBP2016002). Raw data from the murine and TMA-1 analyses are provided in supplementary material, Tables S1 and S2. The large human adenocarcinoma TMA collection (TMA-2; 970 samples) was obtained from a continuous retrospective cohort of archival human lung adenocarcinoma tissue from the University Hospital Leicester, under research ethics committee approval. Clinicopathological and follow-up data were obtained from regional and national databases, and all morphological assessments were performed by a subspecialty consultant diagnostic histopathologist.

Tissue processing, immunohistochemistry (IHC), and image acquisition

Mouse lung tumours were processed and embedded in paraffin as reported in ref 23. FFPE tissue blocks containing whole murine lung or TMA blocks were cut at 4 μm thickness. Two sequential sections were placed on one glass slide (SUPERFROST® PLUS; Thermo Fisher Scientific, Waltham, MA, USA) and labelled to mark the sectioning order. Adjacent tissue sections were each stained with one antibody.

For all antibodies except anti-SPC (which did not require antigen retrieval), antigen retrieval was performed using 10 mM sodium citrate, pH 6.0. All antibodies except anti-SOX2 were incubated for 1.5 h at room temperature, while anti-SOX2 antibody was incubated overnight at 4 °C.

For routine signal development, BrightVision poly-HRP goat anti-rabbit IgG (Immunologic BV, Duiven, The Netherlands) was incubated at room temperature for 30 min. For developing the SOX2 signal, DaG-HRP (sc2020; Santa Cruz Biotechnology, Dallas, TX, USA; 1:100) was incubated at room temperature for 30 min. For signal detection, a Bright-DAB Substrate Kit (Immunologic BV) was used. Whole slide bright-field scans of stained tissue sections were acquired using a 20× objective (PANNORAMIC 250 scanner; 3DHISTECH Ltd, Budapest, Hungary). Sections cut at 3 μm from the TMA-2 human adenocarcinoma cohort were duplex-stained using a Ventana autostainer (Roche Diagnostics, Rotkreuz, Switzerland) with antibodies for cytokeratin AE1/AE3 (# NCL-L-AE1/AE3; Leica Biosystems, Wetzlar, Germany; 1:250) detected with Roche DISCOVERY yellow chromogen and SOX9 (supplementary material, Table S3) with Roche DISCOVERY purple chromogen. The primary antibodies used for mouse samples, TMA-1 are summarised in supplementary material, Table S3. Up to three 1-mm cores per tumour were analysed. Images of haematoxylin-counterstained images were acquired using a Hamamatsu nanozoomer (Hamamatsu Photonics, Shizuoka, Japan).

Region-of-interest (ROI)-specific image analysis

Histopathology classification of murine tumours was performed in conjunction with an expert pathologist as in refs 23 and 24. The sizes of individual tumours and lung tissues were measured using a Pannoramic Case-Viewer (3DHISTECH Ltd). For image analysis, TIFF images of murine lungs with haematoxylin and eosin (H&E) and biomarker staining were exported at a magnification of 1:2. Individual tumours on H&E-stained images were then outlined and annotated by their histotypes using ImageJ/Fiji [28] ('Analyze'→'Tools'→'ROI manager') to generate an ROI library for each murine sample. Images of serial sections stained for SOX9, Ki-67, pAKT, and pERK were registered with the nearby H&E images using our previously reported open-source software Spa-R [29]. For quantifying Ki-67 and SOX9-positive nuclei, individual tumour images were exported as inverted 8-bit grayscale images and processed with a CellProfiler pipeline [30] built with nuclei segmentation, intensity measurement, threshold-based filtration, and percentage calculation modules. The parameters in the pipeline were specified for each staining. To measure pERK and pAKT staining in individual tumours, a workflow using Spa-R and ImageJ/Fiji, and thresholds decided by three independent immunohistochemistry experts, were implemented as before [29]. All raw results can be found in supplementary material, Table S1. To mark collagen IV fibres with red pseudo-colour, our in-house open-source Spa-Q software [29] was applied, using identical thresholds for all samples.

Image analysis of the NSCLC cohort (TMA-1) and the adenocarcinoma cohort (TMA-2)

In both TMA-1 and TMA-2, tumour cell nuclei were categorised as negative, weak, moderate, or strong for SOX9 staining by the same expert pathologist, and an H-score was calculated for each core (% weak + 2× % moderate + 3× % strong nuclei, yielding a score from 0 to 300). The SOX9 expression cut-off, which defined 'SOX9-low' and 'SOX9-high' categories for TMA-2 survival and proliferation analyses, was generated using the Contal and O'Quigley method [31]. This cut-off value corresponded to the most significant difference in the survival curves between the two groups and was the one that provides the minimum P value among all potential cut-off values. The same cut-off was used in all the analyses. Ki-67 staining analysis on TMA-2 was performed with the Ventana pre-diluted ready-to-use kit and assessed by a trained histopathologist.

Statistical analyses

Statistical comparisons and data visualisation for clinical samples (TMA-1) and murine samples were performed using GraphPad Prism 8, where one-way ANOVA multiple comparison (Kruskal–Wallis test, non-parametric test), unpaired t-test (non-parametric Mann–Whitney test), or Fisher's exact test was used. Statistical significance of survival was assessed with a log-rank (Mantel–Cox) test. P values less than 0.05 were considered significant. For TMA-2, statistical analyses were performed in R (version 4.0.4). The Cox proportional hazard model was built through the survival package, and the Kaplan–Meier curves were produced with the *survminer* R package (https://CRAN.R-project.org/package=survminer), which was also used to generate the log–log plots for testing the proportional hazards assumption.

Results

The correlation between SOX9 expression and NSCLC staging is histopathology-selective

To evaluate SOX9 expression in NSCLC histotypes according to their stages, we performed IHC analyses of TMAs encompassing 28 human lung SCCs, 13 human lung ASCs, and 25 human lung PACs. SOX9 expression was not seen in normal lung tissues but was observed in the tumour regions in most of the samples (SCC: 96%; ASC: 100%; PAC: 88%). This significantly contrasts with the expression of SOX2, a well-established genetic driver of SCC, which was more commonly detected in SCCs and ASCs than in PACs (SCC: 85%; ASC: 100%; PAC: 8%) (supplementary material, Figure S1A). The SOX9 protein was mainly located in the nucleoplasm, but cytoplasmic/membranous expression was also observed, particularly in PACs (Figure 1A).

Based on visual scoring, we assigned tumour nuclei to 'SOX9-negative', 'SOX9-weak', 'SOX9-moderate' or 'SOX9-high' categories, and calculated H-scores to quantify the relative SOX9 expression level of each sample. The samples were then grouped by their clinical stages (Figure 1B). Among the three NSCLC histotypes, PACs showed a tendency towards a positive correlation between SOX9 expression and clinical stage (grouped by stages I & II and stages III & IV, $p = 0.041$). In contrast, ASCs and SCCs frequently showed high SOX9 expression also in the low-stage tumours.

To more deeply explore the relationship between SOX9 and the phenotype of human lung adenocarcinoma, we quantified the expression of SOX9 expression in epithelial tumour cell nuclei in a large cohort of 970 resected primary lung adenocarcinomas in tissue microarrays (supplementary material, Figure S1B). Corroborating previous studies [19,20,32], epithelial tumour cell proliferation assessed by the percentage of tumour nuclei expressing Ki-67 was significantly related to SOX9 expression ($p = 0.003$), and in analysis of patient survival time after surgery, higher median levels of SOX9 were related to relatively poor outcome (Cox model HR = 1.2, $p = 0.044$) (supplementary material, Figure S1C,D). Interestingly, nuclear SOX9 levels were higher in mucinous than in non-mucinous tumours ($p < 0.001$) (supplementary material, Figure S1C), and cores with invasive growth patterns showed a non-significant trend towards higher levels of SOX9

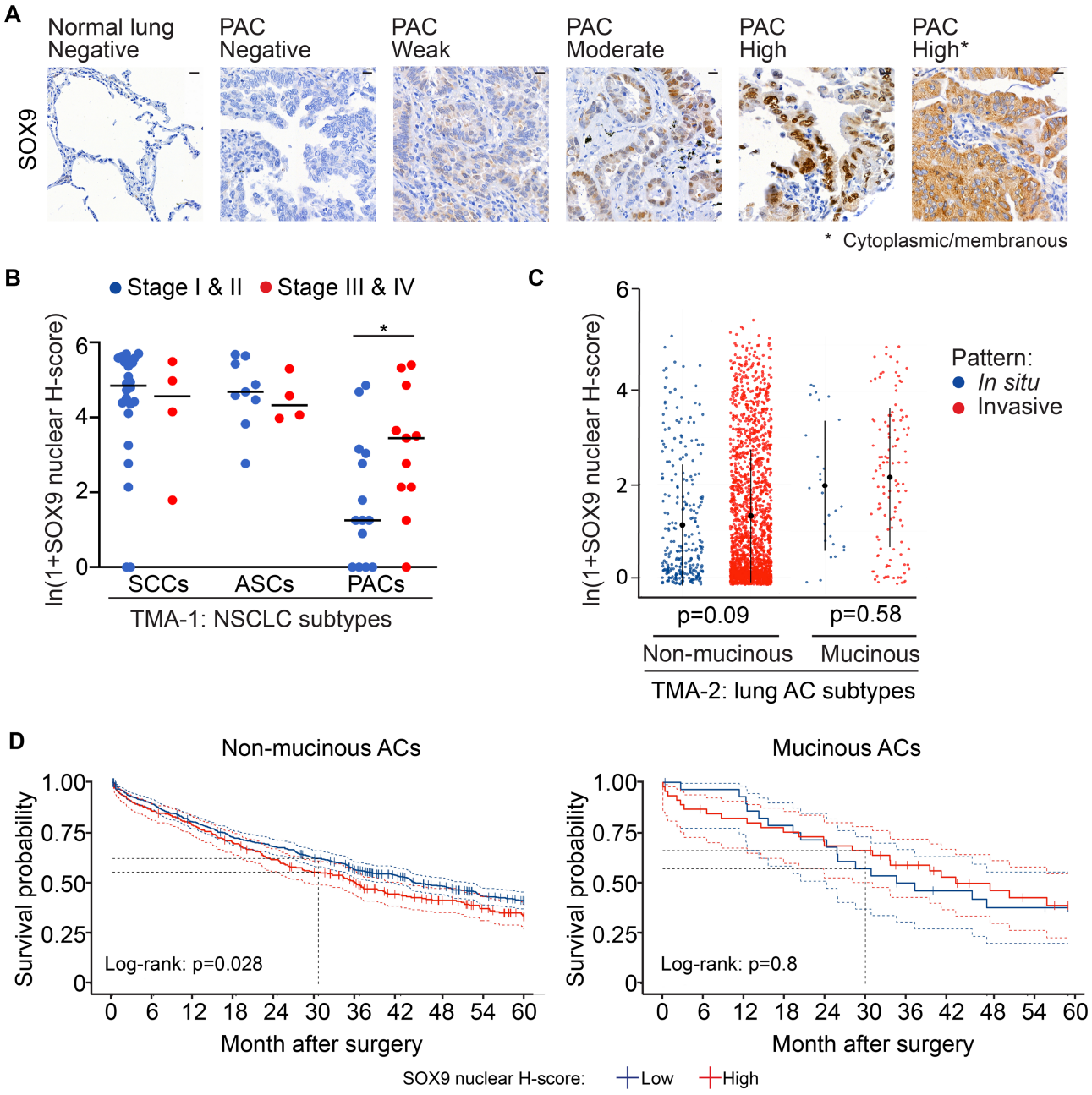

Figure 1. Histopathology-selective correlation between SOX9 expression and NSCLC progression. (A) Representative images illustrating SOX9 expression in normal and malignant epithelial lung tissues. Scale bar: 20 μm. (B) Dot plot showing SOX9 expression (assessed as H-scores) in SCC, ASC, and PAC NSCLC subtypes, grouped as stage I & II and stage III & IV, with staging categories determined following American Joint Committee on Cancer (AJCC) criteria. Each dot represents one sample. Statistical significance was assessed with unpaired *t*-test (non-parametric Mann–Whitney test): $^*p = 0.041$. (C) Dot plot showing the relationship between SOX9 and mucinous phenotype and local invasiveness. (D) Cox proportional hazards model of survival after surgery in patients with non-mucinous AC (left) and mucinous AC (right) related to SOX9 expression. Higher median nuclear SOX9 expression is significantly associated with relatively poor overall survival in patients with non-mucinous lung AC after surgery.

expression selectively in non-mucinous samples ($p = 0.09$ in non-mucinous; $p = 0.58$ in mucinous) (Figure 1C). Finally, a higher median level of SOX9 was related to relatively poor outcome in patients with non-mucinous but not with mucinous AC ($p = 0.028$ in non-mucinous; $p = 0.8$ in mucinous) (Figure 1D). This therefore tentatively suggested a non-mucinous AC histotype-selective impact of SOX9 expression on tumour invasiveness and patient survival.

SOX9 is commonly expressed in advanced murine *Kras*^{G12D};*Lkb1*$^{fl/fl}$ tumours

We subsequently applied animal models to study the role of SOX9 in NSCLC histotype formation. As a first step, we examined SOX9 expression in tumours from *Kras*^{G12D};*Lkb1*$^{fl/fl}$ mice (KL mice characterised in ref 25 and in this study named KLS^{wt}). Most KLS^{wt} histotypes, including ASC, MAC, and PAC, displayed high

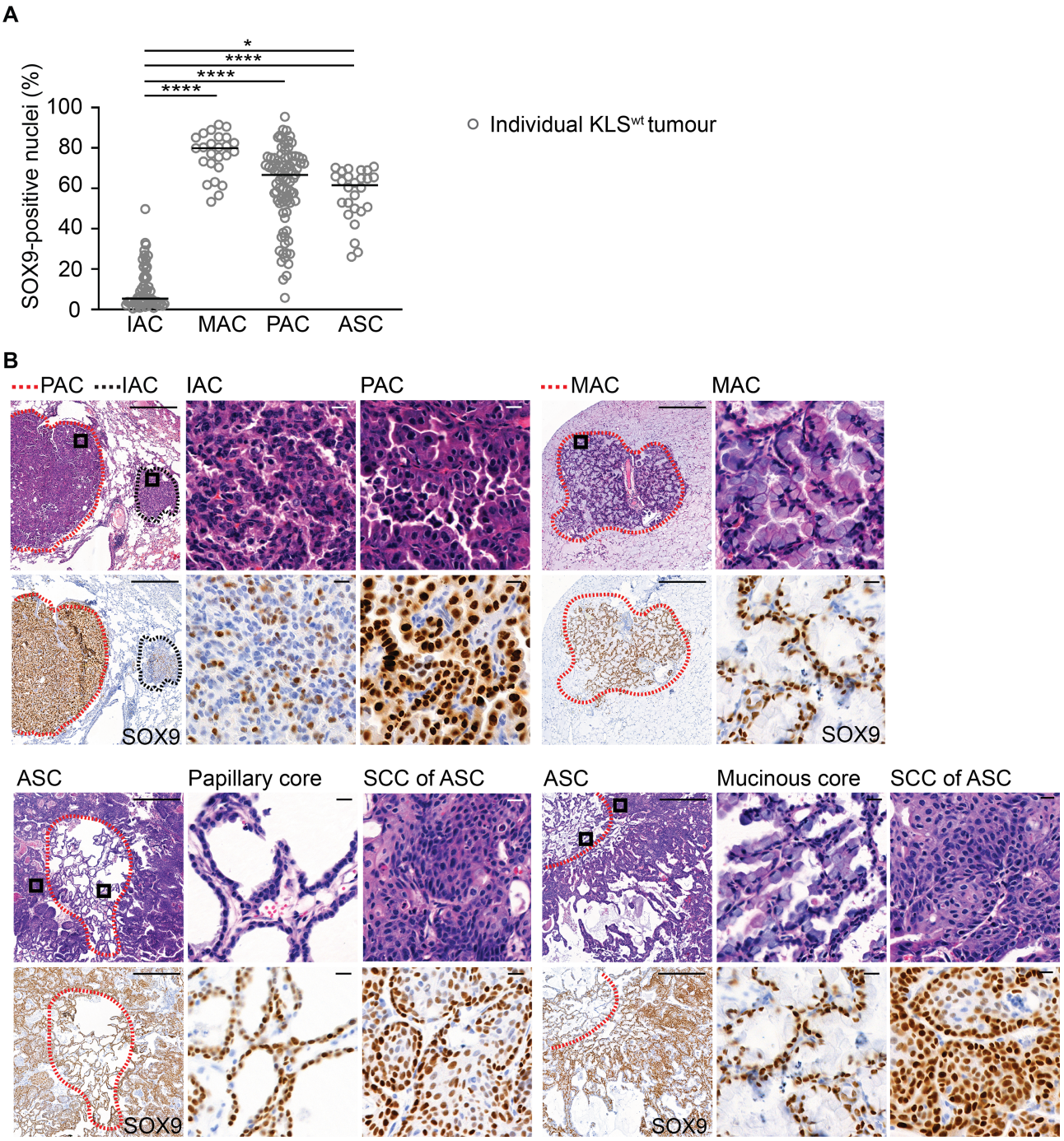

Figure 2. SOX9 expression is highest in advanced *Kras*^{G12D};*Lkb1*$^{fl/fl}$;*SOX9*$^{+/+}$ murine NSCLC tumours. (A) Scatter dot plot showing SOX9 expression in individual tumours, grouped by histotype and calculated as percentages of SOX9-postive cells in total number of tumour cells, from six KLS^{wt}-CC10 and eight KLS^{wt}-SPC mice. Statistical significance was assessed with one-way ANOVA (non-parametric Kruskal–Wallis test): *$p < 0.05$; ****$p < 0.0001$. (B) Representative H&E and SOX9 staining illustrating SOX9 expression in each histotype. Areas depicted at greater magnification are outlined with black squares in the H&E images. Red dashed lines outline a PAC, MAC, or separate SCC and AC regions within an ASC; the black dashed line outlines an IAC. Scale bars: 500 μm for original; 10 μm for magnified images.

SOX9 expression, but its expression was fully absent or low in IACs, which are small and represent the least aggressive lesion type with low proliferation and oncogenic signalling activities (Figure 2 and supplementary material, Figure S2A). Therefore, SOX9 expression levels in murine NSCLC histotypes were comparable to those in clinical NSCLC, showing absent or low expression in IAC or *in situ* non-mucinous adenocarcinomas and wide expression across all advanced NSCLC histotypes.

SOX9 expression is essential for papillary structure formation in *Kras*^{G12D};*Lkb1*$^{fl/fl}$ murine tumours

To investigate the impact of *Sox9* loss on histotype formation, we established *Sox9*-deficient *Kras*^{G12D};*Lkb1*$^{fl/fl}$

cohorts, including KLS^{het} (losing one floxed *Sox9* allele) and KLS^{null} (losing two floxed alleles), following infection with either Ad5-CC10-Cre or Ad5-SPC-Cre viruses (Figure 3A).

Kaplan–Meier survival analysis showed significantly longer survival of KLS^{het}-SPC and KLS^{null}-SPC mice, as well as KLS^{null}-CC10 mice, compared with the relative KLS^{wt} controls (Figure 3B). The increased life span

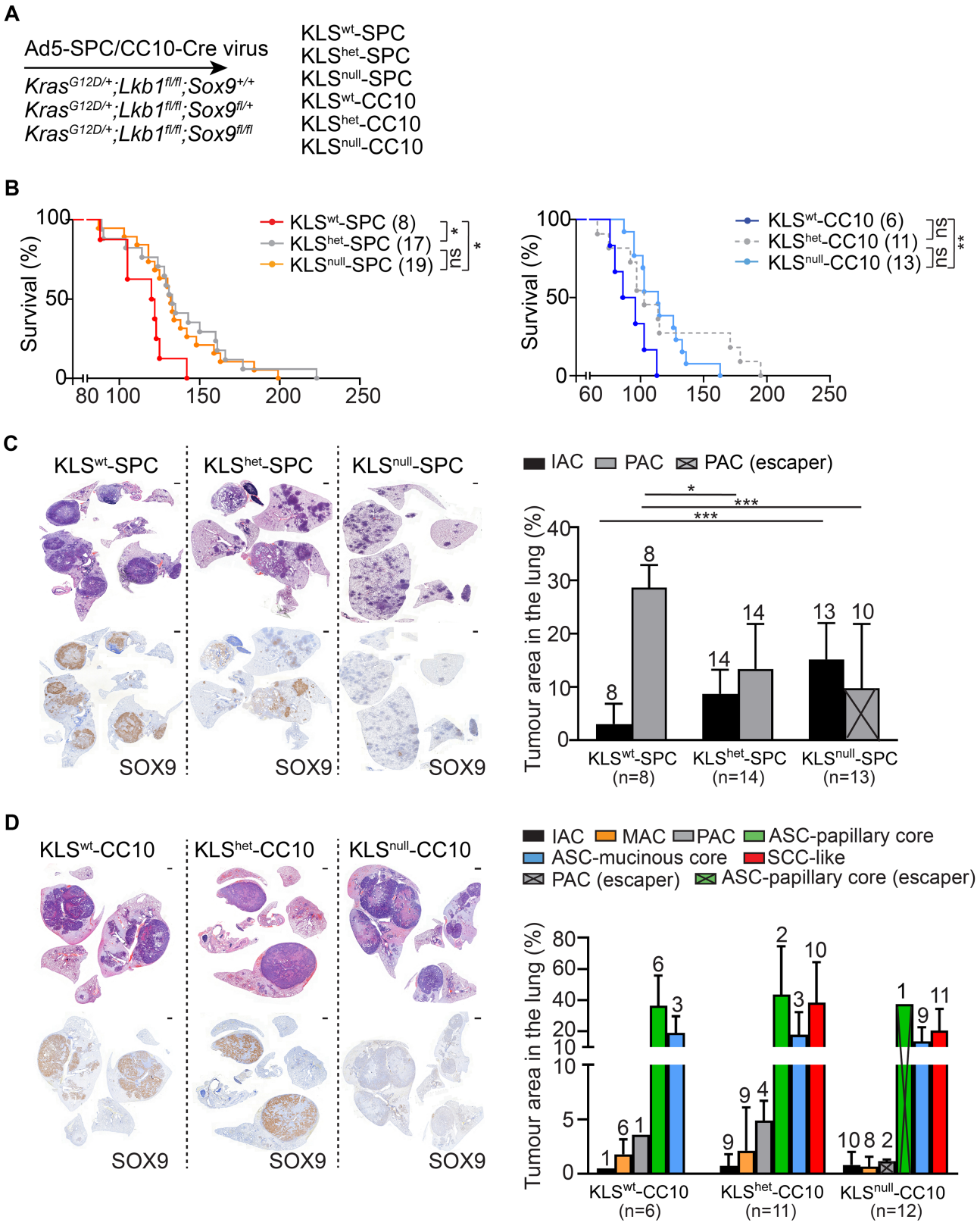

Figure 3 Legend on next page.

of KLS^{het}-SPC and KLS^{null}-SPC cohorts corresponded to a reduced PAC burden, and no papillary tissue structure was observed in SOX9-negative tumours in these mice (Figure 3C). Notably, recombination escape is known to arise in three-allelic compound mice [33], and SOX9-positive tumours were indeed observed at a recombination escape rate of around 30% (Figure 3C and supplementary material, Figure S3A,C and Table S1). Interestingly, these SOX9-positive tumours in the KLS^{null}-SPC cohort always showed a papillary tissue structure. On the other hand, the growth of smaller, less proliferative, and minimally invasive IACs (evidenced by lack of nuclear high mobility group AT-hook 2, or HMGA2, expression) was not affected by *Sox9* loss (Figure 3C and supplementary material, Figure S2A,B).

In Ad5-CC10-Cre cohorts, *Sox9* loss was associated with a shift in the major tumour burden from ASC to SCC tissue with large necrotic regions (Figure 3D). These SCC tumours were positive for the AC marker NKX2-1, lacked SOX2 expression, and strongly expressed nuclear HMGA2 (supplementary material, Figure S2B). This staining profile matched those of ASCs in the KLS^{wt}-CC10 cohort, suggesting that SCC tumours derived from ASCs; we hence called these 'SCC-like' from here onwards. Interestingly, only the formation of ASC with papillary AC cores, but not that with mucinous AC cores, was disrupted (Figure 3D). Additionally, PAC was the only AC subtype normally present in KLS^{wt}-CC10 mice that was absent following *Sox9* loss. Again, SOX9-positive lesions (ASC with papillary AC cores or PACs) were observed in the KLS^{null}-CC10 cohort, here at a recombination escape rate of around 9% (Figure 3D and supplementary material, Figure S3B,D and Table S1).

Taken together, these data suggested that SOX9 expression was selectively required for the growth of papillary histotype tissue. In support of this conclusion, SOX9-positive tumours in both the *Sox9^{null}*-SPC and the *Sox9^{null}*-CC10 cohorts, resulting from *Sox9* recombination escape, always showed a papillary tissue structure (supplementary material, Figure S3). As expected, based on the obligate role of LKB1 as a tumour suppressor in this model, we found no evidence for escape of *Lkb1* recombination (IHC staining; data not shown). Importantly, in both the CC10-infected and the SPC-infected cohorts, KLS^{het} resembled KLS^{null} mice in terms of survival, histopathology spectra, and tumour burden (Figure 3B–D and supplementary material, Figure S3C,D), together indicating that both heterozygous and homozygous *Sox9* loss selectively disrupted papillary NSCLC histotype formation.

SOX9 expression is essential for the oncogenic transformation of cancer-initiating SPC$^+$ cells

To understand why SOX9 selectively regulated the formation of papillary tissue structure, we studied the spatial expression of cell lineage markers and SOX9 in different tumour histotypes. In the Ad5-SPC-Cre cohorts, SPC, which marks AT2 progenitor cells and differentiated AT2 cells, was highly expressed in IACs, while SPC was absent or expressed at low levels in PACs (Figure 4A,B). Also in the Ad5-CC10-Cre cohorts, IACs were positive for SPC, whereas PACs and papillary AC cores of ASCs showed negative to low SPC expression (Figure 4C,D and supplementary material, Figure S4A), indicating that also CC10-Cre induced the transformation of SPC$^+$ cells. Other tumour histotypes found in Ad5-CC10-Cre cohorts that were not affected by *Sox9* loss (Figure 3D), particularly mucinous and squamous tissue, were always SPC-negative (Figure 4D and supplementary material, Figure S4A). Interestingly, we further found that SPC expression was mutually exclusive with SOX9 expression in PACs as well as papillary AC cores of larger ASC tumours, and marked tumour subregions that lacked a prominent papillary structure (Figure 4B,D). The concomitant appearance of papillary structure together with SPC loss upon SOX9 expression, together with the absence of PAC and papillary AC cores following *Sox9* loss as shown in the section above, suggested that SOX9 promotes tumour progression from SPC$^+$ progenitor cells, resulting in papillary tissue formation in the KL NSCLC model (illustrated in Figure 4E).

SOX9 expression differently regulates AC and SCC tissue metastases

Given the previous links between SOX9 expression and invasion, we next asked how *Sox9* loss affected the metastatic potential of NSCLC histotypes by analysing the mediastinal lymph nodes (supplementary material, Figure S4B,C). Metastases in KLS^{wt}-SPC (observed in

Figure 3. *Sox9* loss disrupts the formation of papillary histotype tissue in *Kras^{G12D};Lkb1$^{fl/fl}$* NSCLC tumours. (A) Murine *Sox9* loss-of-function and control cohorts established for this study. (B) Kaplan–Meier survival curves for the indicated cohorts. Statistical significance was assessed with log-rank (Mantel–Cox) test: *$p < 0.05$, **$p < 0.01$. (C) Left: representative images illustrating whole lung sections of individual mice from the indicated Ad5-SPC-Cre cohorts, depicting H&E staining and corresponding SOX9 expression. Scale bars: 1000 μm. Right: bar chart showing the histotype-selective tumour burden (% total histotype-specific tumour region per lung) in the indicated cohorts. Numbers of mice that carry the specified histotype are listed above each bar. All tumours were plotted; PACs found in the KLS^{null}-SPC cohort were due to *Sox9* recombination escape. A plot depicting the data of all tumours per individual mouse can be found in supplementary material, Figure S3C. Statistical significance was assessed with one-way ANOVA and non-parametric Kruskal–Wallis test: *$p < 0.05$, ***$p < 0.001$. (D) Left: representative images illustrating whole lung sections of individual mice from the indicated Ad5-CC10-Cre cohorts, depicting H&E staining and corresponding SOX9 expression. Scale bars: 1000 μm. Right: bar chart showing the histotype-selective tumour burden (% total histotype-specified tumour region per lung) in the indicated cohorts. Numbers of mice that carry the specified histotype are listed above each bar. All tumours were plotted; PACs and ASCs with papillary AC cores found in the KLS^{null}-CC10 cohort expressed SOX9, indicating recombination escape. A plot depicting all tumours per individual mouse can be found in the supplementary material, Figure S3D.

3/8 KLS^{wt}-SPC mice) were similar to those described previously [32] and represented AC histotype tissue with papillary structure (supplementary material, Figure S4C). Metastases that eventually developed in KLS^{null}-SPC mice (4/19 mice, including two minimal metastases) were always SOX9-positive, suggesting that they had

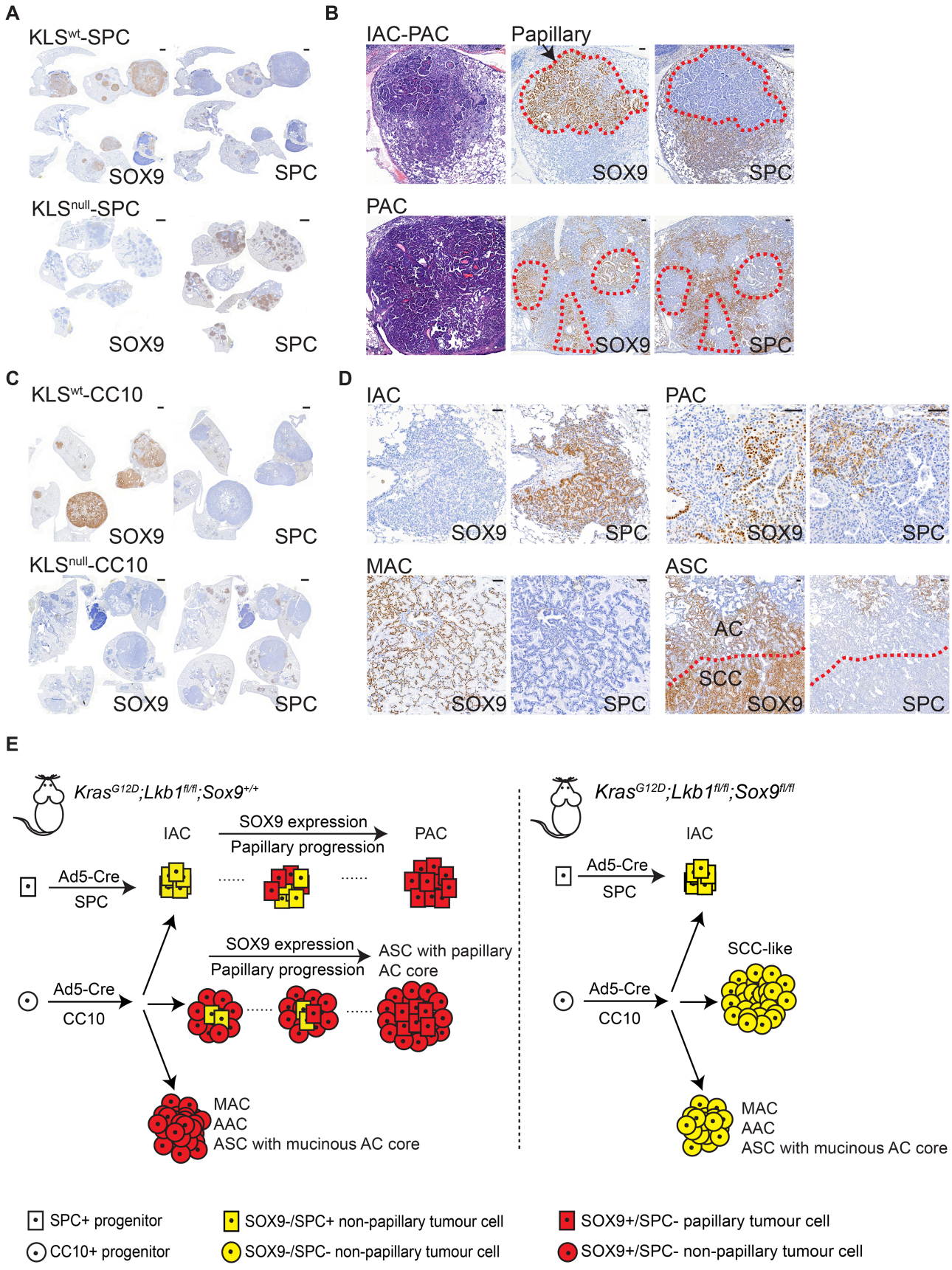

Figure 4 Legend on next page.

originated from PACs that had escaped *Sox9* recombination (Figure 5A,C). Also, the KLS^{het}-SPC mice developed frequent metastasis (12/17 mice, including six minimal metastases), which could not, however, be linked to SOX9 status due to the heterozygous SOX9 expression. Importantly, as reported previously [32], squamous metastases were never detected in KLS^{wt}-CC10 mice but were observed in 2/11 KLS^{het}-CC10 and 2/13 KLS^{null}-CC10 mice (Figure 5B,D and supplementary material, Figure S4D). Three of the four mice carrying squamous metastases succumbed to their disease relatively early and within the survival range of control KLS^{wt} mice. This analysis therefore suggested an overall reduction in AC metastases following *Sox9* loss, linked to a block in PAC progression, mirrored by a notable increase in SCC-like metastases following *Sox9* loss in murine NSCLC.

As SOX9 is known to regulate extracellular matrix (ECM) deposition, we next examined ECM features following Sox9 loss by staining of collagen IV (COLIV), a central component of the tumour basement membrane. COLIV was expressed in both IAC and PAC but only in the latter did it form a linear structure, which was also observed in papillary SOX9-expressing tissue at metastatic sites (Figure 6A,B). This suggested that a precise alignment of COLIV, regulated by SOX9 expression, accompanied the actual growth and later invasiveness of PACs. On the other hand, in *Sox9^{wt}* ASCs, COLIV formed continuous boundaries around squamous tumour cell pockets (Figure 6C), while *Sox9^{null}* SCC-like tumours and the SCC regions of ASCs with mucinous AC cores showed thinner parallel COLIV fibres as well as granular and focal depositions (Figure 6D). A similar alteration in COLIV rearrangement was also observed in the SCC-like LN metastases (Figure 6D). Taken together, in murine PACs, *Sox9* loss reduced primary papillary tumour progression and consequently their invasion, possibly related to an alteration in ECM deposition, while in the squamous histotype the altered COLIV deposition following *Sox9* loss in fact was associated with an increase in tumour invasiveness and metastasis (illustrated in Figure 6E).

Discussion

Re-expression of the lung developmental transcription factor SOX9 specifically in malignant lung tissues, in conjunction with the reported positive correlation of SOX9 expression levels with clinical NSCLC staging, renders SOX9 a good candidate for prognostic and therapeutic interrogation. We studied here the previously unexplored expression and role of SOX9 in NSCLC histotypes. We identified a trend towards elevated SOX9 expression in non-mucinous clinical invasive AC compared with *in situ* samples, as well as in murine PAC compared with IAC, with IAC constituting an early form of PAC. SOX9 was widely expressed across multiple invasive stage murine and clinical NSCLC histotypes, yet murine *Sox9* loss only affected papillary structure formation, not the initiation or progression of other histotype lesions. The overarching conclusion of our work is therefore that the role of SOX9 in NSCLC is to be interpreted in a histotype-selective manner.

The expression of *Kras^{G12D}* combined with *Lkb1* loss in alveolar AT2 or bronchial club lung progenitor cells gives rise to a range of NSCLC histotype tumours. While the alveolar differentiation marker SPC was uniformly expressed in SOX9-negative IAC, its downregulation coincided with the appearance of a papillary tissue structure, and SPC was completely absent in mature SOX9-positive papillary tumour regions. A transition from alveolar de-differentiation to gradual SOX9 expression was similarly seen following oncogenic KRAS expression in organoids [34] and corroborates data showing that ectopic expression of SOX9 in embryonic lungs prevented AT2 cell differentiation [12,13]. This indicates that activation of the KRAS–SOX9 axis is intrinsically associated with AT2 lineage exit, suggestive of an overall model in which cancer development may resemble a reverse process of normal lung tissue development. Use of a conditional *Sox9^{null}* model to address the role of SOX9 in NSCLC progression showed that *Sox9* loss disrupted the formation of PAC and papillary AC cores of ASC lesions but not IAC or histotypes derived from bronchial progenitors. This therefore

Figure 4. SOX9 expression drives the oncogenic transformation of SPC$^+$ progenitors. (A) Representative images illustrating SOX9 and SPC staining in the lung of a KLS^{wt}-SPC and a KLS^{null}-SPC mouse. More IACs were observed in the KLS^{null}-SPC mouse, which all expressed SPC. Scale bars: 1000 μm. (B) Upper: representative images of a tumour from a KLS^{null}-SPC mouse that had partially escaped *Sox9* recombination. The papillary area outlined by the red dashed line contained SOX9-positive regions that were SPC-negative; the remaining IAC-like region was SPC-positive. Lower: representative images of a PAC from a KLS^{wt} mouse depicting mutual exclusive expression of SOX9 and SPC. Scale bars: 100 μm. (C) Representative images illustrating SOX9 and SPC staining in the lung of a KLS^{wt}-CC10 and a KLS^{null}-CC10 mouse, where IACs were SPC-negative. Scale bars: 1000 μm. (D) Representative images showing spatial SOX9 and SPC expression in tumours of the indicated histotypes detected in KLS^{wt}-CC10 mice. The red dashed line separates the SCC and AC regions within an ASC tumour. Scale bars: 100 μm. (E) Schematic explaining how SOX9 regulates murine NSCLC progression in a histotype-selective manner. Left: in KLS^{wt} mice, SPC$^+$ AT2 cells, targeted by Ad5-SPC-Cre virus, give rise to SPC$^+$ IACs. These IAC lesions gradually start to express SOX9, and this is essential for papillary tumour progression; the SPC$^+$ identity is gradually lost when papillary tissue structure appears. Upon Ad5-CC10-Cre infection, SPC$^+$ IACs again develop, either from AT2 cells that are targeted by the CC10-Cre virus or from trans-differentiated club or dual positive SPC$^+$/CC10$^+$ cells. SOX9 expression in papillary tissue corresponds with loss of SPC expression, giving rise to PACs or the papillary AC core of ASC. Lastly, transformation of CC10$^+$ progenitor cells, which are mostly Clara cells, gives rise to MAC, AAC, and ASC with mucinous AC core. Right: *Sox9* loss selectively disrupts papillary histotype tissue formation in which transformed cells exhibit progression from an SPC-positive to an SPC-negative status.

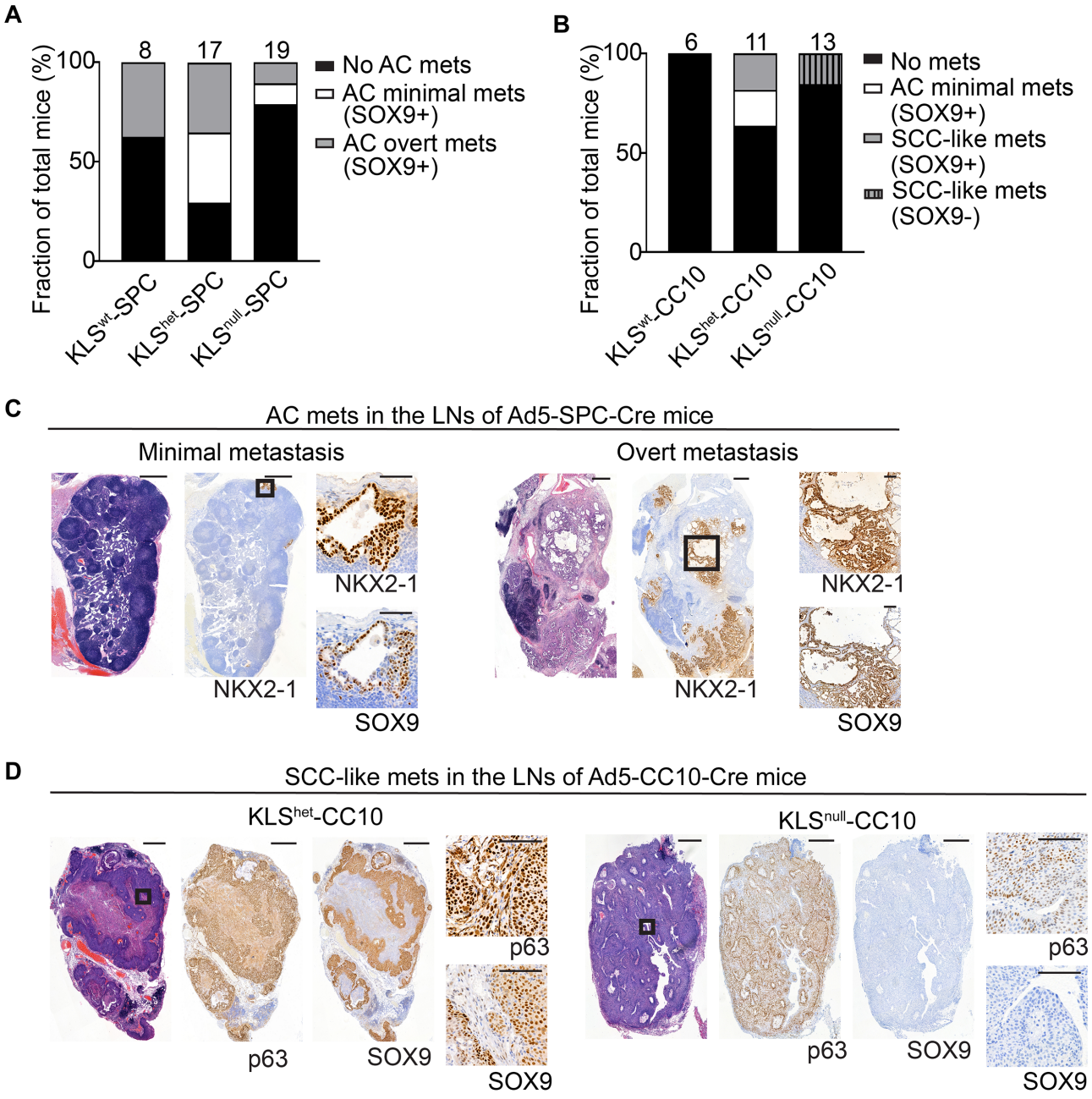

Figure 5. SOX9 expression differently regulates AC and SCC metastases. (A) Bar chart illustrating the proportion of total animals exhibiting no metastasis, minimal metastasis, or overt metastasis in the mediastinal lymph nodes (LNs) in the indicated Ad5-SPC-Cre cohorts. Total numbers of mice are shown above each bar. Metastases were all SOX9-positive, and all were adenocarcinoma. (B) Bar chart illustrating the proportion of total animals exhibiting no, minimal, or overt metastasis in the mediastinal LNs, subgrouped by SOX9 status, in the indicated Ad5-CC10-Cre cohorts. Total numbers of mice are shown above the bars. (C) Representative images illustrating minimal metastasis (only a small LN region infiltrated with tumour cells) and overt metastasis (LN structure severely disrupted) in Ad5-SPC-Cre cohorts; NKX2-1 indicates AC histotype tissue. Areas depicted at greater magnification are outlined with black squares. Scale bars: 500 μm for original images; 100 μm for magnified images. (D) Representative images illustrating SCC-like metastases in Ad5-CC10-Cre cohorts; p63 indicates SCC histotype tissue. Areas depicted at greater magnification are outlined with black squares. Scale bars: 500 μm for original images; 100 μm for magnified images.

suggests that SOX9 expression is specifically required to drive the progression of lesions developed from an SPC^+ cell of origin. Precisely how SOX9 drives AT2 lineage exit remains unclear, and an intriguing future direction would hence be to identify the lineage-specific SOX9 transcriptional network, and particularly its interplay with transcription factors known to regulate alveolar differentiation, such as FOXA2 [20,35,36]. Our study thus reveals a role of SOX9 in NSCLC progression that is selectively linked to the SPC^+/AT2 lineage of cancer-initiating lung progenitor cells.

We showed that *Sox9* loss decreased AC metastasis, which is consistent with a reported pro-metastatic role of SOX9 in multiple cancer types [37–39]. But surprisingly, the incidence of SCC metastasis increased following *Sox9* loss. One mechanism in which SOX9 is

involved in tumour metastasis is through regulating the expression and deposition of ECM components, including multiple collagens and laminins [14]. Our findings importantly suggest that ECM rearrangements regulated by SOX9 expression may influence tumour invasiveness in different and microenvironmental context-dependent

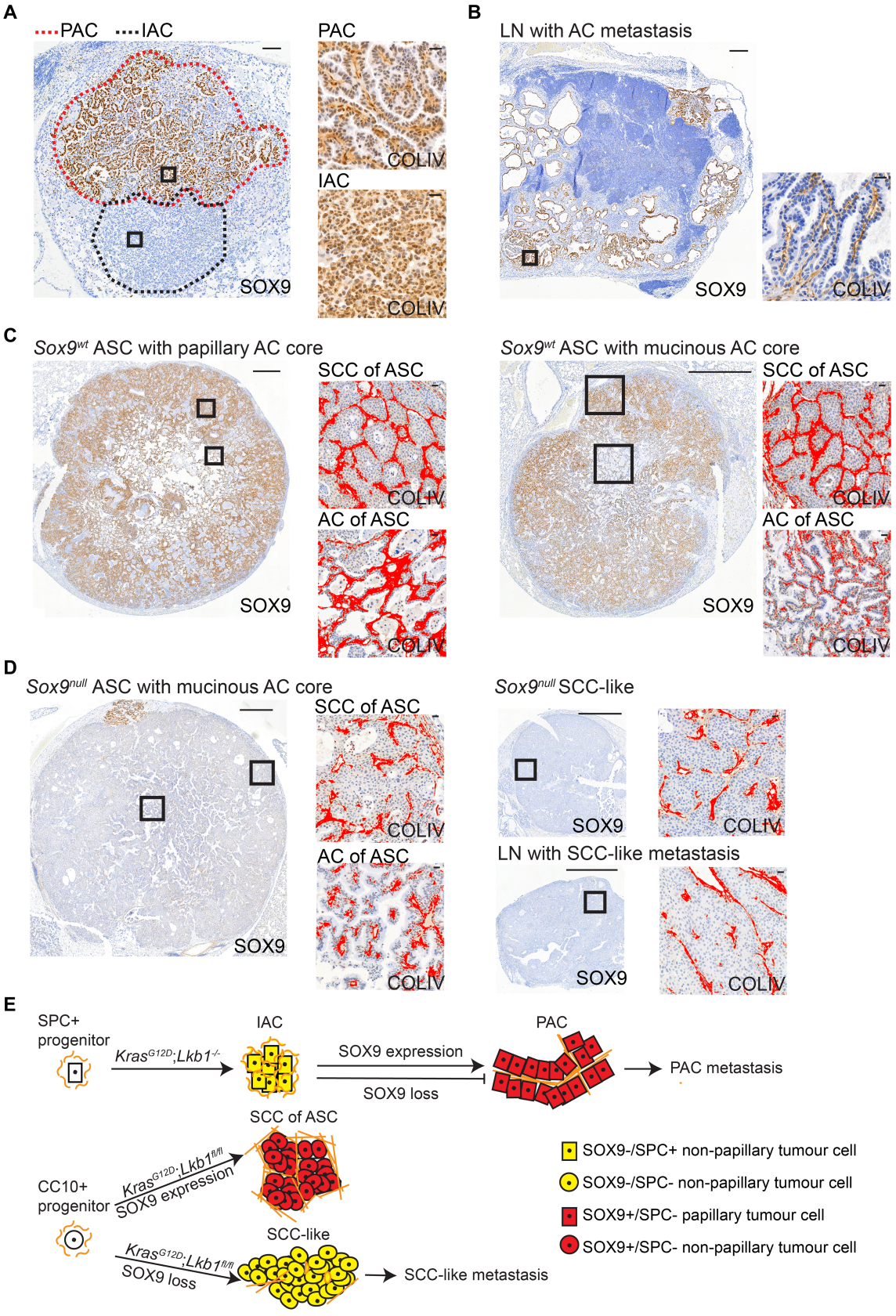

Figure 6 Legend on next page.

manners. We identified differential COLIV deposition, which marks tumour basement membranes, across NSCLC histotypes. Linear COLIV was observed in SOX9-expressing primary PACs and papillary metastatic LNs but not in non-metastatic SOX9-negative IACs, suggesting that it is part of a supportive structure for papillary growth and metastasis upon SOX9 expression. On the other hand, in the SOX9-expressing squamous regions, COLIV formed pocket structures surrounding the tumour islets, and these became discontinuous and granular in the SOX-negative squamous region. Similar COLIV discontinuity has been previously linked to the progression of oral SCC [40], suggesting that disrupted squamous pocket structure is commonly associated with increased squamous invasiveness. Our study therefore depicts functional heterogeneity of SOX9 in NSCLC metastasis, namely a requirement for SOX9-associated matrix deposition in papillary histotype metastasis but an opposing metastasis-suppressing function in squamous histotype tissue.

Analysis of *Sox9* heterozygous null mice showed that loss of one *Sox9* allele already disrupted papillary tissue formation. While *Sox9* haploinsufficiency has not been previously described in cancer studies, loss of one *Sox9* allele has been associated with defects in beta-cell development during murine pancreas organogenesis [41], as well as with defects in murine chondrocyte differentiation [42]. In humans, *Sox9* haploinsufficiency caused by *de novo* heterozygous *Sox9* mutation leads to a rare but severe form of skeletal dysplasia called campomelic dysplasia (CD) [43], and CD mutations can act in a dominant negative manner by disrupting the capacity of SOX9 to dimerise [44]. While our study demonstrates a similar allelic dosage effect of *Sox9* in regulating NSCLC progression and metastasis, further studies are needed to understand whether this is associated with defective SOX9 dimerisation, and which are the target genes that essentially regulate histotype-selective effects during NSCLC progression.

Targeting transcription factors in cancer has been a long-standing challenge, yet the tumour cell-specific SOX9 expression renders it an attractive therapeutic target in NSCLC, in particular for ACs with co-occurring *KRAS* and *LKB1* (*STK11*) mutations, which show low expression of programmed death-ligand 1 and demonstrate poor response towards checkpoint-based immunotherapy [45–48]. Our results indicate that the inhibition of SOX9 could hamper the progression of early-stage ACs with SPC^{+}/AT2 identity. Further study should address the effect of *Sox9* loss subsequent to PAC formation, for example using a flippase-FRT and Cre-loxP dual-recombinase system [49], to learn whether SOX9 is similarly required for the progression or metastasis of already established PACs. If so, then this would provide additional support favouring the development of SOX9 inhibitors to treat advanced AC that derived from AT2 progenitors, particularly in the context of co-occurring *KRAS* and *LKB1* (*STK11*) mutations. Small molecules that disrupt the protein-interaction function of the SOX18 vascular development regulator have recently become available [50]. A similar strategy could possibly be applied to identify small molecules that disrupt SOX9 dimerisation or its binding to other lineage-specific transcription factors. However, while our study implicates SOX9 as a potential anti-cancer target in papillary histotype NSCLC from AT2 cell of origin, *Sox9* loss was also associated with enhanced squamous tissue metastasis, emphasising that therapeutic considerations ought to consider the tumour's histotype context and cannot simply adopt SOX9 positivity as a biomarker for response.

In conclusion, we demonstrate opposing requirements for SOX9 in different NSCLC histotypes and shine light on a pleiotropic SOX9 function that links to different tumour progenitor cells and extracellular microenvironmental features. This provides a strong motivation to evaluate the prognostic value and therapeutic application of SOX9 in the context of NSCLC histopathology types.

Acknowledgements

We thank the Digital and Molecular Pathology Unit supported by the University of Helsinki and Biocenter

Figure 6. Loss of SOX9 affects collagen IV deposition in a histotype-selective manner. (A) Representative images showing SOX9 and collagen IV (COLIV) expression in a KLS^{null}-SPC tumour with both PAC (outlined by red dashed line) and IAC (outlined by black dashed line) features. Squares mark the areas depicted at greater magnification. Scale bars: 100 μm for original images; 20 μm for magnified images. (B) Representative images showing SOX9 and COLIV expression in a lymph node metastasis with PAC features. Squares mark the areas depicted at greater magnification. Scale bars: 100 μm for original images; 20 μm for magnified images. (C) Representative images depicting the COLIV arrangement in a $Sox9^{wt}$ ASC with papillary AC core and a $Sox9^{wt}$ ASC with mucinous AC core. Squares mark the areas depicted at higher magnification. Scale bars: 1000 μm for original images; 20 μm for magnified images. (D) Representative images depicting the COLIV arrangement in a $Sox9^{null}$ ASC with mucinous AC core, a $Sox9^{null}$ SCC-like tumour, and a $Sox9^{null}$ SCC metastasis in the mediastinal LN (p63 staining of the same LN metastasis is shown in Figure 5D). Altered COLIV deposition is detected selectively in SCC of ASC with mucinous AC core and SCC-like tissue. Squares mark the areas depicted at greater magnification. Scale bars: 1000 μm for original images; 20 μm for magnified images. (E) Schematic illustrating how SOX9 may regulate histotype-specific metastasis by influencing ECM (marked by COLIV) secretion/deposition. Linearised COLIV backbones form during IAC to PAC progression, which requires SOX9 expression. *Sox9* loss disrupts PAC formation, and hence their subsequent metastasis. In contrast, while COLIV fibres form continuous boundaries around squamous tumour cell pockets in $Sox9^{wt}$ SCC of ASC that may confine tumour invasiveness, they are re-oriented, or become granular/discontinuous following *Sox9* loss, in SCC-like tumours or SCCs of ASCs with mucinous AC cores. This is associated with an increase in the incidence of SCC-like metastasis, and these metastases similarly show altered COLIV deposition.

Finland, and the Laboratory Animal Centre for husbandry support. We thank Zhangyi He for performing statistical analysis on the TMA-2 dataset. Katja Välimäki is thanked for material support, Ashwini Nagaraj for advice, and Iris Lähdeniemi for reviewing the manuscript. Samples/data used for the research were obtained from the Helsinki Biobank, and we warmly thank all patients for consenting to archive their samples in the Biobank. This research was supported by the University of Helsinki Doctoral Programme in Biomedicine (JB); The Finnish Medical Foundation (MIM); MRC Toxicology Unit core funding (JLQ); Innovative Medicines Initiative Joint Undertaking grant agreement No 115188, the resources of which are composed of a financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in-kind contribution (EWV); and the Academy of Finland grant 307111 (EWV).

Author contributions statement

JB, KN and EWV designed the study and planned the experiments. KN established the animal cohorts and started the experiments. JB conducted further experiments and performed histopathology-selective analyses. AT performed immunohistochemistry and quantification on the large clinical adenocarcinoma cohort. AH and NML assisted with animal health checks, genotyping, and tissue processing. AH performed image scanning. KS provided histopathological review of the murine tumours. MIM provided the grade information of clinical NSCLC samples. JLQ supervised the large clinical adenocarcinoma cohort analysis. JB and EWV wrote the manuscript, and EWV provided supervision.

Data availability statement

The data that support the findings of this study, including murine tumour and TMA-1 analyses, are available as supplementary material, Tables S1 and S2. Data generated with TMA-2 are accessible by contacting the authors.

References

1. Travis WD, Brambilla E, Riely GJ. New pathologic classification of lung cancer: relevance for clinical practice and clinical trials. *J Clin Oncol* 2013; **31:** 992–1001.
2. Campbell JD, Alexandrov A, Kim J, *et al*. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet* 2016; **48:** 607–616.
3. Yanagisawa K, Shyr Y, Xu BJ, *et al*. Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 2003; **362:** 433–439.
4. Bhattacharjee A, Richards WG, Staunton J, *et al*. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A* 2001; **98:** 13790–13795.
5. Travis WD, Brambilla E, Noguchi M, *et al*. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society International multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol* 2011; **6:** 244–285.
6. Li QK, Singh A, Biswal S, *et al*. *KEAP1* gene mutations and NRF2 activation are common in pulmonary papillary adenocarcinoma. *J Hum Genet* 2011; **56:** 230–234.
7. Kadota K, Yeh YC, D'Angelo SP, *et al*. Associations between mutations and histologic patterns of mucin in lung adenocarcinoma: invasive mucinous pattern and extracellular mucin are associated with *KRAS* mutation. *Am J Surg Pathol* 2014; **38:** 1118–1127.
8. Wegner M. From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res* 1999; **27:** 1409–1420.
9. Akiyama H, Chaboissier MC, Martin JF, *et al*. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of *Sox5* and *Sox6*. *Genes Dev* 2002; **16:** 2813–2828.
10. Foster JW, Dominguez-Steglich MA, Guioli S, *et al*. Campomelic dysplasia and autosomal sex reversal caused by mutations in an *SRY*-related gene. *Nature* 1994; **372:** 525–530.
11. Güven A, Kalebic N, Long KR, *et al*. Extracellular matrix-inducing Sox9 promotes both basal progenitor proliferation and gliogenesis in developing neocortex. *Elife* 2020; **9:** e49808.
12. Chang DR, Martinez Alanis D, Miller RK, *et al*. Lung epithelial branching program antagonizes alveolar differentiation. *Proc Natl Acad Sci U S A* 2013; **110:** 18042–18051.
13. Rockich BE, Hrycaj SM, Shih HP, *et al*. Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. *Proc Natl Acad Sci U S A* 2013; **110:** E4456–E4464.
14. Larsimont JC, Youssef KK, Sánchez-Danés A, *et al*. Sox9 controls self-renewal of oncogene targeted cells and links tumor initiation and invasion. *Cell Stem Cell* 2015; **17:** 60–73.
15. Aguilar-Medina M, Avendaño-Félix M, Lizárraga-Verdugo E, *et al*. SOX9 stem-cell factor: clinical and functional relevance in cancer. *J Oncol* 2019; **2019:** 6754040.
16. Sharma A, Cao EY, Kumar V, *et al*. Longitudinal single-cell RNA sequencing of patient-derived primary cells reveals drug-induced infidelity in stem cell hierarchy. *Nat Commun* 2018; **9:** 4931.
17. Lin SC, Chou YT, Jiang SS, *et al*. Epigenetic switch between SOX2 and SOX9 regulates cancer cell plasticity. *Cancer Res* 2016; **76:** 7036–7048.
18. Wong MP, Fung LF, Wang E, *et al*. Chromosomal aberrations of primary lung adenocarcinomas in nonsmokers. *Cancer* 2003; **97:** 1263–1270.
19. Zhou CH, Ye LP, Ye SX, *et al*. Clinical significance of SOX9 in human non-small cell lung cancer progression and overall patient survival. *J Exp Clin Cancer Res* 2012; **31:** 18.
20. Jiang SS, Fang WT, Hou YH, *et al*. Upregulation of SOX9 in lung adenocarcinoma and its involvement in the regulation of cell growth and tumorigenicity. *Clin Cancer Res* 2010; **16:** 4363–4373.
21. Burdelski C, Bujupi E, Tsourlakis MC, *et al*. Loss of SOX9 expression is associated with PSA recurrence in ERG-positive and PTEN deleted prostate cancers. *PLoS One* 2015; **10:** e0128525.
22. Marcker Espersen ML, Linnemann D, Christensen IJ, *et al*. SOX9 expression predicts relapse of stage II colon cancer patients. *Hum Pathol* 2016; **52:** 38–46.
23. Asselin-Labat ML, Filby CE. Adult lung stem cells and their contribution to lung tumourigenesis. *Open Biol* 2012; **2:** 120094.
24. Ferone G, Lee MC, Song J, *et al*. Cells of origin of lung cancers: lessons from mouse studies. *Genes Dev* 2020; **34:** 1017–1032.
25. Nagaraj AS, Lahtela J, Hemmes A, *et al*. Cell of origin links histotype spectrum to immune microenvironment diversity in non-small-cell lung cancer driven by mutant *Kras* and loss of *Lkb1*. *Cell Rep* 2017; **18:** 673–684.
26. Tochigi N, Dacic S, Nikiforova M, *et al*. Adenosquamous carcinoma of the lung: a microdissection study of *KRAS* and *EGFR* mutational

and amplification status in a western patient population. *Am J Clin Pathol* 2011; **135:** 783–789.
27. Närhi K, Nagaraj AS, Parri E, *et al*. Spatial aspects of oncogenic signalling determine the response to combination therapy in slice explants from *Kras*-driven lung tumours. *J Pathol* 2018; **245:** 101–113.
28. Schindelin J, Arganda-Carreras I, Frise E, *et al*. Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012; **9:** 676–682.
29. Bao J, Walliander M, Kovács F, *et al*. Spa-RQ: an image analysis tool to visualise and quantify spatial phenotypes applied to non-small cell lung cancer. *Sci Rep* 2019; **9:** 17613.
30. McQuin C, Goodman A, Chernyshev V, *et al*. CellProfiler 3.0: next-generation image processing for biology. *PLoS Biol* 2018; **16:** e2005970.
31. Contal C, O'Quigley J. An application of changepoint methods in studying the effect of age on survival in breast cancer. *Comput Stat Data Anal* 1999; **30:** 253–270.
32. Aldaz P, Otaegi-Ugartemendia M, Saenz-Antoñanzas A, *et al*. SOX9 promotes tumour progression through the axis BMI1–$p21^{CIP}$. *Sci Rep* 2020; **10:** 1–2.
33. Eichner LJ, Brun SN, Herzig S, *et al*. Genetic analysis reveals AMPK is required to support tumor growth in murine Kras-dependent lung cancer models. *Cell Metab* 2019; **29:** 285–302.e7.
34. Dost AFM, Moye AL, Vedaie M, *et al*. Organoids model transcriptional hallmarks of oncogenic KRAS activation in lung epithelial progenitor cells. *Cell Stem Cell* 2020; **27:** 663–678.e8.
35. Chung C, Kim T, Kim M, *et al*. Hippo–Foxa2 signaling pathway plays a role in peripheral lung maturation and surfactant homeostasis. *Proc Natl Acad Sci U S A* 2013; **110:** 7732–7737.
36. Tan Z, Niu B, Tsang KY, *et al*. Synergistic co-regulation and competition by a SOX9–GLI–FOXA phasic transcriptional network coordinate chondrocyte differentiation transitions. *PLoS Genet* 2018; **14:** e1007346.
37. Yang H, Geng YH, Wang P, *et al*. Extracellular ATP promotes breast cancer invasion and chemoresistance via SOX9 signaling. *Oncogene* 2020; **39:** 5795–5810.
38. Mateo F, Arenas EJ, Aguilar H, *et al*. Stem cell-like transcriptional reprogramming mediates metastatic resistance to mTOR inhibition. *Oncogene* 2017; **36:** 2737–2749.
39. Malladi S, Macalinao DG, Jin X, *et al*. Metastatic latency and immune evasion through autocrine inhibition of WNT. *Cell* 2016; **165:** 45–60.
40. Agarwal P, Ballabh R. Expression of type IV collagen in different histological grades of oral squamous cell carcinoma: an immunohistochemical study. *J Cancer Res Ther* 2013; **9:** 272–275.
41. Dubois CL, Shih HP, Seymour PA, *et al*. Sox9-haploinsufficiency causes glucose intolerance in mice. *PLoS One* 2011; **6:** e23131.
42. Kohn A, Rutkowski TP, Liu Z, *et al*. Notch signaling controls chondrocyte hypertrophy via indirect regulation of Sox9. *Bone Res* 2015; **3:** 15021.
43. Bi W, Huang W, Whitworth DJ, *et al*. Haploinsufficiency of Sox9 results in defective cartilage primordia and premature skeletal mineralization. *Proc Natl Acad Sci U S A* 2001; **98:** 6698–6703.
44. Sock E, Pagon RA, Keymolen K, *et al*. Loss of DNA-dependent dimerization of the transcription factor SOX9 as a cause for campomelic dysplasia. *Hum Mol Genet* 2003; **12:** 1439–1447.
45. Skoulidis F, Byers LA, Diao L, *et al*. Co-occurring genomic alterations define major subsets of *KRAS*-mutant lung adenocarcinoma with distinct biology, immune profiles, and therapeutic vulnerabilities. *Cancer Discov* 2015; **5:** 860–877.
46. Skoulidis F, Goldberg ME, Greenawalt DM, *et al*. *STK11/LKB1* mutations and PD-1 inhibitor resistance in *KRAS*-mutant lung adenocarcinoma. *Cancer Discov* 2018; **8:** 822–835.
47. Cristescu R, Mogg R, Ayers M, *et al*. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 2018; **362:** eaar3593.
48. Koyama S, Akbay EA, Li YY, *et al*. Adaptive resistance to therapeutic PD-1 blockade is associated with upregulation of alternative immune checkpoints. *Nat Commun* 2016; **7:** 10501.
49. Schönhuber N, Seidler B, Schuck K, *et al*. A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer. *Nat Med* 2014; **20:** 1340–1347.
50. Fontaine F, Overman J, Moustaqil M, *et al*. Small-molecule inhibitors of the SOX18 transcription factor. *Cell Chem Biol* 2017; **24:** 346–359.

SUPPLEMENTARY MATERIAL ONLINE

Supplementary figure legends

Figure S1. SOX9 expression in NSCLC subtypes and its correlation with tumour cell proliferation and survival

Figure S2. Altered histopathology and tumour burden following heterozygous or homozygous Sox9 loss

Figure S3. Papillary histotype formation as a result of escaping *Sox9* recombination in KLS^{null} cohorts

Figure S4. The histopathology-selective effect of SOX9 on the progression and metastasis of mouse NSCLC tumours

Table S1. Mouse tumour analyses

Table S2. Human NSCLC TMA-1 analyses

Table S3. Standard IHC and antibody lists for murine samples and TMA-1